PAPER Special Section on Multi-dimensional Mobile Information Network

Security Issues in Mobile Information Networks

Thomas HARDJONO[†] *and* **Jennifer SEBERRY**[††], *Nonmembers*

SUMMARY During the last decade the decrease in the size of computing machinery, coupled with the increase in their computing power has lend to the development of the concept of mobile computing. Effects of this new vision is currently evident in the flourishing numbers of mobile telephones and portable computing units. In this paper we briefly investigate some issues concerning the security of mobile computing systems, within the framework of the categories of mobility, disconnection, data access modes and scale of operation (Imielinski & Badrinath, 1993). In contrast to previous works which concentrate on security in wireless communications, we focus on the security of interactions which are built upon the underlying wireless communications medium. Some conclusions are presented on the future directions for security research in mobile computing systems.

key words: *information security*

1. Introduction

The birth of "mobile computing" has signalled a new era in the field of computing and information systems. The concept of mobile computing is derived from the realization that as computing machinery decrease in size and increase in computing power users will demand these machinery to be part of their everyday life, accompanying them in the carrying-out of their everyday tasks. Researchers in this new field envisage that mobile computing units, such as today's laptops and palmtops, in the future will be communicating with each other via wireless networks, whilst providing location transparency to the user. This notion of transparency is carried-over from that in distributed computing, in which the user is unaware of the remote physical location of the resources being used by the distributed computing system. In the case of mobile computing, however, several differences emerge. The major difference–and in contrast to distributed computing–is precisely the mobility or the non-fixed positioning of some of the computing elements. This difference in itself presents various new challenges to researchers in the field.

The growing interest in mobile computing as a real possibility for the future has been driven partly by the popularity of mobile telephone systems, which to a certain extent has increased the desire on the part of the users to have computing available in a similarly mobile manner. The nomadic nature of some mobile computing elements have introduced new problems which were non-existent in the traditional areas of computing. New solutions and solutions derived from traditional computing are in demand in order to transform the dreams of mobile computing into a reality.

In the notable work of [2], four categories for future developments in mobile wireless computing have been proposed, namely *mobility*, *disconnection*, *data access modes* and *scale* of operation. These areas present challenges to the traditional approach to distributed computing, which until recently have not included effects of mobility into their design and applications. In this paper we would like to propose *security* as being the fifth category which intersects the four categories. Similar to the early research efforts during the last two decades in the area of traditional computing systems, currently much emphasis have been placed on deriving solutions which provide for acceptable performance and availability, with little stress on the need of security in such solutions.

Unlike previous works which concentrate on communications security in wireless networks [3], [4] such as the European GSM or U.S Digital Cellular (USDC), here we wish to focus more on the security of interactions which are built upon the underlying wireless communications medium. Each of the four categories of [2] introduces differing levels and ranges of security requirements, and in this paper we will briefly discuss some security issues within the framework of these categories. The current work is not aimed at providing a comprehensive investigation into all aspects of security within the mobile computing environment. Rather, this paper aims at identifying the broad areas within mobile computing which are likely to have immediate security concerns, and it hopes to provide some pointers as to the direction for future research in mobile computing security.

For simplicity and generality, we will use the terms *Mobile Unit* (MU) for the users and their portable computing unit, *Mobile Support Station* (MSS) for the stations which maintain communications with the mobile units and *Fixed Host* for the other nodes connected to

Manuscript received November 13, 1995.
Manuscript revised February 20, 1996.

[†]The author is with the Centre for Computer Security Research, University of Wollongong, NSW 2522, and also with the Department of Computing and Information Systems, University of Western Sydney–Macarthur, NSW 2560, Australia.

[††]The author is with the Centre for Computer Security Research, University of Wollongong, NSW 2522, Australia.

