

# **Towards Practical and Fundamental Limits of Anonymity Protection**



Dissertation zur Erlangung des Grades eines  
Doktors der Wirtschaftswissenschaft

Eingereicht an der Fakultät für Wirtschaftswissenschaften  
der Universität Regensburg

Vorgelegt von

**Dipl.-Inform. Dang Vinh Pham**

1. Berichterstatter: Prof. Dr. Doğan Kesdoğan  
University of Regensburg, Germany  
Faculty of Business, Economics and Management Information Systems
2. Berichterstatter: Prof. RNDr. Václav Matyáš, Ph.D.  
Masaryk University, Czech Republic  
Faculty of Informatics

Tag der Disputation: 15. November 2013



*This thesis is dedicated to my mother  
for her love and sacrifice.*

## Acknowledgements

After a long journey, it is my pleasure to thank all people who supported this thesis. First of all, I would like to express my gratitude to Prof. Dr. Doğan Kesdoğan and Prof. Ph.D. Václav Matyáš for being my thesis committee and evaluating this thesis. I would like to thank my supervisor Prof. Dr. Doğan Kesdoğan for the fruitful discussions and criticisms during the PhD time. I am thankful for being given the opportunity to do the thesis at his chair. I am indebted to Fatih Karatas and Dr. Lars Fischer for discussions about implementation issues. In particular for the day, when they released me from the weeks of desperate search for a single bug that is specific to parallel programming. Works in this thesis were challenged by exchanges with other colleagues and I would like to thank all of them for providing this research environment. Special thanks to Dr. Benedikt Westermann, Dr. Tobias Mömke and Soujen Chung, who squeezed time from their busy schedule, or spent their vacation to read the thesis. They contributed plenty of comments to help me improving this thesis. Last but not least, I am grateful to my parents, my uncle and my girlfriend for their support and understanding. The PhD work took a great part of my life, leaving little time left to spend with them.

## Abstract

A common function of anonymity systems is the embedding of subjects that are associated to some attributes in a set of subjects, the *anonymity set*. Every subject within the anonymity set appears to be possibly associated to attributes of every other subject within it. The anonymity set covers the associations between the subjects and their attributes. The limit of anonymity protection basically depends on the hardness of disclosing those hidden associations from the anonymity sets. This thesis analyses the protection limit provided by anonymity sets by studying a practical and widely deployed anonymity system, the *Chaum Mix*. A Mix is an anonymous communication system that embeds senders of messages in an anonymity set to hide the association to their recipients (i.e., attributes), in each communication round. It is well known that traffic analyses can uniquely identify a user's recipients by evaluating the sets of senders (i.e., the sender anonymity set) and recipients using the Mix in several rounds. The least number of rounds for that identification represents a fundamental limit of anonymity protection provided by the anonymity sets, similar to Shannon's unicity-distance. That identification requires solving NP-complete problems and was believed to be computationally infeasible.

This thesis shows by a new and optimised algorithm that the unique identification of a user's recipients is for many realistic Mix configurations computational feasible, in the average case. It contributes mathematical estimates of the mean least number of rounds and the mean time-complexity for that unique identification. These measure the fundamental, as well as the practical protection limit provided by the anonymity sets of a Mix. They can be applied to systematically identify Mix configurations that lead to a weak anonymity of a user's recipients. To the best of our knowledge,

this has not been addressed yet, due to the computational infeasibility of past algorithms. All before-mentioned algorithms and analyses can be adapted to deduce information about a user's recipients, even in cases of incomplete knowledge about the anonymity sets, or a low number of observed anonymity sets.

# Contents

<b>Contents</b>	<b>v</b>
<b>List of Figures</b>	<b>xi</b>
<b>List of Tables</b>	<b>xv</b>
<b>Nomenclature</b>	<b>xviii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Anonymous Communication Protection Model . . . . .	5
1.1.1 Attacker Model . . . . .	6
1.1.2 Anonymity Terminology . . . . .	6
1.1.3 Basic Anonymity Techniques . . . . .	10
1.2 The Mix for Anonymous Communication . . . . .	11
1.2.1 Embedding Function . . . . .	12
1.2.1.1 Perfect Mix Concept for Closed Environment . . .	12
1.2.1.2 Chaum Mix Concept for Open Environment . . . .	14
1.2.1.3 Sequence of Mixes . . . . .	16
1.2.2 Group Function . . . . .	17
1.2.3 Mix Variants for Practice-oriented Attacker Models . . . . .	19
1.2.3.1 Stop-and-Go Mix . . . . .	20
1.2.3.2 Pool-Mix . . . . .	21
1.2.3.3 Onion Routing and Non-Disclosing-Method . . . .	22
1.3 Structure . . . . .	22

## CONTENTS

---

<b>2</b>	<b>Combinatorial Attack</b>	<b>25</b>
2.1	Formal Mix and Attacker Model . . . . .	27
2.1.1	Formal Model of Chaum Mix . . . . .	28
2.1.2	Formal Attacker Model . . . . .	29
2.1.2.1	Attacker's Goal . . . . .	29
2.1.2.2	Attack Scheme . . . . .	30
2.1.2.3	Hitting-Set Attack . . . . .	31
2.1.2.4	Learning Number of Alice's Friends . . . . .	33
2.2	Hitting-Set Attack Based on ExactHS . . . . .	35
2.2.1	ExactHS Algorithm . . . . .	36
2.2.1.1	Identification of Hitting-Sets – Examples . . . . .	39
2.2.2	Soundness and Completeness . . . . .	44
2.2.2.1	Properties of Hitting-Sets . . . . .	45
2.2.2.2	Soundness . . . . .	48
2.2.2.3	Completeness . . . . .	49
2.2.3	Worst Case Complexity . . . . .	51
2.2.3.1	Time-Complexity . . . . .	51
2.2.3.2	Space-Complexity . . . . .	55
2.2.4	Evaluation . . . . .	55
2.2.4.1	Communication Traffic . . . . .	57
2.2.4.2	Simulation . . . . .	59
2.3	Approximation of Hitting-Set-Attack . . . . .	63
2.3.1	Classification of Hitting-Sets and Hypotheses . . . . .	63
2.3.2	Approximation Based on No-Trivial-Disproof . . . . .	64
2.3.2.1	Complexity . . . . .	65
2.3.2.2	Relation to $2\times$ -Exclusivity . . . . .	65
2.3.3	Evaluation . . . . .	66
2.4	Summary . . . . .	68
<b>3</b>	<b>Theoretical Limit of Anonymity Protection</b>	<b>71</b>
3.1	Mean Number of Observations . . . . .	72
3.1.1	Bounds of Mean Number of Observations for $2\times$ -Exclusivity	73
3.1.1.1	Relating $1\times$ -Exclusivity and $2\times$ -Exclusivity . . . . .	74



3.1.1.2	Estimation of $2\times$ -Exclusivity Based on $1\times$ -Exclusivity	74
3.1.2	Mean Number of Observations for $k\times$ -Exclusivity . . . . .	76
3.1.2.1	Estimation of $k\times$ -Exclusivity . . . . .	77
3.1.2.2	Comparison of Estimates for $2\times$ -Exclusivity . . . . .	79
3.1.2.3	Effect of Alice's Traffic Distribution on $2\times$ -Exclusivity	80
3.1.3	Relation to Statistical Disclosure Attack . . . . .	80
3.2	Evaluation . . . . .	81
3.2.1	Estimated Number of Observations Required by HS-Attack .	81
3.2.2	Number of Observations Required by HS-attack and SDA . .	83
3.3	Summary . . . . .	85
<b>4</b>	<b>Practical Limit of Anonymity Protection</b>	<b>87</b>
4.1	Upper Bound of Mean Time-Complexity . . . . .	88
4.1.1	Potential – Estimate of Number of Observations Hit by a Hypothesis . . . . .	89
4.1.1.1	Potential in Case of no Chosen Recipient . . . . .	91
4.1.1.2	Potential in General Case . . . . .	92
4.1.1.3	Difference Between Potential and Number of Observations . . . . .	93
4.1.2	Mean Difference Between Potential and Number of Observations	95
4.1.2.1	Expectation of the Difference . . . . .	95
4.1.2.2	Relation of Mean Difference to Number of Chosen Recipients . . . . .	97
4.1.2.3	Relation of Mean Difference to Order of Recipient Choice . . . . .	98
4.1.3	Maximal Mean Number of Recipient Choices for Disproofs .	99
4.1.3.1	Local Maximal Mean . . . . .	100
4.1.3.2	Maximal Mean with respect to Hypothesis Class . .	101
4.1.3.3	Global Maximal Mean . . . . .	102
4.1.4	Maximal Mean Time-Complexity . . . . .	102
4.1.4.1	Estimate . . . . .	103
4.1.5	Evaluation . . . . .	104
4.2	Impact of Traffic Distribution on Mean Time-Complexity . . . . .	107

## CONTENTS

---

4.2.1	Refined Potential – Estimate of Number of Observations Hit by a Hypothesis . . . . .	108
4.2.2	Mean of Potential . . . . .	112
4.2.2.1	Simplified Analysis . . . . .	114
4.2.3	Minimal Mean Number of Recipient Choices for Disproofs . .	114
4.2.3.1	Deriving Maximal Minimal Conditions . . . . .	115
4.2.3.2	Comparing Uniform and Non-Uniform Distribution	123
4.2.3.3	Approaching Optimistic Case Strategy . . . . .	125
4.2.4	Refined Mean Time-Complexity . . . . .	129
4.2.5	Evaluation . . . . .	131
4.2.5.1	Solving Number of Recipient Choices for Disproofs	131
4.3	Summary . . . . .	134
<b>5</b>	<b>Extension</b>	<b>137</b>
5.1	Partial Information . . . . .	138
5.1.1	Quantification of Minimal-Hitting-Sets in a Class . . . . .	139
5.1.1.1	Computing Minimal-Hitting-Sets in a Class . . . .	139
5.1.1.2	Maximal Number of Minimal-Hitting-Sets in a Class	140
5.1.2	Description of Minimal-Hitting-Sets by Extensive-Hypotheses	142
5.1.2.1	Evolution of Extensive-Hypotheses . . . . .	143
5.1.2.2	Construction of Extensive-Hypotheses . . . . .	145
5.1.3	Modelling Evolution of Extensive-Hypotheses . . . . .	150
5.1.3.1	Mean Number of Extensive-Hypotheses . . . . .	150
5.1.3.2	Partial Disclosure . . . . .	153
5.1.3.3	Beyond Unambiguous Information . . . . .	156
5.1.4	Evaluation . . . . .	157
5.2	Vague Information . . . . .	158
5.2.1	Application of Hitting-Set Attack on Vague Observations . . .	159
5.2.1.1	Vague and Erroneous Observations . . . . .	159
5.2.1.2	Applicability of ExactHS . . . . .	161
5.2.2	Analytical Analyses of Conditions for Unique Identification .	164
5.2.2.1	Properties of Ordinary Observations for Unique Identification . . . . .	164

5.2.2.2	Probability Bound of Erroneous Observations for Unique Identification . . . . .	166
5.2.3	Evaluation . . . . .	169
5.3	Summary . . . . .	171
<b>6</b>	<b>Related Works</b>	<b>173</b>
6.1	Combinatorial Analyses . . . . .	174
6.1.1	Attacks for Unique Identification . . . . .	175
6.1.1.1	Intersection Attack . . . . .	175
6.1.1.2	Disclosure Attack . . . . .	175
6.1.1.3	Hitting-Set Attack . . . . .	176
6.1.2	Least Number of Observations for Unique Identification . . .	178
6.1.2.1	Unicity-Distance . . . . .	178
6.1.2.2	2×-Exclusivity . . . . .	178
6.2	Heuristic Analyses . . . . .	179
6.2.1	Likely Set of Friends . . . . .	180
6.2.1.1	Statistical-Hitting-Set Attack . . . . .	180
6.2.1.2	Variants of Statistical-Hitting-Set Attack . . . . .	181
6.2.1.3	HS*-Attack . . . . .	182
6.2.2	Likely Profiles . . . . .	183
6.2.2.1	Statistical Disclosure Attack . . . . .	183
6.2.2.2	Variants of Statistical Disclosure Attack . . . . .	184
6.2.2.3	Bayesian-Interference . . . . .	185
6.2.2.4	Least Square Disclosure Attack . . . . .	185
6.3	Summary . . . . .	186
<b>7</b>	<b>Conclusion</b>	<b>189</b>
7.1	Future Works . . . . .	190
<b>A</b>	<b>Simulation Condition</b>	<b>195</b>
A.1	Hardware . . . . .	195
A.2	Software . . . . .	195
A.3	Computational Time and Memory Usage . . . . .	196

## **CONTENTS**

---

**References** **197**

**Index** **205**

# List of Figures

1.1	Anonymity sets. Sender and recipients are denoted by $s$ and $r$ . Illustrated messages are unlinkable to any subjects framed by ellipses. . .	10
1.2	Processing of messages by embedding function in a round: Each shape around a message $x$ represents a layer of encryption. The outermost encryption layer is removed, when passing a Mix. . . . .	13
1.3	Embedding function in a 2 Mix cascade. Shapes around a message $x$ draw layers of encryptions. Each Mix removes the outermost layer encrypted for it. . . . .	17
1.4	Mix concept complying to CUVE requirement. All Mixes in the cascade have a common list of authenticated senders in a round. Every batch relayed by a Mix must be verified and acknowledged by the senders, using the loop-back function. . . . .	19
1.5	Classification of Mix techniques with respect to their attacker and protection model. . . . .	20
2.1	Mix model. . . . .	28
2.2	Basic scheme in analyses of attacks: Variables $a, r$ stand for arbitrary Alice's friend $a \in {}_A\mathcal{H}$ and recipient $r \in R$ . . . . .	31
2.3	Sequence of collected observations from time point 1 to $t$ . . . . .	39
2.4	Sequence of observations collected by attacker. . . . .	54
2.5	Zipf( $m, \alpha$ ) distribution of Alice's friends, for $m = 23$ . . . . .	58
2.6	Uniformly distributed communication. Left: Mean number of observations to succeed HS-attack. Right: Mean number of finalised sets when succeeding HS-attack. . . . .	60

## LIST OF FIGURES

---

2.7	Uniformly distributed communication. Left: Mean number of observations to succeed HS-attack. Right: Mean number of finalised sets when succeeding HS-attack. . . . .	61
2.8	Zipf-distributed communication. Left: Mean number of observations to succeed HS-attack. Right: Mean number of finalised sets when succeeding HS-attack. . . . .	62
2.9	Zipf-distributed communication. Left: Mean number of observations to succeed HS-attack. Right: Mean number of finalised sets when succeeding HS-attack. . . . .	62
2.10	Mean number of observations until: succeeding HS-attack (HS), fulfilling $2\times$ -exclusivity property (2x-excl) and disappearing of no-trivial-disproofs ( $h_1$ ). . . . .	67
3.1	Mean number of observations: to succeed HS-attack (HS) and to fulfil $2\times$ -exclusivity (2x-excl) versus estimated mean for $2\times$ -exclusivity (2x-excl-est). . . . .	82
3.2	Mean number of observations: to succeed HS-attack (HS) and to fulfil $2\times$ -exclusivity (2x-excl) versus estimated mean for $2\times$ -exclusivity (2x-excl-est). . . . .	83
3.3	Estimated number of required observations: HS-attack (2x-excl-est) versus SDA with 95% true-positive classification ( $SDA_{95\%}$ ), for $u = 400, b = 10, m = 23$ . . . . .	83
3.4	Estimated number of required observations: HS-attack (2x-excl-est) versus SDA with 95% true-positive classification ( $SDA_{95\%}$ ), for $u = 20000, b = 50, m = 40$ . . . . .	84
4.1	Overestimation by potential. Light grey area represents overestimation. <i>Left:</i> $Po(\{r_1, r_2, r_3\}, \{\})$ ; all recipients are non-chosen. <i>Right:</i> $Po(\{r_1, r_2, r_3\}, \{r_1\})$ ; $r_1$ is chosen. . . . .	91
4.2	Unique identification of ${}_A\mathcal{H}$ for $c_{um} = 2$ and $u, b, m$ chosen by (4.15). <i>Left:</i> Number of observations in HS-attack. <i>Right:</i> Level of recursion for disproofs by ExactHS. . . . .	106
4.3	Mean number of finalised sets for unique identification of ${}_A\mathcal{H}$ : $u$ is determined by (4.15) for $c_{um} = 2, b = 50$ and varying value of $m$ . . .	106

## LIST OF FIGURES

---

4.4	Potential of $\mathcal{H} = \{a_1, n_1\}$ is sum of observations containing $a_1$ (horizontal line area) and observations containing $n_1$ (vertical line area). Left: $Po(\mathcal{H}, \{\})$ . Right: $Po(\mathcal{H}, \{a_1\})$ . . . . .	111
4.5	Choosing/removing recipients at $i$ -th level of recursion in enhanced ExactHS as implemented in Algorithm 2. . . . .	120
4.6	Number of recipient choices for disproofs and of finalised sets. <i>Left</i> : Number of choices $c_N^{min}(\cdot)$ . <i>Right</i> : Theoretical vs. empirical number of finalised sets. . . . .	132
4.7	Number of recipient choices for disproofs and of finalised sets. <i>Left</i> : Number of choices $c_N^{min}(\cdot)$ . <i>Right</i> : Theoretical vs. empirical number of finalised sets. . . . .	133
5.1	Number of observation to reduce number of minimal-hitting-sets in $\mathfrak{H}_i$ below 1 (HS1) and 0.1 (HS0.1). . . . .	157
5.2	Number of observations for: full disclosure in simulation (HS), disclosure of at least one recipient (MTTD-1), reduction of number of minimal-hitting-sets below 2 (HS2). . . . .	158
5.3	Unique identification despite erroneous observations. . . . .	163
5.4	Limit of probability of erroneous observations for full disclosure. . . .	170
5.5	Limit of probability of erroneous observations for full disclosure. . . .	171

## **LIST OF FIGURES**

---



# List of Tables

2.1	HS-attack applied to collected observations, given knowledge of $m$ . 3 <sup>rd</sup> and 4 <sup>th</sup> column show minimal-hitting-sets (MHS) and hypotheses computed by HS-attack. . . . .	34
2.2	ExactHS applied on $\mathcal{OS} = \{\mathcal{O}_1, \dots, \mathcal{O}_5\}$ that contains no unique minimum- hitting-set. The same line colour highlights the same level of recursion.	40
2.3	ExactHS applied to $\mathcal{OS} = \{\mathcal{O}_1, \dots, \mathcal{O}_6\}$ , the least observation set con- taining a unique minimum-hitting-set. The same line colour highlights the same level of recursion. . . . .	44
2.4	Naive Computation of hitting-sets of at most size $m$ by combination of recipients from distinct observations. . . . .	54
2.5	Worst case number of finalised sets: HS-algorithm $\binom{u}{m}$ versus Ex- actHS $b^m$ . . . . .	56
5.1	Evolution of minimal-hitting-sets (MHS) and extensive-hypotheses for Alice's set of friends ${}_A\mathcal{H} = \{1, 2, 3\}$ and batch size $b = 2$ . . . . .	144
5.2	Disproof of extensive-hypotheses with exceptions for Alice's set of friends ${}_A\mathcal{H} = \{1, 2, 3\}$ and batch size $b = 3$ . Minimal-hitting-sets is abbreviated by (MHS). . . . .	149
6.1	Mean number of observations for identification of Alice's set of friends with 99% chance by attack (A1), (A2), versus unique identification by HS-attack (HS). . . . .	182

## **LIST OF TABLES**

---

# Nomenclature

## Roman Symbols

$R$	Set of all recipients
$S$	Set of all senders
$u$	Number of users (recipients)
$b$	Batch size
$m$	Number of Alice's friends
$t$	Number of observations collected by the attacker
$g$	Number of finalised sets computed by an attack
$a$	Alice's friend
$n$	Alice's non-friend
$r$	Recipient, without specifying whether it is Alice's friend or non-friend
$\mathcal{O}$	Observation, i.e., a set of recipients that contains a recipient contacted by Alice
$\mathcal{OS}$	Set of observations
$\mathcal{H}$	Hypothesis, i.e., a set consisting of $m$ recipients suspected to be Alice's friends
${}_A\mathcal{H}$	Set of all Alice's friends, i.e., a hypothesis consisting of all Alice's friends
$\mathcal{H}_A$	Set of all Alice's friends in hypothesis $\mathcal{H}$ , i.e., $\mathcal{H}_A = \mathcal{H} \cap {}_A\mathcal{H}$

## LIST OF TABLES

---

$\mathcal{H}_N$	Set of all non-friends in hypothesis $\mathcal{H}$ , i.e., $\mathcal{H}_N = \mathcal{H} \setminus {}_A\mathcal{H}$
$\mathfrak{H}$	Set containing all possible hypotheses
$\mathfrak{H}_j$	Set containing all hypotheses that do not contain $0 \leq j \leq m$ Alice's friends
$\mathcal{C}$	Set of chosen recipients
$P_A(a)$	Probability mass function used by Alice to contact her friend $a$
$P_N(r)$	Probability that any non-Alice sender contacts recipient $r$ in an observation

### Greek Symbols

$\gamma$	Euler-Mascheroni constant $\approx 0.577$
----------	---

### Acronyms

<i>HS-attack</i>	Minimal-hitting-set attack
<i>ExactHS</i>	Exact Hitting Set Algorithm
<i>UMHS</i>	Unique Minimum-hitting-set

# Chapter 1

## Introduction

The importance of protecting privacy was originally addressed by Warren and Brandeis [1890]. They define privacy as “the right to be let alone” to allow human beings to have intimacy. Westin [1970] outlines in a broader context that privacy is crucial for a society like a democracy, which is based on the citizens’ freedom to make individual choices. In order to make such choices, a person needs the opportunity to be let alone to avoid external influences. Therefore, privacy protection is addressed by the United States Privacy Act<sup>1</sup> of 1974. Similarly, the German Constitution comprises the right for “informationelle Selbstbestimmung”<sup>2</sup> (informational self-determination regarding personal data).

Anonymity helps to protect privacy, as it hides the link to the attributes associated to a person. Means to set up anonymity and privacy are well known to humans in the physical world and people insist on them. If we cover ourselves, or talk behind walls, we can usually intuitively and reasonably estimate the provided privacy protection.

Nowadays, a great part of our interactions in the physical world is mapped to the digital world in the Internet. These include, for example, leisure activities, political activities, business activities and interactions with public authorities. Therefore, privacy should also be provided in the digital world. There are several techniques that help to provide privacy in the Internet, cf. Danezis and Gürses [2010]. However, there is still a lack of understanding the fundamental limit of protection provided by many of those techniques. In this thesis, we investigate the anonymity protection provided by

---

<sup>1</sup>Privacy Act. 5 U.S.C. §552a. 1974.

<sup>2</sup>BVerfG, Urteil v. 15.12.1983, Az. 1 BvR 209, 269, 362, 420, 440, 484/83. 1983.

## 1. INTRODUCTION

---

the concept of anonymity sets for anonymous communication.

Pfitzmann and Hansen [2010, p. 10] formulates that “*Anonymity* of a subject from an attacker’s perspective means that the attacker cannot sufficiently identify the subject within a set of subjects, the *anonymity set*.” Anonymity systems seek to establish anonymity sets. This thesis analyses anonymity sets provided by a concrete anonymity system, the Chaum Mix, cf. Chaum [1981], to evaluate the protection they provide. Many anonymity systems are based on the Chaum Mix, cf. Danezis and Diaz [2008]; Edman and Yener [2009]. Some of them, like JAP, Tor and Mixminion, cf. Berthold et al. [2001a]; Danezis et al. [2003]; Dingledine et al. [2004], are currently used by many users. Thus the Chaum Mix is the theoretical basis for many conceptional and practical anonymity systems.

The Chaum Mix hides the association between the senders and recipients of messages (i.e., the attributes) in a network. In every round of communication, it embeds the senders in sender-anonymity sets. According to Pfitzmann [1990], the Mix can be realised in its strongest form in a closed environment<sup>1</sup>, that we call the *perfect Mix*. The perfect-Mix provides information theoretical protection<sup>2</sup> against a *strong attacker*, as in the Dolev-Yao model by Dolev and Yao [1983]. That is an attacker who can observe any link in the network and can delay, inject, or modify messages on that link. However, when adapting the perfect Mix to an open environment<sup>3</sup> like the Internet, the provided protection against a strong attacker is not information theoretically perfect.

The goal of this thesis is to analyse the fundamental limit of anonymity provided by the Chaum Mix in open environments against a strong attacker. Therefore, we analyse the Chaum Mix that would result from applying the perfect Mix to open environments. This can be abstractly modelled by a so-called *threshold Mix* without dummy traffic and broadcast<sup>4</sup>. The model considers the cryptography as secure and assumes that active attacks (i.e., those that delay, inject, mark messages) are avoided by the perfect Mix protocol. Therefore the only information immanently leaked by the Mix to a strong attacker is the sender-anonymity set and the recipient set of the

---

<sup>1</sup>In a closed environment, the set of communicating users is static.

<sup>2</sup>That is a strong attacker cannot gain any new information, despite unlimited computing resources.

<sup>3</sup>That is the set of communicating users is dynamic. Users only join the set if they need to send, or receive a message and leave it otherwise.

<sup>4</sup>Since dummy traffic and broadcast are hard to realise in large open environments.

---

messages relayed in each round. Our analyses refer to the open environment, as it describes the predominant way of communication in the Internet and consequently the most common user communication in practice. Anonymity should be provided with respect to arbitrary communication patterns. Motivated by the observation that users tend to have a persistent set of friends, we analyse the anonymity for the special case where a user recurrently contacts her friends from a static set of friends, over several rounds of communication. We consider the least number of rounds, such that the set of friends is uniquely identifiable, as the fundamental limit of anonymity protection. This is in accordance to Shannon’s unicity-distance, cf. Shannon [1949], that measures the confidentiality provided by a cipher by the number of intercepted cipher text bits to uniquely identify a plain text.

Combinatorial analyses of anonymity sets and recipient sets of messages relayed by the Mix in each round have shown that it is possible to uniquely identify a user’s set of friends. These analyses represent a fundamental threat to the Mix in open environments, as they passively exploit information that is inevitably leaked by the Mix without a malicious behaviour. The Hitting-Set attack (HS-attack) introduced by Kesdogan and Pimenidis [2004] is based on such combinatorial analyses. It is until now the only known attack<sup>1</sup> that identifies a user’s set of friends with the provably least number of rounds, as shown by Kesdogan et al. [2006]. The thesis therefore analyses this attack to measure the fundamental limit of anonymity protection. However, the unique identification of a user’s set of friends requires solving NP-complete problems, as outlined by Agrawal et al. [2003a]; Kesdogan and Pimenidis [2004]. In case of the HS-attack, the underlying problem is the *unique minimum-hitting-set* (UMHS) problem. The algorithm deployed by the HS-attack requires exponential time- and space-complexity to solve the UMHS problem.<sup>2</sup> It is computational infeasible for realistic parameters of the Chaum Mix. These parameters refer to the user’s traffic distributions, the number of friends, the number of users and the size of the anonymity sets and determine the achievable anonymity protection.

In this thesis, we contribute a new algorithm to solve the UMHS problem that provides a significantly lower time-complexity than the old one and only a linear space-

---

<sup>1</sup>This excludes attacks contributed by this thesis.

<sup>2</sup>This refers to the worst and average case complexities.

## 1. INTRODUCTION

---

complexity. Its mean time-complexity is in particular computationally feasible for many realistic parameters of the Chaum Mix. The HS-attack using this new algorithm uniquely identifies a user's set of friends with the least number of rounds. Therefore, this redesign of the HS-attack, instead of the original version, is analysed in this thesis.

The thesis contributes analytical analyses to estimate the least number of rounds and the mean time-complexity for the unique identification of a user's set of friends. As opposed to existing works, our analyses apply to non-uniform traffic distributions and thus cover more realistic cases. We show for reasonable parameters of the Chaum Mix that the mean time-complexity for the unique identification of a user's set of friends is maximal (e.g., exponential) if that user's communication is uniformly distributed. This complexity approaches a polynomial function for various non-uniform distributions of the user's traffic. This relation is particularly proved for the Zipf distribution that is known to closely model a user's Internet and email traffic, as confirmed by Adamic and Huberman [2002]; Almeida et al. [1996]; Breslau et al. [1999]. Therefore, our analytical analyses show that the least number of rounds to uniquely identify a user's set of friends is in many cases even a practical limit of anonymity protection, as opposed to Shannon's unicity-distance Shannon [1949]. We prove that this limit is asymptotically lower than the number of rounds required by the well known *statistical disclosure attack* (SDA) of Danezis [2003] to guess a user's friends with any probability<sup>1</sup>. This limit is bounded by  $O(\frac{1}{p})$ , while the number of rounds required by SDA is  $O(\frac{1}{p^2})$ , where  $p$  is the least probability in the traffic distribution that the attacked user applies to contact her friends.

These relations have not been revealed by past works yet, due to two reasons: The first is the computational infeasibility of the original HS-attack of Kesdogan and Pimenidis [2004]. The second is the restriction of related analytical analyses, cf. Agrawal et al. [2003a]; Kesdogan and Pimenidis [2006]; Kesdogan et al. [2006]; O'Connor [2008] to solely uniform traffic distributions.

Finally, the thesis proposes extensions of analyses of the HS-attack to two novel cases. In the first case, it is assumed that the attacker does not aggregate information from sufficiently many rounds to uniquely identify a user's set of friends. It provides

---

<sup>1</sup>We consider SDA, as its analytical analyses are applicable to non artificial cases, as opposed to heuristic approaches not based on it, as outlined in Section 6.2.



---

analytical analyses of the number of rounds to identify subsets of a user's friends with some probability. The second case assumes that the attacker aggregates in some rounds erroneous information about a user's friends. It suggests a modified application of the HS-attack to identify a user's set of friends despite erroneous information. The analyses derive the maximal probability for erroneous information in a round that is tolerable for a unique identification with a high certainty by that attack.

Erroneous information could be caused by two cases: First, a user randomly contacts recipients who are not her friends, e.g., due to dummy traffic. Second, a user appears to contact a recipient who is not her friend, due to incomplete knowledge about sender anonymity sets and recipient sets processed by the Mix in a round. The suggested extension thus aids studying the HS-attack for conditions that are closer to the real world user behaviour, as well as for Mix variants with indeterministic relay of messages, like the pool-Mix.

The remaining of this Introduction is structured as follows. Section 1.1 introduces the general attacker model and protection model for anonymous communication and the common terminologies. It discusses basic anonymity techniques with respect to these models and concludes that the Chaum Mix technique is the most practical one among them. Section 1.2 details the Chaum Mix technique. It outlines that for open environments, the Chaum Mix provides anonymity against a strong attacker as in the model of Dolev and Yao [1983], as opposed to its more practice oriented variants. This justifies the focus on the Chaum Mix, to study the limit of anonymity protection provided by the Mix technique in open environments in this thesis. Section 1.3 finally depicts the structure of this thesis.

## **1.1 Anonymous Communication Protection Model**

Anonymous communication systems aim at hiding the senders, or recipients of messages, or the communication relations between users in a communication network to provide confidentiality of traffic data with respect to a considered attacker. Anonymity protection is thus considered with respect to an attacker's capability.

Section 1.1.1 introduces descriptions of the attacker model. This is followed by the

## 1. INTRODUCTION

---

definition of anonymity and its variants in Section 1.1.2. Section 1.1.3 provides a brief discussion of basic anonymity techniques and their applicability in practice.

### 1.1.1 Attacker Model

In the following, some terms are introduced that help modelling the attacker. This is necessary, as there is no protection against an almighty attacker, who is not restricted by any model assumptions and can thus control all users, network links and nodes.

A first attribute refers to the role of the attacker with respect to the considered communication event. The attacker is called *involved*, if he participates the communication as a member, either sender, or recipient. Otherwise he is called *uninvolved*. Such an attack is, for example, the operator of the network. The remaining three attributes describe the general capability of an attacker and are adopted from Pfitzmann [1990]:

**Resource:** The attacker can either be *computationally restricted* with limited time and memory resources, or *computationally unrestricted* with unlimited resources.

**Action:** If an attacker can modify, inject, or delay messages, he is called *active*. Otherwise, if he only observes the communication, he is called *passive*.

**Location:** If the attacker can apply his attack on arbitrary network links and nodes, he is called a *global* attacker, otherwise he is called a *local* attacker.

We define a global active attacker, who might be involved, or not involved in a communication as a *strong attacker*. This definition is consistent to the fundamental Dolev-Yao attacker model for cryptography, as described by Dolev and Yao [1983]. In order to study the fundamental limit of anonymity protection, we assume a strong attacker.

However, this thesis aims at analysing confidentiality of traffic data, which is regardless of the content of the messages. Therefore, we exclude for simplicity cryptographic considerations and thus assume that the attacker cannot break cryptographic protocols.

### 1.1.2 Anonymity Terminology

The anonymity terminology that is commonly used in the privacy research area was introduced by Pfitzmann and Köhntopp [2001]. In this thesis, we use the terminol-

---

ogy in its most recent version, as provided by Pfitzmann and Hansen [2010]. Since anonymity depends on the attacker's capability and knowledge, all definitions refer to a considered attacker.

**Anonymity:** “*Anonymity* of a subject from an attacker's perspective means that the attacker cannot sufficiently identify the subject within a set of subjects, the *anonymity set*.” (Pfitzmann and Hansen [2010, p. 10])

**Undetectability:** “*Undetectability*<sup>1</sup> of an item of interest from an attacker's perspective means that the attacker cannot sufficiently distinguish whether it exists or not.” (Pfitzmann and Hansen [2010, p. 16])

**Unlikability:** “*Unlinkability* of two or more items of interest ... from an attacker's perspective means that ... the attacker cannot sufficiently distinguish whether these [items] are related or not.” (Pfitzmann and Hansen [2010, p. 12])

An *item of interest* might be an entity, an action, an identifier, or combinations thereof, c.f. Pfitzmann and Hansen [2010]. According to Pfitzmann and Hansen [2010], anonymity can be described in terms of unlinkability, which supports a more fine grained description of properties than anonymity. Undetectability implies anonymity, but additionally requires the protection of the item of interest as such.

These definitions are coarse and generic. In order to consider anonymity and undetectability in the special context of network communications, we will use the more concrete definitions of these terms as proposed by Pfitzmann [1990]. A network communication system contains a set of all subjects that can use that system. If there is a communication event  $\mathcal{E}$ , e.g., a sending, or receiving of a message, then some of these subjects are involved in the event in the role  $\mathcal{R}$  of, e.g., a sender, or a recipient.

**Undetectability:** An event  $\mathcal{E}$  (e.g., sending/receiving/exchanging of message) is undetectable for an attacker  $\mathcal{A}$  (who is not involved<sup>2</sup> in  $\mathcal{E}$ ), if the a-posteriori probability that  $\mathcal{E}$  exists, after any possible observation  $\mathcal{O}$  of  $\mathcal{A}$ , is  $0 < P(\mathcal{E}|\mathcal{O}) < 1$ .

---

<sup>1</sup>Undetectability was formerly called *unobservability* by Pfitzmann [1990]; Pfitzmann and Köhn-topp [2001].

<sup>2</sup>The event of sending/receiving a message is detectable by the subjects (sender/recipient) involved in that event, therefore undetectability of events does not hold for subjects involved in that events.

## 1. INTRODUCTION

---

If for every possible observation  $\mathcal{O}$  of  $\mathcal{A}$ , the a-priori and a-posteriori probability that  $\mathcal{E}$  exists are additionally equal, that is  $\forall_{\mathcal{O}} P(\mathcal{E}) = P(\mathcal{E}|\mathcal{O})$ , then we call this *perfect preservation of undetectability*<sup>1</sup> according to Pfizmann and Hansen [2010].

*Anonymity*: A subject is anonymous in the role  $\mathcal{R}$  (e.g., sender/recipient) with respect to an event  $\mathcal{E}$  (e.g., sending/receiving of a message) and an attacker  $\mathcal{A}$  (who can even be involved in  $\mathcal{E}$ ), if the a-posteriori probability that it takes the role  $\mathcal{R}$  in  $\mathcal{E}$ , after any possible observation of  $\mathcal{A}$  is greater than 0 and lower than 1.

If for every possible observation  $\mathcal{O}$  of  $\mathcal{A}$ , the a-priori and a-posteriori probability that that subject takes the role  $\mathcal{R}$  in  $\mathcal{E}$  are additionally equal, then we call this *perfect preservation of a subject's anonymity*<sup>2</sup> according to Pfizmann and Hansen [2010].

The set of all subjects who are anonymous in the role  $\mathcal{R}$  with respect to an event  $\mathcal{E}$  and an attacker  $\mathcal{A}$  is the corresponding *anonymity set*. Perfect preservation of a subject's anonymity implies that the anonymity set remains the same for all possible observations  $\mathcal{O}$ .

Anonymity can be further specified by *sender-anonymity*, *recipient-anonymity* and *relationship-anonymity*, c.f Pfizmann [1990]; Pfizmann and Hansen [2010]; Pfizmann and Köhntopp [2001]. These terms refer to following events: sending a message, receiving a message, exchanging any message and corresponding roles. Relationship-anonymity extends the anonymity definition for subjects to that for pairs of subjects. It refers to the anonymity of a pair of (sender, recipient) with respect to the role  $\mathcal{R}$  of communication partners and the event  $\mathcal{E}$  of message exchange. Note that sender-anonymity, or recipient-anonymity is sufficient to imply relationship-anonymity, as stated by Pfizmann [1990]; Pfizmann and Hansen [2010]; Pfizmann and Köhntopp [2001].

According to Shannon [1949], “perfect” protection always means that it is not possible to gain any additional information by observing a system, even if the attacker is computationally unrestricted. The anonymity protection model of Pfizmann [1990] allows specifying the perfect confidentiality of traffic data. That is perfect preservation

---

<sup>1</sup>This was formerly called *perfect unobservability* by Pfizmann [1990].

<sup>2</sup>This was formerly called *perfect anonymity* by Pfizmann [1990].

---

of undetectability against an uninvolved attacker and perfect preservation of anonymity against an involved attacker, which is called *perfect preservation of unobservability* by Pfizmann and Hansen [2010]. That protection refers to a computationally unrestricted attacker.

Note that in our context, the term “perfect” solely characterises the protection of the traffic data. This is regardless of the underlying cryptographic protocols that are used to implement an anonymity concept. Cryptographic considerations are excluded, as we aim to study the protection of traffic data, but not of message content.

**Example 1.1.1** (Anonymity Set). *Figure 1.1 draws the state of an anonymity system with respect to a strong attacker, who can observe all subjects’ actions, but is uninvolved<sup>1</sup> in the communication. It illustrates messages with hidden content in the anonymity system that are unlinkable to any active senders and recipients framed by the ellipses  $S'$  and  $R'$ , with respect to the attacker. The attacker’s knowledge allows excluding the remaining subjects for being senders, or recipients of these messages.*

*Let us consider the events of sending and receiving the message with the grey content, as an example to illustrate anonymity sets. Every subject  $s \in S'$  has a probability of  $0 < p_s < 1$  of being in the role  $\mathcal{R}$  =sender with respect to the event  $\mathcal{E}$  =(send, grey content) and this attacker, so that  $S'$  is a sender-anonymity set. Similarly,  $R'$  is the recipient-anonymity set, as every subject  $r \in R'$  has a probability of  $0 < p_r < 1$  of being in the role  $\mathcal{R}$  =recipient with respect to the event  $\mathcal{E}$  =(receive, grey content) and this attacker. Every pair of sender  $s \in S'$  and recipient  $r \in R'$  has a probability of  $0 < p_{s,r} < 1$  of being the sender and recipient in the event  $\mathcal{E}$  =(exchange, any message) and this attacker. The relationship-anonymity set is thus the cross product of subjects in  $S'$  and  $R'$ , that is:*

$$\{(s_8, r_3), (s_8, r_7), (s_8, r_2), (s_4, r_3), (s_4, r_7), (s_4, r_2), (s_1, r_3), (s_1, r_7), (s_1, r_2)\} .$$

*Assume that the attacker was involved, like the recipients of all messages in this example. The relationship-anonymity set would remain the same, as sender-anonymity, respectively recipient-anonymity implies relationship-anonymity, according to Pfizmann [1990]; Pfizmann and Hansen [2010]; Pfizmann and Köhntopp [2001].*

---

<sup>1</sup>The attacker is neither the sender nor the recipient of the observed messages.

## 1. INTRODUCTION

---

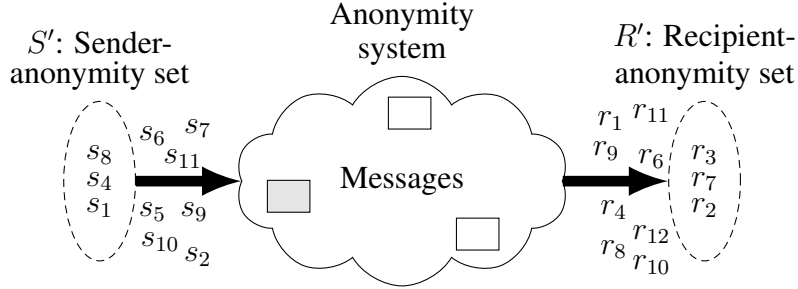


Figure 1.1: Anonymity sets. Sender and recipients are denoted by  $s$  and  $r$ . Illustrated messages are unlinkable to any subjects framed by ellipses.

### 1.1.3 Basic Anonymity Techniques

The following basic anonymous communication techniques against a strong attacker are known:<sup>1</sup>

- *Broadcast* with invisible implicit address for recipient-anonymity and relationship-anonymity, as presented by Pfitzmann and Waidner [1986].
- The *Chaum Mix* for sender-anonymity and relationship-anonymity introduced by Chaum [1981].
- The *DC-Net* for sender-anonymity and relationship-anonymity proposed by Chaum [1988].

As stated by Pfitzmann [1990], broadcast with invisible implicit address only provides anonymity against a strong attacker, if the broadcast medium is secure. This means that the medium must ensure the integrity and the timely delivery of all messages to all recipients, in spite of attacks that might try to delay, drop, or manipulate the network traffic. The attacker might perturb the physical medium itself, or the messages transmitted through it in these attacks. As there is a lack of quality of service guarantee in the Internet, broadcast is impractical in this prevalent communication network.

The DC-Net is based on one-time-pad and requires synchronised computations between the users to transmit a message, as specified by Chaum [1988]. Therefore, this technique is less suitable for spontaneous communication in the Internet.

---

<sup>1</sup>Anonymity systems like Crowds are omitted, as they do not provide protection against a strong attacker, c.f. Reiter and Rubin [1998].

---

In contrast to these techniques, the Chaum Mix is applicable to various network structures and also enables spontaneous communication, making it the most widely deployed technique in practice. A detailed evaluation of these basic anonymity techniques is provided by Kesdogan and Palmer [2006].

The Mix technique also provides an intuitive and transparent construction of anonymity sets as illustrated in Section 1.2. It allows studying the anonymity protection provided by the abstract and general concept of anonymity sets by analysing the protection provided by the Mix technique, as outlined by Kesdogan [2006]. Due to this theoretical and practical property of the Mix, this thesis exclusively focuses on analyses of the Mix technique.

## 1.2 The Mix for Anonymous Communication

The Chaum Mix (also known as *threshold Mix*) introduced by Chaum [1981] represents the base of many popular services offering anonymity in open environments like the Internet. It is a concrete concept that serves to establish anonymity sets in order to provide sender- and relationship-anonymity against a strong attacker. For closed environments, the Chaum Mix concept combined with dummy traffic and broadcast of messages, even provides perfect preservation of unobservability against a computationally unrestricted strong attacker, as shown by Pfitzmann [1990]. We call that concept the *perfect Mix* concept.

The general task of an anonymity system as formulated by Kesdogan and Palmer [2006], is to provide a *group function* and an *embedding function* to build anonymity sets from the users' communication traffic.

**Embedding function:** The embedding function embeds the communication traffic of distinct users in an anonymity set by unifying and hiding their message characteristics (appearance and time), such that they are indistinguishable from each other.

**Group function:** As there is no anonymity against a strong attacker, if there is only a single communication event, the group function serves to enforce that there is additional communication traffic from distinct users, that we call *cover-traffic*. It aims at avoiding that an attacker can control the cover-traffic.

## 1. INTRODUCTION

---

Section 1.2.1 describes the embedding function for the perfect Mix in a closed environments and for the Chaum Mix in an open environment with respect to a strong attacker. The common group function with respect to a strong attacker is described in Section 1.2.2. Combining these two functions provides the corresponding Mix concepts against a strong attacker for the closed, as well as for the open environment.

Section 1.2.3 summarises main variants of Mix concepts for open environments that assume an attacker who is weaker than the strong attacker.

### 1.2.1 Embedding Function

Chaum [1981] introduced a concept for sender and relationship-anonymity, in that messages are prepared by their initiating senders to be grouped and relayed by a Mix to the final recipients, instead of being directly delivered to the final recipients. The purpose of the Mix is to hide the correlation between its incoming and outgoing messages and thus the linking of the sender and recipient of a message.

Section 1.2.1.1 describes the embedding function of the perfect Mix concept for closed environments that aids perfect preservation of unobservability against a strong attacker. The embedding function of the Chaum Mix is a relaxation of that embedding function, such that it is applicable to open environments and aids sender and relationship anonymity against a strong attacker. It is described in Section 1.2.1.2. These sections describe the basic embedding function for the relay of messages through one Mix. Section 1.2.1.3 discusses the straight forward extension of this basis to relay messages through chains of Mixes to reduce the probability that all Mixes are compromised.

#### 1.2.1.1 Perfect Mix Concept for Closed Environment

We use the term *round* to refer to the process of collecting messages and forwarding messages in an anonymity system<sup>1</sup>, as illustrated in Figure 1.2. The perfect Mix concept requires every sender to contribute real or dummy messages in every round. In each round, every sender prepares the same number of messages fixed by the system to be relayed by the Mix, following a *sender protocol*. The Mix processes these messages

---

<sup>1</sup>If messages are processed continuously, a round could refer to a process within a given time window.



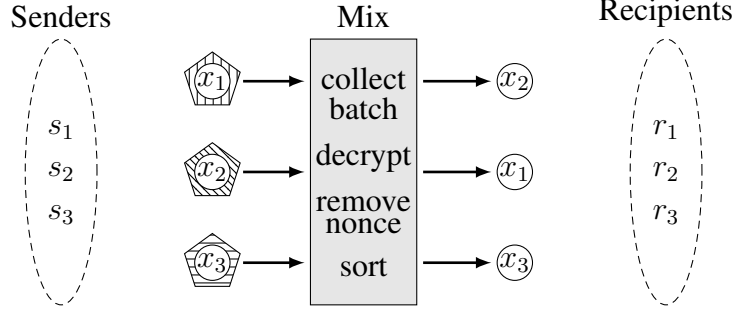


Figure 1.2: Processing of messages by embedding function in a round: Each shape around a message  $x$  represents a layer of encryption. The outermost encryption layer is removed, when passing a Mix.

according to the *Mix protocol* in the same round, as illustrated in Figure 1.2. These protocols conceal the *appearance* and *time characteristic* of messages that might link their senders and recipients. The appearance refers to information about the sender and recipient address, the bit pattern, the size, as well as the number of the message transmitted between a sender and a recipient. The time characteristic addresses the information about the transmission time and the order of transmitted messages.

**Sender Protocol:** The sender prepares his messages to be of constant length either by splitting long messages, or by padding short messages to a length that is fixed by the anonymity system for all messages. If there is no message, then he generates dummy messages, in case of the perfect Mix concept. A message  $mess$  is first encrypted with a *nonce* (number used once) for the final recipient and addressed to that recipient (represented by a circular shape in Figure 1.2). Let us refer to the nonce, address and message for that recipient by the subscript 2, then

$$mess_2 = Enc_2(mess, nonce_2), addr_2 .$$

In case of the perfect Mix concept,  $addr_2$  is an *invisible implicit address* that is used to broadcast the message as proposed by Pfitzmann and Waidner [1986]. Otherwise it is the address of the final recipient of  $mess$ .

This resulting message is then encrypted with a nonce for the Mix and addressed to that Mix (represented by a pentagon shape in Figure 1.2). Let us refer to the

## 1. INTRODUCTION

---

nonce, address and message for that Mix by the subscript 1, then

$$mess_1 = Enc_1(mess_2, nonce_1), addr_1 .$$

**Mix protocol:** A Mix collects the same number of messages from every sender in a *batch* of a fixed size  $b$ , decrypts the message part encrypted only for him and removes the nonce. Given a message  $mess_1$ , by decrypting the part encrypted for him, the Mix thus obtains

$$mess_2, nonce_1 = Dec_1(Enc_1(mess_2, nonce_1))$$

and removes  $nonce_1$ .

The bit representations of the resulting messages are lexicographically sorted in the output batch to shuffle the input and output order of the messages (as illustrated by the Mix in Figure 1.2). If the Mix is the last Mix, then it broadcasts in case of the perfect Mix each output message to all recipients in the anonymity system, otherwise the message is simply forwarded to the recipient.

Perfect preservation of unobservability requires applying these protocols to a *closed environment*, where the set of senders and recipients are static. In that case, the set of all senders and all recipients are in every round the same, so that an attacker cannot gain any new information from observing the anonymity system.

### 1.2.1.2 Chaum Mix Concept for Open Environment

Communication networks in practice like the Internet are *open environments*, where users can join and leave the network at any time and exchange messages, whenever they need to. Consequently, the set of potential users are not known a-priori and not static, as in the closed environment. Therefore, although the Chaum Mix concept provides unlinkability between the input and output messages, the anonymity sets are not static in an open environment, so that perfect preservation of unobservability or perfect preservation of anonymity cannot be provided against a strong attacker.

Networks like the Internet are also very large, so that broadcasting of messages would be impracticable. It is thus reasonable to assume that at the utmost we can ap-

---

ply the perfect Mix embedding function without broadcast to open environments. Yet no theoretic basis for the effective deployment of dummy traffic to increase anonymity in open environments exists, as outlined by Köpsell [2010, pp. 302 – 305]; Berthold and Langos [2003]; Diaz and Preneel [2004]. On the contrary, dummy traffic can cause a significant traffic overhead, as evaluated by Kesdogan [1999, p. 68]; Kesdogan and Palmer [2006], without a clear gain of protection, so that it is mainly omitted. It is further reasonable to assume that the Mix collects exactly one message from each participating sender in every round to minimise the delay of relaying messages, according to analyses of Kesdogan and Palmer [2006]. Otherwise, if the Mix collected, for example, two messages instead of one from each sender<sup>1</sup>, then more time would pass to complete a batch to forward it. While this would decrease the network performance, it might not increase the anonymity protection, as the size of the sender-anonymity-set in a round would remain the same

From the theoretical point of view, we need to understand the anonymity protection provided by the Mix concept without dummy traffic in open environments, before we can understand how dummy traffic can be effectively used to increase that protection. We will follow this line of analysis in this thesis. Therefore, we consider the embedding function of the perfect Mix “without dummy traffic and broadcast” as the basic embedding function of the Chaum Mix in open environments<sup>2</sup>. With the group function in Section 1.2.2, this Mix concept provides sender and relationship-anonymity against a strong attacker in open environments, as shown by Chaum [1981]; Pfizmann [1990].

In Figure 1.2, the sender-anonymity set corresponds to the set of all active senders in a Mix round, i.e.,  $\{s_1, s_2, s_3\}$ . The relationship-anonymity set is represented by all pairs of active senders and recipients in a round, that is:

$$\{(s_1, r_1), (s_1, r_2), (s_1, r_3), (s_2, r_1), (s_2, r_2), (s_2, r_3), (s_3, r_1), (s_3, r_2), (s_3, r_3)\} .$$

The set of active recipients  $\{r_1, r_2, r_3\}$  is no recipient-anonymity set with respect to a strong attacker, as a sender of a message could follow his message to the recipient. However, it is a recipient-anonymity set with respect to a global passive attacker who

---

<sup>1</sup>Note that the Mix must collect the same number of messages from each sender to provide unlinkability between messages in the input and output batch.

<sup>2</sup>This consideration is also in accordance with Chaum [1981].

## 1. INTRODUCTION

---

is not involved in any message transmission.

### 1.2.1.3 Sequence of Mixes

The anonymity protection of the embedding function relies on trusting that the Mix is not compromised by the attacker. However, in the strong attacker model, the Mix itself might be colluded. In order to not be dependent on trusting a single Mix, Chaum [1981] suggests using a sequence of Mixes<sup>1</sup> as illustrated in Figure 1.3, so that the anonymity protection is not affected, as long as at least one of the Mixes is honest. If messages have to pass a fixed sequence of Mixes determined by the anonymity system, then we call it a *Mix-cascade*, while we call it a *Mix-net*, if senders can choose their sequence of Mixes themselves.

In case of using Mixes in open environments, the sender-anonymity set and recipient-set of a particular Mix-cascade consists of all subjects sending and receiving message through that cascade in a round. The sender-anonymity set and recipient-set are known to the Mix-cascade, so that verification of the anonymity protection mechanism by the Mixes and the users is possible. This is necessary to counter attacks by a strong attacker as outlined in Section 1.2.2. If for every sequence of Mixes, the subjects using that sequence are known, then the sender-anonymity set and recipient-set size provided by the Mix-net does not improve over that provided by the Mix-cascade, as outlined by Berthold et al. [2001b]. But if the subjects using a particular Mix sequence were unknown, a Mix-net could extend the sender-anonymity set and recipient-set of that Mix sequence to the set of all senders and recipients who are active in the entire Mix-net, in the same round. However, this would lack verification of anonymity mechanisms to provide protection against a strong attacker, as in Section 1.2.2. Therefore, we only consider Mix-cascades in this thesis. A comprehensive discussion about advantages and disadvantages of Mix-cascades and Mix-nets is provided by Berthold et al. [2001b].

---

<sup>1</sup>The sender and Mix protocol straight forwardly apply to a sequence of Mixes, according to Chaum [1981]; Pfitzmann [1990].

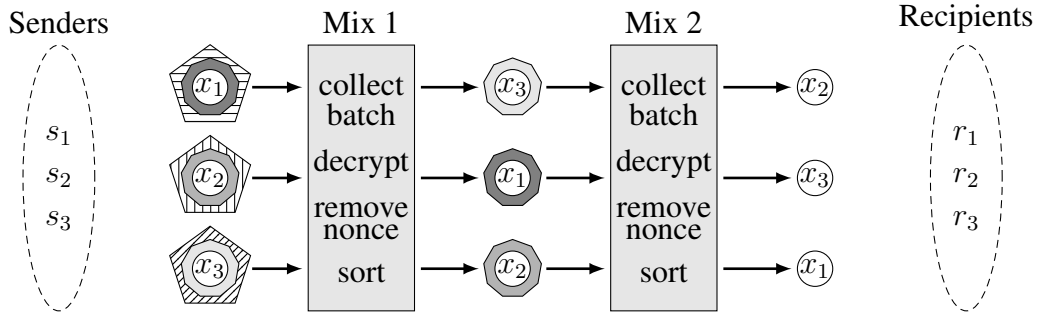


Figure 1.3: Embedding function in a 2 Mix cascade. Shapes around a message  $x$  draw layers of encryptions. Each Mix removes the outermost layer encrypted for it.

## 1.2.2 Group Function

Anonymity and unobservability require several users to cooperate to provide cover-traffic to each other, as there is no anonymity against a strong attacker if there is only one message transmission between a sender and a recipient. Consequently, a strong attacker might try to manipulate the cover-traffic, such that he can isolate the message transmission of a single honest sender from the cover-traffic.

To avoid these kinds of attacks by a strong attacker, Kesdogan and Palmer [2006] require the group function to fulfil the so-called *CUVE* requirement that also comprises all requirements mentioned by Chaum [1981]. *CUVE* requires all Mixes to know all senders participating in a round, i.e., the sender-anonymity set. This information, however, is anyway available to a strong attacker. *CUVE* is an acronym for following requirements:

**Completeness:** All users can verify that their messages have been sent, received, or transmitted.

**Un-reusability:** No user can participate more than an allowed number of times in a round.

**Verifiability:** Messages of non colluded users cannot be changed by attackers without being detected.

**Eligibility:** Only authorised users can participate in a round.

## 1. INTRODUCTION

---

Figure 1.4 illustrates the embedding function of the Mix with the group function that complies to the CUVE requirement. We next describe this group function in detail.

Completeness avoids an attacker from, e.g., blocking messages of  $(b - 1)$  senders of a batch and replacing them by his own messages, cf. Serjantov et al. [2003], such that only the transmission of one honest sender, that is distinguishable to the attacker, remains. Completeness can be technically realised by requiring the Mix to sign the hash of every message relayed in a round and broadcast them to all senders in that round, cf. Chaum [1981]. This is also called the *loop-back function*, cf. Kesdogan and Palmer [2006], and enables senders to verify the correct transmission of their messages to detect blocking of messages<sup>1</sup>. A Mix only relays a batch, if all senders in a round anonymously acknowledge that their messages are correctly included in that batch, as described by Chaum [1981].

Un-reusability avoids an attacker from duplicating and retransmitting a user's message, cf. Chaum [1981], such that a message that is transmitted  $x$  times is linkable to the recipient who receives a message  $x$  times. Un-reusability can be technically realised by time stamping the messages and requiring the Mix to keep a database of relayed messages, so that duplicated messages are omitted, cf. Chaum [1981].

Verifiability avoids an attacker from marking messages by, e.g., manipulating some bits, cf. Berthold et al. [2001b]. Verifiability is technically provided by the loop-back function.

Eligibility avoids masquerading attacks, cf. Dorothy [1982], where an attacker takes several identities and participates as  $(b - 1)$  senders in a batch, such that only the transmission of one honest sender remains, cf. Serjantov et al. [2003]. Eligibility can be technically realised by authenticating every sender in a round and accepting only authenticated messages. Attacks that enable the attacker to know the traces of all messages, apart from that of the message he wants to deanonymise, are generally called  $(b - 1)$ -attacks<sup>2</sup>, cf. Berthold et al. [2001b]; Serjantov et al. [2003].

Note that the Chaum Mix and perfect Mix concept cannot provide anonymity protection, if the attacker controls all Mixes, or if all senders and recipients apart from

---

<sup>1</sup>We focus on protecting anonymity of communication but not on protecting availability, i.e., against denial-of-service attacks.

<sup>2</sup>They are also called  $(n - 1)$ -attack in the literature.

one honest user cooperate with the attacker.

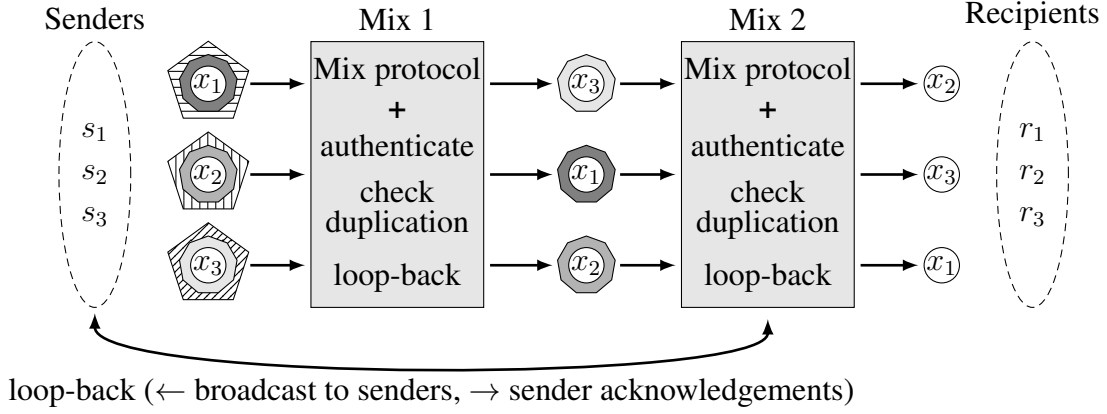


Figure 1.4: Mix concept complying to CUVe requirement. All Mixes in the cascade have a common list of authenticated senders in a round. Every batch relayed by a Mix must be verified and acknowledged by the senders, using the loop-back function.

### 1.2.3 Mix Variants for Practice-oriented Attacker Models

Since the introduction of the Chaum Mix concept by Chaum [1981], several variants of that concept have been suggested for open environments, e.g., Danezis et al. [2003]; Dingledine et al. [2004]; Fasbender et al. [1996]; Kesdogan et al. [1998]; Syverson et al. [1997], assuming that attackers in practice are weaker than those specified by the strong attacker model. These variants commonly avoid the overhead of sender authentication and verification of correct message transmission for the CUVe requirement and of some more protection functions of the Mix. They aim at increasing the flexibility of using anonymity systems in open environments and the performance of those systems. As trade-off, a weakened and restricted attacker model against which they can provide anonymity is accepted, cf. Berthold et al. [2001b]; Serjantov et al. [2003]. A classification of these Mix variants with respect to the perfect Mix and the Chaum Mix is illustrated in Figure 1.5. Readers interested in a more comprehensive description of Mix variants are referred to Danezis [2004]; Danezis and Diaz [2008]; Edman and Yener [2009]; Serjantov [2004].

The strong attacker model is based on the Dolev-Yao model, cf. Dolev and Yao

## 1. INTRODUCTION

---


	Attacker	Protection	strong
Perfect Mix	strong attacker	perfect preservation of anonymity	
Chaum Mix	strong attacker	anonymity	
Mix variants	practical attacker	anonymity	
			weak

Figure 1.5: Classification of Mix techniques with respect to their attacker and protection model.

[1983] that assumes that the network is basically insecure. Considering such an attacker seems to be overcautious, but it is reasonable, as we do not know a-priori all application scenarios and attacker capabilities a system will have to face. That is by designing a system based on assumptions for weaker attackers, that system might provide protection in one situation, but it might fail in another situation, where the assumptions turns out to be invalid. Therefore, it is probably hard to find a meaningful quantification of anonymity protection with respect to practice oriented attacker models. Attacks on Tor, as summarised by Westermann [2012, pp. 20 – 40] particularly show that it is risky to base the protection of a system on assumptions for practice-oriented attacker models, as it might not be possible to assure the validity of such assumptions. We shortly describe the main variants of Mix concepts for practice oriented attacker models in the next sections.

### 1.2.3.1 Stop-and-Go Mix

In the *Stop-and-Go Mix* (SG-Mix) concept proposed by Kesdogan et al. [1998], the function of collecting and shuffling messages is shifted from the Mix to the senders. A sender prepares his message for relaying by attaching to it a random route through a Mix-net of SG-Mixes and the delay of the message at each of the SG-Mixes on that route. Every sender draws the delay at a SG-Mix randomly from an exponential distribution that is dependent on the service rate  $\mu$  of that SG-Mix. That way, it is likely



---

that there will be sufficiently many messages in a SG-Mix. Due to the memoryless property of the exponential distribution, every message in a SG-Mix is equally likely to leave it next, regardless of its arrival time at that SG-Mix. This provides a shuffling of messages, as long as there is more than one message in the SG-Mix.

The random collection of messages in a SG-Mix and thus the provided anonymity depends on the service rate  $\mu$  that relies on estimates of the network traffic. As there is no authentication of senders and no loop-back function in the SG-Mix concept, a strong attacker could control the network traffic, thus undermining the protection provided by SG-Mixes, as mentioned by Kesdogan [1999].

### 1.2.3.2 Pool-Mix

The *pool-Mix* concept tries to increase the anonymity set size by indeterministic relay of messages. This is realised by extending the embedding function of the Chaum Mix to keep a *pool* of  $\eta$  random messages that are not forwarded in a round. In every round, a pool-Mix chooses to output a batch of  $b$  messages from the set of  $(b + \eta)$  messages from the pool and the input batch at that round. This leads to a non-zero probability that a message is not relayed in the same round. As the sender of a message received in one round might have already sent it in a previous round, the sender-anonymity set theoretically covers all senders in the current and past rounds. However, it is more likely that a message is sent in the perimeter close to the same round, than in a round further back in the past. Therefore, Serjantov and Danezis [2003] suggest the so-called “effective” anonymity set size that accounts this likelihood. This supports comparisons of anonymity sets between distinct Mix techniques and variants. Mixmaster and Mixminion, cf. Danezis et al. [2003], are implementations of the pool-Mix concept.

The advantage of the pool-Mix, which is its drawback at the same time is its indeterministic relay of messages. To preserve indeterminism of message relay, authentication of the messages’ senders by the Mix would be an issue, as it would reveal whose messages remain in the pool, thus undermining indeterminism. Verifying the correct message transmission would also be an issue, as a delay of a message could be due to the indeterministic relay, or due to a  $(b - 1)$  attack as well. A  $(b - 1)$  attack is possible, as there is no authentication of the senders, as outlined by Berthold et al. [2001b]; Serjantov et al. [2003]. Therefore the pool-Mix concept does not provide protection

## 1. INTRODUCTION

---

against a strong attacker.

### 1.2.3.3 Onion Routing and Non-Disclosing-Method

The *onion routing* and *non-disclosing-method* (NDM) proposed by Syverson et al. [1997] and by Fasbender et al. [1996] assume a local attacker, who is not able to observe sufficiently many network links to link the messages sent and received by the senders and recipients. In these approaches, messages are relayed through Mix-nets with rudimentary Mix functions that apply the embedding function in Section 1.2.1 without shuffling the input and output order of relayed messages. They assume that a large number of relay nodes and the resulting huge number of Mix-net routes that can be chosen by the senders to transmit their messages provides sufficient anonymity with respect to their local attacker. It is particularly assumed that the attacker does not observe the first and the last relay node on the route of the messages exchanged between the sender and the recipient.

A popular implementation based on these concepts is Tor, which is proposed by Dingledine et al. [2004] and is widely used in practice. Bauer et al. [2007] shows that Tor's attacker model is questionable by undermining its assumptions. They demonstrate an attack that increases the likelihood of observing the first and last Tor relay node of the message route between a sender and a recipient. This attack can be combined with, *traffic confirmation attacks* to identify the sender of a message, c.f. Wang et al. [2007]. However, the anonymity provided by Tor can be even threatened by attacks within its assumed local attacker model, c.f. Evans et al. [2009]; Mittal et al. [2011]; Murdoch and Danezis [2005]. These attacks show that designing the protection of a system for a practice-oriented attacker model, i.e., the local attacker, is risky.

## 1.3 Structure

Chapter 2 deals with a formal description of the Chaum Mix model and attacker model, as well as the scheme of the combinatorial attack based on these models. It introduces a redesign of the HS-attack that is for many realistic Mix parameters computationally feasible. This is confirmed by empirical results that apply the new HS-attack on simulated random observations of Mix rounds. The observed relations between the

---

parameters of the Chaum Mix, the traffic distributions, the number of observations and the complexity to identify a user's set of friends motivates analytical analysis to explain those relations, in this thesis.

Chapter 3 investigates the mean least number of observations required to uniquely identify a user's set of friends. It provides a closed formula for the estimate of that number, even for non-uniform traffic distributions in the Chaum Mix.

Chapter 4 considers the case, where sufficiently many observations have been aggregated to uniquely identify a user's set of friends. It provides analytical analyses of the mean time-complexity for that identification that also applies to non-uniformly distributed traffic.

Chapter 5 studies the disclosure of information about a user's friends for two cases. In the first case, the situation is considered, where the number of observations aggregated by the attacker is not sufficient for a unique identification of Alice's set of friends. It analyses the number of observations, such that a subset of Alice's set of friends can be uniquely identified, or guessed with a certain probability. In the second case, it is assumed that some unknown observations aggregated by the attack are erroneous. It determines the maximal rate of erroneous observations, such that Alice's set of friends can be uniquely identified by an adapted HS-attack with a high certainty. Analyses in the first case only apply to uniformly distributed traffic, while the analyses in the second case also refer to non-uniformly distributed traffic.

Chapter 6 provides an overview over traffic analysis attacks and analyses that refer to the attacker and Chaum Mix model considered in this thesis. That is a strong attacker who aggregates information leaked by the Chaum Mix.

Chapter 7 summarises the results in this theses and outlines open problems for future works.

Appendix A describes the hardware and software deployed by the simulations presented in this thesis.

## **1. INTRODUCTION**

---

## Chapter 2

# Combinatorial Attack

The combinatorial attack considered in this thesis is a traffic analysis attack that applies to any anonymity system that provides sender- and recipient-anonymity sets that are observable to an attacker<sup>1</sup>, cf. Kesdogan [2006]. That is, we consider an ideal anonymity system, such that all information that is available to even a strong attacker are solely the anonymity sets.

Based on the observation that users tend to have a persistent set of friends, the attack aims at disclosing a user's possible set of friends (i.e., her relationship-anonymity). It applies combinatorial analyses on that user's sender- and recipient-anonymity sets, accumulated over several rounds of the anonymity system. For anonymity systems in open environments, this even allows exact identification of a user's set of friends<sup>2</sup>. This identification can be achieved, if the attack accumulates sufficiently many anonymity sets and the user's set of friends remains persistent during that accumulation. As this attack evaluates all information available within the anonymity system (i.e., the anonymity sets), the least number of anonymity sets accumulated to identify a user's set of friends provides a hard limit for the anonymity protection provided by the anonymity system. That user's set of friends cannot be uniquely identified, if fewer anonymity sets are observed than determined by the limit. Therefore this limit represents a pendant to Shannon's unicity-distance for cryptography that is an information theoretic measure

---

<sup>1</sup>Note that this is a sufficient condition for the attack, thus it also applies if, e.g., the recipient-set would not be anonymous, as this would not decrease the information available to the attacker.

<sup>2</sup>According to Pfitzmann and Hansen [2010], sender- and recipient-anonymity implies relationship-anonymity, thus disclosing relationship-anonymity implies the disclosure of the former.

