David Martin
Andrei Serjantov (Eds.)

LNCS 3424

Privacy Enhancing Technologies

4th International Workshop, PET 2004
Toronto, Canada, May 2004
Revised Selected Papers

Springer

David Martin Andrei Serjantov (Eds.)

Privacy Enhancing Technologies

4th International Workshop, PET 2004
Toronto, Canada, May 26-28, 2004
Revised Selected Papers

Springer

Volume Editors

David Martin
University of Massachusetts Lowell, Department of Computer Science
One University Ave., Lowell, Massachusetts 01854, USA
E-mail: dm@cs.uml.edu

Andrei Serjantov
University of Cambridge, Computer Laboratory
William Gates Building, 15 JJ Thomson Avenue, Cambridge CB3 0FD, UK
E-mail: aas@arachsys.com

Library of Congress Control Number: 2005926701

CR Subject Classification (1998): E.3, C.2, D.4.6, K.6.5, K.4, H.3, H.4, I.7

ISSN 0302-9743
ISBN-10 3-540-26203-2 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-26203-9 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springeronline.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11423409 06/3142 5 4 3 2 1 0

Lecture Notes in Computer Science 3424

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Lecture Notes in Computer Science

For information about Vols. 1–3422

please contact your bookseller or Springer

Vol. 3532: A. Gómez-Pérez, J. Euzenat (Eds.), The Semantic Web: Research and Applications. XV, 728 pages. 2005.

Vol. 3526: S.B. Cooper, B. Löwe, L. Torenvliet (Eds.), New Computational Paradigms. XVII, 574 pages. 2005.

Vol. 3525: A.E. Abdallah, C.B. Jones, J.W. Sanders (Eds.), Communicating Sequential Processes. XIV, 321 pages. 2005.

Vol. 3524: R. Barták, M. Milano (Eds.), Integration of AI and OR Techniques in Constraint Programming for Combinatorial Problems. XI, 320 pages. 2005.

Vol. 3523: J.S. Marques, N.P. de la Blanca, P. Pina (Eds.), Pattern Recognition and Image Analysis, Part II. XXVI, 733 pages. 2005.

Vol. 3522: J.S. Marques, N.P. de la Blanca, P. Pina (Eds.), Pattern Recognition and Image Analysis, Part I. XXVI, 703 pages. 2005.

Vol. 3520: O. Pastor, J. Falcão e Cunha (Eds.), Advanced Information Systems Engineering. XVI, 584 pages. 2005.

Vol. 3518: T.B. Ho, D. Cheung, H. Li (Eds.), Advances in Knowledge Discovery and Data Mining. XXI, 864 pages. 2005. (Subseries LNAI).

Vol. 3517: H.S. Baird, D.P. Lopresti (Eds.), Human Interactive Proofs. IX, 143 pages. 2005.

Vol. 3516: V.S. Sunderam, G.D. van Albada, P.M.A. Sloot, J.J. Dongarra (Eds.), Computational Science – ICCS 2005, Part III. LXIII, 1143 pages. 2005.

Vol. 3515: V.S. Sunderam, G.D. van Albada, P.M.A. Sloot, J.J. Dongarra (Eds.), Computational Science – ICCS 2005, Part II. LXIII, 1101 pages. 2005.

Vol. 3514: V.S. Sunderam, G.D. van Albada, P.M.A. Sloot, J.J. Dongarra (Eds.), Computational Science – ICCS 2005, Part I. LXIII, 1089 pages. 2005.

Vol. 3513: A. Montoyo, R. Mu\¯noz, E. Métais (Eds.), Natural Language Processing and Information Systems. XII, 408 pages. 2005.

Vol. 3510: T. Braun, G. Carle, Y. Koucheryavy, V. Tsaoussidis (Eds.), Wired/Wireless Internet Communications. XIV, 366 pages. 2005.

Vol. 3509: M. Jünger, V. Kaibel (Eds.), Integer Programming and Combinatorial Optimization. XI, 484 pages. 2005.

Vol. 3508: P. Bresciani, P. Giorgini, B. Henderson-Sellers, G. Low, M. Winikoff (Eds.), Agent-Oriented Information Systems II. X, 227 pages. 2005. (Subseries LNAI).

Vol. 3507: F. Crestani, I. Ruthven (Eds.), Information Context: Nature, Impact, and Role. XIII, 253 pages. 2005.

Vol. 3505: V. Gorodetsky, J. Liu, V. A. Skormin (Eds.), Autonomous Intelligent Systems: Agents and Data Mining. XIII, 303 pages. 2005. (Subseries LNAI).

Vol. 3503: S.E. Nikoletseas (Ed.), Experimental and Efficient Algorithms. XV, 624 pages. 2005.

Vol. 3502: F. Khendek, R. Dssouli (Eds.), Testing of Communicating Systems. X, 381 pages. 2005.