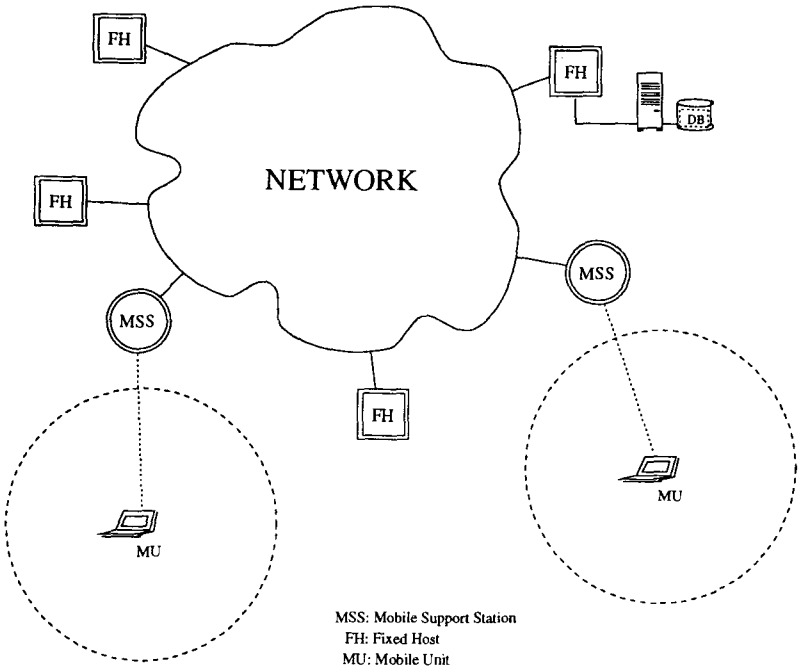

Fig. 1 The mobile computing network and its components (after[1]).

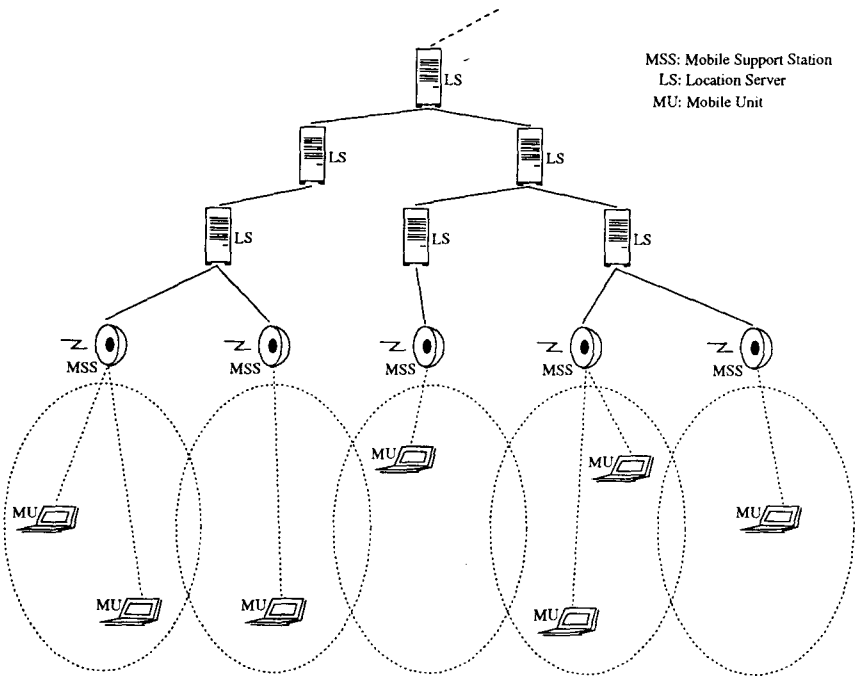

Fig. 2 Location servers in the mobile network.

the mobile support stations[1]. The basic network architecture is shown in Fig. 1. In addition, *Location Servers* (LS) are employed to keep track of the position and actions of each mobile unit. Hence, a given location server holds a database containing information about a mobile unit registered as having a "home-base" in the area or zone managed by the location server and about visitors in the area (cell). These location sever databases are used to manage and operate the mobile wireless network[5], and as such, they are not directly accessible to the mobile units. The data stored in these database are used and inter-exchanged only among location servers, and between location servers and mobile support stations (Fig. 2). The location servers typically correspond to mobile switching offices, with approximately 60 to 100 mobile support stations being catered for by one location server[6]

In the next section we discuss the security issues

pertaining to the mobility factor in mobile computing. This is followed in Sect. 3 with a discussion on the effects of disconnections on security. This followed by security issues relating to the new access modes within mobile computing in Sect. 4 and scale in Sect. 5. Finally, Sect. 6 closes the paper with some remarks and conclusions.