## 2. COMBINATORIAL ATTACK

---

of protection, introduced by Shannon [1949]. In contrast to this, given that there are sufficiently many anonymity sets for the disclosure of a user's set of friends, the computational complexity of the combinatorial attack to disclose those friends represents a practical limit of anonymity protection.

The combinatorial attack allows classifying users according to the persistence of their set of friends, in those users, whose relationship-anonymity can be exactly disclosed and those for that this is not possible. We study the combinatorial attack to understand how the sender- and recipient-anonymity sets constructed by an anonymity system affect this classification. This should aid understanding how to narrow the class of users susceptible to the combinatorial attack, as well as to increase the computation complexity of that attack to rise the practical protection of the users. In other words, our study aids increasing the limit of theoretical and practical anonymity protection for distinct communication patterns of users.

We study in this thesis the application and analysis of this combinatorial attack on a concrete anonymity technique, the Chaum Mix in open environments. To solely focus on the anonymity protection provided by that concept, we assume an ideal implementation of the Chaum Mix protocols and of the CUBE requirements. As described in Section 1.2.1.2, the information that is leaked by such a Chaum Mix to a strong attacker is the set of active senders and recipients, i.e., the sender-anonymity set and the recipient set in a Mix round.

Section 2.1 provides a simple formal description of the Chaum Mix and attacker model that is the underlying model for all analyses in this thesis. Using this simple Mix and attacker model, we describe the combinatorial analysis based on computing minimal-hitting-sets by the HS-attack introduced by Kesdogan and Pimenidis [2004]. This analysis is considered, because it allows the unambiguous disclosure of a user's friends by accumulating the provably least number of observations of sender-anonymity sets and recipient sets, as proved by Kesdogan et al. [2006]. This is in contrast to other related attacks, as outlined in Section 6.

However, the HS-attack requires solving an NP-complete problem and its complexity is originally exponential in time and space, cf. Kesdogan and Pimenidis [2004]. We propose in Section 2.2 a new variant of the HS-attack that contributes a minimised worst case complexity and tractable mean time complexities for practical Mix con-

---

figurations, while providing exactly the same results as the HS-attack. Therefore, all remaining analyses of the HS-attack in this thesis will refer to this more efficient variant.

Although Section 4 will show that there are realistic cases for which our HS-attack variant is tractable, the core of the HS-attack and its variants, is the solving of an NP-complete problem, so that there are always intractable cases. For those cases, Section 2.3 contributes an efficient algorithm that empirically estimates the least number of observations required to uniquely identify Alice’s set of friends more closely than the previous algorithm of Kesdogan et al. [2006]. Although this algorithm is no attack, users can use it to estimate by themselves, when their set of friends can be disclosed by the HS-attack, as suggested by Kesdogan [2006, pp. 31 – 37].

## 2.1 Formal Mix and Attacker Model

The way how anonymity sets are established by the Chaum Mix is similar to the principle of ballot boxes for elections, cf. Chaum [1981]. The ballot box hides the caster of a specific vote within the set of all casters (anonymity set) by shuffling the votes collected in the ballot box (similar to mixing messages within the Mix). This model is sufficiently general to derive more complicated models and simple enough to obtain analytical results. We can clearly measure the effect of distinct parameters, like the size of the anonymity set, or the underlying traffic distribution to the provided anonymity protection by combinatorial approaches.<sup>1</sup> These Mix parameters that specify the properties of the sender-anonymity sets and recipient sets provided by an ideal implementation of the Chaum Mix is defined in this section.

Section 1.2.1.2 has outlined that the Chaum Mix does not prevent passive attackers from observing the sender-anonymity sets and recipient sets for traffic analyses. As a passive strong attacker does not provide any malicious behaviour that is detectable by the Mix<sup>2</sup> and is sufficient for these attacks, we consider without loss of generality solely the global passive attacker in this thesis. That is an attacker who wire-taps every connection in the Mix system and thus observes the sender-anonymity sets and

---

<sup>1</sup>The same analysis is difficult if heuristic approaches are applied, as stated by Pérez-González and Troncoso [2012].

<sup>2</sup>Because the sender-anonymity set and recipient set are immanently leaked by the system.

## 2. COMBINATORIAL ATTACK

recipient sets. We formally describe the information accumulated by this attacker, as well as the evaluation of that information.

### 2.1.1 Formal Model of Chaum Mix

The Chaum Mix was already introduced in Section 1.2.1.2 with technical details. Now we want to focus on formal aspects and abstract from technical details. The Mix system is considered as a black box that outputs information that is visible to the attacker (i.e., the sender-anonymity sets and recipient sets), as illustrated in Figure 2.1. It represents a generalised and simplified model of real-world threshold Mixes that can be adapted to model more complex Mixes, cf. Serjantov and Danezis [2003].

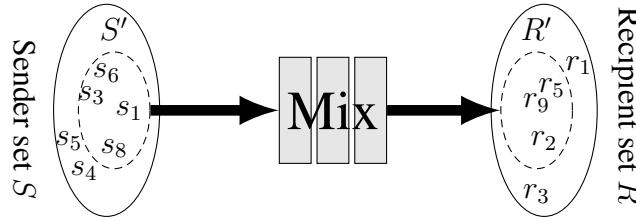


Figure 2.1: Mix model.

The Mix is abstractly described as follow:

- A communication system consists of a set of all senders,  $S$ , a set of all recipients,  $R$ , and a Mix node or Mix cascade as shown in Figure 2.1.  $S$  and  $R$  represent all users with the ability to send or receive messages in the system<sup>1</sup>. If a sender  $s \in S$  communicates with a recipient  $r \in R$ , then we say that  $r$  is a recipient of  $s$ .
- In each communication round a subset  $S' \subseteq S$  of all senders each send precisely one message to their recipients. Let  $R' \subseteq R$  be the set of intended recipients. The act of sending or receiving a message is not hidden to the attacker, therefore  $(S', R')$  represents the information leakage available to an attacker in each round.<sup>2</sup>

<sup>1</sup>This definition allows for cases of  $S \neq R$ , as well as  $S = R$ , i.e., the sender and recipient set might be distinct or identical.

<sup>2</sup>Note that a sender can send to multiple recipients in distinct rounds, but cannot send multiple messages in a single round.



- 
- We call  $S'$  the *sender-anonymity set*, which is the set of all senders who may have sent a given message in a round. The *recipient set*<sup>1</sup>  $R'$  is the set of all recipients who have received a message in a round.
  - We label the size of the sender-anonymity set,  $|S'|$ , as  $b$  which is also called the *batch size*.
  - The size of the *recipient set*,  $|R'|$ , is less than or equal to  $b$ , as each sender sends exactly one message per round, but several senders may communicate with the same recipient. The size of the set of all recipients is  $|R| = u$ .

## 2.1.2 Formal Attacker Model

We consider a *global passive* attacker who observes the traffic on all links between the users and the Mixes in the network. Such an attacker can observe all sending and receiving events in the Mix system, that is he has full knowledge about the pair  $(S', R')$  in every Mix round. We assume that the Mix system is not entirely compromised. This means that at least one of the Mixes in the Mix cascade is honest, so that the Mix system still provides unlinkability of the message entering and leaving the Mix system in a round, complying with the formal Mix model described in Section 2.1.1.

The anonymity protection provided by the Mix system is bypassed, if the contacts of any user can be unambiguously disclosed from the observations, that is from the accumulation of pairs  $(S', R')$ . It is thus sufficient to analyse the disclosure of a single user's contacts to provide measurements for the protection provided by the Mix system. The attacker is assumed to apply the best possible attack, such that it is not possible to gain unambiguous information about a user's contacts with fewer observations than required by that attack.

### 2.1.2.1 Attacker's Goal

The goal of the attacker is to compute, from a set of traffic observations, all possible sets of friends of a target sender  $Alice \in S$ . These possibilities form *hypotheses* for the true set of the sender's friends,  ${}_A\mathcal{H}$ . For a start we assume that Alice contacts a

---

<sup>1</sup>We called it the recipient-anonymity in the past, which is true for passive attackers not involved in the sending or receiving of any messages. However, the term recipient set is in general more precise.

## 2. COMBINATORIAL ATTACK

---

static set of  $m = |{}_A\mathcal{H}|$  friends during the attack<sup>1</sup>. As a convention, we always address a possible set of  $m$  Alice's friends by the term hypothesis, if not otherwise specified. The consideration of non-static communication is discussed as an extension of the attacker model in Section 5.2.

We call a recipient  $r \in {}_A\mathcal{H}$  an *Alice's friend*, or just a *friend*, where  ${}_A\mathcal{H} \subseteq R$ . A recipient who does not communicate with Alice, that is  $r \in R \setminus {}_A\mathcal{H}$ , is called a *non-friend*. If no distinction is required, then  $r$  is simply called a *recipient*.

The attacker focuses on revealing Alice's friends by observing only those pairs  $(S', R')$ , where Alice participates as a sender. Under this condition we refer to the corresponding recipient set  $R'$  as an *observation*,  $\mathcal{O}$ . The set of all observations collected during  $t$  communication rounds is referred to as the *observation set*  $\mathcal{OS} = \{\mathcal{O}_1, \dots, \mathcal{O}_t\}$ .

### 2.1.2.2 Attack Scheme

The attacks considered in this thesis aid investigating the effort to break the anonymity protection provided by the Mix concept and thus by the concept of anonymity sets. In our Mix and attacker model, this effort is dependent on the *Mix parameters*  $(u, b, m)$  and the distribution of the traffic in the considered Mix. To be more precise, the distribution of the traffic refers to the distribution of the recipients addressed by that traffic and we further distinguish between the distribution of *Alice's traffic* and of the *cover-traffic* induced by all senders of the cover-traffic in the Mix system. We use the term *Mix configuration* to refer to such a combination of Mix parameters and traffic distributions. The basic scheme underlying all analyses of these attacks in this thesis is illustrated in Figure 2.2.

We aim at understanding and proving the relation between the effort to reveal information about Alice's friends (represented by the hypotheses) and the considered Mix configurations, based on the attacks. The knowledge of this relation will be applied to derive mathematical measurements of the anonymity protection with respect to the Mix configurations.

---

<sup>1</sup>That is Alice is assumed to contact only friends in  ${}_A\mathcal{H}$  during the attack.

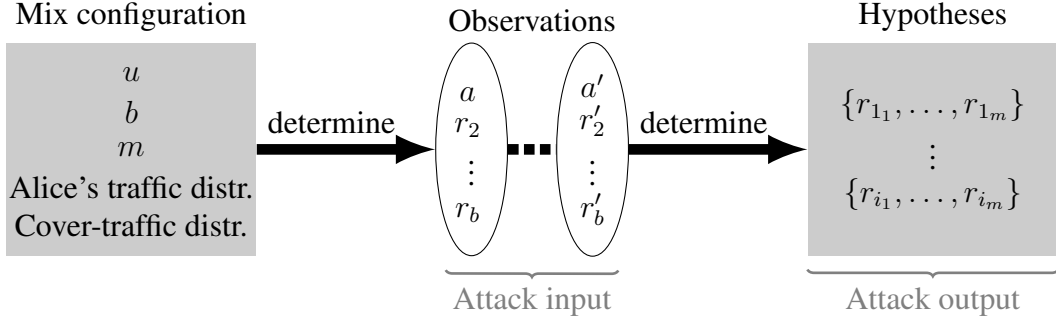


Figure 2.2: Basic scheme in analyses of attacks: Variables  $a, r$  stand for arbitrary Alice’s friend  $a \in {}_A\mathcal{H}$  and recipient  $r \in R$ .

### 2.1.2.3 Hitting-Set Attack

Given that Alice communicates with a static set of  $m$  friends during the attack, the hypotheses for Alice’s possible set of friends in Figure 2.2 can be specified by computing all *hitting-sets* of size  $m$  with respect to the observation set  $\mathcal{OS}$  collected by the attacker. We refer to these hypotheses by the term *specified hypotheses*, if a distinction to the general term “hypothesis” is necessary that refers to every set of  $m$  recipients, regardless whether it is a hitting-set or not. A hitting-set is a set that intersects with all observations<sup>1</sup> in  $\mathcal{OS}$ . A hitting-set is a *minimal-hitting-set* if no proper subset of it is a hitting-set. We call a hitting-set  $\mathcal{H}$  a *unique minimum-hitting-set*<sup>2</sup>, if all hitting-sets  $\mathcal{H}' \neq \mathcal{H}$  in  $\mathcal{OS}$  fulfil the condition  $|\mathcal{H}| < |\mathcal{H}'|$ .

By collecting sufficiently many observations, until  $\mathcal{OS}$  contains a unique minimum-hitting-set, the attacker can unambiguously identify Alice’s set of friends  ${}_A\mathcal{H}$ . This attack is known as the *Hitting-Set attack* (HS-attack), which is introduced by Kesdogan and Pimenidis [2004]. The intuition behind this attack is that at least one Alice’s friends in  ${}_A\mathcal{H}$  appears in each observation<sup>3</sup>, while this does not hold for any set  $\mathcal{H} \neq {}_A\mathcal{H}$ , where  $|\mathcal{H}| \leq |{}_A\mathcal{H}|$ . This applies to non-pathological network traffic, where there are no recipients that are contacted all the time in the anonymity system and thus appear in every observation. Therefore, if there are sufficiently many observations, then  ${}_A\mathcal{H}$  becomes a unique minimum-hitting-set. We call the problem of identifying the unique

<sup>1</sup>Due to the definition of observations,  ${}_A\mathcal{H}$  is a hitting-set in  $\mathcal{OS}$ , that is for all  $\mathcal{O} \in \mathcal{OS}$ ,  ${}_A\mathcal{H} \cap \mathcal{O} \neq \emptyset$ .

<sup>2</sup>Every unique minimum-hitting-set is a minimal-hitting-set, but the reverse conclusion is invalid.

<sup>3</sup>Due to the definition of observations.

## 2. COMBINATORIAL ATTACK

---

minimum-hitting-set, respectively its absent in a given observation set  $\mathcal{OS}$ , the *unique minimum-hitting-set problem* (UMHS problem) . The computation of all minimal-hitting-sets of a size that do not exceed  $m$  is dual to solving the UMHS problem, which is known to be NP-complete, cf. Garey and Johnson [1990]. By solving the UMHS problem, we either obtain a unique minimum-hitting-set, or a proof that there are several minimal-hitting-sets of a minimal size in  $\mathcal{OS}$ .

The HS-attack consists of collecting and adding new observations to  $\mathcal{OS}$  and solving the UMHS problem for that observation set, until  $\mathcal{OS}$  contains a unique minimum-hitting-set. We call the corresponding algorithm that solves the UMHS problem the *HS-algorithm*. If the HS-algorithm finds a unique minimum-hitting-set, then this equals Alice's set of friends  ${}_A\mathcal{H}$  and thus uniquely identifies that set<sup>1</sup>. In that case we say that the attack *succeeds*. It was proven by Kesdogan et al. [2006] that the HS-attack requires the least number of observations to unambiguously identify  ${}_A\mathcal{H}$  with respect to the given Mix and attacker model. Therefore, the mean number of observations required to succeed the HS-attack<sup>2</sup> measures the maximal achievable protection provided by the Mix system against a strong attacker, similar to Shannon's unicity-distance in cryptography, cf. Shannon [1949]. Other combinatorial attacks are either a special case of the HS-attack, or require more observations for the identification of Alice's friends as outlined in Section 6.1.1. We therefore refer to the HS-attack as the basis for combinatorial analyses in this thesis.

In applying the HS-attack, we assume that the size of Alice's set of friends,  $m$ , is known, since  $m$  can be easily learned, as discussed in Section 2.1.2.4.

**Example 2.1.1** (Unique Identification by HS-attack). *We illustrate the unique identification of Alice's set of friends  ${}_A\mathcal{H}$  by the HS-attack. In this example, the set of all recipients is  $R = \{1, 2, 3, 4, 5, 6\}$  and Alice's set of friends is  ${}_A\mathcal{H} = \{1, 2, 3\}$ , thus  $u = |R| = 6$  and  $m = |{}_A\mathcal{H}| = 3$ . The set of observations collected by the attacker is initially  $\mathcal{OS}_0 = \{\}$ . When collecting the  $i$ -th observation  $\mathcal{O}_i$  the attacker extends this set of observations to  $\mathcal{OS}_i = \mathcal{OS}_{i-1} \cup \mathcal{O}_i$  and computes all specified hypotheses in  $\mathcal{OS}_i$  as shown in Table 2.1.*

*The fourth column in Table 2.1 shows the specified hypotheses in  $\mathcal{OS}_i$ , determined by the original HS-algorithm of Kesdogan and Pimenidis [2004]. If no observations*

<sup>1</sup>The equality might not hold in pathological cases that are unlikely as discussed in Section 2.1.2.4.

<sup>2</sup>That is the average size of the observation set  $\mathcal{OS}$  collected by the attacker to succeed the HS-attack.

---

(i.e.,  $i = 0$ ) have been collected, then every set of size  $m$  is a hitting-set. In that case, all hypotheses are specified hypotheses and the number of hypotheses is  $\binom{u}{m}$ , which is 20 in our example. The HS-attack excludes previously specified hypotheses by each collected observation, until a single specified hypothesis remains. That specified hypothesis equals Alice's set of friends.

The third column in Table 2.1 represents all minimal-hitting-sets (MHS) in  $\mathcal{OS}_i$ . We can see that if a single specified hypothesis remains, then this hypothesis is also a unique minimum-hitting-set. Thus computing all specified hypotheses by the HS-algorithm also provides a solution to the UMHS problem for hitting-sets of size  $m$ .

This example also illustrates that the number of specified hypotheses is never lower than the number of minimal-hitting-sets of at most size  $m$  in every sequence of observations collected by the attacker. There are particularly more specified hypotheses than minimal-hitting-sets prior to a certain number of observations, while they become equal after that number of observations.

#### 2.1.2.4 Learning Number of Alice's Friends

The intuition behind our attack is that at least one of Alice's friends must appear in each observation<sup>1</sup>, while this does not hold for any other set  $\mathcal{H}$ , where  $\mathcal{H} \not\subseteq {}_A\mathcal{H}$ . Therefore, after a sufficient number of  $t$  observations, Alice's set of friends  ${}_A\mathcal{H}$  remains the unique minimum-hitting-set.

Assume the existence of a set in which  $\mathcal{H} \neq {}_A\mathcal{H}$ , where  $|\mathcal{H}| < m$  happens to be a unique minimum-hitting-set<sup>2</sup>. Let  $q$  be the probability that any recipient in  $\mathcal{H}$  appears in a random observation. The probability that  $\mathcal{H}$  remains a hitting-set after  $x$  observations decreases according to an exponential function  $q^x$  and is thus negligible. As defined in cryptography, a function  $f(x) : \mathbb{N} \mapsto \mathbb{R}$  is *negligible*, if for every positive polynomial  $pol(x)$ , there is a  $c > 0$  such that  $|f(x)| < \frac{1}{pol(x)}$  for all  $x > c$ . Consequently, collecting sufficiently many observations when applying the HS-attack leads to a probability of learning the wrong set of Alice's friends and the wrong value of  $m$  that is very small to matter. This is even the case for moderate number of observations  $x$ .

---

<sup>1</sup>Recall that an "observation" refers to a round in which Alice participates.

<sup>2</sup>This case is pathological, as it requires an uncommon combination of two coincidences, a hitting-set  $\mathcal{H}$  of a size less than  $m$  and the absence of any other hitting-sets of the same size as  $|\mathcal{H}|$ .

## 2. COMBINATORIAL ATTACK

$i$	$\mathcal{O}_i$	MHS	Hypotheses
0	—	—	$\{1, 2, 3\},$ $\{1, 2, 4\}, \{1, 2, 5\}, \{1, 2, 6\}, \{1, 3, 4\}, \{1, 3, 5\},$ $\{1, 3, 6\}, \{2, 3, 4\}, \{2, 3, 5\}, \{2, 3, 6\},$ $\{1, 4, 5\}, \{1, 4, 6\}, \{1, 5, 6\}, \{2, 4, 5\}, \{2, 4, 6\},$ $\{2, 5, 6\}, \{3, 4, 5\}, \{3, 4, 6\}, \{3, 5, 6\},$ $\{4, 5, 6\}$
1	$\{1, 4\}$	$\{1\}, \{4\}$	$\{1, 2, 3\},$ $\{1, 2, 4\}, \{1, 2, 5\}, \{1, 2, 6\}, \{1, 3, 4\}, \{1, 3, 5\},$ $\{1, 3, 6\}, \{2, 3, 4\},$ $\{1, 4, 5\}, \{1, 4, 6\}, \{1, 5, 6\}, \{2, 4, 5\}, \{2, 4, 6\},$ $\{3, 4, 5\}, \{3, 4, 6\},$ $\{4, 5, 6\}$
2	$\{2, 5\}$	$\{1, 2\}, \{1, 5\},$ $\{4, 2\}, \{4, 5\}$	$\{1, 2, 3\},$ $\{1, 2, 4\}, \{1, 2, 5\}, \{1, 2, 6\}, \{1, 3, 5\},$ $\{2, 3, 4\},$ $\{1, 4, 5\}, \{1, 5, 6\}, \{2, 4, 5\}, \{2, 4, 6\},$ $\{3, 4, 5\},$ $\{4, 5, 6\}$
3	$\{1, 6\}$	$\{1, 2\}, \{1, 5\},$ $\{4, 2, 6\}, \{4, 5, 6\}$	$\{1, 2, 3\},$ $\{1, 2, 4\}, \{1, 2, 5\}, \{1, 2, 6\}, \{1, 3, 5\},$ $\{1, 4, 5\}, \{1, 5, 6\}, \{2, 4, 6\},$ $\{4, 5, 6\}$
4	$\{3, 4\}$	$\{1, 2, 3\}, \{1, 2, 4\},$ $\{1, 5, 3\}, \{1, 5, 4\},$ $\{4, 2, 6\}, \{4, 5, 6\}$	$\{1, 2, 3\},$ $\{1, 2, 4\}, \{1, 3, 5\},$ $\{1, 4, 5\}, \{2, 4, 6\},$ $\{4, 5, 6\}$
5	$\{2, 6\}$	$\{1, 2, 3\}, \{1, 2, 4\},$ $\{4, 2, 6\}, \{4, 5, 6\}$	$\{1, 2, 3\},$ $\{1, 2, 4\},$ $\{2, 4, 6\},$ $\{4, 5, 6\}$
6	$\{1, 3\}$	$\{1, 2, 3\}, \{1, 2, 4\}$	$\{1, 2, 3\},$ $\{1, 2, 4\}$
7	$\{3, 5\}$	$\{1, 2, 3\}$	$\{1, 2, 3\}$

Table 2.1: HS-attack applied to collected observations, given knowledge of  $m$ . 3<sup>rd</sup> and 4<sup>th</sup> column show minimal-hitting-sets (MHS) and hypotheses computed by HS-attack.

---

We can easily learn  $m$  by applying the HS-attack for the possible set of Alice's friends of size  $m' = 1, \dots, m$ .<sup>1</sup> If the HS-attack is applied for  $m' < m$ , then there will be no hitting-set of size  $m'$  after collecting some set of observations  $\mathcal{OS}'$ , where  $t' = |\mathcal{OS}'|$  is the number of observations. Let  $t = |\mathcal{OS}|$  be the least number of observations to succeed the HS-attack for  $m$ , then  $t' \leq t$ ,<sup>2</sup> where  $\mathcal{OS} = \mathcal{OS}' \cup \{\mathcal{O}_{t'+1}, \dots, \mathcal{O}_t\}$ . If  $m' < m$ , then  $t'$  is usually low<sup>3</sup>, such that the correct value  $m$  can be learned fast and reliably. The MHS column in Table 2.1 of Example 2.1.1 illustrates, when values of  $m'$  that underestimate  $m$  are detected.

$m' = 1$ : After collecting the second observation ( $i = 2$ ), all MHS in  $\mathcal{OS}_2$  are of at least size 2. This proves that there is no hitting-set of a size lower 2 in  $\mathcal{OS}_2$ . The HS-attack applied for  $m' = 1$  will fail to find any hitting-sets of size 1 and thus detects that the size of Alice's set of friends cannot be 1.

$m' = 2$ : After collecting the fourth observation ( $i = 4$ ), all MHS in  $\mathcal{OS}_4$  are of at least size 3. This proves that there is no hitting-set of size smaller than 3 in  $\mathcal{OS}_2$ . The HS-attack applied for  $m' = 2$  will fail to find any hitting-sets of size 2 and thus detects that the size of Alice's set of friends cannot be 2. The attacker learns the correct value of  $m$  after 4 observations in this example.

## 2.2 Hitting-Set Attack Based on ExactHS

ExactHS is an algorithm that allows solving the UMHS problem by computing all minimal-hitting-sets in a given observation set  $\mathcal{OS}$  that does not exceed the size  $m$ . Its worst case time-complexity is proportional to  $b^m$ , while its space-complexity is linear, in contrast to the original HS-algorithm of Kesdogan and Pimenidis [2004] that requires  $O(\binom{u}{m})$  time and space-complexity. By using ExactHS to solve the UMHS problem in the HS-attack, instead of using the original HS-algorithm, we obtain our variant of the HS-attack. We refer to that variant by the term *HS-attack using ExactHS*. The practical advantage of the HS-attack using ExactHS is its low mean time-

---

<sup>1</sup>By applying the HS-attack for  $m'$ , the size of all computed hitting-sets do not exceed  $m'$ .

<sup>2</sup>This inequality is obvious. As long as there is a hitting-set  $\mathcal{H}'$  of size  $m' < m$ , there must be several hitting-sets  $\mathcal{H} = \mathcal{H}' \cup \{r_{m'+1}, \dots, r_m\}$  of size  $m$  for arbitrary  $\{r_{m'+1}, \dots, r_m\} \subseteq R \setminus \mathcal{H}'$ .

<sup>3</sup>In general, the smaller the size  $m'$  of a set  $\mathcal{H}' \neq {}_A\mathcal{H}$ , the smaller is its probability to remain a hitting-set when new observations are collected and thus the smaller is  $t'$ .

## 2. COMBINATORIAL ATTACK

---

complexity that is with respect to many realistic Mix configurations, computationally feasible. This feasibility is despite the NP-completeness of the UMHS problem, as we will demonstrate in Section 2.2.4 and prove in Chapter 4.

The computational advantage of ExactHS over HS-algorithm is due to two reasons: Firstly, ExactHS computes all minimal-hitting-sets instead of all specified hypotheses with respect to the attacker's observation set  $\mathcal{OS}$  and  $m$ .<sup>1</sup> It makes use of the observation that the number of minimal-hitting-sets of at most size  $m$  (i.e., at most  $b^m$ ) is lower than the number of specified hypotheses of size  $m$  (i.e.,  $\binom{u}{m}$ ) in every observation set  $\mathcal{OS}$ , as demonstrated in Table 2.1 of Example 2.1.1. Secondly, the mean time-complexity of ExactHS approaches a lower bound<sup>2</sup>, if applied to sufficiently many collected observations. This is due to the decrease of the number of minimal-hitting-sets of size  $m$  with respect to the number of observations. Therefore, ExactHS can be applied, when the UMHS problem becomes computationally less hard. This is the case, when the number of observations collected by the attacker is close to that to uniquely identify Alice's set of friends.

This is not possible with HS-algorithm, as it requires the computation of all possible  $\binom{u}{m}$  hypotheses prior to determining which hypotheses can be excluded by the collected observations.

Section 2.2.1 introduces the ExactHS algorithm. The correctness and completeness of that algorithm in computing all minimal-hitting-sets in any given observation set is proved in Section 2.2.2. Section 2.2.3 then determines a tight bound for the worst case time and space-complexity of ExactHS.

### 2.2.1 ExactHS Algorithm

ExactHS recursively computes all minimal-hitting-sets of at most size  $m$  in a given observation set  $\mathcal{OS}$ . Applying ExactHS thus allows proving, whether there is a unique minimum-hitting-set of size  $m$  in a given observation set  $\mathcal{OS}$ . We use the following notation:

---

<sup>1</sup>The MHS column in Table 2.1.1 equals the result of ExactHS with respect to  $\mathcal{OS}_i$  and  $m$ . If there is only a single MHS, then it is a UMHS, otherwise there is more than one MHS of at most size  $m$ .

<sup>2</sup>This bound is dependent on the Mix configuration and is analytically determined in Chapter 4.



- 
- $\mathcal{C}$ : Set of at most  $m$  suspected<sup>1</sup> recipients representing a subset of a possible hitting-set. We initially set  $\mathcal{C} = \{\}$ .
  - $\mathcal{HS}$ : Set that stores all hitting-sets identified by ExactHS, which is initially  $\mathcal{HS} = \{\}$ . It represents a global variable that is a non integral part of ExactHS, like an output stream on the display<sup>2</sup>.
  - $\mathcal{OS}[r]$ : Set of observations containing recipient  $r$ , that is  $\mathcal{OS}[r] = \{\mathcal{O} \in \mathcal{OS} \mid r \in \mathcal{O}\}$ . We call  $|\mathcal{OS}[r]|$  the *frequency* of  $r$ , where  $|\mathcal{OS}[r]|$  is 0, if  $r$  is not in any observations of  $\mathcal{OS}$ .
  - $\mathcal{OS}[\{r_1, \dots, r_k\}]$ : Set of observations containing any recipient  $r_1, \dots, r_k$ , that is the observation set resulting from  $\bigcup_{i=1}^k \mathcal{OS}[r_i]$ .

The basic scheme of ExactHS was first proposed by Pham and Kesdogan [2009]; Pham [2008]. It is presented in Algorithm 1, solely to illustrate the difference to its enhanced version as implemented in Algorithm 2, that is the focus of this thesis. While

---

**Algorithm 1** Basic Scheme of ExactHS.

---

```

1: procedure EXACTHS( $\mathcal{OS}'$ ,  $m'$ ,  $\mathcal{C}$ )
2:   if  $\mathcal{OS}' = \{\}$  then
3:      $\mathcal{HS} = \mathcal{HS} \cup \mathcal{C}$  ▷  $\mathcal{C}$  is a hitting-set
4:   else if  $m' \geq 1$  then ▷  $\mathcal{C}$  can be extended, if  $|\mathcal{C}| < m$ 
5:     fix  $\mathcal{O}' \in \mathcal{OS}'$ 
6:     while  $(\{\} \notin \mathcal{OS}') \wedge (|\mathcal{O}'| \geq 1)$  do
7:       choose  $r \in \mathcal{O}'$  ▷  $r$  will be added to  $\mathcal{C}$ 
8:       EXACTHS( $\mathcal{OS}' \setminus \mathcal{OS}'[r]$ ,  $m' - 1$ ,  $\mathcal{C} \cup \{r\}$ ) ▷ select remaining  $(m' - 1)$ 
       recipients
9:        $\mathcal{OS}' \leftarrow \bigcup_{\mathcal{O}'_l \in \mathcal{OS}'} \{\mathcal{O}'_l \setminus \{r\}\}$  ▷ remove  $r$  from all observ. in  $\mathcal{OS}'$ 
10:       $\mathcal{O}' \leftarrow \mathcal{O}' \setminus \{r\}$  ▷ do not choose  $r$  in this recursion level again

```

---

both ExactHS versions provide the same worst case complexities, Algorithm 2 contains tiny, but effective changes in Lines 5 – 7 to reduce the mean time-complexity. These changes are based on analytical results that will be introduced in Chapter 4. We anticipate the effect of the lower mean time-complexity due to applying those results to ExactHS, by evaluating Algorithm 2 in this chapter.

---

<sup>1</sup>During execution,  $\mathcal{C}$  either becomes a minimal-hitting-set, or it will be proved not to be a subset of any minimal-hitting-sets that has not been evaluated.

<sup>2</sup>The space-complexity of ExactHS is therefore independent of  $\mathcal{HS}$ .

## 2. COMBINATORIAL ATTACK

---

All evaluations, descriptions and analyses in this remaining chapter refer by default to the enhanced version of ExactHS, whenever the term ExactHS is referred, if not otherwise stated. However, analyses of the worst cases complexities apply to both versions of ExactHS. We now describe in detail the steps taken by ExactHS, when jumping from one level of recursion to the succeeding level of recursion and back.

---

### Algorithm 2 ExactHS.

---

```

1: procedure EXACTHS( $\mathcal{OS}', m', \mathcal{C}$ )
2:   if  $\mathcal{OS}' = \{\}$  then
3:      $\mathcal{HS} = \mathcal{HS} \cup \mathcal{C}$  ▷  $\mathcal{C}$  is a hitting-set
4:   else if  $m' \geq 1$  then ▷  $\mathcal{C}$  can be extended, if  $|\mathcal{C}| < m$ 
5:     fix  $\mathcal{O}' \in \mathcal{OS}'$  that contains the most frequent recipient
6:     while  $(\{\} \notin \mathcal{OS}') \wedge (\max_{\{r_1, \dots, r_{m'}\} \subseteq \bigcup_{\mathcal{O} \in \mathcal{OS}'} \mathcal{O}} \{\sum_{l=1}^{m'} |\mathcal{OS}'[r_l]|\} \geq |\mathcal{OS}'|)$  do
7:       choose  $r \in \mathcal{O}'$  that is most frequent ▷  $r$  will be added to  $\mathcal{C}$ 
8:       EXACTHS( $\mathcal{OS}' \setminus \mathcal{OS}'[r], m' - 1, \mathcal{C} \cup \{r\}$ ) ▷ select remaining  $(m' - 1)$ 
       recipients
9:        $\mathcal{OS}' \leftarrow \bigcup_{\mathcal{O}'_l \in \mathcal{OS}'} \{\mathcal{O}'_l \setminus \{r\}\}$  ▷ remove  $r$  from all observ. in  $\mathcal{OS}'$ 
10:       $\mathcal{O}' \leftarrow \mathcal{O}' \setminus \{r\}$  ▷ do not choose  $r$  in this recursion level again

```

---

The computation of the minimal-hitting-sets is initially invoked by calling the algorithm  $ExactHS(\mathcal{OS}, m, \mathcal{C})$  and setting  $\mathcal{C} = \{\}$  and  $\mathcal{HS} = \{\}$ .

For ease of reference we address by the subscript  $i$  the state of a set, when entering the  $i$ -th level of recursion. Thus  $\mathcal{C}_i, \mathcal{OS}'_i$  are the sets prior to any modifications that could appear within level  $i$ , where  $\mathcal{C}_0 = \{\}$  and  $\mathcal{OS}'_0 = \mathcal{OS}$  are the sets in the initial call of ExactHS.

At each level  $i$  of recursion in the algorithm, jumping to the next level  $i + 1$  extends the current set of recipients  $\mathcal{C}_i$  by exactly one recipient,  $r$ , chosen at Line 7 in Algorithm 2. This recipient is element of a fixed observation  $\mathcal{O}' \in \mathcal{OS}'_i$ , determined by the algorithm in Line 5 in level  $i$ . The set of chosen recipients when entering the  $(i + 1)$ -th level of recursion is thus  $\mathcal{C}_{i+1} = \mathcal{C}_i \cup \{r\}$ .

$\mathcal{OS}'_{i+1}$ , determined at Line 8, results from removing all observations intersecting with  $r$  in  $\mathcal{OS}'_i$  and removing all recipients who have been previously evaluated in level  $i$  (within the while loop). That way, we focus in the  $(i + 1)$ -th recursion level only on those observations that are not hit by recipients in  $\mathcal{C}_{i+1}$ .

If, at Line 2, the algorithm detects that all remaining observations in  $\mathcal{OS}'_{i+1}$  intersect with  $\mathcal{C}_{i+1}$ ,  $\mathcal{C}_{i+1}$  is proven to be a hitting-set. This hitting-set is added to  $\mathcal{HS}$

and ExactHS immediately return back to level  $i$ , so that no set containing  $\mathcal{C}_{i+1}$  will be computed in the future<sup>1</sup>. Otherwise, Line 6 enables detecting the case that every set consisting of  $\mathcal{C}_{i+1}$  and any  $m - (i + 1)$  remaining recipients cannot hit all observations in  $\mathcal{OS}$ . This disproves all hypotheses containing  $\mathcal{C}_{i+1}$  that have not yet been evaluated by ExactHS, therefore ExactHS immediately returns back to level  $i$ , so that any set containing  $\mathcal{C}_{i+1}$  will be ignored in the future<sup>2</sup>. We refer to sets excluded by the algorithm, because they are proven to be minimal-hitting-sets, or to be no subset of any hitting-set by the term *finalised sets*.

After choosing and evaluating recipient  $r$  in recursion level  $i$ , ExactHS removes, at Line 9,  $r$  from all observations of  $\mathcal{OS}'_i$  and, at Line 10, from the designated observation  $\mathcal{O}'$ . Thus  $\mathcal{C}_i$  will not be extended by  $r$  in any succeeding recursion level again.

ExactHS stops choosing new recipients in level  $i$ , if Line 6 detects that all recipients in the designated observation  $\mathcal{O}'$  have been chosen (first condition), or if  $\mathcal{C}_i$  and any  $m - i$  remaining recipients cannot hit all observations in  $\mathcal{OS}$  (second condition).

### 2.2.1.1 Identification of Hitting-Sets – Examples

In order to support understanding Algorithm 2, we demonstrate the identification of all minimal-hitting-sets of at most size  $m$  in an observation set by the ExactHS algorithm in two examples. Example 2.2.1 shows ExactHS applied to an observation set that does not contain a unique minimum-hitting-set. Example 2.2.2 shows ExactHS applied to an observation set that contains a unique minimum-hitting-set.

In these examples, Alice's set of friends is  ${}_A\mathcal{H} = \{1, 2, 3\}$ , and the batch size of the Mix is  $b = 2$ . We demonstrate the application of Algorithm 2 on sequences of observations  $\mathcal{O}_1, \dots, \mathcal{O}_t$  collected by the attacker at distinct time points  $1 \leq t' \leq t$ , that are enlisted in Figure 2.3.

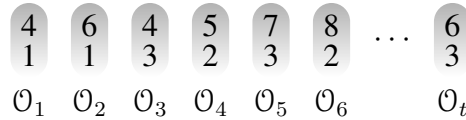


Figure 2.3: Sequence of collected observations from time point 1 to  $t$ .

<sup>1</sup>Although these sets would be hitting-sets in  $\mathcal{OS}$ , they would not be minimal.

<sup>2</sup>This avoids evaluating sets for that we know that they cannot be hitting-sets in  $\mathcal{OS}$ . The lower  $|\mathcal{C}|$  is, the more computation can be avoided.

## 2. COMBINATORIAL ATTACK

As a convention, we attach a numerical subscript  $i$  to a set in Algorithm 2 to denote the state of that set when entering the  $i$ -th level of recursion in Algorithm 2, where  $\mathcal{OS}'_0 = \mathcal{OS}, \mathcal{C}_0 = \{\}$  are the initial states. In each level  $i$  of recursion in Algorithm 2, an observation  $\mathcal{O}'_i \in \mathcal{OS}'_i$  is fixed and ExactHS chooses one of the two recipients in  $\mathcal{O}'_i$ . We number the choices alphabetically by  $a, b$ . The state of  $\mathcal{O}'_i$  when conducting the first choice in level  $i$  is labelled  $\mathcal{O}'_{i.a}$ .

We assume without loss of generality that Algorithm 2 evaluates observations in  $\mathcal{OS}'$  from left to right in Line 5. Similarly recipients in a set  $\mathcal{O}'$  are assumed to be evaluated from left to right in Line 7.

**Example 2.2.1** (Observation Set without Unique Minimum-Hitting-Set). *In this example, ExactHS is applied to identify all minimal-hitting-sets of at most size  $m = 3$  in the observation set  $\mathcal{OS} = \{\mathcal{O}_1, \dots, \mathcal{O}_5\}$ , that does not contain a unique minimum-hitting-set.*

$i.j$	$\mathcal{O}'_{i,j}$	$\mathcal{OS}'_{i,j}$	$\mathcal{C}_{i+1}$	$\mathcal{OS}'_{i+1}$	$\mathcal{HS}$
0.a	{4, 1}	{{4, 1}, {6, 1}, {4, 3}, {5, 2}, {7, 3}}	{4}	{{6, 1}, {5, 2}, {7, 3}}	{}
1.a	{6, 1}	{{6, 1}, {5, 2}, {7, 3}}	{4, -}	-	{}
0.b	{1}	{{1}, {6, 1}, {3}, {5, 2}, {7, 3}}	{1}	{{3}, {5, 2}, {7, 3}}	{}
1.a	{3}	{{3}, {5, 2}, {7, 3}}	{1, 3}	{{5, 2}}	{}
2.a	{5, 2}	{{5, 2}}	{1, 3, 5}	{{}}	{}
3.a	-	{{}}	-	-	{{1, 3, 5}}
2.b	{2}	{{2}}	{1, 3, 2}	{{}}	{{1, 3, 5}}
3.a	-	{{}}	-	-	{{1, 3, 5}, {1, 3, 2}}

Table 2.2: ExactHS applied on  $\mathcal{OS} = \{\mathcal{O}_1, \dots, \mathcal{O}_5\}$  that contains no unique minimum-hitting-set. The same line colour highlights the same level of recursion.

*Each line in Table 2.2 shows the state of sets in a level of recursion that result from ExactHS constructing a particular set  $\mathcal{C}$  that is suspected to be subset of a hitting-set in  $\mathcal{OS}$ . We explain Table 2.2 on a line by line basis with respect to steps in Algorithm 2, next.*

*0.a: This is the state, when Algorithm 2 is initially invoked by  $\text{ExactHS}(\mathcal{OS}, m, \mathcal{C})$ ,*

---

thus  $\mathcal{OS}'_{0.a} = \mathcal{OS}'_0 = \mathcal{OS}$  and  $\mathcal{C}_0 = \mathcal{C} = \{\}$  is the first state in level 0. Line 5 in Algorithm 2 fixes  $\mathcal{O}'_0 = \{4, 1\}$ , as it is the first observation in  $\mathcal{OS}'_0$  that contains a recipient with maximal frequency<sup>1</sup>. Line 6 in Algorithm 2 is fulfilled, as  $\{\} \notin \mathcal{OS}'_{0.a}$  and  $|\mathcal{OS}'_{0.a}[4]| + |\mathcal{OS}'_{0.a}[1]| + |\mathcal{OS}'_{0.a}[3]| \geq |\mathcal{OS}'_{0.a}|$ . Algorithm 2 thus chooses the first recipient 4 in  $\mathcal{O}'_{0.a}$  to extend  $\mathcal{C}$ , such that  $\mathcal{C}_1 = \{\} \cup \{4\}$  and  $\mathcal{OS}'_1 = \mathcal{OS}'_{0.a} \setminus \mathcal{OS}'_{0.a}[4]$ , when Algorithm 2 enters the next level of recursion.

1.a:  $\mathcal{O}'_1 = \{6, 1\}$  is fixed, as it is the first observation in  $\mathcal{OS}'_1$  that contains a recipient with maximal frequency. There are no recipients  $r_1, r_2$  in  $\mathcal{OS}'_{1.a}$ , such that  $|\mathcal{OS}'_{1.a}[r_1]| + |\mathcal{OS}'_{1.a}[r_2]| \geq |\mathcal{OS}'_{1.a}|$ . This proves that  $\mathcal{C}_1 = \{4\}$  is no subset of any hitting-set of size  $m = 3$  that has not been evaluated and prevents Algorithm 2 from choosing the next recipient. Instead, Algorithm 2 terminates this level and returns to level 0. ( $\{4\}$  is a finalised set.)

0.b: Due to Lines 9, 10 in Algorithm 2, recipient 4 that was previously chosen in level 0, is removed from all observations in  $\mathcal{OS}'_{0.a}$  and from  $\mathcal{O}'_{0.a}$ , resulting in  $\mathcal{OS}'_{0.b}$  and  $\mathcal{O}'_{0.b}$ . Only recipient 1 in  $\mathcal{O}'_{0.b}$  remains to be evaluated<sup>2</sup> in the current level. As  $\{\} \notin \mathcal{OS}'_{0.b}$  and  $|\mathcal{OS}'_{0.b}[1]| + |\mathcal{OS}'_{0.b}[3]| + |\mathcal{OS}'_{0.b}[6]| \geq |\mathcal{OS}'_{0.b}|$ , Algorithm 2 chooses recipient 1 to extend  $\mathcal{C}$ , such that  $\mathcal{C}_1 = \{\} \cup \{1\}$  and  $\mathcal{OS}'_1 = \mathcal{OS}'_{0.b} \setminus \mathcal{OS}'_{0.b}[1]$ , when Algorithm 2 enters the next level of recursion.

1.a:  $\mathcal{O}'_1 = \{3\}$  is fixed, as it is the first observation in  $\mathcal{OS}'_1$  that contains a recipient with maximal frequency. Only recipient 3 in  $\mathcal{O}'_{1.a}$  remains to be evaluated in the current level. As  $\{\} \notin \mathcal{OS}'_{1.a}$  and  $|\mathcal{OS}'_{1.a}[3]| + |\mathcal{OS}'_{1.a}[5]| \geq |\mathcal{OS}'_{1.a}|$ , Algorithm 2 chooses recipient 3 to extend  $\mathcal{C}$ , such that  $\mathcal{C}_2 = \{1\} \cup \{3\}$  and  $\mathcal{OS}'_2 = \mathcal{OS}'_{1.a} \setminus \mathcal{OS}'_{1.a}[3]$ , when Algorithm 2 enters the next level of recursion.

2.a:  $\mathcal{O}'_2 = \{5, 2\}$  is fixed and the only observation in  $\mathcal{OS}'_2$ . As  $\{\} \notin \mathcal{OS}'_{2.a}$  and  $|\mathcal{OS}'_{2.a}[5]| \geq |\mathcal{OS}'_{2.a}|$ , Algorithm 2 chooses recipient 5 to extend  $\mathcal{C}$ , such that  $\mathcal{C}_3 = \{1, 3\} \cup \{5\}$  and  $\mathcal{OS}'_3 = \mathcal{OS}'_{2.a} \setminus \mathcal{OS}'_{2.a}[5]$ , when Algorithm 2 enters the next level of recursion.

---

<sup>1</sup>We always refer to the frequencies of recipients in the observation set in the considered level of recursion.

<sup>2</sup>Note that  $\mathcal{O}'_0$  is fixed in Line 5 in Algorithm 2, therefore only recipients in  $\mathcal{O}'_0$  are evaluated in the 0-th level of recursion.

## 2. COMBINATORIAL ATTACK

---

3.a:  $\mathcal{O}'_3 = \{\}$ , thus  $\mathcal{C}_3 = \{1, 3, 5\}$  is a hitting-set and is added to  $\mathcal{HS}$ . After that Algorithm 2 terminates this level and returns to level 2. ( $\{1, 3, 5\}$  is a finalised set.)

2.b: Recipient 5, that was previously chosen in level 2, is removed from all observations in  $\mathcal{OS}'_{2,a}$  and from  $\mathcal{O}'_{2,a}$ , resulting in  $\mathcal{OS}'_{2,b}$  and  $\mathcal{O}'_{2,b}$ . As  $\{\} \notin \mathcal{OS}'_{2,b}$  and  $|\mathcal{OS}'_{2,b}[2]| \geq |\mathcal{OS}'_{2,b}|$ , Algorithm 2 chooses recipient 2 to extend  $\mathcal{C}$ , such that  $\mathcal{C}_3 = \{1, 3\} \cup \{2\}$  and  $\mathcal{OS}'_3 = \mathcal{OS}'_{2,b} \setminus \mathcal{OS}'_{2,b}[2]$ , when Algorithm 2 enters the next level of recursion.

3.a:  $\mathcal{O}'_3 = \{\}$ , thus  $\mathcal{C}_3 = \{1, 3, 2\}$  is a hitting-set and is added to  $\mathcal{HS}$ . After that Algorithm 2 terminates this level and returns to level 2. ( $\{1, 3, 2\}$  is a finalised set.)

Algorithm 2 terminates, when returning from the last level 3 above, as no recipients remain to be chosen in all levels below 3. The attacker can conclude from the hitting-sets  $\{1, 3, 5\}, \{1, 3, 2\}$  in  $\mathcal{HS}$  that there is no unique minimum-hitting-set in  $\mathcal{OS}$ . Thus more observations have to be collected to succeed the HS-attack.

**Example 2.2.2** (Observation Set with Unique Minimum-Hitting-Set). In this example, *ExactHS* is applied to identify all minimal-hitting-sets of at most size  $m = 3$  in the observation set  $\mathcal{OS} = \{\mathcal{O}_1, \dots, \mathcal{O}_6\}$ , that contains a unique minimum-hitting-set. In comparison to Example 2.2.1, the attacker collects one more observation, which is  $\mathcal{O}_6$ .

We explain Table 2.3 on a line by line basis with respect to steps in Algorithm 2, next.

0.a: This is the first state in level 0, thus  $\mathcal{OS}'_{0,a} = \mathcal{OS}'_0 = \mathcal{OS}$  and  $\mathcal{C}_0 = \{\}$ .  $\mathcal{O}'_0 = \{4, 1\}$  is fixed, as it is the first observation in  $\mathcal{OS}'_0$  that contains a recipient with maximal frequency. As  $\{\} \notin \mathcal{OS}'_{0,a}$  and  $|\mathcal{OS}'_{0,a}[4]| + |\mathcal{OS}'_{0,a}[1]| + |\mathcal{OS}'_{0,a}[3]| \geq |\mathcal{OS}'_{0,a}|$ , Algorithm 2 chooses the first recipient 4 in  $\mathcal{O}'_{0,a}$  to extend  $\mathcal{C}$ , such that  $\mathcal{C}_1 = \{\} \cup \{4\}$  and  $\mathcal{OS}'_1 = \mathcal{OS}'_{0,a} \setminus \mathcal{OS}'_{0,a}[4]$ , when Algorithm 2 enters the next level of recursion.

1.a:  $\mathcal{O}'_1 = \{5, 2\}$  is fixed, as it is the first observation in  $\mathcal{OS}'_1$  that contains a recipient with maximal frequency<sup>1</sup>. There are no recipients  $r_1, r_2$  in  $\mathcal{OS}'_{1,a}$ , such that

---

<sup>1</sup> $\mathcal{O}'_1$  is not the first observation in  $\mathcal{OS}'_1$ , but the first containing a recipient with maximal frequency.

---

$|\mathcal{OS}'_{1.a}[r_1]| + |\mathcal{OS}'_{1.a}[r_2]| \geq |\mathcal{OS}'_{1.a}|$ . This proves that  $\mathcal{C}_1 = \{4\}$  is no subset of any hitting-set of size  $m = 3$  that has not been evaluated and prevents Algorithm 2 from choosing the next recipient. Instead, Algorithm 2 terminates this level and returns to level 0. ( $\{4\}$  is a finalised set.)

0.b: Recipient 4, that was previously chosen in level 0, is removed from all observations in  $\mathcal{OS}'_{0.a}$  and from  $\mathcal{O}'_{0.a}$ , resulting in  $\mathcal{OS}'_{0.b}$  and  $\mathcal{O}'_{0.b}$ . Only recipient 1 in  $\mathcal{O}'_{0.b}$  remains to be evaluated in Algorithm 2 in the current level. As  $\{\}$   $\notin \mathcal{OS}'_{0.b}$  and  $|\mathcal{OS}'_{0.b}[1]| + |\mathcal{OS}'_{0.b}[3]| + |\mathcal{OS}'_{0.b}[2]| \geq |\mathcal{OS}'_{0.b}|$ , Algorithm 2 chooses recipient 1 to extend  $\mathcal{C}$ , such that  $\mathcal{C}_1 = \{\} \cup \{1\}$  and  $\mathcal{OS}'_1 = \mathcal{OS}'_{0.b} \setminus \mathcal{OS}'_{0.b}[1]$ , when Algorithm 2 enters the next level of recursion.

1.a:  $\mathcal{O}'_1 = \{3\}$  is fixed, as it is the first observation in  $\mathcal{OS}'_1$  that contains a recipient with maximal frequency. Only recipient 3 in  $\mathcal{O}'_{1.a}$  remains to be evaluated in the current level. As  $\{\}$   $\notin \mathcal{OS}'_{1.a}$  and  $|\mathcal{OS}'_{1.a}[3]| + |\mathcal{OS}'_{1.a}[2]| \geq |\mathcal{OS}'_{1.a}|$ , Algorithm 2 chooses recipient 3 to extend  $\mathcal{C}$ , such that  $\mathcal{C}_2 = \{1\} \cup \{3\}$  and  $\mathcal{OS}'_2 = \mathcal{OS}'_{1.a} \setminus \mathcal{OS}'_{1.a}[3]$ , when Algorithm 2 enters the next level of recursion.

2.a:  $\mathcal{O}'_2 = \{5, 2\}$  is fixed, as it is the first observation in  $\mathcal{OS}'_2$  that contains a recipient with maximal frequency. As  $\{\}$   $\notin \mathcal{OS}'_{2.a}$  and  $|\mathcal{OS}'_{2.a}[2]| \geq |\mathcal{OS}'_{2.a}|$ , Algorithm 2 chooses recipient<sup>1</sup> 2 to extend  $\mathcal{C}$ , such that  $\mathcal{C}_3 = \{1, 3\} \cup \{2\}$  and  $\mathcal{OS}'_3 = \mathcal{OS}'_{2.a} \setminus \mathcal{OS}'_{2.a}[2]$ , when Algorithm 2 enters the next level of recursion.

3.a:  $\mathcal{O}'_3 = \{\}$ , thus  $\mathcal{C}_3 = \{1, 3, 2\}$  is a hitting-set and is added to  $\mathcal{HS}$ . After that Algorithm 2 terminates this level and returns to level 2. ( $\{1, 3, 2\}$  is a finalised set.)

2.b: Recipient 2, that was previously chosen in level 2, is removed from all observations in  $\mathcal{OS}'_{2.a}$  and from  $\mathcal{O}'_{2.a}$ , resulting in  $\mathcal{OS}'_{2.b}$  and  $\mathcal{O}'_{2.b}$ . There is no recipient  $r$  in  $\mathcal{O}'_{2.b}$ , such that  $|\mathcal{OS}'_{2.b}[r]| \geq |\mathcal{OS}'_{2.b}|$ . This proves that  $\mathcal{C}_2 = \{1, 3\}$  is no subset of any hitting-set of size  $m = 3$  that has not been evaluated and prevents Algorithm 2 from choosing the next recipient. Instead, Algorithm 2 terminates this level and returns to level 1. ( $\{1, 3\}$  is a finalised set.)

---

<sup>1</sup>The frequency of recipient 2 is maximal, therefore it is the first recipient chosen from  $\mathcal{O}'_2$ .

## 2. COMBINATORIAL ATTACK

Algorithm 2 terminates, when returning from the last level 2 above, as no recipients remain to be chosen in all levels below 2. The attacker thus identifies the unique minimum-hitting-set  $\{1, 3, 2\}$  in  $\mathcal{HS}$ .

This example illustrates that the number of observations required by ExactHS to uniquely identify  ${}_A\mathcal{H}$  is the least possible. There is no unique minimum-hitting-set, if the attacker collects just one fewer observation, as shown in Example 2.2.1. Clearly,  ${}_A\mathcal{H}$  remains a unique minimum-hitting-set, even if more than the least number of observations are collected, if Alice does not change her set of friends.

$i.j$	$\mathcal{O}'_{i,j}$	$\mathcal{OS}'_{i,j}$	$\mathcal{C}_{i+1}$	$\mathcal{OS}'_{i+1}$	$\mathcal{HS}$
0.a	$\{4, 1\}$	$\{\{4, 1\}, \{6, 1\}, \{4, 3\}, \{5, 2\}, \{7, 3\}, \{8, 2\}\}$	$\{4\}$	$\{\{6, 1\}, \{5, 2\}, \{7, 3\}, \{8, 2\}\}$	$\{\}$
1.a	$\{5, 2\}$	$\{\{6, 1\}, \{5, 2\}, \{7, 3\}, \{8, 2\}\}$	$\{4, -\}$	$-$	$\{\}$
0.b	$\{1\}$	$\{\{1\}, \{6, 1\}, \{3\}, \{5, 2\}, \{7, 3\}, \{8, 2\}\}$	$\{1\}$	$\{\{3\}, \{5, 2\}, \{7, 3\}, \{8, 2\}\}$	$\{\}$
1.a	$\{3\}$	$\{\{3\}, \{5, 2\}, \{7, 3\}, \{8, 2\}\}$	$\{1, 3\}$	$\{\{5, 2\}, \{8, 2\}\}$	$\{\}$
2.a	$\{5, 2\}$	$\{\{5, 2\}, \{8, 2\}\}$	$\{1, 3, 2\}$	$\{\{\}\}$	$\{\}$
3.a	$-$	$\{\{\}\}$	$-$	$-$	$\{\{1, 3, 2\}\}$
2.b	$\{5\}$	$\{\{5\}, \{8\}\}$	$\{1, 3, -\}$	$-$	$\{\{1, 3, 2\}\}$

Table 2.3: ExactHS applied to  $\mathcal{OS} = \{\mathcal{O}_1, \dots, \mathcal{O}_6\}$ , the least observation set containing a unique minimum-hitting-set. The same line colour highlights the same level of recursion.

### 2.2.2 Soundness and Completeness

We prove that ExactHS is sound and complete with respect to the computation of minimal hitting-sets in the given observation set  $\mathcal{OS}$  and the maximal-hitting-set size  $m$ . *Soundness* means that ExactHS only computes hitting-sets of at most size  $m$  in  $\mathcal{OS}$ . *Completeness* means that ExactHS identifies all minimal-hitting-sets of at most size  $m$  in  $\mathcal{OS}$ .

Section 2.2.2.1 proofs properties of hitting-sets that are crucial for the proofs of soundness and completeness in Section 2.2.2.2 and Section 2.2.2.3.



---

### 2.2.2.1 Properties of Hitting-Sets

In this section, we show that the hitting- respectively non hitting-set property of a set  $\mathcal{H}$  in an observation set  $\mathcal{OS}$  is inherited by particular subsets of  $\mathcal{H}$  and subsets of  $\mathcal{OS}$  and constraints thereof. These inheritance relations are formulated in Claim 1 and Claim 2 and will be required in the proof of soundness and completeness of ExactHS.

**Claim 1.** *Let  $\mathcal{H} \subseteq R$  be any set of recipients and  $\mathcal{OS}$  a given set of observations. For every set  $\{r_1, \dots, r_k\} \subseteq R \setminus \mathcal{H}$  and  $k \in \mathbb{N}$ ,  $\mathcal{H}$  is a hitting-set in the observation set  $\mathcal{OS}$ , if and only if  $\mathcal{H}$  is a hitting-set in  $\mathcal{OS}' = \{\mathcal{O} \setminus \{r_1, \dots, r_k\} \mid \mathcal{O} \in \mathcal{OS}\}$ . If  $\mathcal{H}$  is a minimal-hitting-set, then this equivalence relation preserves that property.*<sup>1</sup>

**Claim 2.** *Let  $\mathcal{H} \subseteq R$  be any set of recipients and  $\mathcal{OS}$  be a given set of observations. For every set  $\mathcal{C} \subseteq \mathcal{H}$ ,  $\mathcal{H}$  is a hitting-set in  $\mathcal{OS}$  if and only if  $\mathcal{H} \setminus \mathcal{C}$  is a hitting-set in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$ . If  $\mathcal{H}$  is a minimal-hitting-set in  $\mathcal{OS}$ , then  $\mathcal{H} \setminus \mathcal{C}$  is a minimal-hitting-set in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$ .*<sup>2</sup>

By combining these two claims, we can deduce for every  $\mathcal{C} \subseteq \mathcal{H}$  and  $\{r_1, \dots, r_k\} \subseteq R \setminus \mathcal{H}$ , that  $\mathcal{H}$  is a hitting-set in  $\mathcal{OS}$ , if and only if  $\mathcal{H} \setminus \mathcal{C}$  is a hitting-set in  $\{\mathcal{O} \setminus \{r_1, \dots, r_k\} \mid \mathcal{O} \in \mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]\}$ . And, if  $\mathcal{H}$  is a minimal-hitting-set in  $\mathcal{OS}$ , then  $\mathcal{H} \setminus \mathcal{C}$  is a minimal-hitting-set in  $\{\mathcal{O} \setminus \{r_1, \dots, r_k\} \mid \mathcal{O} \in \mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]\}$ .

**Proof of Claim 1** The proof of Claim 1 is decomposed into two proofs. The first proof shows the “equivalence relation for the hitting-set property” in Claim 1, that is:

$$\begin{aligned} \forall \mathcal{H} \subseteq R \forall \{r_1, \dots, r_k\} \subseteq R \setminus \mathcal{H} : \mathcal{H} \text{ is hitting-set in } \mathcal{OS} \Leftrightarrow \\ \mathcal{H} \text{ is hitting-set in } \{\mathcal{O} \setminus \{r_1, \dots, r_k\} \mid \mathcal{O} \in \mathcal{OS}\} \end{aligned} \quad (2.1)$$

The second proof shows that the equivalence relation in (2.1) also applies to the minimal-hitting-set property, that is:

$$\begin{aligned} \forall \mathcal{H} \subseteq R \forall \{r_1, \dots, r_k\} \subseteq R \setminus \mathcal{H} : \mathcal{H} \text{ is minimal-hitting-set in } \mathcal{OS} \Leftrightarrow \\ \mathcal{H} \text{ is minimal-hitting-set in } \{\mathcal{O} \setminus \{r_1, \dots, r_k\} \mid \mathcal{O} \in \mathcal{OS}\} \end{aligned} \quad (2.2)$$

---

<sup>1</sup>If  $\mathcal{H}$  is no hitting-set in  $\mathcal{OS}$ , then it does not mean that there is no hitting-set in  $\mathcal{OS}$ . However, if  $\{\}$   $\in \mathcal{OS}$ , then there are no hitting-sets in  $\mathcal{OS}$ .

<sup>2</sup>We define that every set  $\mathcal{H} \subseteq R$  is a hitting-set in  $\mathcal{OS} = \{\}$ , including  $\mathcal{H} = \{\}$ .

## 2. COMBINATORIAL ATTACK

---

*Proof (first).* We prove the equivalence relation “ $\Leftrightarrow$ ” in (2.1) for each direction of the relation “ $\Leftarrow$ ” and “ $\Rightarrow$ ” separately. The proof of (2.1) is done by contradiction. Therefore, it is assumed that there are sets  $\mathcal{H}$ ,  $\{r_1, \dots, r_k\} \subseteq R \setminus \mathcal{H}$  and  $\mathcal{OS}$ , such that the relation that  $\mathcal{H}$  is a hitting-set in  $\mathcal{OS}$ , if and only if  $\mathcal{H}$  is a hitting-set in  $\mathcal{OS}' = \{\mathcal{O} \setminus \{r_1, \dots, r_k\} \mid \mathcal{O} \in \mathcal{OS}\}$  is invalid.

$\Rightarrow$ : Assume that there is a set  $\{r_1, \dots, r_k\} \subseteq R \setminus \mathcal{H}$ , such that  $\mathcal{H}$  is a hitting-set in  $\mathcal{OS}$  and no hitting-set in  $\mathcal{OS}' = \{\mathcal{O} \setminus \{r_1, \dots, r_k\} \mid \mathcal{O} \in \mathcal{OS}\}$ . Then there is an observation  $\mathcal{O}' \in \mathcal{OS}'$ , such that  $\mathcal{H} \cap \mathcal{O}' = \emptyset$ . By definition of  $\mathcal{OS}'$ , the observation  $\mathcal{O} = \mathcal{O}' \cup \{r_1, \dots, r_k\}$  must be in  $\mathcal{OS}$  and thus  $\mathcal{H} \cap \mathcal{O} = \emptyset$ . This is a contradiction to the initial assumption that  $\mathcal{H}$  is a hitting-set in  $\mathcal{OS}$  and thus proves that: If  $\mathcal{H}$  is a hitting-set in  $\mathcal{OS}$ , then it is a hitting-set in  $\mathcal{OS}'$ .

$\Leftarrow$ : Assume that there is a set  $\{r_1, \dots, r_k\} \subseteq R \setminus \mathcal{H}$ , such that  $\mathcal{H}$  is a hitting-set in  $\mathcal{OS}' = \{\mathcal{O} \setminus \{r_1, \dots, r_k\} \mid \mathcal{O} \in \mathcal{OS}\}$  and no hitting-set in  $\mathcal{OS}$ . Then for all  $\mathcal{O}' \in \mathcal{OS}'$ ,  $\mathcal{H} \cap \mathcal{O}' \neq \emptyset$ . Since for every  $\mathcal{O} \in \mathcal{OS}$ ,  $\mathcal{O} \supseteq \mathcal{O}'$ , we conclude that  $\mathcal{H} \cap \mathcal{O} \neq \emptyset$ . This is a contradiction to the initial assumption that  $\mathcal{H}$  is no hitting-set in  $\mathcal{OS}$  and thus proves that: If  $\mathcal{H}$  is a hitting-set in  $\mathcal{OS}'$ , then it is a hitting-set in  $\mathcal{OS}$ .

These two proofs, “ $\Rightarrow$ ” and “ $\Leftarrow$ ” complete the proof of (2.1). □

*Proof (second).* We prove the equivalence relation “ $\Leftrightarrow$ ” in (2.2) for each direction of the relation “ $\Leftarrow$ ” and “ $\Rightarrow$ ” separately. Let  $\mathcal{H} \subseteq R$ ,  $\{r_1, \dots, r_k\} \subseteq R \setminus \mathcal{H}$  be any set of recipients and  $\mathcal{OS}$  be any given observation set, where  $\mathcal{OS}' = \{\mathcal{O} \setminus \{r_1, \dots, r_k\} \mid \mathcal{O} \in \mathcal{OS}\}$ .

Recall from Section 2.1.2.3 that a minimal-hitting-set is a hitting-set in a given observation set, such that no proper subset of it is a hitting-set in that observation set.

$\Rightarrow$ : Let  $\mathcal{H}$  be a minimal-hitting-set in  $\mathcal{OS}$ , then it is due to (2.1) a hitting-set in  $\mathcal{OS}'$ . Since no subset  $\mathcal{H}' \subset \mathcal{H}$  is a hitting-set in  $\mathcal{OS}$ , expression (2.1) implies that no  $\mathcal{H}' \subset \mathcal{H}$  is a hitting-set in  $\mathcal{OS}'$ . Therefore  $\mathcal{H}$  is a minimal-hitting-set in  $\mathcal{OS}'$ .

$\Leftarrow$ : Let  $\mathcal{H}$  be a minimal-hitting-set in  $\mathcal{OS}'$ , then it is a hitting-set in  $\mathcal{OS}$ . Since no subset  $\mathcal{H}' \subset \mathcal{H}$  is a hitting-set in  $\mathcal{OS}'$ , expression (2.1) implies that no  $\mathcal{H}' \subset \mathcal{H}$  is a hitting-set in  $\mathcal{OS}$ . Therefore  $\mathcal{H}$  is a minimal-hitting-set in  $\mathcal{OS}$ .

---

These two proofs, “ $\Rightarrow$ ” and “ $\Leftarrow$ ” complete the proof of (2.2).  $\square$

**Proof of Claim 2** The proof of Claim 2 is decomposed into two proofs. The first proof shows the equivalence relation for the hitting-set property in Claim 2, that is:

$$\forall \mathcal{H} \in \mathcal{R} \forall \mathcal{C} \subseteq \mathcal{H} : \mathcal{H} \text{ is hitting-set in } \mathcal{OS} \Leftrightarrow \mathcal{H} \setminus \mathcal{C} \text{ is hitting-set in } \mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}] \quad (2.3)$$

The second proof shows that:

$$\begin{aligned} \forall \mathcal{H} \in \mathcal{R} \forall \mathcal{C} \subseteq \mathcal{H} : \mathcal{H} \text{ is minimal hitting-set in } \mathcal{OS} \Rightarrow \\ \mathcal{H} \setminus \mathcal{C} \text{ is minimal hitting-set in } \mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}] \end{aligned} \quad (2.4)$$

*Proof (first).* We prove the equivalence relation “ $\Leftrightarrow$ ” in (2.3) for each direction of the relation “ $\Leftarrow$ ” and “ $\Rightarrow$ ” separately. The proof of (2.3) is done by contradiction. Therefore, it is assumed that there are sets  $\mathcal{H}, \mathcal{C} \subseteq \mathcal{H}$  and  $\mathcal{OS}$ , such that the relation that  $\mathcal{H}$  is a hitting-set in  $\mathcal{OS}$ , if and only if  $\mathcal{H} \setminus \mathcal{C}$  is a hitting-set in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$  is invalid.

$\Rightarrow$ : Assume that there is a set  $\mathcal{H}$  and a non-empty set  $\mathcal{C} \subseteq \mathcal{H}$ , such that  $\mathcal{H}$  is a hitting-set in  $\mathcal{OS}$  and  $\mathcal{H} \setminus \mathcal{C}$  is no hitting-set in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$ . Therefore, there must be an observation  $\mathcal{O} \in \mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$ , such that  $(\mathcal{H} \setminus \mathcal{C}) \cap \mathcal{O} = \emptyset$ . However  $\mathcal{C}$  does not intersect any observation in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$  either, that is  $\forall \mathcal{O}' \in \mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}] : \mathcal{C} \cap \mathcal{O}' = \emptyset$ . This implies that  $\mathcal{H} \cap \mathcal{O} = \emptyset$ , which is a contradiction to the initial assumption that  $\mathcal{H}$  is a hitting-set in  $\mathcal{OS}$ . This proves that: If  $\mathcal{H}$  is a hitting-set in  $\mathcal{OS}$ , then  $\mathcal{H} \setminus \mathcal{C}$  is a hitting-set in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$ .