Vol. 3501: B. Kégl, G. Lapalme (Eds.), Advances in Artificial Intelligence. XV, 458 pages. 2005. (Subseries LNAI).

Vol. 3500: S. Miyano, J. Mesirov, S. Kasif, S. Istrail, P. Pevzner, M. Waterman (Eds.), Research in Computational Molecular Biology. XVII, 632 pages. 2005. (Subseries LNBI).

Vol. 3499: A. Pelc, M. Raynal (Eds.), Structural Information and Communication Complexity. X, 323 pages. 2005.

Vol. 3498: J. Wang, X. Liao, Z. Yi (Eds.), Advances in Neural Networks – ISNN 2005, Part III. L, 1077 pages. 2005.

Vol. 3497: J. Wang, X. Liao, Z. Yi (Eds.), Advances in Neural Networks – ISNN 2005, Part II. L, 947 pages. 2005.

Vol. 3496: J. Wang, X. Liao, Z. Yi (Eds.), Advances in Neural Networks – ISNN 2005, Part II. L, 1055 pages. 2005.

Vol. 3495: P. Kantor, G. Muresan, F. Roberts, D.D. Zeng, F.-Y. Wang, H. Chen, R.C. Merkle (Eds.), Intelligence and Security Informatics. XVIII, 674 pages. 2005.

Vol. 3494: R. Cramer (Ed.), Advances in Cryptology – EUROCRYPT 2005. XIV, 576 pages. 2005.

Vol. 3493: N. Fuhr, M. Lalmas, S. Malik, Z. Szlávik (Eds.), Advances in XML Information Retrieval. XI, 438 pages. 2005.

Vol. 3492: P. Blache, E. Stabler, J. Busquets, R. Moot (Eds.), Logical Aspects of Computational Linguistics. X, 363 pages. 2005. (Subseries LNAI).

Vol. 3489: G.T. Heineman, I. Crnkovic, H.W. Schmidt, J.A. Stafford, C. Szyperski, K. Wallnau (Eds.), Component-Based Software Engineering. XI, 358 pages. 2005.

Vol. 3488: M.-S. Hacid, N.V. Murray, Z.W. Raś, S. Tsumoto (Eds.), Foundations of Intelligent Systems. XIII, 700 pages. 2005. (Subseries LNAI).

Vol. 3486: T. Helleseth, D. Sarwate, H.-Y. Song, K. Yang (Eds.), Sequences and Their Applications - SETA 2004. XII, 451 pages. 2005.

Vol. 3483: O. Gervasi, M.L. Gavrilova, V. Kumar, A. Laganà, H.P. Lee, Y. Mun, D. Taniar, C.J.K. Tan (Eds.), Computational Science and Its Applications – ICCSA 2005, Part IV. XXVII, 1362 pages. 2005.

Vol. 3482: O. Gervasi, M.L. Gavrilova, V. Kumar, A. Laganà, H.P. Lee, Y. Mun, D. Taniar, C.J.K. Tan (Eds.), Computational Science and Its Applications – ICCSA 2005, Part III. LXVI, 1340 pages. 2005.

Vol. 3481: O. Gervasi, M.L. Gavrilova, V. Kumar, A. Laganà, H.P. Lee, Y. Mun, D. Taniar, C.J.K. Tan (Eds.), Computational Science and Its Applications – ICCSA 2005, Part II. LXIV, 1316 pages. 2005.

Vol. 3480: O. Gervasi, M.L. Gavrilova, V. Kumar, A. Laganà, H.P. Lee, Y. Mun, D. Taniar, C.J.K. Tan (Eds.), Computational Science and Its Applications – ICCSA 2005, Part I. LXV, 1234 pages. 2005.

Vol. 3479: T. Strang, C. Linnhoff-Popien (Eds.), Location- and Context-Awareness. XII, 378 pages. 2005.

Vol. 3478: C. Jermann, A. Neumaier, D. Sam (Eds.), Global Optimization and Constraint Satisfaction. XIII, 193 pages. 2005.

Vol. 3477: P. Herrmann, V. Issarny, S. Shiu (Eds.), Trust Management. XII, 426 pages. 2005.

Vol. 3475: N. Guelfi (Ed.), Rapid Integration of Software Engineering Techniques. X, 145 pages. 2005.

Vol. 3468: H.W. Gellersen, R. Want, A. Schmidt (Eds.), Pervasive Computing. XIII, 347 pages. 2005.

Vol. 3467: J. Giesl (Ed.), Term Rewriting and Applications. XIII, 517 pages. 2005.

Vol. 3465: M. Bernardo, A. Bogliolo (Eds.), Formal Methods for Mobile Computing. VII, 271 pages. 2005.

Vol. 3464: S.A. Brueckner, G.D.M. Serugendo, A. Karageorgos, R. Nagpal (Eds.), Engineering Self-Organising Systems. XIII, 299 pages. 2005. (Subseries LNAI).

