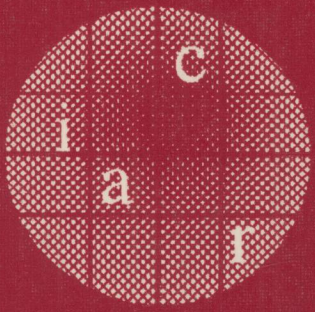
Christian Cachin
Jan Camenisch (Eds.)

LNCS 3027

Advances in Cryptology – EUROCRYPT 2004

International Conference on the Theory
and Applications of Cryptographic Techniques
Interlaken, Switzerland, May 2004, Proceedings



Springer

Christian Cachin Jan Camenisch (Eds.)

Advances in Cryptology - EUROCRYPT 2004

International Conference on the Theory
and Applications of Cryptographic Techniques
Interlaken, Switzerland, May 2-6, 2004
Proceedings

Springer

Volume Editors

Christian Cachin
Jan Camenisch
IBM Zurich Research Laboratory
Säumerstrasse 4, CH-8803 Rüschlikon, Switzerland
E-mail: {cca,jca}@zurich.ibm.com

Library of Congress Control Number: Applied for

CR Subject Classification (1998): E.3, F.2.1-2, G.2.1, D.4.6, K.6.5, C.2, J.1

ISSN 0302-9743
ISBN 3-540-21935-8 Springer-Verlag Berlin Heidelberg New York

Springer-Verlag is a part of Springer Science+Business Media

springeronline.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Boller Mediendesign
Printed on acid-free paper SPIN: 10999516 06/3142 5 4 3 2 1 0

Lecture Notes in Computer Science 3027

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

Editorial Board:

Springer
Berlin
Heidelberg
New York
Hong Kong
London
Milan
Paris
Tokyo

Preface

These are the proceedings of Eurocrypt 2004, the 23rd Annual Eurocrypt Conference. The conference was organized by members of the IBM Zurich Research Laboratory in cooperation with IACR, the International Association for Cryptologic Research.

The conference received a record number of 206 submissions, out of which the program committee selected 36 for presentation at the conference (three papers were withdrawn by the authors shortly after submission). These proceedings contain revised versions of the accepted papers. These revisions have not been checked for correctness, and the authors bear full responsibility for the contents of their papers.

The conference program also featured two invited talks. The first one was the 2004 IACR Distinguished Lecture given by Whitfield Diffie. The second invited talk was by Ivan Damgård who presented "Paradigms for Multiparty Computation." The traditional rump session with short informal talks on recent results was chaired by Arjen Lenstra.

The reviewing process was a challenging task, and many good submissions had to be rejected. Each paper was reviewed independently by at least three members of the program committee, and papers co-authored by a member of the program committee were reviewed by at least six (other) members. The individual reviewing phase was followed by profound and sometimes lively discussions about the papers, which contributed a lot to the quality of the final selection. Extensive comments were sent to the authors in most cases. At the end, the comments and electronic discussion notes filled more than 32,000 lines of text! We would like to thank the members of the program committee for their hard work over the course of several months; it was a pleasure for us to work with them and to benefit from their knowledge and insight. We are also very grateful to the external reviewers who contributed with their expertise to the selection process. Their work is highly appreciated.

The submission of all papers was done using the electronic submission software written by Chanathip Namprempre with modifications by Andre Adelsbach. During the review process, the program committee was mainly communicating using the Web-based review software developed by Bart Preneel, Wim Moreau, and Joris Claessens. We would like to thank Roger Zimmermann for his help with installing and running the software locally, and for solving many other problems, not the least of which was the assembly of these proceedings. The final decisions were made at a meeting in Rüschlikon at the IBM Zurich Research Laboratory. Helga Steimann helped us with the organization and also made sure there was enough coffee and food available so that we could concentrate on the papers and were not distracted by empty stomachs. Thanks a lot!

We are grateful to Endre Bangerter, Martin Hirt, Reto Strobl, and Roger Zimmermann for their help with the local arrangements of the conference.