$\Leftarrow$ : Assume that there is a set  $\mathcal{H}$  and  $\mathcal{C} \subseteq \mathcal{H}$ , such that  $\mathcal{H} \setminus \mathcal{C}$  is a hitting-set in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$  and  $\mathcal{H}$  is no hitting-set in  $\mathcal{OS}$ . Note that  $\mathcal{C}$  is a hitting-set in  $\mathcal{OS}[\mathcal{C}]$  due to the definition of  $\mathcal{OS}[\mathcal{C}]$ , while  $\mathcal{H} \setminus \mathcal{C}$  is a hitting-set in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$  due to the initial assumption. These two facts imply that  $\mathcal{H}$  is a hitting-set in  $\mathcal{OS}$ , which is a contradiction to the initial assumption that  $\mathcal{H}$  is no hitting-set in  $\mathcal{OS}$  and proves that: If  $\mathcal{H} \setminus \mathcal{C}$  is a hitting-set in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$ , then  $\mathcal{H}$  is a hitting-set in  $\mathcal{OS}$ .

These two proofs “ $\Rightarrow$ ” and “ $\Leftarrow$ ” complete the proof of (2.3).  $\square$

## 2. COMBINATORIAL ATTACK

---

*Proof (second).* We now prove (2.4). Let  $\mathcal{H}$  be a minimal-hitting-set in  $\mathcal{OS}$ , we show that for every  $\mathcal{C} \subseteq \mathcal{H}$ ,  $\mathcal{H} \setminus \mathcal{C}$  is a minimal-hitting-set in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$ . This requires proving the following two conditions:

- For every  $\mathcal{C} \subseteq \mathcal{H}$ ,  $\mathcal{H} \setminus \mathcal{C}$  is a hitting-set in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$ .
- For every  $\mathcal{C} \subseteq \mathcal{H}$ , every proper subset  $\mathcal{H}_{\mathcal{C}} \subset \mathcal{H} \setminus \mathcal{C}$  is no hitting-set in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$ .

The first condition is already proven by (2.3), therefore we only prove the second condition. Let us define for convenience the term  $\mathcal{H}' = \mathcal{H}_{\mathcal{C}} \cup \mathcal{C}$ . Since  $\mathcal{H}' \subset \mathcal{H}$ , we conclude that  $\mathcal{H}'$  is no hitting-set<sup>1</sup> in  $\mathcal{OS}$ . The equivalence relation in (2.3) implies that for every  $\mathcal{C}' \subseteq \mathcal{H}'$ ,  $\mathcal{H}' \setminus \mathcal{C}'$  is no hitting-set in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}']$ . Using the substitution  $\mathcal{C}' = \mathcal{C}$ , this particularly proves that  $\mathcal{H}' \setminus \mathcal{C}$  is no hitting-set in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$  and completes the proof of (2.4).<sup>2</sup>  $\square$

### 2.2.2.2 Soundness

*Proof.* We prove the soundness of Algorithm 2 by contradiction and therefore assume that ExactHS adds a set  $\mathcal{H}$  to  $\mathcal{HS}$  that is no hitting-set in  $\mathcal{OS}$ , where  $|\mathcal{H}| = i \leq m$ . This means that there is an observation  $\mathcal{O} \in \mathcal{OS}$ , such that  $\mathcal{H} \cap \mathcal{O} = \emptyset$ .

The proof is based on the analysis of the trace of computing  $\mathcal{H} = \{r_1, \dots, r_i\}$  by ExactHS. To aid this, we define without loss of generality that  $\mathcal{C}_j = \{r_1, \dots, r_j\} \subseteq R$  is the set of distinct recipients when ExactHS enters the  $j$ -th level of recursion, where  $r_j$  is the recipient who was added when entering the  $j$ -th level of recursion on the trace of computing  $\mathcal{H}$ , for  $j \in \{1, \dots, i\}$ ,  $\mathcal{C}_0 = \{\}$  and  $\mathcal{C}_i = \mathcal{H}$ .

Observe that after choosing  $r_j$  in the  $(j - 1)$ -th level of recursion, Line 8 invokes the  $j$ -th level of recursion by submitting  $\mathcal{OS}'_j = \{\mathcal{O}' \setminus \{r_{j_1}, \dots, r_{j_k}\} \mid \mathcal{O}' \in \mathcal{OS}'_j \setminus \mathcal{OS}'_j[r_j], \{r_{j_1}, \dots, r_{j_k}\} \subseteq R \setminus \mathcal{H}\}$ , where  $r_{j_1}, \dots, r_{j_k}$  are any recipients evaluated and removed<sup>3</sup> in the while loop in the  $(j - 1)$ -th level of recursion<sup>4</sup> prior to choosing  $r_j$ . According to Claim 2,  $\mathcal{H}$  is no hitting-set in  $\mathcal{OS}$ , if and only if  $\mathcal{H} \setminus \{r_1, \dots, r_j\}$  is no

<sup>1</sup>Because  $\mathcal{H}$  is a minimal-hitting-set in  $\mathcal{OS}$ , no proper subset of  $\mathcal{H}$  must be a hitting-set in  $\mathcal{OS}$ .

<sup>2</sup>It is possible that  $\mathcal{H} \setminus \mathcal{C}$  is minimal-hitting-set in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$ , whereas  $\mathcal{H}$  is a hitting-set, but no minimal-hitting-set in  $\mathcal{OS}$ . Therefore (2.4) contains no equivalence relation for the minimal-hitting-set property.

<sup>3</sup>The index  $k$  refers to distinct values in different levels of recursion and is reused for convenience.

<sup>4</sup>Since we only trace the case, when ExactHS computes  $\mathcal{H}$ , no  $r \in \mathcal{H}$  must be removed on that trace, therefore  $\{r_{j_1}, \dots, r_{j_k}\} \cap \mathcal{H} = \emptyset$  in all level  $j \in \{0, \dots, i\}$  of recursion.

---

hitting-set in  $\mathcal{OS} \setminus \{r_1, \dots, r_j\}$  for  $j \in \{1, \dots, i\}$ . By combining this with Claim 1, we deduce from the transitivity of the equivalence relation, that  $\mathcal{H} \setminus \{r_1, \dots, r_j\}$  is no hitting-set in  $\mathcal{OS} \setminus \{r_1, \dots, r_j\}$ , if and only if  $\mathcal{H} \setminus \{r_1, \dots, r_j\}$  is no hitting-set in  $\mathcal{OS}'_j$  for  $j \in \{1, \dots, i\}$ .

From this and our initial assumption that  $\mathcal{H}$  is no hitting-set, we conclude that when ExactHS enters the  $i$ -th level of recursion that is also the last level,  $\mathcal{C}_i \setminus \{r_1, \dots, r_i\}$  is no hitting-set in  $\mathcal{OS}'_i$ . Therefore  $\mathcal{OS}'_i \neq \{\}$  in Line 2 in Algorithm 2, so that  $\mathcal{C}_i$  and thus  $\mathcal{H}$  cannot be added to  $\mathcal{HS}$ . This contradicts the initial assumption that ExactHS would accept  $\mathcal{H}$  and thus proves that ExactHS is sound and only adds hitting-sets to  $\mathcal{HS}$ .  $\square$

### 2.2.2.3 Completeness

*Proof.* We prove the completeness of Algorithm 2 by induction on the size of minimal-hitting-sets. It will be shown that ExactHS identifies all minimal-hitting-sets of any given size in any given observation set  $\mathcal{OS}$ .

Superscripts and subscripts attached to sets will have following meanings: The superscript in  $\mathcal{OS}^i$  and  $\mathcal{H}^i$  annotate that there is a minimal-hitting-set of size  $i$  in  $\mathcal{OS}^i$  and that  $\mathcal{H}^i$  is one of that set. We address by the subscript  $j$  the state of a set, when entering the  $j$ -th level of recursion.

Ind. basis: Let  $i = |\mathcal{H}^1| = 1$  and  $\mathcal{H}^1$  be a minimal-hitting-set in the observation set  $\mathcal{OS}^1$ . We prove that the invocation of  $ExactHS(\mathcal{OS}, 1, \mathcal{C})$ , where  $\mathcal{OS} = \mathcal{OS}^1$ ,  $\mathcal{C} = \{\}$  will identify<sup>1</sup>  $\mathcal{H}^1$  as a hitting-set in  $\mathcal{OS}^1$ .

Since  $\mathcal{H}^1$  is a minimal-hitting-set in  $\mathcal{OS}^1$ , every observation determined in Line 5 of Algorithm 2 must contain  $r \in \mathcal{H}^1$  in the 0-th level of recursion. Let  $\mathcal{O}_0$  be the observation fixed in Line 5 in the 0-th level of recursion. Without loss of generality,  $r_{1_1}, \dots, r_{1_k} \in R \setminus \mathcal{H}^1$  are any  $k \geq 0$  recipients<sup>2</sup> evaluated and removed within the while-loop in Algorithm 2 before ExactHS chooses  $r \in \mathcal{O}_0 \cap \mathcal{H}^1$  in Line 7. In that case, Line 8 invokes the next level of recursion, such that  $\mathcal{OS}'_1 = \{\mathcal{O}' \setminus \{r_{1_1}, \dots, r_{1_k}\} \mid \mathcal{O}' \in \mathcal{OS}'_0 \setminus \mathcal{OS}'_0[r]\}$  is the observation set, when entering the 1-st level of recursion. Since every observation  $\mathcal{O} \in \mathcal{OS}^1$  contains

---

<sup>1</sup>If hitting-set is identified by ExactHS, then that set is added to the set of identified hitting-sets  $\mathcal{HS}$ .

<sup>2</sup>The index  $k$  refers to distinct values in different levels of recursion and is reused for convenience.

## 2. COMBINATORIAL ATTACK

---

$r$ , we conclude that  $\mathcal{OS}'_1 = \{\}$ . This will be detected in Line 2 in level 1 and  $\mathcal{C}_1 = \{r\} = \mathcal{H}^1$  will be added as a hitting set to  $\mathcal{HS}$ .<sup>1</sup>

Ind. step: Assume for  $i = |\mathcal{H}^i| \geq 1$ , where  $\mathcal{H}^i$  is a minimal-hitting-set in the observation set  $\mathcal{OS}^i$ , that the invocation of  $ExactHS(\mathcal{OS}, i, \mathcal{C})$ , where  $\mathcal{OS} = \mathcal{OS}^i$ ,  $\mathcal{C} = \{\}$ , will identify  $\mathcal{H}^i$  as a hitting-set. We prove that this assumption also applies to the case of  $i + 1 = |\mathcal{H}^{i+1}|$ , where  $\mathcal{H}^{i+1}$  is a minimal-hitting-set in the observation set  $\mathcal{OS}^{i+1}$ .

Since  $\mathcal{H}^{i+1}$  is a minimal-hitting-set in  $\mathcal{OS}^{i+1}$ , every observation determined in Line 5 of Algorithm 2 must contain any  $r \in \mathcal{H}^{i+1}$  in the 0-th level of recursion. Let  $\mathcal{O}_0$  be the observation fixed in Line 5 in level 0 and  $r_{1_1}, \dots, r_{1_k} \in R \setminus \mathcal{H}^{i+1}$  be any recipients evaluated and removed prior to choosing  $r \in \mathcal{O}_0 \cap \mathcal{H}^{i+1}$  in Line 7. In that case, Line 8 invokes the next level of recursion, such that  $\mathcal{OS}'_1 = \{\mathcal{O}' \setminus \{r_{1_1}, \dots, r_{1_k}\} \mid \mathcal{O}' \in \mathcal{OS}_0' \setminus \mathcal{OS}'_0[r]\}$  is the observation set, when entering the 1-st level of recursion. Due to Claim 1 and Claim 2,  $\mathcal{H}^{i+1} \setminus \{r\}$  is a minimal-hitting-set of size  $i$  in  $\mathcal{OS}'_1$ . According to the induction assumption, any minimal-hitting set of size  $|\mathcal{H}^{i+1} \setminus \{r\}| = i$  in  $\mathcal{OS}'_1$  will be identified by  $ExactHS$ . Therefore, the recursive invocation of  $ExactHS(\mathcal{OS}'_1, i, \{r\})$  in Line 8 in level 0 will identify the minimal-hitting-set  $\mathcal{H}^{i+1}$  in  $\mathcal{OS}^{i+1}$ .

We conclude from the induction steps that for every observation set  $\mathcal{OS}$  and minimal-hitting-set  $\mathcal{H}^i$  of size  $i$  in  $\mathcal{OS}$ , the invocation of  $ExactHS(\mathcal{OS}, i, \mathcal{C})$ , where  $\mathcal{C} = \{\}$ , will identify  $\mathcal{H}^i$ . If invoking  $ExactHS(\mathcal{OS}, i, \mathcal{C})$  identifies  $\mathcal{H}^i$ , then invoking  $ExactHS(\mathcal{OS}, i, \mathcal{C})$  also identifies  $\mathcal{H}^i$ . This and the induction steps proves that the invocation of  $ExactHS(\mathcal{OS}, m, \mathcal{C})$  identifies all minimal-hitting-sets  $\mathcal{H}$  in  $\mathcal{OS}$ , where  $|\mathcal{H}| \leq m$  and thus proves the completeness of  $ExactHS$ .  $\square$

Note that given  $m$  and  $\mathcal{OS}$ , the above proof does not exclude, that  $ExactHS$  might also add hitting-sets of at most size  $m$  to  $\mathcal{HS}$  that are not minimal in  $\mathcal{OS}$ . However, even in special cases, where all hitting-sets of at most size  $m$  in  $\mathcal{OS}$  are minimal-hitting-sets, such that all hitting-sets added to  $\mathcal{HS}$  by  $ExactHS$  are inevitably minimal, the maximal

---

<sup>1</sup> $ExactHS$  identifies  $\mathcal{H}^1$  already in the 1-st level of recursion, thus invoking  $ExactHS(\mathcal{OS}, m, \{\})$  for  $m \geq 1$  would provide a different result.

---

value of  $|\mathcal{HS}|$ <sup>1</sup> would be identical to that in the general case<sup>2</sup>. Section 2.2.3.1 proves this equality and shows that the maximal value of  $|\mathcal{HS}|$  primarily determines the worst case complexity of ExactHS. Therefore, we do not care about excluding non minimal-hitting-sets in  $\mathcal{HS}$ .

## 2.2.3 Worst Case Complexity

This section derives the worst case time and worst case space-complexity of ExactHS with respect to the parameters  $u, b, m$  in the Mix system and the number  $t$  of observations collected by the attacker. We prove that the worst case time-complexity of ExactHS is  $O(b^m t b m)$ , while its time-complexity is  $O((m+1)tb)$ .<sup>3</sup> As will be proven in Section 2.2.3.1, the maximal number of minimal-hitting-sets of at most size  $m$  is  $b^m$ , so that the worst case time-complexity of ExactHS is minimal except for the minor factor  $(tbm)$  that is the size of the input (i.e., the observation set) to ExactHS.

### 2.2.3.1 Time-Complexity

Finalised sets are those sets  $\mathcal{C}$  in Algorithm 2 in Section 2.2.1 that ExactHS has proved to be a hitting-set, or to be no subset of any hitting-set that it has not yet computed, during its computation of hitting-sets of at most size  $m$  in  $\mathcal{OS}$ . The time-complexity of ExactHS is primarily determined by the computation of these finalised sets.

**Claim 3.** *For given parameters  $u, b, m$  in the Mix system, where  $u \geq bm$ , the maximal number of finalised sets computed by ExactHS is  $b^m$  and is a tight maximum of the number of minimal-hitting-sets of at most size  $m$ .*

**Claim 4.** *For given parameters  $u, b, m$  in the Mix system, where  $u \geq bm$ , the worst case time-complexity of ExactHS is  $O(b^m t b m)$ .*

*Proof (Claim 3).* ExactHS is initially invoked by calling  $ExactHS(\mathcal{OS}, m, \mathcal{C})$ , where  $\mathcal{C} = \{\}$ . Since  $\mathcal{C}$  is extended by exactly one recipient in each level of recursion<sup>4</sup>. In

---

<sup>1</sup>That is the maximal number of minimal-hitting-sets of at most size  $m$ .

<sup>2</sup>That is the maximal number of hitting-sets of at most size  $m$  added to  $\mathcal{HS}$  by ExactHS, regardless whether they are minimal, or not.

<sup>3</sup>Note that the worst case complexity of ExactHS is invariant to  $u$ .

<sup>4</sup>The level of executing  $ExactHS(\mathcal{OS}, m, \mathcal{C})$  is counted as level 0. ExactHS extends  $\mathcal{C}$  by the  $i$ -th recipient, when entering the  $i$ -th level of recursion, where  $\mathcal{C} = \{\}$  for  $i = 0$ .

## 2. COMBINATORIAL ATTACK

---

every level  $i$  of recursion, ExactHS chooses in Line 7 one of the  $b$  recipients  $r_{i+1}$  in an observation  $\mathcal{O}_i$ , that was fixed at Line 5, to extend  $\mathcal{C}$  at the  $(i + 1)$ -th level of recursion.

The bound  $m$  for the number of successive recursive invocations to choose a recipient to extend  $\mathcal{C}$ , and the bound  $b$  for the number of choices of recipients in each level of recursion, proves that ExactHS computes at most  $b^m$  finalised sets.

To prove that  $b^m$  is the tight maximum of the number minimal-hitting-set of at most size  $m$ , it is sufficient to show that a set of observations  $\mathcal{OS}$  exists, such that ExactHS applied to  $\mathcal{OS}$  computes exactly  $b^m$  distinct minimal-hitting-sets<sup>1</sup>. Consider the set of  $m$  pairwise disjoint observations  $\mathcal{OS} = \{\mathcal{O}_0, \dots, \mathcal{O}_{m-1}\}$ ,<sup>2</sup> where  $|\mathcal{O}_i| = b$  for  $0 \leq i \leq m - 1$ . We assume without loss of generality, that  $\mathcal{O}_i \in \mathcal{OS}$  is fixed in the  $i$ -th level of recursion of ExactHS. The finalised sets computed by ExactHS thus have the structure  $\mathcal{C} = \{r_1, \dots, r_m\}$ , where  $r_j \in \mathcal{O}_{j-1}$  for  $1 \leq j \leq m$  is a recipient who ExactHS chooses in the  $(j - 1)$ -th level of recursion. In this case, all  $b^m$  finalised sets  $\mathcal{C}$  are minimal-hitting-sets, proving that the maximum of the number of minimal-hitting-sets is exactly  $b^m$ .  $\square$

*Proof (Claim 4).* As a first step, the proof shows that computing a single finalised set  $\mathcal{H}$  requires at most  $O(tbm)$  operations in ExactHS, as represented by Algorithm 2. We assume for simplicity and without loss of generality for the proof, that  $\mathcal{H} = \{r_1, \dots, r_m\} = \mathcal{C}_m$ , where  $r_i$  is the first recipient who ExactHS chooses in Line 7 in the  $(i - 1)$ -th level of recursion, for  $1 \leq i \leq m$ .

When ExactHS chooses  $r_i$  in the  $(i - 1)$ -th level of recursion, invoking the next level of recursion by calling *ExactHS*( $\mathcal{OS}_i, m - i, \mathcal{C}_{i-1} \cup \{r_i\}$ ) in Line 8, where  $\mathcal{OS}_i = \mathcal{OS}_{i-1} \setminus \mathcal{OS}_{i-1}[r_i]$  requires removing all observations containing  $r_i$  from  $\mathcal{OS}_{i-1}$ . Since  $|\mathcal{OS}_{i-1}| \leq |\mathcal{OS}| = t$ , we have to evaluate at most  $t$  observations for containment of  $r_i$ . Evaluating containment of  $r_i$  in a single observation requires comparison of  $r_i$  to at most  $b$  recipients in that observation. Thus extending  $\mathcal{C}_{i-1}$  by a single recipient  $r_i$  requires  $O(tb)$  operations. Consequently, computing a single set  $\mathcal{H}$  requires  $O(tbm)$  operations.

---

<sup>1</sup>Remember that ExactHS is complete (see Section 2.2.2.3) and thus identifies all minimal-hitting-sets.

<sup>2</sup>In context of the observations collected by the attacker,  $\mathcal{O}_i$  denotes the  $i$ -th collected observation, for  $i > 0$ . To ease references, we use in this proof a deviant notation, where  $\mathcal{O}_{i-1}$  is the  $i$ -th collected observation.



---

Combining this result with the maximal number of finalised sets in Claim 3 proves that the worst case time-complexity of ExactHS is  $O(b^m t b m)$ . A similar proof was provided by Pham and Kesdogan [2009]; Pham [2008].  $\square$

**Example 2.2.3** (Maximal Number of Minimal-Hitting-Sets). *This example illustrates a case, where ExactHS computes the maximal number of  $b^m$  minimal-hitting-sets and thus requires the worst case time-complexity of  $O(b^m t b m)$ . In that case, the attacker collects a set of observations  $\{\mathcal{O}_0, \dots, \mathcal{O}_{m-1}\}$ , where all observations are pairwise disjoint.*

*Let the set of all recipients in this example be  $R = \{1, \dots, 12\}$ , where  $b = 2$  is the batch size of the Mix. Let Alice's set of friends be  ${}_A\mathcal{H} = \{1, 2, 3\}$ .*

*The information available to the attacker are the following collected observations:*

$$\{1, 4\}, \{2, 5\}, \{3, 6\}$$

*By applying ExactHS on this set of observations for  $m = 3$ , we obtain  $b^m = 8$  minimal-hitting-sets, which is the maximal number of minimal-hitting-sets of at most size 3 in any observation set. These minimal-hitting-sets have the structure  $\mathcal{H} = \{r_1, \dots, r_m\}$ , where  $r_j \in \mathcal{O}_{j-1}$  for  $1 \leq j \leq m$ . They are in this example the following minimal-hitting-sets:*

$$\{1, 2, 3\}, \{1, 2, 6\}, \{1, 5, 3\}, \{1, 5, 6\}, \{4, 2, 3\}, \{4, 2, 6\}, \{4, 5, 3\}, \{4, 5, 6\}$$

*Note that the number of minimal-hitting-sets of at most size  $m$  never exceeds  $b^m$ , even if there are more observations, and that this number does not depend on the number of recipients  $u$ .*

**Non Triviality of ExactHS** We illustrate that providing an algorithm that computes at most  $b^m$  hitting-sets, while solving the UMHS problem is not trivial by an example.

In essence, a hitting-set is a set of recipients, each is picked from a distinct observation, and every observation must have one recipient in the hitting-set. Let us solve the unique minimum-hitting-set problem from this simple point of view. That is we compute the combination of the recipients in distinct observations.

## 2. COMBINATORIAL ATTACK

**Example 2.2.4** (Naive Computation of Hitting-Sets). *Let Alice's set of friends be  ${}_A\mathcal{H} = \{1, 2, 3\}$ ,  $b = 2$  and  $R = \{1, \dots, 12\}$  as in Example 2.2.3. Assume that the observations collected by the attacker in  $\mathcal{OS}$  are*

<div style="border: 1px solid black; border-radius: 50%; padding: 5px; display: inline-block; text-align: center;">4 1</div>	<div style="border: 1px solid black; border-radius: 50%; padding: 5px; display: inline-block; text-align: center;">5 2</div>	<div style="border: 1px solid black; border-radius: 50%; padding: 5px; display: inline-block; text-align: center;">6 1</div>	<div style="border: 1px solid black; border-radius: 50%; padding: 5px; display: inline-block; text-align: center;">7 2</div>	<div style="border: 1px solid black; border-radius: 50%; padding: 5px; display: inline-block; text-align: center;">8 1</div>	<div style="border: 1px solid black; border-radius: 50%; padding: 5px; display: inline-block; text-align: center;">9 2</div>	<div style="border: 1px solid black; border-radius: 50%; padding: 5px; display: inline-block; text-align: center;">10 1</div>	<div style="border: 1px solid black; border-radius: 50%; padding: 5px; display: inline-block; text-align: center;">11 2</div>	<div style="border: 1px solid black; border-radius: 50%; padding: 5px; display: inline-block; text-align: center;">3 1</div>
$\mathcal{O}_1$	$\mathcal{O}_2$	$\mathcal{O}_4$	$\mathcal{O}_3$	$\mathcal{O}_5$	$\mathcal{O}_6$	$\mathcal{O}_7$	$\mathcal{O}_8$	$\mathcal{O}_9$

Figure 2.4: Sequence of observations collected by attacker.

*The computation of the hitting-sets by combination of recipients from each observation is represented in Table 2.4. It shows the computed hitting-sets for distinct sequences of observations in Figure 2.4.*

$i$	$\mathcal{O}_i$	Hypotheses
1	$\{1, 4\}$	$\{1\}, \{4\}$
2	$\{2, 5\}$	$\{1, 2\}, \{1, 5\}, \{4, 2\}, \{4, 5\}$
3	$\{1, 6\}$	$\{1, 2\}, \{1, 2, 6\}, \{1, 5\}, \{1, 5, 6\}, \{4, 2, 6\}, \{4, 5, 6\}$
4	$\{2, 7\}$	$\{1, 2\}, \{1, 2, 7\}, \{1, 2, 6\}, \{1, 2, 5\}, \{1, 5, 7\}, \{4, 2, 6\}$
...	...	...
9	$\{1, 3\}$	$\{1, 2\}, \{1, 2, 3\}, \{1, 2, 4\}, \{1, 2, 5\}, \{1, 2, 6\}, \{1, 2, 7\}, \{1, 2, 8\}, \{1, 2, 9\}, \{1, 2, 10\}, \{1, 2, 11\},$

Table 2.4: Naive Computation of hitting-sets of at most size  $m$  by combination of recipients from distinct observations.

*The maximal number of minimal-hitting-sets of at most size  $m$  is due to Claim 3  $b^m$ , which is in this example  $2^3 = 8$ . However considering all hitting-sets of at most size 3 resulting from any combination of the 9 observations will lead to at least 9 hitting-sets. These are at least the sets  $\mathcal{H} = \{1, 2, x\}$  for  $x \in \{3, \dots, 11\}$ .*

The structure of observations depicted in Example 2.2.4 straight forwardly applies to any larger number of hitting-sets, for  $m = 3, b = 2$ . We can also extend the structure for arbitrary values of  $b$  and  $m$ . This thus demonstrates that finding an algorithm like the ExactHS that only evaluates  $b^m$  finalised sets to solve the UMHS problem is not trivial.

---

### 2.2.3.2 Space-Complexity

We observe that in each level  $i$  of recursion, Algorithm 2 keeps a copy of the actual observation set  $\mathcal{OS}_i$  in the memory, for  $0 \leq i \leq m$ . Therefore, there are at most  $m + 1$  observation sets that simultaneously allocate memory. This implies, cf. Pham [2008], that the worst case space-complexity of Algorithm 2 is bounded by:

$$O((m + 1)tb), \quad (2.5)$$

where  $t = |\mathcal{OS}|$  is the number of observations collected by the attacker and  $tb$  is the maximal number of recipients in  $\mathcal{OS}$ . Hence ExactHS is a linear space algorithm.

Note that we do not account the size of the set  $\mathcal{HS}$  in Algorithm 2, since the hitting-sets in  $\mathcal{HS}$  are not an integral part of Algorithm 2. We can consider  $\mathcal{HS}$  as a system output. Even deleting  $\mathcal{HS}$  would have no effect on the continuative execution of Algorithm 2.

Also note that the worst case space-complexity of the original HS-algorithm as presented by Kesdogan and Pimenidis [2004] is at least  $O(\binom{u}{m})$ . That is because the algorithm requires computing the initial set of all  $\binom{u}{m}$  hypotheses and then excludes those hypotheses that are no hitting-sets in the sequence of observations collected by the attacker. In general, as illustrated in Example 2.1.1, the number of specified hypotheses can significantly exceed the number of minimal-hitting-sets, so that algorithms based on computing all specified hypotheses lead to a higher complexity than computing all minimal-hitting-sets.

### 2.2.4 Evaluation

We apply the (enhanced) ExactHS, as implemented in Algorithm 2 to randomly generated observations. To illustrate the efficiency of the HS-attack using ExactHS for practical cases, we consider the empirical least number of observations to succeed the HS-attack and the time-complexity of ExactHS for a range of Mix parameters  $u, b, m$ . These include Mix parameters that are infeasible for the original HS-algorithm, as well as those that are also infeasible with respect to the worst case time-complexity of ExactHS. The worst case complexity is determined by the number of finalised sets, as enlisted in Table 2.5 for some Mix parameters  $u, b, m$  considered in the simulations.

## 2. COMBINATORIAL ATTACK

---

The delay of the traffic relayed by the Mix increases with the number of message it has to collect for a batch, as analysed by Kesdogan and Palmer [2006]. Therefore, we only consider batch sizes  $b \leq 85$  that provide reasonable practical delays specified by Kesdogan and Palmer [2006].

$u$	$b$	$m$	$\binom{u}{m}$	$b^m$
100	10	10	$1.7 \times 10^{13}$	$1.0 \times 10^{10}$
400	10	10	$2.6 \times 10^{19}$	$1.0 \times 10^{10}$
400	40	10	$2.6 \times 10^{19}$	$1.0 \times 10^{16}$
400	10	23	$1.4 \times 10^{37}$	$1.0 \times 10^{23}$
10000	50	20	$4.0 \times 10^{61}$	$9.5 \times 10^{33}$
20000	50	20	$4.3 \times 10^{67}$	$9.5 \times 10^{33}$
20000	85	20	$4.3 \times 10^{67}$	$3.9 \times 10^{38}$
20000	50	40	$1.3 \times 10^{124}$	$9.1 \times 10^{67}$

Table 2.5: Worst case number of finalised sets: HS-algorithm  $\binom{u}{m}$  versus ExactHS  $b^m$ .

Alice’s communication traffic is modelled by a Zipf distribution that is known to closely model e-mail and internet traffic, cf. Adamic and Huberman [2002]; Almeida et al. [1996]; Breslau et al. [1999]; Glassman [1994]. This reveals the impact of Alice’s communication on the feasibility of ExactHS in practical cases. Each observation contains an Alice’s friend that is drawn from a  $\text{Zipf}(m, \alpha)$  distribution and  $(b - 1)$  recipients of cover-traffic that are drawn uniformly from the set of  $|R| = u$  possible recipients. We model for the sake of simplicity the cover-traffic by a uniform distribution that bound various cover-traffic distributions of interest, instead of analysing those distributions separately.

The HS-attack is *successful* (or *succeeds*) if ExactHS can uniquely identify Alice’s set of friends  ${}_A\mathcal{H}$ . The simulation generates new random observations until the HS-attack is successful and we call this an *experiment*. The average number of observations required by an attack is therefore the mean of the number of observations of all successful attacks (i.e., of all experiments). To ensure that our results are statistically significant, experiments are repeated until 95% of the results fall within 5% of the empirically observed mean. Every experiment is repeated at least 300 times.

---

### 2.2.4.1 Communication Traffic

The distribution of Alice's communication traffic and that of the cover-traffic result from the probability distribution of the recipients of each sender. We model the distribution of a sender by the Zipf distribution as it allows an easy formulation of non-uniform distributions and closely models e-mail and Internet traffic distributions, cf. Adamic and Huberman [2002]; Almeida et al. [1996]; Breslau et al. [1999]; Glassman [1994].

**Zipf Distribution** Let  $Y$  be a random variable with values in the discrete state space  $\Omega = \{1, \dots, v\}$ ,  $|\Omega| = v$ .  $Y$  is  $\text{Zipf}(v, \alpha)$  distributed, if its probability mass function  $P_z^{v, \alpha}$  and cumulative distribution function  $F_z^{v, \alpha}$  are

$$P_z^{v, \alpha}(Y = i) = \frac{i^{-\alpha}}{\sum_{l=1}^v l^{-\alpha}},$$

$$F_z^{v, \alpha}(Y = i) = \frac{\sum_{k=1}^i k^{-\alpha}}{\sum_{l=1}^v l^{-\alpha}},$$

for all  $i \in \Omega$ .

Let  $U = \{r_1, \dots, r_v\}$  be the ordered set of all recipients of a sender and  $|\Omega| = |U|$ , where for all  $i \in \Omega$ ,  $r_i \in U$  is the recipient who is  $i$ -th most frequently contacted by that sender. If for all  $i \in \Omega$ , the probability that this sender contacts his  $i$ -th most frequently contacted recipient  $r_i \in U$  is  $P_z^{v, \alpha}(i)$ , then we say that his recipients are *Zipf( $v, \alpha$ ) distributed*, or that he contacts his recipients according to the *Zipf( $v, \alpha$ ) distribution*. In case that  $\alpha = 0$  in that probability distribution, we say that the sender's recipients are *uniformly distributed*, or that the sender contacts his recipients according to a *uniform distribution*. That is  $P_z^{v, 0}(i) = \frac{1}{v}$  for  $i \in \Omega$  and  $r_i \in U$ . We generally call all distributions that are not uniform, *non-uniform* distributions.

For example, let  ${}_A\mathcal{H} = \{a_1, \dots, a_m\}$  be Alice's set of friends ordered by the frequency of being contacted by Alice and  $\Omega = \{1, \dots, m\}$  and  $|\Omega| = |{}_A\mathcal{H}|$ . Alice's friends are  $\text{Zipf}(m, \alpha)$ -distributed, if the probability that she contacts her  $i$ -th most frequently contacted friend is  $P_z^{m, \alpha}(i)$ , for  $i \in \Omega$  and  $a_i \in {}_A\mathcal{H}$ . This distribution is illustrated in Figure 2.5.

## 2. COMBINATORIAL ATTACK

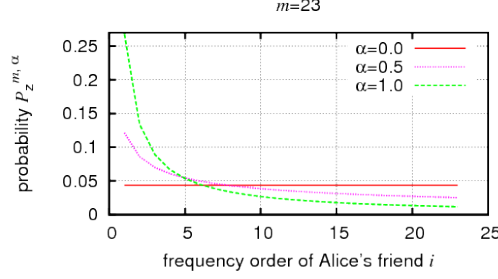


Figure 2.5: Zipf( $m, \alpha$ ) distribution of Alice's friends, for  $m = 23$ .

**Cover-traffic** We consider in this thesis that the cumulative communication traffic of all senders other than Alice leads to a distribution of the cover-traffic that is bounded from above by a uniform distribution. This thesis mainly focuses on the effect of Alice's communication on its relationship anonymity. To ease those analyses, we choose to simplify the cover-traffic by a uniform distribution of the recipients in  $R$ , such that without Alice's traffic, each recipient  $r \in R$  appears with the same probability  $P_N$  in an observation. Each sender can select its recipient according to an arbitrary distribution provided that  $\forall r \in R : P(r \in \mathcal{OS}') = P_N$ , where  $P(r \in \mathcal{OS}')$  denotes the probability that  $r$  appears in the observations in  $\mathcal{OS}$  without considering Alice's communication. That is  $\mathcal{OS}'$  result from removing all of Alice's communications in the observations in  $\mathcal{OS}$ , collected by the attacker.

As the overall distribution of the cover-traffic, but not the individual senders' traffic distributions in the cover-traffic are relevant for the HS-attack, we further assume for simplicity and without loss of generality that each of the  $(b - 1)$  non-Alice senders of a batch chooses its recipient from the set  $R$  of  $u$  recipients according to a uniform distribution in every round. The corresponding probability mass function is  $\frac{1}{u}$  for every  $r \in R$ , thus  $P_N = 1 - (\frac{u-1}{u})^{b-1}$  is the probability that a recipient  $r \in R$  is contacted by any sender other than Alice in an observation.

As a start, we only simulate and analyse uniformly distributed cover-traffic in this thesis for a better comparability of the simulative and analytical results. However, assume, for example, that the Mix parameters  $(\tilde{u}, b, m)$  are given, where  $|\tilde{R}| = \tilde{u}$  and  $\tilde{R}$  is the set of all recipients in the Mix system and we want to bound a specific (non-uniformly) distributed cover-traffic with given probability  $P(r \in \mathcal{OS}')$  for all  $r \in \tilde{R}$  by

---

$P_N$ . We can set  $P_N = 1 - \left(\frac{u-1}{u}\right)^{b-1}$  for a  $u$ , such that  $\forall r \in \tilde{R} : P(r \in \mathcal{OS}') \leq P_N$ .<sup>1</sup>

#### 2.2.4.2 Simulation

**Uniformly Distributed Alice’s Traffic** Figure 2.6 and Figure 2.7 draw the mean number of observations and of finalised sets evaluated by ExactHS when uniquely identifying Alice’s set of friends. They refer to the case that Alice’s communication and the cover-traffic is uniformly distributed. This allows the exclusive analysis of the influence of each the Mix parameters  $u, b, m$  on these quantities, that is independent of the underlying traffic distribution. We observe that changing a parameter, such that the mean number of observations to succeed the HS-attack is increased, also increases the corresponding number of finalised sets and thus the time-complexity of ExactHS. However, Figure 2.6 and Figure 2.7 also show that the relation between the Mix parameters  $u, b, m$  and the mean number of observations and mean time-complexity when succeeding the HS-attack is non-trivial. Low Mix parameters (e.g.,  $(u = 400, b = 10, m = 20)$  in Figure 2.6) can lead to a higher mean number of observations and of finalised sets than large Mix parameters (e.g.,  $(u = 20000, b = 50, m = 20)$  in Figure 2.7). This is even more surprising, as the corresponding worst case number of finalised sets for the former Mix parameters (i.e.,  $10^{20}$ ) is lower than for the latter Mix parameters (i.e.,  $50^{20}$ ). The estimate (4.16) of the mean number of finalised sets in Section 4.1.4.1 will mathematically show this non-trivial relation to the worst case number of finalised sets determined by Claim 3 in Section 2.2.3.1.

---

<sup>1</sup>Analyses of ExactHS would be applied to the corresponding Mix parameters  $(u, b, m)$  and a uniformly distributed cover-traffic resulting from every sender other than Alice contacting its recipients in  $R$  with the probability mass function  $\frac{1}{u}$ , where  $|R| = u$ ,  ${}_A\mathcal{H} \subseteq R$  and  $R \subseteq \tilde{R}$  if  $|R| \leq |\tilde{R}|$ , else  $R \supseteq \tilde{R}$ .

## 2. COMBINATORIAL ATTACK

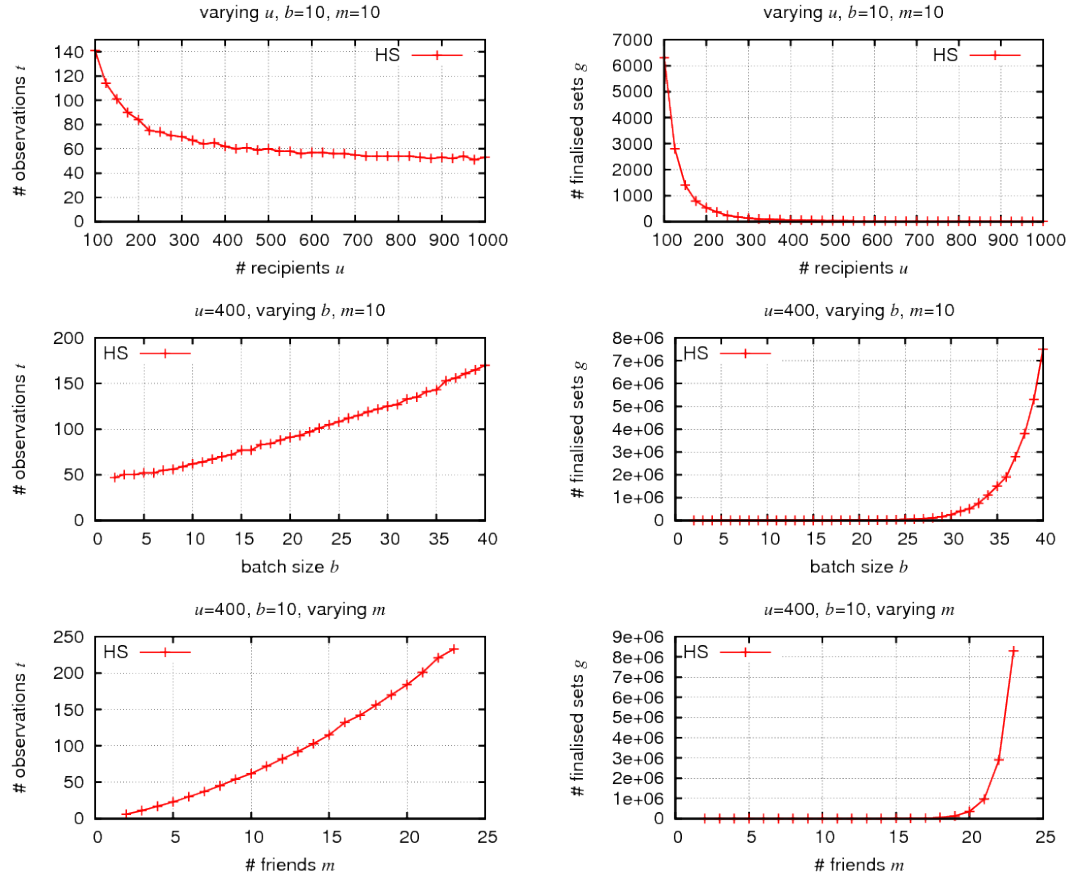


Figure 2.6: Uniformly distributed communication. Left: Mean number of observations to succeed HS-attack. Right: Mean number of finalised sets when succeeding HS-attack.



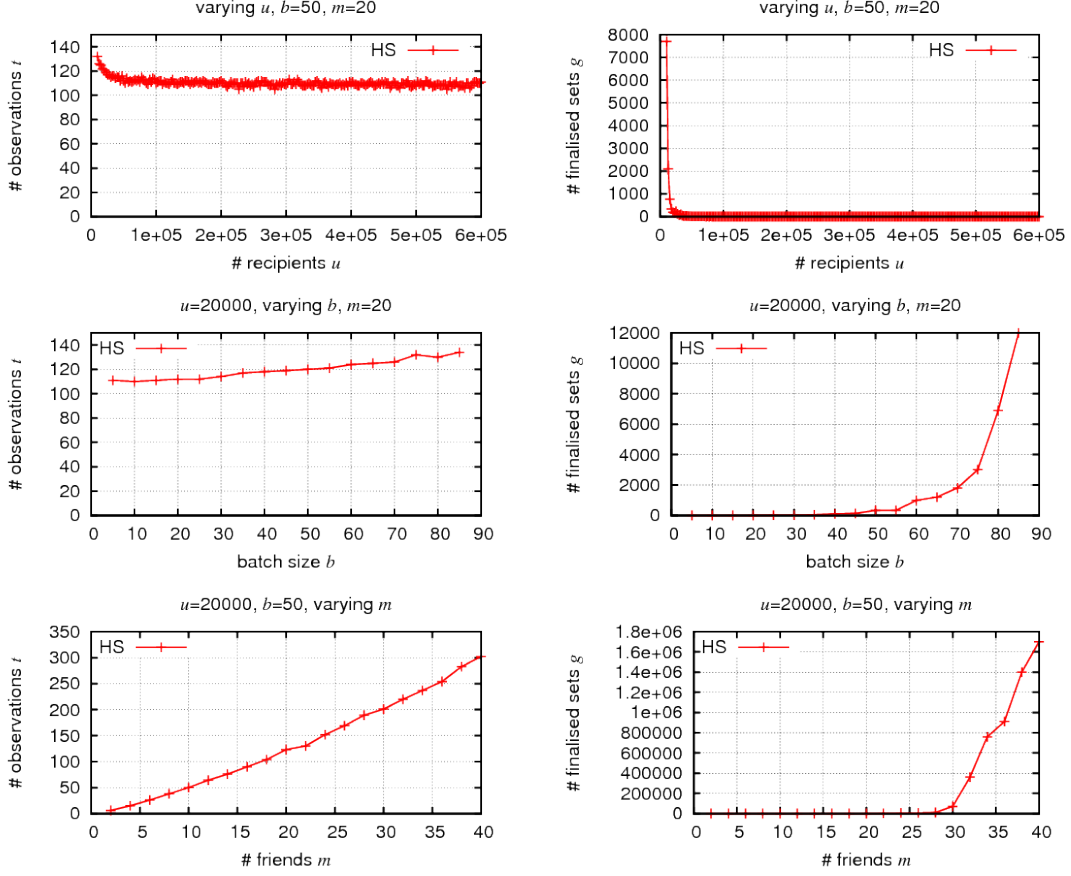


Figure 2.7: Uniformly distributed communication. Left: Mean number of observations to succeed HS-attack. Right: Mean number of finalised sets when succeeding HS-attack.

**Zipf-Distributed Alice's Traffic** Figure 2.8 and Figure 2.9 study the mean number of observations and of the finalised sets exclusively with respect to the Zipf distribution of Alice's friends for the Mix parameters ( $u = 400, b = 10, m = 23$ ) respectively ( $u = 20000, b = 50, m = 40$ ). It leaves the Mix parameters  $u, b, m$  unchanged, while varying the weight  $\alpha$  of the Zipf distribution Alice use to contact her  $m$  friends. We observe that increasing the non-uniformity of Alice's communication (i.e., the weight  $\alpha$ ) increases the mean number of observations, while decreasing the mean number of finalised sets when succeeding the HS-attack. For the Mix parameters ( $u = 400, b = 10, m = 23$ ) in Figure 2.8, the number of observations is for ( $\alpha = 1$ ) by a factor

## 2. COMBINATORIAL ATTACK

of 3 higher than that for ( $\alpha = 0$ ), while the number of finalised sets is for ( $\alpha = 1$ ) 58000, which is by a factor of  $\frac{1}{143}$  lower than that for ( $\alpha = 0$ ). For the Mix parameters ( $u = 20000, b = 50, m = 40$ ) in Figure 2.9, the number of observations is for ( $\alpha = 1$ ) by a factor of 3 higher than that for ( $\alpha = 0$ ), while the number of finalised sets is for ( $\alpha = 1$ ) 20000, which is by a factor of  $\frac{1}{85}$  lower than that for ( $\alpha = 0$ ).

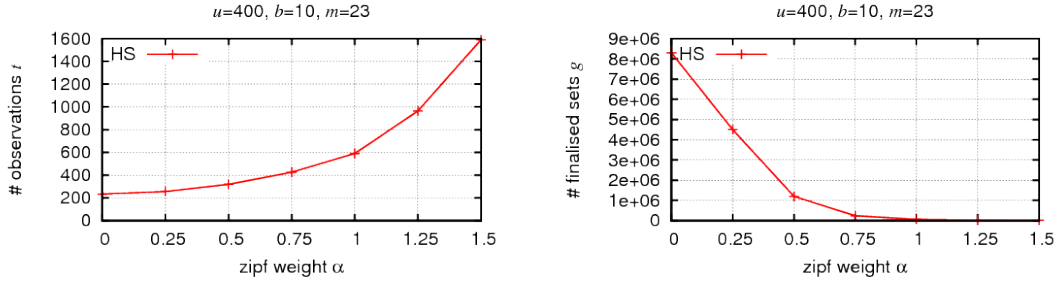


Figure 2.8: Zipf-distributed communication. Left: Mean number of observations to succeed HS-attack. Right: Mean number of finalised sets when succeeding HS-attack.

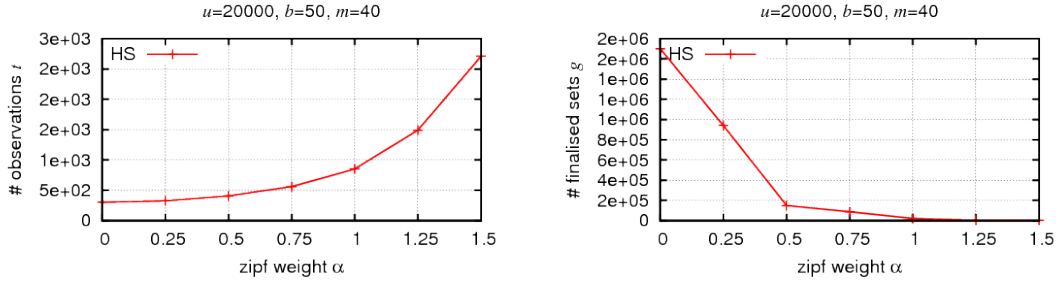


Figure 2.9: Zipf-distributed communication. Left: Mean number of observations to succeed HS-attack. Right: Mean number of finalised sets when succeeding HS-attack.

**Summary** In summary, Table 2.5 illustrates that the worst case number of finalised sets computed by HS-algorithm is significantly higher than that computed by ExactHS. The simulation results in Figure 2.6 and Figure 2.7 reveal that the mean number of finalised sets computed by ExactHS when succeeding the HS-attack is even significantly lower than the corresponding worst case number of finalised sets with respect to the Mix parameters. We further observe that regardless of the Mix parameters and of the

---

worst case time-complexity of ExactHS, the mean number of finalised sets and thus the mean time-complexity of ExactHS decreases, the more non-uniform Alice contacts her friends, as illustrated by the Zipf distribution of Alice's friends in Figure 2.8 and Figure 2.9. While these simulations illustrate complex relations between the Mix parameters and the mean number of observations and of the finalised sets computed by ExactHS when succeeding the HS-attack, we will derive these relations mathematically in Chapter 3 and Chapter 4.

## 2.3 Approximation of Hitting-Set-Attack

The theoretical limit of anonymity protection is reasonably represented by the least number of observations for unique identification of a user's contacts, as this limit is also valid for attackers with unlimited computing resources<sup>1</sup>.

This section provides an approximation that allows an efficient empirical estimate of the least number of observations to uniquely identify Alice's friends by the HS-attack from below. It applies to arbitrary concrete Mix parameters and communication traffics. The approximation requires the a-priori knowledge of Alice's set of friends  ${}_A\mathcal{H}$ . Therefore, it is solely an estimate of the number of observations to succeed the HS-attack, but no replacement of that attack. This approximation allows Alice to estimate in real time her current anonymity protection by herself to immediately adjust that protection, by, e.g., decent usage of dummy traffic as suggested by Kesdogan [2006, pp. 31 – 37].

### 2.3.1 Classification of Hitting-Sets and Hypotheses

The lower bound for the least number of observations is based on analysing the number of observations, until a particular class of hitting-sets disappears. Therefore we define a classification of hitting-sets that are computed by ExactHS in this section.

Let  $\mathfrak{H}$  be the set of all hypotheses<sup>2</sup>. We classify each hypothesis  $\mathcal{H} \in \mathfrak{H}$  according to the number of non-friends in it, such that  $\mathcal{H}$  is assigned one of the  $m + 1$  disjoint

---

<sup>1</sup>The theoretic limit of protection always refers to an attacker with unlimited computing resources.

<sup>2</sup>That are all hitting-sets of size  $m$

## 2. COMBINATORIAL ATTACK

---

classes  $\mathfrak{H}_0, \dots, \mathfrak{H}_m$ , where

$$\begin{aligned}\mathfrak{H}_0 &= \{{}_A\mathcal{H}\} \\ \mathfrak{H}_i &\subseteq (R \setminus {}_A\mathcal{H})^i \times {}_A\mathcal{H}^{m-i}, \text{ for } 0 < i < m. \\ \mathfrak{H}_m &\subseteq (R \setminus {}_A\mathcal{H})^m\end{aligned}$$

A hypothesis  $\mathcal{H}$  belongs to the class  $\mathfrak{H}_i$  that is stated by  $\mathcal{H} \in \mathfrak{H}_i$ , if it contains exactly  $(m - i)$  distinct Alice's friends. The class  $\mathfrak{H}_0$  contains exactly one set, Alice's set of friends  ${}_A\mathcal{H}$ . The class  $\mathfrak{H}_m$  represents hypotheses consisting of only Alice's non-friends.

This classification thus applies to every hitting-set of size  $m$ . To apply this classification to hitting-set  $\mathcal{H}$ , where  $|\mathcal{H}| < m$ , we assign  $\mathcal{H}$  to the class  $\mathfrak{H}_i$ , if it contains exactly  $i$  friends, so that  $\mathcal{H} \cup \{n_{|\mathcal{H}|+1}, \dots, n_m\} \in \mathfrak{H}_i$  is a hypothesis for every  $\{n_{|\mathcal{H}|+1}, \dots, n_m\} \in R \setminus {}_A\mathcal{H}$ . Thus for every  $\mathcal{H}$  that belongs<sup>1</sup> to  $\mathfrak{H}_i$ , there is a hypothesis  $\mathcal{H}' \supset \mathcal{H}$ , where  $\mathcal{H}' \in \mathfrak{H}_i$ .

### 2.3.2 Approximation Based on No-Trivial-Disproof

The UMHS problem remains NP-complete, even though Alice's set of friends is a-priori known, thus motivating the need for efficient approximations. It is clear that Alice's set of friends  ${}_A\mathcal{H}$  is a unique minimum-hitting-set in  $\mathcal{OS}$  if and only if there is no specified hypothesis in any of the classes  $\mathfrak{H}_1, \dots, \mathfrak{H}_m$  in  $\mathcal{OS}$ . We call the existence of a specified hypothesis in the class  $\mathfrak{H}_1$  in a given observation set, a *trivial-disproof* that  ${}_A\mathcal{H}$  is not unique. Our approximation consists of determining the least number of observations  $t_a = |\mathcal{OS}|$ , such there is no specified hypothesis in the class  $\mathfrak{H}_1$ , that is there is no trivial-disproof.

The trivial-disproofs can be efficiently identified by Algorithm 3 that we explain in a line by line basis, next. Each specified hypothesis  $\mathcal{H} \in \mathfrak{H}_1$  contains a subset  $\mathcal{C} \subset \mathcal{H}$ , consisting of exactly  $(m - 1)$  friends, that is  $\mathcal{C} \subset {}_A\mathcal{H}$ , where  $|\mathcal{C}| = m - 1$  and a non-friend  $n \in R \setminus {}_A\mathcal{H}$ . Line 2 in Algorithm 3 thus loops over all possible sets  $\mathcal{C}$  specified in the former sentence. For each  $\mathcal{C}$ , Line 4 determines by  $\mathcal{J}$  all possible non-friends for

---

<sup>1</sup>The term “belongs to” is not strictly defined, but it provides a consistent classification of hitting-sets.

---

**Algorithm 3** Trivial-Disproof.

---

```
1: procedure TRIVDISPROOF( ${}_A\mathcal{H}, m, \mathcal{OS}$ )
2:   for  $\mathcal{C} \in \{\{a_1, \dots, a_{m-1}\} \subseteq {}_A\mathcal{H}\}$  do            $\triangleright$  Choose distinct sets of  $m - 1$  friends
3:      $\mathcal{OS}' \leftarrow \{\mathcal{O} \setminus {}_A\mathcal{H} \mid \mathcal{O} \in \mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]\}$ 
4:      $\mathcal{I} \leftarrow \bigcap_{\mathcal{O}' \in \mathcal{OS}'} \mathcal{O}'$ 
5:     if  $(|\mathcal{OS}'| = 0) \vee (\mathcal{I} \neq \{\})$  then
6:       Print  $\{\mathcal{C} \cup \{n\} \mid n \in \mathcal{I}\}$             $\triangleright$  All identified trivial-disproofs
```

---

$n$  that hit those observations  $\mathcal{OS}' = \mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$  that are not hit by  $\mathcal{C}$ . If  $\mathcal{C}$  is a subset of a hitting-set of size  $m$  in  $\mathcal{OS}$ , then Line 6 provides the following two possible outputs. If  $\mathcal{C}$  is a hitting-set in  $\mathcal{OS}$ , then  $\mathcal{C} \cup \{n\} \in \mathfrak{H}_1$  for every  $n \in R \setminus {}_A\mathcal{H}$ , therefore it is sufficient to output  $\mathcal{C}$ , otherwise  $\mathcal{C} \cup \{n\}$ , for every  $n \in \mathcal{I}$  is provided as the output. Algorithm 3 terminates without any output, if there is no trivial-disproof.

Since *no-trivial-disproof* is a necessary, but not sufficient condition for  ${}_A\mathcal{H}$  to be a unique minimum-hitting-set,  $t_a$  is thus a provable lower bound for the least number of observations to succeed the HS-attack. Indeed, we can observe in Figure 2.10 that specified hypotheses in the class  $\mathfrak{H}_1$  remain hitting-sets for a large number of observations  $|\mathcal{OS}|$  that is close to the least number of observations  $t$  to succeed the HS-attack.

### 2.3.2.1 Complexity

In Line 2 of Algorithm 3, the number of subsets  $\mathcal{C} \subset {}_A\mathcal{H}$ , where  $|\mathcal{C}| = m - 1$  is  $\binom{m}{m-1} = m$ . Computing  $\mathcal{OS}'$  in Line 3 requires comparing the  $m$  friends in a given  $\mathcal{C}$  and  ${}_A\mathcal{H} \setminus \mathcal{C}$  with the  $b$  recipients in the observations  $\mathcal{O} \in \mathcal{OS}$ , which leads to  $O(mtb)$  operations, where  $t = |\mathcal{OS}|$ . Computing  $\mathcal{I}$  in Lines 4 requires  $O(btb)$  operations. The overall time-complexity of Algorithm 3 is thus  $m(O(mtb) + O(btb))$ , that is

$$O(m^2tb^2) \text{ .}$$

The space-complexity of Algorithm 3 is obviously linear.

### 2.3.2.2 Relation to 2 $\times$ -Exclusivity

According to Kesdogan et al. [2006] a friend appears *exclusively* in an observation, if there is no other friend in that observation. A friend is *2 $\times$ -exclusive*, if it appears at least two times exclusively in observations, or at least one time alone (i.e., without

## 2. COMBINATORIAL ATTACK

---

any other recipient) in an observation. The  $2\times$ -*exclusivity* property is fulfilled, if all Alice's friends are  $2\times$ -exclusive. It was shown by Kesdogan et al. [2006] that this property is a good approximation for the lower bound of the number of observations to fully disclose Alice's set of friends. We prove that no-trivial-disproof is an even better approximation than  $2\times$ -exclusivity by showing the following two cases. Firstly, there are observations, where the  $2\times$ -exclusivity property is fulfilled, although there is still a hitting-set that belongs to  $\mathfrak{H}_1$ . Secondly, if the  $2\times$ -exclusivity property is not fulfilled, then there is always a hitting-set that belongs to  $\mathfrak{H}_1$ .

*Proof (Case 1).* Since this is an existence proof, we only have to construct an example, where  $2\times$ -exclusivity is given, although a hitting-set that belongs to  $\mathfrak{H}_1$  exists. Let Alice's set of friends be  ${}_A\mathcal{H} = \{1, 3\}$ . Assume that the following four observations are given:

$$\{3, 5, 6\}, \{3, 1\}, \{1\}, \{6, 3\}.$$

All friends in  ${}_A\mathcal{H}$  are obviously  $2\times$ -exclusive, but there is a hitting-set  $\{1, 6\}$  that belongs to  $\mathfrak{H}_1$ .  $\square$

*Proof (Case 2).* Let  ${}_A\mathcal{H} = \{a_1, \dots, a_m\}$  be the Alice's set of friends. Assume wlog. that the friend  $a_1$  appears only one time exclusively in a given observation set  $\mathcal{OS}$ . Let  $\mathcal{O} \in \mathcal{OS}$  be the observation, in that  $a_1$  appears exclusively. Then the set  $\mathcal{H} = \{n, a_2, \dots, a_m\}$ , where  $n \in \mathcal{O} \setminus {}_A\mathcal{H}$  is a hitting-set belonging to  $\mathfrak{H}_1$ .  $\square$

### 2.3.3 Evaluation

The plots from Figure 2.10 represent results of applying the HS-attack on simulated Mix traffic. In this simulation we assume that in each round, Alice chooses her recipient uniformly distributed from one of her  $m$  friends, i.e., with probability  $\frac{1}{m}$ . All the other  $b - 1$  senders are assumed to choose their recipients uniformly distributed from the set of all  $|R| = u$  recipients, i.e., with probability  $\frac{1}{u}$ .

Each of the figures varies one of the parameter  $u, m, b$ , while all the other parameters remain fixed. The standard parameter of this simulation is  $u = 400, m = 10, b = 10$ . The line labelled (HS) shows the mean number of observations of the HS-attack to reveal the Alice's set of friends. Accordingly the line labelled (2x-excl) is the mean

number of observation to fulfil the  $2\times$ -exclusivity criterion, while the line labelled  $(h_1)$  visualizes the mean number of observations until there is no specified hypothesis in the class  $\mathfrak{H}_1$ , that is there is no trivial-disproof.

The simulation results confirm our proof in Section 2.3.2.2. The no-trivial-disproof criterion provides a better estimate of the number of observations required to succeed the HS-attack than the  $2\times$ -exclusivity criterion. The top right plot in Figure 2.10 reveals that the superiority of the estimate by no-trivial-disproof becomes more significant, compared to the estimate by  $2\times$ -exclusivity, if the batch size is large. This is because a larger batch size increases the probability that observations that contain an Alice's friend, e.g.,  $a_i \in {}_A\mathcal{H}$  exclusively, also contain some common non-friend  $r$ . Consequently it becomes more likely that  $\{a_1, \dots, a_{i-1}, r, a_{i+1}, \dots, a_m\} \in \mathfrak{H}_1$  is a hitting-set, although all Alice's friends are  $2\times$ -exclusive, for larger batch sizes.

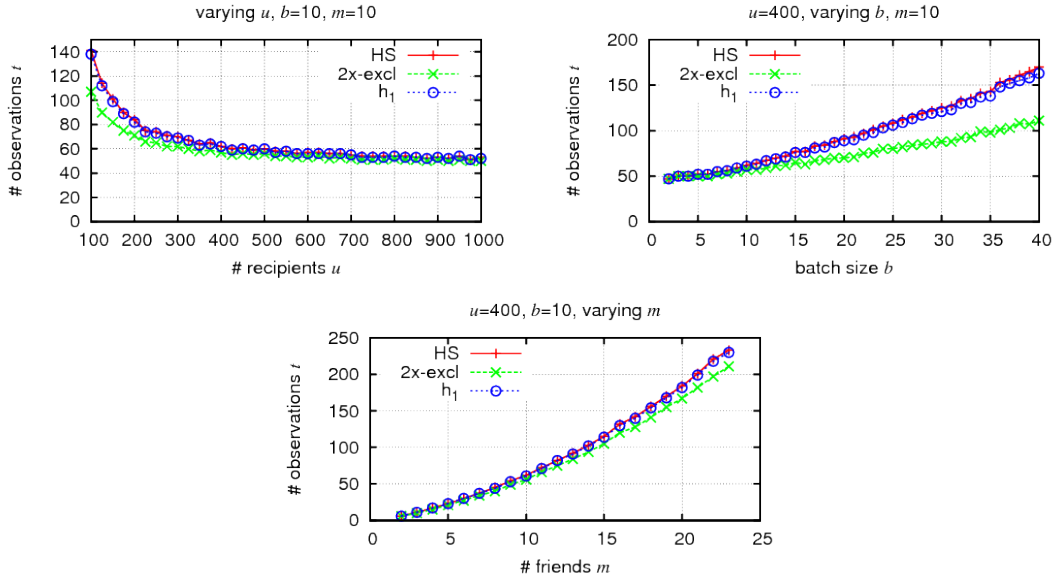


Figure 2.10: Mean number of observations until: succeeding HS-attack (HS), fulfilling  $2\times$ -exclusivity property (2x-excl) and disappearing of no-trivial-disproofs ( $h_1$ ).

### 2.4 Summary

This chapter drew the basic scheme underlying the analyses of our combinatorial attack on the Chaum Mix in this thesis. As illustrated in Figure 2.2 in Section 2.1.2.2, the Mix configuration models the input of the attack (the random observations), while the output of the attack is determined by the hypotheses. We consider the effort of an attack to uniquely identify Alice’s set of friends as a measure for the limit of anonymity protection provided by the Mix system. Therefore we studied the least number of observations and the computational complexity required for that unique identification by an attack with respect to distinct Mix configurations. The analyses in this chapter were empirical and based on simulations.

We chose to consider the HS-attack introduced by Kesdogan and Pimenidis [2004], as it provably requires the least number of observations to uniquely identify Alice’s set of friends, as proved by Kesdogan et al. [2006]. This least number of observations to disclose Alice’s relationship anonymity is a pendant to Shannon’s unicity-distance that measures the information theoretic least number of intercepted cipher text bits to disclose an encrypted message, cf. Shannon [1949]. The original HS-attack of Kesdogan and Pimenidis [2004] uniquely identifies Alice’s set of friends by solving the UMHS problem with the HS-algorithm that is exponential in (mean) time and space. However, the UMHS can be solved by orders of magnitudes more efficiently by our ExactHS algorithm in Section 2.2.1. This was illustrated by Table 2.5 in Section 2.2.4. Section 2.2.3 proved that ExactHS provides a linear worst case space-complexity and minimises the worst case time-complexity. We demonstrated in Section 2.2.4 that it can even provide feasible mean time complexities. The HS-attack using ExactHS to solve the UMHS problem, instead of using the original HS-algorithm identifies Alice’s set of friends with the least number of observations. This thus contributes a more efficient variant of the HS-attack and we use and analyse this efficient variant by default in this thesis.

We further observed that the mean time-complexity itself can be decreased without changing the Mix parameters, as in case of changing the distribution of Alice’s friends from a uniform distribution to a non-uniform distribution. In that case, we modelled the distribution of Alice’s friends by a  $\text{Zipf}(m, \alpha)$  distribution, that is known to closely model e-mail and Internet traffic Adamic and Huberman [2002]; Almeida



---

et al. [1996]; Breslau et al. [1999]; Glassman [1994]. It could be observed that increasing the weight  $\alpha$  and thus increasing the non-uniformity of the distribution of Alice's friends significantly decreases the mean time-complexity of ExactHS, while increasing the least number of observations to uniquely identify Alice's set of friends by the HS-attack. The increase of the number of observations is due to recipients who are rarely contacted by Alice.

The thesis additionally contributed in Section 2.3 an efficient empirical approximation of the number of observations required to succeed the HS-attack. This approximation is based on a necessary condition for the unique identification of Alice's set of friends that subsumes the  $2\times$ -exclusivity condition of Kesdogan et al. [2006]. We call this the no-trivial-disproof condition and proved that it provides a lower bound for the least number of observations to uniquely identify Alice's set of friends that is closer than the bound provided by the  $2\times$ -exclusivity. This approximation aids efficient empirical analyses of the HS-attack, but it is no attack, as it requires the a-priori knowledge of Alice's set of friends.

The empirical evaluations in Section 2.2.4 indicated complex relations between the Mix parameters, the traffic distributions, the mean time-complexity of ExactHS and the least number of observations to succeed the HS-attack. However, analytical analyses are crucial to understand the systematic behind those relations, in order to explain the limit of anonymity protection provided by the Chaum Mix. Therefore, Chapter 3 and Chapter 4 will analyse that relations analytically.

## **2. COMBINATORIAL ATTACK**

---

## Chapter 3

# Theoretical Limit of Anonymity Protection

The theoretical limit refers to the anonymity protection provided by a system against an attacker with unlimited computing resources<sup>1</sup>. In accordance to Shannon’s unicity-distance, cf. Shannon [1949], we consider the least number of observations to uniquely identify Alice’s friends as the theoretical limit of anonymity protection.

The HS-attack and its approximations in Chapter 2 allow empirical analyses of the least number of observations for the unique identification of Alice’s friends for single Mix configurations<sup>2</sup>. However, this cannot replace analytical analyses, which allow proving systematic relationships between Mix configurations and the least number of observations for the unique identifications.

Due to the complexity of analytical analyses, past works, as Kesdogan and Pimenidis [2006]; Kesdogan et al. [2006]; O’Connor [2008] solely focused on analysing the Chaum Mix for the simple case of uniformly distributed traffic. In contrast to past works, we contribute a closed formula that estimates the mean of the least number of observations to uniquely identify Alice’s set of friends with respect to arbitrary distribution of Alice’s traffic and Mix parameters. It can also be applied to non-uniformly distributed cover traffic. Our estimate thus allows studying more realistic traffic distributions and comparing the number of observations required by the HS-attack with

---

<sup>1</sup>The theoretical limit means that the protection against a computationally unlimited attacker cannot exceed this limit.

<sup>2</sup>The HS-attack using ExactHS is computationally tractable in some cases.

### 3. THEORETICAL LIMIT OF ANONYMITY PROTECTION

---

those of other attacks.

The estimate shows that the mean number of observations to uniquely identify Alice's set of friends is inverse proportional to the least probability  $p$  in the distribution of Alice's traffic. Firstly, this implies that the uniformly distributed Alice's traffic minimises the number of observations to uniquely identify Alice's set of friends, that was as yet unproven, but believed, cf. Kesdogan and Pimenidis [2006]; Kesdogan et al. [2006, 2009]. Secondly, this shows that the HS-attack requires only  $O(\frac{1}{p})$  observations to exactly identify all Alice's friends, whereas  $O(\frac{1}{p^2})$  observations are required by the Statistical Disclosure attack (SDA) of Danezis [2003] to inexactly classify Alice's friends.

To the best of our knowledge, we are the first to show the mentioned relationships analytically.

Section 3.1 mathematically estimates the mean of the least number of observations to uniquely identify Alice's set of friends. This results in a closed formula that determines that mean with respect to the parameters in the Mix configuration.

Section 3.2 compares the mathematically estimated mean of the least number of observations with that obtained by applying the HS-attack on random observations in simulations. This reveals that our mathematically estimated mean is reasonable close to the empirical mean. Additionally, we graphically compare the estimated mean least number of observations required to succeed the HS-attack with that required by the SDA to guess Alice's friends with a true-positive rate of 95%.

#### 3.1 Mean Number of Observations

The work of Kesdogan et al. [2006] has shown that the least number of observations, such that the  $2\times$ -exclusivity property is fulfilled, provides a close lower bound for the least number of observations to uniquely identify Alice's set of friends, cf. Section 2.3.2.2. We propose an approximation of the mean of the least number of observations to fulfil the  $2\times$ -exclusivity property. This provides a closed formula that estimates the mean of the least number of observations to uniquely identify Alice's set of friends with respect to arbitrary distribution of Alice's traffic and Mix parameters. It is more general than the mean number of observations for  $2\times$ -exclusivity of Kesdogan

---

et al. [2006], since the mean of Kesdogan et al. [2006] is a non-closed formula that is only valid for uniform traffic distributions. The advantages of Kesdogan et al. [2006] is the exact computation of the mean number of observations for  $2\times$ -exclusivity, but we can see in Section 3.2 that our estimate is reasonable close to that exact mean.

