AFRL-OSR-VA-TR-2013-0205

Techniques for Secure and Reliable Computational Outsourcing

Mikhail Atallah
Purdue University

April 2013
Final Report

DISTRIBUTION A: Approved for public release.

AIR FORCE RESEARCH LABORATORY
AF OFFICE OF SCIENTIFIC RESEARCH (AFOSR)
ARLINGTON, VIRGINIA 22203
AIR FORCE MATERIEL COMMAND

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)* 27/02/2013	**2. REPORT TYPE** FINAL REPORT	**3. DATES COVERED** *(From - To)* 03/11/2009-03/11/2012

4. TITLE AND SUBTITLE

Techniques for Secure and Reliable Computational Outsourcing

5a. CONTRACT NUMBER

5b. GRANT NUMBER

FA9550-09-1-0223

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Mikhail Atallah

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Purdue University
Hovde Hall, 610 Purdue Mall, West Lafayette, IN 47907

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFOSR
875 N Randolph St
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

AFRL-OSR-VA-TR-2013-0205

12. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRIBUTION A: APPROVED FOR PUBLIC RELEASE

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Techniques were developed that make it possible to use remote servers without having to reveal to them either (i) the con
fidential inputs and outputs of computations; or (in the case of information storage and retrieval) the confidential data
and queries thereupon. The techniques also make cheating by the remote untrusted servers detectable; here cheating mean
s "not carrying out the expected computational and storage duties". Significant progress was also made in the direction
of hiding from the remote servers the access patterns to the encrypted data that they store, a potentially important cons
ideration in situations where it is not enough to hide the data (e.g., when the access patterns reveal too much about the
nature of how the data is being used). The progress in this area brings closer the day when remote cloud servers can be used for the most confidential tasks, without worry about confidentiality being compromised by security breaches occurring at the cloud service providers.

15. SUBJECT TERMS

Information security, cryptographic protocols, computational outsourcing, storage outsourcing, cloud computing

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Mikhail Atallah
U	U	U	SAR	7	**19b. TELEPHONE NUMBER** *(Include area code)* 765-463-7310

Reset

February 2013 Final Performance Report

Proposal title: Techniques for Secure and Reliable Computational Outsourcing

Attention: Dr. Robert L. Herklotz, Program Manager
Security and Information Operations
Air Force Office of Scientific Research

PI: Mikhail Atallah
Department of Computer Science
Purdue University
West Lafayette, IN 47906
mja@cs.purdue.edu

Co-PI: Marina Blanton
Department of Computer Science & Engineering
University of Notre Dame
Notre Dame, IN 46556
mblanton@cse.nd.edu

Abstract

Techniques were developed that make it possible to use remote servers without having to reveal to them either (i) the confidential inputs and outputs of computations; or (in the case of information storage and retrieval) the confidential data and queries thereupon. The techniques also make cheating by the remote untrusted servers detectable; here cheating means "not carrying out the expected computational and storage duties". Significant progress was also made in the direction of hiding from the remote servers the access patterns to the encrypted data that they store, a potentially important consideration in situations where it is not enough to hide the data (e.g., when the access patterns reveal too much about the nature of how the data is being used).

The contributions of the work can be categorized as either (i) being the first to achieve the confidentiality-preserving outsourcing for the computational and data structuring problems considered; or (ii) achieving significantly better performance than the previously published schemes for the problems considered. The progress in this area brings closer the day when

remote cloud servers can be used for the most confidential tasks, without worry about confidentiality being compromised by security breaches occurring at the cloud service providers.

Summary of Technical Results Achieved

The major results obtained are now briefly summarized, categorized according to the main theme of each. In what follows, we use *secure outsourcing* to refer to the use of remote servers that are not cleared to view any confidential data and computations, so that the client avails itself of their computational and storage without revealing anything to them about either the confidential inputs or outputs they helped compute.

Sequence comparisons

Protocols were given for securely outsourcing the most important of all distance metrics between two sequences: The edit distance, which, given two sequences x and y of respective lengths n and m, is the cost of a minimum-cost sequence of insertions, deletions, and substitutions that transform x into y. This computation is expensive, and securely outsourcing it is a significant achievement. The previous method of achieving this was far less efficient, both from a theoretical point of view and because it used homomorphic encryption. By utilizing garbled circuit evaluation techniques in a novel way, the new method avoids the use of public-key cryptography and uses only symmetric encryption. The advantages of the new scheme over the previous best known protocols for for this problem are summarized below.