Eurocrypt 2004 was supported by the IBM Zurich Research Laboratory, Crypto AG, Omnisec, MediaCrypt, HP, Microsoft Research, and Swiss International Air Lines.

Our most important thanks go to our families for bearing with us through this busy period, for their support, and for their love.

Last but not least, we thank all the authors from all over the world who submitted papers. It is due to them and their work that the conference took place.

February 2004 Christian Cachin and Jan Camenisch

EUROCRYPT 2004

May 2–6, 2004, Interlaken, Switzerland

Sponsored by the
International Association for Cryptologic Research (IACR)

in cooperation with the
IBM Zurich Research Laboratory, Switzerland

Program and General Chairs

Christian Cachin and Jan Camenisch
IBM Zurich Research Laboratory, Switzerland

Program Committee

Alex Biryukov ..... Katholieke Universiteit Leuven, Belgium
John Black ..... University of Colorado at Boulder, USA
Christian Cachin ..... IBM Zurich Research Laboratory, Switzerland
Jan Camenisch ..... IBM Zurich Research Laboratory, Switzerland
Jean-Sébastien Coron ..... Gemplus Card International, France
Claude Crépeau ..... McGill University, Canada
Ivan Damgård ..... Aarhus University, Denmark
Juan Garay ..... Bell Labs - Lucent Technologies, USA
Rosario Gennaro ..... IBM T.J. Watson Research Center, USA
Alain Hiltgen ..... UBS, Switzerland
Thomas Johansson ..... Lund University, Sweden
Antoine Joux ..... DCSSI Crypto Lab, France
Joe Kilian ..... NEC Laboratories America, USA
Arjen Lenstra ..... Citibank, USA & TU Eindhoven, The Netherlands
Yehuda Lindell ..... IBM T.J. Watson Research Center, USA
Anna Lysyanskaya ..... Brown University, USA
Daniele Micciancio ..... UC San Diego, USA
Omer Reingold ..... Weizmann Institute of Science, Israel
Vincent Rijmen ..... Cryptomathic and IAIK, Belgium
Phillip Rogaway ..... UC Davis, USA & Chiang Mai University, Thailand
Igor Shparlinski ..... Macquarie University, Australia
Edlyn Teske ..... University of Waterloo, Canada
Rebecca Wright ..... Stevens Institute of Technology, USA

External Reviewers

Adi Akavia
Joy Algesheimer
Jee Hea An
Siddhartha Annapureddy
Giuseppe Ateniese
Endre Bangerter
Lejla Batina
Amos Beimel
Mihir Bellare
Siddika Berna Ors
Simon Blackburn
Carlo Blundo
Alexandra Boldyreva
Dan Boneh
Colin Boyd
Xavier Boyen
An Braeken
Thomas Brochman
Ran Canetti
Scott Contini
Don Coppersmith
Nora Dabbous
Christophe De Cannière
Alex Dent
Giovanni Di Crescenzo
Christophe Doche
Yevgeniy Dodis
Patrik Ekdahl
Nelly Fazio
Serge Fehr
Marc Fischlin
Matthias Fitzi
Scott Fluhrer
Matt Franklin
Martin Gagne
Steven Galbraith
M. I. Gonzáles Vasco
Jens Groth
Jaime Gutierrez
Stuart Haber
Shai Halevi
Helena Handschuh
Darrel Hankerson
Danny Harnik
Jonathan Herzog
Florian Hess
Alejandro Hevia
Jason Hinek
Susan Hohenberger
Nicholas Hopper
Nick Howgrave-Graham
Jim Hughes
Yuval Ishai
Markus Jakobsson
Stas Jarecki
Eliane Jaulmes
Fredrik Jönsson
Marc Joye
Yael Tauman Kalai
Aggelos Kiayias
Neal Koblitz
David Kohel
Yoshi Kohno
Hugo Krawczyk
Ted Krovetz
Sébastien Kunz-Jacques
John Langford
Joseph Lano
Moses Liskov
Benjamin Lynn
Philip MacKenzie
Chip Martel
Alex May
Dominic Mayers
Ralph C. Merkle
Sara Miner
Ilya Mironov
Siguna Müller
Frédéric Muller
Sean Murphy
Chanathip Namprempre
Moni Naor
Mats Näslund
Phong Nguyen
Antonio Nicolosi
Svetla Nikova
Kobbi Nissim
Luke O'Connor
Roberto Oliveira
Pascal Paillier
Adriana Palacio
Kenneth Paterson
Souradyuti Paul
Thomas Pedersen
Chris Peikert
Erez Petrank
Birgit Pfitzmann
Benny Pinkas
David Pointcheval
Jonathan Poritz
John Proos
Michael Quisquater
Tal Rabin
Zulfikar Ramzan
Leonid Reyzin
Pierre-Michel Ricordel
Alon Rosen
Amit Sahai
Louis Salvail
Palash Sarkar
Jasper Scholten
Hovav Shacham
Taizo Shirai
Thomas Shrimpton
Alice Silverberg
Adam Smith
Patrick Solè
Jessica Staddon
Markus Stadler
Martijn Stam
Andreas Stein
Ron Steinfeld
Reto Strobl
Frédéric Valette
Bart Van Rompay
Luis von Ahn
Shabsi Walfish
Huaxiong Wang
Bogdan Warinschi
John Watrous
Christopher Wolf
Ke Yang