2. Mobility and Security

The fact that both users and the data that they carry have become a mobile component in computing has in itself introduced a set of security problems different to that in traditional computing. In the traditional case of fixed (non-mobile) computing physical protection could easily be afforded by making a computer and database system physically isolated from the other components in the environment. In such a configuration it was possible to make the system self-sufficient, without any need to communicate with the external world. More recent firewall techniques may also be applied to achieve the same effect.

In mobile computing this form of isolation and self-sufficiency is difficult to achieve due the relatively limited resources available to a mobile unit, thereby necessitating it to communicate with the mobile support station. The mobility of users and the data that they carry introduces security problems from the point of view of the *existence* and *location* of a user (which is deemed to be data in themselves) and from the point of view the privacy and authenticity of the data exchanged between users and between a user and a fixed host. More specifically, a user on a mobile wireless network may choose to have the information concerning his or her existence treated as being confidential. That is, a user may choose to remain anonymous to the majority of other users on the network, with the exception of a select number with whom the user often interacts.

This problem of user anonymity in mobile computing is related to a more difficult problem of the *trust* level afforded by each node in the wireless network and the problem of the security of location data concerning a user when the location data is stored or transferred between nodes as the user moves in a nomadic fashion. These nodes must provide some assurance to the user about his or her anonymity, independent of the differing levels of trust that may exist for each node. This requirement is of particular importance in the case of a user that crosses between two zones which are under two nodes respectively, each having a different trust level. Equally important is the secure transfer of data between databases at nodes which hold location data and other information or parameters in the user profile. Here all traffic internal to the network and transparent to the nomadic user must be maintained private and authentic. Any notion of trust in this case must be clearly expressed and formulated (such as that in [7]), and new formulations may be required specifically for security in the mobile computing environment.

Another potential security problem lies in the possibility of information leakage through the inference made by an attacker masquerading as a mobile support station with or without the aid of a subverted mobile support station. The attacker which masquerades as a mobile support station may issue a number of queries to the database at the user's home node or to database at other nodes, with the aim of deducing parts of the user profile containing the patterns and history of the user's movements. Here again, security techniques are required both for the databases and for the identification of users and mobile support stations. Any scheme to be used must ensure that any queries submitted to a given database at a user's home-base is accompanied by sufficient proof that the user approves of the queries submitted by the (foreign) mobile support station controlling the zone under which the user is currently roaming or passing through. It is, therefore, not unreasonable to assume that some method of delegation of rights [8] will be employed between the user and the mobile support stations (and fixed hosts) in the network.

Related to the management of these databases and the provision of performance transparency [6] for the nomadic user is the issue of the replication of certain parameters and user profiles with the aim of replicating the environments surrounding the user. Thus, as the user roams across zones, the user must not experience a degradation in the access and latency times. Such degradations will reduce the mobile unit response time to the user, and thus affect the work of the user. Reasonable response times must be maintained to provide mobility transparency to the user. Again, security must be considered in the context of replication, both from the trust level of the mobile support stations and fixed hosts and from the point of view of data leakage. In general, as sensitive data is replicated across several sites, the security risks are also increased due to the multiplication of the points of attack [9].

3. Security in Disconnections

Another major issue in mobile computing which arises from mobility and power (battery) restrictions is disconnections. The disconnection of a mobile unit from a mobile support station is necessary for the conservation of power of the mobile unit. A mobile unit can typically find itself running on a temporary form of power supply (e.g. spare battery) while its main power source is being recovered (ie. recharged). In this situation differing levels of disconnection may be introduced, ranging from the normal connection to connections using low bandwidth channels.

A crucial aspect of disconnection is the *elective* or *non-elective* nature of a disconnection [2]. The non-elective disconnection refers to the cases when a mo-

bile unit disconnects due to an unforeseen event, such as system crashes or total communications break-down when moving into certain geographic regions. The elective disconnection refers to the user who deliberately chooses or elects to disconnect. That is, the planned and desired disconnection on the part of the owner of the mobile unit who wishes to temporarily limit remote access to his or her mobile unit. This type of disconnection typically covers instances where power availability is degrading, where the owner wishes to employ the maximum resources (e.g. CPU) available for a short task, or where the owner simply desires the mobile unit to be put into a "sleep" state.