- The client does only $O(m+n)$ work and communication, as opposed to the previous $O(mn)$.
- The round complexity has been reduced to 1, as opposed to the previous $O(mn)$.
- The space used at the servers has been reduced to $O(m + n)$, as opposed to the previous $O(mn)$.
- The cryptography used in the new scheme is only of the symmetric kind, whereas the previous used homomorphic encryption and oblivious transfer.

Biometrics

The first protocols for securely outsourcing biometric comparisons and searching were designed (for iris identification). The protocols were validated experimentally on a database of iris codes. This is important because, unlike passwords, biometrics cannot be modified if they are leaked to an adversary in digital form.

Information storage and retrieval

Novel techniques were designed for storing, at a remote server, an encrypted database such that confidentiality-preserving remote query processing by weak clients is supported even for complex queries. Techniques for hiding the query access patterns were also designed (hiding which encrypted data items are the target of the various queries). Significantly, the approach uses only inexpensive (symmetric) encryption.

Finite Automata

Protocols were given for the problem of secure outsourcing of error-resilient DNA searching via oblivious evaluation of finite automata, where a client has a DNA sequence, and a service provider has a pattern that corresponds to a genetic test. Error-resilient searching is achieved by representing the pattern as a nite automaton and evaluating it on the DNA sequence (which is treated as the input), where confidentiality of both the pattern and the DNA sequence must be preserved. The techniques are applicable to any type of finite automata (e.g., signature-based intrusion detection automata), but the optimizations were tailored to the setting of DNA searching.

Linear-algebra computations

Protocols were designed for a client to securely outsource expensive algebraic computations (like the multiplication of large matrices) to a remote server, such that the server learns nothing about the client's input or the result of the computation, and any attempted corruption of the answer by the server is detected with high probability. The computational work performed at the client was linear in the size of its input (which is unavoidable) and did not require the client to locally carry out any expensive encryptions of such input. The computational burden on the server was proportional to the time complexity of the best practically used algorithms for solving the algebraic problem (e.g., cubic time for multiplying two matrices). The improvements given include the option of using a single server, avoiding the use of any expensive cryptographic primitives (no homomorphic encryption), resilience to collusion between the remote servers (hence the ability to detect any attempt by the servers at collusive and coordinated corruption of the answer).

Algebraic computations over closed semi-rings

The above algebraic outsourcing techniques were significantly extended to no longer hinge on the existence of additive and multiplicative inverses for the familiar matrix multiplication over the (+,*) ring – they work when one (or both) of these inverses do not exist, as happens for many practically important algebraic structures (including closed semi-rings) when one or both of the two operations in the matrix multiplication is the ”min” or ”max” operation. Such matrix multiplications are very common in optimization. The protocols designed were for the cases of (+,min) multiplication, (min,max) multiplication, and of (min,+) multiplication; the last two cases are particularly important primitives in many combinatorial optimization problems.

Pattern matching in the Hamming distance with thresholds

An efficient solution was given to a significant generalization of the classic pattern matching problem, motivated by the situation where the entries in the text and pattern are analog, or distorted by additive noise, or imprecisely given for some other reason: In any alignment of the pattern with the text, two aligned symbols contribute 1 to the similarity score if they differ by no more than a given threshold, otherwise they contribute zero (the classic Hamming distance matching problem is the special case of zero threshold).

Storage of a total order relationship

Protocols were designed for storage outsourcing where is a total order on n items that are stored with a remote server called the dealer, and a user query consists of a pair of items whose relative ordering should be revealed along with a proof that the result is correct. The proof is generated using the dealer's local data (i.e., without bothering the data owner). The main difficulty was achieving efficient storage and query-processing while achieving the desiderata that (i) the user should learn nothing other than the answer to their query, and (ii) that a misbehaving dealer should not be able to convince a user of a wrong ordering. The scheme was generalized to partial orders that can be decomposed into a number of total orders, in which case a user either learns the ordering of the two queried items, or learns that they are incomparable.