Vol. 3463: M. Dal Cin, M. Kaâniche, A. Pataricza (Eds.), Dependable Computing - EDCC 2005. XVI, 472 pages. 2005.

Vol. 3462: R. Boutaba, K.C. Almeroth, R. Puigjaner, S. Shen, J.P. Black (Eds.), NETWORKING 2005. XXX, 1483 pages. 2005.

Vol. 3461: P. Urzyczyn (Ed.), Typed Lambda Calculi and Applications. XI, 433 pages. 2005.

Vol. 3460: Ö. Babaoglu, M. Jelasity, A. Montresor, C. Fetzer, S. Leonardi, A. van Moorsel, M. van Steen (Eds.), Self-star Properties in Complex Information Systems. IX, 447 pages. 2005.

Vol. 3459: R. Kimmel, N.A. Sochen, J. Weickert (Eds.), Scale Space and PDE Methods in Computer Vision. XI, 634 pages. 2005.

Vol. 3458: P. Herrero, M.S. Pérez, V. Robles (Eds.), Scientific Applications of Grid Computing. X, 208 pages. 2005.

Vol. 3456: H. Rust, Operational Semantics for Timed Systems. XII, 223 pages. 2005.

Vol. 3455: H. Treharne, S. King, M. Henson, S. Schneider (Eds.), ZB 2005: Formal Specification and Development in Z and B. XV, 493 pages. 2005.

Vol. 3454: J.-M. Jacquet, G.P. Picco (Eds.), Coordination Models and Languages. X, 299 pages. 2005.

Vol. 3453: L. Zhou, B.C. Ooi, X. Meng (Eds.), Database Systems for Advanced Applications. XXVII, 929 pages. 2005.

Vol. 3452: F. Baader, A. Voronkov (Eds.), Logic for Programming, Artificial Intelligence, and Reasoning. XI, 562 pages. 2005. (Subseries LNAI).

Vol. 3450: D. Hutter, M. Ullmann (Eds.), Security in Pervasive Computing. XI, 239 pages. 2005.

Vol. 3449: F. Rothlauf, J. Branke, S. Cagnoni, D.W. Corne, R. Drechsler, Y. Jin, P. Machado, E. Marchiori, J. Romero, G.D. Smith, G. Squillero (Eds.), Applications of Evolutionary Computing. XX, 631 pages. 2005.

Vol. 3448: G.R. Raidl, J. Gottlieb (Eds.), Evolutionary Computation in Combinatorial Optimization. XI, 271 pages. 2005.

Vol. 3447: M. Keijzer, A. Tettamanzi, P. Collet, J.v. Hemert, M. Tomassini (Eds.), Genetic Programming. XIII, 382 pages. 2005.

Vol. 3444: M. Sagiv (Ed.), Programming Languages and Systems. XIII, 439 pages. 2005.

Vol. 3443: R. Bodik (Ed.), Compiler Construction. XI, 305 pages. 2005.

Vol. 3442: M. Cerioli (Ed.), Fundamental Approaches to Software Engineering. XIII, 373 pages. 2005.

Vol. 3441: V. Sassone (Ed.), Foundations of Software Science and Computational Structures. XVIII, 521 pages. 2005.

Vol. 3440: N. Halbwachs, L.D. Zuck (Eds.), Tools and Algorithms for the Construction and Analysis of Systems. XVII, 588 pages. 2005.

Vol. 3439: R.H. Deng, F. Bao, H. Pang, J. Zhou (Eds.), Information Security Practice and Experience. XII, 424 pages. 2005.

Vol. 3438: H. Christiansen, P.R. Skadhauge, J. Villadsen (Eds.), Constraint Solving and Language Processing. VIII, 205 pages. 2005. (Subseries LNAI).

Vol. 3437: T. Gschwind, C. Mascolo (Eds.), Software Engineering and Middleware. X, 245 pages. 2005.

Vol. 3436: B. Bouyssounouse, J. Sifakis (Eds.), Embedded Systems Design. XV, 492 pages. 2005.

Vol. 3434: L. Brun, M. Vento (Eds.), Graph-Based Representations in Pattern Recognition. XII, 384 pages. 2005.

Vol. 3433: S. Bhalla (Ed.), Databases in Networked Information Systems. VII, 319 pages. 2005.

Vol. 3432: M. Beigl, P. Lukowicz (Eds.), Systems Aspects in Organic and Pervasive Computing - ARCS 2005. X, 265 pages. 2005.

Vol. 3431: C. Dovrolis (Ed.), Passive and Active Network Measurement. XII, 374 pages. 2005.

Vol. 3430: S. Tsumoto, T. Yamaguchi, M. Numao, H. Motoda (Eds.), Active Mining. XII, 349 pages. 2005. (Subseries LNAI).