Lecture Notes in Computer Science

For information about Vols. 1–2898

please contact your bookseller or Springer-Verlag

Vol. 3027: C. Cachin, J. Camenisch (Eds.), Advances in Cryptology - EUROCRYPT 2004. XII, 628 pages. 2004.

Vol. 3025: G.A. Vouros, T. Panayiotopoulos (Eds.), Methods and Applications of Artificial Intelligence. XV, 546 pages. 2004. (Subseries LNAI).

Vol. 3015: C. Barakat, I. Pratt (Eds.), Passive and Active Network Measurement. XI, 300 pages. 2004.

Vol. 3011: J.-C. Régin, M. Rueher (Eds.), Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems. XI, 415 pages. 2004.

Vol. 3010: K.R. Apt, F. Fages, F. Rossi, P. Szeredi, J. Váncza (Eds.), Recent Advances in Constraints. VIII, 285 pages. 2004. (Subseries LNAI).

Vol. 3009: F. Bomarius, H. Iida (Eds.), Product Focused Software Process Improvement. XIV, 584 pages. 2004.

Vol. 3007: J.X. Yu, X. Lin, H. Lu, Y. Zhang (Eds.), Advanced Web Technologies and Applications. XXII, 936 pages. 2004.

Vol. 3006: M. Matsui, R. Zuccherato (Eds.), Selected Areas in Cryptography. XI, 361 pages. 2004.

Vol. 3005: G.R. Raidl, S. Cagnoni, J. Branke, D.W. Corne, R. Drechsler, Y. Jin, C.G. Johnson, P. Machado, E. Marchiori, F. Rothlauf, G.D. Smith, G. Squillero (Eds.), Applications of Evolutionary Computing. XVII, 562 pages. 2004.

Vol. 3004: J. Gottlieb, G.R. Raidl (Eds.), Evolutionary Computation in Combinatorial Optimization. X, 241 pages. 2004.

Vol. 3003: M. Keijzer, U.-M. O'Reilly, S.M. Lucas, E. Costa, T. Soule (Eds.), Genetic Programming. XI, 410 pages. 2004.

Vol. 3001: A. Ferscha, F. Mattern (Eds.), Pervasive Computing. XIII, 358 pages. 2004.

Vol. 2999: E.A. Boiten, J. Derrick, G. Smith (Eds.), Integrated Formal Methods. XI, 541 pages. 2004.

Vol. 2998: Y. Kameyama, P.J. Stuckey (Eds.), Functional and Logic Programming. X, 307 pages. 2004.