Section 3.1.1 estimates the mean least number of observations to fulfil  $2\times$ -exclusivity from the top and from the bottom. This is followed by the estimate of the mean least number of observations for  $k\times$ -exclusivity, for arbitrary  $k \in \mathbb{N}$  in Section 3.1.2. The bounds in Section 3.1.1 show that Section 3.1.2 provides a closer estimate of the mean least number of observations for  $2\times$ -exclusivity, for  $k = 2$ . Therefore we prefer in this thesis estimates based on Section 3.1.2. All our estimates show that the mean number of observations to uniquely identify Alice's set of friends is minimised, if Alice's traffic is uniformly distributed.

Section 3.1.3 compares the estimated mean least number of observations for  $2\times$ -exclusivity with the number of observations required by the SDA of Danezis [2004] to guess Alice's friend with some true-positive rates<sup>1</sup>. It shows that the former mean is in  $O(\frac{1}{p})$ , while the latter mean is in  $O(\frac{1}{p^2})$ , where  $p$  denotes the least probability  $p < 1$  in the distribution of Alice's friends.

### 3.1.1 Bounds of Mean Number of Observations for $2\times$ -Exclusivity

As a first estimate of the mean of the number of observations for  $2\times$ -exclusivity for arbitrary distribution of Alice's traffic we propose estimates of the lower and upper bound of that mean in this section. The idea is to estimate the mean of the least number of observations, such that all Alice's friends are  $1\times$ -exclusive, that we called *1-exclusivity*. This mean can be straight forwardly estimated and provides a lower bound of the corresponding mean for  $2\times$ -exclusivity.

Section 3.1.1.1 shows the relation between  $1\times$ -exclusivity and  $2\times$ -exclusivity. This allows estimating the lower and upper bound of the mean of the least number of observations for  $2\times$ -exclusivity, based on the corresponding mean for  $1\times$ -exclusivity. The closed formula for the estimate of that means is derived in Section 3.1.1.2.

---

<sup>1</sup>The SDA was introduced by Danezis [2003]. But the estimate of the number of observations required by the SDA was corrected by Danezis [2004] and is thus more recent.

### 3. THEORETICAL LIMIT OF ANONYMITY PROTECTION

---

#### 3.1.1.1 Relating $1\times$ -Exclusivity and $2\times$ -Exclusivity

Remember that an Alice's friend  $a$  is exclusive in an observation, if the other senders do not contact any of the other  $(m - 1)$  friends in  ${}_A\mathcal{H}$ , in that observation. It is  $1\times$ -exclusive, respectively  $2\times$ -exclusive, if it is exclusive in at least one, respectively two observations.

Let  $T_{1\times e}$ ,  $T_{2\times e}$  be the random variables for the least number of observations, until all friends are  $1\times$ -exclusive, respectively  $2\times$ -exclusive with means  $E(T_{1\times e})$ ,  $E(T_{2\times e})$ . All senders are assumed to choose their recipients statistically independently, so that the probability that a particular friend is exclusive in a given observation is not dependent on any past observations. This therefore implies the following relations:

$$E(T_{1\times e}) \leq E(T_{2\times e}) \leq 2E(T_{1\times e}).$$

#### 3.1.1.2 Estimation of $2\times$ -Exclusivity Based on $1\times$ -Exclusivity

We estimate  $E(T_{2\times e})$  for arbitrary distributions of Alice's friends by providing a closed formula for  $E(T_{1\times e})$ . Most importantly, we prove that  $E(T_{1\times e}) \propto 1/p$  and thus  $E(T_{2\times e}) \propto 1/p$ , where  $p = \min\{P_A(a) \mid a \in {}_A\mathcal{H}\}$  is the least probability in the considered distribution. Changing Alice's traffic distribution and  $p$ , while leaving  $u, b, m$  and the cover-traffic distribution unchanged, only changes the number of observations to identify  ${}_A\mathcal{H}$  by ExactHS linearly with  $1/p$ . The proof derives the following estimate of the expectation of the least number of observations for unique identification of all Alice's friends:

$$E(T_{2\times e}) \leq \frac{1}{p} \frac{2(\ln(m) + \gamma)}{\left(\frac{u-(m-1)}{u}\right)^{b-1}} \quad \text{and} \quad E(T_{2\times e}) \propto \frac{1}{p}.$$

*Proof.* Let  $T_{1\times, i}$  be the random variable for the least number of observations passed between Alice contacting the  $(i - 1)$ -th friend and the  $i$ -th friend, for  $i \in \{1, \dots, m\}$ . This is regardless whether the observations are exclusive, or not and thus independent of the behaviour of senders other than Alice.

Let  $T_{e, a_j}$  be the random variable for the least number of times Alice has to contact  $a_j$ ,<sup>1</sup> until  $a_j$  is exclusive. Obviously,  $T_{e, a_j}$  is only dependent on the behaviour

---

<sup>1</sup>This only counts observations in  $\mathcal{OS}_A[a_j]$ .

---

of the senders other than Alice and independent of Alice's behaviour. Let us set  $a = \operatorname{argmax}_{a_j \in {}_A\mathcal{H}} E(T_{e,a_j})$  and  $T_e = T_{e,a}$ , then  $T_{1 \times, i}$  and  $T_e$  are statistically independent and we obtain

$$E(T_{1 \times e}) \leq \sum_{i=1}^m E(T_{1 \times, i}) E(T_e) .$$

$E(T_e)$ : Assume that every  $r \in R$ ,  $|R| = u$  is contacted uniformly distributed by any  $(b-1)$  non-Alice senders in every observation, then the probability that  $r$  is contacted by any non-Alice sender is  $P_N(r) = P_N = 1 - (\frac{u-1}{u})^{b-1}$ . Given Alice contacts  $a_j \in {}_A\mathcal{H}$  and the remaining  $(b-1)$  non-Alice senders do not, then  $a_j$  is exclusive. That probability is  $P_e(a_j) = (\frac{u-(m-1)}{u})^{b-1}$ .  $T_{e,a_j}$  is geometrically distributed;  $E(T_{e,a_j}) = \frac{1}{P_e(a_j)} = (\frac{u-(m-1)}{u})^{1-b}$  for  $j = 1, \dots, m$ , thus  $E(T_e) = E(T_{e,a_j})$ . This  $E(T_e)$  serves as an upper bound for  $E(T'_{e,a_j})$  of all cases, where  $r' \in R'$  is non-uniformly contacted with  $P'_N(r')$  and  $\max_{r' \in {}_A\mathcal{H}} \{P'_N(r')\} \leq P_N$ , for any recipient sets  $R' \supset {}_A\mathcal{H}$ .

$\sum_{i=1}^m E(T_{1 \times, i})$ : Alice contacts (arbitrarily distributed) friends  $a_j \in {}_A\mathcal{H}$  with  $P_A(a_j)$  in each observation. Let  $T_{1 \times} = \sum_{i=1}^m T_{1 \times, i}$ , then  $E(T_{1 \times}) = \sum_{i=1}^m E(T_{1 \times, i})$  is the mean number of observations until Alice contacts each of her friends at least once. This equals the *Coupon Collector Problem* (CCP) Boneh and Hofri [1997], where  $a_1, \dots, a_m$  are  $m$  coupon types and  $P_A(a_j)$  is the probability of getting a type  $a_j$  coupon. According to Boneh and Hofri [1997, Eq. (30)]:

$$E(T_{1 \times}) = \int_{t=0}^{\infty} (1 - \prod_{j=1}^m (1 - e^{-P_A(a_j)t})) dt$$

is the mean number of coupons collected to obtain all types.

Let  $p = \min_{a_j \in {}_A\mathcal{H}} \{P_A(a_j)\}$  and  $m' = 1/p$ , then  $m' \geq m$ . Let us further define  $U = \{a_1, \dots, a_m, n_{m+1}, \dots, n_{m'}\}$  as the type set of a CCP, where every type has the probability  $p$  for integral  $m'$ . Collecting all types in  ${}_A\mathcal{H} \subseteq U$  requires

### 3. THEORETICAL LIMIT OF ANONYMITY PROTECTION

---

according to Boneh and Hofri [1997, Eq. (39)]

$$E(T_{1 \times \mathcal{H}}) = \int_{t=0}^{\infty} (1 - \prod_{j=1}^m (1 - e^{-pt})) dt \quad (3.1)$$

$$= \frac{1}{p} \sum_{i=1}^m \frac{1}{i} \quad (\text{using Flajolet et al. [1992]}) \quad (3.2)$$

$$\approx \frac{1}{p} (\ln(m) + \gamma) \quad (3.3)$$

coupons, where  $\gamma \approx 0.577$  is the Euler-Mascheroni number. Note that  $E(T_{1 \times}) \leq E(T_{1 \times \mathcal{H}})$ , as  $P_A(a_j) \geq p$  for  $j = 1, \dots, m$ . Collecting all types in  ${}_A\mathcal{H}$  requires for arbitrary distributions fewer coupons than collecting all types in  ${}_A\mathcal{H} \subseteq U$ , where every type in  $U$  is of probability  $p$  and  $|U| = \frac{1}{p}$ .

Let  $p = \min_{a_j \in {}_A\mathcal{H}} \{P_A(a_j)\}$ , we conclude from this proof that:

$$\begin{aligned} E(T_{1 \times e}) &\leq \frac{1}{p} (\ln(m) + \gamma) \left( \frac{u - (m-1)}{u} \right)^{1-b} & \text{and } E(T_{1 \times e}) &\propto \frac{1}{p} \\ E(T_{2 \times e}) &\leq \frac{1}{p} \frac{2(\ln(m) + \gamma)}{\left( \frac{u - (m-1)}{u} \right)^{b-1}} & \text{and } E(T_{2 \times e}) &\propto \frac{1}{p} \end{aligned} \quad (3.4)$$

□

Inequality (3.4) thus estimates the expectation of the least number of observations to uniquely identify Alice's set of friends for arbitrary distribution of Alice's friends. It can be observed, that the number of observations to succeed the HS-attack is proportional to  $\frac{1}{p}$ . Thus the least probability in the distribution of Alice's friends determines how long Alice can contact her set of friends, without them being uniquely identifiable.

#### 3.1.2 Mean Number of Observations for $k \times$ -Exclusivity

We propose a closed formula for the estimated mean of the least number of observations, such that all Alice's friends are  $k \times$ -exclusive, for arbitrary  $k \in \mathbb{N}$  and distributions of Alice's traffic. In accordance to the definition of Kesdogan et al. [2006], a friend  $a \in {}_A\mathcal{H}$  is  $k \times$ -exclusive, if it appears at least  $k$  times exclusively, or at least



---

one time alone<sup>1</sup> in the considered observations. The  $k \times$ -*exclusivity* is fulfilled, if all Alice's friends are  $k \times$ -exclusive. While this chapter mainly discusses  $k \times$ -exclusivity for  $k = 2$ , the general case of  $k \in \mathbb{N}$  will be applied in Section 5.2.2.1 in Chapter 5.

Section 3.1.2.1 derives the estimate of the mean of the least number of observations  $E(T_{2 \times e})$  for  $k \times$ -exclusivity, based on a close estimate of the *general CCP* by Brayton [1963]. The general CCP can be used to estimate the mean number of coupon collections  $E(T_{k \times})$ , such that every coupon type appears at least  $k$  times, for  $k \in \mathbb{N}$ . The resulting estimate of  $E(T_{2 \times e})$  by (3.5) (for  $k = 2$ ) might over, or under estimate  $E(T_{2 \times e})$ , whereas (3.4) estimates the upper bound of  $E(T_{2 \times e})$ .

Section 3.1.2.2 shows that the estimate of  $E(T_{2 \times e})$  in (3.4) exceeds that in (3.5) (for  $k = 2$ ) by a factor that is lower than 2, while the estimate of  $E(T_{1 \times e})$  in (3.4) is lower than (3.5) (for  $k = 2$ ). This allows concluding that (3.5) provides reasonable estimate of  $E(T_{2 \times e})$  for all Mix configurations and a closer estimate of  $E(T_{2 \times e})$  than (3.4). The estimate (3.5) is therefore preferred in the remaining thesis.

Section 3.1.2.3 analyses the effect of Alice's traffic distribution on the mean number of observations for  $2 \times$ -exclusivity. This concludes that the uniform distribution of Alice's traffic minimises that mean in the average case.

### 3.1.2.1 Estimation of $k \times$ -Exclusivity

**Claim 5.** Let  $E(T_{e,a})$  be the mean number of observations Alice has to contact a friend  $a \in {}_A\mathcal{H}$ ,<sup>2</sup> until  $a$  is exclusive and  $E(T_e) = \max_{a \in {}_A\mathcal{H}} E(T_{e,a})$ , as defined in Section 3.1.1.2. Let  $E(T_{k \times})$  be the mean number of coupon collections to obtain every coupon type at least  $k$  times in the general CCP, cf. Brayton [1963]. The set of  $m$  coupon types is represented by  ${}_A\mathcal{H}$ , where  $P_A(a)$  is the probability that a coupon is of type  $a \in {}_A\mathcal{H}$ . The distribution specified by  $P_A(a)$  can be arbitrary.

We denote the mean number of observations until all Alice's friends are  $k \times$ -exclusive for arbitrary integer  $k \geq 1$ , by  $E(T_{k \times e})$ . Its estimate is

$$\begin{aligned} E(T_{k \times e}) &\leq E(T_{k \times})E(T_e) \\ &\approx \left( \frac{1}{p}(\ln m + \gamma) + (k-1)\frac{1}{p} \ln \ln m \right) \left( \frac{u - (m-1)}{u} \right)^{1-b}, \end{aligned} \quad (3.5)$$

---

<sup>1</sup>This is an observation  $\mathcal{O}' = \{a\}$ .

<sup>2</sup>This only refers to observations, in that Alice contacts  $a$ , that is  $\mathcal{OS}_A[a]$ .

### 3. THEORETICAL LIMIT OF ANONYMITY PROTECTION

---

where  $p = \min_{a \in {}_A\mathcal{H}} P_A(a)$  and  $\gamma \approx 0,57721$  is the Euler-Mascheroni constant.

The  $1 \times$ -exclusivity of all Alice's friends requires on average  $\frac{1}{p}(\ln m + \gamma)E(T_e)$  observations, while every additional exclusivity of all Alice's friends requires on average  $(\frac{1}{p} \ln \ln m)E(T_e)$  additional observations.

Claim 3.5 generalises the mathematical estimate of  $2 \times$ -exclusivity in Section 3.1.1.2 to the estimate of  $k \times$ -exclusivity, for arbitrary integer  $k \geq 1$ . We therefore reuse the notations and definitions provided in Section 3.1.1.2.

*Proof of Claim 5.* Recall from Section 3.1.1.2 that  $T_e$  is the random variable for the number of times Alice has to contact a particular friend  $a \in {}_A\mathcal{H}$ ,<sup>2</sup> until  $a$  is exclusive, where  $a = \operatorname{argmax}_{a' \in {}_A\mathcal{H}} E(T_{e,a'})$ . The random variable  $T_{k \times}$  denotes the number of observations<sup>1</sup> until Alice contacts each of her friends at least  $k$  times. These random variables are according to Section 3.1.1.2 statistically independent, so that the mean number of observations  $E(T_{k \times e})$  until all Alice's friends are  $k \times$ -exclusive can be estimated by the product of the expectations  $E(T_e), E(T_{k \times})$  of these two random variables, that is:

$$E(T_{k \times e}) \leq E(T_{k \times})E(T_e) . \quad (3.6)$$

The equation  $E(T_e) = \left(\frac{u-(m-1)}{u}\right)^{1-b}$  was derived in Section 3.1.1.2, therefore it remains to show the estimate of  $E(T_{k \times})$  for integer  $k \geq 1$ .

The expression for  $E(T_{k \times})$  was derived by Brayton [1963], which is the following equality for large value of  $m$ , that is for  $m \rightarrow \infty$ :

$$E(T_{k \times}) = \frac{m}{\delta}(\ln \kappa m + \gamma) + (k-1)\frac{m}{\delta}(\ln \ln \kappa m + \ln \frac{1}{\delta}) + o(1) .$$

The variables in this equation have the following meaning in our context:

- $m = |{}_A\mathcal{H}|$  is the number of coupon types, where w.l.o.g.  ${}_A\mathcal{H} = \{1, \dots, m\}$ .
- $\delta = \min_{x \in (0,1]} f(x) \leq 1$ , where  $P_A(a) = \int_{(a-1)/m}^{a/m} f(x)dx$  and  $\int_0^1 f(x)dx = 1$ .  $\delta$  is the continuous counterpart of the discrete probability  $\min_{a \in {}_A\mathcal{H}} P_A(a)$ . We therefore set  $f(x) = mP_A(\lceil xm \rceil)$ . Therefore  $\delta = m \min_{a \in {}_A\mathcal{H}} P_A(a)$ .

---

<sup>1</sup>For clarity, this refers to all observations.

- 
- $\kappa = \gamma_1 \frac{\delta^{k-1}}{(k-1)!} \leq 1$ , where  $0 < \gamma_1 \leq 1$  is the size of the interval, where  $f(x) = \delta$ .
  - $o(1)$  is a negligible value.

Let  $p = \min_{a \in \mathcal{H}} P_A(a)$ , then  $\delta = mp$ . We simplify and approximate the above equation by

$$\begin{aligned} E(T_{k \times}) &= \frac{1}{p} \left( \ln \frac{\gamma_1}{(k-1)!} m + \gamma \right) + (k-1) \frac{1}{p} \ln \ln \kappa m + o(1) \\ &\approx \frac{1}{p} (\ln m + \gamma) + (k-1) \frac{1}{p} \ln \ln m . \end{aligned} \quad (3.7)$$

The last estimate result from approximating  $\frac{\gamma_1}{(k-1)!}$  and  $\kappa$  by its upper bound 1. Applying (3.7) to inequality (3.6) result in (3.5) and completes the proof.  $\square$

### 3.1.2.2 Comparison of Estimates for $2 \times$ -Exclusivity

Note that  $E(T_{1 \times}) \leq E(T_{2 \times}) \leq 2E(T_{1 \times})$ . This is similar for the case of  $2 \times$ -exclusivity in Section 3.1.1.1, below:

$$E(T_{1 \times e}) \leq E(T_{2 \times e}) \leq 2E(T_{1 \times e}) .$$

The middle expression in this inequality is estimated by (3.5) for  $k = 2$ , while the left and right expression is estimated by (3.4) in Section 3.1.1.2. We can observe a corresponding inequality for these estimates, as illustrated below:

$$\underbrace{\frac{1}{p} (\ln m + \gamma) E(T_e)}_{\approx E(T_{1 \times})} \leq \underbrace{\left( \frac{1}{p} (\ln m + \gamma) + \frac{1}{p} \ln \ln m \right) E(T_e)}_{\approx E(T_{2 \times})} \leq \underbrace{2 \frac{1}{p} (\ln m + \gamma) E(T_e)}_{\approx 2E(T_{1 \times})} ,$$

where  $p = \min_{a \in \mathcal{H}} P_A(a)$  and  $E(T_e) = \left( \frac{u-(m-1)}{u} \right)^{1-b}$ , as in Section 3.1.1.2.

Note that the closeness of the estimates of  $E(T_{2 \times e})$  is dependent on the closeness of the estimate of  $E(T_{2 \times})$ . According to Brayton [1963], the estimate (3.7) approaches  $E(T_{2 \times})$  for large value of  $m$ . This is therefore more precise than estimating  $E(T_{2 \times})$  by  $2E(T_{1 \times})$  and thus leads to a closer estimate of  $E(T_{2 \times e})$  by (3.5) than by (3.4), for large value of  $m$ .

### 3. THEORETICAL LIMIT OF ANONYMITY PROTECTION

---

However, we need to show that (3.7) also leads to reasonable estimates for even small value of  $m$ . The inequality above shows that the estimate of  $E(T_{2 \times e})$  by (3.4) exceeds that by (3.5) (for  $k = 2$ ) by a factor that is lower than 2, while the estimate of  $E(T_{1 \times e})$  is lower than (3.5) (for  $k = 2$ ).<sup>1</sup> The estimate (3.5) for  $k = 1$  is even equal to the estimate of  $E(T_{1 \times e})$  in (3.4). This shows that (3.5) (for  $k = 2$ ) reasonably estimates  $E(T_{2 \times e})$  for all Mix configurations.

#### 3.1.2.3 Effect of Alice's Traffic Distribution on $2 \times$ -Exclusivity

The estimates show that  $E(T_{2 \times}) \propto \frac{1}{p}$ , where  $p = \min_{a \in {}_A\mathcal{H}} P_A(a)$ , thus implying that  $E(T_{2 \times})$  is minimal if Alice's traffic is uniformly distributed. Consequently the mean number of observations for  $2 \times$ -exclusivity with respect to a random set of Alice's friends  ${}_A\mathcal{H}$  is in the general case minimal, if Alice's traffic is uniformly distributed. This is because  $E(T_{2 \times e}) \approx E(T_{2 \times})E(T_e) \propto \frac{1}{p}$ .

The  $2 \times$ -exclusivity is a necessary condition for the unique identification of Alice's set of friends and determines a close lower bound for the number of observations for that identification, cf. Kesdogan et al. [2006]. Therefore, we expect that the uniform distribution of Alice's traffic also minimises the mean number of observations for the unique identification of Alice's set of friends. This is also confirmed by our simulation results, as demonstrated in Section 3.2.

#### 3.1.3 Relation to Statistical Disclosure Attack

While combinatorial attacks like the HS-attack aim at exact identification of friends, heuristic attacks, as introduced by the Statistical Disclosure attack, cf. Danezis [2004], aim at correct classification of friends with certain probabilities. Although these two approaches are orthogonal, as we will outline in Section 6, our estimate (3.5) allows relating the number of observations required by these attacks, analytically. This was not possible in the past.

---

<sup>1</sup>Note that the estimate of  $E(T_{1 \times})$  in (3.4) was derived for arbitrary integer value of  $m$ , with no particular assumption on the scope of  $m$ .

---

The *Statistical Disclosure attack* (SDA), cf. Danezis [2004] requires at least

$$\frac{1}{p^2} l^2 \left[ \sqrt{\frac{u-1}{u^2}}(b-1) + \sqrt{\frac{u-1}{u^2}}(b-1) + p^2 \left( \frac{1}{p} - 1 \right) \right]^2 \quad (3.8)$$

observations to classify recipients whom Alice contacts with the least probability  $p = \min_{a \in {}_A\mathcal{H}} P_A(a)$ . Setting  $l = 2$ ,  $l = 3$  leads to a correct classification in 95%, respectively 99% of the cases<sup>1</sup>. In case of uniformly distributed Alice’s traffic,  $p = \frac{1}{m}$  according to Danezis [2004]. For invariant  $u, b, m, l$  and non-Alice sender behaviour, the SDA requires  $O(\frac{1}{p^2})$  observations to classify all Alice’s friends while ExactHS only requires  $O(\frac{1}{p})$  observations to uniquely identify all Alice’s friends, according to (3.5).

## 3.2 Evaluation

This section demonstrates the estimate (3.5) for the mean of the least number of observations to uniquely identify Alice’s set of friends for concrete Mix configurations. Section 3.2.1 compares this estimate with the corresponding empirical mean of the least number of observations to succeed the HS-attack to confirm the closeness of the estimate. We also apply this estimate to compare the mean number of observations required to succeed the HS-attack with that to guess Alice’s friends by the SDA in Section 3.2.2

The evaluations are applied to the same Mix configurations as in the simulations in Section 2.2.4 to supports the comparability of evaluation results. We therefore assume that Alice’s traffic to her friends in  ${}_A\mathcal{H}$  is Zipf( $m, \alpha$ ) distributed, while the cover-traffic to recipients in  $R$  is uniformly distributed, as described in Section 2.2.4.1.

### 3.2.1 Estimated Number of Observations Required by HS-Attack

Figure 3.1 and Figure 3.2 visualise the empirical mean number of observations to succeed the HS-attack, labelled (HS) and to fulfil  $2 \times$ -exclusivity labelled (2x-excl), obtained from simulations. These are compared with the estimate (3.5) of the mean of

---

<sup>1</sup>This refers to the true-positive rate of the SDA. The false-positive rate might be arbitrary large and has (to the best of our knowledge) not been mathematically analysed for the general case, yet.

### 3. THEORETICAL LIMIT OF ANONYMITY PROTECTION

the least number of observations for  $2\times$ -exclusivity, labelled (2x-excl-est), which is:

$$E(T_{2\times e}) \approx \frac{1}{p} ((\ln m + \gamma) + \ln \ln m) \left(1 - \frac{(m-1)}{u}\right)^{1-b}. \quad (3.9)$$

Since Alice's traffic is  $\text{Zipf}(m, \alpha)$  distributed according to the probability mass function  $P_z^{m,\alpha}$  in Section 2.2.4.1,  $p = \min_{a \in \mathcal{H}} P_A(a) = P_z^{m,\alpha}(m)$ .

The plots provide these comparisons for distinct Mix configurations that are modelled by the parameters  $u, b, m, \alpha$ . The y-axis always shows the mean number of observations, while the x-axis vary one of the parameters  $u, b, m, \alpha$ . Counting from left to right and from top to bottom, Alice's traffic is uniformly distributed in the first three plots (i.e.,  $\alpha = 0$ ) and non-uniformly distributed in the fourth plot in Figure 3.1 and Figure 3.2. We can observe that the estimate (3.5) provides reasonable approximations, even for the more complex cases, where Alice's traffic is non-uniformly distributed.

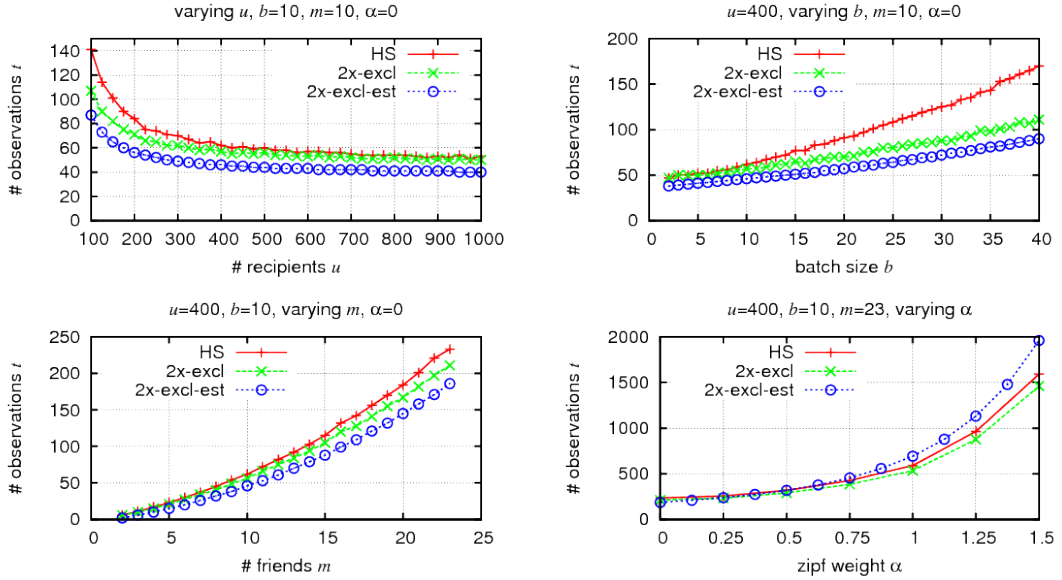


Figure 3.1: Mean number of observations: to succeed HS-attack (HS) and to fulfil  $2\times$ -exclusivity (2x-excl) versus estimated mean for  $2\times$ -exclusivity (2x-excl-est).

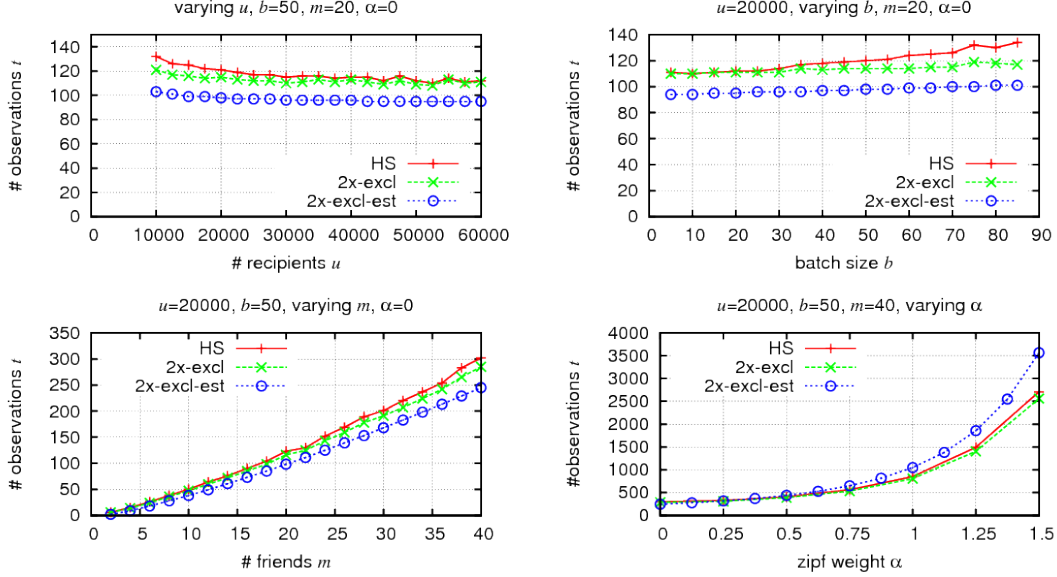
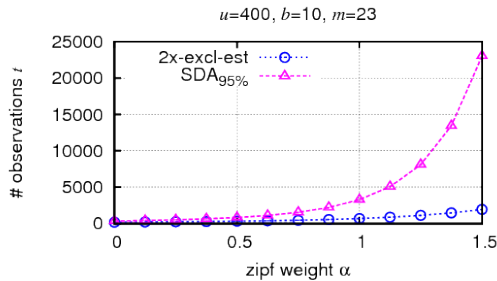


Figure 3.2: Mean number of observations: to succeed HS-attack (HS) and to fulfil  $2\times$ -exclusivity (2x-excl) versus estimated mean for  $2\times$ -exclusivity (2x-excl-est).

### 3.2.2 Number of Observations Required by HS-attack and SDA

We compare the mean number of observations to succeed the HS-attack as estimated by (3.5) with the number of observations required by the SDA to correctly classify Alice's friends with a true-positive rate of 95%, as provide by (3.8), cf. Danezis [2004]. This serves to concretely illustrate that the SDA requires a number of observations that is by the factor of  $O(\frac{1}{p})$  higher than those required by ExactHS, where  $p$  is the least probability in the distribution of Alice's friends.



$\alpha$	$p$	2x-excl-est	SDA <sub>95%</sub>
0.0	0.0435	186	343
0.5	0.0253	319	840
1.0	0.0116	693	3282
1.5	0.0041	1960	23036

$p = \min_{a \in A^{\mathcal{H}}} P_A(a) = P_z^{23, \alpha}(23)$   
in Zipf(23,  $\alpha$ ) distribution

Figure 3.3: Estimated number of required observations: HS-attack (2x-excl-est) versus SDA with 95% true-positive classification (SDA<sub>95%</sub>), for  $u = 400$ ,  $b = 10$ ,  $m = 23$ .

### 3. THEORETICAL LIMIT OF ANONYMITY PROTECTION

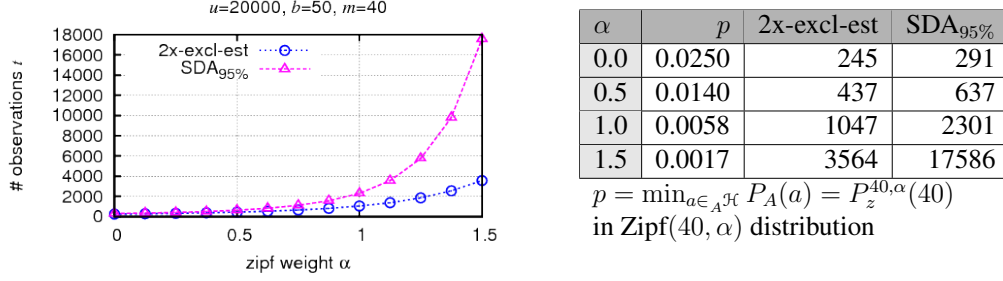


Figure 3.4: Estimated number of required observations: HS-attack (2x-excl-est) versus SDA with 95% true-positive classification (SDA<sub>95%</sub>), for  $u = 20000$ ,  $b = 50$ ,  $m = 40$ .

Figure 3.3 and Figure 3.4 provide evaluations for the Mix parameters ( $u = 400$ ,  $b = 10$ ,  $m = 23$ ), respectively ( $u = 20000$ ,  $b = 50$ ,  $m = 40$ ) and Zipf( $m, \alpha$ ) distributed Alice's traffic. These are the same parameters as considered in the bottom right plots in Figure 3.1 and Figure 3.2. The plots draw the estimated number of observations to succeed the HS-attack based on (3.5) labelled by (2x-excl-est) and to classify Alice's friends with a true-positive rate of 95% by the SDA based on (3.8) (for  $l = 2$ ) labelled by (SDA<sub>95%</sub>).

We observe that the number of observations required by the SDA increasingly exceeds that required by the HS-attack for increasing value of  $\alpha$ , as  $p$  decreases with increasing  $\alpha$ .

Note that (3.8) solely considers the true-positive rate of the SDA, that is given an Alice's friend, that friend is classified as Alice's friend with a certain rate (e.g., 95% in Figure 3.3 and Figure 3.4). However, the false-positive rate can be arbitrary large and estimating it for general cases is hard, as outlined by Pérez-González and Troncoso [2012] and in Section 6.2.2.4. When the SDA terminates, there is thus an unknown number of non-friends who are classified as Alice's friends, whereas there is a unique identification of Alice's set of friends, when the HS-attack terminates. We ignore in this evaluation that the SDA requires a learning phase to estimate the distribution of the cover-traffic and assume that this information is a-priori known to the attacker.



---

### 3.3 Summary

In this chapter, we analytically revealed the relation between the least number of observations required by the HS-attack to uniquely identify Alice’s set of friends and the Mix configuration. This relation is mathematically described by (3.5) that provides an estimate of the least number of observations to succeed the HS-attack with respect to arbitrary Mix parameters  $u, b, m$ , uniformly distributed cover-traffic and arbitrary distribution of Alice’s friends. As discussed in Section 3.1.1.2 and Section 2.2.4.1, the uniform distribution of the cover-traffic can model a bound of a real (non-uniformly distributed) cover-traffic, so that our estimate also applies to non-uniformly distributed cover-traffic. The loss of precision of the estimate in that case is not focused in this thesis and is left for future works.

Considering non-uniformly distributed traffic is important, as the real user traffic is not uniformly distributed. We could see by the evaluation in Section 3.2 and by our analytical estimate (3.5) that non-uniformly distributed communication of Alice leads to an increase of the number of observations to succeed the HS-attack. As shown in (3.5) that number of observations is proportional to  $\frac{1}{p}$ , where  $p$  is the least probability in the distribution of Alice’s friends. This supports Mix designers to adjust, when particular communication patterns (i.e., static set of friends) can be uniquely identified for traffic distributions of interest, at the design phase of the Mix system. Alice might also apply this estimate to adjust her traffic distribution to avoid the unique identifiability of her set of friends.

The closed formula (3.5) also allows comparing the number of observations to succeed the HS-attack with that required by the SDA, cf. Danezis [2004]. We chose to compare with the SDA, as it is the basis of many heuristic attacks, and efficiently applicable to non-artificial Mix configurations, cf. Section 6.2.2. Additionally, the SDA also provides a clear mathematical analysis of the true-positive rate for the classification of Alice’s friends that facilitates the comparison. One of the advantage of the HS-attack is that it requires the least number of observations to uniquely identify Alice’s set of friends. Comparing (3.5) with (3.8) shows that this number of observations is even by a factor of  $O(\frac{1}{p})$  lower than that required by the SDA to guess Alice’s friends with any bias. We also observed that other heuristic approaches, cf. Kesdogan and Pimenidis [2004]; Kesdogan et al. [2009], require significantly more observations

### 3. THEORETICAL LIMIT OF ANONYMITY PROTECTION

---

than the HS-attack. As we will outline in Section 6.2.1, these attacks must collect sufficiently many observations, such that some statistical properties required by the heuristics in those attacks become significant. The HS-attack provides a clear and precise classification of users according to the duration of the persistence of their set of friends, into those users whose friends can be uniquely identified and those for whom this is not possible with respect to the considered Mix configurations. As implied by our analyses, users whose friends can be uniquely identified by the HS-attack might be too short-lived to obtain reasonable guesses by the SDA.

The advantage of heuristic attacks is that they do not focus on providing exact identification of Alice's friends, but just on providing good guesses of friends. These guesses are based on some heuristic and usually allow computationally efficient attacks that do not rely on solving NP-complete problems. However, as illustrated in the simulations in Section 2.2.4.2, there are Mix configurations that are also computationally feasible for the HS-attack using ExactHS that will be analytically analysed in Chapter 4.

Our estimate leaves some research issues for future works. One issue is to obtain a closer estimate of the least number of observations to uniquely identify Alice's set of friends by deploying the non-trivial-disproof, instead of the  $2\times$ -exclusivity criteria, as outlined in Section 2.3. Deriving a closed formula for that estimate seems to be more complex than deriving (3.5). Another issue is the analytical estimate of the number of observations to uniquely identify any Alice's friend, instead of Alice's set of friends, in case of non-uniformly distributed traffic. We conjecture that non-uniform distributions of Alice's traffic increase the number of observations to uniquely identify Alice's set of friends, while decreasing the number of observations to identify any of Alice's friends at the same time. Section 5.1 will propose analyses towards this direction.

## Chapter 4

# Practical Limit of Anonymity Protection

Whereas the previous chapter dealt with anonymity protection against an attacker with unlimited computing resources, the practical limit considers an attacker with limited resources<sup>1</sup>. It refers to the computational complexity required by an attacker to break a system. Therefore, we assume that the attacker has collected sufficiently many observations, such that Alice’s friends are uniquely identifiable by the HS-attack.

Chapter 2 empirically showed that the practical limit of anonymity protection is more realistically measured by the “mean time”, than by the “worst case” complexity of an attack. It demonstrated the tractability of the mean time-complexity of the enhanced ExactHS in Algorithm 2 and thus of the HS-attack using ExactHS for several realistic Mix configurations, despite the exponential worse case complexity of ExactHS. This chapter proposes analytical analyses of the relations between the Mix parameters, the traffic distributions and the mean time-complexity of the enhanced ExactHS to explain those empirical observations. It allows measuring the practical limit provided by distinct Mix configurations. To the best of our knowledge, we are the first to analyse the practical limit of anonymity protection against the HS-attack that is more fine-granular than the purely theoretic worst case analyses.

Section 4.1 provides an estimate of the upper bound on the mean time-complexity

---

<sup>1</sup>The practical limit means that the protection cannot exceed that limit.

## 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

of the enhanced ExactHS that is invariant to Alice’s traffic distribution. This allows identifying Mix parameters, where ExactHS is tractable and efficient, regardless of the distribution of Alice’s friends.

The specific effect of Alice’s traffic distribution on the mean time-complexity of ExactHS is analytically analysed in Section 4.2. It proves that non-uniform distribution of Alice’s traffic can lead to an efficient mean time-complexity of ExactHS, despite Mix parameters that are intractable for the HS-attack in case of uniformly distributed traffic.

### 4.1 Upper Bound of Mean Time-Complexity

This section contributes a closed formula to estimate the upper bound of the mean time-complexity of ExactHS that is independent of the distribution of Alice’s friends, as proposed by Pham et al. [2011]. We show that the mean time-complexity of ExactHS is linearly bounded, if for every non-friend, the probability that it appears in a random observation does not exceed  $\frac{1}{m^2}$ .

Section 4.2 will show that the uniform distribution of Alice’s traffic maximises the mean time-complexity for ExactHS. Therefore, we analyse the upper bound of the mean time-complexity of ExactHS for Alice’s uniform traffic distribution to provide an estimate of the upper bound that is independent of the distribution of Alice’s friends.

Alice’s set of friends  ${}_A\mathcal{H}$  is a unique minimum-hitting-set in a given observation set  $\mathcal{OS}$ , if no other set  $\mathcal{H} \neq {}_A\mathcal{H}$ , where  $|{}_A\mathcal{H}| = m$  is a hitting-set in  $\mathcal{OS}$ . The basic scheme of ExactHS tries to construct a disproof of  $\mathcal{H}$  by choosing one recipient in  $\mathcal{H}$  in each level of recursion, thus obtaining a set of *chosen* recipients  $\mathcal{C} \subseteq \mathcal{H}$ .

A chosen recipient set  $\mathcal{C} \subseteq \mathcal{H}$  is sufficient to disprove  $\mathcal{H}$ , if the following inequality (that corresponds to Line 6 of Algorithm 2) is true:

$$\underbrace{\max_{\{r'_1, \dots, r'_{m-|\mathcal{C}|}\} \subseteq \bigcup_{\mathcal{O} \in \mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]} \mathcal{O}} \sum_{i=1}^{m-|\mathcal{C}|} |\mathcal{OS}[r'_i]|}_{\text{\# observations potentially hit by } \mathcal{H} \setminus \mathcal{C}} < \underbrace{|\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]|}_{\text{\# observations not hit by } \mathcal{C}}. \quad (4.1)$$

That is if the cumulative frequency of every  $m - |\mathcal{C}|$  distinct recipients in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$

---

is lower than the number of observations that are not yet hit by  $\mathcal{C}$ , so that  $\mathcal{H} \setminus \mathcal{C}$  cannot be a hitting set in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$ .<sup>1</sup> Provided the inequality is true, it directly follows from Claim 2 in Section 2.2.2.1 that  $\mathcal{H}$  is no hitting set in  $\mathcal{OS}$ .

If we denote by  $\mathcal{C}$  a set of chosen recipients that is sufficient to disprove  $\mathcal{H}$ , then deploying the disproof condition (4.1) would allow ExactHS to disprove  $\mathcal{H}$  at the  $|\mathcal{C}|$ -th level of recursion. We derive an upper bound  $c_{um}$  for the mean of  $|\mathcal{C}|$  with respect to all  $\mathcal{H} \neq {}_A\mathcal{H}$ , to obtain a bound of  $O(b^{c_{um}})$  for the mean time-complexity of ExactHS.

Section 4.1.1 derives an estimate of the number of observations hit by any single set  $\mathcal{H} \neq {}_A\mathcal{H}$  in  $\mathcal{OS}$ , given any chosen set  $\mathcal{C} \subseteq \mathcal{H}$ , that is called the *potential*  $Po(\mathcal{H}, \mathcal{C})$ . This potential never underestimates the number of observations hit by  $\mathcal{H}$ , so that  $\mathcal{H}$  is provably no hitting-set in  $\mathcal{OS}$ , if the difference  $Po(\mathcal{H}, \mathcal{C}) - |\mathcal{OS}|$  is negative<sup>2</sup>.

The mean of this difference is determined with respect to the uniform traffic distribution in Section 4.1.2. Section 4.1.3 derives the upper bound of the number of recipient choices  $c_{um}$ , such that this mean difference is negative with respect to all  $\mathcal{H} \neq {}_A\mathcal{H}$  and a  $\mathcal{C} \subseteq \mathcal{H}$  that maximises  $Po(\mathcal{H}, \mathcal{C})$ ,<sup>3</sup> for  $|\mathcal{C}| \geq c_{um}$ .

The closed formula for the upper bound of the mean time-complexity of ExactHS (as implemented in Algorithm 2) with respect to the Mix parameters and  $c_{um}$  is provided in Section 4.1.4. This mathematical bound is confirmed by simulation results in Section 4.1.5.

### 4.1.1 Potential – Estimate of Number of Observations Hit by a Hypothesis

In this section we introduce the definition of the *potential* our estimate of the number of distinct observations hit by a set  $\mathcal{H} \neq {}_A\mathcal{H}$  in a given observation set  $\mathcal{OS}$ . This value allows us to estimate the number of recipient choices required to disprove a set, and thus to understand the mean complexity of ExactHS. Note that this “estimate” is part of our analysis of the complexity, and does not affect the exactness of ExactHS itself.

---

<sup>1</sup>A recipient  $r \in R$  is in an observation set  $\mathcal{OS}$ , if there is an observation  $\mathcal{O} \in \mathcal{OS}$ , where  $r \in \mathcal{O}$ .

<sup>2</sup>This is a sufficient, but not necessary condition for the disproof of  $\mathcal{H}$  that is similar to (4.1), but more fine grained.

<sup>3</sup>This assumes that ExactHS prefers choosing recipients who maximise the number of recipient choices to disprove a hypothesis.

#### 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

To ease the notation, we assume without loss of generality that all considered sets are of the structure

$$\mathcal{H} = \left\{ \underbrace{r_1, \dots, r_c}_{\text{chosen recipients}}, \underbrace{r_{c+1}, \dots, r_m}_{\text{non-chosen recipients}} \right\} .$$

Each  $r_i$  represents a distinct recipient, and the number of recipients is  $|\mathcal{H}| = m$ . The first  $c$  recipients constitute the set  $\mathcal{C} = \{r_1, \dots, r_c\}$  that we call the *chosen recipient set*, where each recipient  $r_i \in \mathcal{C}$ , for  $1 \leq i \leq c$  is called a *chosen recipient* and the variable  $c = |\mathcal{C}|$  always denotes the number of chosen recipients. The remaining  $(m - c)$  recipients in  $\mathcal{H}$  are called *non-chosen* recipients. We define that the set of all observations hit by the chosen recipients in  $\mathcal{OS}$ , that is  $\mathcal{OS}[\mathcal{C}]$ , is exactly known, whereas, only the frequency of every non-chosen recipient is known. That is for every  $r_i \in \mathcal{C}$ ,  $1 \leq i \leq c$ , the frequency  $|\mathcal{OS}[r_i]|$  and  $|\mathcal{OS}[r_i] \setminus \mathcal{OS}[\mathcal{C}]|$  is known, where the latter results from knowing  $\mathcal{OS}[\mathcal{C}]$ .

The potential of  $\mathcal{H}$  with respect to  $\mathcal{C}$  is denoted  $Po(\mathcal{H}, \mathcal{C})$ .

$$Po(\mathcal{H}, \mathcal{C}) = \underbrace{|\mathcal{OS}[\mathcal{C}]|}_{\text{\# observ. hit by all chosen recipients}} + \sum_{i=c+1}^m \underbrace{|\mathcal{OS}[r_i] \setminus \mathcal{OS}[\mathcal{C}]|}_{\text{\# observ. hit by non-chosen recipients in } \mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]} \quad (4.2)$$

There are two extreme cases. If all recipients are chosen, that is if  $c = m$ , then the observations hit by  $\mathcal{H}$  are exactly known. The potential is therefore the number of observations hit by  $\mathcal{H}$ .

If all recipients are non-chosen, that is  $c = 0$ , then it is not known, which observations contain recipients in  $\mathcal{H}$ . The potential is therefore the cumulative frequency of the recipients in  $\mathcal{H}$  in  $\mathcal{OS}$ . This case is analysed in Section 4.1.1.1 as a simple starting point to analyse the general case.

In the general case,  $Po(\mathcal{H}, \mathcal{C})$  is the sum of the exact number of observations  $|\mathcal{OS}[\mathcal{C}]|$  containing any of the  $c$  chosen recipients and the cumulative frequency of each non-chosen recipient in  $\mathcal{H} \setminus \mathcal{C}$  in the remaining observations  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$ . We derive the potential function for this case in Section 4.1.1.2.

The more recipients are chosen in  $\mathcal{H}$ , the more accurately the potential represents the number of distinct observations intersecting with  $\mathcal{H}$ . While never underestimating

the set of observations intersecting with  $\mathcal{H}$  the potential might overestimate due to observations that cover more than one non-chosen recipient, as visualised in Figure 4.1 and explained in the sequel.

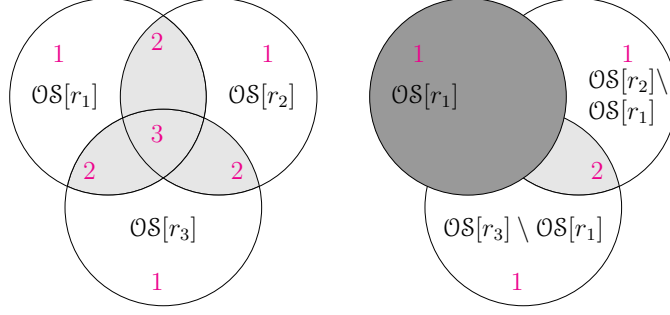


Figure 4.1: Overestimation by potential. Light grey area represents overestimation. *Left:*  $Po(\{r_1, r_2, r_3\}, \{\})$ ; all recipients are non-chosen. *Right:*  $Po(\{r_1, r_2, r_3\}, \{r_1\})$ ;  $r_1$  is chosen.

#### 4.1.1.1 Potential in Case of no Chosen Recipient

The set of observations covered by recipient  $r_i$  is represented by a circle around  $OS[r_i]$  for  $i = 1, 2, 3$  in the left-hand picture in Figure 4.1. The grey area represents those observations that are covered by at least two recipients  $r_i, r_j$  for  $i \neq j$ . The number in the area shows the number of times observations in that area are counted in the potential. In this example, the considered set is  $\mathcal{H} = \{r_1, r_2, r_3\}$ , where  $\mathcal{C} = \{\}$  is the chosen recipient set. We study this marginal case, as a simple starting point to analyse the general case of the potential function.

The left picture in Figure 4.1 illustrates how  $Po(\mathcal{H}, \{\})$  overestimates  $|OS[\mathcal{H}]|$ , which is the number of observations covered by  $\mathcal{H}$ . The overestimation is caused by those observations that are covered by more than one of the recipients  $r_1, r_2, r_3$ . The exact number of observations covered by  $\mathcal{H}$  in the left picture in Figure 4.1 can be computed by the following inclusion-exclusion formula:

$$|OS[\mathcal{H}]| = |OS[r_1]| + |OS[r_2]| + |OS[r_3]| - |OS[r_1] \cap OS[r_2]| - |OS[r_1] \cap OS[r_3]| - |OS[r_2] \cap OS[r_3]| + |OS[r_1] \cap OS[r_2] \cap OS[r_3]|. \quad (4.3)$$

In this case, all recipients in  $\mathcal{H}$  are non-chosen, therefore the potential as formu-

#### 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

lated in (4.2) is the frequency of each recipient in  $\mathcal{OS}$ , which is

$$Po(\mathcal{H}, \{\}) = |\mathcal{OS}[r_1]| + |\mathcal{OS}[r_2]| + |\mathcal{OS}[r_3]| .$$

Combining (4.3) and the simplification of (4.2) above, we reformulate the potential, such that the overestimation of  $\mathcal{OS}[\mathcal{H}]$  is mathematically directly visible, next:

$$Po(\mathcal{H}, \{\}) \leq |\mathcal{OS}[\mathcal{H}]| + \underbrace{|\mathcal{OS}[r_1] \cap \mathcal{OS}[r_2]| + |\mathcal{OS}[r_1] \cap \mathcal{OS}[r_3]| + |\mathcal{OS}[r_2] \cap \mathcal{OS}[r_3]|}_{\text{overestimation}}$$

This estimation by combining (4.2) and the inclusion-exclusion formula provides the basic idea to reveal the overestimation that is applied to the general case in the next section.

##### 4.1.1.2 Potential in General Case

The case when one recipient is chosen, that is  $\mathcal{C} = \{r_1\}$ , while the other recipients in  $\mathcal{H}$  are non-chosen is illustrated by the right-hand picture of Figure 4.1. Applying the formulation (4.2) of the potential for  $Po(\{r_1, r_2, r_3\}, \{r_1\})$  specifies that choosing  $r_1$  causes all observations containing it, represented by the dark circle, to be removed in the frequency consideration of the non-chosen recipients. In this case  $Po(\mathcal{H}, \mathcal{C})$  overestimates  $|\mathcal{OS}[\mathcal{H}]|$  by double-counting the grey area that represents observations that are covered by  $r_2$  and  $r_3$  but not by  $r_1$ . Combining (4.3) and (4.2) result in the following estimation of the potential and its overestimate of  $\mathcal{OS}[\mathcal{H}]$  in this example:

$$Po(\mathcal{H}, \{r_1\}) \leq |\mathcal{OS}[\mathcal{H}]| + |\mathcal{OS}[r_2] \cap \mathcal{OS}[r_3]| .$$

After illustrating the idea of estimating the potential for the concrete cases in Figure 4.1, we are now ready to derive the estimation of the potential in the general case. In general, given a set  $\mathcal{H}$  and a subset  $\mathcal{C} = \{r_1, \dots, r_c\}$  of  $0 \leq c \leq m$  chosen recipients, the overestimation of the number of covered observations result from the non-chosen recipients  $r_k, r_l$  for  $c < k, l \leq m$  in  $\mathcal{H} \setminus \mathcal{C}$ .

The overestimation is bounded by the size of the  $\binom{m-c}{2}$  pairwise intersections  $\mathcal{OS}[r_k] \cap \mathcal{OS}[r_l]$ . This results in the following simplified estimation of the potential



---

and its overestimation of  $\mathcal{OS}[\mathcal{H}]$  in the general case:

$$Po(\mathcal{H}, \mathcal{C}) \leq |\mathcal{OS}[\mathcal{H}]| + \underbrace{\sum_{c < k, l \leq m; k \neq l} |\mathcal{OS}[r_k] \cap \mathcal{OS}[r_l]|}_{\text{overestimation}} . \quad (4.4)$$

#### 4.1.1.3 Difference Between Potential and Number of Observations

A set  $\mathcal{H}$  is sufficiently disproved by  $\mathcal{C}$  with respect to a given observation set  $\mathcal{OS}$ , if  $Po(\mathcal{H}, \mathcal{C}) < |\mathcal{OS}|$ . That is if the difference between the potential and the number of observations collected by the attacker is less than 0. This section derives a formula for the difference between the potential and the number of observations, allowing us to determine the number of recipient choices  $c = |\mathcal{C}|$  to disprove  $\mathcal{H}$ .

In order to distinguish the effect of Alice's friends and non-friends to  $Po(\mathcal{H}, \mathcal{C})$ , each recipient  $r \in \mathcal{H}$  is relabelled  $n$  for non-friend, and  $a$  for Alice's friend. Without loss of generality, every  $\mathcal{H} \in \mathfrak{H}_j$ , where  $|\mathcal{H}| = m$  from now on has the following structure:

$$\mathcal{H} = \underbrace{\{n_1, \dots, n_{c_N}, a_1, \dots, a_{c_A}\}}_{c \text{ chosen recipients}} , \underbrace{\{n_{c_N+1}, \dots, n_j, a_{c_A+1}, \dots, a_{m-j}\}}_{(m-c) \text{ non-chosen recipients}} .$$

The number of chosen recipients is  $c = c_N + c_A$ , where  $c_N \leq j$  denotes the number of chosen non-friends, while  $c_A \leq m - j$  denotes the number of chosen friends. The variable  $j$  denotes the number of non-friends in hitting sets of the structure  $\mathfrak{H}_j$ . We still use the notation  $r_i$  to address the  $i$ -th recipient in  $\mathcal{H}$  if a distinction is not important. As before, the first  $c$  recipients  $r_1, \dots, r_c \in \mathcal{H}$  are chosen, while the remaining  $(m - c)$  recipients are non-chosen. We define  $\mathcal{H}_A = \mathcal{H} \cap {}_A\mathcal{H}$  as the subset containing only Alice's friends and  $\mathcal{H}_N = \mathcal{H} \setminus {}_A\mathcal{H}$  as the subset consisting of only non-friends.

The following estimates for  $|\mathcal{OS}[\mathcal{H}]|$  and  $|\mathcal{OS}|$  will be used next in inequality (4.8):

$$|\mathcal{OS}[\mathcal{H}]| \leq |\mathcal{OS}[\mathcal{H}_A]| + \sum_{n \in \mathcal{H}_N} |\mathcal{OS}[n] \setminus \mathcal{OS}[\mathcal{H}_A]| \quad (4.5)$$

$$|\mathcal{OS}| \geq |\mathcal{OS}[\mathcal{H}_A]| + \sum_{a \in ({}_A\mathcal{H} \setminus \mathcal{H}_A)} \underbrace{|\mathcal{OS}[a] \setminus \mathcal{OS}[_A\mathcal{H} \setminus \{a\}]|}_{\text{observ. containing } a \text{ exclusively}} . \quad (4.6)$$

Remember from Section 2.3.2.2 that an observation contains Alice's friend  $a \in {}_A\mathcal{H}$

#### 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

exclusively, cf. Kesdogan et al. [2006], if it does not contain any other friend of Alice.

We next formulate the difference between the potential and the number of observations in the inequality (4.7) to derive the number of recipient choices  $c = |\mathcal{C}|$  to disprove  $\mathcal{H}$ . If  $Po(\mathcal{H}, \mathcal{C}) < |\mathcal{OS}|$ , so that inequality (4.7) is invalid, we can conclude that  $c = |\mathcal{C}|$  is sufficient large to disprove  $\mathcal{H}$ . This difference is simplified in (4.8) that result from applying (4.4) and (4.5) to  $Po(\mathcal{H}, \mathcal{C})$ , and applying (4.6) to  $|\mathcal{OS}|$ .

$$\begin{aligned}
0 &\leq Po(\mathcal{H}, \mathcal{C}) - |\mathcal{OS}| & (4.7) \\
&\leq \sum_{c_A < k, l \leq m-j; k \neq l} |\mathcal{OS}[a_k] \cap \mathcal{OS}[a_l]| + \sum_{c_A < k \leq m-j; c_N < l \leq j} |\mathcal{OS}[a_k] \cap \mathcal{OS}[n_l]| \\
&\quad + \sum_{c_N < k, l \leq j; k \neq l} |\mathcal{OS}[n_k] \cap \mathcal{OS}[n_l]| + \sum_{n \in \mathcal{H}_N} |\mathcal{OS}[n] \setminus \mathcal{OS}[\mathcal{H}_A]| \\
&\quad - \sum_{a \in ({}_A\mathcal{H} \setminus \mathcal{H}_A)} |\mathcal{OS}[a] \setminus \mathcal{OS}[_A\mathcal{H} \setminus \{a\}]| & (4.8)
\end{aligned}$$

For simplicity we restrict our analysis to those cases where the probability that a particular recipient  $r \in \mathcal{H}$  is contacted by a sender other than Alice, within a given observation  $\mathcal{O}$ , is significantly lower than the probability that Alice's friend is contacted by Alice. This allows us to ignore the possibility that some pair of recipients  $r_k, r_l \in \mathcal{H}$  is contacted by senders other than Alice in the same  $\mathcal{O}$ . This allows us to ignore counting the observations described below in (4.8):

$$\{\mathcal{O} \in \mathcal{OS}[r_k] \cap \mathcal{OS}[r_l] \mid r_k, r_l \in \mathcal{H} \text{ chosen by non-Alice senders in } \mathcal{O}\} . \quad (4.9)$$

We call the resulting simplified estimation of (4.8) the *difference* function  $D(c, c_N, c_A, j)$ , which is

$$\begin{aligned}
&\sum_{c_A < k, l \leq m-j; k \neq l} |\mathcal{OS}[a_k] \cap \mathcal{OS}[a_l]| + \sum_{c_A < k \leq m-j; c_N < l \leq j} |\mathcal{OS}[a_k] \cap \mathcal{OS}[n_l]| + \\
&\sum_{n \in \mathcal{H}_N} |\mathcal{OS}[n] \setminus \mathcal{OS}[\mathcal{H}_A]| - \sum_{a \in ({}_A\mathcal{H} \setminus \mathcal{H}_A)} |\mathcal{OS}[a] \setminus \mathcal{OS}[_A\mathcal{H} \setminus \{a\}]| . \quad (4.10)
\end{aligned}$$

---

### 4.1.2 Mean Difference Between Potential and Number of Observations

In this section, we analyse the expectation of the difference function for a concrete and simplified communication traffic model of Alice and the other senders, which we call the *uniform communication model*. In the uniform communication model, the cover-traffic and Alice's traffic is uniformly distributed, as defined in Section 2.2.4.1.

The upper bound of the mean time-complexity of ExactHS estimated in Section 4.1.4 will be inferred from our analyses of the expectation of the difference function for this uniform communication model.

It is sufficient to derive the mean time-complexity of ExactHS for this uniform communication model to obtain an upper-bound of that complexity that is also valid for non-uniformly distributed traffic. Provided that we have the mean time-complexity analysis of ExactHS for the uniform communication model, this can be used to bound the mean time-complexity for non-uniformly distributed traffic. We can analyse given Mix parameters  $(\tilde{u}, b, m)$  with non-uniformly distributed cover-traffic by analysing the corresponding Mix parameters  $(u, b, m)$  of a uniformly distributed cover-traffic, as sketched in Section 2.2.4.1. In that analysis, considering a uniform distribution of Alice's traffic provides an upper bound of the mean time-complexity of ExactHS. That is a non-uniform distribution of Alice's traffic would not lead to a higher mean time-complexity, as will be shown in Section 4.2.

#### 4.1.2.1 Expectation of the Difference

In the uniformly distributed cover-traffic (which is without Alice's traffic), each recipient  $r \in R$  appears with the same *cumulative probability* of  $P_N$  in an observation. That is the probability that any of the  $(b - 1)$  senders of the cover-traffic contacts a given recipient  $r \in R$  in an observation is the cumulative probability  $P_N$ , for every  $r \in R$ . As mentioned in Section 2.2.4.1 each sender can select its recipient according to an individual distribution, provided that the cumulative probability is  $P_N$ . To simplify our analysis we assume that, in every batch, each of the  $(b - 1)$  non-Alice senders choose their recipients uniformly from the set  $R$  of  $u$  recipients with probability  $\frac{1}{u}$ . The cumulative probability is thus  $P_N = 1 - (\frac{u-1}{u})^{b-1}$  for every  $r \in R$ .

The uniform distribution of Alice's traffic means that in every observation, Alice

#### 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

contacts one of her  $m$  friends  $a \in {}_A\mathcal{H}$  with the probability mass function  $P_A = \frac{1}{m}$ .

We apply the above description of the cover-traffic and Alice's traffic in the uniform communication model to (4.10). This results in the next four equations that represent the expectation of the four expressions in (4.10), where the number of observations collected by the attacker is  $t = |\mathcal{OS}|$ .

$$\begin{aligned} E_1(c, c_N, c_A, j) &= t \binom{m-j-c_A}{2} \frac{2}{m} P_N \\ E_2(c, c_N, c_A, j) &= t(j-c_N)(m-j-c_A) \frac{1}{m} P_N \\ E_3(c, c_N, c_A, j) &= tj \frac{j}{m} P_N \\ E_4(c, c_N, c_A, j) &= tj \frac{1}{m} \left(1 - \frac{m-1}{u}\right)^{b-1} \end{aligned}$$

The sum of these expectations is the estimate of the mean of  $Po(\mathcal{H}, \mathcal{C}) - |\mathcal{OS}|$ .

The expressions following  $t$  in  $E_1, E_2, E_3, E_4$  are significant and are discussed next.

$E_1$ : For Alice's friends  $a_k, a_l \in \mathcal{H}_A$ , where  $a_k \neq a_l$ , the probability that Alice contacts  $a_k$  and one of the other  $(b-1)$  senders contact  $a_l$  in an observation is  $\frac{1}{m} P_N$ . Due to symmetry, the probability that  $a_k$  and  $a_l$  appear in an observation is  $\frac{2}{m} P_N$ . This is multiplied by the number of possible pairs of non-chosen Alice's friends  $\binom{m-j-c_A}{2}$ .

$E_2$ : For recipients  $a_k \in \mathcal{H}_A$  and  $n_l \in \mathcal{H}_N$ , the probability that Alice contacts  $a_k$  and one of the other  $(b-1)$  senders contacts  $n_l$  is  $\frac{1}{m} P_N$ . The factor  $(m-j-c_A)$  shows the number of non-chosen Alice's friends  $a_k$  while the factor  $(j-c_N)$  represents the number of non-chosen non-friends  $n_l$ .

$E_3$ : Let  $a_1, \dots, a_j \in ({}_A\mathcal{H} \setminus \mathcal{H})$  be the  $j$  Alice's friends who are not in  $\mathcal{H}$ . The probability that a given non-friend  $n_k \in \mathcal{H}_N$  appears in an observation where Alice contacts one of  $a_1, \dots, a_j$  is  $\frac{j}{m} P_N$ . The final factor  $j$  accounts for the fact that there are  $j$  non-friends  $n_k$  in  $\mathcal{H}_N$ .

$E_4$ : Alice's friend  $a \in ({}_A\mathcal{H} \setminus \mathcal{H})$  is exclusive in an observation if Alice contacts  $a$  and

---

none of the other  $(b - 1)$  senders contact any of the recipients  $a' \in ({}_A\mathcal{H} \setminus \{a\})$ . The probability that  $a$  is exclusive is therefore  $\frac{1}{m} \left(1 - \frac{m-1}{u}\right)^{b-1}$ . The factor  $j$  accounts for this exclusivity probability for the  $j$  Alice's friends  $a_1, \dots, a_j \in ({}_A\mathcal{H} \setminus \mathcal{H})$  not appearing in  $\mathcal{H}$ .

The sum of these these expectations results is the mean of the difference function  $D(c, c_N, c_A, j)$ , which is

$$E_D(c, c_N, c_A, j) = \frac{t}{m} \left[ ((m - c - 1)(m - j - c_A) + j^2)P_N - j \left(1 - \frac{m-1}{u}\right)^{b-1} \right] . \quad (4.11)$$

#### 4.1.2.2 Relation of Mean Difference to Number of Chosen Recipients

This section shows for any hypothesis  $\mathcal{H} \in \mathfrak{H}_j$ , that adding more chosen recipients to every chosen recipient set  $\mathcal{C} \subset \mathcal{H}$  decreases the mean (4.11) of the difference function.

**Claim 6.** *Let  $\mathcal{H} \in \mathfrak{H}_j$  be any hypothesis and  $\mathcal{C} = \mathcal{C}_i$  be the subset of chosen recipients in any sequence  $\mathcal{C}_1 \subset \dots \subset \mathcal{C}_m \subseteq \mathcal{H}$ , for  $1 \leq i \leq m$ . Given  $\mathcal{H}$  and any sequence of chosen recipients, the expectation  $E_D(c, c_N, c_A, j)$  is a monotonically decreasing function with respect to the number of chosen recipients  $c = |\mathcal{C}|$ , where  $1 \leq c \leq m - \frac{1}{2}$ .*

The proof consists of two parts. We will show that  $E_D(c, c_N, c_A, j)$  is monotonically decreasing given that  $c_N$  is fixed and then for the case that  $c_A$  is fixed.

*Monotonicity of  $E_D(c, c_N, c_A, j)$  given fixed  $c_N$ .* This analysis refers to the case that the number of chosen non-friends  $c_N$  is fixed in the number of chosen recipients  $c$ . By definition  $c_A = (c - c_N)$ , therefore we replace all  $c_A$  in (4.11) by  $(c - c_N)$ . The following function determines the gradient of the resulting function by computing its partial derivative with respect to  $c$ :  $\frac{\partial E_D(c, c_N, c - c_N, j)}{\partial c} = \frac{tP_N}{m} (2c - 2m - c_N + j + 1)$  .

This equation is less-than or equal 0, if

$$c \leq m + \frac{1}{2}(c_N - j) - \frac{1}{2} . \quad (4.12)$$

We consider the inequality (4.12) for different cases of  $(c_N - j)$ . The condition  $c_N \leq j$  is given by definition , therefore only the following cases exist:

#### 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

$c_N = j$ : In this case  $E_D$  is a decreasing function if  $c \leq m - \frac{1}{2}$ .

$c_N < j$ : In this case  $E_D$  is always a decreasing function. The proof derives from the definition  $c = (c_N + c_A)$ , where  $c_A \leq (m - j)$ . Replacing  $c_A$  in the first equation by the latter inequality, we obtain:

$$c \leq m + (c_N - j) \Rightarrow c \leq m + \frac{1}{2}(c_N - j) - \frac{1}{2}, \text{ since } c_N - j \leq -1.$$

Therefore (4.12) is always fulfilled in this case.

This proves that  $E_D(c, c_N, c_A, j)$  is a monotonically decreasing function with respect to the number of chosen recipients  $c$ , where  $1 \leq c \leq m - \frac{1}{2}$ , given that  $c_N$  is fixed.  $\square$

*Monotonicity of  $E_D(c, c_N, c_A, j)$  given fixed  $c_A$ .* We now consider the case that the number of Alice's friends is fixed in the number of chosen recipients  $c$ . The gradient of  $E_D(c, c_N, c_A, j)$  with respect to  $c$  is now:  $\frac{\partial E_D(c, c_N, c_A, j)}{\partial c} = \frac{tP_N}{m}(-m + c_A + j)$ .

The relation  $(c_A + j) \leq m$  is given by definition, therefore the gradient is always less-than or equal to 0. This proves that  $E_D(c, c_N, c_A, j)$  is a monotonically decreasing function, given that  $c_A$  is fixed.  $\square$

We conclude from these two proofs that given a hypothesis  $\mathcal{H}$  and any sequence of chosen recipients  $\mathcal{C}_1 \subset \dots \subset \mathcal{C}_m \subseteq \mathcal{H}$ ,  $E_D(c, c_N, c_A, j)$  is a monotonically decreasing function with respect to  $c = |\mathcal{C}|$  for  $\mathcal{C} \in \{\mathcal{C}_1, \dots, \mathcal{C}_m\}$  in the sequence, where  $1 \leq c \leq m - \frac{1}{2}$ . This completes the proof of Claim 6. All remaining analyses within Section 4.1 implicitly assume  $c \in [1, \dots, m - 1]$ .