Vol. 3429: E. Andres, G. Damiand, P. Lienhardt (Eds.), Discrete Geometry for Computer Imagery. X, 428 pages. 2005.

Vol. 3428: Y.-J. Kwon, A. Bouju, C. Claramunt (Eds.), Web and Wireless Geographical Information Systems. XII, 255 pages. 2005.

Vol. 3427: G. Kotsis, O. Spaniol (Eds.), Wireless Systems and Mobility in Next Generation Internet. VIII, 249 pages. 2005.

Vol. 3424: D. Martin, A. Serjantov (Eds.), Privacy Enhancing Technologies. VIII, 345 pages. 2005.

Vol. 3423: J.L. Fiadeiro, P.D. Mosses, F. Orejas (Eds.), Recent Trends in Algebraic Development Techniques. VIII, 271 pages. 2005.

Preface

The first workshop in this series was held at the International Computer Science Institute in Berkeley and was published as LNCS 2009 under the name "Workshop on Design Issues in Anonymity and Unobservability." Subsequent Privacy Enhancing Technologies (PET) workshops met in San Francisco in 2002 (LNCS 2482) and Dresden in 2003 (LNCS 2760). This volume, LNCS 3424, holds the proceedings from PET 2004 in Toronto. Our 2005 meeting is scheduled for Dubrovnik, and we hope to keep finding new and interesting places to visit on both sides of the Atlantic – or beyond.

An event like PET 2004 would be impossible without the work and dedication of many people. First and foremost we thank the authors, who wrote and submitted 68 full papers or panel proposals, 21 of which appear herein.

The Program Committee produced 163 reviews in total. Along the way, they were assisted in reviewing by Steven Bishop, Rainer Bohme, Sebastian Clauß, Claudia Díaz, Richard E. Newman, Ulrich Flegel, Elke Franz, Stefan Kopsell, Thomas Kriegelstein, Markus Kuhn, Stephen Lewis, Luc Longpre, Steven Murdoch, Shishir Nagaraja, Thomas Nowey, Peter Palfrader, Lexi Pimenidis, Klaus Ploessl, Sivaramakrishnan Rajagopalan, Marc Rennhard, Leo Reyzin, Pankaj Rohatgi, Naouel Ben Salem, Sandra Steinbrecher, Mike Szydlo, Shabsi Walfish, Jie Wang, Brandon Wiley, and Shouhuai Xu.

We invited two prominent speakers to speak at the workshop: Ross Anderson explained the "Economics of Security and Privacy", and Andreas Pfitzmann covered "Research on Anonymous Communication in German(y) 1983–1990." In addition, we held two panel discussions, two lively rump sessions, and we enjoyed a number of memorable social activities. Slides from many of the presentations are available at http://petworkshop.org/

A successful workshop depends not only on an interesting program, but also on a hospitable venue and attention to detail. Alison Bambury did a fantastic job coordinating the local arrangements. Roger Dingledine managed the stipend pool, funded by Microsoft Corporation, the Information and Privacy Commissioner's Office (Ontario), the Centre for Innovation Law and Policy at the University of Toronto, and Bell University Labs. The stipend fund helped more than 20 people attend the workshop. Paul Syverson led the committee to determine the recipients of the PET Award for Outstanding Research, also funded by Microsoft Corporation. Finally, Richard Owens, PET 2004's General Chair, oversaw the whole event and ensured that everything happened as planned and within budget. We offer our thanks to all of you, and to everyone who contributed their time, interest, and resources to the 2004 PET Workshop.

January 2005

David Martin
Andrei Serjantov
Program Committee Co-chairs

Privacy Enhancing Technologies 2004
Toronto, Canada
May 26–28, 2004
http://petworkshop.org/

Program Committee

Alessandro Acquisti, Heinz School, Carnegie Mellon University, USA
Caspar Bowden, Microsoft EMEA, UK
Jean Camp, Kennedy School, Harvard University, USA
Richard Clayton, University of Cambridge, UK
Lorrie Cranor, School of Computer Science, Carnegie Mellon University, USA
George Danezis, University of Cambridge, UK
Roger Dingledine, The Free Haven Project, USA
Hannes Federrath, Universität Regensburg, Germany
Ian Goldberg, Zero Knowledge Systems, Canada
Philippe Golle, Palo Alto Research Center, USA
Marit Hansen, Independent Centre for Privacy Protection Schleswig-Holstein, Germany
Markus Jakobsson, RSA Laboratories, USA
Dogan Kesdogan, Rheinisch-Westfälische Technische Hochschule Aachen, Germany
Brian Levine, University of Massachusetts, Amherst, USA
David Martin, University of Massachusetts, Lowell, USA
Andreas Pfitzmann, Dresden University of Technology, Germany
Matthias Schunter, IBM Zurich Research Lab, Switzerland
Andrei Serjantov, University of Cambridge, UK
Adam Shostack, Informed Security Inc., Canada
Paul Syverson, Naval Research Lab, USA