Vol. 2997: S. McDonald, J. Tait (Eds.), Advances in Information Retrieval. XIII, 427 pages. 2004.

Vol. 2996: V. Diekert, M. Habib (Eds.), STACS 2004. XVI, 658 pages. 2004.

Vol. 2995: C. Jensen, S. Poslad, T. Dimitrakos (Eds.), Trust Management. XIII, 377 pages. 2004.

Vol. 2994: E. Rahm (Ed.), Data Integration in the Life Sciences. X, 221 pages. 2004. (Subseries LNBI).

Vol. 2993: R. Alur, G.J. Pappas (Eds.), Hybrid Systems: Computation and Control. XII, 674 pages. 2004.

Vol. 2992: E. Bertino, S. Christodoulakis, D. Plexousakis, V. Christophides, M. Koubarakis, K. Böhm, E. Ferrari (Eds.), Advances in Database Technology - EDBT 2004. XVIII, 877 pages. 2004.

Vol. 2991: R. Alt, A. Frommer, R.B. Kearfott, W. Luther (Eds.), Numerical Software with Result Verification. X, 315 pages. 2004.

Vol. 2989: S. Graf, L. Mounier (Eds.), Model Checking Software. X, 309 pages. 2004.

Vol. 2988: K. Jensen, A. Podelski (Eds.), Tools and Algorithms for the Construction and Analysis of Systems. XIV, 608 pages. 2004.

Vol. 2987: I. Walukiewicz (Ed.), Foundations of Software Science and Computation Structures. XIII, 529 pages. 2004.

Vol. 2986: D. Schmidt (Ed.), Programming Languages and Systems. XII, 417 pages. 2004.

Vol. 2985: E. Duesterwald (Ed.), Compiler Construction. X, 313 pages. 2004.

Vol. 2984: M. Wermelinger, T. Margaria-Steffen (Eds.), Fundamental Approaches to Software Engineering. XII, 389 pages. 2004.

Vol. 2983: S. Istrail, M.S. Waterman, A. Clark (Eds.), Computational Methods for SNPs and Haplotype Inference. IX, 153 pages. 2004. (Subseries LNBI).

Vol. 2982: N. Wakamiya, M. Solarski, J. Sterbenz (Eds.), Active Networks. XI, 308 pages. 2004.

Vol. 2981: C. Müller-Schloer, T. Ungerer, B. Bauer (Eds.), Organic and Pervasive Computing – ARCS 2004. XI, 339 pages. 2004.

Vol. 2980: A. Blackwell, K. Marriott, A. Shimojima (Eds.), Diagrammatic Representation and Inference. XV, 448 pages. 2004. (Subseries LNAI).

Vol. 2978: R. Groz, R.M. Hierons (Eds.), Testing of Communicating Systems. XII, 225 pages. 2004.

Vol. 2977: G. Di Marzo Serugendo, A. Karageorgos, O.F. Rana, F. Zambonelli (Eds.), Engineering Self-Organising Systems. X, 299 pages. 2004. (Subseries LNAI).

Vol. 2976: M. Farach-Colton (Ed.), LATIN 2004: Theoretical Informatics. XV, 626 pages. 2004.

Vol. 2973: Y. Lee, J. Li, K.-Y. Whang, D. Lee (Eds.), Database Systems for Advanced Applications. XXIV, 925 pages. 2004.

Vol. 2972: R. Monroy, G. Arroyo-Figueroa, L.E. Sucar, H. Sossa (Eds.), MICAI 2004: Advances in Artificial Intelligence. XVII, 923 pages. 2004. (Subseries LNAI).

Vol. 2971: J.I. Lim, D.H. Lee (Eds.), Information Security and Cryptology -ICISC 2003. XI, 458 pages. 2004.

Vol. 2970: F. Fernández Rivera, M. Bubak, A. Gómez Tato, R. Doallo (Eds.), Grid Computing. XI, 328 pages. 2004.