Students

- Hao Yuan (Purdue)
  Ph.D. Thesis: Security and privacy techniques for outsourced and distributed databases
  http://docs.lib.purdue.edu/dissertations/AAI3413915/
  (the thesis contents were published in the articles co-authored by Hao Yuan that are listed in the next section)
- Timothy Duket (Purdue)
- Mehrdad Aliasgari (Notre Dame)
- Yihua Zhang (Notre Dame)

Refereed Journal and Conference Papers that Acknowledge the AFOSR Award

1. E. Aguiar, Y. Zhang, and M. Blanton, "An Overview of Issues and Recent Developments in Cloud Computing and Storage Security," Book chapter in High Performance Semantic Cloud Auditing, B.-Y. Choi, K. Han, and S. Song (Editors), Springer, 2013.
2. M. Blanton, M. Atallah, K. Frikken, and Q. Malluhi, "Secure and Efficient Outsourcing of Sequence Comparisons," European Symposium on Research in Computer Security (ESORICS 2012), pp. 505–522, Sep. 2012.
3. Mikhail J. Atallah, Keith B. Frikken, Shumiao Wang, "Private Outsourcing of Matrix Multiplication over Closed Semi-rings," SECRYPT 2012, pp. 136-144.
4. M. Blanton and M. Aliasgari, "Secure Outsourced Computation of Iris Matching," Journal of Computer Security, Vol. 20, No. 2–3, pp. 259–305, 2012.
5. Mohamed Yakout, Mikhail J. Atallah, Ahmed K. Elmagarmid, "Efficient and Practical Approach for Private Record Linkage," J. Data and Information Quality 3(3): 5 (2012).
6. Ashish Kundu, Mikhail Atallah, and Elisa Bertino, "Leakage-Free Redactable Signatures," Proc. ACM Conference on Data and Application Security and Privacy, San Antonio, Texas, February 2012, pp. 307-316.

7. Kai Christian Bader, Mikhail J. Atallah, Christian Grothoff, "Efficient relaxed search in hierarchically clustered sequence datasets," ACM Journal of Experimental Algorithmics 17(1) (2012).

8. Mehdi Bentounsi, Salima Benbernou, Cheikh S. Deme, Mikhail J. Atallah, "Anonyfrag: an anonymization-based approach for privacy-preserving BPaaS," Cloud-I 2012: 9.

9. Mehdi Bentounsi, Salima Benbernou, Mikhail J. Atallah, "Privacy-Preserving Business Process Outsourcing," ICWS 2012, pp. 662-663.

10. M. Blanton, Y. Zhang, and K. Frikken, "Secure and Verifiable Outsourcing of Large-Scale Biometric Computations," IEEE International Conference on Information Privacy, Security, Risk and Trust (PASSAT '11), pp. 1185–1991, Oct. 2011. (Full version of this paper is also under submission at TISSEC.)

11. Vinayak Deshpande, Leroy B. Schwarz, Mikhail J. Atallah, Marina Blanton, Keith B. Frikken, Outsourcing Manufacturing, "Secure Price Masking Mechanisms for Purchasing Component Parts," Production and Operations Management (POMS), 20 (2) (2011), pp. 165-180.

12. Marina Blanton, Everaldo Aguiar, "Private and Oblivious Set and Multiset Operations," IACR Cryptology ePrint Archive 2011: 464 (2011)

13. Keith Frikken. Hao Yuan, and Mikhail J. Atallah, "Secure Authenticated Comparisons," Proc. 9th Conference on Applied Cryptography and Network Security (ACNS 11), Nerja, Spain, June 2011, pp. 514-531.

14. Marina Blanton, Paolo Gasti, "Secure and Efficient Protocols for Iris and Fingerprint Identification," ESORICS 2011, pp. 190-209.

15. Marina Blanton, Mehrdad Aliasgari, "On the (Non-)reusability of Fuzzy Sketches and Extractors and Security in the Computational Setting," SECRYPT 2011, pp. 68-77.

16. Marina Blanton: Achieving Full Security in Privacy-Preserving Data Mining. SocialCom/PASSAT 2011, pp. 925-934.

17. Yuqing Sun, Qihua Wang, Ninghui Li, Elisa Bertino, Mikhail J. Atallah, "On the Complexity of Authorization in RBAC under Qualification and Security Constraints," IEEE Trans. Dependable Sec. Comput. 8(6), 2011, pp. 883-897.

