Computers & Security, 14 (1995) 465-472

0167-4048(95)00011-9

Refereed paper

Applications of smartcards for anonymous and verifiable databases

Thomas Hardjono and
Jennifer Seberry

Centre for Computer Security Research, Department of Computer Science, University of Wollongong, Wollongong, NSW 2522, Australia

In this paper we describe a practical solution towards anonymous and verifiable databases based on the use of smartcards and the recent Improved Leighton–Micali protocol for the distribution of keys. The scheme is addressed particularly to public data held in separate government databases with the aim of preventing unauthorized government institutions from gathering and merging private data concerning individuals from these separate containers. The solution can be realized through the recent Clipper Chip and smartcard technology, and its security relies on the strength of these technologies. The scheme is also extendible to mobile computing environments.

Keywords: Security and protection, Database management, Network protocols.

1. Introduction

Security of public data represents an issue which is increasingly becoming important and relevant to all individuals within society. Public data can range from statistics which bear no direct impact on any given individual in society, to medical and financial information whose disclosure may affect an individual's standing within society. In traditional paper-based societies the gathering of such personal information concerning a particular individual was difficult to perform due to the shear amount of manual work involved. Hence, only certain government bodies could afford such data gathering based on some legal warrant.

In today's computerized world the collection and transfer of voluminous amounts of information over wide geographic distances has been accepted as a common everyday occurrence. With recent advances in fibre optics technology, the notion of superhighways for data is becoming a reality. With this increasing ease at which voluminous data can be transferred and the increasing speed of data processing systems, the capacity for data gathering and intelligent computerized processing has also been significantly increased. These advances, which are beneficial to society from one point of

0167-4048/95/$9.50 © 1995, Elsevier Science Ltd

view, have raised questions from the opposite point of view, namely, of whether such computing power can be misused against society both by certain individuals within the society and by the very government upon which members of society have placed their trust. Accepting that for the functioning of a society some trust must be placed by the society on its government, a method of assurance must still be used to guarantee that an individual's personal details, which are spread across different government institutions, cannot be illegally gathered and merged together to give a total picture of that individual's private life.

One of the earliest efforts directed into finding possible solutions to this problem is that by Brandt *et al.* [1]. This effort recognized that databases belonging to different institutions must provide to the individual users the properties of the users being *anonymous* and the databases being *verifiable*. More specifically, when different data items are given by an individual to these distinct and separate institutions, these data items should not be identifiable by others as having come from the one same individual. The true identity of each individual must remain unknown to other individuals and to each institution. Each individual must also have the ability to verify that his or her personal details held by an institution are correct.

In the current work we investigate the issue of anonymous and verifiable databases in the context of recent developments in tamper-proof hardware, with the aim of presenting some practical solutions to the need of such databases. Our approach is founded on the use of smartcard technology coupled with an improved version of the recent Leighton–Micali key distribution protocol [2, 3]. In the next section we discuss the motivations of the current work, with the aim of placing Brandt's work [1] in our context. This is followed in Section 3 by the description of the scheme, focusing on the Improved Leighton–Micali protocol and on the issues of anonymity, data storage and data verifiability. Section 4 closes with some brief remarks and conclusions.

2. Motivations

The need of a practical scheme to realize the notion of anonymous and verifiable databases is becoming self-evident in computerized nations. One recent example in Australia was the public debate over the Australian Identity Card [4] by which every Australian resident would be assigned a unique number as an identifying piece of information. This number would then be used as a pointer to cross-reference data in various government institutions which held information concerning the owner of the number. Although this move by the government was defeated, in actuality the government later proceeded to use the citizens' taxation file numbers more or less as a substitute for the proposed identity card.

One important recent development in the United States which has again brought the debate about citizens' right to privacy into the foreground is the introduction of the Clipper Chip [5, 6] and its related technology. The Clipper Chip is a high-speed and high-security encryption device to be used by the US Government for its telephone and other networking equipment. The chip has a classified encipherment algorithm and contains a secret key. Through a 'key-escrow' system an appointed government agency can obtain a legal warrant to wiretap communications between any two parties that are using the device. The main idea behind this notion is to provide secure communications to the users of the Chip against external attacks, while at the same time allowing the government to monitor communications that are suspected of being a threat to national security or to society in general (e.g. drug traffickers, industrial espionage).

This paper extends the notions embodied in the Clipper Chip concept towards another area, namely that of providing ways to achieve anonymous and verifiable databases. We require the appointed agency or authority to be a trusted adjudicator between the members of society and the other ordinary government institutions. In this

way sensitive data concerning citizens in general may be guarded against illegal access while data concerning suspected citizens can be made readily available to the appointed authority. In the following discussion we will denote the appointed government agency as the *Trusted Authority* (TA). We assume that each institution holds a database containing every individual's details which are relevant to the functioning of the institution. Any exchange of data between departments must be through the Trusted Authority who regulates as to which details are exchangeable and who enforces the chosen policies. Thus, for example, the taxation department holds taxation-related information, while the health department has a health record of individuals that obtain medical service from the government hospitals. An individual is able to submit new details to each institution, respectively, and each individual can query each database independently without his or her identity being revealed.