Vol. 2964: T. Okamoto (Ed.), Topics in Cryptology – CT-RSA 2004. XI, 387 pages. 2004.

Vol. 2963: R. Sharp, Higher Level Hardware Synthesis. XVI, 195 pages. 2004.

Vol. 2962: S. Bistarelli, Semirings for Soft Constraint Solving and Programming. XII, 279 pages. 2004.

Vol. 2961: P. Eklund (Ed.), Concept Lattices. IX, 411 pages. 2004. (Subseries LNAI).

Vol. 2960: P.D. Mosses (Ed.), CASL Reference Manual. XVII, 528 pages. 2004.

Vol. 2958: L. Rauchwerger (Ed.), Languages and Compilers for Parallel Computing. XI, 556 pages. 2004.

Vol. 2957: P. Langendoerfer, M. Liu, I. Matta, V. Tsaoussidis (Eds.), Wired/Wireless Internet Communications. XI, 307 pages. 2004.

Vol. 2954: F. Crestani, M. Dunlop, S. Mizzaro (Eds.), Mobile and Ubiquitous Information Access. X, 299 pages. 2004.

Vol. 2953: K. Konrad, Model Generation for Natural Language Interpretation and Analysis. XIII, 166 pages. 2004. (Subseries LNAI).

Vol. 2952: N. Guelfi, E. Astesiano, G. Reggio (Eds.), Scientific Engineering of Distributed Java Applications. X, 157 pages. 2004.

Vol. 2951: M. Naor (Ed.), Theory of Cryptography. XI, 523 pages. 2004.

Vol. 2949: R. De Nicola, G. Ferrari, G. Meredith (Eds.), Coordination Models and Languages. X, 323 pages. 2004.

Vol. 2948: G.L. Mullen, A. Poli, H. Stichtenoth (Eds.), Finite Fields and Applications. VIII, 263 pages. 2004.

Vol. 2947: F. Bao, R. Deng, J. Zhou (Eds.), Public Key Cryptography – PKC 2004. XI, 455 pages. 2004.

Vol. 2946: R. Focardi, R. Gorrieri (Eds.), Foundations of Security Analysis and Design II. VII, 267 pages. 2004.

Vol. 2943: J. Chen, J. Reif (Eds.), DNA Computing. X, 225 pages. 2004.

Vol. 2941: M. Wirsing, A. Knapp, S. Balsamo (Eds.), Radical Innovations of Software and Systems Engineering in the Future. X, 359 pages. 2004.

Vol. 2940: C. Lucena, A. Garcia, A. Romanovsky, J. Castro, P.S. Alencar (Eds.), Software Engineering for Multi-Agent Systems II. XII, 279 pages. 2004.

Vol. 2939: T. Kalker, I.J. Cox, Y.M. Ro (Eds.), Digital Watermarking. XII, 602 pages. 2004.

Vol. 2937: B. Steffen, G. Levi (Eds.), Verification, Model Checking, and Abstract Interpretation. XI, 325 pages. 2004.

Vol. 2936: P. Liardet, P. Collet, C. Fonlupt, E. Lutton, M. Schoenauer (Eds.), Artificial Evolution. XIV, 410 pages. 2004.

Vol. 2934: G. Lindemann, D. Moldt, M. Paolucci (Eds.), Regulated Agent-Based Social Systems. X, 301 pages. 2004. (Subseries LNAI).

Vol. 2930: F. Winkler (Ed.), Automated Deduction in Geometry. VII, 231 pages. 2004. (Subseries LNAI).

Vol. 2929: H. de Swart, E. Orlowska, G. Schmidt, M. Roubens (Eds.), Theory and Applications of Relational Structures as Knowledge Instruments. VII, 273 pages. 2003.

Vol. 2926: L. van Elst, V. Dignum, A. Abecker (Eds.), Agent-Mediated Knowledge Management. XI, 428 pages. 2004. (Subseries LNAI).