General Chair

Richard Owens, University of Toronto, Canada

Sponsoring Institutions

Microsoft Corporation
Information and Privacy Commissioner's Office (Ontario)
Centre for Innovation Law and Policy at the University of Toronto
Bell University Labs

Table of Contents

Anonymity and Covert Channels in Simple Timed Mix-Firewalls*

Richard E. Newman[1], Vipan R. Nalla[1], and Ira S. Moskowitz[2]

[1] CISE Department,
University of Florida,
Gainesville, FL 32611-6120
{nemo, vreddy}@cise.ufl.edu
[2] Center for High Assurance Computer Systems, Code 5540,
Naval Research Laboratory,
Washington, DC 20375
moskowitz@itd.nrl.navy.mil

Abstract. Traditional methods for evaluating the amount of anonymity afforded by various Mix configurations have depended on either measuring the size of the set of possible senders of a particular message (the anonymity set size), or by measuring the entropy associated with the probability distribution of the messages possible senders. This paper explores further an alternative way of assessing the anonymity of a Mix system by considering the capacity of a covert channel from a sender behind the Mix to an observer of the Mix's output.

Initial work considered a simple model [4], with an observer (Eve) restricted to counting the number of messages leaving a Mix configured as a firewall guarding an enclave with one malicious sender (Alice) and some other naive senders (Clueless$_i$'s). Here, we consider the case where Eve can distinguish between multiple destinations, and the senders can select to which destination their message (if any) is sent each clock tick.

1 Introduction

In [4] the idea of measuring the lack of perfect anonymity (quasi-anonymity) via a covert channel was initiated. This idea was formalized in [5]. Our concern in this paper is to identify, and to calculate the capacity of, the covert channels that arise from the use of a Mix [8, 6] as an exit firewall from a private enclave (as briefly addressed in [4–Sec. 4].) In general, we refer to a covert channel that arises, due to a state of quasi-anonymity, as a quasi-anonymous channel [5]. The quasi-anonymous channel also serves the dual role of being a measure of the lack of perfect anonymity. [1] uses a similar model for statistical attacks in which Eve correlates senders' actions with observed output.

* Research supported by the Office of Naval Research.

D. Martin and A. Serjantov (Eds.): PET 2004, LNCS 3424, pp. 1–16, 2005.

2 Exit Mix-Firewall Model

There are $N+1$ senders in a private enclave. Messages pass one way from the private enclave to a set of M receivers. The private enclave is behind a firewall which also functions as a timed Mix [6] that fires every tick, t, hence we call it a simple timed Mix-firewall. For the sake of simplicity we will refer to a simple timed Mix-firewall as a Mix-firewall in this paper. One of the $N+1$ senders, called Alice, is malicious. The other N clueless senders, $\text{Clueless}_i, i = 1, \ldots, N$, are benign. Each sender may send at most one message per unit time t to the set of receivers. All messages from the private enclave to the set of receivers pass through public lines that are subject to eavesdropping by an eavesdropper called Eve. The only action that Eve can take is to count the number of messages per t going from the Mix-firewall to each receiver, since the messages are otherwise indistinguishable. Eve knows that there are $N+1$ possible senders. The N clueless senders act in an independent and identical manner (i.i.d.) according to a fixed distribution $C_i, i = 1, \ldots, N$. Alice, by sending or not sending a message each t to at most one receiver, affects Eve's message counts. This is how Alice covertly communicates with Eve via a quasi-anonymous channel [5].

Alice acts independently (through ignorance of the clueless senders) when deciding to send a message; we call this the *ignorance assumption*. Alice has the same distribution each t. Between Alice and the N clueless senders, there are $N+1$ possible senders per t, and there are $M+1$ possible actions per sender (each sender may or may not transmit, and if it does transmit, it transmits to exactly one of M receivers).

We consider Alice to be the input to the quasi-anonymous channel, which is a proper communications channel [9]. Alice can send to one of the M receivers or not send a message. Thus, we represent the inputs to the quasi-anonymous channel by the $M+1$ input symbols $0, 1, \ldots, M$, where $i = 0$ represents Alice not sending a message, and $i \in \{1, \ldots, M\}$ represents Alice sending a message to the ith receiver R_i. The "receiver" in the quasi-anonymous channel is Eve.