##### 4.1.2.3 Relation of Mean Difference to Order of Recipient Choice

This section shows that, in general, if one prefers to chose non-friends in  $\mathcal{H} \in \mathfrak{H}_j$  first and then the remaining friends in  $\mathcal{H}$ , then the number of choices required to disprove  $\mathcal{H}$  is maximised.

**Claim 7.** *Let  $c$  be a fixed number of chosen recipients and  $c_N$  be the number of chosen non-friends, where  $c_N \leq j \leq c$ . The expectation  $E_D(c, c_N, c_A, j)$  with respect to  $c_N$  is a monotonically increasing function.*

---

*Proof.* To analyse how  $E_D$  is related to the number of non-friend choices  $c_N$ , we compute the partial derivative of  $E_D(c, c_N, c - c_N, j)$  with respect to  $c_N \leq j \leq c$ , where  $c$  is fixed. This is:  $\frac{\partial E_D(c, c_N, c - c_N, j)}{\partial c_N} = \frac{tP_N}{m}(m - c - 1)$  .

This equation is clearly greater than 0 (since  $c \leq m - 1$  is assumed), therefore  $E_D(c, c_N, c - c_N, j)$  is a monotonically increasing function for  $c_N$  in the complete interval  $[0, \dots, j]$ .  $\square$

Note that  $E_D(c, c_N, c - c_N, j)$  for  $c_N > j$  is, by definition of  $c_N$ , not defined. Let  $\mathcal{H}$  and the number of chosen recipients  $c \geq j$  be fixed, while  $\mathcal{C} \subseteq \mathcal{H}$  is variable for  $|\mathcal{C}| = c$ . Claim 7 implies that  $Po(\mathcal{H}, \mathcal{C})$  is in most of the cases maximal if  $c_N = j$  of the chosen recipients in  $\mathcal{C}$  are non-friends. Disproving  $\mathcal{H}$  therefore requires the maximal number of chosen recipients in most of the cases, if the non-friends are chosen first. To simplify the notation and because of the importance of the number of non-friends, we will replace the notation  $E_D(c, c_N, c - c_N, j)$  by the shorter notation  $E_D(c, c_N, j)$  in the sequel.

### 4.1.3 Maximal Mean Number of Recipient Choices for Disproofs

We analyse in this section the maximum of the mean number of recipient choices to disprove hypotheses. This maximum corresponds to modelling the case that ExactHS would always choose a recipient in a hypothesis  $\mathcal{H} \neq {}_A\mathcal{H}$ , such that the mean number of recipient choices required to disprove  $\mathcal{H}$  is maximal. While this assumption about the choices of recipients by ExactHS is pessimistic and not general, it is considered to provide the maximal mean number of recipient choices to disprove a single hypothesis, as well as that to disprove all hypotheses.

Section 4.1.3.1 derives a formula for the maximal mean number of recipient choices to disprove single hypotheses. The relation of this mean number is analysed with respect to the class of the hypothesis in Section 4.1.3.2. This allows identifying the hypothesis class containing hypotheses that require the maximal mean number of recipient choices for disproofs with respect to all hypotheses classes. Section 4.1.3.3 derives a closed formula for the maximal number of recipient choices to disprove all hypotheses from the results of these analyses.

## 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

### 4.1.3.1 Local Maximal Mean

In this section, we derive the maximal mean number of recipient choices required to disprove a single hypothesis  $\mathcal{H} \in \mathfrak{H}_j$ , where  $\mathcal{H} \neq {}_A\mathcal{H}$ . We obtain this by evaluating structures of chosen recipient sets  $\mathcal{C} \subseteq \mathcal{H}$ , that immanently lead to a high number of recipient choices for disproofs in most of the cases. It straight forwardly follows from Claim 7 in Section 4.1.3 that such a structure is obtained, if  $\mathcal{C}$  preferably consist of non-friends.

**Claim 8.** *Let  $\frac{u}{b-1} \geq 2(m-1)$ . The maximal mean number of recipient choices  $c$  to disprove a hypothesis  $\mathcal{H} \in \mathfrak{H}_j$ , such that  $E_D(c, c_N, j) = 0$ , is with respect to  $u, b, m, j$ :*

$$\begin{aligned} c_m(u, b, m, j) &= \operatorname{argmax}_{0 \leq c \leq m} \{E_D(c, c_N, j) = 0 \mid 0 \leq c_N \leq j\} \\ &= m - \frac{1}{2} - \sqrt{j \frac{u}{b-1} - j^2 + j - mj + \frac{1}{4}}. \end{aligned} \quad (4.13)$$

We call  $c_m(u, b, m, j)$  the *maximal mean number of recipient choices* to disprove a hypothesis. If  $(u, b, m, j)$  is clear from the context, or if their particular values are not important, we also refer to  $c_m(u, b, m, j)$  shortly by  $c_m$ .

*Proof.* In order to ensure that all non-friends are chosen first, we set  $c_N = j$ . Given this, the maximal number of recipient choices is the value  $c$ , such that  $E_D(c, c_N, j)$  in (4.11) is 0.

$$\begin{aligned} 0 &= E_D(c, j, j) \\ &\leq \frac{t}{m} \left[ ((m-c-1)(m-c) + j^2) \left(1 - \left(1 - \frac{b-1}{u}\right)\right) - j \left(1 - (b-1) \frac{m-1}{u}\right) \right]. \end{aligned}$$

We obtain (4.13) by computing the positive root of the last right hand side function for the variable  $c$ . Equation (4.13) is valid if the expression within the square root is at least 0. That is, if

$$0 \leq ju(b-1)^{-1} - j^2 + j - mj + \frac{1}{4}.$$



---

Since  $j \leq m$  the above equation holds, if

$$u(b-1)^{-1} \geq 2(m-1) \ .$$

□

Note that it is sufficient to assume  $c_N = j$  and  $c \geq j$  for the proof. There is no need to consider the case  $c < j$  for the maximal mean number of recipient choices, where  $c_N < j$ , separately. For an intuitive explanation, we assume a set  $\mathcal{H} \in \mathfrak{H}_j$  for a maximal value  $j$ , such that  $c_N = j = c$  is the maximal number of non-friend choices to disprove  $\mathcal{H}$ . Let  $\mathcal{H}' \in \mathfrak{H}_{j'}$  be another set, where  $j' > j$ . Since we assume that each Alice's friend are more frequently observed by the attacker than any non-friend, the relation  $Po(\mathcal{H}', \{\}) < Po(\mathcal{H}, \{\})$  holds in most of the cases. We can particularly deduce that  $E_D(c, c_N, j') < E_D(c, c_N, j)$ , which implies that the maximal number of recipient choices to disprove  $\mathcal{H}$ , as well as  $\mathcal{H}'$ , is  $c$ . It is thus sufficient to analyse the case  $c_N = j$  and  $c \geq j$ .

#### 4.1.3.2 Maximal Mean with respect to Hypothesis Class

We now analyse the relation of the maximal mean number of recipient choices (4.13) with respect to the hypotheses classes. It allows determining conditions, where (4.13) is a monotonically decreasing function with respect to  $\mathfrak{H}_j$ . Under these conditions, the identification of the hypothesis class containing hypotheses that requires the globally maximal mean number of recipient choices for disproofs is greatly simplified.

The next equation is the partial derivative of (4.13) with respect to  $j$ .

$$\frac{\partial c_m}{\partial j} = -\frac{1}{2} \frac{(u(b-1)^{-1} - 2j + 1 - m)}{(ju(b-1)^{-1} - j^2 + j - mj + \frac{1}{4})^{\frac{1}{2}}} \quad (4.14)$$

The value of  $c_m$  is thus monotonically decreasing, if the numerator in the above is at least 0.

$$0 \leq u(b-1)^{-1} - 2j + 1 - m \quad \Rightarrow \quad j \leq 0.5 (u(b-1)^{-1} - m + 1)$$

Thus, if the maximal number of non-friend choices in a disproof is not larger than

## 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

$\frac{1}{2}(\frac{u}{b-1} - m + 1)$ , then (4.13) is a monotonically decreasing function. If  $\frac{u}{b-1} \geq 3m - 1$ , then this case is necessarily fulfilled and we assume this condition for the remaining analyses in Section 4.1.3.

### 4.1.3.3 Global Maximal Mean

This section provides a simple closed formula for the global maximal mean number of recipient choices to disprove every hypothesis. It is the maximum of (4.13) for given Mix parameters  $u, b, m$ .

We derive the maximum of (4.13) under the condition in Section 4.1.3.2, so that (4.13) is monotonically decreasing. As shown below, obtaining a closed formula for the maximum of (4.13) with respect to  $u, b, m$  is straight forward under this condition.

**Claim 9.** *Let  $\frac{u}{b-1} \geq 3m - 1$  and  $c_m(u, b, m, j)$  be the maximal mean number of recipient choices for single hypotheses. The maximal value of  $c_m$  for fixed  $u, b, m$  is*

$$\begin{aligned} c_{um}(u, b, m) &= \max_j c_m(u, b, m, j) \\ &= m - \frac{1}{2} - \sqrt{\frac{u}{b-1} - m + \frac{1}{4}} \quad , \text{ where } 0 \leq c_{um} \leq m \quad . \end{aligned} \quad (4.15)$$

We call  $c_{um}$  the *upper bound of the mean number of recipient choices* for disproofs.

*Proof.* Let  $\frac{u}{b-1} \geq 3m - 1$ , then  $c_m$  is monotonically decreasing with respect to  $j$ . It is therefore maximal if we set  $j = 1$  in (4.13) and thus obtain (4.15).  $\square$

In case of  $\frac{u}{b-1} < 3m - 1$ , the right hand side of equation (4.15) might not provide a maximal value for  $c_{um}$ . However, we can conclude from that equation that, if  $\frac{u}{b-1} = m - \frac{1}{4}$ , then  $c_{um} \geq m - \frac{1}{2}$  and that  $c_{um}$  increases if the value of  $\frac{u}{b-1}$  decreases.

### 4.1.4 Maximal Mean Time-Complexity

The last section derives the upper bound of the mean number of recipient choices required to disprove all hypotheses  $\mathcal{H} \neq_A \mathcal{H}$ . This section applies that result to estimate the upper bound of the mean time-complexity. We outline the implementation of the potential in Algorithm 2, such that analyses of the potential and of the mean time-complexity to disprove hypotheses directly apply to ExactHS.

---

The basic ExactHS scheme in Algorithm 1 that is proposed by Pham [2008], focuses solely on minimising the worst case complexity of solving the unique minimum hitting set problem. Each recipient choice in a level of recursion in Algorithm 1 specifies a branch that directs to distinct proofs or disproofs of hypotheses for possible minimal-hitting-sets. That number of possible minimal-hitting-sets is according to Claim 3 in Section 2.2.3.1 tightly bounded by  $b^m$ . Therefore, even if Alice's set of friends  ${}_A\mathcal{H}$  is a unique minimum hitting set, there are  $b^m - 1$  hypotheses for possible minimal-hitting-sets that need to be disproved to prove the uniqueness of  ${}_A\mathcal{H}$ .

However, estimating the number of observations hit by hypotheses based on the potential and the difference function in the enhanced version of ExactHS allows cutting branches for the disproof of hypotheses. This avoids computing each of the  $b^m - 1$  single hypotheses of minimal-hitting-sets in ExactHS for disproofs and thus reduces the practical mean time-complexity of ExactHS. The application of the potential is implemented in Line 6 of Algorithm 2, that is marked by the grey colour to highlight the difference to Algorithm 1. In that line,  $\mathcal{C}$  is the set of chosen recipients, where  $|\mathcal{C}| = m - m'$ ,  $|\mathcal{OS}'| = |\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]| = |\mathcal{OS}| - |\mathcal{OS}[\mathcal{C}]|$  and  $\sum_{l=1}^{m'} |\mathcal{OS}'[r_l]| = Po(\mathcal{C} \cup \{r_1, \dots, r_{m'}\}, \mathcal{C}) - |\mathcal{OS}[\mathcal{C}]|$  for any possible  $m'$  non-chosen recipients  $r_1, \dots, r_{m'}$ . Therefore, the inequality in Line 6 of Algorithm 2 is equivalent to that in (4.7) in the analysis of the difference function in Section 4.1.1.3. These inequalities are enlisted below.

$$\sum_{l=1}^{m'} |\mathcal{OS}'[r_l]| \geq |\mathcal{OS}'| \quad (\text{Line 6 of Algorithm 2})$$

$$\Leftrightarrow Po(\mathcal{C} \cup \{r_1, \dots, r_{m'}\}, \mathcal{C}) \geq |\mathcal{OS}| \quad (\text{Inequality (4.7)})$$

The upper bound  $c_{um}$  of the mean number of recipient choices to disprove hypotheses bounds the size  $|\mathcal{C}|$  and thus the level of recursions required by the enhanced ExactHS implemented by Algorithm 2 to disprove hypotheses.

#### 4.1.4.1 Estimate

Since  $c_{um}$  is the upper bound of the mean number of recipient choices in Algorithm 2, we estimate from (4.15) in Section 4.1.3.3 that  $b^{c_{um}}$  is the upper bound of the mean number of finalised sets computed by ExactHS for the unambiguous identification of

## 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

${}_A\mathcal{H}$ . Applying (4.15), results in the following closed formula that estimates this mean:

$$b^{m-\frac{1}{2}-\sqrt{\frac{u}{b-1}-m+\frac{1}{4}}} \approx b^{m-\frac{1}{2}-\sqrt{\frac{1}{P_N}-m+\frac{1}{4}}} . \quad (4.16)$$

The estimate of the upper bound of the mean time-complexity of ExactHS results from the product of the above expression and  $tbm$ . This estimate is  $O(b^{c_{um}}tbm)$  and corresponds to the worst case time-complexity of  $O(b^mtbm)$ , if  $c_{um} = m$ .

From the relations  $P_N = 1 - (1 - \frac{1}{u})^{b-1} \approx \frac{b-1}{u}$  and  $P_A = \frac{1}{m}$  we conclude from (4.16), that:

- If every recipient not contacted by Alice is at least as likely to appear in an observation as recipients contacted by Alice, the upper bound of the mean time-complexity roughly equals the worst case complexity  $O(b^mtbm)$ . That is if  $P_N = \frac{1}{m-\frac{1}{4}}$ .
- The upper bound of the mean time-complexity becomes linear  $O(tbm)$  if every recipient not contacted by Alice appears in observations with a probability close to  $\frac{1}{m^2}$ .

### 4.1.5 Evaluation

To support our mathematical analysis, we now show the ExactHS applied to randomly generated observations. These observations are generated under the uniform communication model as analysed in Section 4.1.2. This is chosen to allow direct comparison between the simulative and our theoretical results.

In each experiment, we apply the HS-attack on random observations that are generated according to the uniform communication model until the HS-attack is successful, so that Alice's set of friends is uniquely identified by ExactHS. The average of the number of observations required in the experiments is the mean number of observation to succeed the HS-attack. We repeat the experiments until 95% of these results fall within 5% of the empirically observed mean; thereby each experiment is repeated at least 300 times.

To demonstrate that our estimate closely predicts the empirical upper bound of the mean time-complexity of ExactHS, we apply the HS-attack on observations of a Mix with parameters  $u, b, m$  that are chosen according to (4.15), where  $c_{um} = 2$ . It

---

is therefore expected that when the HS-attack succeeds, the mean time-complexity of ExactHS will be polynomial in  $O(b^2tbm)$ , while its upper bound of the mean number of recipient choices (i.e., the level of recursion) will be bounded by 2. This illustrates a range of concrete non trivial Mix parameters, where ExactHS is computationally feasible despite its infeasible worst case complexity.

**Upper Bound of Mean Number of Recipient Choices** The left plot in Figure 4.2 draws the mean number of observations to uniquely identify Alice’s set of friends for the same Mix parameters as in Figure 4.3. The straight line (HS) represents the mean of the least number of observations required by the HS-attack. The HS-attack requires very few observations when it succeeds, so that the empirical distribution of the cover-traffic strongly diverges from the function  $P_N \approx \frac{b-1}{u}$  from which they are drawn<sup>1</sup>. Due to the law of large numbers, this side effect diminishes for large number of observations. We therefore additionally consider the application of ExactHS where the number of observations is twice that required by (5.8). This is shown in the plot by the dashed line (HS2).

The right plot in Figure 4.2 draws the mean maximal number of recipient choices (i.e., level of recursion) for disproofs by ExactHS under the conditions represented by the lines (HS) and (HS2). The line (HS) shows that the number of recipient choices is notably higher than  $c_{um}$  if ExactHS identifies  ${}_A\mathcal{H}$  with the least number of observations. This is because the probabilities of many non-friends exceed  $P_N$  due to a low number of observations, and because we consider the maximal number of recipient choices for disproofs in each experiment. With more observations, as in (HS2), we can see that the mean maximal number of recipient choices is about  $c_{um}$  for all selected  $u, b, m$  as predicted by (4.15). Collecting even more additional observations when applying ExactHS, does not noticeably change the mean maximal number of recipient choices.

**Upper Bound of Mean Time-Complexity** Figure 4.3 draws the mean number of finalised sets in the experiments. The number of recipients  $u$  is determined by (4.15) with respect to a fixed  $c_{um} = 2$ , batch size  $b = 50$  and the number of Alice’s friends  $m$

---

<sup>1</sup>For example, assume that  $P_N = 1/400$ , but the HS-attack succeeds after  $|\mathcal{OS}| = 100$  observations. The empirical probability of every recipient observed in  $\mathcal{OS}$  exceeds  $P_N$  by at least a factor of 4.

#### 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

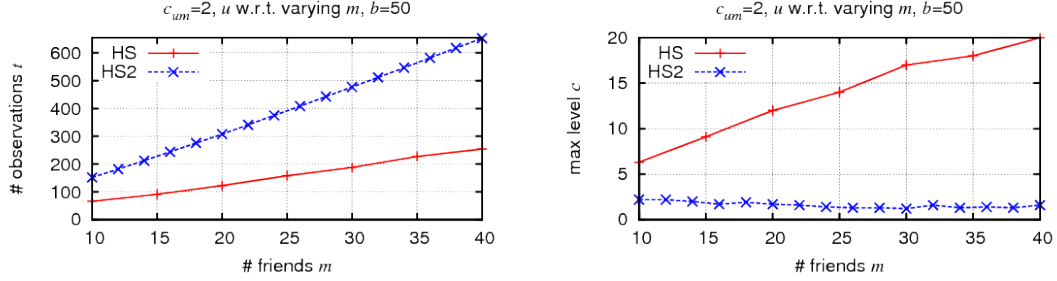


Figure 4.2: Unique identification of  ${}_A\mathcal{H}$  for  $c_{um} = 2$  and  $u, b, m$  chosen by (4.15). *Left*: Number of observations in HS-attack. *Right*: Level of recursion for disproofs by ExactHS.

on the x-axis. The value of  $u$  ranges from 3200 for  $m = 10$  to 70000 for  $m = 40$ . We can observe that the mean number of finalised sets to uniquely identify Alice's set of friends is below  $b^{c_{um}} = 2500$ , as predicted by (4.15). This is considered for the least number of observations required by ExactHS, labelled by (HS) and for the number of observations that is twice that required by (5.8), labelled by (HS2). The mean number of finalised sets computed under the condition HS2 is in the range of 14 – 36.

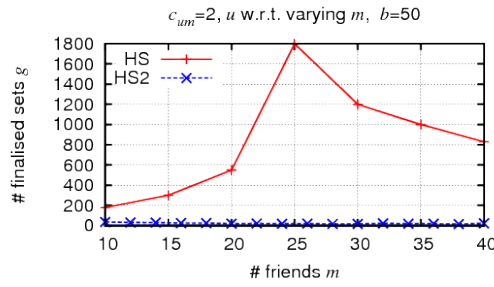


Figure 4.3: Mean number of finalised sets for unique identification of  ${}_A\mathcal{H}$ :  $u$  is determined by (4.15) for  $c_{um} = 2, b = 50$  and varying value of  $m$ .

---

## 4.2 Impact of Traffic Distribution on Mean Time-Complexity

We analyse the impact of Alice's traffic distribution, based on an estimate of a *hypothetical minimal mean time-complexity* of ExactHS. This complexity assumes that ExactHS disproves every hypothesis  $\mathcal{H} \neq {}_A\mathcal{H}$  by choosing recipients, such that the mean size of the set of chosen recipients  $\mathcal{C} \subseteq \mathcal{H}$  for that disproof is minimal. In general, the lower the required number of recipient choices to disprove hypotheses, the lower is the mean time-complexity of ExactHS. However, our assumption is hypothetical, because it is probably computational infeasible to know for every hypothesis, the recipient choices that minimises the mean number of recipient choices to disprove it. Therefore, we consider the weaker assumption that ExactHS prefers choosing in each level of recursion, the recipient first, that is most frequently contacted by Alice in that level. We call the latter assumption the *optimistic case* assumption and propose analyses of the mean time-complexity of ExactHS under this assumption. This leads to an estimate of the hypothetical minimal mean time-complexity of ExactHS from above that is dependent on the traffic distributions.

We show that the optimistic case assumption can be approached by the enhanced ExactHS, if the probability that any recipient  $r \in R$  is contacted by any sender other than Alice in an observation is below  $\frac{1}{m}$ . This is achieved by the modifications (Lines 5 to 7), provided by Algorithm 2. Therefore, the mean time-complexity of ExactHS under the optimistic case assumption is approached by the mean time-complexity of Algorithm 2, which we also confirm in simulations. The optimistic case assumption is thus not purely hypothetical and can be approached efficiently.

We prove that the mean time-complexity of ExactHS under the optimistic case assumption is maximal, if Alice's traffic is uniformly distributed. Analyses of the  $\text{Zipf}(m, \alpha)$  distribution of Alice's traffic particularly show that this complexity approaches a polynomial function for increasing value of  $\alpha$ . The Zipf distribution is known to closely model Internet traffic, as evaluated by Almeida et al. [1996]; Breslau et al. [1999]; Cunha et al. [1995]; Glassman [1994].

Let  $Po(\mathcal{H}, \mathcal{C})$  be the potential of any hypothesis  $\mathcal{H} \neq {}_A\mathcal{H}$  and chosen set  $\mathcal{C} \subseteq \mathcal{H}$  in an observation set  $\mathcal{OS}$  accumulated by the attacker, introduced in Section 4.1.1.

## 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

Section 4.2.1 refines the formulation of  $Po(\mathcal{H}, \mathcal{C})$  to distinguish the effect of the single Alice's friends and non-friends in  $\mathcal{H}$  and  $\mathcal{C}$  on the value of  $Po(\mathcal{H}, \mathcal{C})$ . This is applied in Section 4.2.2 to determine a mean of the difference  $Po(\mathcal{H}, \mathcal{C}) - |\mathcal{OS}|$  that is specific to the distribution of Alice's set of friends.

Section 4.2.3 analyses for every hypothesis  $\mathcal{H} \neq {}_A\mathcal{H}$  the pair  $(c_N, c_A)$  of minimal number of choices of non-friends and friends, where  $\mathcal{C} = \mathcal{C}_N \cup \mathcal{C}_A$ ,  $|\mathcal{C}_N| = c_N$  and  $|\mathcal{C}_A| = c_A$ , such that the mean of  $Po(\mathcal{H}, \mathcal{C}) - |\mathcal{OS}|$  is negative.<sup>1</sup> It determines the maximal minimal number of choices of non-friends  $c_N^{min}(c_A)$  in those pairs, for the values of  $c_A = 0, \dots, m - 1$ , with respect to the traffic distributions. We show that the mean number of non-friends  $c'_N$  and friends  $c'_A$  to disprove a hypothesis  $\mathcal{H}' \neq {}_A\mathcal{H}$  by ExactHS, provided the optimistic case assumption, is approximately bounded by  $(c_N^{min}(c'_A), c'_A)$ .<sup>2</sup>

Section 4.2.4 applies this information to estimate the minimal mean time-complexity of ExactHS from above that is dependent on the traffic distributions. That is the mean time-complexity of ExactHS under the optimistic case assumption.

### 4.2.1 Refined Potential – Estimate of Number of Observations Hit by a Hypothesis

In this section, we reformulate the potential  $Po$  in Section 4.1.1, such that the effect of Alice's traffic distribution and that of the cover-traffic are analytically distinguishable. This provides the basis for an in depth analysis of the relation between the required number of recipient choices to disprove hypotheses by ExactHS and Alice's traffic distribution, for the succeeding sections. Section 4.2.3.3 will deploy this analysis to derive a strategy to choose recipients to approach the mean number of recipient choices to disprove hypotheses by ExactHS under the optimistic case assumption. This strategy is implemented in the enhanced version of ExactHS in Algorithm 2 and leads to a significant lower mean time-complexity than Algorithm 1.

As defined in Section 4.1.1.3, each hypothesis  $\mathcal{H} \neq {}_A\mathcal{H}$  is described by the following

---

<sup>1</sup>The number of these pairs  $(c_N, c_A)$  would determine the hypothetical minimal mean time-complexity of ExactHS.

<sup>2</sup>We show that Algorithm 2 approaches this optimistic case assumption under realistic conditions.



---

structure:

$$\{n_1, \dots, n_{c_N}, a_1, \dots, a_{c_A}, n_{c_N+1}, \dots, n_j, a_{c_A+1}, \dots, a_{m-j}\} \ .$$

We denote by the set  $\mathcal{C} \subseteq \mathcal{H}$  the chosen recipients in  $\mathcal{H}$ , where  $\mathcal{C} = \mathcal{C}_N \cup \mathcal{C}_A$ . The set  $\mathcal{C}_N = \mathcal{C} \setminus {}_A\mathcal{H} = \{n_1, \dots, n_{c_N}\}$  represents all chosen non-friends, while  $\mathcal{C}_A = \mathcal{C} \cap {}_A\mathcal{H} = \{a_1, \dots, a_{c_A}\}$  represents all chosen Alice's friends in  $\mathcal{H}$ . The size of  $\mathcal{C}$  is denoted by  $c = c_N + c_A$ , where  $c_N = |\mathcal{C}_N|$  and  $c_A = |\mathcal{C}_A|$ .

We define the set of observations, in which Alice contacts a particular friend  $a \in {}_A\mathcal{H}$  by

$$\mathcal{OS}_A[a] = \{\mathcal{O} \in \mathcal{OS} \mid a \in \mathcal{O} \wedge \text{Alice contacts } a \text{ in } \mathcal{O}\} \ .$$

This set has the following properties:

- $\mathcal{OS}_A[a] \subseteq \mathcal{OS}[a]$ , due to observations  $\mathcal{O} \in \mathcal{OS}[a]$  in which  $a$  was chosen by a non-Alice sender, but not by Alice.
- For  $a_i \neq a_j$ ,  $\mathcal{OS}_A[a_i] \cap \mathcal{OS}_A[a_j] = \emptyset$  and  $\bigcup_{i=1}^m \mathcal{OS}_A[a_i] = \mathcal{OS}$ .

We define the set of observations, in which any non-Alice sender contacts a particular recipient  $r \in R$  by

$$\mathcal{OS}_N[r] = \{\mathcal{O} \in \mathcal{OS} \mid r \in \mathcal{O} \wedge \text{a non-Alice sender contacts } r \text{ in } \mathcal{O}\} \ .$$

This set has the following properties:

- For  $r_i \neq r_j$ ,  $\mathcal{OS}_N[r_i] \cap \mathcal{OS}_N[r_j] \subseteq \mathcal{OS}[r_i, r_j]$ .
- $\mathcal{OS}_N[n] = \mathcal{OS}[n]$ , but  $\mathcal{OS}_N[a] \subseteq \mathcal{OS}[a]$ . Note that  $\mathcal{OS}[a] = \mathcal{OS}_A[a] \cup \mathcal{OS}_N[a]$ .

Reformulating the potential  $Po$  with respect to  $\mathcal{OS}_A$  and  $\mathcal{OS}_N$ , results in (4.17). This equation is equivalent to (4.2) in Section 4.1.1, but explicitly formulates the proportion of Alice's traffic and that of the other senders in the observations evaluated by the potential.

#### 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

$$\begin{aligned}
Po(\mathcal{H}, \mathcal{C}) &= |\mathcal{OS}_A[a_1, \dots, a_{c_A}]| \\
&+ |\mathcal{OS}_N[n_1, \dots, n_{c_N}, a_1, \dots, a_{c_A}] \setminus \mathcal{OS}_A[a_1, \dots, a_{c_A}]| \\
&+ \sum_{i=c_A+1}^{m-j} |\mathcal{OS}_A[a_i] \setminus \mathcal{OS}_N[n_1, \dots, n_{c_N}, a_1, \dots, a_{c_A}]| \\
&+ \sum_{i=c_A+1}^{m-j} |\mathcal{OS}_N[a_i] \setminus \mathcal{OS}_N[n_1, \dots, n_{c_N}, a_1, \dots, a_{c_A}] \setminus \mathcal{OS}_A[a_1, \dots, a_{c_A}, a_i]| \\
&+ \sum_{i=c_N+1}^j |\mathcal{OS}_N[n_i] \setminus \mathcal{OS}_N[n_1, \dots, n_{c_N}, a_1, \dots, a_{c_A}] \setminus \mathcal{OS}_A[a_1, \dots, a_{c_A}]|
\end{aligned} \tag{4.17}$$

The first two lines counts all chosen observations, that is, the set of observations  $|\mathcal{OS}[\mathcal{C}]| = |\mathcal{OS}[n_1, \dots, n_{c_N}, a_1, \dots, a_{c_A}]|$ . The third and fourth lines counts non-chosen observations that contain an as-yet non-chosen Alice's friend in  $\mathcal{H}$ , i.e.,  $\sum_{i=c_A+1}^{m-j} |\mathcal{OS}[a_i] \setminus \mathcal{OS}[n_1, \dots, n_{c_N}, a_1, \dots, a_{c_A}]|$ . The last line counts non-chosen observations that contain an as-yet non-chosen non-friend in  $\mathcal{H}$ . This equals the number represented by  $\sum_{i=c_N+1}^j |\mathcal{OS}[n_i] \setminus \mathcal{OS}[n_1, \dots, n_{c_N}, a_1, \dots, a_{c_A}]|$ .

**Example 4.2.1** (Refined Potential). *Figure 4.4 visualises the potential function on the set  $\mathcal{H} = \{a_1, n_1\}$ . The square frame covers all observations in  $\mathcal{OS}$ . Observations in which Alice contacts one of her two friends  $a_1, a_2$  are represented by triangles with the labels  $\mathcal{OS}_A[a_1]$  and  $\mathcal{OS}_A[a_2]$  that partitions  $\mathcal{OS}$ . The set of observations in which any non-Alice sender contacts the recipient  $a_1$ , respectively  $n_1$ , is marked by circles.*

*The horizontal pattern marks all observations containing  $a_1$ , that is  $\mathcal{OS}[a_1]$ . The vertical pattern marks all observations in  $\mathcal{OS} \setminus \mathcal{C}$  containing  $n_1$ . (These are  $\mathcal{OS}[n_1]$  for  $\mathcal{C} = \{\}$  in the first picture and  $\mathcal{OS}[n_1] \setminus \{a_1\}$  for  $\mathcal{C} = \{a_1\}$  in the second picture.)*

*The first picture in Figure 4.4 shows the case that no recipients in  $\mathcal{H}$  are chosen, i.e.,  $|\mathcal{C}| = |\{\}| = c = 0$ .*

$$\begin{aligned}
Po(\mathcal{H}, \{\}) &= |\mathcal{OS}[a_1]| + |\mathcal{OS}[n_1]| \\
&= |\mathcal{OS}_A[a_1]| + |\mathcal{OS}_N[a_1] \setminus \mathcal{OS}_A[a_1]| + |\mathcal{OS}_N[n_1]|
\end{aligned}$$

*The first equation above is formulated in term of (4.2), while the second equation is in*

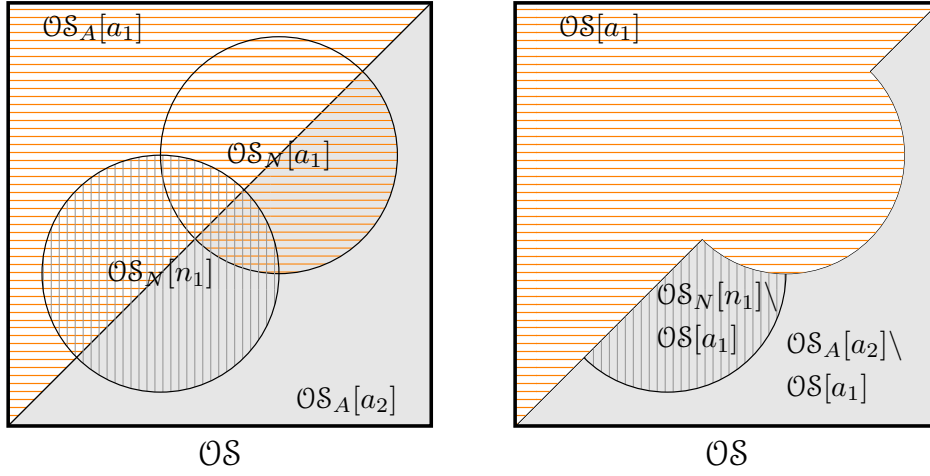


Figure 4.4: Potential of  $\mathcal{H} = \{a_1, n_1\}$  is sum of observations containing  $a_1$  (horizontal line area) and observations containing  $n_1$  (vertical line area). Left:  $Po(\mathcal{H}, \{\})$ . Right:  $Po(\mathcal{H}, \{a_1\})$ .

term of (4.17).

The area of overlapping horizontal and vertical patterns visualises those observations that intersect with more than one recipient in  $\mathcal{H}$ . These observations are counted twice in  $Po(\mathcal{H}, \mathcal{C})$  and thus lead to an overestimation of the real number of observations hit by  $\mathcal{OS}[\mathcal{H}]$ .

The second picture in Figure 4.4 shows the case that  $\mathcal{C} = \{a_1\}$  is chosen, while  $n_1$  is non-chosen. That is  $|\mathcal{C}| = c = 1$ .

$$\begin{aligned} Po(\mathcal{H}, \{a_1\}) &= |\mathcal{OS}[a_1]| + |\mathcal{OS}[n_1] \setminus \mathcal{OS}[a_1]| \\ &= |\mathcal{OS}_A[a_1]| + |\mathcal{OS}_N[a_1] \setminus \mathcal{OS}_A[a_1]| + |\mathcal{OS}_N[n_1] \setminus \mathcal{OS}_N[a_1] \setminus \mathcal{OS}_A[a_1]| \end{aligned}$$

It illustrates that the estimated number of observations hit by  $\mathcal{H}$  becomes more precise the more recipients are chosen, i.e., the larger  $|\mathcal{C}| = c$  is. This is because we know  $\mathcal{OS}[\mathcal{C}]$  and can thus exclude it in the estimation of the number of observations hit by the remaining recipients  $\mathcal{H} \setminus \mathcal{C}$ . In our example  $\mathcal{OS}[\mathcal{C}] = \mathcal{OS}[a_1]$  is exactly known in  $Po(\mathcal{H}, \mathcal{C})$ . Only the number of observations hit by  $\mathcal{H} \setminus \{a_1\} = \{n_1\}$  in  $\mathcal{OS} \setminus \mathcal{OS}[a_1]$  is estimated in  $Po(\mathcal{H}, \mathcal{C})$ .

## 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

### 4.2.2 Mean of Potential

In this section we determine the mean of the potential with respect to the distribution of recipients contacted by non-Alice senders and the distribution of Alice's friends. This mean is applied to derive the mean number of recipient choices to disprove all hypotheses  $\mathcal{H} \neq {}_A\mathcal{H}$  in Section 4.2.3.

The probabilities that are known in our analysis are:

$P_A(a)$ : Probability that Alice contacts recipient  $a$  in an observation, i.e.  $\frac{|\mathcal{OS}_A[a]|}{|\mathcal{OS}|}$ . Note that  $\sum_{a \in {}_A\mathcal{H}} P_A(a) = 1$ , since Alice contacts one friend in each observation.

$P_N(r)$ : Probability that any of the  $(b - 1)$  non-Alice senders contact recipient  $r$  in an observation, i.e.,  $\frac{|\mathcal{OS}_N[r]|}{|\mathcal{OS}|}$ . Note that  $\sum_{r \in R} P_N(a) \geq 1$ , since each observation contains several recipients.

We consider the Chaum Mix in an open environment, so that every sender can contribute messages to a batch. Since these senders are random individuals, we assume that each single sender contacts its recipients independently from the behaviour of the other senders.

Therefore, and to ease the maths, we assume that the distribution of Alice's communication to her friends is independent of the distribution of the traffic of other senders, so that  $P_A$  and  $P_N$  are statistically independent. We similarly assume for the sake of simplicity that  $P_N$  is statistically independent with respect to any set of recipients, thus assuming that  $P_N(r_1 \wedge \dots \wedge r_k) = \prod_{l=1}^k P_N(r_l)$ .

Let  $t = |\mathcal{OS}|$  be the number of observations collected by the attacker. Let the probability  $P(r_1, \dots, r_l)$  abbreviates  $P(r_1 \vee \dots \vee r_l)$  and  $P(r_1 \dots r_l)$  abbreviates  $P(r_1 \wedge \dots \wedge r_l)$ . The expectation of the potential of a given set  $\mathcal{H}$  and  $\mathcal{C} = \{n_1, \dots, n_{c_N}, a_1, \dots, a_{c_A}\}$

---

that result from (4.17) is

$$\begin{aligned}
E(Po(\mathcal{H}, \mathcal{C})) &= t \left[ P_A(a_1, \dots, a_{c_A}) \right. \\
&\quad + P_N(n_1, \dots, n_{c_N}, a_1, \dots, a_{c_A}) (1 - P_A(a_1, \dots, a_{c_A})) \\
&\quad + \sum_{i=c_A+1}^{m-j} P_A(a_i) [1 - P_N(n_1, \dots, n_{c_N}, a_1, \dots, a_{c_A})] \\
&\quad + \sum_{i=c_A+1}^{m-j} P_N(a_i) [1 - P_N(n_1, \dots, n_{c_N}, a_1, \dots, a_{c_A})] [1 - P_A(a_1, \dots, a_{c_A}, a_i)] \\
&\quad \left. + \sum_{i=c_N+1}^j P_N(n_i) [1 - P_N(n_1, \dots, n_{c_N}, a_1, \dots, a_{c_A})] [1 - P_A(a_1, \dots, a_{c_A})] \right].
\end{aligned}$$

Note that  $\mathcal{OS}_A[a_i] \cap \mathcal{OS}_A[a_j] = \emptyset$ , therefore  $P_A(a_1, \dots, a_{c_A}) = \sum_{i=1}^{c_A} P_A(a_i)$ . The disjunct probability  $P_N(n_1, \dots, n_{c_N}, a_1, \dots, a_{c_A})$  is denoted by  $P_{cN}$  for brevity, as its specific value will be irrelevant to our succeeding analysis.

$$\begin{aligned}
E(Po(\mathcal{H}, \mathcal{C})) &\leq t \left[ \sum_{i=1}^{c_A} P_A(a_i) + P_{cN} \left( 1 - \sum_{i=1}^{c_A} P_A(a_i) \right) \right. \\
&\quad \left. + (1 - P_{cN}) \left( 1 - \sum_{i=1}^{c_A} P_A(a_i) \right) * \right. \\
&\quad \left. \underbrace{\left( \sum_{i=c_A+1}^{m-j} \frac{P_A(a_i)}{1 - \sum_{i=1}^{c_A} P_A(a_i)} + \sum_{i=c_A+1}^{m-j} P_N(a_i) + \sum_{i=c_N+1}^j P_N(n_i) \right)}_{\text{if } E(Po) \geq t, \text{ then this expression } \geq 1} \right] \\
&\tag{4.18}
\end{aligned}$$

Note that the first and second line right of  $(\leq)$  within the outermost square bracket of (4.18) adds up to 1 if the expression in the third line does not exist. Therefore if  $E(Po(\mathcal{H}, \mathcal{C})) \geq t$  then the third expression must be at least 1. This criterion can be deployed to determine the mean number of recipient choices to disprove  $\mathcal{H}$  and we show how to derive that mean number next.

## 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

### 4.2.2.1 Simplified Analysis

Our purpose in analysing the expectation of the potential of a set  $\mathcal{H} \in \mathfrak{H}_j$  is to determine the mean number of non-friends  $|\mathcal{C}_N| = c_N$  and Alice's friends  $|\mathcal{C}_A| = c_A$  that needs to be chosen to disprove  $\mathcal{H}$ , where  $\mathcal{C} \subseteq \mathcal{H}$ . This is the number of chosen recipients such that  $E(Po(\mathcal{H}, \mathcal{C})) < t$ . Using the last factor expression in (4.18), we simplify this case by analysing the following basic equality in this section, instead:

$$\begin{aligned} 1 &= \frac{\sum_{k=c_A+1}^{m-j} P_A(a_k)}{1 - \sum_{l=1}^{c_A} P_A(a_l)} + \sum_{i=c_A+1}^{m-j} P_N(a_i) + \sum_{i=c_N+1}^j P_N(n_i) \\ &= f_{Po}(\mathcal{H}, \mathcal{C}, c_N, c_A, j) . \end{aligned} \quad (4.19)$$

The three (+) separated expressions on the right hand side of (=) above are (from left to right) the probabilities that:

1. Alice contacts one of the  $(m - j - c_A)$  non-chosen Alice's friends  $\mathcal{H}_A \setminus \mathcal{C} = \{a_{c_A+1}, \dots, a_{m-j}\}$  in  $\mathcal{OS} \setminus \mathcal{OS}_A[\mathcal{C}_A]$
2. Senders other than Alice contact one of the non-chosen Alice's friends  $\mathcal{H}_A \setminus \mathcal{C}$  in  $\mathcal{OS}$
3. Senders other than Alice contact one of the  $(j - c_N)$  non-chosen non-friends  $\mathcal{H}_N \setminus \mathcal{C} = \{n_{c_N+1}, \dots, n_j\}$  in  $\mathcal{OS}$

### 4.2.3 Minimal Mean Number of Recipient Choices for Disproofs

This section analyses the hypothetical minimal mean number of recipient choices to disprove hypotheses by ExactHS. We relate this with the mean number of recipient choices to disprove hypotheses by ExactHS under the optimistic case assumption. It is proved that the latter mean number of recipient choices is maximal for uniformly distributed Alice's traffic. We outline that the optimistic case assumption can be approximated by the implementation of ExactHS in Algorithm 2 in practice. This is provided that the probability that any recipient  $r \in R$  is contacted by any sender other than Alice in an observation is below  $\frac{1}{m}$ .

The function  $f_{Po}(\mathcal{H}, \mathcal{C}, c_N, c_A, j)$  reflects the mean difference between the potential  $Po(\mathcal{H}, \mathcal{C})$  and the number of observations as outlined in Section 4.2.2.1, for

---

$\mathcal{H} \in \mathfrak{H}_j$ ,  $\mathcal{C} \subseteq \mathcal{H}$ ,  $|\mathcal{C}_N| = c_N$ ,  $|\mathcal{C}_A| = c_A$ . If the value of this function is below 1, then  $|\mathcal{C}|$  recipient choices are sufficient to disprove  $\mathcal{H}$  in most of the cases. Therefore, the pair  $(c_N, c_A)$  of minimal number of choices of non-friends and friends, such that  $\min_{\mathcal{C} \subseteq \mathcal{H}} f_{Po}(\mathcal{H}, \mathcal{C}, c_N, c_A, j) < 1$  determines the hypothetical minimal mean number of recipient choices to disprove  $\mathcal{H}$  by ExactHS. Provided these pairs  $(c_N, c_A)$  for every hypothesis  $\mathcal{H} \neq {}_A\mathcal{H}$ , all pairs with the same value of  $c_A$  represents a class induced by  $c_A$ , for  $c_A = 0, \dots, m-1$ . Each class induced by  $c_A$  contains a pair  $(c_N, c_A)$  with the highest value of  $c_N$  in that class and we refer to that value with respect to  $c_A$  by the function  $c_N^{min}(c_A)$ .

Let  $(c_N, c_A)$  be the hypothetical minimal mean number of recipient choices to disprove a hypothesis  $\mathcal{H} \neq {}_A\mathcal{H}$  by ExactHS. We show that the mean number of recipient choices to disprove  $\mathcal{H}$  by ExactHS under the optimistic case assumption is approximately bounded by  $(c_N^{min}(c_A), c_A)$ . This is the relation between these two means.

Section 4.2.3.1 contributes an estimate of  $(c_N^{min}(c_A), c_A)$  and deduces the choices of recipients according to the optimistic case assumption from that. This assumption constrains the mean number of choices of non-friends  $c_N$  and friend  $c_A$  to disprove a hypothesis  $\mathcal{H}$  by ExactHS, such that it is approximately bounded by  $(c_N^{min}(c_A), c_A)$ .

Section 4.2.3.2 provides analyses of  $c_N^{min}(c_A)$  with respect to Alice's traffic distribution. It shows for the Zipf distribution, that the mean number of recipient choices to disprove hypotheses by ExactHS under the optimistic case assumption can be reduced close to 0.

Section 4.2.3.3 shows that the optimistic case assumption can be approached by the enhanced ExactHS, implemented by Algorithm 2.

#### 4.2.3.1 Deriving Maximal Minimal Conditions

We define a classification for the hypothetical minimal number  $(c_N, c_A)$  of choices of non-friends and friends to disprove hypotheses by ExactHS, according to the value of  $c_A$ . This is used to estimate the *maximal minimal number of non-friend choices*  $c_N^{min}(c_A)$  in each of these classes, for  $c_A = 0, \dots, m-1$ . We follow from this the optimistic case assumptions that leads to a mean number of choices of non-friends  $c'_N$  and friends  $c_A$  to disprove hypotheses by ExactHS that do not exceed  $(c_N^{min}(c_A), c_A)$ .

#### 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

Let a particular value of  $c_A$  be given, we call a hypothesis  $\mathcal{H} \in \mathfrak{H}_j$ , for  $j \leq m - c_A$ ,  $c_A$ -disprovable<sup>1</sup>, if there is a  $\mathcal{C} \subseteq \mathcal{H}$ , where  $|\mathcal{C}_A| = c_A$ ,  $|\mathcal{C}_N| = c_N$  and  $Po(\mathcal{H}, \mathcal{C}) < t$ . It is called  $c_A$ -average-disprovable if  $E(Po(\mathcal{H}, \mathcal{C})) < t$ . For the sake of simplicity the conditions  $f_{Po}(\mathcal{H}, \mathcal{C}, c_N, c_A, j) \leq 1$  will be analysed instead of the latter condition.

For each value  $c_A$ ,  $c_N^{min}(c_A)$  is such defined that there exists a  $c_A$ -disprovable hypothesis  $\mathcal{H} \in \mathfrak{H}_j$  (for any  $j \leq m - c_A$ ), where  $\min_{\mathcal{C} \subseteq \mathcal{H}} \{f_{Po}(\mathcal{H}, \mathcal{C}, c_N, c_A, j)\} = 1$  and  $c_N = c_N^{min}(c_A)$ . And  $\mathcal{H}$  is selected, such that all  $c_A$ -disprovable hypotheses  $\mathcal{H}' \in \mathfrak{H}_i$  (for  $i \leq m - c_A$ ), fulfils  $\min_{\mathcal{C}' \subseteq \mathcal{H}'} \{f_{Po}(\mathcal{H}', \mathcal{C}', c'_N, c_A, i)\} \leq 1$  and  $c'_N \leq c_N^{min}(c_A)$ .

Using the definition of the function  $f_{Po}$  in (4.19), we can estimate  $c_N^{min}(c_A)$  by the following three steps.

1. The function  $\max_{\mathcal{H} \in \mathfrak{H}_j} \min_{\mathcal{C} \subseteq \mathcal{H}} \{f_{Po}(\mathcal{H}, \mathcal{C}, c_N, c_A, j)\}$  with respect to given value of  $c_N, c_A, j$  is estimated, for  $c_N \leq j \leq m - c_A$  and  $|\mathcal{C}_A| = c_A, |\mathcal{C}_N| = c_N$ .<sup>2</sup> We will refer to this estimated function by  $f_{Po}^{min}(c_N, c_A, j)$ . This is analysed in Paragraph “Simple Analysis of  $f_{Po}(c_N, c_A, j)$ ” and Paragraph “Refined Analysis of  $f_{Po}(c_N, c_A, j)$ ”.
2. The function  $\max_{j: c_N \leq j \leq m - c_A} \{f_{Po}^{min}(c_N, c_A, j)\}$  with respect to  $c_N, c_A$  is determined in Paragraph “Deriving  $f_{Po}^{min}(c_N, c_A)$ ”. It equals  $f_{Po}^{min}(c_N, c_A, c_N)$ , to that we refer to by the term  $f_{Po}^{min}(c_N, c_A)$  for brevity.
3. To estimate  $c_N^{min}(c_A)$ , the equation  $f_{Po}^{min}(c_N, c_A) = 1$  is solved for  $c_N$ . This will be demonstrated in Section 4.2.5.1.

**Simple Analysis of  $f_{Po}^{min}(c_N, c_A, j)$**  This analysis refers to step one. We estimate the maximum of  $f_{Po}(\mathcal{H}, \mathcal{C}, c_N, c_A, j)$  given that the chosen recipient set  $\mathcal{C}$  minimise

---

<sup>1</sup>If  $j > m - c_A$ , then  $\mathcal{H}$  is not  $c_A$ -disprovable, but it is  $c'_A$ -disprovable with respect to another value  $c'_A < c_A$ . Note that every set  $\mathcal{H} \neq \mathcal{H}$  is disprovable.

<sup>2</sup>Without violation of our definition of  $c_N^{min}(c_A)$ , there is no need to restrict  $\mathcal{H}$  to  $c_A$ -average-disprovable sets here.



$f_{Po}$  with respect to given values of  $c_N, c_A, j$ .

$$\begin{aligned}
& \max_{\mathcal{H} \in \mathfrak{H}_j} \min_{\mathcal{C} \subseteq \mathcal{H}} \{f_{Po}(\mathcal{H}, \mathcal{C}, c_N, c_A, j)\} \\
& \leq \max_{\mathcal{H}_A = \{a_1, \dots, a_{m-j}\}} \min_{\{a_1, \dots, a_{c_A}\} \subseteq \mathcal{H}_A} \left\{ \frac{\sum_{k=c_A+1}^{m-j} P_A(a_k)}{1 - \sum_{l=1}^{c_A} P_A(a_l)} \right\} \\
& \quad + \max_{r_1, \dots, r_{m-1-c_A}} \left\{ \sum_{l=1}^{m-c_N-c_A} P_N(r_l) \right\} \tag{4.20}
\end{aligned}$$

To simplify notation, the first  $c_A$  recipients in  $\mathcal{H}_A$  in the max function are the recipients in the min function in the first expression right of ( $\leq$ ), without restriction of generality.

The expressions within the big brackets in (4.20) are relaxations of the expressions in  $f_{Po}(\mathcal{H}, \mathcal{C}, c_N, c_A, j)$ . The second and third sums  $\sum_{l=c_A+1}^{m-j} P_N(a_l) + \sum_{l=c_N+1}^j P_N(n_l)$  in  $f_{Po}(\mathcal{H}, \mathcal{C}, c_N, c_A, j)$  are simplified by the single expression  $\sum_{l=1}^{m-c_N-c_A} P_N(r_l)$  for any  $r_l \in R$ , thus ignoring the requirement of  $\{r_1, \dots, r_{m-c_N-c_A}\} \subseteq \mathcal{H}$ .

By relabelling the indices of the first and second expressions on the right side of ( $\leq$ ), such that  $P_A(a_k)$  and  $P_N(r_l)$  are the  $k$ -th and  $l$ -th most likely recipients with respect to the probability functions  $P_A$  and  $P_N$ , we obtain the following equivalent formula for (4.20):

$$\widehat{f_{Po}^{min}}(c_N, c_A, j) = \frac{\sum_{k=c_A+1}^{m-j} P_A(a_k)}{1 - \sum_{l=1}^{c_A} P_A(a_l)} + \sum_{l=1}^{m-c_N-c_A} P_N(r_l). \tag{4.21}$$

*Proof.* We now prove that the right hand side of (4.20) and (4.21) are equal. It is obvious that the second expression in (4.20) and in (4.21) are equal. Therefore, it remains to show that the first expression in (4.20) and (4.21) are equal.

The first expression within the brackets in (4.20) is of the structure  $\frac{z-y}{1-y}$ , where  $z = \sum_{a_k \in \mathcal{H}_A} P_A(a_k)$ ,  $y = \sum_{a_l \in \mathcal{C}_A} P_A(a_l)$ ,  $\mathcal{C}_A \subseteq \mathcal{H}_A$ ,  $|\mathcal{H}_A| = m-j$ ,  $|\mathcal{C}_A| = c_A$  and  $1 \geq z \geq y \geq 0$ . The function  $\frac{z-y}{1-y}$  is monotonically increasing with respect to  $z$  and monotonically decreasing with respect to  $y$ . Since  $\mathcal{C}_A$  is determined by the min function in (4.20), it must consist of the most likely recipients in  $\mathcal{H}_A$  to obtain a large value of  $y$ .

Now let us assume that  $\mathcal{H}_A$  consists of the most likely  $m-j$  Alice's friends, so that  $z$  is maximal. We proof by contradiction that  $\frac{z-y}{1-y}$  is maximal, given that  $\mathcal{C}_A$  consists

#### 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

of the most likely  $c_A$  Alice's friends in  $\mathcal{H}_A$ . Let us assume that there is a set  $\mathcal{H}'_A$  with  $z' = z - \delta$  and most likely recipients  $\mathcal{C}'_A \subseteq \mathcal{H}'_A$  with  $y' = y - \epsilon$  for  $\delta > 0$  and  $\delta \geq \epsilon \geq 0$ , such that  $\frac{z'-y'}{1-y'} > \frac{z-y}{1-y}$ .

To maximise  $\frac{z'-y'}{1-y'}$ , the value of  $y'$  should be low, while the value of  $z'$  should be large. Therefore we consider the marginal case of  $\delta = \epsilon$ , so that  $z' = z - \epsilon$  and  $y' = y - \epsilon$ .

$$\frac{z' - y'}{1 - y'} = \frac{(z - \epsilon) - (y - \epsilon)}{1 - (y - \epsilon)} = \frac{z - y}{1 - y + \epsilon} \leq \frac{z - y}{1 - y}.$$

This inequality however contradicts our initial assumption and thus proves  $\frac{z-y}{1-y}$  to be maximal. It follows from the proof that (4.20) and (4.21) are equal.  $\square$

This analysis shows that a strategy that always prefers choosing those recipients who are most frequently contacted by Alice leads to a mean number of non-friend choices to disprove  $c_A$ -average-disprovable hypothesis that does not exceed  $c_N^{min}(c_A)$ . We call this the *optimistic case strategy*.

As will be shown in Section 4.2.3.3 and in Example 4.2.2, the optimistic case strategy can be approached, if in every level of recursion, ExactHS always chooses the recipient first, that is most frequent in the observations remaining in that level of recursion. This strategy is implemented in the enhanced version of ExactHS in Line 7 of Algorithm 2.

**Refined Analysis of  $f_{Po}^{min}(c_N, c_A, j)$**  We assume in this analysis that the optimistic case strategy can be realised in ExactHS. This implies that Alice's friends who are most frequently contacted by Alice would always be chosen and removed first when constructing chosen sets  $\mathcal{C}$  in ExactHS. We outline that this induces some constraints on the structure of hypotheses that remain to be disproved by ExactHS. These constraints are accounted for in the estimation of  $\max_{\mathcal{H} \in \mathfrak{H}_j} \min_{\mathcal{C} \subseteq \mathcal{H}} f_{Po}(\mathcal{H}, \mathcal{C}, c_N, c_A, j)$  and leads to a more precise estimate of the number of recipient choices to disprove hypotheses. While this refined estimate leads to a lower number of recipient choices for disproofs, it does not change the optimistic case strategy of choosing recipients, as we will see.

Let  $\mathcal{C}_i = \{r_1, \dots, r_i\}$  be the set of chosen recipients when entering the  $i$ -th level

---

of recursion in ExactHS. We assume w.l.o.g. in the remaining sections that  $\mathcal{C}_i$  is an ordered set, where the  $l$ -th recipient  $r_l \in \mathcal{C}_i$  represents the recipient added, when entering the  $l$ -th level of recursion, for  $l = 1, \dots, i$ . The set of observations, when entering the  $i$ -th level of recursion the first time<sup>1</sup> is  $\mathcal{OS}_i$ , given that  $\mathcal{C}_i$  was chosen.

Consider the case that the extension of  $\mathcal{C}_i$  at the  $(i + 1)$ -th level of recursion is  $\mathcal{C}_{i+1} = \mathcal{C}_i \cup \{a_{i+1}\}$ , where  $a_{i+1} \in \mathcal{O}_i$  is an Alice's friend chosen at the  $i$ -th level of recursion. Then, due to the optimistic case strategy,  $a_{i+1}$  must be the recipient who is most frequently contacted by Alice in  $\mathcal{O}_i$ .

Now, consider the other case that the extension at the  $(i + 1)$ -th level of recursion is  $\mathcal{C}_{i+1} = \mathcal{C}_i \cup \{n_{i+1}\}$ , where  $n_{i+1} \in \mathcal{O}_i$  is a non-friend in the  $i$ -th level of recursion. Then, due to the optimistic case strategy and Line 9 in Algorithm 1,  $a_{i+1} \in \mathcal{O}_i$  must be removed from all observations in  $\mathcal{OS}_i$  in level  $i$ , before a non-friend  $n_{i+1} \in \mathcal{O}_i$  can be added in the  $(i + 1)$ -th level of recursion. Accounting this will lead to a refined formulation of (4.21) as represented in (4.24) and a more precise estimate of the number of recipient choices to disprove hypotheses.

Each chosen set  $\mathcal{C}_i$  in ExactHS is therefore *associated* with a particular set of Alice's friends  $\mathcal{A}_i = \{a_1, \dots, a_i\}$  that we denote by  $\mathcal{C}_i \sim \mathcal{A}_i$ . Each recipient  $a_l \in \mathcal{A}_i$  represents Alice's most frequent recipient contact in  $\mathcal{OS}_{l-1}$  and there is an observation  $\mathcal{O}_{l-1}$  selected by ExactHS in the  $(l - 1)$ -th level of recursion, such that  $a_l \in \mathcal{O}_{l-1}$  for  $l = 1, \dots, i$ . If a recipient  $r_l \in \mathcal{C}_i$  is an Alice's friend then it equals  $a_l \in \mathcal{A}_i$ . If it is a non-friend, then  $a_l, r_l \in \mathcal{O}_{l-1}$  and  $a_l$  has been removed from all observations in  $\mathcal{OS}_{l-1}$  prior to adding  $r_l$  in the  $l$ -th recursion level.

In the optimistic case, every set  $\mathcal{C}_i$  computed by ExactHS is associated to  $\mathcal{A}_i$  consisting of the  $i$  recipients who are most frequently contacted by Alice. Therefore, the hypotheses  $\mathcal{H} \supseteq \mathcal{C}_i$  that are disproved by ExactHS are restricted by the condition  $\mathcal{H} \cap \mathcal{A}_i \setminus \mathcal{C}_i = \emptyset$ . That means that the non-chosen recipients in  $\mathcal{H}$  must not be any of the remaining most likely recipients  $\mathcal{A}_i \setminus \mathcal{C}_i$  and that the maximal value of  $f_{Po}(\mathcal{H}, \mathcal{C}, c_N, c_A, j)$  is accordingly restricted.

**Example 4.2.2 (Optimistic Case).** *This example illustrates the optimistic case strategy applied to ExactHS. We demonstrate at the same time that this optimistic case strategy*

---

<sup>1</sup>After entering the  $i$ -th level of recursion, recipients will be removed from observations in  $\mathcal{OS}_i$ . The first time refers to the state of  $\mathcal{OS}_i$  prior to any modifications.

#### 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

is approached by the enhanced version of *ExactHS* as represented by Algorithm 2.<sup>1</sup> Applying Algorithm 2 to the observations in this example will lead to the same recipient choices as under the optimistic case strategy.

Let  ${}_A\mathcal{H} = \{1, 2\}$ ,  $m = 2$ ,  $b = 2$ . Let the attacker collect following 5 observations  $\mathcal{OS} = \mathcal{OS}_0$ :

$$\{1, 3\}, \{2, 4\}, \{1, 4\}, \{2, 3\}, \{1, 5\}$$

that are sufficient to uniquely identify  ${}_A\mathcal{H}$  by *ExactHS*. Hypotheses are referred to by  $\mathcal{H}$ . The frequency of a recipient is denoted by the pair (recipient [frequency]). They are in  $\mathcal{OS}$ :

$$1[3], 2[2], 3[2], 4[2], 5[1] .$$

Recursion level 0 refers to the initial invocation of Alg. 2 by *ExactHS*( $\mathcal{OS}_0, m = 2, \mathcal{C}_0 = \{\}$ ). In each level  $i$ , an observation  $\mathcal{O}_i \in \mathcal{OS}_i$  is fixed and *ExactHS* chooses one of the 2 recipients in  $\mathcal{O}_i$ . We number the choices by  $a, b$ , the first choice leads to  $\mathcal{O}_{i.a}$ .

$i.j$	$\mathcal{O}_{i.j}$	$\mathcal{OS}_{i.j}$	$\mathcal{C}_{i+1}$	$\mathcal{A}_{i+1}$	$\mathcal{OS}_{i+1}$	$\max_{\mathcal{H}} Po(\mathcal{H}, \mathcal{C}_i)$
0.a	$\{1, 3\}$	$\{1, 3\}, \{2, 4\}, \{1, 4\}, \{2, 3\}, \{1, 5\}$	$\{1\}$	$\{1\}$	$\{2, 4\}, \{2, 3\}$	$5(\geq  \mathcal{OS} )$
1.a	$\{2, 4\}$	$\{2, 4\}, \{2, 3\}$	$\{1, 2\}$	$\{1, 2\}$	$\{\}$	$5(\geq  \mathcal{OS} )$
1.b	$\{4\}$	$\{4\}, \{3\}$	$\{1, -\}$	$\{1, 2\}$	$-$	$4(<  \mathcal{OS} )$
0.b	$\{3\}$	$\{3\}, \{2, 4\}, \{4\}, \{2, 3\}, \{5\}$	$\{-\}$	$\{1\}$	$-$	$4(<  \mathcal{OS} )$

Figure 4.5: Choosing/removing recipients at  $i$ -th level of recursion in enhanced *ExactHS* as implemented in Algorithm 2.

The enhanced version of *ExactHS* represented by Algorithm 2 approaches the optimistic case strategy by choosing and removing in every level of recursion, the recipient first that is most frequent in the observations at that level of recursion in Line 7.

We show the state of the chosen recipients and the potential in each level of recursion in Algorithm 2 next.

0.a:  $\mathcal{O}_0 = \{1, 3\}$  is fixed by *ExactHS*, as it contains the most frequent recipient 1 in  $\mathcal{OS}_0$ .  $\max_{\mathcal{H}} Po(\mathcal{H}, \{\}) \geq |\mathcal{OS}|$ , thus recipient 1 is chosen and will be added at recursion level 1, i.e  $\mathcal{C}_1 = \{\} \cup \{1\}$ . See Fig. 4.5.

<sup>1</sup>A proof of this is provided in Section 4.2.3.3.

---

1.a:  $\mathcal{O}_1 = \{2, 4\}$  is fixed by ExactHS, as it contains the most frequent recipient 2 in  $\mathcal{OS}_1$ .  $\max_{\mathcal{H}} Po(\mathcal{H}, \{1\}) \geq |\mathcal{OS}|$ , thus recipient 2 is chosen and will be added at recursion level 2, i.e.,  $\mathcal{C}_2 = \{1\} \cup \{2\}$ . Level 2 ends with Hitting set  $\mathcal{C}_2$ . See Figure 4.5.

1.b: Recipient 2 is removed from all observations in  $\mathcal{OS}_1$ , therefore  $\max_{\mathcal{H}} Po(\mathcal{H}, \{1\}) < |\mathcal{OS}|$  in level 1;  $\mathcal{A}_2 = \{1, 2\}$ . All hypotheses  $\{\mathcal{H} \mid \{1\} \in \mathcal{H} \wedge \{2\} \notin \mathcal{H}\}$  are disproved. No recipient will be added (-).

0.b: Recipient 1 is removed from all observations in  $\mathcal{OS}_0$ , therefore  $\max_{\mathcal{H}} Po(\mathcal{H}, \{1\}) < |\mathcal{OS}|$  in level 0 and  $\mathcal{A}_1 = \{1\}$ . All  $\mathcal{H}$ , where  $\{1\} \notin \mathcal{H}$  are disproved. No recipient will be added (-).

In this example, Alice's friends are chosen and removed first in every recursion level in Algorithm 2, as in each level, one of Alice's friends becomes the most frequent recipient. Thus Algorithm 2 chooses those recipients who would be chosen if the optimistic case strategy would be assumed and Section 4.2.3.3 shows that this case is approached in realistic Mix configurations.

We can see in level 0 that all hypotheses that will be proved and disproved in level 1 fulfil  $\mathcal{H} \cap \mathcal{A}_1 \setminus \mathcal{C}_1 = \emptyset$ , where  $\mathcal{A}_1 = \{1\}$  and  $\mathcal{A}_i$  consists of the  $i$  most frequent Alice's friends in  $\mathcal{OS}$ . That is the non-chosen recipients of any potential hypotheses do not contain  $\mathcal{A}_1 \setminus \mathcal{C}_1$ . In level 0.a, the succeeding chosen recipient set will be  $\mathcal{C}_1 = \{1\}$ , thus  $\mathcal{H} \cap \mathcal{A}_1 \setminus \mathcal{C}_1 = \emptyset$ . In level 0.b, the succeeding chosen recipient set will be  $\mathcal{C}_1 = \{3\}$  and 1 has been removed from all observations in  $\mathcal{OS}_0$ , thus  $\mathcal{H} \cap \mathcal{A}_1 \setminus \mathcal{C}_1 = \emptyset$ . We see in level 1 similarly  $\mathcal{H} \cap \mathcal{A}_2 \setminus \mathcal{C}_2 = \emptyset$ .

Let  $c_N, c_A, j, \mathcal{A}_{c_N+c_A}$  be given, where  $\mathcal{A}_{c_N+c_A}$  consists of the  $c = c_N + c_A$  Alice's friends who are most frequently contacted by Alice in  $\mathcal{OS}$ . In the optimistic case, every hypothesis  $\mathcal{H}$  and chosen recipient set  $\mathcal{C} \subseteq \mathcal{H}$  must fulfil  $\mathcal{C} \sim \mathcal{A}_{c_N+c_A}$ ,  $|\mathcal{C}_A| = c_A$  and  $\mathcal{H} \cap \mathcal{A}_{c_N+c_A} \setminus \mathcal{C} = \emptyset$ .<sup>1</sup> Therefore, provided  $\mathcal{H}$  and  $\mathcal{A}_{c_N+c_A}$ , the set of chosen Alice's friends is determined by  $\mathcal{H} \cap \mathcal{A}_{c_N+c_A} = \mathcal{C}_A$ . Applying these constraints to

---

<sup>1</sup>This means that the  $c_N + c_A$  friends who are most frequently contacted by Alice must either be chosen, or already removed.

#### 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

$\max_{\mathcal{H} \in \mathfrak{H}_j} \min_{\mathcal{C} \subseteq \mathcal{H}} f_{Po}(\mathcal{H}, \mathcal{C}, c_N, c_A, j)$  leads to the following formulation:

$$\begin{aligned} & \max_{\mathcal{H} \in \mathfrak{H}_j} \min_{\mathcal{C} \subseteq \mathcal{H}: \mathcal{C}_A = \mathcal{H} \cap \mathcal{A}_{c_N+c_A}} \{f_{Po}(\mathcal{H}, \mathcal{C}, c_N, c_A, j)\} \\ & \leq \max_{\substack{\mathcal{C}_A = \{a_1, \dots, a_{c_A}\} \subseteq \mathcal{A}_{c_N+c_A}, \\ \mathcal{H}_A \setminus \mathcal{C} = \{a_{c_A+1}, \dots, a_{m-j}\} \subseteq \mathcal{H} \setminus \mathcal{A}_{c_N+c_A}}} \left\{ \frac{\sum_{k=c_A+1}^{m-j} P_A(a_k)}{1 - \sum_{l=1}^{c_A} P_A(a_l)} \right\} \\ & \quad + \max_{r_1, \dots, r_{m-c_N-c_A} \in R} \left\{ \sum_{l=1}^{m-c_N-c_A} P_N(r_l) \right\}. \end{aligned} \quad (4.22)$$

The expression within the first brackets in (4.22) is of the structure

$$\frac{z - u}{1 - y}, \quad (4.23)$$

for invariant  $u = \sum_{a_r \in \mathcal{A}_{c_N+c_A}} P_A(a_r)$  and variable  $y = \sum_{a_l \in \mathcal{C}_A} P_A(a_l)$  and  $z = \sum_{i=1}^{m-j+c_N} P_A(a_i)$  (for arbitrary  $a_1, \dots, a_{m-j+c_N} \in {}_A\mathcal{H}$ ).<sup>1</sup> The numerator in that expression is the cumulative probability  $\sum_{a \in \mathcal{H}_A \setminus \mathcal{C}_A: |\mathcal{H}_A \setminus \mathcal{A}_{c_N+c_A}|=m-j-c_A} P_A(a)$  which is simplified by the less constrained expression  $z - u$ . The denominator in that expression is  $1 - y$ . Since (4.23) is monotonically increasing with respect to increasing value of  $z$  and  $y$ , it is maximal, if the value of  $z$  and  $y$  are maximal.