In both types of disconnections, a number of potential security loop-holes may be introduced. The transition from one level of disconnection to another may present an opportunity for an attacker to masquerade either the mobile unit or the mobile support station. Any disconnection transition scheme must ensure that an attacker cannot mimic the mobile unit and then present the mobile support station with a false cancellation command with regard to the disconnection. That is, an attacker should not be able to "hi-jack" the communications of a mobile unit which is stepping-down its level of connection and then masquerading as the mobile unit. Similarly, an attacker must not masquerade as a mobile support station to a mobile unit that is about to step-up its level of connection.

One possible solution to this problem is for the mobile unit and the mobile support station to agree upon a *secret* before any transition in the levels of elective disconnections. When at a later time the mobile unit wishes to upgrade the connection level with the mobile support station, both parties can start by exchanging the shared secret in a secure manner. A good candidate for a solution to this problem is zero-knowledge protocols, in which the mobile unit and the mobile support station can convince each other that they hold the shared secret without having to transmit the full secret [10]. Other schemes based on smartcards may also be a good candidate (e.g. [11]).

Another potential security problem related to disconnections is the leaking of information through the use of inappropriate concurrency control techniques in the database of the mobile unit. During the past few years research work in the area of secure concurrency control techniques have indicated that *covert channels* exist in many standard concurrency control algorithms [12], [13]. A covert channel is created in multilevel database systems when one user can indirectly control the abort/commit patterns exhibited by the transaction scheduler. These patterns are then experienced by other transactions belonging to a second user. In this manner, the first user can leak data to the second user who may not actually have access to the data readable by the first user. If a database in a mobile unit is to be accessible by a remote user in the form of the remote user submitting small transactions, then the potential for covert channels will exist unless the mobile unit employs a suitable and secure concurrency control algorithm.

Related to this issue is the problem of incomplete transactions caused by elective and non-elective disconnections. A number of scenarios present themselves in this context. Security and integrity problems may occur in the case when hand-offs occur between two mobile support stations as the mobile unit crosses zones (cells). Other security problem may occur when a mobile unit deposits a "timed" transaction at a fixed-node, which begins to execute when certain conditions are met (e.g. time of day, availability of raw data) and which transmits the results back to the nomadic mobile unit. These scenarios represent only a few potential security problems among many others in the context of levels of disconnections and transaction management in mobile computing.

4. Secure Data Access Methods

One of the advantages of mobile communications derives from the possible use of broadcasting techniques to provide services to varying sizes of audience groups (users) with minimal change in the delivery cost of the services. The work of [14] identifies two methods of delivering information to the mobile unit of users by a broadcast server, namely through *data broadcasting* and *interactive requests*. The possibility of continuous broadcasting of ever-changing data lends to the attractive notion of data broadcasting being a public form of "memory", where mobile units periodically refresh their limited memories (caches) using "data on the air".

Two important parameters related to the broadcasting of data are *access time* and *tuning time*, the first referring to the time taken for a reply to be received by a client (mobile unit) from the broadcast server, the later referring to the amount of time taken by the client in "listening" to the channel in order to obtain the selected data. Here, the mobile unit will first listen to an "index channel" that delivers a directory related to the broadcasted data, then it will proceed to use the directory information as a guide as to when the mobile unit should access the stream of data. Ideally, the mobile unit should remain in "doze" mode until the required data is being broadcasted, at which time it should automatically wake itself up from this mode. Such a scenario is within reach of today's technology (with the minor exception of the battery technology that must still be improved). However, there are a number of issues related to the privacy, authenticity and integrity of the broadcasted data that need to be addressed and solved.

The first and foremost is the authentication of the source of the broadcast (e.g. broadcast server) by the mobile units. Since such a broadcast may carry public data whose accuracy is paramount (e.g. stock exchange

data) and whose authority for publication (e.g. NYSE) is accepted by the community, source authentication and integrity becomes crucial as minor inaccuracies–accidental or deliberate–may result in great losses on the part of the users. Therefore, methods are needed for initial source authentication by the mobile units, and for the periodic source re-authentication by the mobile units in such a way that it consumes less power than the initial authentication. Such methods would be attractive if they use (cryptographic) parameters which are embedded within the stream of broadcast data and which are recognizable at a given time only by the mobile units that require to perform re-authentication at that time. This will also lend to the ability of targeting re-authentication procedures for different groups of users (mobile units) at different moments in time.