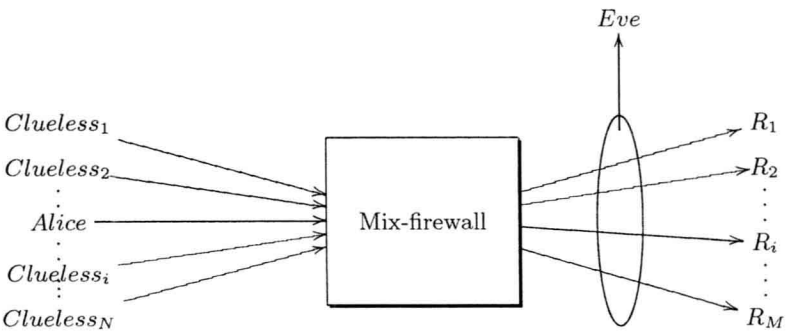

Fig. 1. Exit Mix-firewall model with N clueless senders and M distinguishable receivers

Eve receives the output symbols $e_j, j = 1, \ldots, K$. Eve receives e_1 if no sender sends a message. The other output symbols correspond to all the different ways the $N+1$ senders can send or not send, at most one message each, out of the private enclave, provided at least one sender does send a message.

For the sake of simplicity we introduce a dummy receiver R_0 (not shown above). If a sender does not send a message we consider that to be a "message" to R_0. For $N+1$ senders and M receivers, the output symbol e_j observed by Eve is an $M+1$ vector $\langle a_0^j, a_1^j, ..., a_M^j \rangle$, where a_i^j is how many messages the Mix-firewall sends to R_i. Of course it follows that $\sum_{i=0}^{M} a_i^j = N+1$.

The quasi-anonymous channel that we have been describing is a discrete memoryless channel (DMC). We define the channel matrix M as an $(M+1) \times K$ matrix, where M$[i,j]$ represents the conditional probability that Eve observes the output symbol e_j given that Alice input i. We model the clueless senders according to the i.i.d. C_i for each period of possible action t:

$$P(Clueless_i \; doesn't \; send \; a \; message) = p$$

$$P(Clueless_i \; sends \; a \; message \; to \; any \; receiver) = \frac{q}{M} = \frac{1-p}{M}$$

where in keeping with previous papers, $q = 1-p$ is the probability that Clueless$_i$ sends a message to any one of the M receivers. When Clueless$_i$ *does* send a message, the destination is uniformly distributed over the receivers $R_1, \ldots, R_M$. We call this the **semi-uniformity assumption**. Again, keep in mind that each clueless sender has the same distribution each t, but they all act independently of each other.

We model Alice according to the following distribution each t:

$$P(Alice \; sends \; a \; message \; to \; R_i) = x_i$$

Of course, this tells us that

$$x_0 = P(Alice \; doesn't \; send \; a \; message) = 1 - \sum_{i=1}^{M} x_i \; .$$

We let A represent the distribution for Alice's input behavior, and we denote by E the distribution of the output that Eve receives. Thus, the channel matrix M along with the distribution A totally determine the quasi-anonymous channel. This is because the elements of M take the distributions C_i into account, and M and A let one determine the distribution describing the outputs that Eve receives, P(Eve receives e_j).

Now that we have our set-up behind our exit Mix-firewall model, we may now go on to analyze various cases in detail. Additional cases and more detail are available in [7].

3 Capacity Analyses of the Exit Mix-Firewall Model

The mathematics of the problem gets quite complex. Therefore, we start with some simple special cases before attempting to analyze the problem in general.

The mutual information between A and E is given by

$$I(A,E) = H(A) - H(A|E) = H(E) - H(E|A) = I(E,A).$$

The capacity of the quasi-anonymous channel is given by [9]

$$C = \max_{A} I(A,E) \ ,$$

where the maximization is over the different possible values that the x_i may take (of course, the x_i are still constrained to represent a probability distribution). Recall $\mathrm{M}[i,j] = P(E = e_j | A = i)$, where $\mathrm{M}[i,j]$ is the entry in the i^{th} row and j^{th} column of the channel matrix, M. To distinguish the various channel matrices, we will adopt the notation that $\mathrm{M}_{N.M}$ is the channel matrix for N clueless senders and M receivers.

3.1 One Receiver ($M = 1$)

Case 1 — No clueless senders and one receiver ($N = 0,\ M = 1$)
Alice is the only sender, and there is only one receiver R_1. Alice sends either 0 (by not sending a message) or 1 (by sending a message). Eve receives either $e_1 = \langle 1,0 \rangle$ (Alice did nothing) or $e_2 = \langle 0,1 \rangle$ (Alice sent a message to the receiver). Since there is no noise (there are no clueless senders) the channel matrix M is the 2×2 identity matrix and it trivially follows that $P(E = e_1) = x_0$, and that $P(E = e_2) = x_1$.

$$\mathrm{M}_{0.1} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} e_1 & e_2 \end{array} \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{array}$$

Since $x_0 = 1 - x_1$, we see that[1] $H(E) = -x_0 \log x_0 - (1 - x_0) \log(1 - x_0)$. The channel matrix is an identity matrix, so the conditional probability distribution $P(E|A)$ is made up of zeroes and ones, therefore $H(E|A)$ is identically zero. Hence, the capacity is the maximum over x_0 of $H(E)$, which is easily seen to be unity[2] (and occurs when $x_0 = 1/2$). Of course, we could have obtained this capacity[3] without appealing to mutual information since we can noiselessly send one bit per tick, but we wish to study the non-trivial cases and use this as a starting point.