Let us relabel the indices in (4.22), such that  $P_A(a_k)$  and  $P_N(r_l)$  are probabilities of the  $k$ -th and  $l$ -th most likely recipients with respect to  $P_A(\cdot)$  and  $P_N(\cdot)$ . We obtain the following formula that is equivalent to the right hand side of (4.22):

$$f_{Po}^{min}(c_N, c_A, j) = \frac{\sum_{k=c_N+c_A+1}^{m-j+c_N} P_A(a_k)}{1 - \sum_{l=1}^{c_A} P_A(a_l)} + \sum_{l=1}^{m-c_N-c_A} P_N(r_l). \quad (4.24)$$

The expression left of (+) in (4.24) equals (4.23) for maximal value of  $z$  and  $y$ . The expression right of (+) in (4.24) equals the expression right of (+) in (4.22). The condition that the friends in  $\mathcal{A}_{c_N+c_A}$  must either be chosen, or removed, such that  $\mathcal{H}_A \setminus \mathcal{C} \subseteq {}_A\mathcal{H} \setminus \mathcal{A}_{c_N+c_A}$ , is fulfilled by the numerator in the expression left of (+) in (4.24). Therefore we can conclude that (4.24) and the right hand side of (4.22) are equal.

<sup>1</sup>Expression (4.23) is less constrained than the expression in the brackets left of (+) in (4.22). We will see that the conditions to maximise (4.23) also maximises the latter expression in (4.22).

---

The equations (4.24) and (4.21) shows that  $f_{Po}^{mmin}(c_N, c_A, j) \leq \widehat{f_{Po}^{mmin}}(c_N, c_A, j)$ . Accounting the effect of the optimistic case strategy on hypotheses in  $f_{Po}^{mmin}(c_N, c_A, j)$  thus leads to a more precise and lower estimate of the number of recipient choices to disprove hypotheses than without accounting that effect. Therefore we only consider  $f_{Po}^{mmin}(c_N, c_A, j)$  in the remaining analyses.

**Deriving  $f_{Po}^{mmin}(c_N, c_A)$**  Given that the values of  $c_N, c_A$  are fixed, the numerator of the left expression in  $f_{Po}^{mmin}(c_N, c_A, j)$  is monotonically decreasing with respect to  $j$ . Since  $c_N \leq j \leq m - c_A$ , we conclude that  $f_{Po}^{mmin}(c_N, c_A, j)$  is maximal, if  $j = c_N$ . We refer to this case by  $f_{Po}^{mmin}(c_N, c_A)$ , where

$$f_{Po}^{mmin}(c_N, c_A) = \frac{\sum_{k=c_N+c_A+1}^m P_A(a_k)}{1 - \sum_{l=1}^{c_A} P_A(a_l)} + \sum_{l=1}^{m-c_N-c_A} P_N(r_l) . \quad (4.25)$$

To obtain  $c_N^{mmin}(c_A)$ , we need to solve  $1 = f_{Po}^{mmin}(c_N, c_A)$ , which is by reformulation equivalent to solving the more convenient equation  $1 = f_{Po}^{mmin}(c_N, c_A)$ , where  $f_{Po}^{mmin}(c_N, c_A)$  equals

$$1 - \sum_{k=c_A+1}^{c_N+c_A} P_A(a_k) + (1 - \sum_{l=1}^{c_A} P_A(a_l)) \sum_{s=1}^{m-c_N-c_A} P_N(r_s) . \quad (4.26)$$

Note that if  $P_N(r) < \frac{1}{m^2}$  (as considered in Section 4.1.4.1), then (4.26) is less than 1, for  $c_N \geq 1$ . This is because the expression right of (+) is at most  $\frac{1 - \sum_{l=1}^{c_A} P_A(a_l)}{m}$ , which is less than  $\sum_{k=c_A+1}^{c_N+c_A} P_A(a_k)$ , for  $c_N \geq 1$ . We therefore deduce that  $c_N^{mmin}(c_A) \leq 1$  for  $c_A = 0, \dots, m-1$ , which is regardless of Alice's traffic distribution. This implies due to (4.29) in Section 4.2.4 a polynomially bounded mean time-complexity of ExactHS, which confirms the corresponding analysis in Section 4.1.

#### 4.2.3.2 Comparing Uniform and Non-Uniform Distribution

Provided the optimistic case assumption, we prove that if Alice's traffic is non-uniformly distributed, then  $f_{Po}^{mmin}(c_N, c_A)$  in (4.25) is not higher than in the case of Alice's uniform traffic distribution. All chosen sets constructed by ExactHS are thus assumed to fulfil  $\mathcal{C} \sim \mathcal{A}_{c_N+c_A}$ , for  $|\mathcal{C}| = c_N + c_A$ .

#### 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

We additionally analyse  $f_{Po}^{min}(c_N, c_A)$  for the Zipf( $m, \alpha$ ) distribution of Alice's friends. Our result proves that the mean time-complexity of ExactHS decreases with respect to the weight  $\alpha$  and approaches linear time for moderately large value of  $\alpha$ .

The values of  $f_{Po}^{min}$  for the non-uniform and uniform communication only differs by the expression left of (+) in (4.25). Let us call that expression  $f_1^n$  in the non-uniform case and  $f_1^u$  in the uniform case. We prove next that  $f_1^n(c_N, c_A) \leq f_1^u(c_N, c_A)$  which is of the structure

$$\frac{1 - (\frac{c_N}{m} + \delta) - (\frac{c_A}{m} + \epsilon)}{1 - (\frac{c_A}{m} + \epsilon)} \leq \frac{\frac{m-c_N}{m} - \frac{c_A}{m}}{1 - \frac{c_A}{m}} \quad (4.27)$$

$$\Leftrightarrow \delta \geq -\epsilon \frac{c_N}{m - c_A} . \quad (4.28)$$

If inequality (4.28) is fulfilled, then the number of recipient choices to disprove  $c_A$ -average-disprovable sets does not exceed the number of recipient choices, if Alice communicates uniformly distributed. Assume that  $\delta, c_N$  and  $\epsilon, c_A$  in (4.28) are substituted by some other values  $\delta', c'_N$  and  $\epsilon', c'_A$  and that the resulting inequality is valid, then we say that  $\delta', c'_N$  and  $\epsilon', c'_A$  applies to (4.28).

*Proof.* Let  $\mathcal{C} \sim \mathcal{A}_{c_N+c_A}$  be the chosen recipients who disprove a hypothesis  $\mathcal{H}$ .  $\mathcal{C}_A \subseteq \mathcal{A}_{c_N+c_A}$ , where  $|\mathcal{C}_A| = c_A$  and  $P_A(\mathcal{C}_A) = \frac{c_A}{m} + \epsilon$  corresponds to the second bracketed expression in the numerator left of ( $\leq$ ) in (4.27). The first bracketed expression in the numerator left of ( $\leq$ ) in (4.27) corresponds to  $P_A(\mathcal{A}_{c_N+c_A} \setminus \mathcal{C}_A) = \frac{c_N}{m} + \delta$ .

Imagine that ExactHS computes  $\mathcal{C} \sim \mathcal{A}_{c_N+c_A}$ , but then keeps choosing additional recipients, such that it arrives at  $\mathcal{C}' \sim \mathcal{A}'_{c'_N+c_A}$ , where  $\mathcal{C}'_A = \mathcal{C}_A$  and  $\mathcal{A}'_{c'_N+c_A} \setminus \mathcal{C}'_A = {}_A\mathcal{H} \setminus \mathcal{C}_A$ . Since  $P_A({}_A\mathcal{H}) = 1$ , we deduce that  $P_A(\mathcal{A}'_{c'_N+c_A} \setminus \mathcal{C}'_A) = \frac{m-c_A}{m} - \epsilon = \frac{c'_N}{m} + \delta'$  and that inequality (4.28) applies to  $c'_N, \delta'$  and  $c_A, \epsilon$ . The overall average probability of every recipient in  $\mathcal{A}'_{c'_N+c_A} \setminus \mathcal{C}'_A$  is  $\bar{P}_A = \frac{P_A(\mathcal{A}'_{c'_N+c_A} \setminus \mathcal{C}'_A)}{m - c_A} = \frac{1}{m} - \frac{\epsilon}{m - c_A}$ . Since  $\mathcal{A}_{c_N+c_A}$  consists of the most frequent Alice's friends and  $\mathcal{A}_{c_N+c_A} \setminus \mathcal{C}_A \subseteq {}_A\mathcal{H} \setminus \mathcal{C}_A$ , we can deduce that  $P_A(\mathcal{A}_{c_N+c_A} \setminus \mathcal{C}_A) = \frac{c_N}{m} + \delta \geq c_N \bar{P}_A = \frac{c_N}{m} - \frac{c_N \epsilon}{m - c_A}$ . Therefore inequality (4.28) also applies to  $c_N, \delta$  and  $c_A, \epsilon$ .  $\square$

**Zipf Distribution** Given that Alice follows a Zipf( $m, \alpha$ ) distribution when choosing her friend in each round, we prove that  $c_N^{min}(c_A)$  decreases with respect to  $\alpha \geq 0$ .



---

Assume w.l.o.g. that  $c_A$  is fixed, then  $c_N = c_N^{min}(c_A)$  decrease with respect to  $\alpha$ , if the condition  $f_{Po}^{min}(c_N, c_A) < 1$  can be fulfilled by lower values of  $c_N = c_N^{min}(c_A)$  for increasing  $\alpha$ .

*Proof.* Note that  $f_{Po}^{min}(c_N, c_A) < 1$  if and only if the expressions right of the leading number 1 in (4.26) is negative. Let w.l.o.g.  $P_z^{m,\alpha}(y)$  be the probability of Alice's  $y$ -th most frequently contacted friend  $a_y$  for  $y \in \{1, \dots, m\}$ . By applying the Zipf( $m, \alpha$ ) distribution of Alice's friends to (4.26), that is setting  $P_A(a_y) = P_z^{m,\alpha}(y)$ , we obtain the following expression that is equivalent to the expression right of the leading number 1 in (4.26):

$$-\underbrace{(F_z^{m,\alpha}(c_N + c_A) - F_z^{m,\alpha}(c_A))}_{G_1(\alpha)} + \underbrace{(1 - F_z^{m,\alpha}(c_A))}_{G_2(\alpha)} \sum_{s=1}^{m-c_N-c_A} P_N(r_s) .$$

It is known that  $F_z^{m,\alpha}(y)$  is monotonically increasing with respect to  $\alpha \geq 0$ , where  $\lim_{\alpha \rightarrow \infty} F_z^{m,\alpha}(y) = 1$  for any  $y \in \{1, \dots, m\}$ . Therefore  $G_2(\alpha)$  is monotonically decreasing with respect to  $\alpha \geq 0$ . Note that  $G_2(\alpha) \geq G_1(\alpha) \geq 0$  and we prove by

$$G_2(\alpha) - G_1(\alpha) = 1 - F_z^{m,\alpha}(c_N + c_A)$$

that the difference  $G_2(\alpha) - G_1(\alpha)$  decreases monotonically with respect to  $\alpha$ . Therefore the value of  $c_N = c_N^{min}(c_A)$ , such that  $-G_1(\alpha) + G_2(\alpha) \sum_{s=1}^{m-c_N-c_A} P_N(r_s) < 0$ , decreases with respect to  $\alpha$ .  $\square$

Note that increasing  $\alpha$  (i.e., the non-uniformness in the Zipf distribution) decreases the number of recipient choices for disproofs and the exponent of ExactHS's time-complexity, due to (4.29) in Section 4.2.4. That is  $c_N^{min}(\cdot)$  approaches 0, so that the mean time-complexity of ExactHS approaches a polynomial function.

#### 4.2.3.3 Approaching Optimistic Case Strategy

This section analyses the mean time-complexity of ExactHS for the practical case, as opposed to the optimistic case as yet considered in Section 4.2.3. In the *practical case*, ExactHS might construct chosen sets  $\mathcal{C} \sim \mathcal{A}_{c_N+c_A}$ , where  $\mathcal{A}_{c_N+c_A}$  does not consists of the  $c_N + c_A$  recipients who are most frequently contacted by Alice. We proof that the

#### 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

mean time-complexity of the practical case approaches that of the optimistic case, for realistic Mix configurations.

For the formal correctness of the proofs, it is required that the traffic volume of the cover-traffic to each single recipient is less than  $\frac{1}{m}$  of all observations, so that  $P_N(r) < \frac{1}{m}$  for all  $r \in R$ . This condition is therefore assumed in the remaining of this section. Note that Alice's traffic volume averaged over her  $m = |{}_A\mathcal{H}|$  friends covers  $\frac{1}{m} = \frac{1}{m} \sum_{i=1}^m P_A(a_i)$  of all observations, even if  $P_A(a) < P_N(r)$  for many  $r \in R$ ,  $a \in {}_A\mathcal{H}$ .

We resolve the apparent caveat that ExactHS often prefers recipients  $a \in {}_A\mathcal{H}$  that are most frequently contacted by Alice, even if  $P_A(a) < P_N(r)$  for many  $r \in R$ . This results in the mean number of recipient choices in the practical case approaching that in the optimistic case. Most importantly, we prove that if Alice traffic is non-uniformly distributed, then the number of recipient choices for disproofs by ExactHS is not greater than in the case of uniformly distributed traffic.

Note that the mean time-complexity of ExactHS is identical in the practical and optimistic case, if Alice traffic is uniformly distributed. That is if  $P_A(a) = \frac{1}{m}$  for all  $a \in {}_A\mathcal{H}$ .

**Claim 10.** *Let  $\mathcal{C} \sim \mathcal{A}_{c_N+c_A}$  be any set constructed and disproved by ExactHS. Let  $P_A(\mathcal{C}_A) = \frac{c_A}{m} + \epsilon$  and  $P_A(\mathcal{A}_{c_N+c_A} \setminus \mathcal{C}_A) = \frac{c_N}{m} + \delta$ , where  $-1 < \epsilon < 1$ ,  $-1 < \delta < 1$ . If the probability  $P_N(r)$  that any recipient  $r$  contacted by senders, other than Alice is below  $\frac{1}{m}$ , then inequality (4.28) applies to  $c_N, \delta$  and  $c_A, \epsilon$ .*

Since non-friends are less likely observed than  $\frac{1}{m}$ , we deduce that the first recipients  $a \in \mathcal{A}_{c_N+c_A}$  are the most likely Alice's friends of probability  $P_A(a) \geq \frac{1}{m}$ . If for all recipients  $a \in \mathcal{A}_{c_N+c_A}$ ,  $P_A(a) > \max_{r \in R} P_N(r)$ , then inequality (4.28) obviously applies as shown in Section 4.2.3.2. Therefore the remaining analyses of this section refers to the case that  $\mathcal{A}_{c_N+c_A} = \{a_1, \dots, a_i, a_{i+1}, \dots, a_{c_N+c_A}\}$ , contains all Alice's friends  $P_A(a_1) \geq \dots \geq P_A(a_i) \geq \frac{1}{m}$  and some other Alice's friends  $P_A(a_l) \leq \max_{r \in R} P_N(r) < \frac{1}{m}$  for  $l = i+1, \dots, c_N + c_A$ .

**Impact of Non Uniform Distribution** Let  $\mathcal{C}$  be the recipients chosen by ExactHS that disproves a hypothesis  $\mathcal{H}$ , where  $\mathcal{C} \sim \mathcal{A}_{c_N+c_A}$ . We define  $\mathcal{C}_i \sim \mathcal{A}_{c_N^i+c_A^i}$ , where

$\mathcal{A}_{c_N+c_A}^i$  consists of the first  $c_N^i + c_A^i$  Alice's friends<sup>1</sup> who were added to  $\mathcal{A}_{c_N+c_A}$ . The corresponding chosen Alice's friend are  $\mathcal{C}_{Ai}$ , where  $|\mathcal{C}_{Ai}| = c_A^i \leq c_A$  and the chosen non-friends are  $\mathcal{C}_{Ni}$ , where  $|\mathcal{C}_{Ni}| = c_N^i \leq c_N$ .

Let  $P_A^i(a) = \frac{|\mathcal{OS}_{Ai}[a]|}{|\mathcal{OS}_i|}$  denote the probability that Alice contacts  $a$  in  $\mathcal{OS}_i$ , where  $\mathcal{OS}_i$  is the set of observations when ExactHS enters the  $i$ -th recursion level the first time after choosing the  $i$  recipients in  $\mathcal{C}_i$ . Since the communication of other senders are statistically independent of Alice's communication, we follow  $\frac{|\mathcal{OS}_{Ai}[a]|}{|\mathcal{OS}_i|} = \frac{|\mathcal{OS}_A[a] \setminus \mathcal{OS}[\mathcal{C}_{Ai}]|}{|\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}_{Ai}]|} = \frac{|\mathcal{OS}_A[a]|}{|\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}_{Ai}]|}$ . This means that  $P_A^i(a)$  is only affected by the chosen Alice's friends  $\mathcal{C}_{Ai}$  and that  $P_A^i(a) \geq P_A(a) = \frac{|\mathcal{OS}_A[a]|}{|\mathcal{OS}|}$ . Thus the more Alice's friends  $\mathcal{C}_{Ai}$  are chosen, the higher is the probability  $P_A^i(a)$  of the remaining recipients  $a \in \mathcal{C}_{Ai} \setminus \mathcal{A}_{c_N+c_A}^i$ . We define  $\mathcal{OS}_0 = \mathcal{OS}$  and  $P_A^0(a) = P_A(a)$ , therefore these notations are used interchangeably. We further define  $P_A^i(\mathcal{C}_{Ai}) = \frac{c_A^i}{m} + \epsilon_i$  and  $P_A^i(\mathcal{A}_{c_N+c_A}^i \setminus \mathcal{C}_{Ai}) = \frac{c_N^i}{m} + \delta_i$ .

Let  $P_N^i(r) = \frac{|\mathcal{OS}_{Ni}[r]|}{|\mathcal{OS}_i|}$  denote the probability that senders other than Alice contact  $r$  in  $\mathcal{OS}_i$ . We assume for large number of receivers  $u \gg m$  that  $P_N^i(r) \approx P_N(r)$ . To give an intuitive justification, we consider the case that every single sender other than Alice contacts a recipient with the uniform probability  $\frac{1}{u}$ . The probability that a recipient is contacted by any sender other than Alice in an observation is  $P_N = 1 - (1 - \frac{1}{u})^{b-1}$ . If  $i$  recipients are chosen, then  $P_N^i = 1 - (1 - \frac{1}{u+1-i})^{b-1}$ . For example, let us assume an extreme case with parameters  $m = 10$ ,  $b = 400$ ,  $u = 4000$ , then  $P_N = 0.0949$  and choosing  $i = 10$  recipients marginally changes the probability to  $P_N^{10} = 0.0951$ .

We define  $P^i(r) = \frac{|\mathcal{OS}_i(r)|}{|\mathcal{OS}_i|}$  as the probability that recipient  $r$  is observed in observations in  $\mathcal{OS}_i$ . Notably  $P^i(n) = P_N(n)$  and  $P^i(a) = \frac{|\mathcal{OS}_{Ai}(a) \cup \mathcal{OS}_{Ni}(a)|}{|\mathcal{OS}_i|} \leq P_A^i(a) + P_N(a)$ .

Let  $\mathcal{C} \sim \mathcal{A}_{c_N+c_A}$  be chosen by ExactHS to disprove some hypotheses in practice. For  $i = 1, \dots, c_N + c_A$ ,  $a_i \in \mathcal{A}_{c_N+c_A}$  result from ExactHS preferring to choose and remove the most likely recipient  $P^{i-1}(r)$  in  $\mathcal{OS}_{i-1}$  first. In general,  $a_i$  will be in many cases the recipient with the maximal probability  $P^{i-1}(a_i)$  in  $\mathcal{OS}_{i-1}$  and close to the  $i$ -th most likely Alice's friend in  $\mathcal{OS}$ , so that ExactHS approaches the optimistic case. We assume as the general case that the communication of other senders to Alice's friends do not significantly skew the relative order of Alice's communication, so that for any  $a_k, a_l$ , if  $P_A^{i-1}(a_k) \geq P_A^{i-1}(a_l)$ , then  $P^{i-1}(a_k) \geq P^{i-1}(a_l)$  in most of the cases. A justification is provided in the next Paragraph "Perturbations" in this section. For simplicity we only use  $P_A^{i-1}(a)$  instead of  $P^{i-1}(a)$  in the remaining analyses to also

---

<sup>1</sup> $i = c_N^i + c_A^i$ .

#### 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

cover the case that senders other than Alice do not contact Alice's friends, which is  $P_N(a) = 0$  for all Alice's friends  $a$ .

We define  $\bar{P}_A^i = \frac{P_A(\mathcal{H} \setminus \mathcal{C}_{Ai})}{m - c_A^i} = \frac{1}{m} - \frac{\epsilon_i}{m - c_A^i}$  as the *overall average probability* of an Alice's friend who is not in  $\mathcal{C}_{Ai}$ . Since  $\mathcal{C}_0 = \{\}$ ,  $c_A^0 = 0$ ,  $\epsilon_0 = 0$ , we obtain  $\bar{P}_A^0 = \frac{1}{m}$ . Most importantly, if there is a recipient  $a$  in  $\mathcal{OS}_i$ , where  $P_A(a) \geq \bar{P}_A^i$ , then  $P_A^i(a) = \frac{P_A(a)}{1 - (\frac{c_A^i}{m} + \epsilon_i)} \geq \frac{\bar{P}_A^i}{1 - (\frac{c_A^i}{m} + \epsilon_i)} = \frac{1}{m - c_A^i}$ . Therefore, given that  $\mathcal{C}_i$  is chosen, ExactHS would in general still choose the most frequent Alice's friend  $a$  first, although  $P_A(a) < \frac{1}{m}$ , as long as  $P_A(a) \geq \bar{P}_A^i$ .<sup>1</sup> This is because the probability  $P_N(n)$  of every non-friend  $n$  remains lower than  $\frac{1}{m}$  and thus appears less frequently than  $a$  in  $\mathcal{OS}_i$ .

*Proof.* (Claim 10) Imagine that ExactHS chooses to successively extend  $\mathcal{C}_i$  by  $j$  Alice's friend  $a'_1, \dots, a'_j$ , where  $P_A(a'_l) > \bar{P}_A^i$  for  $l = 1, \dots, j$ . We assume for simplicity and without restriction of generality that this result in  $\mathcal{C}_{i+j} = \mathcal{C}_i \cup \{a'_1, \dots, a'_j\}$  and  $\mathcal{C}_{Ai+j} = \mathcal{C}_{Ai} \cup \{a'_1, \dots, a'_j\}$ . We prove next that choosing Alice's friends  $a'$ , where  $P_A(a') \geq \bar{P}_A^i$  leads to a lower overall average probability  $\bar{P}_A^{i+j} \leq \bar{P}_A^i$ .

$$\begin{aligned} P_A(\mathcal{H} \setminus \mathcal{C}_{Ai+j}) &= P_A(\mathcal{H} \setminus \mathcal{C}_{Ai}) - P_A(a'_1) - \dots - P_A(a'_j) \\ &\leq \frac{m - c_A^i}{m} - \epsilon_i - j\bar{P}_A^i \\ &= \frac{m - c_A^i - j}{m} - \frac{\epsilon_i(m - c_A^i - j)}{m - c_A^i} \end{aligned}$$

This therefore proves  $\bar{P}_A^{i+j} = \frac{P_A(\mathcal{H} \setminus \mathcal{C}_{Ai+j})}{m - c_A^{i+j}} \leq \bar{P}_A^i$ . Accordingly, if we imagine that ExactHS chooses to extend  $\mathcal{C}_i$  by Alice's friend  $a$ , where  $P_A(a) \leq \bar{P}_A^i$  instead, the overall probability would increase, so that  $\bar{P}_A^{i+j} \geq \bar{P}_A^i$ .

Without restriction of generality, for any  $\mathcal{C} \sim \mathcal{A}_{c_N + c_A}$  there is a largest  $\mathcal{C}_i \sim \mathcal{A}_{c_N^i + c_A^i}$ , such that  $\mathcal{A}_{c_N^i + c_A^i}$  consists of only the most frequently contacted Alice's friends, where  $P_A(a_l) \geq \bar{P}_A^i$  for  $l = 1, \dots, i$  and  $a_l \in \mathcal{A}_{c_N^i + c_A^i}$ . And there must be no  $a \in \mathcal{H} \setminus \mathcal{A}_{c_N^i + c_A^i}$ , where  $P_A(a) \geq \bar{P}_A^i$ , unless  $\mathcal{A}_{c_N^i + c_A^i} = \mathcal{A}_{c_N + c_A}$  and thus  $\mathcal{C}_i = \mathcal{C}$ . The latter case, is the optimistic case for ExactHS, as considered in the Section 4.2.3.2, hence (4.28) applies to  $c_N, \delta, c_A, \epsilon$ . The remaining analysis in this section therefore addresses only the first case, which is  $\mathcal{A}_{c_N^i + c_A^i} \subset \mathcal{A}_{c_N + c_A}$ .

<sup>1</sup>To be more precise, ExactHS still chooses the most frequent Alice's friend  $a$ , even if  $P_A(a) \leq \bar{P}_A^i$ , as long as  $P_A^i(a) > \max_n P_N(n)$ . However it is sufficient to consider  $P_A(a) \geq \bar{P}_A^i$  to prove Claim 10.

---

To prove that inequality (4.28) applies to  $c_N, \delta, c_A, \epsilon$ , it is necessary and sufficient to show that  $P_A(\mathcal{A}_{c_N+c_A} \setminus \mathcal{C}_A) \geq c_N \bar{P}_A^{c_N+c_A}$ . This inequality is sufficiently proved, if  $\mathcal{A}_{c_N+c_A} \setminus \mathcal{C}_A$  consists of only recipients  $a \in {}_A\mathcal{H} \setminus \mathcal{C}_A$ , where  $P_A(a) \geq \bar{P}_A^{c_N+c_A}$ , or if all  $a$ , where  $P_A(a) \geq \bar{P}_A^{c_N+c_A}$  are in  $\mathcal{A}_{c_N+c_A} \setminus \mathcal{C}_A$ .

The relations  $\bar{P}_A^i \leq \bar{P}_A^{c_N+c_A}$  and  $(\mathcal{A}_{c_N+c_A}^i \setminus \mathcal{C}_{Ai}) \subseteq (\mathcal{A}_{c_N+c_A} \setminus \mathcal{C}_A)$  implies that  $\mathcal{A}_{c_N+c_A} \setminus \mathcal{C}_A$  contains all Alice's friends  $a$ , where  $P_A(a) \geq \bar{P}_A^i$  and thus also all recipients  $a'$ , where  $P_A(a') \geq \bar{P}_A^{c_N+c_A}$  for  $a, a' \in {}_A\mathcal{H} \setminus \mathcal{C}_{Ai}$ . This proves  $P_A(\mathcal{A}_{c_N+c_A} \setminus \mathcal{C}_A) = \frac{c_N}{m} + \delta \geq c_N \bar{P}_A^{c_N+c_A} = \frac{c_N}{m} - \frac{c_N \epsilon}{m-c_A}$  and that inequality (4.28) applies to  $c_N, \delta, c_A, \epsilon$ .

This means that if ExactHS prefers to choose and remove the most frequent recipients first, then it will in general compute  $\mathcal{C}_i \sim \mathcal{A}_{c_N+c_A}^i$  at the  $i$ -recursion level. No matter which recipients ExactHS will chose or remove at the succeeding levels, it will not require more recipient choices to disprove  $\mathcal{H}$  then in the case that Alice communicates uniformly.  $\square$

**Perturbations** There might be specific matches of communication behaviour of Alice and other senders, so that the time-complexity of ExactHS could potentially be higher than predicted by our analyses. Such a case might appear if the order of choosing Alice's friends by ExactHS is skewed by the traffic of other senders to Alice's friends. However, if  $c$  recipients are required for a disproof, then the relative order of recipient choices within these  $c$  recipients has little effects on the time-complexity of ExactHS. Perturbations also becomes negligible in cases of low communication frequency of senders other than Alice, that is if  $\max_{r \in R} P_N(r) \ll \frac{1}{m}$ .

## 4.2.4 Refined Mean Time-Complexity

Section 4.1.4 provides an estimate of the upper bound the mean time-complexity of ExactHS with respect to the Mix parameters, which is invariant to the traffic distributions. We estimate in this section the mean time-complexity of ExactHS under the optimistic case assumption which models the effect of traffic distributions.

**Number of Finalised Sets and Time-Complexity** We classify each hypothesis according to pair  $(c_N, c_A)$  of hypothetical minimal mean number of choices of non-friends and Alice's friends to disprove it. To avoid evaluating single pairs  $(c_N, c_A)$

#### 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

for each hypothesis, we evaluate the pairs  $(c_N^{mmin}(c_A), c_A)$  instead, where  $c_N^{mmin}(c_A)$  is the number of non-friends to disprove  $c_A$ -average-disprovable hypotheses as defined in Section 4.1.4. The value of  $c_N^{mmin}$  and  $c_A$  determines the structure of the finalised sets computed by ExactHS. Applying the analyses of the quantity of these structures of Section 5.1.1.2 finally derives the number of finalised sets computed by ExactHS and thus its complexity with respect to distinct Alice's traffic distributions.

For each pair  $(c_N^{mmin}(c_A), c_A)$ ,  $c_A \in \{0, \dots, m-1\}$ , ExactHS chooses at most  $c_N^{mmin}(c_A) + c_A$  recipients to disprove  $c_A$ -average-disprovable hypotheses. There are exactly  $c_A$  Alice's friend choices that takes place in some of the  $c_N^{mmin}(c_A) + c_A$  the levels of recursion in ExactHS. We assume for simplicity that there is only one Alice's friend in each observation<sup>1</sup>, and thus one possible choice of Alice's friend in each level of recursion. This result in  $\binom{c_N^{mmin}(c_A) + c_A}{c_A}$  possibilities of choosing a set of  $c_A$  Alice's friends<sup>2</sup> in the  $c_N^{mmin}(c_A) + c_A$  levels of recursion in ExactHS.

Let us consider a particular chosen set of  $c_A$  friends and fix the level of recursion in that these recipients are chosen. ExactHS must then choose a non-friend in each of those levels of recursion, where no Alice's friends were chosen. In each of those levels of recursion, there are at most  $(b-1)$  possible non-friends from which ExactHS can choose one. This result in  $(b-1)^{c_N^{mmin}(c_A)}$  possibilities of choosing a set of  $c_N^{mmin}(c_A)$  non-friends.<sup>3</sup>

Combining the possible  $c_A$  Alice's friend choices and  $c_N^{mmin}$  non-friend choices, we obtain the next estimate of the mean number of finalised sets computed by ExactHS for the unique identification of  ${}_A\mathcal{H}$ :

$$g(b, m) = \sum_{l=0}^m \binom{l + c_N^{mmin}(l)}{l} (b-1)^{c_N^{mmin}(l)} . \quad (4.29)$$

The mean time-complexity of ExactHS is thus  $O(g(b, m)mtb)$ , where  $t$  is the number of observations to identify  ${}_A\mathcal{H}$  derived in Section 3.1. To clarify that the number of finalised sets is dependent on the function  $c_N^{mmin}(\cdot)$ , we also use the notation

---

<sup>1</sup>Observations with more than one friend can lead to some increase, or decrease of the number of finalised sets.

<sup>2</sup>If ExactHS only chooses  $c'_N < c_N^{mmin}(c_A)$  non-friends to disprove  $c_A$ -disprovable hypotheses, then  $\binom{c'_N + c_A}{c_A} < \binom{c_N^{mmin}(c_A) + c_A}{c_A}$ .

<sup>3</sup>Readers interested in a more detailed proof are referred to Section 5.1.1.2.

---


$$g(b, m, c_N^{mmin}(\cdot)).$$

## 4.2.5 Evaluation

This section evaluates mathematically and simulatively the mean number of observations and time-complexity to unambiguously identify Alice's set of friends using the enhanced version of ExactHS in Algorithm 2. We consider the practical case, where  $\max_{r \in R} P_N(r) < \frac{1}{m}$ . The evaluations illustrate the effect of Alice contacting her friends according to a Zipf( $m, \alpha$ ) distribution, chosen due to its known similarity to Internet traffic Adamic and Huberman [2002]; Almeida et al. [1996]; Breslau et al. [1999]; Glassman [1994].

### 4.2.5.1 Solving Number of Recipient Choices for Disproofs

In our evaluations, Alice chooses her friend according to the Zipf( $m, \alpha$ ) distribution in each observation. The remaining  $(b-1)$  senders are assumed to choose their recipients arbitrarily, but such that the cumulative probability that a recipient is contacted by a non-Alice sender is at most  $P_N = 1 - (\frac{u-1}{u})^{b-1}$ .

To simplify the maths, we approximate the cumulative Zipf distribution function  $F_z^{m,\alpha}(y)$  described in Section 2.2.4.1 by considering it as a continuous function. That is  $F_z^{m,\alpha}(y) \approx \frac{\int_{i=1}^{y+1} i^{-\alpha} di}{\int_{j=1}^{m+1} j^{-\alpha} dj} = \frac{(y+1)^{1-\alpha}-1}{(m+1)^{1-\alpha}-1}$  for  $\alpha \neq 1$ . Applying this approximation to (4.26) allows deriving the following closed formula:

$$f_{Po}^{mmin}(c_N, c_A) \approx 1 - \frac{(c_N + c_A + 1)^{1-\alpha} - (c_A + 1)^{1-\alpha}}{(m + 1)^{1-\alpha} - 1} + \left(1 - \frac{(c_A + 1)^{1-\alpha} - 1}{(m + 1)^{1-\alpha} - 1}\right)(m - c_N - c_A)P_N.$$

To simplify discussions, we consider  $f_{Po}^{mmin}(c_N, c_A)$  as being equal the formula on the right side of ( $\approx$ ) in the remaining of this section. We obtain the function  $c_N^{mmin}(c_A)$  by

#### 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

solving the next equation for  $c_N$ , as sketched in Section 4.2.3:

$$\begin{aligned}
 1 &= f_{Po}^{min}(c_N, c_A) \\
 &= 1 - \frac{(c_N + c_A + 1)^{1-\alpha} - (c_A + 1)^{1-\alpha}}{(m + 1)^{1-\alpha} - 1} + \\
 &\quad \left(1 - \frac{(c_A + 1)^{1-\alpha} - 1}{(m + 1)^{1-\alpha} - 1}\right)(m - c_N - c_A)P_N .
 \end{aligned} \tag{4.30}$$

We used Maple to solve (4.30) for  $c_N$ . This result in the function  $c_N^{min}(c_A)$  that is plotted on the left plot in Figure 4.6 and Figure 4.7.

**Theoretical Number of Recipient Choices for Disproofs** The left plot in Figure 4.6 and Figure 4.7 shows the number of recipient choices to disprove hypotheses by ExactHS, when the HS-attack succeeds, for the Mix parameters ( $u = 400, b = 10, m = 23$ ) respective ( $u = 20000, b = 50, m = 40$ ). It shows  $c_N^{min}(c_A)$  on the y-axis for the values of  $\alpha = 0, 0.5, 1, 1.5$  and with respect to the values of  $c_A$  on the x-axis. In all plots,  $\alpha$  is the weight in the  $\text{Zipf}(m, \alpha)$  distribution of Alice's traffic defined in Section 2.2.4.1.

We observe that the number of recipient choices to disprove  $c_A$ -average-disprovable sets is maximal, if Alice's traffic is uniformly distributed ( $\alpha = 0$ ) and decreases with respect to  $\alpha$ . This observation complies to our proof in Section 4.2.3.2 and confirms it.

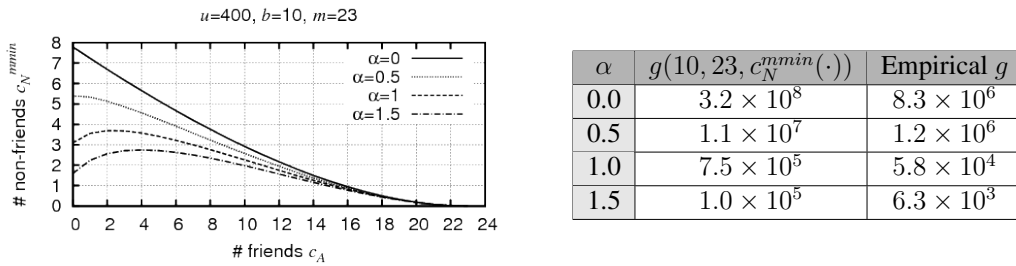


Figure 4.6: Number of recipient choices for disproofs and of finalised sets. *Left:* Number of choices  $c_N^{min}(\cdot)$ . *Right:* Theoretical vs. empirical number of finalised sets.

**Theoretical and Empirical Number of Finalised Sets** The right table in Figure 4.6 and Figure 4.7 enlists the theoretical and empirical number of finalised sets computed



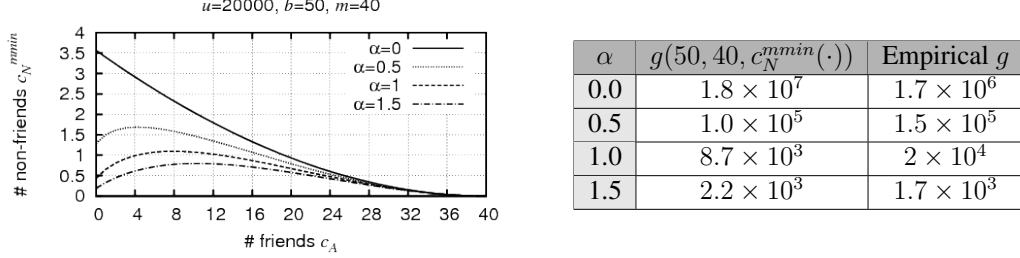


Figure 4.7: Number of recipient choices for disproofs and of finalised sets. *Left*: Number of choices  $c_N^{min}(\cdot)$ . *Right*: Theoretical vs. empirical number of finalised sets.

by ExactHS, when the HS-attack succeeds, for the Mix parameters ( $u = 400, b = 10, m = 23$ ) respective ( $u = 20000, b = 50, m = 40$ ).

The second column in the table determines the theoretical mean number of finalised sets by (4.29), for the given values of  $u, b, m, \alpha$  and the corresponding function  $c_N^{min}(\cdot)$  illustrated on the left plot in the same figure.

The third column in the table represents the empirical mean number of finalised sets obtained from simulations described in Section 2.2.4. These simulations apply the HS-attack to the same Mix configurations as considered in the theoretical analyses. Their results are graphically illustrated in Figure 2.8 (for  $u = 400, b = 10, m = 23$ ) and in Figure 2.9 (for  $u = 20000, b = 50, m = 40$ ) in Section 2.2.4.2.

We observe that the theoretical and empirical mean number of finalised sets decreases with respect to  $\alpha$ . This complies to our analytical analysis of the function  $c_N^{min}(\cdot)$  represented on the left plot. That is  $c_N^{min}(c_A)$  decreases for all  $c_A$  with respect to  $\alpha$  and therefore implies that the number of finalised sets (4.29) also decreases.

Let without loss of generality Alice's set of friends be  ${}_A\mathcal{H} = \{1, \dots, m\}$ , where  $a \in {}_A\mathcal{H}$  is the  $a$ -th most frequently contacted by Alice. For the Mix parameters ( $u = 400, b = 10, m = 23$ ), the traffic volume of the cover-traffic to any recipient  $r \in R$  is  $P_N \approx 0.02$ . This is higher than Alice's traffic volume to her friends  $a \geq 13$  in case of  $\alpha = 1$  and  $a \geq 8$  in case of  $\alpha = 1.5$ . That is  $P_N > P_z^{m,1}(a)$  for  $a \geq 13$  and  $P_N > P_z^{m,1.5}(a)$  for  $a \geq 8$ . Figure 4.6 thus reveals that the mean number of finalised sets computed by ExactHS decreases with respect to  $\alpha$ , although Alice's traffic volume to an increasingly number of friends becomes lower than the cover-traffic volume. This relation seems to be robust, as it mainly remains, even in simulation where

## 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

we attempted to perturb the recipient choices by ExactHS to increase its mean time-complexity. In those simulations that are not presented in this thesis, we considered non-uniformly distributed cover-traffics that skew the order of the recipients chosen by ExactHS. However, due to a lack of analytical analyses for those simulations we do not detail them in this thesis and just mention them as a side note.

### 4.3 Summary

Provided that the attacker has collected sufficiently many observations for a unique identification of Alice's set of friends, we provided analytical analyses of the mean time-complexity of ExactHS for that identification. They allow mathematical identification of non trivial and practical Mix configurations, where the UMHS problem can be efficiently solved by ExactHS, despite the exponential worst case time-complexity of ExactHS. This shows that the anonymity protection provided by an anonymity system is more appropriately measured by the mean complexity of an attack, than by its worst case complexity. At the same time our results proves that the HS-attack is not a purely theoretic measure of the least number of observations to uniquely identify a user's set of friends, but also a practical one, as it is in many case computationally feasible.

Section 4.1 contributed an estimate of the upper bound of the mean time-complexity of ExactHS with respect to the Mix parameters. This bound is invariant to Alice's traffic distribution and can be applied to obtain bounds for non-uniformly distributed cover-traffic. Our analyses allow mathematical identification of those Mix parameters that lead to a mean time-complexity of ExactHS that is bounded by an arbitrary desired upper bound. Section 4.1.4.1 showed that ExactHS succeeds within linear mean time-complexity, if for every recipient, the probability  $P_N$  that it is contacted in an observation by any sender of the cover-traffic does not exceed  $\frac{1}{m^2}$ . In other words, ExactHS can uniquely identify  ${}_A\mathcal{H}$  within linear mean time-complexity, for any Mix configuration, where the distribution of the cover-traffic provides a probability  $P_N \leq \frac{1}{m^2}$ . This time-complexity is independent of Alice's traffic distribution<sup>1</sup>.

Simulations in Section 4.1.5 revealed that the variance of the mean time-complexity

---

<sup>1</sup>That is Alice's traffic distribution does not increase this time-complexity.

---

of ExactHS might be high, in the case that  ${}_A\mathcal{H}$  is identified with the least number of observations. This was caused by high variances of the traffic distributions, due to the low number of observations required to succeed the HS-attack. This variance becomes negligible, if ExactHS is reapplied with more than the least number of observations to succeed the HS-attack, by collecting additional observations. However, the mean time-complexity of ExactHS does not decrease, even for significantly more additional observations. This indicates that that complexity is mainly independent of the number of observations that is significantly beyond the least number of observations to succeed the HS-attack.

Section 4.2 investigated the effect of the distribution of Alice’s traffic on the mean time-complexity of ExactHS, thus extending and refining Section 4.1. It provided a mathematical estimate of that complexity with respect to various Mix parameters and traffic distributions. Section 4.2.3.2 proved that, if for every recipient, the probability  $P_N$  that it is contacted in an observation by any sender of the cover-traffic is below  $\frac{1}{m}$ , then the mean time-complexity of ExactHS is maximal, if Alice’s traffic distribution is uniform. Although our analyses apply to arbitrary distributions, we chose the Zipf( $m, \alpha$ ) distribution of Alice’s friends to study the decrease of ExactHS’s mean time-complexity in detail. We proved in Section 4.2.3.2 that the mean time-complexity of ExactHS decreases with respect to  $\alpha$  and approaches a polynomially bounded complexity. That is despite the computational infeasibility of the mean time-complexity of ExactHS, when Alice’s traffic is uniformly distributed.

It could be concluded from Section 3.1.1.2 and Section 4.2.4 that the mean number of observations to uniquely identify  ${}_A\mathcal{H}$  decreases with respect to  $\alpha$ , while the mean time-complexity of that identification decreases. This proved that the uniform distribution of Alice’s traffic minimises the mean number of observations to uniquely identify  ${}_A\mathcal{H}$ , while maximising the mean time-complexity of that identification.

We also observed that the number of observations increases reasonably with  $\alpha$ , while the mean time-complexity seems to decrease exponentially fast with increasing value of  $\alpha$ . Although we can estimate the mean time-complexity mathematically with respect to various Mix configurations, the equations are yet not simple enough to provide a direct proof of that observation. Proving this relation is left for future works.

The mathematical analyses in this chapter referred to the case that  $P_N(r) < \frac{1}{m}$ . Our simulations (that are not evaluated in this thesis) indicated cases where ExactHS

#### 4. PRACTICAL LIMIT OF ANONYMITY PROTECTION

---

can be efficiently applied, even if  $P_N(r) \geq \frac{1}{m}$ . The mathematical analyses of these cases remain for future works.

It cannot be excluded that there will be an algorithm in the future that solves the UMHS problem with a lower mean time-complexity than ExactHS. Therefore, our analyses of the practical limit of anonymity protection estimates the maximal, but not the least anonymity protection provided by a Mix configuration.

# Chapter 5

## Extension

The previous chapters analysed the least number of observations to uniquely identify Alice’s set of friends, as well as the worst case and mean time-complexity of ExactHS for that identification. These analyses are based on the following two assumptions: First, the attacker can evaluate sufficiently many observations for the unique identification of Alice’s set of friends. Second, the observed information about the senders and recipients of the messages relayed in a Mix round is exact and complete<sup>1</sup>. This chapter investigates the HS-attack for the case that one of these two assumptions is omitted, which we call the case of “relaxed assumptions”. We show that the basic idea of the HS-attack, the evaluation of minimal-hitting-sets, remains applicable in that case and allows deducing knowledge about Alice’s friends. This only requires slight adaptations of the HS-attack described in Chapter 2.

By relaxing the first assumption, we aim to study the gain of partial or likely information about Alice’s friends, when the attacker can only observe a restricted number of observations. We investigate by the relaxation of the second assumption, the identification of Alice’s friends, even if some of the attacker’s observations are erroneous. This addresses practice-oriented Mix and attacker models, where an attacker has incomplete knowledge about the sender-anonymity set or recipient set in around, cf. Section 1.2.3.

Section 5.1 analyses the information that can be disclosed about Alice’s friends, if there are not sufficiently many observations for a unique identification of Alice’s set of

---

<sup>1</sup>Recall that the links between the senders and recipients of single messages are hidden by the Mix.

## 5. EXTENSION

---

friends. Section 5.2 considers the case that the attacker occasionally collects so-called *erroneous observations* that miss Alice’s friends. It analyses conditions, where Alice’s set of friends can still be identified with a high certainty despite these errors.

### 5.1 Partial Information

The HS-attack determines all specified hypotheses for Alice’s set of friends. The past chapters evaluated these specified hypotheses for the unique identification of Alice’s set of friends that we call the *full disclosure*.

This section studies the case, where the number of observations might not be sufficient for a full disclosure by the HS-attack. For this case, we evaluate the specified hypotheses for the unique identification of subsets of Alice’s set of friends, that we call the *partial disclosure*.

In essence, each of Alice’s friend is uniquely identifiable if he is contained by every specified hypothesis. This is equivalent to being contained by every (minimal-) hitting-set computed by ExactHS, since every specified hypothesis is a superset of a minimal-hitting-set<sup>1</sup>. Therefore, by computing the intersection  $\bigcap_{\mathcal{H} \in \mathcal{HS}} \mathcal{H}$  of all (minimal-) hitting-sets computed by ExactHS (Algorithm 1, 2), we obtain the unique identification of a subset of Alice’s friends, a partial disclosure. A partial disclosure can be illustrated by Example 2.1 in Section 2.1.2.3. In that example, applying the intersection of all minimal-hitting-sets, respectively of all specified hypotheses in  $i = 6$  leads to the unique identification of the subset  $\{1, 2\}$  of Alice’s set of friends  $\{1, 2, 3\}$ .

This section contributes analytical analyses of the number of observations for the partial disclosure of Alice’s friends for the uniform communication model, as a first step in that direction. We provide a mathematical description of the evolution of the minimal-hitting-sets (i.e., the minimal-hitting-sets that remain after applying ExactHS) with respect to the number of observations collected by the attacker. This allows concluding the information that can be partially disclosed, as it is derived from those remaining minimal-hitting-sets.

---

<sup>1</sup>We refer to minimal-hitting-set of at most size  $m$  and hypothesis definitions in Section 2.1.2.3.

---

To aid describing the evolution of minimal-hitting-sets, Section 5.1.1 provides a quantification of all distinct minimal-hitting-set classes. This reveals that the quantity of the minimal-hitting-sets is non monotonous with respect to the number of observations collected by the attacker, so that it would be difficult to describe the minimal-hitting-sets mathematically.

As a solution to this problem, Section 5.1.2 introduces an abstract structure called extensive-hypotheses. The set of extensive-hypotheses covers all information represented by the minimal-hitting-sets, while being monotonous with respect to the number of observations. Section 5.1.3 outlines that the evolution of the extensive-hypothesis can be easily described and provides conclusions about the evolution of the minimal-hitting-sets.

### 5.1.1 Quantification of Minimal-Hitting-Sets in a Class

Section 2.3.1 introduced a classification of minimal-hitting-sets. We are interested in an analytical quantification of the maximal number of minimal-hitting-sets in each of those classes, in this section.

Section 5.1.1.1 determines for each class  $\mathfrak{H}_i$ , for  $i = 0, \dots, m$ , all those minimal-hitting-sets computed by ExactHS that are in that class, with respect to the attacker's observation set  $\mathcal{OS}$ . This is realised by a modified invocation of Algorithm 2. While this modified invocation allows computing all minimal-hitting-sets as before, it also enables a simple quantification of the number of minimal-hitting-sets in each class  $\mathfrak{H}_i$ , as shown in Section 5.1.1.2.

#### 5.1.1.1 Computing Minimal-Hitting-Sets in a Class

In this section, we provide an algorithm to reveal all minimal-hitting-sets in a particular class, that are computed by ExactHS with respect to a given observation set  $\mathcal{OS}$ . While this algorithm is based on a modified invocation of ExactHS, it solely serves to aid analytical analyses and should not be confused with an attack.

Let a particular subset  $\mathcal{C} \subseteq {}_A\mathcal{H}$  of Alice's friends of size  $m - i$  be fixed and  $\mathcal{OS}$  be the observations collected by the attacker, for  $i = 0, \dots, m$ . Executing the following three steps computes all minimal-hitting-sets of at most size  $m$ , that contains

## 5. EXTENSION

---

exactly all Alice's friends in  $\mathcal{C}$ , by using Algorithm 2:

1. All observations containing any friend in  $\mathcal{C}$  are removed from the initial observation-set  $\mathcal{OS}$ , thus obtaining  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$ .
2. All remaining friends in  ${}_A\mathcal{H} \setminus \mathcal{C}$  are removed from observations in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$ , resulting in the modified observation set

$$\mathcal{OS}' = \{\mathcal{O} \setminus {}_A\mathcal{H} \mid \mathcal{O} \in \mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]\}.$$

3. Algorithm 2 is applied on  $\mathcal{OS}'$  to compute all minimal-hitting-sets containing  $\mathcal{C}$  and at most  $i$  non-friends, by invoking  $ExactHS(\mathcal{OS}', i, \mathcal{C})$ .

Step 1 aids the computation of minimal-hitting-sets  $\mathcal{H}'$  in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$ . Due to Claim 2 in Section 2.2.2.1, this allows determining all minimal-hitting sets  $\mathcal{H} \supseteq \mathcal{C}$  in  $\mathcal{OS}$  by setting  $\mathcal{H} = \mathcal{H}' \cup \mathcal{C}$ . Step 2 constrains  $\mathcal{H}'$  to not contain any Alice's friends, so that every  $\mathcal{H} = \mathcal{H}' \cup \mathcal{C}$  is a hitting-set in  $\mathcal{OS}$  that contains exactly  $|\mathcal{C}| = m - i$  Alice's friends.

Applying the above three steps for all possible sets  $\mathcal{C} \subseteq R \setminus {}_A\mathcal{H}$ , where  $|\mathcal{C}| = m - i$  thus determines the subset of all minimal-hitting-sets in  $\mathfrak{H}_i$ , for  $i \in \{0, \dots, m\}$ , by  $ExactHS$ . Note that we do not exclude the computation of non minimal-hitting-sets for the same reason as in Section 2.2.2.3, as it does not change the maximal number of minimal-hitting-sets in the classes as shown in the next section.

### 5.1.1.2 Maximal Number of Minimal-Hitting-Sets in a Class

There are  $\binom{m}{m-i}$  possibilities to choose  $(m - i)$  different friends out of Alice's set of friends  ${}_A\mathcal{H}$ . Calling Algorithm 2 in Step 3 will compute at most  $(b - 1)^i$  different minimal-hitting-sets. The base of  $(b - 1)$  is due to Step 2, that removes all friends from all observations in  $\mathcal{OS} \setminus \mathcal{OS}[\mathcal{C}]$ . Let  $\mathfrak{H}_i^* \subseteq \mathfrak{H}_i$  denote the maximal set of all minimal-hitting-sets in  $\mathfrak{H}_i$  that  $ExactHS$  can compute, then

$$|\mathfrak{H}_i^*| = \binom{m}{m-i} (b - 1)^i = \binom{m}{i} (b - 1)^i \quad (5.1)$$



---

is the maximal number of minimal-hitting-sets in the class  $\mathfrak{H}_i$ . We prove that this bound is tight by showing observation sets, where the number of all minimal-hitting-sets is exactly  $b^m$ , while the number of minimal-hitting-set in each class  $\mathfrak{H}_i$  is exactly  $|\mathfrak{H}_i^*|$  for  $i = 0, \dots, m$ , next.

*Proof.* We prove that the maximal number of minimal-hitting-sets  $|\mathfrak{H}_i^*|$  in each class  $\mathfrak{H}_i$ , for  $i = 0, \dots, m$  is tight, by considering the same observation sets as in the proof of Claim 3 in Section 2.2.3.1. That is the set of  $m$  pairwise disjoint observations  $\mathcal{OS} = \{\mathcal{O}_0, \dots, \mathcal{O}_{m-1}\}$ , where  $|\mathcal{O}_i| = b$ , for  $i = 0, \dots, m-1$ . Due to the proof of Claim 3, there are exactly  $b^m$  minimal-hitting-sets of at most size  $m$  in  $\mathcal{OS}$ . Each of these minimal-hitting-sets  $\mathcal{H}$  consists of  $m$  recipients and is of the structure  $\mathcal{H} = \{r_1, \dots, r_m\}$ , where  $r_j \in \mathcal{O}_{j-1}$  for  $j = 1, \dots, m$ . Since each  $\mathcal{O} \in \mathcal{OS}$  contains exactly one Alice's friend, the observations  $\mathcal{O}_0, \dots, \mathcal{O}_{m-1}$  contain all  $m$  distinct Alice's friends. Let  $\mathcal{H} \in \mathfrak{H}_i$  be a minimal-hitting-set in  $\mathcal{OS}$ , then the  $(m-i)$  Alice's friends in  $\mathcal{H}_A$  can be chosen from  $\binom{m}{m-i}$  possible sets of  $(m-i)$  observations in  $\mathcal{OS}$ . The remaining  $i$  non-friends in  $\mathcal{H}_N$  can be chosen from the remaining  $i$  observations. Since each observation contains  $(b-1)$  non-friends, there are  $(b-1)^i$  sets of  $i$  non-friends for  $\mathcal{H}_N$ . Therefore, there are  $|\mathfrak{H}_i^*| = \binom{m}{m-i}(b-1)^i$  minimal-hitting-sets  $\mathcal{H}$  in  $\mathfrak{H}_i$ .  $\square$

Note that the cumulative sum of the maximal number of minimal-hitting-sets in all classes  $\mathfrak{H}_i$  is exactly the maximal number of minimal-hitting-sets of  $b^m$ . That is

$$\sum_{i=0}^m |\mathfrak{H}_i^*| = \sum_{i=0}^m \binom{m}{m-i} (b-1)^i = \sum_{i=0}^m \binom{m}{i} (b-1)^i 1^{m-i} = b^m .$$

**Example 5.1.1** (Tightness of Estimated Number Minimal-Hitting-Sets). *We show a concrete set of observations that contains exactly  $b^m$  minimal-hitting-sets, where the number of minimal-hitting-sets in each class is exactly  $|\mathfrak{H}_i^*|$  for  $i = 0, \dots, m$ . Let Alice's set of friends be  ${}_A\mathcal{H} = \{1, 2, 3\}$ ,  $m = 3$  and  $b = 3$  in our example. The  $m$  disjoint observations  $\{\mathcal{O}_0, \dots, \mathcal{O}_{m-1}\}$  collected by the attacker are:*

$$\{1, 4, 5\}, \{2, 6, 7\}, \{3, 8, 9\} .$$

## 5. EXTENSION

---

The minimal-hitting-sets in the classes  $\mathfrak{H}_0, \dots, \mathfrak{H}_m$  are:

$$\begin{aligned}\mathfrak{H}_0^* &= \{\{1, 2, 3\}\} \\ \mathfrak{H}_1^* &= \{\{1, 2, 8\}, \{1, 2, 9\}, \{1, 6, 3\}, \{1, 7, 3\}, \{4, 2, 3\}, \{5, 2, 3\}\} \\ \mathfrak{H}_2^* &= \{\{1, 6, 8\}, \{1, 6, 9\}, \{1, 7, 8\}, \{1, 7, 9\}, \{4, 2, 8\}, \{4, 2, 9\}, \{5, 2, 8\}, \{5, 2, 9\}, \\ &\quad \{4, 6, 3\}, \{4, 7, 3\}, \{5, 6, 3\}, \{5, 7, 3\}\} \\ \mathfrak{H}_3^* &= \{\{4, 6, 8\}, \{4, 6, 9\}, \{4, 7, 8\}, \{4, 7, 9\}, \{5, 6, 8\}, \{5, 6, 9\}, \{5, 7, 8\}, \{5, 7, 9\}\}\end{aligned}$$

The reader can verify for all classes  $\mathfrak{H}_i^*$ , for  $i = 0, \dots, m$  that  $|\mathfrak{H}_i^*| = \binom{m}{m-i} b^i$ , which is  $|\mathfrak{H}_i^*| = \binom{3}{3-i} 2^i$  in this example. The number of all minimal-hitting-sets is  $\sum_{i=0}^3 \binom{3}{3-i} 2^i = 1 + 6 + 12 + 8 = 3^3$ , which is  $b^m$ .

### 5.1.2 Description of Minimal-Hitting-Sets by Extensive-Hypotheses

This section provides a description of the minimal-hitting-sets of at most size  $m$  in the sequence of observations sets  $\mathcal{OS}_0, \mathcal{OS}_1, \dots, \mathcal{OS}_t$  collected by the attacker. We define that  $\mathcal{OS}_0 = \{\}$ ,  $\mathcal{OS}_i = \mathcal{OS}_{i-1} \cup \mathcal{O}_i$ , where  $\mathcal{O}_i$  is the  $i$ -th observation collected by the attacker, for  $i = 1, \dots, t$ . For all  $i = 1, \dots, t$ , the minimal-hitting-sets in  $\mathcal{OS}_{i+1}$  result from removing and extending the minimal-hitting-sets in  $\mathcal{OS}_i$ , such that all resulting hitting-sets hit the new observation  $\mathcal{O}_{i+1}$ . A hitting set  $\mathcal{H}$  in  $\mathcal{OS}_i$  that is no hitting-set in  $\mathcal{OS}_{i+1}$ , is removed, if  $|\mathcal{H}| = m$ , otherwise it is replaced by an extension  $\mathcal{H} \cup r$  for any  $r \in \mathcal{O}_{i+1}$ . We call these changes of the minimal-hitting-sets due to new observations, an *evolution* of the minimal-hitting-sets. Our description of the minimal-hitting-sets is called *extensive-hypotheses*. It aids deriving a simple mathematical model for the evolution of the minimal-hitting-sets with respect to the number of observations collected by the attacker.

We observe that the number of minimal-hitting-sets of at most size  $m$  tends to increase at the beginning with respect to the number of observations collected by the attacker. At some turning point, that number tends to decrease for increasing number of observations, until it arrives at a unique minimum-hitting-set that identifies all Alice's friends. At the same time the sizes of the minimal-hitting-sets approach  $m$  with increasing

---

number of observations. Due to the alternating quantity of the minimal-hitting sets and their changing sizes, modelling the evolution of the minimal-hitting-sets with respect to the attacker's observations, would be mathematically complex.

However, analyses of ExactHS in Section 5.1.1 provide information about the classification of all potential minimal-hitting-sets of size  $m$  as well as their quantity in each class, even if they are not yet computed by ExactHS. We call these potential minimal-hitting-sets of size  $m$  *extensive-hypotheses*. They will be such defined, that every minimal-hitting-set of at most size  $m$  in an observation set is a subset of an extensive-hypothesis. Therefore, the set of extensive-hypotheses that remain valid in a given observation set  $\mathcal{OS}$  will be dual to the set of minimal-hitting-set in  $\mathcal{OS}$ , so that partial information that are derivable by minimal-hitting-sets are derivable by extensive-hypotheses.

As illustrated by the example in Table 5.1, the number of extensive-hypotheses only decreases with respect to the number of observations collected by the attack, while the size of an extensive-hypothesis is always  $m$ . This allows for mathematical description of the evolution of the extensive-hypotheses that is less complex than describing minimal-hitting-sets.

### 5.1.2.1 Evolution of Extensive-Hypotheses

Table 5.1 shows the set of all minimal-hitting-sets  $\mathcal{M}_i$ , the set of all extensive-hypotheses  $\mathcal{L}_i$  and the disproved sets that result from evaluating the observation set  $\mathcal{OS}_i = \{\mathcal{O}_1, \dots, \mathcal{O}_i\}$  collected by the attacker. It is an example, where Alice's set of friends is  ${}_A\mathcal{H} = \{1, 2, 3\}$  and the batch size is  $b = 2$ .

For  $i = 0$  there is no observation and thus no minimal-hitting-set, however, we know by Section 2.3.1 the initial classes of all extensive-hypotheses, as represented in  $\mathcal{L}_0$  in Table 5.1. Each element  $\mathcal{H} \in \mathcal{L}_i$  is called an *extensive-class*<sup>1</sup>. We will index these classes by their order from left to right and from top to bottom by subscripts, such that  $\mathcal{H}_u$  is the  $u$ -th class in the set of all extensive-hypotheses.

An extensive-class is *unspecified*, if it contains a variable  $x$ ,<sup>2</sup> otherwise it is *speci-*

---

<sup>1</sup>Writing  $\mathcal{H} \subseteq \mathcal{L}_0$  would be formally more correct, as  $\mathcal{H}$  represents several extensive-hypotheses, but we entirely use the element-notation and -operations with  $\mathcal{H}$  to simplify notations and explanations.

<sup>2</sup>We address by  $x$  any indexed variable  $x_u^v$  thereof.

## 5. EXTENSION

$i$	$\mathcal{O}_i$	MHS $\mathcal{M}_i$	Extensive-hypotheses $\mathcal{L}_i$	Disproved
0			$\mathfrak{H}_0 : \{1, 2, 3\};$ $\mathfrak{H}_1 : \{1, 2, x_1^2\}, \{1, 3, x_1^3\}, \{2, 3, x_1^4\};$ $\mathfrak{H}_2 : \{1, x_1^5, x_2^5\}, \{2, x_1^6, x_2^6\}, \{3, x_1^7, x_2^7\};$ $\mathfrak{H}_3 : \{x_1^8, x_2^8, x_3^8\}$	
1	$\{1, 4\}$	$\{1\},$ $\{4\}$	$\mathfrak{H}_0 : \{1, 2, 3\};$ $\mathfrak{H}_1 : \{1, 2, x_1^2\}, \{1, 3, x_1^3\}, \{2, 3, 4\};$ $\mathfrak{H}_2 : \{1, x_1^5, x_2^5\}, \{2, 4, x_1^6\}, \{3, 4, x_1^7\};$ $\mathfrak{H}_3 : \{4, x_1^8, x_2^8\}$	
2	$\{2, 5\}$	$\{1, 2\},$ $\{1, 5\},$ $\{4, 2\},$ $\{4, 5\}$	$\mathfrak{H}_0 : \{1, 2, 3\};$ $\mathfrak{H}_1 : \{1, 2, x_1^2\}, \{1, 3, 5\}, \{2, 3, 4\};$ $\mathfrak{H}_2 : \{1, 5, x_1^5\}, \{2, 4, x_1^6\}, \{3, 4, 5\};$ $\mathfrak{H}_3 : \{4, 5, x_1^8\}$	
3	$\{1, 6\}$	$\{1, 2\},$ $\{1, 5\},$ $\{4, 2, 6\},$ $\{4, 5, 6\}$	$\mathfrak{H}_0 : \{1, 2, 3\};$ $\mathfrak{H}_1 : \{1, 2, x_1^2\}, \{1, 3, 5\};$ $\mathfrak{H}_2 : \{1, 5, x_1^4\}, \{2, 4, 6\};$ $\mathfrak{H}_3 : \{4, 5, 6\}$	$\{2, 3, 4\},$ $\{3, 4, 5\}$
4	$\{3, 4\}$	$\{1, 2, 3\},$ $\{1, 2, 4\},$ $\{1, 5, 3\},$ $\{1, 5, 4\},$ $\{4, 2, 6\},$ $\{4, 5, 6\}$	$\mathfrak{H}_0 : \{1, 2, 3\};$ $\mathfrak{H}_1 : \{1, 2, 4\}, \{1, 3, 5\};$ $\mathfrak{H}_2 : \{1, 5, 4\}, \{2, 4, 6\};$ $\mathfrak{H}_3 : \{4, 5, 6\}$	

Table 5.1: Evolution of minimal-hitting-sets (MHS) and extensive-hypotheses for Alice's set of friends  ${}_A\mathcal{H} = \{1, 2, 3\}$  and batch size  $b = 2$ .

---

*fied*. Each variable  $x$  represents any  $(b-1)$  *unspecified recipients*<sup>1</sup> who are non-friends  $n \in R \setminus {}_A\mathcal{H}$ , while numbers or  $a, r, n$  denote concrete recipients who we also call *specified recipients* for clarity. Variables  $x_v^u, x_w^u \in \mathcal{H}_u$ , where  $v \neq w$  always represent distinct unspecified recipients. Therefore, each unspecified extensive-class determines the structure of extensive-hypotheses that are not yet fully specified, while a specified extensive-class describes a single fully specified extensive-hypothesis and we interchangeably refer to the latter by the term *specified extensive-hypothesis*. The only specified extensive-hypothesis in  $\mathcal{L}_0$  is  $\mathcal{H}_1 = \{1, 2, 3\}$ . Recipients who are Unspecified and become specified when evaluating the  $i$ -th observation are bold highlighted in Table 5.1.

The benefit of considering extensive-hypotheses is illustrated in Table 5.1 by the case of  $i = 3$ . It reveals the disproof of extensive-hypotheses (i.e.,  $\{2, 3, 4\}$  and  $\{3, 4, 5\}$ ) and thus the decrease of the number of extensive-hypotheses that are not apparent if considering the minimal-hitting-sets and their quantity. Firstly, we observe that the number of minimal-hitting-sets is increasing with respect to  $i = 1, \dots, 4$ , while the number of extensive-hypotheses and thus the number of potential minimal-hitting-sets of size  $m$  is actually decreasing. Secondly, we can see that given sufficient many observations (i.e.,  $i = 4$  in this table), the set of extensive-hypotheses equals the set of all minimal-hitting-sets. In summary, the evolution of extensive-hypotheses are monotonous and equal that of minimal-hitting-set for large number of observations, and remains monotonous even for lower number of observations, where the quantities and sizes of minimal-hitting-sets alternate.

### 5.1.2.2 Construction of Extensive-Hypotheses

We show the construction of  $\mathcal{L}_i$  with respect to the observation set  $\mathcal{OS}_i = \{\mathcal{O}_1, \dots, \mathcal{O}_i\}$  and the minimal-hitting-sets  $\mathcal{M}_i$  for  $i \geq 1$ . It is based on the adapted application of ExactHS and allows the classification and quantification of extensive-hypotheses according to the classification and quantification of minimal-hitting-set and hypotheses in Section 2.3.1 and in Section 5.1.1, by slight adaptations of the definitions.

---

<sup>1</sup>That is the concrete recipient is as yet unknown.

## 5. EXTENSION

---

Let w.l.o.g. every extensive-class be described by

$$\mathcal{H} = \{a_1, \dots, a_{m-j}, n_1, \dots, n_l, x_{l+1}, \dots, x_j\} ,$$

for  $0 \leq j \leq m, 0 \leq l \leq j$  and  $|\mathcal{H}| = m$ .

We adapt the terminology for hypotheses and minimal-hitting-sets to extensive-classes and -hypotheses as follow:

$\mathcal{H} \in \mathfrak{H}_j$ : If  $\mathcal{H}$  does not contain  $j$  Alice's friends<sup>1</sup>.

$\mathcal{H}_N$ : Set of specified non-friends in  $\mathcal{H}$ , that is  $\mathcal{H}_N = \mathcal{H} \cap R \setminus {}_A\mathcal{H} = \{n_1, \dots, n_l\}$ .

$\mathcal{H}_A$ : Set of Alice's friends in  $\mathcal{H}$ , that is  $\mathcal{H}_A = \mathcal{H} \cap {}_A\mathcal{H} = \{a_1, \dots, a_{m-j}\}$ .

$\mathcal{L}_i$ : For every subset of Alice's friends  $\{a_1, \dots, a_{m-j}\} \subseteq {}_A\mathcal{H}$ , the set  $\mathcal{H} = \{a_1, \dots, a_{m-j}, n_1, \dots, n_l, x_{l+1}, \dots, x_j\}$  is in  $\mathcal{L}_i$  that is denoted  $\mathcal{H} \in \mathcal{L}_i$ , if and only if  $\{n_1, \dots, n_l\}$  is a minimal-hitting-set in  $\mathcal{OS}'_i = \{\mathcal{O} \setminus {}_A\mathcal{H} \mid \mathcal{O} \in \mathcal{OS}_i \setminus \mathcal{OS}_i[a_1, \dots, a_{m-j}]\}$ , for  $0 \leq j \leq m, 0 \leq l \leq j$ . We call these extensive-classes *minimal extensive-classes* (if clarity is needed), thus  $\mathcal{L}_i$  consists of solely minimal extensive-classes. Note, if  $\mathcal{OS}'_i = \{\}$ , then  $\{\}$  is by definition a minimal-hitting-set in that set. Therefore  $\mathcal{L}_0$  contains exactly the  $m + 1$  extensive-classes  $\mathcal{H} = \{a_1, \dots, a_{m-j}, x_1, \dots, x_j\}$  for  $j = 0, \dots, m$ .