Together with source authentication comes the need to maintaining the integrity of the broadcast data stream. Here, a number of possible attack scenarios present themselves, one being the denial of service attack ranging from crude channel interference to the sophisticated modifications of the index channel resulting in the mobile unit listening to incorrect (useless or undesirable) parts of the broadcast stream. More sophisticated attacks may even attempt to substitute segments from both the index channel and the data stream in such a way that the mobile unit is unaware of the attack.

The notion of a continuous broadcast of data being a public "memory"[14] together with the limited physical memory (cache) available at the mobile unit leads to the important issue of trust accorded by the users to the source of broadcast. Since the "public memory" will become a fleeting entity of a short lifetime, accountability of the source and the auditability of the broadcast data becomes necessary to prevent fraud when the broadcasted data carries commercial value and has commercial impact on its recipients.

5. Scaling Security

The specific issues relating to security increases in complexity as various components within the mobile network increase in number and in their mode of interaction. The increase in the number of mobile units and their wider geographic distribution across regional and political boundaries will result in the need for new solutions specific to mobile computing. The potential for the proliferation of mobile units may result in the need for increase in the size and capacity of the infrastructure supporting the network.

With the increase in number and geographic distribution of mobile units, some basic security functionalities will be required to be provided by the Mobile Support Stations and Location Servers. Examples of these functionalities include large scale key distribution and key management solutions, the provision of secrecy and authentication across large geographic boundaries with minimal delay, and the secure management of parts of the mobile network which are under different management bodies (national and international). International security policies to regulate trans-border data flows will also need to be established as nomadic mobile units wander in and out of countries and sensitive regions.

Currently a number of solutions exist in the area of distributed computing systems for the provision of security in a globalized computing environment. Hierarchical key generation, distribution and management techniques already exist which are aimed for the traditional stationary computing systems[15],[16]. Extensions to these techniques for the mobile environment may be a first step towards satisfying the security needs of mobile computing. However, new solutions will also need to be designed and implemented if security is to scale properly in the mobile network.

6. Remarks and Conclusion

There is still a long way for research to proceed before mobile computing will become a daily reality in society. Although considerable effort is being focused towards research in mobile computing, much of it is concentrating on the performance and availability of mobile computing, with comparatively little attention being given to the security issues in such an environment.

In this paper we have proposed *security* to be a major category for future developments in mobile computing. We have discussed briefly the issues of security in the context of mobility, disconnection and data access methods, presenting a number of potential problems in the security of a mobile computing environment.

The mobile computing environment and its security presents a new ground for further research, with some problems which are non-existent in the traditional non-mobile computing environment. Future work on the security of mobile computing must address the problems pertaining to the security of information within the three sub-areas of the mobile environment:

- The security of information residing in the mobile units, and the correctness and integrity of data in these mobile units.

- The security of information as it travels "over the air" between mobile units and mobile support stations. An important consideration in this area is the power consumption of the algorithms and schemes that implement this secure data transfer. New secure data storage schemes and data organization techniques will be required to facilitate rapid searching and transfer of data to and from mobile units.

- The security of information within the mobile wireless network. This includes the security of

databases holding control data used for the operations and management of the mobile wireless network.

These three sub-areas of research will be crucial if mobile computing is to be a reality in the future.

Acknowledgement

This work has been supported in part by the Australian Research Council (ARC) under the reference number A49232172 and the University of Wollongong *Computer Security: Technical and Social Issues* research program. The second author has received additional funding from the ARC under the reference numbers A49130102 and A49131885.