Case 2 — N clueless senders and one receiver ($M = 1$)
This case reduces to the *indistinguishable receivers* case with N senders. This is the situation analyzed in [4] with both an exit Mix-firewall that we have

[1] All logarithms are base 2.

[2] The units of capacity are bits per tick t, but we will take the units as being understood for the rest of the paper. Note that all symbols take one t to pass through the channel.

[3] This uses Shannon's [9] asymptotic definition of capacity, which is equivalent for noiseless channels (in units of bits per symbol).

been discussing and an entry Mix-firewall, with the receivers behind the latter. Alice can either send or not send a message, so the input alphabet again has two symbols. Eve observes $N+2$ possible output symbols. That is, Eve sees $e_1 = \langle N+1, 0\rangle$, $e_2 = \langle N, 1\rangle$, $e_3 = \langle N-1, 2\rangle$, $\cdots$, $e_{N+2} = \langle 0, N+1\rangle$. A detailed discussion of this case can be found in [4].

3.2 Some Special Cases for Two Receivers ($M = 2$)

There are two possible receivers. Eve has knowledge of the network traffic, so Alice can signal Eve with an alphabet of three symbols: 1 or 2, if Alice transmits to R_1 or R_2, respectively, or the symbol 0 for not sending a message. Let us analyze the channel matrices and the entropies for different cases of senders.

The symbol e_j that Eve receives is an 3-tuple of the form $\langle a_0^j, a_1^j, a_2^j\rangle$, where a_i^j is the number of messages received by i^{th} receiver.[4] The index $i = 0$ stands for Alice not sending any message. The elements of the 3-tuple must sum to the total number of senders, $N+1$,

$$\sum_{i=0}^{2} a_i = N + 1 \ .$$

Case 3 — No clueless senders and two receivers ($N = 0,\ M = 2$)
Alice is the only sender and can send messages to two possible receivers. The channel matrix is trivial and there is no anonymity in the channel.

$$\mathrm{M}_{0.2} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} \langle 1,0,0\rangle & \langle 0,1,0\rangle & \langle 0,0,1\rangle \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{array}$$

The subscript 0.2 represents one sender (Alice alone) and two receivers. The 3×3 channel matrix $\mathrm{M}_{0.2}[i, j]$ represents the conditional probability of Eve receiving the symbol e_j, when Alice sends to the Receiver i. '0' stands for not sending a message.

The mutual information I is given by the entropy $H(E)$ describing Eve

$$I(E, A) = H(E) = -x_1 \log x_1 - x_2 \log x_2 - (1 - x_1 - x_2)\log(1 - x_1 - x_2).$$

The capacity of this noiseless covert channel is $\log 3 \approx 1.58$ (at x_i=1/3, $i = 0, 1, 2$). This is the maximum capacity, which we note corresponds to zero anonymity.

[4] Recall that the a_i^j's of the output symbol are not directly related to A, which denotes the distribution of Alice.

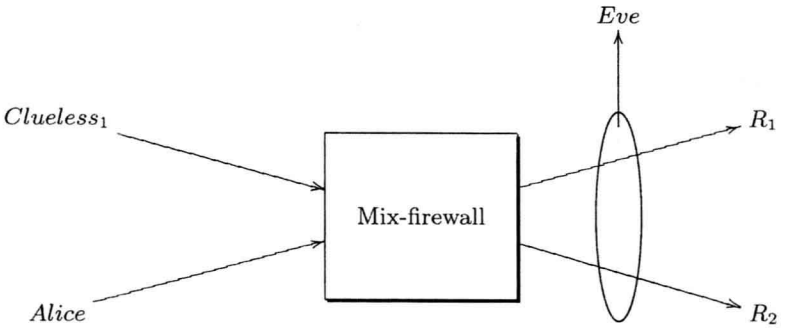

Fig. 2. Case 4: system with $N = 1$ clueless sender and $M = 2$ receivers

Case 4 — $N = 1$ clueless sender and $M = 2$ receivers
There are only six symbols that Eve may receive since there are six ways to put two indistinguishable balls into three distinct urns.

Let us consider the channel matrix.

$$\mathrm{M}_{1.2} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cccccc} \langle 2,0,0\rangle & \langle 1,1,0\rangle & \langle 1,0,1\rangle & \langle 0,2,0\rangle & \langle 0,1,1\rangle & \langle 0,0,2\rangle \end{array} \\ \begin{pmatrix} p & q/2 & q/2 & 0 & 0 & 0 \\ 0 & p & 0 & q/2 & q/2 & 0 \\ 0 & 0 & p & 0 & q/2 & q/2 \end{pmatrix} \end{array}$$

The 3×6 channel matrix $\mathrm{M}_{1.2}[i,j]$ represents the conditional probability of Eve receiving the symbol e_j when Alice sends to R_i. As noted, the dummy receiver R_0 corresponds to Alice not sending to any receiver (however this is still a transmission to Eve via the quasi-anonymous channel).