Vol. 2923: V. Lifschitz, I. Niemelä (Eds.), Logic Programming and Nonmonotonic Reasoning. IX, 365 pages. 2004. (Subseries LNAI).

Vol. 2919: E. Giunchiglia, A. Tacchella (Eds.), Theory and Applications of Satisfiability Testing. XI, 530 pages. 2004.

Vol. 2917: E. Quintarelli, Model-Checking Based Data Retrieval. XVI, 134 pages. 2004.

Vol. 2916: C. Palamidessi (Ed.), Logic Programming. XII, 520 pages. 2003.

Vol. 2915: A. Camurri, G. Volpe (Eds.), Gesture-Based Communication in Human-Computer Interaction. XIII, 558 pages. 2004. (Subseries LNAI).

Vol. 2914: P.K. Pandya, J. Radhakrishnan (Eds.), FST TCS 2003: Foundations of Software Technology and Theoretical Computer Science. XIII, 446 pages. 2003.

Vol. 2913: T.M. Pinkston, V.K. Prasanna (Eds.), High Performance Computing - HiPC 2003. XX, 512 pages. 2003. (Subseries LNAI).

Vol. 2911: T.M.T. Sembok, H.B. Zaman, H. Chen, S.R. Urs, S.H. Myaeng (Eds.), Digital Libraries: Technology and Management of Indigenous Knowledge for Global Access. XX, 703 pages. 2003.

Vol. 2910: M.E. Orlowska, S. Weerawarana, M.M.P. Papazoglou, J. Yang (Eds.), Service-Oriented Computing - ICSOC 2003. XIV, 576 pages. 2003.

Vol. 2909: R. Solis-Oba, K. Jansen (Eds.), Approximation and Online Algorithms. VIII, 269 pages. 2004.

Vol. 2908: K. Chae, M. Yung (Eds.), Information Security Applications. XII, 506 pages. 2004.

Vol. 2907: I. Lirkov, S. Margenov, J. Wasniewski, P. Yalamov (Eds.), Large-Scale Scientific Computing. XI, 490 pages. 2004.

Vol. 2906: T. Ibaraki, N. Katoh, H. Ono (Eds.), Algorithms and Computation. XVII, 748 pages. 2003.

Vol. 2905: A. Sanfeliu, J. Ruiz-Shulcloper (Eds.), Progress in Pattern Recognition, Speech and Image Analysis. XVII, 693 pages. 2003.

Vol. 2904: T. Johansson, S. Maitra (Eds.), Progress in Cryptology - INDOCRYPT 2003. XI, 431 pages. 2003.

Vol. 2903: T.D. Gedeon, L.C.C. Fung (Eds.), AI 2003: Advances in Artificial Intelligence. XVI, 1075 pages. 2003. (Subseries LNAI).

Vol. 2902: F.M. Pires, S.P. Abreu (Eds.), Progress in Artificial Intelligence. XV, 504 pages. 2003. (Subseries LNAI).

Vol. 2901: F. Bry, N. Henze, J. Ma luszyński (Eds.), Principles and Practice of Semantic Web Reasoning. X, 209 pages. 2003.

Vol. 2900: M. Bidoit, P.D. Mosses (Eds.), Casl User Manual. XIII, 240 pages. 2004.

Vol. 2899: G. Ventre, R. Canonico (Eds.), Interactive Multimedia on Next Generation Networks. XIV, 420 pages. 2003.