References

[1] T. Imielinski and B.R. Badrinath, "Mobile wireless computing: Solutions and challenges in data management," Communications of the ACM, vol.37, no.10, pp.18–28, 1994.
[2] T. Imielinski and B.R. Badrinath, "Data management for mobile computing," SIGMOD RECORD, vol.22, no.1, pp.34–39, 1993.
[3] D. Brown, "Security planning for personal communications," Proc. the 1st ACM Conference on Computer and Communications Security, pp.107–111, ACM Press, 1993.
[4] M.J. Beller, L.-F. Chang, and Y. Yacobi, "Privacy and authentication on a portable communications system," IEEE Journal on Selected Areas in Communications, vol.11, no.6, pp.821–829, 1993.
[5] W. van den Broek and E. Buitenwerf, "Distributed databases for third generation mobile systems," Proc. the International Council for Computer Communication Intelligent Networks Conference, ed. P.W. Bayliss, pp.333–347, IOS Press, Tampa, Florida, 1992.
[6] B.R. Badrinath and T. Imielinski, "Replication and mobility," Proc. the 2nd IEEE Workshop on Management of Replicated Data, pp.9–12, IEEE, Nov. 1992.
[7] Department of Defense, "Trusted Computer System Evaluation Criteria–Orange Book," Pub. DOD 5200.28-STD, U.S. Department of Defense, 1985.
[8] B. Lampson, M. Abadi, M. Burrows, and E. Wobber, "Authentication in distributed systems: Theory and practice," Technical Report 83, Digital Systems Research Center, Feb. 1992.
[9] L. Gong, "Increasing availability and security of an authentication service," IEEE Journal on Selected Areas in Communications, vol.11, no.5, pp.657–662, 1993.
[10] J. Seberry and J. Pieprzyk, "Cryptography: An Introduction to Computer Security," Prentice Hall, Sydney, 1989.
[11] T. Leighton and S. Micali, "Secret-key agreement without public-key cryptography," Advances in Cryptology–Proceedings of Crypto '93, ed. D.R. Stinson, vol.773 of Lecture Notes in Computer Science, pp.456–479, Springer-Verlag, 1993.
[12] T.F. Keefe, W.T. Tsai, and J. Srivastava, "Multilevel secure database concurrency control," Technical Report TR 89-45, University of Minnesota, July 1989.
[13] S. Jajodia and B. Kogan, "Transaction processing in multilevel-secure databases using replicated architecture," Proc. the 1990 IEEE Symposium on Security and Privacy, pp.360–368, IEEE Computer Society, Oakland, CA, 1990.
[14] T. Imielinski, S. Viswanathan, and B.R. Badrinath, "Data on the air-organization and access," technical report, Department of Computer Science, Rutgers University, NJ, 1992. (available from ftp `riches.rutgers.edu`).
[15] A.D. Birrell, B.W. Lampson, R.M. Needham, and M.D. Schroeder, "A global authentication service without global trust," Proc. the 1986 IEEE Symposium on Security and Privacy, pp.156–172, IEEE Computer Society, Oakland, CA, 1986.
[16] J.J. Tardo and K. Alagappan, "SPX: Global authentication using public-key certificates," Proc. the 1991 IEEE Symposium on Research in Security and Privacy, pp.232–244, IEEE Computer Society, Oakland, CA, 1991.

Thomas Hardjono obtained the B.Sc. (Hons) degree in Computer Science from The University of Sydney and the Ph.D. degree also in Computer Science from The University of New South Wales, Australia, in 1987 and 1991 respectively. From 1992 to 1993 he was a Research Scientist at the ATR Communication Systems Research Laboratories, in Kyoto, Japan. Currently he is a Research Fellow at the Centre for Computer Security Research at the University of Wollongong, Wollongong, Australia. He also lectures at the University of Western Sydney – Macarthur. His research interests includes security in network and distributed systems, mobile computing security, database security and cryptography. He has published over 50 conference and journal papers in the area of security. Thomas Hardjono is a member of the ACM, the IEEE and the IACR, and also of the *Standards Australia* Committee IT 12/4 on Computer Security Techniques.

Jennifer Seberry was born in Sydney, Australia in 1944. She received the B.Sc. degree in mathematics from the University of New South Wales, Sydney, Australia in 1966, the M.Sc. and Ph.D. degrees in mathematics from La Trobe University, Melbourne, Australia in 1969 and 1971, respectively. She was the first woman awarded the Ph.D. degree at La Trobe University. She worked for English Electric-LEO Computers before graduating from New South Wales. She also has worked as a programmer and teacher both at high school and university levels. She has lectured in many European, Asian, North American, and Australian cities and worked in Canada, the United States and India. In 1988, she founded the Centre for Communication Security Research funded by Telecom, Australia, to carry out research in different areas of Security and where she has been a director ever since. She and her colleagues have recently received considerable media attention for their research into "anti-hacking systems" or "User Unique Identification." Prof. Seberry has coauthored six books and has published over 200 scholary papers in international journals and conference proceedings. Her research interests include Hadamard matrices, finite mathematics, cryptography, and computer security. She is a member of the IACR, the IEEE and the ACM.