Given the above channel matrix we have:

$$\begin{aligned} H(E) = -\{ & px_0 \log[px_0] \\ & + [qx_0/2 + px_1] \log[qx_0/2 + px_1] \\ & + [qx_0/2 + px_2] \log[qx_0/2 + px_2] \\ & + [qx_1/2] \log[qx_1/2] + [qx_1/2 + qx_2/2] \log[qx_1/2 + qx_2/2] \\ & + [qx_2/2] \log[qx_2/2] \}. \end{aligned}$$

The conditional entropy is given by

$$H(E|A) = -\sum_{i=0}^{2} \left[p(x_i) \sum_{j=1}^{6} p(e_j|x_i) \log p(e_j|x_i) \right] = h_2(p) \ ,$$

where $h_2(p)$ denotes the function

$$\begin{aligned} h_2(p) \ &= \ -(1-p)/2 \log((1-p)/2) - (1-p)/2 \log((1-p)/2) - p \log p \\ &= -(1-p) \log((1-p)/2) - p \log p \ . \end{aligned}$$

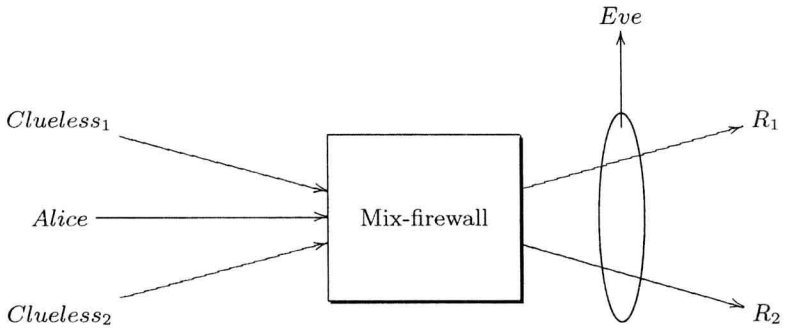

Fig. 3. Case 5: system with $N = 2$ clueless senders and $M = 2$ receivers

The mutual information between Alice and Eve is given by:

$$I(A, E) = H(E) - H(E|A) .$$

and the channel capacity is given by:

$$\begin{aligned} C &= \max_{A} I(A, E) \\ &= \max_{x_1, x_2} -\{px_0 \log[px_0] \\ &+[qx_0/2+px_1] \log[qx_0/2+px_1] \\ &+[qx_0/2+px_2] \log[qx_0/2+px_2] \\ &+[qx_1/2] \log[qx_1/2]+[qx_1/2+qx_2/2] \log[qx_1/2+qx_2/2] \\ &+[qx_2/2] \log[qx_2/2]\}-h_2(p). \end{aligned}$$

Note that the maximization is over x_1 and x_2, since x_0 is determined by these two probabilities (holds for any N). This equation is very difficult to solve analytically and requires numerical techniques. Figure 4 shows the capacity for this case with the curve labeled $N = 1$. From the plot the minimum capacity is approximately 0.92, when $p = 1/3$.

Case 5 — $N = 2$ clueless senders and $M = 2$ receivers

With two clueless senders and two receivers, Eve may receive ten symbols since there are ten different ways to put three indistinguishable balls into three distinct urns.

$$M_{2.2} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cccccccccc} \langle 3,0,0\rangle & \langle 2,1,0\rangle & \langle 2,0,1\rangle & \langle 1,2,0\rangle & \langle 1,1,1\rangle & \langle 1,0,2\rangle & \langle 0,1,2\rangle & \langle 0,3,0\rangle & \langle 0,2,1\rangle & \langle 0,0,3\rangle \end{array} \\ \begin{pmatrix} p^2 & pq & pq & q^2/4 & q^2/2 & q^2/4 & 0 & 0 & 0 & 0 \\ 0 & p^2 & 0 & pq & pq & 0 & q^2/4 & q^2/4 & q^2/2 & 0 \\ 0 & 0 & p^2 & 0 & pq & pq & q^2/2 & 0 & q^2/4 & q^2/4 \end{pmatrix} \end{array}$$

The 3×10 channel matrix $M_{2.2}[i, j]$ represents the conditional probability of Eve receiving e_j when Alice sends a message to receiver R_i.

Figure 4 shows the capacity for this case in the curve labeled $N = 2$. Again, the minimum capacity is found at $p = 1/3 = 1/(M + 1)$. From the plot the minimum capacity is approximately 0.62, when $p = 1/3$.