Table of Contents

Foundations I

Identity-Based Encryption

Elliptic Curves

Signatures II

Public-Key Cryptography

Foundations II

Multiparty Computation

Cryptanalysis

New Applications

Algorithms and Implementation

Anonymity

Efficient Private Matching and Set Intersection

Michael J. Freedman[1]*, Kobbi Nissim[2]**, and Benny Pinkas[3]

[1] New York University
(mfreed@cs.nyu.edu)
[2] Microsoft Research SVC
(kobbi@microsoft.com)
[3] HP Labs
(benny.pinkas@hp.com)

Abstract. We consider the problem of computing the intersection of private datasets of two parties, where the datasets contain lists of elements taken from a large domain. This problem has many applications for online collaboration. We present protocols, based on the use of homomorphic encryption and balanced hashing, for both semi-honest and malicious environments. For lists of length k, we obtain $O(k)$ communication overhead and $O(k \ln \ln k)$ computation. The protocol for the semi-honest environment is secure in the standard model, while the protocol for the malicious environment is secure in the random oracle model. We also consider the problem of approximating the size of the intersection, show a linear lower-bound for the communication overhead of solving this problem, and provide a suitable secure protocol. Lastly, we investigate other variants of the matching problem, including extending the protocol to the multi-party setting as well as considering the problem of approximate matching.

1 Introduction

This work considers several two-party set-intersection problems and presents corresponding secure protocols. Our protocols enable two parties that each hold a set of inputs – drawn from a *large* domain – to jointly calculate the intersection of their inputs, without leaking any additional information. The set-intersection primitive is quite useful as it is extensively used in computations over databases, *e.g.*, for data mining where the data is vertically partitioned between parties (namely, each party has different attributes referring to the same subjects).

One could envision the usage of efficient set-intersection protocols for online recommendation services, online dating services, medical databases, and many other applications. We are already beginning to see the deployment of such applications using either trusted third parties or plain insecure communication.

* Research partially done while the author was visiting HP Labs.
** Research done while the author was at NEC Labs.

C. Cachin and J. Camenisch (Eds.): EUROCRYPT 2004, LNCS 3027, pp. 1–19, 2004.

Contributions. We study private two-party computation of set intersection, which we also denote as *private matching* (PM):

- Protocols for computing private matching, based on homomorphic encryption and balanced allocations: (i) a protocol secure against semi-honest adversaries; and (ii) a protocol, in the random oracle model, secure against malicious adversaries.[4] Their overhead for input lists of length k is $O(k)$ communication and $O(k \ln \ln k)$ computation, with small constant factors. These protocols are more efficient than previous solutions to this problem.
- Variants of the private matching protocol that (i) compute the intersection *size*, (ii) decide whether the intersection size is greater than a threshold, or (iii) compute some other function of the intersection set.
- We consider private approximation protocols for the intersection size (similar to the private approximation of the Hamming distance by [10]). A simple reduction from the communication lower-bound on disjointness shows that this problem cannot have a sublinear *worst-case* communication overhead. We show a sampling-based private approximation protocol that achieves instance-optimal communication.
- We extend the protocol for set intersection to a multi-party setting.
- We introduce the problem of secure approximate (or "fuzzy") matching and search, and we present protocols for several simple instances.

2 Background and Related Work

Private equality tests (PET). A simpler form of private matching is where each of the two datasets has a single element from a domain of size N. A circuit computing this function has $O(\log N)$ gates, and therefore can be securely evaluated with this overhead. Specialized protocols for this function were also suggested in [9, 18, 17], and they essentially have the same overhead. A solution in [3] provides fairness in addition to security.

A circuit-based solution for computing PM of datasets of length k requires $O(k^2 \log N)$ communication and $O(k \log N)$ oblivious transfers. Another trivial construction compares all combinations of items from the two datasets using k^2 instantiations of a PET protocol (which itself has $O(\log N)$ overhead). The computation of this comparison can be reduced to $O(k \log N)$, while retaining the $O(k^2 \log N)$ communication overhead [18]. There are additional constructions that solve the private matching problem at the cost of only $O(k)$ exponentiations [12, 8]. However, these constructions were only analyzed in the random oracle model, against semi-honest parties.

Disjointness and set intersection. Protocols for computing (or deciding) the intersection of two sets have been researched both in the general context of communication complexity and in the context of secure protocols. Much attention has been given to evaluating the communication complexity of the disjointness problem, where the two parties in the protocol hold subsets a and b of

[4] For malicious clients, we present a protocol that is secure in the standard model.