Let  $\mathcal{H}'$  be a minimal-hitting-set in  $\mathcal{OS}_i$ , then due to Claim 1 and Claim 2 in Section 2.2.2.1,  $\mathcal{H}' \setminus \mathcal{H}'_A$  is a minimal-hitting-set in  $\mathcal{OS}'_i = \{\mathcal{O} \setminus {}_A\mathcal{H} \mid \mathcal{O} \in \mathcal{OS}_i \setminus \mathcal{OS}_i[\mathcal{H}'_A]\}$ . Therefore, for every minimal-hitting-set  $\mathcal{H}'$  in  $\mathcal{OS}_i$ , there is a  $\mathcal{H} \in \mathcal{L}_i$ , such that  $\mathcal{H}_A = \mathcal{H}'_A$  and  $\mathcal{H}_N = \mathcal{H}'_N$ .

For every subset of Alice's friends  $\{a_1, \dots, a_{m-j}\} \subseteq {}_A\mathcal{H}$  and observation set  $\mathcal{OS}$ , we can compute all minimal-hitting-sets  $\{n_1, \dots, n_l\}$ , for  $l \leq j$  in  $\mathcal{OS}' = \{\mathcal{O} \setminus {}_A\mathcal{H} \mid \mathcal{O} \in \mathcal{OS} \setminus \mathcal{OS}[a_1, \dots, a_{m-j}]\}$  as in Section 5.1.1.1. This allows quantifying the maximal number of extensive-hypotheses represented by each minimal extensive-class based on Section 5.1.1.2. The initial number of extensive-hypotheses in  $\mathcal{L}_0$  that is in  $\mathfrak{H}_j$  is equal

---

<sup>1</sup>Again, writing  $\mathcal{H} \subseteq \mathfrak{H}_j$ , would be formally more correct, but we entirely use the set-notation and -operation with  $\mathcal{H}$  to simplify notations and discussions.

---

the maximal number of minimal-hitting-sets in  $\mathfrak{H}_j$ , which is due to (5.1),

$$\binom{m}{j} (b-1)^j ,$$

for  $j = 0, \dots, m$ . And the number of extensive-hypotheses represented by each minimal extensive-class  $\mathcal{H} = \{a_1, \dots, a_{m-j}, n_1, \dots, n_l, x_{l+1}, \dots, x_j\}$  is

$$(b-1)^{j-l} , \text{ for } 0 \leq l \leq j .$$

However, to provide a better illustration of the evolution of the minimal extensive-classes and the extensive-hypotheses represented by it, we additionally provide an iterative computation of  $\mathcal{L}_i$ , for  $i \geq 1$ . This algorithm computes  $\mathcal{L}_i$  in  $\mathcal{OS}_i = \mathcal{OS}_{i-1} \cup \mathcal{O}_i$ , for given  $\mathcal{L}_{i-1}$  and is presented next.

1. Set  $\mathcal{L}_i = \{\}$  before the start of its construction below.
2. For each extensive-class  $\mathcal{H} \in \mathcal{L}_{i-1}$ , where  $\mathcal{H} \in \mathfrak{H}_j$  for some  $0 \leq j \leq m$ , let  $\{r_1, \dots, r_k\} = \mathcal{H} \cap R$  be the set of all specified recipients for  $k \leq m$ . Apply either 3. or 4. to each  $\mathcal{H}$ .
3. If  $\{r_1, \dots, r_k\} \cap \mathcal{O}_i \neq \emptyset$ , then add  $\mathcal{H}$  to  $\mathcal{L}_i$ , i.e.,  $\mathcal{L}_i = \mathcal{L}_i \cup \{\mathcal{H}\}$ , since all extensive-hypotheses described by  $\mathcal{H}$  remain valid in  $\mathcal{OS}_i$ .
4. Else if  $\{r_1, \dots, r_k\} \cap \mathcal{O}_i = \emptyset$  and  $k < m$ , then add for each non-friend  $n \in \mathcal{O}_i \setminus {}_A\mathcal{H}$  the extensive-class  $\mathcal{H}' = \{r_1, \dots, r_k, n, x_1, \dots, x_{m-k-1}\} \in \mathfrak{H}_j$  to  $\mathcal{L}_i$ , only if  $\mathcal{H}'$  is a minimal extensive-class in  $\mathcal{OS}_i$ .

**Exclusion of Extensive-Hypotheses** We can observe from the algorithm above and in Table 5.1 that unspecified recipients in an extensive-class  $\mathcal{H}$  become specified whenever an observation  $\mathcal{O}$  is collected, where  $\mathcal{H} \cap \mathcal{O} = \emptyset$ , unless  $\mathcal{H}$  is already a specified extensive-hypothesis, so that it becomes disproved. In the usual case, all extensive-classes become specified extensive-hypotheses prior to being disproved. Since there are initially  $b^m$  extensive-hypotheses, we can mathematically model their disproofs with respect to the attacker's number of observations to predict the partial information that can be revealed by the attacker, as will be presented in Section 5.1.3.

## 5. EXTENSION

---

However, in some exceptional cases, some extensive-hypotheses can be excluded, even if they are not specified, thus leading to a faster exclusion of extensive-hypotheses and a faster disclosure of partial information. Modelling these exceptional cases would refine our analyses, but increases the complexity of the analyses. This thesis provides the basic to analyse partial information and therefore leaves this refinement for future works. Nevertheless we outline the exceptional cases next for completeness.

An *exception* can only arise in point 4 in the computation of  $\mathcal{L}_i$  and consists of following cases:

Case 1: The extensive-class  $\mathcal{H}' \in \mathfrak{H}_j$  resulting from specifying a recipient in  $\mathcal{H} \in \mathfrak{H}_j$  is not minimal in  $\mathcal{OS}_i$ .

Case 2: There are fewer than  $(b - 1)$  non-friends in the next observation  $\mathcal{O}_i$ .

These exceptions are illustrated in the example in Table 5.2, where Alice's set of friends is  ${}_A\mathcal{H} = \{1, 2, 3\}$  and the batch size is  $b = 3$ .

For  $i = 2$ , the extensive-classes  $\{3, 4, 8\}, \{4, 8, x\}$  that are highlighted by the grey colour are not minimal and therefore they are excluded, as stated in exception Case 1. These extensive-classes are not minimal, as  $\{3, 4, 8\} \cap R \setminus {}_A\mathcal{H} = \{4, 8\}$  and  $\{3, 4, x\} \cap R \setminus {}_A\mathcal{H} = \{4, 8\}$  are not minimal-hitting-sets in  $\mathcal{OS}_2$ . Note that these extensive-classes hit the observations in  $\mathcal{OS}_2$ .

For  $i = 3$ , the observation  $\mathcal{O}_i = \{1, 6\}$  only contains 2 recipients instead of 3 and leads to a reduction of the number of extensive-hypotheses that remain to be disproved, as stated in exception Case 2. Observe that the extensive-classes  $\{2, 4, x_1^9\}, \{2, 8, x_1^{10}\}, \{3, 8, x_1^{12}\}, \{4, 5, x_1^{13}\}, \{8, x_1^{14}, x_2^{14}\}$  in  $\mathcal{OS}_2$  do not hit  $\mathcal{O}_3$ , so that due to Step 4, an unspecified recipient in each of these classes must be replaced by a non-friend in  $\mathcal{O}_3$  in  $i = 3$ . In the non exceptional case, an observation would contain  $b - 1$  non-friends, so that an unspecified extensive-class  $\mathcal{H}$  would result in at most  $b - 1$  extensive-classes by specifying an unspecified recipient in  $\mathcal{H}$ . However,  $\mathcal{O}_2$  contains only one non-friend, so that each unspecified extensive-class only result in only one extensive-class by specifying an unspecified recipient.

The last column in Table 5.2 enlists the extensive-classes that are explicitly excluded by the algorithm described in “Construction of Extensive-Hypotheses”. These are exclusions due to the exceptional Case 1 (as in  $i = 2$ ) and due to normal disproofs (as in  $i = 4$ ). In opposite to that, exclusions due to Case 2 are not explicitly computed



$i$	$\mathcal{O}_i$	MHS $\mathcal{M}_i$	Extensive-hypotheses $\mathcal{L}_i$	Excluded
0			$\mathfrak{H}_0 : \{1, 2, 3\};$ $\mathfrak{H}_1 : \{1, 2, x_1^2\}, \{1, 3, x_1^3\}, \{2, 3, x_1^4\};$ $\mathfrak{H}_2 : \{1, x_1^5, x_2^5\}, \{2, x_1^6, x_2^6\}, \{3, x_1^7, x_2^7\};$ $\mathfrak{H}_3 : \{x_1^8, x_2^8, x_3^8\}$	
1	$\{1, 4, 8\}$	$\{1\},$ $\{4\},$ $\{8\}$	$\mathfrak{H}_0 : \{1, 2, 3\};$ $\mathfrak{H}_1 : \{1, 2, x_1^2\}, \{1, 3, x_1^3\}, \{2, 3, 4\},$ $\{2, 3, 8\};$ $\mathfrak{H}_2 : \{1, x_1^6, x_2^6\}, \{2, 4, x_1^7\}, \{2, 8, x_1^8\},$ $\{3, 4, x_1^9\}, \{3, 8, x_1^{10}\};$ $\mathfrak{H}_3 : \{4, x_1^{11}, x_2^{11}\}, \{8, x_1^{12}, x_2^{12}\}$	
2	$\{2, 5, 8\}$	$\{1, 2\},$ $\{1, 5\},$ $\{4, 2\},$ $\{4, 5\},$ $\{8\}$	$\mathfrak{H}_0 : \{1, 2, 3\};$ $\mathfrak{H}_1 : \{1, 2, x_1^2\}, \{1, 3, 5\}, \{1, 3, 8\},$ $\{2, 3, 4\}, \{2, 3, 8\};$ $\mathfrak{H}_2 : \{1, 5, x_1^7\}, \{1, 8, x_1^8\}, \{2, 4, x_1^9\},$ $\{2, 8, x_1^{10}\}, \{3, 4, 5\}, \{3, 8, x_1^{12}\};$ $\mathfrak{H}_3 : \{4, 5, x_1^{13}\}, \{8, x_1^{14}, x_2^{14}\}$	$\{3, 4, 8\},$ $\{4, 8, x\}$
3	$\{1, 6\}$	$\{1, 2\},$ $\{1, 5\},$ $\{4, 2, 6\},$ $\{4, 5, 6\},$ $\{8, 1\},$ $\{8, 6\}$	$\mathfrak{H}_0 : \{1, 2, 3\};$ $\mathfrak{H}_1 : \{1, 2, x_1^2\}, \{1, 3, 5\}, \{1, 3, 8\},$ $\mathfrak{H}_2 : \{1, 5, x_1^5\}, \{1, 8, x_1^6\}, \{2, 4, 6\},$ $\{2, 8, 6\}, \{3, 8, 6\};$ $\mathfrak{H}_3 : \{4, 5, 6\}, \{8, 6, x_2^{11}\}$	$\{2, 3, 4\},$ $\{2, 3, 8\},$ $\{3, 4, 5\}$
4	$\{3, 4, 7\}$	$\{1, 2, 3\},$ $\{1, 2, 4\},$ $\{1, 2, 7\},$ $\{1, 5, 3\},$ $\{1, 5, 4\},$ $\{1, 5, 7\},$ $\{4, 2, 6\},$ $\{4, 5, 6\}$ $\{8, 1, 3\},$ $\{8, 1, 4\},$ $\{8, 1, 7\},$ $\{8, 6, 1\},$ $\{8, 6, 3\},$ $\{8, 6, 7\}$	$\mathfrak{H}_0 : \{1, 2, 3\};$ $\mathfrak{H}_1 : \{1, 2, 4\}, \{1, 2, 7\}, \{1, 3, 5\},$ $\{1, 3, 8\},$ $\mathfrak{H}_2 : \{1, 5, 4\}, \{1, 5, 7\}, \{1, 8, 4\},$ $\{1, 8, 7\}, \{2, 4, 6\}, \{3, 8, 6\};$ $\mathfrak{H}_3 : \{4, 5, 6\}, \{8, 6, 4\}, \{8, 6, 7\}$	$\{2, 8, 6\}$

Table 5.2: Disproof of extensive-hypotheses with exceptions for Alice's set of friends  ${}_A\mathcal{H} = \{1, 2, 3\}$  and batch size  $b = 3$ . Minimal-hitting-sets is abbreviated by (MHS).

## 5. EXTENSION

---

and are therefore not enlisted in the last column.

Let  $k = |\mathcal{H} \cap R|$  be the number of specified recipients in an unspecified extensive-class  $\mathcal{H}$  and  $\mathcal{H} \cap \mathcal{O}_i = \emptyset$  as in Step 4. We determine the maximal number of excluded extensive-hypotheses in each exceptional case below.

Case 1: Let specifying a recipient in  $\mathcal{H}$  result in an extensive-class  $\mathcal{H}'$  that is not minimal in  $\mathcal{OS}_i$ .  $\mathcal{H}'$  contains  $(m - k - 1)$  unspecified recipients, each representing  $(b - 1)$  non-friends. Therefore, excluding  $\mathcal{H}'$  leads to an exclusion of at most  $(b - 1)^{m-k-1}$  distinct extensive-hypotheses.

Case 2: Let  $l = (b - 1) - |\mathcal{O}_i \setminus {}_A\mathcal{H}|$  be the number of non-friends missed in  $\mathcal{O}_i$ . As in Case 1, each non-friend missed in  $\mathcal{O}_i$  would determine an extensive-class that represents at most  $(b - 1)^{m-k-1}$  distinct extensive-hypotheses. Therefore at most  $l(b - 1)^{m-k-1}$  extensive-hypotheses are excluded in total.

We observe that most extensive-classes become specified very fast and logically at least as fast as minimal-hitting-sets reach the size  $m$ , so that  $k = |\mathcal{H} \cap R|$  is usually close to  $m$ . The impact of exclusions by exceptions on the number of the extensive-hypotheses is therefore moderate in comparison to normal (non-exceptional) exclusions. For the sake of simplicity, we only mathematically model the normal exclusion of extensive-hypotheses from the initial  $b^m$  extensive-hypotheses in this thesis.

### 5.1.3 Modelling Evolution of Extensive-Hypotheses

The last section shows that we can model the evolution of minimal-hitting-sets by the evolution of minimal extensive-classes. In this section, we simplify modelling the evolution of minimal extensive-classes by mathematically modelling the evolution of those specified extensive-hypotheses that would result from specifying the minimal extensive-classes. That is we consider all extensive-hypotheses as being specified initially, so that we solely model the evolution of specified extensive-hypotheses.

#### 5.1.3.1 Mean Number of Extensive-Hypotheses

In this section, we derive closed formulas for the mean number of specified extensive-hypotheses that remain in distinct classes after a given number of  $t$  observations. Since

---

all extensive-hypotheses are considered to be specified initially, the number of specified extensive-hypotheses for  $t = 0$  is  $b^m$  (i.e., the maximal number).

Let  $V_i$  be a random variable, where  $V_i = 1$  is the event that a particular specified extensive-hypothesis  $\mathcal{H} \in \mathfrak{H}_i$  remains valid after  $t$  observations, while  $V_i = 0$  denotes the inverse event. Note that a specified extensive-hypothesis is a hitting-set. To simplify the maths, we assume that Alice contacts her friends statistically independently and that the observations are also statistically independent. The probability of  $V_i = 1$  therefore corresponds to  $t$  statistically independent Bernoulli trials, where the outcome of each of the  $t$  trials shows that  $\mathcal{H}$  remains a hitting-set. Thereby a single Bernoulli trial corresponds to the outcome, whether  $\mathcal{H}$  remains a hitting-set at the next collected observation.

$$P(V_i = 1) = [P(\mathcal{H} \text{ remains hitting-set in next observation})]^t$$

The probability that a specified extensive-hypothesis  $\mathcal{H} \in \mathfrak{H}_i$  is excluded in the next observation is:

$$p_{exc}(u, b, m, i) = \frac{i}{m} \left(1 - \frac{m}{N}\right)^{b-1} . \quad (5.2)$$

The first factor  $\frac{i}{m}$  is the probability that Alice contacts any of her  $i = |\mathcal{H} \setminus \mathcal{H}_A|$  friends who are not in  $\mathcal{H}$ . The probability that a non Alice sender does not contact any of the recipients in  $\mathcal{H}$  is  $1 - \frac{m}{N}$ . Therefore, the second factor is the probability that the remaining  $(b - 1)$  senders of a batch who are other than Alice, do not contact any recipients in  $\mathcal{H}$ .

By formulating the probability  $P(V_i = 1)$  in terms of  $p_{exc}(u, b, m, i)$ , we obtain the following equation:

$$P(V_i = 1) = (1 - p_{exc}(u, b, m, i))^t .$$

Let  $\mathfrak{H}_i^* = \{\mathcal{H}_1, \dots, \mathcal{H}_{|\mathfrak{H}_i^*|}\}$  be the maximal set of all specified extensive-hypotheses in  $\mathfrak{H}_i$  and  $V_{i_j}$  be the event that  $\mathcal{H}_j \in \mathfrak{H}_i^*$  remains valid after  $t$  observations, for  $j = 1, \dots, |\mathfrak{H}_i^*|$ . The expectation  $E$  of the number of specified extensive-hypotheses in  $\mathfrak{H}_i^*$  after  $t$  observations is thus the expectation of the convolution of the random variables

## 5. EXTENSION

---

$V_{i_1}, \dots, V_{i_{|\mathfrak{H}_i^*|}}$ . This is described by the following equation:

$$E(V_{i_1}, \dots, V_{i_{|\mathfrak{H}_i^*|}}) = \sum_{j=1}^{|\mathfrak{H}_i^*|} E(V_{i_j}) \quad (5.3)$$

$$\begin{aligned} &= \sum_{j=1}^{|\mathfrak{H}_i^*|} P(V_{i_j} = 1) \\ &= |\mathfrak{H}_i^*| P(V_i = 1) . \end{aligned} \quad (5.4)$$

In (5.3), the additivity of the expectation function allows splitting the complex expectation on the left of ( $=$ ) to a sum of expectations of single  $V_{i_j}$  events on the right side of ( $=$ ). The probability of the outcome  $P(V_{i_j} = 1)$  is identical for every fixed specified extensive-hypothesis  $\mathcal{H}_j \in \mathfrak{H}_i^*$  (i.e.,  $P(V_{i_j} = 1) = P(V_i = 1)$ ), hence the right side of the former equation can be simplified to (5.4).

To clarify that (5.4) depends on the parameter  $u$ ,  $b$ ,  $m$  and  $t$  we denote by the function

$$E_{\mathfrak{H}_i}(u, b, m, t) = |\mathfrak{H}_i^*| (1 - p_{exc}(u, b, m, i))^t = \binom{m}{i} (b-1)^i \left(1 - \frac{i}{m} \left(1 - \frac{m}{u}\right)^{b-1}\right)^t , \quad (5.5)$$

the mean number of specified extensive-hypotheses that remain in  $\mathfrak{H}_i^*$  and thus in  $\mathfrak{H}_i$  after  $t$  observations, for given Mix parameters  $u, b, m$ .

Formula (5.3) can be easily extended to cover the mean number of specified extensive-hypotheses for any combination of classes including the consideration of all classes. The mean number of specified extensive-hypotheses in  $\mathfrak{H}$  (i.e., in all classes) that remain after  $t$  observations for given Mix parameters  $u, b, m$ , is

$$\begin{aligned} E_{\mathfrak{H}}(u, b, m, t) &= \sum_{i=0}^m \binom{m}{i} (b-1)^i \left(1 - \frac{i}{m} \left(1 - \frac{m}{u}\right)^{b-1}\right)^t \\ &\leq ((b-1)e^{-\frac{t}{m}(1-\frac{m}{u})^{b-1}} + 1)^m . \end{aligned} \quad (5.6)$$

**Number of Observations and Threshold of Hypotheses** The expectations  $E_{\mathfrak{H}_i}$  and  $E_{\mathfrak{H}}$  of the number of specified extensive-hypotheses after  $t$  observations can be easily

---

reformulated to derive the number of observations, such that  $\delta$  specified extensive-hypotheses remains on average.

By a transformation of (5.5), where  $\delta$  denotes the left side of the equation, we obtain

$$t_{\mathfrak{H}_i} = \frac{\ln \delta - \ln \binom{m}{i} - i \ln (b-1)}{\ln \left(1 - \frac{i}{m} \left(1 - \frac{m}{u}\right)^{b-1}\right)}, \text{ for } \delta > 0. \quad (5.7)$$

This equation represents the number of observations, such that at most  $\delta$  specified extensive-hypotheses remain on average in the class  $\mathfrak{H}_i$  for  $i \geq 1$ .

Similarly we reformulate (5.6) to obtain the number of observations  $t_{\mathfrak{H}}$ , such that there are on average fewer than  $\delta$  specified extensive-hypotheses in  $\mathfrak{H}$ . Alice's set of friends  ${}_A\mathcal{H}$  always remains in  $\mathfrak{H}$ , therefore  $\delta > 1$  in the expression below.

$$t_{\mathfrak{H}} \leq \frac{m(\ln(b-1) - \ln(\delta^{1/m} - 1))}{\left(1 - \frac{m}{N}\right)^{b-1}}, \text{ for } \delta > 1 \quad (5.8)$$

Note that modelling the quantity of the classes of specified extensive-hypotheses also models that of the classes of the minimal-hitting-set computed by ExactHS. Therefore the analyses in this section and the remaining sections also refer to the minimal-hitting-sets (and hypotheses) computed by ExactHS. We keep use the terms extensive-hypotheses and specified extensive-hypothesis, instead of the plain term hypothesis solely to remain precise and consistent in their usage.

### 5.1.3.2 Partial Disclosure

The *partial disclosure* is the unambiguous identification of a subset  ${}_A\mathcal{H}' \subseteq {}_A\mathcal{H}$  of Alice's set of friends. The full disclosure introduced in Chapter 2 and analysed in the previous chapters is a special case of the partial disclosure.

**Probability to Identify  $k$  Particular Recipients.** The probability to identify  $k$  particular friends  ${}_A\mathcal{H}' \subseteq {}_A\mathcal{H}$  after at most  $t$  observations is the probability that all hypotheses are disproved that do not contain all of these  $k = |{}_A\mathcal{H}'|$  friends after at most  $t$  observations. That probability specifies a discrete distribution with respect to  $t$  and we refer to it by the term  $f_{id}$ .

## 5. EXTENSION

---

To obtain  $f_{id}$ , we determine the number of all specified extensive-hypotheses in each class  $\mathfrak{H}_i$ , for  $i = 1, \dots, m$  that have to be disproved to identify  ${}_A\mathcal{H}'$ . According to (5.1),  $|\mathfrak{H}_i^*| = \binom{m}{i}(b-1)^i$ , so that  $\binom{m-k}{i}(b-1)^i$  is the number of specified extensive-hypotheses in  $\mathfrak{H}_i$  that contain all  $k$  recipients in  ${}_A\mathcal{H}'$ . The number of specified extensive-hypotheses in  $\mathfrak{H}_i$  that have to be excluded to enable the identification of  ${}_A\mathcal{H}'$  is therefore

$$exNo_i(b, m, k, i) = \left( \binom{m}{i} - \binom{m-k}{i} \right) (b-1)^i. \quad (5.9)$$

For the sake of simplicity, we assume for any distinct specified extensive-hypotheses  $\mathcal{H}_u, \mathcal{H}_v$ , that the probability of their disproof by a random observation is statistically independent. The probability to disprove a given specified extensive-hypothesis  $\mathcal{H} \in \mathfrak{H}_i$  is provided by  $p_{exc}(u, b, m, i)$  in (5.2). Applying this probability function to the single specified extensive-hypotheses that have to be disproved to identify  ${}_A\mathcal{H}'$  results in the function  $f_{id}$  with respect to the parameters  $u, b, m, t$  and  $k = |{}_A\mathcal{H}'|$ .

$$f_{id}(u, b, m, k, t) = \prod_{i=1}^{m-k} (1 - (1 - p_{exc}(u, b, m, i))^t)^{\left(\binom{m}{i} - \binom{m-k}{i}\right)(b-1)^i} \prod_{i=m-k+1}^m (1 - (1 - p_{exc}(u, b, m, i))^t)^{\binom{m}{i}(b-1)^i} \quad (5.10)$$

**Probability to Identify at Least  $k$  Recipients.** Based on the function  $f_{id}$  of the last section, we derive the probability distribution  $f_{id_{any}}$  that at least  $k$  of Alice's friends can be identified after at most  $t$  observations. In contrast to the previous section we are not focusing on identifying particular recipients, but on the probability to identify a certain number of recipients.

Let  $Y^k$  be a random variable denoting the event that particular  $k$  friends in Alice's set of friends  ${}_A\mathcal{H}$  are identified. That is  $Y^k = 1$  if the designated recipients are identified else  $Y^k = 0$  for the inverse case. To simplify the notation we will abbreviate the probability  $P(Y^k = 1)$  by the term  $P(Y^k)$ .

Let  $Y_1^k, \dots, Y_{\binom{m}{k}}^k$  be  $\binom{m}{k}$  distinct random variables. Each of this variable represents the event that distinct subsets of  ${}_A\mathcal{H}$  of cardinality  $k$  are identified. In order to compute the probability that at least  $k$  of Alice's friends can be identified, we have to determine

---

the probability that any of these  $Y_i^k$  events, for  $i \in \{1, \dots, \binom{m}{k}\}$  takes place. Thereby it would be imprecise to simply sum up the probabilities  $P(Y_i^k)$  for  $i \in \{1, \dots, \binom{m}{k}\}$ , because the events  $Y_i^k$  are not statistically independent. We can solve this problem by applying the inclusion-exclusion-formula below.

$$\begin{aligned}
P(Y_1^k \vee \dots \vee Y_{\binom{m}{k}}^k) &= P(Y_1^k) + \dots + P(Y_{\binom{m}{k}}^k) \\
&\quad - P(Y_1^k, Y_2^k) - \dots - P(Y_{\binom{m}{k}-1}^k, Y_{\binom{m}{k}}^k) \\
&\quad + \dots + \dots \\
&\quad - \dots - \dots \\
&\quad \vdots
\end{aligned} \tag{5.11}$$

Assume that  $\{a_{i_1}, a_{i_2}\}$  and  $\{a_{j_1}, a_{j_2}\}$  are those recipients who are identified by the event  $Y_i^k$  respectively  $Y_j^k$  (for  $k = 2$ ). The above expression  $P(Y_i^k, Y_j^k)$  is the probability that all recipients of the joint set  $\{a_{i_1}, a_{i_2}, a_{j_1}, a_{j_2}\}$  are identified<sup>1</sup>. Let us denote the joint event by the term  $Y^{k'}$ , where  $k' = |\{a_{i_1}, a_{i_2}, a_{j_1}, a_{j_2}\}| \leq 2k$ , then  $P(Y_i^k, Y_j^k) = P(Y^{k'})$  can be computed by (5.10). It is also obvious that this transformation can even be applied to an arbitrary number of joints of events, i.e., we can transform  $P(Y_1^k, \dots, Y_z^k)$  to  $P(Y^{k'})$  for any  $z \geq 1$  accordingly.

The next formula is an elaborate formulation of (5.11) for the special case of  $k = 1$ . It is the probability to identify at least one of Alice's friends after at most  $t$  observations.

$$f_{id_{any}}(u, b, m, t, 1) = \sum_{s=1}^m \left( (-1)^{s-1} \binom{m}{s} f_{id}(u, b, m, s, t) \right) \tag{5.12}$$

The general probability for arbitrary values of  $k$  is shown below. It is the probability to identify at least  $k \leq m$  of any Alice's friends after at most  $t$  observations.

$$f_{id_{any}}(u, b, m, t, k) = \sum_{i=1}^{\binom{m}{k}} (-1)^{i-1} \sum_{j_1=1}^{\binom{m}{k}-(i-1)} \dots \sum_{j_i=j_{i-1}+1}^{\binom{m}{k}-(i-i)} f_{id}(u, b, m, |\bigcup_{z=1}^i Y_{j_z}^k|, t) ,$$

---

<sup>1</sup>Note that the size of the joint set might be smaller than the sum of the single sets' sizes.

## 5. EXTENSION

---

where  $\bigcup_{z=1}^i Y_{j_z}^k$  is the union of the set of recipients identified by each  $Y_{j_z}^k$ .

Provided the distribution  $f_{id_{any}}(u, b, m, t, k)$ , the probability that at least  $k$  recipients can be identified after exactly  $t$  observation, is

$$p_{id_{any}}(u, b, m, t, k) = f_{id_{any}}(u, b, m, t, k) - f_{id_{any}}(u, b, m, t-1, k) .$$

**Mean Number of Observations for Partial Disclosure** We are now able to provide the formula for the mean number of observations to unambiguously identify at least  $k$  Alice's friends which we call MTDD-k. This mean is expressed by:

$$E_{id_{any}}(u, b, m, t, k) = \sum_{t=1}^{\infty} t p_{id_{any}}(u, b, m, t, k) . \quad (5.13)$$

Note that  $E_{id_{any}}(u, b, m, t, 1)$  particularly determines the mean number of observations to identify at least one of Alice's friends (MTDD-1). This provides a more refined measurement of the information theoretic limit of anonymity protection, based on the least number of observations to reveal the first unambiguous piece of information about Alice's friends as opposed to that considered in Chapter 2.

### 5.1.3.3 Beyond Unambiguous Information

Provided the model for the evolution of the specified extensive-hypotheses in the previous sections, we are also able to analytically analyse the attacker's information that is beyond the condition for unambiguous identification of friends.

As a demonstration, we estimate the probability that a random minimal-hitting-set computed by ExactHS contains a certain number of Alice's friends. This determines the probability to correctly guess any friend, as well as the closeness of the attacker's knowledge about Alice's set of friends, for a given number of observations.

Figure 5.1 plots the number of observations to reduce the number of minimal-hitting-sets in each class  $\mathcal{H}_i$ , for  $i = 1, \dots, m$ , to a value lower than  $\delta$  by using (5.7). The figure shows this for  $\delta = 1$  by the straight line (HS1) and for  $\delta = 0.1$  by the dashed line (HS0.1), for the Mix parameters  $u = 400$ ,  $b = 10$ ,  $m = 10$ .

We can see that the number of those minimal-hitting-sets that contain less of Alice's friends, decreases faster than the number of those minimal-hitting-sets that contain



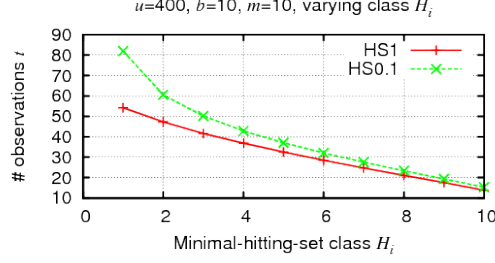


Figure 5.1: Number of observation to reduce number of minimal-hitting-sets in  $\mathfrak{H}_i$  below 1 (HS1) and 0.1 (HS0.1).

more of Alice's friends. Thus after a particular number of observations  $t$ , the number of hypotheses containing fewer than  $k$  Alice's friends are negligible, since  $E_{\mathfrak{H}_i}(t) < \delta$  for  $i > (m - k)$ . This is interesting, as the initial number of minimal-hitting-sets in a class  $\mathfrak{H}_i$  is higher than that in a class  $\mathfrak{H}_j$ , for  $i < j$ , as shown by (5.1).

Figure 5.1 also illustrates that after about  $t = 40$  observations, the attacker will unlikely find a minimal hitting set containing fewer than 7 of Alice's friends. Thus any minimal-hitting-set computed by ExactHS contains with a high probability at least 70% of Alice's friends. If we assume for simplicity that minimal-hitting-sets are excluded statistic independently from each other, then the probability to find at least  $k$  of  $m$  Alice's friends after at most  $t$  observations, is

$$f_{id_k}(N, b, m, k, t) \geq \prod_{i=m-k-1}^m (1 - (1 - p_{exc}(u, b, m, i))^t)^{|\mathfrak{H}_i^*|}.$$

#### 5.1.4 Evaluation

We compare the mathematical mean number of observations for partial disclosure of Alice's friends in Section 5.1.3 with that for the full disclosure of Alice's friends provided by our simulations in Chapter 2.2.4. The analytical analyses in this section are provided for the uniform communication model. Therefore we compare them with the simulation results from applying the HS-attack on observations that are randomly generated from the Alice's uniformly distributed traffic and the uniformly distributed cover-traffic.

## 5. EXTENSION

**Partial Disclosure** Figure 5.2 compares the expected number of observations to disclose at least one recipient (MTTD-1) by using  $E_{id_{any}}(u, b, m, t, 1)$  in (5.13) with the simulation result for the mean number of observations for full disclosure (HS) and the mean number of observation to reduce the set of specified extensive-hypotheses to a size below 2 (HS2) computed by (5.8). The comparison is with respect to different parameters  $u, b, m$ . We can see that the partial disclosure (MTTD-1) appears noticeable earlier than full disclosure (HS) and before the set of specified extensive-hypotheses is reduced to a size below 2. This difference increases, the more observations are required for full disclosure.

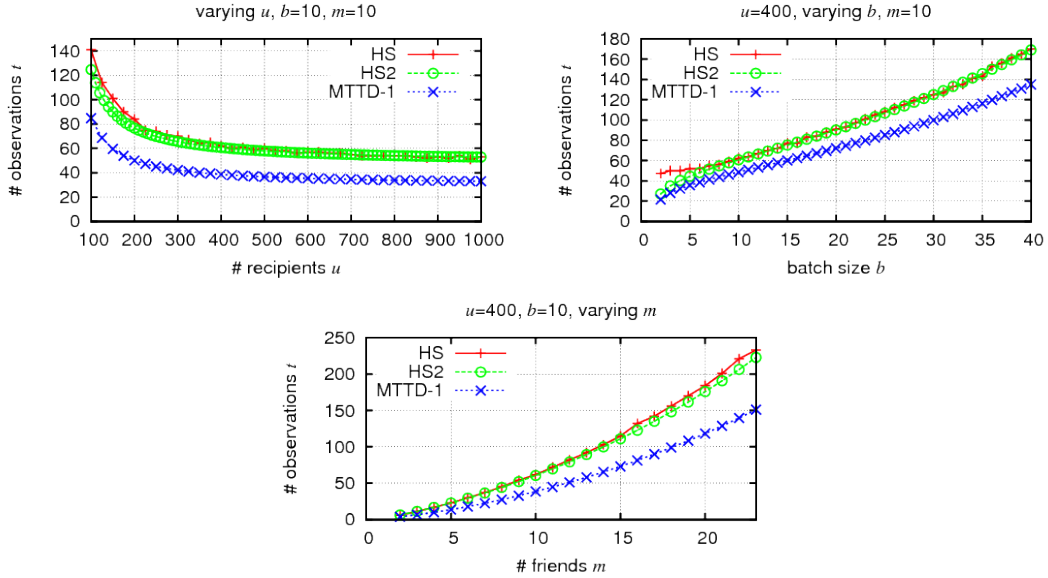


Figure 5.2: Number of observations for: full disclosure in simulation (HS), disclosure of at least one recipient (MTTD-1), reduction of number of minimal-hitting-sets below 2 (HS2).

## 5.2 Vague Information

All analyses so far assume that the attacker has full knowledge about the senders and recipients of the messages relayed in a Mix round. In contrast to that, this section assumes that the attack's information can be incomplete, or erroneous. We consider an adaptation of the HS-attack (using ExactHS) that tolerates random occasional er-

---

erroneous observations that miss Alice’s friends. Those erroneous observations could result from Alice occasionally contacting recipients other than her core friends<sup>1</sup>, or by observations of a non deterministic Mix flushing strategy, like the pool-Mix in Section 1.2.3. The goal of applying the HS-attack is to identify Alice’s core set of friends, despite erroneous observations. Our attack and analytical analyses apply to various Mix parameters and traffic distributions, as considered in Chapter 3.

Section 5.2.1 shows the application of ExactHS to uniquely identify Alice’s core set of friends despite unknown erroneous observations. However, due to the randomness of the number of erroneous observations, there is a certain probability that the result of ExactHS is biased. This bias is caused by an underestimate of the number of erroneous observations<sup>2</sup> when applying ExactHS. Therefore, Section 5.2.2 provides analytical analyses to identify Mix configurations and probability bounds of erroneous observations, such the probability of biases can be monotonically decreased towards 0 by increasing the number of observations.

## 5.2.1 Application of Hitting-Set Attack on Vague Observations

This section introduces a model of the erroneous observations that only assume the knowledge of the probability of erroneous observations in Section 5.2.1.1. That model can be used to model random deviances from normal user behaviour, or to model errors induced by indeterministic Mix flushing strategies.

Section 5.2.1.2 shows that it is possible to apply ExactHS on the set of observations and erroneous observations, such that the result of ExactHS can be evaluated to identify Alice’s set of friends.

### 5.2.1.1 Vague and Erroneous Observations

Let us denote by  $\_ \tilde{\mathcal{O}}$  an *erroneous observation*, that is a recipient set in a Mix round that lacks Alice’s friends, so that  $\_ \tilde{\mathcal{O}} \cap \_ \mathcal{H} = \emptyset$ . The set of all erroneous observations collected by the attacker is denoted by  $\_ \tilde{\mathcal{OS}}$ , that we call the *erroneous observation*

---

<sup>1</sup>We do not consider occasional contacts, or dummy traffic as friends of Alice and therefore do not refer to them by the term friends.

<sup>2</sup>Some friends might be missed, or non-friends might be misidentified as friends.

## 5. EXTENSION

---

set. We refer to the union set of ordinary observations<sup>1</sup>  $\mathcal{OS}$  and erroneous observations by the set  $\tilde{\mathcal{OS}} = \mathcal{OS} \cup \_ \tilde{\mathcal{OS}}$  that we call the *vague observation set*. The observations in the vague observation set are called *vague observations* and they are denoted by  $\tilde{\mathcal{O}}$ . Thus each  $\tilde{\mathcal{O}} \in \tilde{\mathcal{OS}}$  is either an ordinary observation, or an erroneous observation.

**Error Model** This section considers a simple probability model for the event of collecting an erroneous observation. We assume for every vague observation, that the probability that it is an erroneous observation is  $p_{er}$  and that this probability is statistically independent from the past observations. Let  $\tilde{t} = |\tilde{\mathcal{OS}}|$  be the number of vague observations collected by the attacker, the probability that there are  $\epsilon$  erroneous observations in  $\tilde{\mathcal{OS}}$  is binomially distribution with the probability mass function

$$P(|\_ \tilde{\mathcal{OS}}| = \epsilon) = \binom{\tilde{t}}{\epsilon} p_{er}^{\epsilon} p_{er}^{\tilde{t}-\epsilon} \quad (5.14)$$

and mean and variance

$$\begin{aligned} E(|\_ \tilde{\mathcal{OS}}|) &= \tilde{t} p_{er} , \\ Var(|\_ \tilde{\mathcal{OS}}|) &= \tilde{t} p_{er} (1 - p_{er}) . \end{aligned}$$

**Application of Error Model to Pool-Mix** A *pool-Mix* collects  $b$  messages in every round that are prepared similarly for relay like in the threshold Mix. In contrast to threshold Mixes, pool-Mixes always keep a pool of  $\eta$  messages. If a batch of messages arrives at the pool-Mix, then  $b$  out of the  $(b + \eta)$  messages are randomly chosen to be flushed, so that there is a probability greater 0 that Alice's message remains in the pool instead of being flushed. Pool-Mixes thus represent an indeterministic Mix concept.

The probability that Alice's message is flushed after exactly  $l$  rounds<sup>2</sup>, is

$$p_f(l) = \frac{b}{b + \eta} \left( \frac{\eta}{b + \eta} \right)^l , l \geq 0 . \quad (5.15)$$

An attacker might additionally decrease the probability of an erroneous observation as follow: Whenever Alice sends a message, the attacker can merge the pool-Mix

---

<sup>1</sup>That is an observation as defined in Section 2.1.2.1.

<sup>2</sup> $l = 0$  means that a message is flushed at the same round of its arrival.

---

output in that round with those of  $l$  consecutive rounds (where Alice is not active), to one vague observation. The probability that this vague observation is an erroneous observation is

$$p_{er} = 1 - \sum_{i=0}^l p_f(i) .$$

The merged  $(l + 1)$  outputs corresponds to a vague observation of size  $\tilde{b} = (l + 1)b$  for a value of  $l \geq 0$  chosen by the attacker. Thus the drawback of reducing  $p_{er}$  for a pool-Mix is the increase of the size of vague observations.

### 5.2.1.2 Applicability of ExactHS

We determine the value of  $\tilde{m}$ , such that the invocation of  $ExactHS(\tilde{\mathcal{O}}\mathcal{S}, \tilde{m}, \mathcal{C})$ ,  $\mathcal{C} = \{\}$  results in a set of all hitting-sets  $\mathcal{H}\mathcal{S}$ , where  $\bigcap_{\mathcal{H} \in \mathcal{H}\mathcal{S}} \mathcal{H} = {}_A\mathcal{H}$ , if  $\tilde{\mathcal{O}}\mathcal{S}$  is sufficiently large. We call this case the unique identification of  ${}_A\mathcal{H}$  despite erroneous observations, or simply the unique identification of Alice's friends by ExactHS.

In this section, we treat the number of erroneous observations as static during the HS-attack to simplify explanations. However, the analyses provided here also apply to the non static case, if the proportion of erroneous observations  $|\_ \tilde{\mathcal{O}}\mathcal{S}|$  in  $\tilde{\mathcal{O}}\mathcal{S}$  does not exceed some ratio during the HS-attack<sup>1</sup>, as we will prove in Section 5.2.2.

**Claim 11.** *Let  $\tilde{\mathcal{O}}\mathcal{S}$  be a vague observation set collected by the attacker. Given that the number of erroneous observations  $\epsilon = |\_ \tilde{\mathcal{O}}\mathcal{S}|$  in  $\tilde{\mathcal{O}}\mathcal{S}$  is fixed,  ${}_A\mathcal{H}$  becomes the subset of all hitting-sets of at most size  $\tilde{m} = (m + \epsilon)$ , if  $|\tilde{\mathcal{O}}\mathcal{S}|$  is sufficiently large. This is equivalent to  $\bigcap_{\mathcal{H} \in \mathcal{H}\mathcal{S}} \mathcal{H} \supseteq {}_A\mathcal{H}$ , where  $\mathcal{H}\mathcal{S}$  contains all hitting-sets computed by invoking  $ExactHS(\tilde{\mathcal{O}}\mathcal{S}, \tilde{m}, \mathcal{C})$ ,  $\mathcal{C} = \{\}$ .*

**Claim 12.** *Let  $\tilde{\mathcal{O}}\mathcal{S}$  be a vague observation set collected by the attacker and  $\epsilon = |\_ \tilde{\mathcal{O}}\mathcal{S}|$  be the number of erroneous observations in it. Provided that no erroneous observation in  $\_ \tilde{\mathcal{O}}\mathcal{S}$  is a singleton<sup>2</sup>, the intersection of all hitting-set of at most size  $\tilde{m} = (m + \epsilon)$  in  $\tilde{\mathcal{O}}\mathcal{S}$  does not contain any recipient  $r \notin {}_A\mathcal{H}$ . This is equivalent to  $\bigcap_{\mathcal{H} \in \mathcal{H}\mathcal{S}} \mathcal{H} \subseteq {}_A\mathcal{H}$ , where  $\mathcal{H}\mathcal{S}$  contains all hitting-sets computed by invoking  $ExactHS(\tilde{\mathcal{O}}\mathcal{S}, \tilde{m}, \mathcal{C})$ ,  $\mathcal{C} = \{\}$ .*

---

<sup>1</sup>The HS-attack repeats adding new vague observations to  $\tilde{\mathcal{O}}\mathcal{S}$  and applying ExactHS on  $\tilde{\mathcal{O}}\mathcal{S}$ , until it succeeds.

<sup>2</sup>We exclude this case, since it is pathological.

## 5. EXTENSION

---

Claim 11 and Claim 12 show, that if the number of observations in  $\tilde{\mathcal{OS}} = \mathcal{OS} \cup \_ \tilde{\mathcal{OS}}$  can be sufficiently increased, while the number of erroneous observations remains static, then  $\bigcap_{\mathcal{H} \in \mathcal{HS}} \mathcal{H}$  uniquely identifies Alice's set of friends for a sufficient large  $\tilde{\mathcal{OS}}$ . This is for  $\mathcal{HS}$  computed by  $ExactHS(\tilde{\mathcal{OS}}, \tilde{m}, \mathcal{C})$ ,  $\mathcal{C} = \{\}$  and  $\tilde{m} = m + |\_ \tilde{\mathcal{OS}}|$ , where no erroneous observation in  $\_ \tilde{\mathcal{OS}}$  is a singleton.

In the case that  $|\_ \tilde{\mathcal{OS}}|$  is unknown, or non static, it can be estimated by (5.14). Overestimating  $|\_ \tilde{\mathcal{OS}}|$  and  $\tilde{m}$  would just increase the number of vague observations required for the identification of  ${}_A \mathcal{H}$ . Given a sufficiently large  $\tilde{\mathcal{OS}}$ , if  $|\_ \tilde{\mathcal{OS}}|$  and  $\tilde{m}$  is underestimated, then there will be no hitting-set of at most size  $\tilde{m}'$  for some  $m \leq \tilde{m}' < \tilde{m}$  in  $\tilde{\mathcal{OS}}$ , while the result of invoking  $ExactHS(\tilde{\mathcal{OS}}, \tilde{m}'', \mathcal{C})$ ,  $\mathcal{C} = \{\}$ , for  $\tilde{m}' < \tilde{m}'' < \tilde{m}$  will be arbitrary.

*Proof of Claim 11.* Let  $\mathcal{H}'$  be a hitting-set in  $\tilde{\mathcal{OS}}$ , then  $\mathcal{H}'$  must be a hitting set in  $\mathcal{OS}$  and in  $\_ \tilde{\mathcal{OS}}$ .<sup>1</sup> Let the number of erroneous observations  $\epsilon$  be fixed in  $\tilde{\mathcal{OS}}$ . Provided this, We show that all hitting-set of at most size  $\tilde{m} = m + \epsilon$  in  $\tilde{\mathcal{OS}}$  will contain  ${}_A \mathcal{H}$  with almost certainty, if the number of ordinary observations in  $\tilde{\mathcal{OS}}$  can be increased. Let  $\mathcal{H}' \not\supseteq {}_A \mathcal{H}$  be a hitting-set of at most size  $\tilde{m}$  in  $\tilde{\mathcal{OS}}$ , then there is a probability  $q < 1$  that  $\mathcal{H}'$  hits a random ordinary observation, as considered in Section 2.1.2.4. The probability that  $\mathcal{H}'$  remains a hitting-set in  $\mathcal{OS}$  (and thus also in  $\tilde{\mathcal{OS}}$ ) after collecting  $x$  additional observations is  $q^x$  and is negligible even for moderate value of  $x$ .

Consequently, by collecting sufficiently many vague observations when invoking  $ExactHS(\tilde{\mathcal{OS}}, \tilde{m}, \mathcal{C})$ ,  $\mathcal{C} = \{\}$ , we obtain  $\bigcap_{\mathcal{H} \in \mathcal{HS}} \mathcal{H} \supseteq {}_A \mathcal{H}$  with almost certainty, provided that  $\epsilon$  is fixed.  $\square$

*Proof of Claim 12.* It is sufficient to prove that the intersection of all hitting-sets  $\mathcal{H}$  of at most size  $\tilde{m} = (m + \epsilon)$ , where  $\mathcal{H} \supseteq {}_A \mathcal{H}$  cannot contain a recipient  $r \notin {}_A \mathcal{H}$ . This equivalently proves  $\bigcap_{\mathcal{H} \in \mathcal{HS}: \mathcal{H} \supseteq {}_A \mathcal{H}} \mathcal{H} \subseteq {}_A \mathcal{H}$ , where  $\mathcal{HS}$  result from invoking  $ExactHS(\tilde{\mathcal{OS}}, \tilde{m}, \mathcal{C})$ .

Using Claim 2 in Section 2.2.2.1, a set  $\mathcal{H}$  is a hitting-set in  $\tilde{\mathcal{OS}}$  if and only if  $\mathcal{H}_N = \mathcal{H} \setminus {}_A \mathcal{H}$  is a hitting-set in  $\tilde{\mathcal{OS}} \setminus \tilde{\mathcal{OS}}[{}_A \mathcal{H}] = \_ \tilde{\mathcal{OS}}$ . The latter equality holds, since erroneous observations do not contain any Alice's friends by definition. Therefore, let  $\epsilon = |\_ \tilde{\mathcal{OS}}|$  and  $\_ \tilde{\mathcal{OS}} = \{\_ \tilde{\mathcal{O}}_1, \dots, \_ \tilde{\mathcal{O}}_\epsilon\}$ , then every set  $\mathcal{H}_N = \{r_1, \dots, r_\epsilon\}$ , for  $r_i \in \_ \tilde{\mathcal{O}}_i, i = 1, \dots, \epsilon$  is a hitting-set of at most size  $\epsilon$  in  $\_ \tilde{\mathcal{OS}}$  and every  $\mathcal{H} = {}_A \mathcal{H} \cup \mathcal{H}_N$

---

<sup>1</sup>Note that if new observations are collected by the attacker, then  $\tilde{\mathcal{OS}}$  is extended by them.

is a hitting-set of at most size  $\tilde{m}$  in  $\tilde{\mathcal{OS}}$ . In the non-pathological case, every erroneous observation contains at least two recipients, therefore for every recipient  $r_i \in \mathcal{H}_N$ , there is a recipient  $r'_i \neq r_i$ , where  $r_i, r'_i \in \tilde{\mathcal{O}}_i$ , so that for every  $i = 1, \dots, \epsilon$ ,  $\mathcal{H}'_N = \{r_1, \dots, r_{i-1}, r'_i, r_{i+1}, \dots, r_\epsilon\}$  is another hitting-set in  $\tilde{\mathcal{OS}}$ . Therefore, the intersection of all hitting-sets  $\mathcal{H}'_N$  in  $\tilde{\mathcal{OS}}$ , for  $i = 1, \dots, \epsilon$ , results in an empty set. Consequently, the intersection of all hitting-sets  $\mathcal{H} = {}_A\mathcal{H} \cup \mathcal{H}'_N$  in  $\tilde{\mathcal{OS}}$ , for  $i = 1, \dots, \epsilon$ , is  ${}_A\mathcal{H}$ . This implies that  $\bigcap_{\mathcal{H} \in \mathcal{HS}} \mathcal{H} \subseteq {}_A\mathcal{H}$ , where  $\mathcal{HS}$  contains all hitting-sets computed by invoking  $ExactHS(\tilde{\mathcal{OS}}, \tilde{m}, \mathcal{C})$ ,  $\mathcal{C} = \{\}$ .

We show that  $\tilde{m} = (m + |\tilde{\mathcal{OS}}|)$  is even the least value for  $ExactHS(\tilde{\mathcal{OS}}, \tilde{m}, \mathcal{C})$ ,  $\mathcal{C} = \{\}$ , such that  $\bigcap_{\mathcal{H} \in \mathcal{HS}} \mathcal{H} \subseteq {}_A\mathcal{H}$  for every vague observation set, by constructing a vague observation set  $\tilde{\mathcal{OS}} = \mathcal{OS} \cup \tilde{\mathcal{OS}}$ , where all hitting-sets are of at least size  $\tilde{m}$ . Let every erroneous observation  $\tilde{\mathcal{O}} \in \tilde{\mathcal{OS}}$  be disjoint to any other observation  $\tilde{\mathcal{O}}' \in \tilde{\mathcal{OS}}$ , that is  $\tilde{\mathcal{O}} \cap \tilde{\mathcal{O}}' = \emptyset$  for  $\tilde{\mathcal{O}} \neq \tilde{\mathcal{O}}'$ , then the size of the smallest minimal-hitting-sets in  $\tilde{\mathcal{OS}}$  is  $|\tilde{\mathcal{OS}}|$ . Let  ${}_A\mathcal{H}$  be the unique minimum-hitting-set in  $\mathcal{OS}$ , then  $\tilde{m} = (m + |\tilde{\mathcal{OS}}|)$  is the size of the smallest minimal-hitting-sets in  $\tilde{\mathcal{OS}}$  and  $\bigcap_{\mathcal{H} \in \mathcal{HS}} \mathcal{H} = {}_A\mathcal{H}$ . This is illustrated in Example 5.2.1.  $\square$

**Example 5.2.1** (Unique Identification Despite Errors). *This example shows a vague observations set  $\tilde{\mathcal{OS}}$  with  $\epsilon = |\tilde{\mathcal{OS}}|$  erroneous observations, where the size of the smallest minimal-hitting-sets in  $\tilde{\mathcal{OS}}$  is  $\tilde{m} = (m + \epsilon)$  and  $\bigcap_{\mathcal{H} \in \mathcal{HS}} \mathcal{H} = {}_A\mathcal{H}$ , where  $\mathcal{HS}$  results from invoking  $ExactHS(\tilde{\mathcal{OS}}, \tilde{m}, \mathcal{C})$ ,  $\mathcal{C} = \{\}$ .*

*We consider a batch size of  $b = 2$ , Alice's set of friends  ${}_A\mathcal{H} = \{1, 2\}$  and an observations set  $\tilde{\mathcal{OS}} = \mathcal{OS} \cup \tilde{\mathcal{OS}}$  that contains  $\epsilon = 2$  erroneous observations that are illustrated in Figure 5.3.*

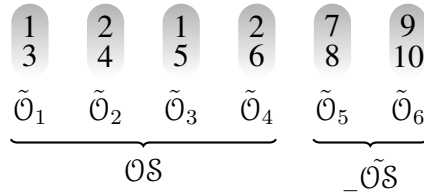


Figure 5.3: Unique identification despite erroneous observations.

*In Figure 5.3, Alice's set of friends  $\{1, 2\}$  is the unique minimum-hitting-set in  $\mathcal{OS}$ . Every erroneous observation  $\tilde{\mathcal{O}}_5, \tilde{\mathcal{O}}_6$  is disjoint to all other vague observation in  $\tilde{\mathcal{OS}}$ .*

## 5. EXTENSION

---

Therefore, every set  $\{r_5, r_6\}$ , where  $r_5 \in \tilde{\mathcal{O}}_5$  and  $r_6 \in \tilde{\mathcal{O}}_6$  is a smallest minimal-hitting-set in  $_{\tilde{\mathcal{O}}\mathcal{S}}$  and every  $\{1, 2\} \cup \{r_5, r_6\}$  is a smallest minimal-hitting-set in  $\tilde{\mathcal{O}}\mathcal{S}$ . These sets are enlisted in

$$\mathcal{HS} = \{\{1, 2, 7, 9\}, \{1, 2, 7, 10\}, \{1, 2, 8, 9\}, \{1, 2, 8, 10\}\} .$$

The size of all smallest minimal-hitting-sets in  $\tilde{\mathcal{O}}\mathcal{S}$  are  $\tilde{m} = (m + \epsilon)$ , which is  $2 + 2$  and  $\bigcap_{\mathcal{H} \in \mathcal{HS}} \mathcal{H} = \{1, 2\}$ , where  $\mathcal{HS}$  results from invoking  $ExactHS(\tilde{\mathcal{O}}\mathcal{S}, 4, \mathcal{C})$ ,  $\mathcal{C} = \{\}$ . This uniquely identifies Alice's set of friends, provided  $\epsilon$  is known.

### 5.2.2 Analytical Analyses of Conditions for Unique Identification

This section considers the HS-attack for the case that the exact number of erroneous observations is unknown, but can be estimated by the attacker's knowledge of the value of  $p_{er}$ . We analytically determine Mix configurations and ranges of probabilities  $p_{er}$ , such that Alice's set of friends can be uniquely identified by the HS-attack with a probability that increases monotonically with the number of vague observations  $|\tilde{\mathcal{O}}\mathcal{S}|$ . This probability approaches 1 for  $|\tilde{\mathcal{O}}\mathcal{S}| \rightarrow \infty$ .

Section 5.2.2.1 identifies properties of ordinary observations, such that Alice's sets of friends can be uniquely identified, despite a given maximal number of erroneous observations.

Section 5.2.2.2 relates those properties of ordinary observations to the probability of erroneous observations to mathematically derive Mix configurations and error probabilities, such that Alice's set of friends can be uniquely identified by  $ExactHS$  with a probability that monotonically increases with respect to  $|\tilde{\mathcal{O}}\mathcal{S}|$ .

#### 5.2.2.1 Properties of Ordinary Observations for Unique Identification

We provide a "sufficient condition" for ordinary observations, that if fulfilled allows the unique identification of all Alice's friends by  $ExactHS$  despite a bounded number of erroneous observations. It is aimed to aid analytical analyses and identification of Mix configurations, where  $ExactHS$  can uniquely identify Alice's friends.

However, this sufficient condition is too complex for mathematical analyses. There-



---

fore, we introduce in Claim 13 the necessary condition of that sufficient condition. The condition provided by Claim 13 will be analysed instead of the sufficient condition to identify Mix configurations that allow unique identifications by ExactHS in the remaining of Section 5.2.2.2.

**Sufficient Condition** Let  $\tilde{\mathcal{OS}} = \mathcal{OS} \cup \_ \tilde{\mathcal{OS}}$  be the vague observations collected by the attacker and  $\epsilon = |\_ \tilde{\mathcal{OS}}|$ . Observe that every hitting-set in  $\tilde{\mathcal{OS}}$  must be a hitting-set in  $\mathcal{OS}$ , which particularly holds for hitting-sets of at most size  $\tilde{m} = m + \epsilon$ . Therefore, if all (minimal-)hitting-sets<sup>1</sup> of at most size  $\tilde{m}$  in the ordinary observation set  $\mathcal{OS}$  contain  ${}_A\mathcal{H}$ , then this must be the case for all hitting-sets of at most size  $\tilde{m}$  in  $\tilde{\mathcal{OS}}$ . This condition is thus a *sufficient condition for the unique identification of Alice's set of friends*  ${}_A\mathcal{H}$ .

It is not a necessary condition, as the identifiability of  ${}_A\mathcal{H}$  in  $\tilde{\mathcal{OS}}$  does not require every (minimal-) hitting-set of at most size  $\tilde{m}$  in  $\mathcal{OS}$  to be a superset of  ${}_A\mathcal{H}$ , as demonstrated by Example 5.2.1. That example shows the unique identification of  ${}_A\mathcal{H} = \{1, 2\}$  in a set  $\tilde{\mathcal{OS}}$ , where  $m = 2$ ,  $\epsilon = 2$ ,  $\tilde{m} = 4$ , whereas the minimal-hitting-sets of at most size  $\tilde{m}$  in  $\mathcal{OS}$  are  $\{1, 2\}, \{1, 4, 6\}, \{2, 3, 5\}, \{3, 4, 5, 6\}$ .

**Relaxation of Sufficient Condition** The above sufficient condition is hard to analyse, as it is equivalent to the unique-minimum-hitting-set problem for the special case of no erroneous observations, that is  $\epsilon = 0$ .

We therefore consider a necessary condition of that sufficient condition instead, that is the  $(2 + \epsilon) \times$ -exclusivity of all Alice's friends in Claim 13. Note that  $(2 + \epsilon) \times$ -exclusivity closely models the unique identification of Alice's set of friends for the case of no erroneous observations  $\epsilon = 0$ , cf. Kesdogan et al. [2006].

**Claim 13.** *If there is no (minimal-) hitting-set  $\mathcal{H}$  of at most size  $\tilde{m}' = (m + \epsilon')$  in  $\mathcal{OS}$  (for integers  $\epsilon' \geq 0$ ), where  $\mathcal{H} \not\supseteq {}_A\mathcal{H}$ , then every Alice's friend must be  $(2 + \epsilon') \times$ -exclusive in  $\mathcal{OS}$ .*

*Proof of Claim 13.* The special case of  $\epsilon' = 0$  is known as  $2 \times$ -exclusivity and was already proven by Kesdogan et al. [2006]. We provide a straight forward proof for integers  $\epsilon' > 0$  below, which is a proof by contradiction.

---

<sup>1</sup>Recall that every hitting-set is a superset of a minimal-hitting-set, so that it is sufficient to consider minimal-hitting-sets.

## 5. EXTENSION

Assume that there is no hitting-set  $\mathcal{H} \not\supseteq {}_A\mathcal{H}$  of at most size  $\tilde{m}' = (m + \epsilon')$  in  $\mathcal{OS}$ , although there is an Alice's friend  $a \in {}_A\mathcal{H}$  that is not at least  $(2 + \epsilon')\times$ -exclusive. Note that  ${}_A\mathcal{H} = \{a_1, \dots, a_m\}$  is a hitting-set of size  $m$  in  $\mathcal{OS}$ . Let without loss of generality  $a = a_1$ , then  $\{a_2, \dots, a_m\}$  hits all observations in  $\mathcal{OS}$  except those observations  $\mathcal{O}_1, \dots, \mathcal{O}_l$  that contains  $a$  exclusively, where  $l \leq (1 + \epsilon')$ . Therefore,  $\mathcal{H}' = \{a_2, \dots, a_m\} \cup \{n_1, \dots, n_l\}$ , for  $n_i \in \mathcal{O}_i \setminus {}_A\mathcal{H}$ ,  $i = 1, \dots, l$  would be a hitting-set of size  $(m - 1 + l) \leq \tilde{m}'$  in  $\mathcal{OS}$ , where  $\mathcal{H}' \not\supseteq {}_A\mathcal{H}$ . This is a contradiction to the initial assumption and thus proves that  $(2 + \epsilon')\times$ -exclusivity of all Alice's friends is a necessary condition, so that every hitting-set of at most size  $\tilde{m}'$  in  $\mathcal{OS}$  is a superset of  ${}_A\mathcal{H}$ .  $\square$

### 5.2.2.2 Probability Bound of Erroneous Observations for Unique Identification

We provide a formula based on Claim 13 that estimates for every given Mix configuration, pairs  $(t, \epsilon')$  of mean number of ordinary observations and maximal number of acceptable erroneous observations such that  ${}_A\mathcal{H}$  can be identified by ExactHS.

This formula is related to the probability of erroneous observations  $p_{er}$  to determine the bound of  $p_{er}$ , such that the probability of at most  $\epsilon'$  erroneous observations monotonically increases with respect to  $t$  and the number of vague observations  $\tilde{t} = t + \epsilon'$ .

#### Mean Number of Ordinary Observations and Number of Acceptable Errors

Based on Claim 13, we provide an estimate (5.16) of the mean number of ordinary observations  $E(T_{(2+\epsilon')\times e})$ , such that all Alice's friends are  $(2 + \epsilon')\times$ -exclusive. It estimates for every given Mix configuration with parameters  $(u, b, m)$  and the least probability  $p = \min_{a \in {}_A\mathcal{H}} P_A(a)$  of Alice's communication, pairs  $(t, \epsilon')$  of mean number of ordinary observations and maximal number of erroneous observations, such that  ${}_A\mathcal{H}$  can be identified by ExactHS. This estimate is due to Claim 5 and is as follow:

$$E(T_{(2+\epsilon')\times e}) \approx \underbrace{\left( \frac{1}{p}(\ln m + \gamma) + \frac{1}{p} \ln \ln m \right) \left( \frac{u - (m - 1)}{u} \right)^{1-b}}_{t_2} + \epsilon' \underbrace{\left( \frac{1}{p} \ln \ln m \right) \left( \frac{u - (m - 1)}{u} \right)^{1-b}}_{t_1}, \quad (5.16)$$

---

where  $t_2$  and  $t_1$  are constants and  $t = t_2 + \epsilon' t_1$  is the estimate of  $E(T_{(2+\epsilon') \times e})$ , so that (5.16) determines pairs of  $(t, \epsilon')$ . Whenever we refer to  $t$  or  $\epsilon'$  in this section, we always mean  $t$  or  $\epsilon'$  determined by (5.16), so that a large value of  $t$  implies a corresponding large value of  $\epsilon'$ .

Note that  $t_2$  estimates the mean number of observations to uniquely identify Alice's set of friends, provided no erroneous observations, in (5.16). Thus, each erroneous observation would require  $t_1$  additional observations for the unique identification by the HS-attack. It can be seen that the value of  $t_1$  is just a logarithmic fraction of the value of  $t_2$ .

**Probability of Erroneous Observations** The mean number of ordinary observations that is needed to identify Alice's friends, if at most  $\epsilon'$  of the attacker's vague observations are erroneous is estimated by the number  $t$  of ordinary observations for  $(2 + \epsilon')$ -exclusivity. Provided such a  $t$  and  $\epsilon'$ , we determine the probability of at most  $\epsilon'$  erroneous observations with respect to the number of vague observations  $\tilde{t} = t + \epsilon'$ . This is deployed to derive the bound  $p_{er}$  of the probability of erroneous observations in (5.17). It shows that increasing the number  $\tilde{t}$  of vague observations monotonically increases the probability of identifying Alice's set of friends by ExactHS, if the probability of erroneous observations is below  $p_{er}$ .

**Claim 14.** *Let  $\tilde{\mathcal{OS}} = \mathcal{OS} \cup \_ \tilde{\mathcal{OS}}$  denote a set of vague observations and  $\epsilon = |\_ \tilde{\mathcal{OS}}|$  the number of erroneous observations. Let  $p_{er}$  be the probability that a random vague observation is an erroneous observation.*

*The sufficient and necessary condition, such that all Alice's friends are  $(2 + \epsilon) \times$ -exclusive in  $\mathcal{OS}$  with a monotonically increasing probability approaching 1 for  $|\tilde{\mathcal{OS}}| \rightarrow \infty$ , is*

$$p_{er} < \frac{1}{\left(\frac{1}{p} \ln \ln m\right) \left(\frac{u-(m-1)}{u}\right)^{1-b} + 1}, \quad (5.17)$$

where  $p = \min_{a \in \_A \mathcal{H}} P_A(a)$ .

*Proof of Claim 14.* Let  $\tilde{t}$  denote the number of vague observations, such that all Alice's

## 5. EXTENSION

---

friends are  $(2 + \epsilon') \times$ -exclusive, while there are  $\epsilon'$  erroneous observations. Therefore,

$$\begin{aligned}\tilde{t} &= t + \epsilon' \\ &= t_2 + \epsilon'(t_1 + 1)\end{aligned}$$

and whenever we refer to  $\tilde{t}$  in this proof, we always mean a  $\tilde{t}$  that result from this equation.

The probability (5.14) that there are  $\epsilon'$  erroneous observations, given  $\tilde{t}$  vague observations and the error probability  $p_{er}$ , specifies a binomial distribution. For large values of  $\tilde{t}$ , that distribution can be closely approximated by the normal distribution  $\mathcal{N}(\mu, \sigma^2)$  with the probability density and cumulative distribution function

$$\begin{aligned}f(\epsilon') &= \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{1}{2}\left(\frac{\epsilon' - \mu}{\sigma}\right)^2} \\ F(\epsilon') &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{\epsilon' - \mu}{\sigma}} e^{-\frac{x^2}{2}} dx, \end{aligned} \quad (5.18)$$

where  $\mu = \tilde{t}p_{er}$  and  $\sigma^2 = \tilde{t}p_{er}(1 - p_{er})$  are the same mean and variance as in (5.14).

We derive the bound for  $p_{er}$ , such that increasing  $\tilde{t} = t + \epsilon'$  increases the probability  $F(\epsilon')$  that there are at most  $\epsilon'$  erroneous observations<sup>1</sup> with  $F(\epsilon') \rightarrow 1$  for  $\tilde{t} \rightarrow \infty$ , while all Alice's friends are  $(2 + \epsilon') \times$ -exclusive.

Observe that the value of  $F(\epsilon')$  is monotonically increasing and approaches 1 if the upper limit of the integral in (5.18),

$$\begin{aligned}\frac{\epsilon' - \mu}{\sigma} &= \frac{\frac{\tilde{t} - t_2}{t_1 + 1} - \tilde{t}p_{er}}{\sqrt{\tilde{t}}\sqrt{p_{er}(1 - p_{er})}} \\ &= \frac{\sqrt{\tilde{t}}\left(\frac{1}{t_1 + 1} - p_{er}\right)}{\sqrt{p_{er}(1 - p_{er})}} - \frac{\frac{t_2}{t_1 + 1}}{\sqrt{\tilde{t}}\sqrt{p_{er}(1 - p_{er})}}\end{aligned} \quad (5.19)$$

is monotonically increasing and approaches infinity for  $\tilde{t} \rightarrow \infty$ . This is only the case, if the factor  $\left(\frac{1}{t_1 + 1} - p_{er}\right)$  in the numerator of the expression left of  $(-)$  is positive. That

---

<sup>1</sup>Recall that  $\tilde{t}$  and  $\epsilon'$  are interlaced, so that increasing  $\tilde{t}$  increases  $\epsilon'$  accordingly and vice versa.

---

is if the following inequality is fulfilled:

$$\begin{aligned}
p_{er} &< \frac{1}{t_1 + 1} \\
&= \frac{1}{\left(\frac{1}{p} \ln \ln m\right) \left(\frac{u-(m-1)}{u}\right)^{1-b} + 1} \quad (t_1 \text{ replaced as in (5.16)}) \quad . \quad (5.17)
\end{aligned}$$

Inequality (5.17) determines the bound for  $p_{er}$ , such that  $F(\epsilon')$  is monotonically increasing with respect to  $\tilde{t}$ , such that  $F(\epsilon') \rightarrow 1$  for  $\tilde{t} \rightarrow \infty$ .