18. Mikhail J. Atallah, Yinian Qi, and Hao Yuan, "Asymptotically Efficient Algorithms for Skyline Probabilities of Uncertain Data," ACM Transactions on Database Systems, 36 (2), 2011.

19. Hao Yuan, Mikhail J. Atallah, "Running Max/Min Filters Using 1+o(1) Comparisons per Sample," IEEE Trans. Pattern Anal. Mach. Intell. 33(12), 2011, pp. 2544-2548.

20. Mikhail J. Atallah, Timothy W. Duket, "Pattern matching in the Hamming distance with thresholds," Inf. Process. Lett. 111(14), 2011, pp. 674-677.

21. Paolo Falcarin, Christian S. Collberg, Mikhail J. Atallah, Mariusz H. Jakubowski, "Guest Editors' Introduction: Software Protection," IEEE Software 28(2), 2011, pp. 24-27.

22. M. Blanton and M. Aliasgari, “Secure Outsourcing of DNA Searching via Finite Automata,” Annual IFIP Conference on Data and Applications Security (DBSec ’10), pp. 49–64, Jun. 2010.

23. Emil Stefanov and Mikhail J. Atallah, “Duress Detection for Authentication Attacks Against Multiple Administrators,” Proc. 2010 ACM CCS Workshop on Insider Threats (WITS 10), Chicago, October 2010, pp. 37-46.

24. Hao Yuan and Mikhail J. Atallah, “Data Structures for Range Minimum Queries in Multidimensional Arrays,” Proc. 21st ACM-SIAM Symp. on Discrete Algorithms (SODA 10), Austin, Texas, January 2010, pp. 150-160.

25. Mikhail J. Atallah and Keith B. Frikken, “Securely Outsourcing Linear Algebra Computations,” Proc. of 5th ACM Symposium on Information, Computer and Communications Security (AsiaCCS 10), Beijing, China, April 2010, pp. 48-59.

26. Yinian Qi and Mikhail J. Atallah, “Identifying Interesting Instances for Probabilistic Skylines,” Proc. 21st International Conference and Workshop on Database and Expert Systems Applications (DEXA 10), Bilbao, Spain, August 2010, pp. 300-314.

27. Yuqing Sun, Qihua Wang, Ninghui Li, Elisa Bertino, and Mikhail Atallah, “On the Complexity of Authorization in RBAC under Qualification and Security Constraints,” IEEE Transactions on Dependable and Secure Computing, Sept. 2010.

28. Hao Yuan, Mikhail J. Atallah, “Data Structures for Range Minimum Queries in Multidimensional Arrays,” Proc. 21st ACM-SIAM Symp. on Discrete Algorithms (SODA 2010), pp. 150-160.

29. Hao Yuan, Mikhail J. Atallah, “Efficient and secure distribution of massive geo-spatial data,” Proc. 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (GIS 2009), pp. 440-443

30. Daniel G. Aliaga, Mikhail J. Atallah, “Genuinity Signatures: Designing Signatures for Verifying 3D Object Genuinity,” Proc. 30th Annual Conference of the European Association for Computer Graphics (Eurographics 09), pp. 437-446 (2009)

31. Mikhail J. Atallah, Yinian Qi, “Computing all skyline probabilities for uncertain data,” Proc. 28th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS 2009), pp. 279-287.

32. Marina Blanton, William M. P. Hudelson, “Biometric-Based Non-transferable Anonymous Credentials,” Proc. 11th International Conference on Information and Communications Security (ICICS 2009), pp. 165-180.

33. Hao Yuan, Mikhail J. Atallah, “Efficient data structures for range-aggregate queries on trees,” Proc. of 12th International Conference on Database Theory (ICDT 2009), pp. 111-120.

34. Keith B. Frikken, Marina Blanton, Mikhail J. Atallah, “Robust Authentication Using Physically Unclonable Functions,” Proc. 12th Information Security Conference (ISC 2009), pp. 262-277.

35. Mohamed Yakout, Mikhail J. Atallah, Ahmed K. Elmagarmid, “Efficient Private Record Linkage,” Proc. 25th International Conference on Data Engineering (ICDE 2009), pp. 1283-1286.