This inequality is thus a sufficient condition, such that the probability that all Alice's friends are  $(2 + \epsilon) \times$ -exclusive is monotonically increasing and approaches 1 for  $|\tilde{\mathcal{O}}\mathcal{S}| \rightarrow \infty$ , where  $\tilde{\mathcal{O}}\mathcal{S} = \mathcal{O}\mathcal{S} \cup \_ \tilde{\mathcal{O}}\mathcal{S}$ ,  $\epsilon = |\_ \tilde{\mathcal{O}}\mathcal{S}|$ . It is even a necessary condition for that, because  $p_{er} = \frac{1}{t_1+1}$  would imply  $\frac{\epsilon' - \mu}{\sigma} \rightarrow 0$  and thus  $F(\epsilon') \rightarrow \frac{1}{2}$ , for  $\tilde{t} \rightarrow \infty$ .  $\square$

### 5.2.3 Evaluation

We mathematically evaluate the limit of the probability that a random vague observation is an erroneous observation, such that Alice's set of friends can be uniquely identified by the HS-attack in Section 5.2.1.2 with a high probability. This limit is determined by (5.17). Provided that the real error probability is below that limit, the attacker can control the probability to succeed the HS-attack estimated by (5.18), by increasing the number of vague observations.

The limit of the probability of erroneous observations is evaluated with respect to the same Mix parameters  $(u, b, m)$  and uniformly distributed cover-traffic and Zipf distributed Alice's traffic, as considered in Section 2.2.4. That is the recipients of the cover-traffic are assumed to be uniformly distributed in the set of all recipients  $R$ . Alice's is assumed to contact her friends in  ${}_A\mathcal{H}$  according to the Zipf( $m, \alpha$ ) distribution.

#### Limit of Probability of Erroneous Observations for Unique Identification

Figure 5.4 and Figure 5.5 draw on the y-axis the limit of the probability of erroneous observations that can be tolerated by ExactHS. We provide this by using (5.17), for various Mix parameters  $u, b, m$  and Zipf weight  $\alpha$  of Alice's traffic distribution.

Comparing the graphs in Figure 5.4 and Figure 5.5 with that of (Figure 2.6, Figure 2.8) and (Figure 2.7, Figure 2.9) in Section 2.2.4.2 reveals the tendency that chang-

## 5. EXTENSION

ing parameters that strongly increases the number of observations required by the HS-attack strongly decreases the probability of erroneous observations tolerable to the HS-attack. Depending on the Mix configurations, the tolerable erroneous observations could cover 0.1% – 28% of the vague observations collect by the attacker, as illustrated in Figure 5.4 and Figure 5.5.

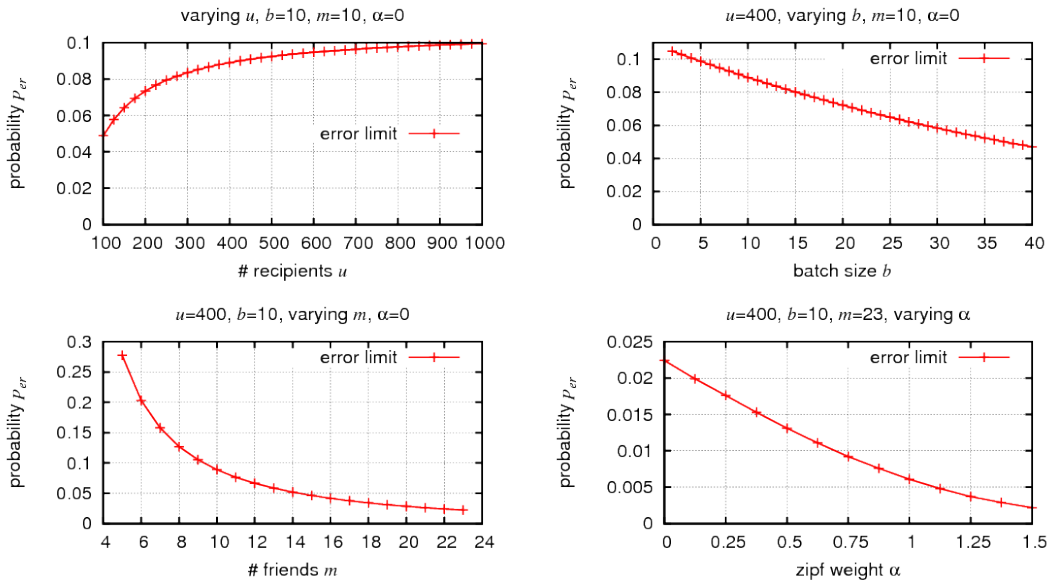


Figure 5.4: Limit of probability of erroneous observations for full disclosure.

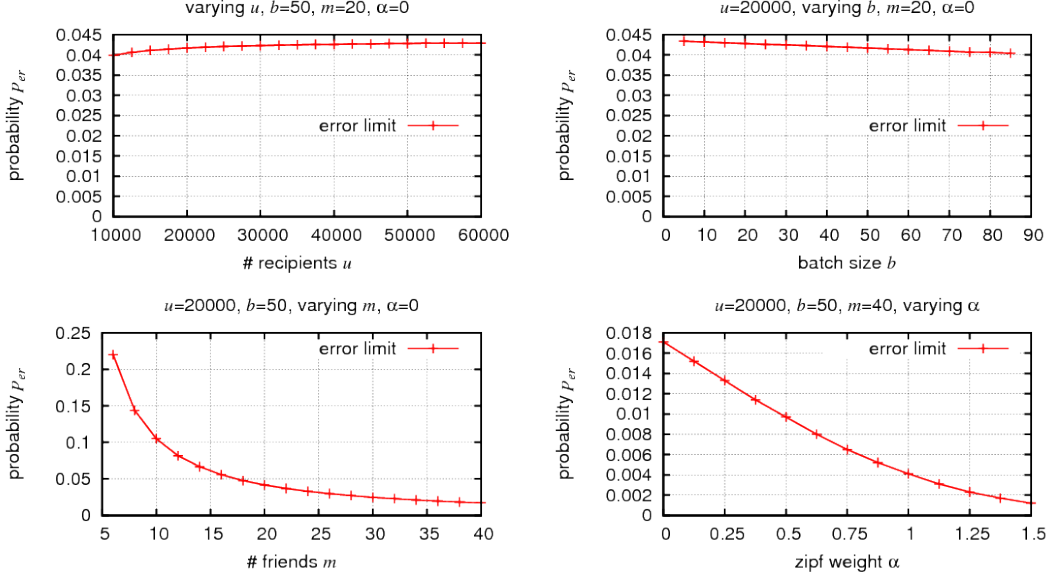


Figure 5.5: Limit of probability of erroneous observations for full disclosure.

### 5.3 Summary

The previous chapters analysed the HS-attack for the case that the attacker can collect sufficiently many observations to uniquely identify Alice's set of friends and that these observations are exact. In this chapter we analysed the HS-attack under the condition that one of these two assumptions is not fulfilled.

Section 5.1 considered the case that the attacker does not collect sufficiently many observations for a unique identification of Alice's set of friends. It provided a mathematical model for the evolution of the minimal-hitting-sets computed by ExactHS on the set of observations available to the attacker. This model solely refers to cases, where Alice traffic and the cover-traffic are uniformly distributed. Deploying this model allows estimating the mean number of observations to uniquely identify subsets of Alice's set of friends, as shown in Section 5.1.3.2.

We also suggested analyses for the case that the number of observations available to the attacker is even too low for a unique identification of any of Alice's friend in Section 5.1.3.3. Our model of the evolution of the minimal-hitting-sets computed by ExactHS shows that the class of minimal-hitting-sets that contain less of Alice's friends, diminishes faster than those that contain more of Alice's friends. This allows

## 5. EXTENSION

---

estimating the number of observations, such that the minimal-hitting-sets computed by ExactHS contain a high number of Alice's friends. That is ExactHS provides several hypotheses for Alice's set of friends, if it cannot find a unique solution. Applying the analysis in Section 5.1.3.3 reveals the probability that the attacker correctly identifies a certain number of Alice's friend by choosing one of those hypotheses.

Section 5.2 studied the case that Alice has a static set of friends, but the attack occasionally collects erroneous observations that miss her friends. The term vague observation comprises erroneous and non-erroneous observations. Section 5.2.1 contributed an adaptation of the HS-attack to identify Alice's set of friends  ${}_A\mathcal{H}$ , despite erroneous observations. That attack uniquely identifies  ${}_A\mathcal{H}$  with a certain probability that is dependent on the number of observations, the fraction of erroneous observation, and the Mix parameters and traffic distributions.

Provided this concrete adaptation of the HS-attack, we contributed analytical analyses of the mentioned dependencies, such that  ${}_A\mathcal{H}$  can be uniquely identified by the HS-attack with a high probability. These analyses apply to various Mix parameters and arbitrary distributions of Alice's traffic. Section 5.2.2.2 derived the limit of the probability of an erroneous observation in each round, such that the probability that  ${}_A\mathcal{H}$  is uniquely identified by the HS-attack approaches 1 for increasing number of vague observations occurred. Equation (5.16) shows that each erroneous observation just requires the HS-attack to collect a constant number of additional observations to uniquely identify  ${}_A\mathcal{H}$ . This additional number of observations is just a logarithmical fraction of the number of observations required to uniquely identify  ${}_A\mathcal{H}$  provided no erroneous observations.

The case that the probability of erroneous observations is too high for a unique identification of  ${}_A\mathcal{H}$  has not been considered yet. This can be considered as a case, where ExactHS provides several hypotheses, each containing some subsets of  ${}_A\mathcal{H}$ . We might obtain those analyses by combining the analyses induced by Section 5.1 and by Section 5.2. This is left for future works.



# Chapter 6

## Related Works

Traffic analyses aim at modelling the information leakage of an anonymity system and its accumulation over time as a means to measure the anonymity protection provided by a system.

The problem of traffic analysis for anonymous communication was presented by Raymond [2001]. In the meantime, a broad range of traffic analysis approaches have been provided by the literature, cf. Danezis [2004]; Kesdogan [2006]; Murdoch and Danezis [2005]; Serjantov [2004]; Troncoso [2011]; Wang et al. [2007]; Wright et al. [2002], that refer to distinct anonymity systems and attacker models. However, this thesis is concerned with the fundamental limit of anonymity protection against a strong attacker, in open environments. Therefore, we only focus on traffic analyses that apply to the Chaum Mix in open environments, with respect to a strong attacker. These approaches deploy the immanent information leakage of the Mix concept, the anonymity sets and recipient sets, accumulated over several rounds. They mainly consider a passive attacker, as such an attacker is sufficient to gain information about a user's profile, while being hard to detect, or to thwart<sup>1</sup>. These attacks thus represent a fundamental threat for the Mix concept in open environments.

We distinguish traffic analyses on the Chaum Mix between *combinatorial* and *heuristic* analyses. Combinatorial analyses are basically concerned with the disclosure of exact information about a user's set of friends that is consistent to the observations of the anonymity system. They can determine all potential set of friends of a

---

<sup>1</sup>Because these attacks just make use of information immanently available to all users.

## 6. RELATED WORKS

---

user from the attacker's observations, as well as the least number of observations to unambiguously identify the friends, as outlined in Section 2.

In contrast to that, heuristic analyses are concerned with providing estimates of a targeted user's set of friends, or of its profile. A *profile* assigns to every observed recipient the probability that it receives a message from the targeted user. Heuristic analyses usually deploy statistical properties of the communication traffic and relax constraints that would be necessary to infer exact information about a user's communication. This can provide computationally efficient and easily applicable attacks. Statistical properties usually become significant, if the number of accumulated observations is large. However, a small number of observations might lead to distorted statistics and less predictive biases, so that heuristic approaches are less suitable to estimate the hard limit of anonymity protection. Section 3.1.3 particularly shows for the SDA that deploying statistics does not generally lead to a number of observations to guess Alice's profile that is lower than, or close to that to uniquely identify Alice's set of friends by the HS-attack.

### 6.1 Combinatorial Analyses

Combinatorial analyses pick up the fact that users tend to have a persistent set of friends. How long the set of friends is persistent with respect to the number of communication rounds in an anonymity system is up to the individual users and might range from one round to all observed rounds.

Given that a user's set of friends is persistent for a sufficient number of rounds, combinatorial approaches can exactly identify that user's friends from observations collected during all the rounds of the anonymity system. This applies to the case of no erroneous observations and non-pathological communication traffic<sup>1</sup>. The number of observations required by the combinatorial approaches for that exact identification determines when an anonymity system fails to protect its users and the class of user, whose friends can be exactly identified<sup>2</sup>.

---

<sup>1</sup>Cases, where the same recipients are contacted all the time and thus appear in every observation are considered pathological, as they are not typical communication traffic.

<sup>2</sup>Those are all users, whose set of friends are persistent for a number of rounds that is not less than the number of observations required to identify those friends.

---

Attacks based on combinatorial analyses are powerful attacks, as they do not rely on any special assumptions about traffic distributions and traffic patterns, apart from the fact that the attacked user's set of friends remains persistent during the attack<sup>1</sup>. They are unbiased under this condition. This thesis follows the line of these analyses and extends them.

## 6.1.1 Attacks for Unique Identification

### 6.1.1.1 Intersection Attack

The *Intersection attack* introduced by Berthold and Langos [2003] applies to the special case of  $m = 1$ . That is a targeted user who we call Alice repeatedly contacts the same recipient, as in case of sending a long stream of data packets to a single recipient. In that case, whenever Alice sends a message through the anonymity system, the set of active recipients at that round always contains Alice's recipient. The attacker therefore computes the intersection of all sets of active recipients, where Alice is an active sender, that is referred to by the term observation in Section 2.1.2.1. As users are only active in an open environment<sup>2</sup>, if they have to send, or receive messages, the intersection of sufficiently many observations, will result in a singleton that uniquely identifies Alice's recipient. This combinatorial analysis of the set of active senders and recipients at several rounds thus uncovers the relationship-anonymity of Alice and exactly identifies her recipient.

### 6.1.1.2 Disclosure Attack

The *Disclosure attack* introduced by Kesdogan et al. [2003]; Agrawal et al. [2003a,b] was the first approach that generalised the idea of the intersection attack to users with a persistent set of friends of any size  $m \in \mathbb{N}$ . Similar to the intersection attack, the attacker targets a user Alice, to uncover her relationship-anonymity, that is her set of friends. Whenever Alice sends a message, the active set of recipients at the same round, i.e., the observation, will contain at least one of her friends.

The Disclosure attack consists of two phases, the *learning phase* and the *excluding*

---

<sup>1</sup>This is motivated by the observation that humans tend to have persistent contacts.

<sup>2</sup>This is in contrast to the closed environment, where all users are always active.

## 6. RELATED WORKS

---

*phase*, in which combinatorial analyses are applied to observations collected by the attack. These phases are described next.

**Learning phase:** The attacker collects  $m$  mutually disjoint observations that is called the *basis set*, so that each of these observations contains exactly one Alice's friend.

**Excluding phase:** The attacker repeats collecting a new observation to exclude those recipients in the basis set that are non-friends, until the basis set entirely consists of singletons<sup>1</sup>. Recipients in a basis set can be excluded, whenever the attacker finds a new observation that intersects exactly one set in the basis set. In that case, he replaces the set in the basis set that is intersected by the new observation, by the intersection of these two sets, thus resulting in an update of the basis set.

The authors derived mathematical estimates of the mean number of observations required by the Disclosure attack to uniquely identify all of Alice's friends, for the case that Alice contacts her friends according to a uniform distribution, cf. Agrawal et al. [2003a,b].

The combinatorial analyses deployed by the Disclosure attack are not specifically restricted to the Chaum Mix, but applies to all anonymity systems that provide sender and recipient anonymity sets. However the Disclosure attack requires solving an NP-complete problem, cf. Agrawal et al. [2003a].

### 6.1.1.3 Hitting-Set Attack

The *Hitting-Set attack* (HS-attack) proposed by Kesdogan and Pimenidis [2004] enables disclosing the relationship-anonymity of a targeted user Alice, who has a persistent set of friends of any size  $m \in \mathbb{N}$ . This attack is applicable to all systems where the Disclosure attack is applicable. In contrast to the Disclosure attack, the HS-attack provably requires the least number of observations to uniquely identify Alice's friends as proved by Kesdogan et al. [2006]. It thus determines a hard limit for the anonymity protection provided by the considered anonymity system.

The idea of the HS-attack is based on the fact that every set of active recipients, where Alice is an active sender (i.e., observation), must contain at least one of Alice's

---

<sup>1</sup>In that case, each singleton represents one of Alice's friends.

---

friends. Consequently, the set of Alice's friends must hit every observation, implying that Alice's set of friends must be a hitting-set. If the attacker collects sufficiently many observations, then Alice's set of friends becomes a unique minimum-hitting-set.

The HS-attack initially computes all  $\binom{u}{m}$  possible sets of  $m$  recipients and excludes all those sets that do not hit all collected observations. The attack repeats this exclusion of possible sets of friends and collection of new observations until only one possible set of friends remains. In that case, this set of friends is a unique minimum-hitting-set with respect to the observations collected by the attacker and uniquely identifies Alice's set of friends. Computing minimum-hitting-sets is known to be an NP-complete problem, according to Garey and Johnson [1990].

In retrospect, the Intersection respectively Disclosure attack also compute the unique minimum-hitting-set for the special case of  $m = 1$ , respectively for the general case of  $m \in \mathbb{N}$ . The Intersection attack is thus equivalent to the HS-attack for  $m = 1$ . However, the Disclosure attack disregards all observations that do not provide  $m$  mutually disjoint sets in the learning-phase and all observations that are not disjoint to  $(m - 1)$  sets in the basis set of the excluding-phase. These are sufficient, but not necessary conditions to identify the unique minimum-hitting-set. In opposite to that, the identification of the unique minimum-hitting-set by the HS-attack is based on a necessary and sufficient condition, cf. Kesdogan et al. [2006], that exploits all observations and thus requires fewer observations than the Disclosure attack. Consider, for example, for  $m = 2$  and  ${}_A\mathcal{H} = \{1, 2\}$  the following observation set:

$$\mathcal{OS} = \{\{1, 3, 4\}, \{2, 4, 5\}, \{1, 5, 6\}, \{1, 2, 7\}, \{2, 3, 6\}\} .$$

The unique minimum-hitting-set in  $\mathcal{OS}$  is  ${}_A\mathcal{H}$ , although there are no two mutually disjoint sets in  $\mathcal{OS}$ , as required by the learning-phase of the Disclosure attack. Since the HS-attack was proven to require the least number of observations for the unique identification of friends by Kesdogan et al. [2006], this example sufficiently proves that the Disclosure attack usually requires more than that least number of observations.

## 6. RELATED WORKS

---

### 6.1.2 Least Number of Observations for Unique Identification

#### 6.1.2.1 Unicity-Distance

An abstract analysis of the least number of observations that are theoretically necessary to uniquely identify Alice's set of friends, based on a probabilistic model of observations and information was provided by Kesdogan and Pimenidis [2006]. That analysis applies Shannon's unicity-distance, cf. Shannon [1949], to the traffic analysis problem.

By applying the Entropy functions of Shannon [1949] to this probabilistic model, the approach of Kesdogan and Pimenidis [2006] measures the theoretical reduction of uncertainty about Alice's set of friends provided by each observation. Following the idea of Shannon's unicity-distance, cf. Shannon [1949], the least number of observations, such that the uncertainty about Alice's set of friends is 0, provides a theoretical lower bound for the number of observations required to uniquely identify Alice's set of friends, cf. Kesdogan and Pimenidis [2006].

This lower bound is solely derived for the case that Alice's traffic and the cover-traffic are uniformly distributed. Another limitation is that it makes no suggestion whether there is an algorithm that can identify Alice's set of friends with a number of observations that is close to that theoretical lower bound. Thus it remains an abstract measure. Indeed, the HS-attack that provably requires the least number of observations to uniquely identify Alice's set of friends, cf. Kesdogan et al. [2006], requires noticeably more observations than this purely theoretical lower bound, as shown by Pham [2006, p. 89].

#### 6.1.2.2 $2\times$ -Exclusivity

The  $2\times$ -*exclusivity* of all Alice's friends, proposed by Kesdogan et al. [2006] and discussed in Section 2.3.2.2 is a necessary but not sufficient condition for the unique identification of Alice's set of friends. Providing an exact formula for the mean number of observations, such that all Alice's friends are  $2\times$ -exclusive determines a provable lower bound for the least number of observations required by the HS-attack to uniquely identify Alice's set of friends Kesdogan et al. [2006]. This bound was mathematically derived, solely for the case that Alice's traffic and the cover-traffic are uniformly dis-

---

tributed. It was introduced by Kesdogan et al. [2006].<sup>1</sup>

The bound derived from the  $2\times$ -exclusivity realistically measures the least number of observations for the unique identification of Alice's set of friends, as the  $2\times$ -exclusivity condition is concrete and verifiable for any set of collected observations. Indeed evaluations of the application of the HS-attack on simulated observations, cf. Kesdogan et al. [2006], confirm that this bound is rather close to the real mean number of observations required by the HS-attack, as opposed to the abstract unicity-distance of Kesdogan and Pimenidis [2006] for traffic analysis, cf. Pham [2006, p. 89].

## 6.2 Heuristic Analyses

Heuristic analyses are concerned with guessing Alice's friends with a probability that is higher than a purely random guess. This is orthogonal to combinatorial analyses that investigate the (least) number of observations to uniquely identify Alice's set of friends to determine a hard limit of anonymity protection provided by a system.

We distinguish between approaches, where a guess specifies a possible set of Alice's friends, i.e., a probable unique minimum-hitting-set, and those where it specifies a possible profile of Alice. In Alice's profile, every friend is assigned to the probability that Alice contacts that friend, while the probability assigned to all non-friends is 0. Given a possible profile of Alice, her possible set of friends can be specified by the set of all recipients in that profile who are assigned to a probability higher than 0. Note that this resulting set might be larger than the number of Alice's friends and might consist of all recipients in the anonymity system in the worst case. A profile thus subsumes the definition of a set of possible Alice's friends. However, this distinction emphasises that guessing the unique minimum-hitting-set is central to the first approach, but not assigning probabilities to observed recipients, as opposed to the second approach.

Heuristic approaches relax the problem of identifying Alice's friends from exactly to with some probabilities. They deploy statistical properties of the traffic and heuristics for the estimate of the possible friends to reduce the complexity of that estimate. Guesses provided by these approaches thus carry some bias that can be complexly interleaved and thus difficult to estimate. However, it is crucial to estimate these biases

---

<sup>1</sup>Note Section 3.1.2 provides a closed formula for this bound that applies to arbitrary non-uniform distributions.

## 6. RELATED WORKS

---

to precisely measure the anonymity provided by a system, such that the measurement is meaningful.

### 6.2.1 Likely Set of Friends

#### 6.2.1.1 Statistical-Hitting-Set Attack

Kesdogan and Pimenidis [2004] suggested a statistical version of the HS-attack, the *Statistical-Hitting-Set* (SHS) attack that provides a low computation complexity at the cost of possibly biased results. The idea of the SHS-attack Kesdogan and Pimenidis [2004] is to avoid the evaluation of all possible sets of  $m$  recipients to find a unique minimum-hitting-set, by just evaluating a fixed number of sets that are assumed to be most likely.

The *basic assumption of the SHS-attack* is that the recipients who are most frequently observed in the observation set aggregated by the attacker, are also most likely to be Alice's friends. Therefore, the SHS-attack prefers evaluating sets of those recipients who appear most frequently in the observations collected by the attacker and have not been evaluated. By evaluating the fixed number of sets of  $m$  recipients, the attack either finds a probable unique minimum-hitting-set<sup>1</sup>, or none, or several hitting-sets of size  $m$ . The attack is terminated in the first case by assuming that the probable unique-minimum-hitting-set is Alice's set of friends, while it is repeated with additional observations in the latter cases<sup>2</sup>.

As the SHS-attack does not evaluate all sets of  $m$  recipients, it might never find any probable unique minimum-hitting-set, or only one that is not Alice's set of friends. These biases arise, if the number of collected observations is too small, or if there are some recipients who are more frequently contacted than some of Alice's friends. This could be due to statistical variances, or due to Alice's traffic and cover-traffic that are non-uniformly distributed.

---

<sup>1</sup>The attacker cannot be sure that it is a unique-minimum-hitting-set, as he does not consider all sets of  $m$  recipients.

<sup>2</sup>The attacker can exit this loop by stopping the attack.



---

### 6.2.1.2 Variants of Statistical-Hitting-Set Attack

Two further variants of the SHS-attack have been suggested by Kesdogan et al. [2009], that are based on the same basic assumption.

The first attack (A1) assumes by collecting a sufficient large number of observations that the smallest hitting-set  $\mathcal{H}$  that consists of the  $|\mathcal{H}|$  most frequently contacted recipients in those observations is Alice's set of friends. This attack can be considered as a special case of the SHS-attack, where only one probable set of Alice's friends is evaluated. For the case of uniform communication of Alice and the other senders, the paper proposes a formula for the sufficient number of observations to find Alice's recipients being the  $m$  most frequently contacted recipients with a high probability.

The second attack (A2) is an extension of the first attack, which additionally evaluates, whether the hitting-set  $\mathcal{H}$  of most frequent recipients in the collected observations is a probable unique minimum-hitting-set. In this evaluation,  $\mathcal{H}$  is identified as a unique minimum-hitting-set, if the cumulative frequency of the recipients of any set that result from replacing a single recipient in  $\mathcal{H}$  is lower than the number of collected observations. However this evaluation of an attack would fail, for example, if  $\mathcal{H}$  is a unique minimum-hitting-set, but the cumulative frequency of a proper subset of  $\mathcal{H}$  is higher than the number of collected observations<sup>1</sup>. Applying this attack, thus requires additional constraints on the statistical properties of the frequencies of the recipients, cf. Kesdogan et al. [2009]. The authors also provide a formula for the number of observations required by the attack to correctly identify Alice's set of friends with a high probability. This formula solely applies to the case that Alice's traffic and the cover-traffic are uniformly distributed and comply with the previously mentioned constraints.

The conditions<sup>2</sup> specified by these variants of the SHS-attacks also apply to the SHS-attack and show that it is under those particular conditions possible to efficiently identify Alice's set of friends with a high probability, if the required number of observations is of no concern. As illustrated by Table 6.1, the average number of observations required by the attacks A1 and A2 to identify Alice's set of friends in 99% of the cases is significantly higher than that required by the HS-attack to exactly identify

---

<sup>1</sup>This is possible, as Alice's friends are also contacted by senders other than Alice, which increases the corresponding frequencies.

<sup>2</sup>Those are the statistical properties of the frequencies of the recipients and the number of observations that should be collected for the attack.

## 6. RELATED WORKS

---

$u$	$b$	$m$	HS	A1	A2
20000	50	20	119	66468	580094
50000	50	20	113	66259	1449190

Table 6.1: Mean number of observations for identification of Alice’s set of friends with 99% chance by attack (A1), (A2), versus unique identification by HS-attack (HS).

Alice’s set of friends, although all conditions required by A1 and A2 are fulfilled.

Due to relying on the basic assumption, the HS-attack as well as its variants prefer evaluating sets of most frequently contacted recipients, so that the more of Alice’s friends are less frequently contacted than other recipients, the more inductive biases are caused by these attacks. This inauspicious situation is likely for realistic non-uniform communication distributions of Alice and the other senders, as the Zipf distribution, cf. Adamic and Huberman [2002]; Breslau et al. [1999]. Therefore more observations might not lead to less biases.

### 6.2.1.3 HS\*-Attack

The *HS\*-attack* proposed by Kesdogan and Pimenidis [2004] aims at reducing the computational complexity of the HS-attack and has a binary outcome: Either it provides a correct disclosure of Alice’s set of friends, or no results.

The HS\*-attack consists of two steps: Firstly, the computation of a minimal-hitting-set and secondly, the evaluation of the uniqueness of the minimal-hitting-set. In the first step, the SHS algorithm of Kesdogan and Pimenidis [2004] is applied to efficiently determine a possible set of Alice’s friends, i.e., a minimal hitting-set  $\mathcal{H}$  in the set of observation  $\mathcal{OS}$  collected by the attack. In the second step, an algorithm is applied that either proves that  $\mathcal{H}$  is a unique minimum-hitting-set in  $\mathcal{OS}$ , or provides another minimal-hitting-set to disprove that  $\mathcal{H}$  is not unique. The HS\*-attack uniquely identifies Alice’s set of friends in the first case and terminates, while it repeats the attack for additional observations in the latter case.

As the HS\*-attack relies on computing minimal-hitting-sets by the SHS-attack, it might require more observations for the unique identification of Alice’s set of friends than the HS-attack. In cases of inauspicious traffic distributions, where the SHS-attack cannot find a minimal-hitting-set, the HS\*-attack might not terminate.

The HS\*-attack is claimed to provide a super-polynomial worst case time-complexity

---

for the identification of Alice’s set of friends, cf. Kesdogan and Pimenidis [2004], that is determined by the complexity of its second step. This is due to the belief that the uniqueness of a given minimal-hitting-set can be verified in super-polynomial runtime, which is unfortunately wrong. As proved by Pham [2006, pp. 27 – 31], the worst case time-complexity of the HS\*-attack remains exponential.

## 6.2.2 Likely Profiles

### 6.2.2.1 Statistical Disclosure Attack

The *Statistical Disclosure* attack (SDA) introduced by Danezis [2003] is solely based on evaluating the ratio of the number of messages received by each recipient in the observations collected by the attacker. It aims at estimating the traffic distribution of Alice that is also called the *profile* of Alice.

Due to the law of large numbers, the ratio of the number of messages sent by Alice to her friends becomes statistically distinguishable from that sent by the other senders to the recipients in the cover-traffic, for large number of observations, if the ratio largely remains static. The SDA deploys this property by estimating the ratios of messages received by recipients in the cover-traffic (i.e., without Alice’s traffic) and computing its difference to the ratios of messages received by the recipients in all collected observations. Evaluating this difference provides an estimate of Alice’s traffic distribution with respect to the set of all recipients in the anonymity system. This thus represents an estimate of Alice’s profile and aids guessing Alice’s likeliest friends, cf. Danezis [2003].

The advantage of the SDA is its efficiency and simplicity by just relying on statistics, thus allowing it to be applied to various anonymity systems, as outlined by Danezis and Serjantov [2005]; Danezis et al. [2007]; Mathewson and Dingledine [2005].

The error probabilities of the guessed Alice’s friends depend on the variance of the real traffic distributions<sup>1</sup> from the estimated traffic distributions, the characteristics of the traffic distributions and the number of collected observations. In the case of static and uniformly distributed cover-traffic and static distribution of Alice’s friends, the number of observations required by the SDA to classify a given Alice’s friend as a friend within a given probability, is mathematically derived by Danezis [2003,

---

<sup>1</sup>Those are Alice’s traffic and that of the other senders.

## 6. RELATED WORKS

---

2004]. This mathematically analyses the *true-positive* rate, the correct classification of an Alice's friends as a friend by the SDA. However, it does not provide analytical analyses of the *false-positive* rate, the classification of a non-friend as an Alice's friend. That rate might be arbitrarily large when applying the SDA. Therefore, the bias of the estimate of Alice's profile by the SDA is still not well understood, as outlined by Pérez-González and Troncoso [2012].

### 6.2.2.2 Variants of Statistical Disclosure Attack

The *Perfect Matching Disclosure attack* (PMDA) and the *Normalised Statistical Disclosure attack* (NSDA) proposed by Troncoso et al. [2008] apply the SDA to initially estimate the profiles of all observed senders. That initial estimate serves as a start point to re-estimate those senders' profiles for new observations. The NSDA relaxes some constraints considered in the PMDA to decrease the complexity of the attack at the expense of increasing biases, as PMDA might be computationally infeasible for large anonymity systems, as outlined by Troncoso et al. [2008].

The SDA estimates Alice's profile from the single observations collected by the attacker, thereby a-priori assuming in each observation that every recipient in that observation has the same probability of being Alice's friend. The re-estimate of Alice's profile by the PMDA and the NSDA is similar to this estimate by the SDA, with the exception that in each single observation, the recipient with the highest probability in the initial profile is assigned a high probability to be Alice's friend, while the remaining recipients are assigned the same low probability.

Troncoso et al. [2008] state that this re-estimate reduces the bias of the estimated sender profiles below the bias caused by the SDA, in some cases. However, Danezis and Troncoso [2009]; Diaz et al. [2008] report that arbitrary small biases in the estimate of a profile can lead to arbitrarily large inductive biases, when reusing that profile. The PMDA and the NSDA do not provide analytical analyses of the bias of the estimate of users' profiles, as outlined by Pérez-González and Troncoso [2012]. In contrast to the SDA, Troncoso et al. [2008] also do not provide analytical analyses of the true-positive rates of their attacks.

---

### 6.2.2.3 Bayesian-Interference

The attack by *Bayesian-interference* of Danezis and Troncoso [2009] provides a co-estimate of the likely friends of every sender and the likely *matching between senders and recipients* of messages in the batches relayed by the Mix. This co-estimate is applied to the sender-anonymity sets and recipient sets accumulated by the attacker. It is based on the similar idea like the PMDA and the NSDA, that gaining information about the users' friends induces information about the likely senders and recipients of messages in a batch and vice versa. According to Danezis and Troncoso [2009], Bayesian-interference avoids inductive errors due to reusing biased profiles in the co-estimate, in contrast to the PMDA and the NSDA.

The attack suggested by Danezis and Troncoso [2009] formulates the estimate of users' profiles and of the matching between senders and recipients of batches from observations collected by the attacker as a Bayesian interference problem. This provides an a-posteriori probability distribution of the users' profiles and of the matchings with respect to the given observations. Describing this a-posteriori probability distribution mathematically is due to its high complexity difficult, so that it is estimated by sampling from marginal distributions of that probability distribution. According to Pérez-González and Troncoso [2012], an analytic analysis of the bias of the estimated profile has not been provided.

### 6.2.2.4 Least Square Disclosure Attack

The *Least Square Disclosure attack* (LSDA) of Pérez-González and Troncoso [2012] suggest an algorithm that aims at estimating a user's profile and quantifying the bias of that estimate analytically.

The LSDA considers the number of messages sent by a sender, e.g., Alice to a recipient in a round as a random event that depends on the unknown probability  $p$  with which Alice contacts that recipient. The attack tries to estimate for every possible recipient of Alice, that corresponding probability, to estimate Alice's profile. The basic idea of the LSDA is to assume that this random event is  $\mathcal{N}(\mu, \sigma^2)$  normally distributed, with unknown mean  $\mu$  and variance  $\sigma^2$  that is determined by  $p$ . This allows applying the well know *least-square approach* to estimate the unknown parameters  $\mu, \sigma^2$  of the normal distribution of the number of messages exchanged between Alice and that

## 6. RELATED WORKS

---

recipient, from the observed number of transmitted messages in the single collected observations, cf. Pérez-González and Troncoso [2012]. Since  $p$  is linked to  $\mu, \sigma^2$ , this also provides an estimate of the probability  $p$  of communications between Alice and the considered recipient.

Applying this estimate to all pairs of Alice and a possible recipient provides an estimate of Alice's profile. The least-square approach provides reliable analytical analyses of the bias of the estimates, if the considered random events are normally distributed.

The LSDA, as evaluated by Pérez-González and Troncoso [2012] analyses a quite "artificial case" of user communication, in which all users have the same number of friends and contact their friends according to the same uniform distribution. The communication distribution in this special case provides mathematical properties that support the application of the least-square approach<sup>1</sup> and thus leads to reasonable estimates of user profiles.

However, as assumed by Pérez-González and Troncoso [2012], the number of messages transmitted between senders and recipients can be arbitrarily distributed. Therefore, more realistic communications would add inductive errors to the estimate of the user profiles that are not accounted for by the least-square approach and by the LSDA. Realistic communications might also be hard to evaluate, since the LSDA requires modelling the traffic distribution of all individual users in the anonymity system, to estimate a single profile, cf. Pérez-González and Troncoso [2012].

### 6.3 Summary

The Intersection attack of Berthold and Langos [2003] and the Disclosure attack of Agrawal et al. [2003a,b]; Kesdogan et al. [2003] are combinatorial attacks that uniquely identifies Alice's set of friends. They are the first attacks that prove that the Chaum Mix does not provide anonymity protection against a strong passive attacker who observes the anonymity and recipient sets generated by the Mix.

Succeeding works on the combinatorial attack increase the effectiveness of the Disclosure attack and provide estimates of the least number of observations required for the unique identification of Alice's set of friends. These are represented by the

---

<sup>1</sup>The number of messages a friend receives from Alice is not normally distributed, but the number of messages it receives from all senders is in this special case similar to a normal distribution.

---

HS-attack of Kesdogan and Pimenidis [2004] and the  $2\times$ -exclusivity of Kesdogan et al. [2006] and the unicity-distance analysis of Kesdogan and Pimenidis [2006]. The HS-attack requires the provably least number of observations for that unique identification, but it is computationally intractable. The mathematical estimates of the least number of observations to uniquely identify Alice's set of friends solely apply to the case of uniformly distributed Alice's traffic and cover-traffic.

Due to the computational infeasibility of the Disclosure attack, heuristic attacks have been suggested. They provide efficient guesses of Alice's friends at the expense of biases. Among those heuristic attacks, we distinguish two categories: Attacks in the first category try to guess a probable unique minimum-hitting-set and are based on the HS-attack. These attacks are represented by the SHS-attack of Kesdogan and Pimenidis [2004], its variants and by the HS\*-attack, cf. Kesdogan and Pimenidis [2004]; Kesdogan et al. [2009]. The heuristics applied by those attacks to decrease the complexity of identifying a probable set of Alice's friends refer to the case of uniformly distributed Alice's traffic and cover-traffic. Analyses of Kesdogan et al. [2009] show that the number of observations required by those attacks to identify Alice's set of friends with a sufficiently high probability is by orders of magnitudes higher than that required by the HS-attack

Attacks in the second category are more general and provide guesses of Alice's traffic distribution, called Alice's profile. They assign each recipient a probability that it is Alice's friend. The idea of those attacks on the Chaum Mix was introduced by the SDA of Danezis [2003]. According to the evaluations published by Danezis [2003]; Troncoso et al. [2008], this attack and its variants provide reasonable true-positive rates for the classification of an Alice's friend as a friend. The false-positive rate in these attacks might be arbitrary and is difficult to determine, as mentioned by Pérez-González and Troncoso [2012]. The Bayesian-interference aims to decrease inductive biases due the heuristics deployed in the former heuristic attacks. However, as outlined by Pérez-González and Troncoso [2012] the heuristic nature of heuristic attacks makes it generally hard to provide a precise estimate of the biases. The LSDA aims at providing analytical analyses of the biases in its guesses of Alice's profile. However, those analyses are solely applied to an artificial Mix configuration. It is not clear from the paper, cf. Pérez-González and Troncoso [2012], how more realistic Mix configurations could be analysed and whether those analyses would be computationally feasible. To

## **6. RELATED WORKS**

---

the best of our knowledge, only the SDA has been demonstrated to be applicable to the pool-Mix, cf. Danezis and Serjantov [2005]; Mathewson and Dingledine [2005], with respect to all attacks mentioned in this section.



# Chapter 7

## Conclusion

This thesis addressed the fundamental limit of anonymity protection provided by anonymity sets constructed with the Chaum Mix technique. It proposed an analytical analysis of the least number of observations<sup>1</sup> and of the time-complexity to uniquely identify Alice's set of friends. These analyses showed that both quantities depend on each of the Mix parameters (i.e., the number of recipients, friends and the batch size), as well as on the traffic distributions (i.e., the cover-traffic and Alice's traffic distribution). This dependency was mathematically derived and described for the case that Alice's set of friends is static. To the best of our knowledge, we are the first to show that dependency with respect to non-uniform traffic distributions. As it turned out in this thesis, non-uniform traffic distributions can significantly effect the least number of observations and the mean time-complexity to uniquely identify Alice's set of friends.

This basic work was extended to address the two novel cases of insufficient number of observations and of erroneous observations. In the first case, we proposed analytical analyses of the number of observations to disclose partial, or likely information about Alice's friends. It addressed the situation, where the attacker's observations are insufficient to uniquely identify Alice's set of friends.

In the second case, we investigated the effect on the full disclosure of Alice's set of friends, if the attacker does not have full knowledge about the anonymity sets and recipient sets (i.e., the observations) in each Mix round. It provided bounds for the rate of erroneous observations, such that a full disclosure of Alice's set of friends is still

---

<sup>1</sup>Observations derived from the attacker's aggregation of the sender anonymity sets and recipient sets of the Chaum Mix.

## 7. CONCLUSION

---

possible.

All analyses in this thesis referred to the HS-attack that represents combinatorial analyses of the observations of the Chaum Mix. The HS-attack provably requires the least number of observations to uniquely identify Alice's set of friends, provided that the set of friends is static, cf. Kesdogan et al. [2006]. Since the observations are the only information leaked by the Chaum Mix to a strong attacker, the HS-attack therefore measures the fundamental limit of anonymity protection provided by the Chaum Mix. We showed by a redesign of the originally computationally infeasible HS-attack of Kesdogan and Pimenidis [2004] that the unique identification of Alice's set of friends is for many realistic Mix configurations, computationally feasible. This was proved by corresponding analyses of the least number of observations and of the mean time-complexity. We considered the mean time-complexity, as it turned out to be by orders of magnitude lower than the worst case time-complexity, in case of our redesigned HS-attack. It thus represents a more realistic measure of the protection against an attacker than the worst case time-complexity analysed in the past. By means of the before mentioned extensions, we demonstrated that our HS-attack and analyses can be even adapted to the more general cases of partial disclosure and of full disclosure, despite erroneous observations. They thus contribute a basis to also analyse more practice-oriented attacker models and Mix models as discussed in Section 1.2.3.

### 7.1 Future Works

**Continuation of Analyses** This thesis contributed analyses of the fundamental limit of anonymity protection provided by the Chaum Mix for various Mix configurations. They mainly referred to the case that Alice's set of friends is static and that sufficiently many observations can be aggregated by the attacker to uniquely identify Alice's set of friends. The extension of those analyses to the more general case of insufficient number of observations and erroneous observations still remains to be completed. That is analysing the least number of observations to gain partial information about Alice's friends with a certain probability, despite erroneous observations. Chapter 5 proposed works towards that direction and we outline those issues that remain for future works next.

The analyses of the disclosure of partial information in Section 5.1 refer to the

---

case, where the attacker's observations are insufficient to uniquely identify Alice's set of friends. Since up to now, these analyses only refer to uniformly distributed traffic, they remain to be extended to non-uniformly distributed traffic.

The information gain about Alice's friends in case of insufficient number of observations for a full disclosure and in case of erroneous observations have been analysed separately in Chapter 5. It remains for future works to provide analyses that also account for the combination of these two cases of insufficient number of observations and erroneous observations.

The mean time-complexity to gain partial information despite erroneous observations has not been investigated yet and remains to be investigated in future works. It might reveal traffic distributions, where gaining that information is computationally feasible. Recall that until now, the feasibility of the HS-attack has only been analysed for the case that there are sufficiently many observations and no erroneous observations.

**Application of Analyses** Assume that analytical analyses of the number of observations, the mean time-complexity, and the probability to gaining correct partial information about Alice's friends by the HS-attack was provided: Then finding a practical strategy for the Mix to efficiently achieve a desired level of anonymity would still be an open issue. The Chaum Mix with continuous dummy traffic and broadcast provides perfect preservation of anonymity, but is not practical, as outlined in Section 1.2.1.2. Finding a practical strategy is thus not trivial.

Kesdogan [2006, pp. 31 – 37] suggested a strategy, where users monitor their anonymity by evaluating the number of observations that remain to be aggregated to succeed the HS-attack. It sketches the idea that Alice's set of friends could remain anonymous, if Alice could selectively send dummy traffic that is dynamically adapted to her current traffic to extend the number of observations to succeed the HS-attack. This generally raises the open question of the rate, the distribution and the injection strategy of dummy traffic that provides an optimal trade off between throughput and anonymity protection. Answering this might not be trivial, as the attempts of users to optimise their own protection might lead to a decrease of the protection of other users. That is due to the dependence of a user's anonymity protection on the cover-traffic that changes, if individual users changes their traffic. Therefore, future works might need

## 7. CONCLUSION

---

to incorporate game theory to find Pareto-optimal strategies to increase the anonymity protection and throughput for all users. Given such a strategy, the coordination and enforcement of users to follow that strategy need to be investigated in future works. This might be an issue in open environments, as the set of users of the anonymity network are volatile.

**Beyond Anonymous Communication** In this thesis, we analysed the protection provided by the Chaum Mix, a concrete technique that constructs anonymity sets on a simple way. As anonymity sets also model other anonymity systems, future works could investigate the applicability of our analyses to other anonymity techniques and that of those techniques to the Chaum Mix. This targets at finding a uniform and comprehensive evaluation of anonymity protection. Since anonymity protection solutions in practice are usually composed of anonymity techniques of distinct layers (e.g., application- and network-layer), providing a uniform evaluation of single components is crucial to evaluate composed systems.

The following two research areas consider, for example, anonymity techniques to that our analyses of anonymity sets might be adaptable. Investigating this adaptation could be the next research step toward finding a uniform anonymity evaluation.

The concept of providing anonymity sets by the Mix of Chaum [1981] is applied in the concept of *Mix-zones* of Beresford and Stajano [2003]; Freudiger et al. [2007] in the research area of *location-privacy*. Location-privacy addresses the hiding of a user's current and past positions, despite the user's use of location-based-services that request position information. A Mix-zone is described by Beresford and Stajano [2003] as a spatial area, where user's cannot be located. The Mix-zone enables users to change their appearance (e.g., pseudonym) and time characteristic (e.g., their positions relative to each other) to provide unlinkability of the users entering and leaving the Mix-zone. This is in accordance to providing unlinkability for messages entering and leaving a Mix. Similar to communication patterns, users might have movement patterns and location preferences that could be disclosed.

Anonymity sets are also investigated in the research area of *privacy-preserving data publishing*, as surveyed by Fung et al. [2010], like the  $k$ -anonymity of Sweeney [2002] and the  $l$ -diversity of Machanavajjhala et al. [2007]. Privacy-preserving data publishing addresses the release of attributes of sets of users for evaluations, such that

---

the exact link between individual users and their privacy sensitive attributes remain hidden to the evaluator. To achieve *k-anonymity* or *l-diversity*, released attributes are modified, such that for each set of privacy sensitive attributes, there is a sufficient large anonymity set of users who are possibly linked to those attributes. As summarised by Fung et al. [2010], all existing privacy-preserving data publishing techniques leak some information about individual users, provided the existence of background knowledge. Therefore observing a certain number of released data would disclose a user's sensitive attributes. This is similar to the case that every observation of a Mix leaks some information out a user's communication pattern.

## **7. CONCLUSION**

---

# Appendix A

## Simulation Condition

### A.1 Hardware

The simulations in this thesis were executed on servers that are equipped with two Intel Xeon X5650 processors that run at 2.67GHz with 12GB RAM. Each of these servers contains in total 12 dedicated cores. We used simultaneously up to 6 servers for the simulations.

### A.2 Software

The simulation of random observations and the HS-attack using the ExactHS algorithm in this thesis were implemented in C++. To ensure that our results are statistically significant, every experiment<sup>1</sup> is repeated until at least 95% of the considered results fall within 5% of the empirically observed mean. The least number of repetitions of an experiment is set to 300 to avoid strong variances of the empirical mean due to a low number of repetitions. We used Pthreads to run several experiments in parallel to mitigate the time required by these large number repetitions.

---

<sup>1</sup>Recall that, in each experiment, we repeatedly apply the HS-attack on simulated random observations until Alice's set of friend can be uniquely identified.

### A.3 Computational Time and Memory Usage

The time required to complete an experiment is dependent on the Mix configurations. For example, the observed maximal time required by an experiment for the Mix parameters ( $u = 20000, b = 5, m = 20$ ) was below 1 second in the case that Alice's traffic and the cover-traffic were uniformly distributed. In contrast to this, we occasionally observed peak times of up to 29 days for an experiment, for the Mix parameters ( $u = 20000, b = 50, m = 40$ ). However, we could complete all (about 430) experiments in less than 2 months with the latter Mix configuration.

The observed maximal memory usage of a single experiment was less than 200MB with respect to all observed experiments. Since each of our servers executed at most 12 experiments in parallel, the memory usage of each server was below 2.4GB.



# References

- Lada A. Adamic and Bernardo A. Huberman. Zipf's Law and the Internet. *Glottometrics*, 3:143 – 150, 2002. 4, 56, 57, 68, 131, 182
- Dakshi Agrawal, Dogan Kesdogan, and Stefan Penz. Probabilistic Treatment of MIXes to Hamper Traffic Analysis. In *Proceedings of the 2003 IEEE Symposium on Security and Privacy*, pages 16 – 27. IEEE, 2003a. 3, 4, 175, 176, 186
- Dakshi Agrawal, Dogan Kesdogan, and Stefan Penz. Probabilistic Treatment of MIXes to Hamper Traffic Analysis. *IEEE Symposium on Security and Privacy*, 0:16, 2003b. 175, 176, 186
- Virgílio Almeida, Azer Bestavros, Mark Crovella, and Adriana de Oliveira. Characterizing Reference Locality in the WWW. In *Proceedings of the fourth international conference on Parallel and distributed information systems, DIS '96*, pages 92 – 103, 1996. 4, 56, 57, 68, 107, 131
- Kevin Bauer, Damon McCoy, Dirk Grunwald, Tadayoshi Kohno, and Douglas Sicker. Low-resource Routing Attacks Against Tor. In *Proceedings of the 2007 ACM workshop on Privacy in electronic society, WPES '07*, pages 11 – 20. ACM, 2007. 22
- Alastair R. Beresford and Frank Stajano. Location Privacy in Pervasive Computing. *IEEE Pervasive Computing*, 2(1):46 – 55, 2003. 192
- Oliver Berthold and Heinrich Langos. Dummy Traffic against Long Term Intersection Attacks. In *Privacy Enhancing Technologies*, volume 2482 of *LNCs*, pages 110 – 128. Springer, 2003. 15, 175, 186

## REFERENCES

---

- Oliver Berthold, Hannes Federrath, and Stefan Köpsell. Web Mixes: A System for Anonymous and Unobservable Internet Access. In *Designing Privacy Enhancing Technologies*, volume 2009 of *LNCS*, pages 115 – 129. Springer, 2001a. 2
- Oliver Berthold, Andreas Pfitzmann, and Ronny Standtke. The Disadvantages of Free MIX Routes and how to Overcome them. In *Designing Privacy Enhancing Technologies*, volume 2009 of *LNCS*, pages 30 – 45. Springer, 2001b. 16, 18, 19, 21
- Arnon Boneh and Micha Hofri. The Coupon-Collector Problem Revisited – A Survey of Engineering Problems and Computational Methods. *Communications in Statistics. Stochastic Models*, 13(1):39 – 66, 1997. 75, 76
- Robert King Brayton. On the Asymptotic Behavior of the Number of Trials Necessary to Complete a Set with Random Selection. *Journal of Mathematical Analysis and Applications*, 7(1):31 – 61, 1963. 77, 78, 79
- Lee Breslau, Pei Cao, Li Fan, Graham Phillips, and Scott Shenker. Web Caching and Zipf-like Distributions: Evidence and Implications. In *Proceedings of Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM '99*, volume 1, pages 126 – 134. IEEE, 1999. 4, 56, 57, 69, 107, 131, 182
- David L. Chaum. Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms. *Communications of the ACM*, 24(2):84 – 88, 1981. 2, 10, 11, 12, 15, 16, 17, 18, 19, 27, 192
- David L. Chaum. The Dining Cryptographers Problem: Unconditional Sender and Recipient Untraceability. *Journal of Cryptology*, 1(1):65 – 75, 1988. 10
- Carlos R. Cunha, Azer Bestavros, and Mark E. Crovella. Characteristics of WWW Client-based Traces. Technical report, Boston University, 1995. 107
- George Danezis. Statistical Disclosure Attacks: Traffic Confirmation in Open Environments. In *Proceedings of Security and Privacy in the Age of Uncertainty*, pages 421 – 426, 2003. 4, 72, 73, 183, 187
- George Danezis. *Better Anonymous Communications*. PhD thesis, University of Cambridge, 2004. 19, 73, 80, 81, 83, 85, 173, 184

## REFERENCES

---

- George Danezis and Claudia Diaz. A Survey of Anonymous Communication Channels. Technical Report MSR-TR-2008-35, Microsoft Research, 2008. 2, 19
- George Danezis and Seda Gürses. A Critical Review of 10 Years of Privacy Technology. In *Proceedings of Surveillance Cultures: A Global Surveillance Society?*, 2010. 1
- George Danezis and Andrei Serjantov. Statistical Disclosure or Intersection Attacks on Anonymity Systems. In *Information Hiding*, volume 3200 of *LNCS*, pages 293 – 308. Springer, 2005. 183, 188
- George Danezis and Carmela Troncoso. Vida: How to Use Bayesian Inference to De-anonymize Persistent Communications. In *Privacy Enhancing Technologies*, volume 5672 of *LNCS*, pages 56 – 72. Springer, 2009. 184, 185
- George Danezis, Roger Dingledine, and Nick Mathewson. Mixminion: Design of a Type III Anonymous Remailer Protocol. In *Proceedings of the 2003 IEEE Symposium on Security and Privacy*, pages 2 – 15. IEEE, 2003. 2, 19, 21
- George Danezis, Claudia Diaz, and Carmela Troncoso. Two-sided Statistical Disclosure Attack. In *Privacy Enhancing Technologies*, volume 4776 of *LNCS*, pages 30 – 44. Springer, 2007. 183
- Claudia Diaz and Bart Preneel. Taxonomy of Mixes and Dummy Traffic. In *Information Security Management, Education and Privacy*, volume 148 of *IFIP International Federation for Information Processing*, pages 217 – 232. Springer, 2004. 15
- Claudia Diaz, Carmela Troncoso, and Andrei Serjantov. On the Impact of Social Network Profiling on Anonymity. In *Privacy Enhancing Technologies*, volume 5134 of *LNCS*, pages 44 – 62. Springer, 2008. 184
- Roger Dingledine, Nick Mathewson, and Paul Syverson. Tor: The Second-Generation Onion Router. In *Proceedings of the 13th USENIX Security Symposium*, pages 303 – 320. USENIX, 2004. 2, 19, 22
- Danny Dolev and Andrew C. Yao. On the Security of Public Key Protocols. *IEEE Transactions on Information Theory*, 29(2):198 – 208, 1983. 2, 5, 6, 19

## REFERENCES

---

- Elizabeth Robling Denning Dorothy. *Cryptography and Data Security*. Addison-Wesley, 1982. 18
- Matthew Edman and Bülent Yener. On Anonymity in an Electronic Society: A Survey of Anonymous Communication Systems. *ACM Computing Surveys*, 42(1):1 – 35, 2009. 2, 19
- Nathan S. Evans, Roger Dingledine, and Christian Grothoff. A Practical Congestion Attack on Tor Using Long Paths. In *Proceedings of the 18th conference on USENIX security symposium*, SSYM '09, pages 33 – 50. USENIX, 2009. 22
- Andreas Fasbender, Dogan Kesdogan, and Olaf Kubitz. Variable and Scalable Security: Protection of Location Information in Mobile IP. In *Vehicular Technology Conference on Mobile Technology for the Human Race*, volume 2, pages 963 – 967, 1996. 19, 22
- Philippe Flajolet, Danièle Gardy, and Loÿs Thimonier. Birthday Paradox, Coupon Collectors, Caching Algorithms and Self-organizing Search. *Discrete Applied Mathematics*, 39(3):207 – 229, 1992. 76
- Julien Freudiger, Maxim Raya, Márk Félegyházi, Panos Papadimitratos, and Jean-Pierre Hubaux. Mix-Zones for Location Privacy in Vehicular Networks. In *Proceedings of the First International Workshop on Wireless Networking for Intelligent Transportation Systems*, 2007. 192
- Benjamin C. M. Fung, Ke Wang, Rui Chen, and Philip S. Yu. Privacy-Preserving Data Publishing: A Survey on Recent Developments. *ACM Computing Surveys*, 42(4):1 – 53, 2010. 192, 193
- Michael R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Freeman, 1990. 32, 177
- Steven Glassman. A Caching Relay for the World Wide Web. *Computer Networks and ISDN Systems*, 27(2):165 – 173, 1994. 56, 57, 69, 107, 131
- Dogan Kesdogan. *Privacy im Internet - Vertrauenswürdige Kommunikation in offenen Umgebungen*. DuD-Fachbeiträge. Vieweg, 1999. 15, 21

## REFERENCES

---

- Dogan Kesdogan. Confidentiality of Traffic Data. Habilitation thesis, RWTH-Aachen University, 2006. 11, 25, 27, 63, 173, 191
- Dogan Kesdogan and Charles Palmer. Technical Challenges of Network Anonymity. *Computer Communications*, 29(3):306 – 324, 2006. 11, 15, 17, 18, 56
- Dogan Kesdogan and Lexi Pimenidis. The Hitting Set Attack on Anonymity Protocols. In *Information Hiding*, volume 3200 of *LNCS*, pages 326 – 339. Springer, 2004. 3, 4, 26, 31, 32, 35, 55, 68, 85, 176, 180, 182, 183, 187, 190
- Dogan Kesdogan and Lexi Pimenidis. The Lower Bound of Attacks on Anonymity Systems – A Unicity Distance Approach. In *Quality of Protection*, volume 23 of *Advances in Information Security*, pages 145 – 158. Springer, 2006. 4, 71, 72, 178, 179, 187
- Dogan Kesdogan, Jan Egner, and Roland Bschkes. Stop-and-Go-MIXes Providing Probabilistic Anonymity in an Open System. In *Information Hiding*, volume 1525 of *LNCS*, pages 83 – 98. Springer, 1998. 19, 20
- Dogan Kesdogan, Dakshi Agrawal, and Stefan Penz. Limits of Anonymity in Open Environments. In *Information Hiding*, volume 2578 of *LNCS*, pages 53 – 69. Springer, 2003. 175, 186
- Dogan Kesdogan, Dakshi Agrawal, Vinh Pham, and Dieter Rauterbach. Fundamental Limits on the Anonymity Provided by the Mix Technique. In *Proceedings of the 2006 IEEE Symposium on Security and Privacy*, pages 86 – 99. IEEE, 2006. 3, 4, 26, 27, 32, 65, 66, 68, 69, 71, 72, 73, 76, 80, 94, 165, 176, 177, 178, 179, 187, 190
- Dogan Kesdogan, Daniel Mölle, Stefan Richter, and Peter Rossmanith. Breaking Anonymity by Learning a Unique Minimum Hitting Set. In *Computer Science - Theory and Applications*, volume 5675 of *LNCS*, pages 299 – 309. Springer, 2009. 72, 85, 181, 187
- Stefan Köpsell. *Entwicklung und Betrieb eines Anonymisierungsdienstes für das WWW*. PhD thesis, Technische Universität Dresden, 2010. German. 15

## REFERENCES

---

- Ashwin Machanavajjhala, Daniel Kifer, Johannes Gehrke, and Muthuramakrishnan Venkitasubramaniam. L-diversity: Privacy beyond k-anonymity. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(1), 2007. 192
- Nick Mathewson and Roger Dingledine. Practical Traffic Analysis: Extending and Resisting Statistical Disclosure. In *Privacy Enhancing Technologies*, volume 3424 of *LNCS*, pages 17–34. Springer, 2005. 183, 188
- Prateek Mittal, Ahmed Khurshid, Joshua Juen, Matthew Caesar, and Nikita Borisov. Stealthy traffic analysis of low-latency anonymous communication using throughput fingerprinting. In *Proceedings of the 18th ACM conference on Computer and communications security*, CCS ’11, pages 215 – 226. ACM, 2011. 22
- Steven J. Murdoch and George Danezis. Low-Cost Traffic Analysis of Tor. In *Proceedings of the 2005 IEEE Symposium on Security and Privacy*, pages 183 – 195. IEEE, 2005. 22, 173
- Luke O’Connor. Entropy Bounds for Traffic Confirmation. Cryptology ePrint Archive, Report 2008/365, August 2008. URL <http://eprint.iacr.org/2008/>. 4, 71
- Fernando Pérez-González and Carmela Troncoso. Understanding Statistical Disclosure: A Least Squares Approach. In *Privacy Enhancing Technologies*, volume 7384 of *LNCS*, pages 38 – 57. Springer, 2012. 27, 84, 184, 185, 186, 187
- Andreas Pfitzmann. *Diensteintegrierende Kommunikationsnetze mit teilnehmerüberprüfbarem Datenschutz*, volume 234 of *Informatik-Fachberichte*. Springer, 1990. 2, 6, 7, 8, 9, 10, 11, 15, 16
- Andreas Pfitzmann and Marit Hansen. Anonymity, Unobservability, Pseudonymity, and Identity Management – A Proposal for Terminology. Version v0.34, August 2010. 2, 7, 8, 9, 25
- Andreas Pfitzmann and Marit Köhntopp. Anonymity, Unobservability, and Pseudonymity - A Proposal for Terminology. In *Designing Privacy Enhancing Technologies*, volume 2009 of *LNCS*, pages 1 – 9. Springer, 2001. 6, 7, 8, 9

## REFERENCES

---

- Andreas Pfitzmann and Michael Waidner. Networks Without User Observability – Design Options. In *Advances in Cryptology – EUROCRYPT’ 85*, volume 219 of *LNCS*, pages 245 – 253. Springer, 1986. 10, 13
- Dang Vinh Pham. Analysis of Attacks on Chaumian Mixes (Analyse von Angriffen auf Chaummixen). Master’s thesis, RWTH-Aachen University, April 2006. 178, 179, 183
- Dang Vinh Pham and Dogan Kesdogan. A Combinatorial Approach for an Anonymity Metric. In *Information Security and Privacy*, volume 5594 of *LNCS*, pages 26 – 43. Springer, 2009. 37, 53
- Dang Vinh Pham, Joss Wright, and Dogan Kesdogan. A Practical Complexity-Theoretic Analysis of Mix Systems. In *Computer Security ESORICS 2011*, volume 6879 of *LNCS*, pages 508 – 527. Springer, 2011. 88
- Vinh Pham. Analysis of the Anonymity Set of Chaumian Mixes. In *13th Nordic Workshop on Secure IT-Systems*, 2008. 37, 53, 55, 103
- Jean-François Raymond. Traffic Analysis: Protocols, Attacks, Design Issues, and Open Problems. In *Designing Privacy Enhancing Technologies*, volume 2009 of *LNCS*, pages 10–29. Springer, 2001. 173
- Michael K. Reiter and Aviel D. Rubin. Crowds: Anonymity for Web Transactions. *ACM Transactions on Information and System Security*, 1(1):66 – 92, 1998. 10
- Andrei Serjantov. On the Anonymity of Anonymity Systems. Technical Report 604, University of Cambridge, 2004. 19, 173
- Andrei Serjantov and George Danezis. Towards an Information Theoretic Metric for Anonymity. In *Privacy Enhancing Technologies*, volume 2482 of *LNCS*, pages 41 – 53. Springer, 2003. 21, 28
- Andrei Serjantov, Roger Dingledine, and Paul F. Syverson. From a Trickle to a Flood: Active Attacks on Several Mix Types. In *Information Hiding*, volume 2578 of *LNCS*, pages 36 – 52. Springer, 2003. 18, 19, 21

## REFERENCES

---

- Claude Shannon. Communication Theory of Secrecy Systems. *Bell System Technical Journal*, 28:656 – 715, 1949. 3, 4, 8, 26, 32, 68, 71, 178
- Latanya Sweeney. k-Anonymity: A Model for Protecting Privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(5):557–570, 2002. 192
- Paul F. Syverson, David M. Goldschlag, and Michael G. Reed. Anonymous Connections and Onion Routing. In *Proceedings of the 1997 IEEE Symposium on Security and Privacy*, pages 44 – 54. IEEE, 1997. 19, 22
- Carmela Troncoso. *Design and Analysis Methods for Privacy Technologies*. PhD thesis, Katholieke Universiteit Leuven, 2011. 173
- Carmela Troncoso, Benedikt Gierlichs, Bart Preneel, and Ingrid Verbauwhede. Perfect Matching Disclosure Attacks. In *Privacy Enhancing Technologies*, volume 5134 of *LNCS*, pages 2 – 23. Springer, 2008. 184, 187
- Xinyuan Wang, Shiping Chen, and Sushil Jajodia. Network Flow Watermarking Attack on Low-Latency Anonymous Communication Systems. In *Proceedings of the 2007 IEEE Symposium on Security and Privacy*, pages 116 – 130. IEEE, 2007. 22, 173
- Samuel D. Warren and Louis D. Brandeis. The Right to Privacy. *Harvard Law Review*, 4(5):193 – 220, 1890. 1
- Benedikt Westermann. *Challenges of Anonymous Communication: Bridging Gaps Between Theory and Practice*. PhD thesis, NTNU – Trondheim, 2012. 20
- Alan F. Westin. *Privacy and Freedom*. Bodley Head, 1970. 1
- Matthew Wright, Micah Adler, Brian Neil Levine, and Clay Shields. An Analysis of the Degradation of Anonymous Protocols. In *Proceedings of the Network and Distributed Security Symposium, NDSS '02*. The Internet Society, 2002. 173



# Index

$(b - 1)$ -attack, *see*  $(n - 1)$ -attack  
 $(n - 1)$ -attack, 18  
 $R$ , 28  
 $S$ , 28  
 $\alpha$ , 57  
 $\mathcal{C}$ , 90  
 $\mathcal{HS}$ , 37  
 $\mathcal{OS}[r]$ , 37  
 ${}_A\mathcal{H}$ , 29  
 $b$ , 14, 29  
 $m$ , 33  
 $r$ , 28  
 $u$ , 29  
  
Alice, 29  
    friend, 30  
    non-friend, 30  
anonymity, 8  
    perfect preservation, 8  
    recipient-anonymity, 8  
    relationship-anonymity, 8, 9  
    sender-anonymity, 8  
anonymity set, 8, 9  
    recipient-anonymity set, 9  
    relationship-anonymity set, 9  
    sender-anonymity set, 9, 29

attacker  
    global passive, 29  
    active, 6  
    computationally restricted, 6  
    computationally unrestricted, 6  
    global, 6  
    involved, 6  
    local, 6  
    passive, 6  
    strong, 6  
    uninvolved, 6  
average-disprovable, 116  
  
basis set, 176  
batch, 14, 29  
Bayesian-interference, 185  
broadcast, 10  
  
CCP, 75  
    general CCP, 77  
Chaum Mix, 11, 15  
CUVE, 17  
  
DC-Net, 10  
disclosure  
    full, 138  
    partial, 153

## INDEX

---

- Disclosure attack, 175
- distribution
  - non-uniform, 57
  - uniform, 57
  - Zipf, 57
- environment
  - closed, 14
  - open, 14
- ExactHS, 35
  - completeness, 44
  - soundness, 44
- excluding phase, 176
- exclusive, 65
  - $k \times$ -exclusive, 76
  - $1 \times$ -exclusive, 73
  - $2 \times$ -exclusive, 65
- exclusivity
  - $k \times$ -exclusivity, 77
  - $1 \times$ -exclusivity, 73
  - $2 \times$ -exclusivity, 66, 178
- experiment, 56
- extensive-class, 143, 146
  - minimal, 146
  - specified, 145
  - unspecified, 143
- false-positive, 184
- finalised set, 39
- frequency, 37
- Heuristic analyses, 179
- hitting-set, 31
  - evolution, 142
  - minimal, 31
  - unique minimum, 31
- HS\*-attack, 182
- HS-algorithm, 32
- HS-attack, 31, 176
  - succeed, 32
  - using ExactHS, 35
- hypothesis, 30
  - extensive, 143
  - specified, 31
  - specified extensive, 145
- Intersection attack, 175
- item of interest, 7
- learning phase, 175
- Least Square Disclosure attack, 185
- loop-back function, 18
- LSDA, *see* Least Square Disclosure attack
- Mix
  - configuration, 30
  - embedding function, 12
  - group function, 17
  - parameter, 30
- Mix protocol, 14
- Mix-cascade, 16
- Mix-net, 16
- negligible, 33
- no-trivial-disproof, 65
- nonce, 13
- Normalised Statistical Disclosure attack, 184

- NSDA, *see* Normalised Statistical Disclosure attack
- observation, 30
  - erroneous, 159
  - vague, 160
- observation set, 30
  - erroneous, 160
  - vague, 160
- optimistic case, 107, 118
- perfect, 8
- Perfect Matching Disclosure attack, 184
- perfect Mix, 11
- PMDA, *see* Perfect Matching Disclosure attack
- pool-Mix, 21
- potential, 89, 91, 92
- profile, 174
- recipient, 28, 30
  - chosen, 90
  - non-chosen, 90
  - specified, 145
  - unspecified, 145
- recipient set, 29
  - chosen recipient set, 90
- round, 12
- SDA, *see* Statistical Disclosure attack
- Sender protocol, 13
- SHS, *see* Statistical-Hitting-Set attack
- Statistical Disclosure attack, 81, 183
- Statistical-Hitting-Set attack, 180
- Stop-and-Go-Mix, 20
- threshold Mix, 11
- traffic, 57
  - Alice's traffic, 30
  - cover-traffic, 11, 30, 57
- trivial-disproof, 64
- true-positive, 184
- UMHS problem, *see* unique minimum-hitting-set problem
- undetectability, 7
- unicity-distance, 178
- unique minimum-hitting-set problem, 32
- unlinkability, 7
- unobservability, 9