Each individual has the duty to initially enrol himself or herself to the Trusted Authority, bringing their personal identification information (e.g. birth certificate, retina scan, fingerprint, DNA sequence). The Trusted Authority creates a *pseudonym* [1] for an individual corresponding to each institution that holds data about the individual. Hence, an individual has a different pseudonym when dealing with each institution. For each individual, the Trusted Authority issues a tamper-proof smartcard containing that individual's set of pseudonyms and other cryptographic parameters. For a given institution, the Trusted Authority also issues cryptographic parameters which are stored in the tamper-proof smartcard belonging to an appointed trusted local authority (person) who is a representative of the institution (e.g. system administrator). Unlike the identity of individuals, each institution has a unique identity which is published.

The database at each institution is assumed to be managed by a trusted DBMS which can be used by staff members at the site only through a number of tamper-free terminals [7, 8]. These tamper-free terminals represent the only valid access points to the database. A number of tamper-free terminals are also provided at the site for use by visiting individuals in the public, while remote tamper-free terminals may also be connected provided that a secure channel can be created between the remote tamper-free terminals and the local tamper-free terminals. The appointed representative for an institution has the duty to periodically load the cryptographic parameters from his or her smartcard to each of the resident tamper-free terminals at that institution. This configuration is shown in Fig. 1.

3. Towards a practical scheme

In this section we present a practical scheme for anonymous and verifiable databases based on the Improved Leighton–Micali (ILM) protocol [3]. The original Leighton–Micali protocol [2] had an inherent flaw which in our context allowed an attacker to read data belonging to an individual when it was in transit between the institution and the individual's terminal. This flaw has subsequently been solved and the protocol improved Zheng's work [3].[1]

Following the requirements of [2, 3], we assume that tamper-proof VLSI chips are readily available to be incorporated into smartcards and tamper-free terminals. We also assume that a publicly known one-way hash function h exists (which may also be replaced with a cryptographically strong pseudo-random function).

When an individual wishes to submit data to an institution or to verify existing data held by an institution, he or she must interact via a tamper-free terminal which establishes a connection with another tamper-free terminal located at the institution. Communications between these two terminals must be via a session key which is selected by either terminal and is transferred securely to

[1]The work of [3] has recently been published in [9].

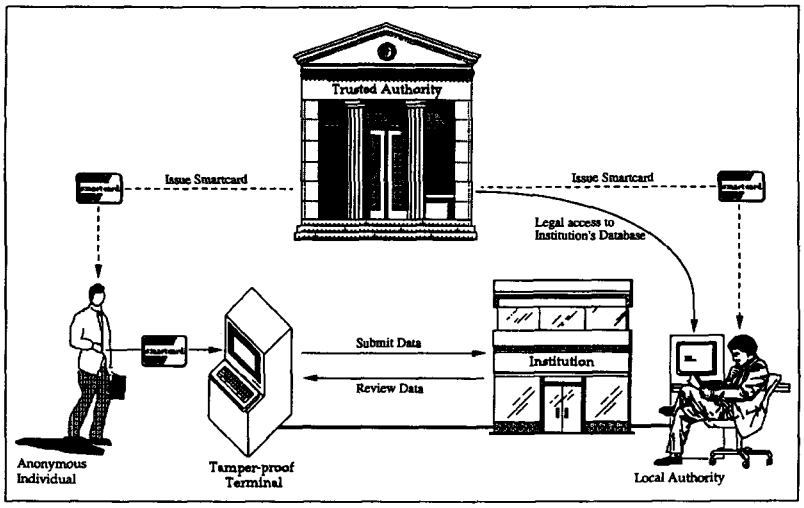

Fig. 1. Anonymous individuals and verifiable databases.

the other. The session key is then discarded by both terminals after the session is over. Newly submitted data is assumed to be placed in a temporary location within the institution's database to be read, verified and classified by one of the institution's staff members. Only then can such data be committed to the database. In the following, we assume that all communications are protected against replays (e.g. via timestamps or nonces).

3.1 Session keys: the ILM protocol

In the ILM protocol it is assumed that the Trusted Authority holds M secret keys $(X_1, \ldots, X_M)$. Each secret key is chosen uniformly at random by the Trusted Authority and is of length k bits. For each user i who is enrolled into the system, the Trusted Authority selects M random integers $(\alpha_1, \ldots, \alpha_M)$ from the interval $[1, L]$, where L is an integer.[2]

The Trusted Authority then employs h to compute $Y_m = h^{\alpha_m}(X_m)$ for all $m = 1, \ldots, M$, where $h^s(X)$ indicates applying consecutively the function h on an input X for s times. That is,

$$h^s(X) = \underbrace{h(\ldots h(h(X))\ldots)}_{s \text{ times}}.$$

Here $(\alpha_1, \ldots, \alpha_M)$ becomes the public key of individual i which is known to all institutions. The corresponding secret key $(Y_1, \ldots, Y_M)$ is then placed by the Trusted Authority into the tamper-proof chip of the individual's smartcard. The smartcard is only delivered to the individual after the secret key has been inserted, and hence no person has access to the secret keys or the other secret parameters within that individual's tamper-proof chip.

Assuming that individual i needs to verify or submit data to institution j, then their respective terminals must establish a secure channel by way of encipherment using key $K_{i,j}$. This secure channel will afterwards be used to transfer the random session key K_s. The crucial requirement at this point is that both terminals must establish the same key $K_{i,j}$ independently *without* previous communications. This process can be done as follows

[2]Leighton and Micali recommended the size of M be of the order O $(B^3 \log N)$, where N is the number of users and B is the upper-bound on the number of dishonest users [2,9].

[3]:

1. The terminal of individual i must obtain the public key $(\beta_1, \ldots, \beta_M)$ of institution j. This public key can be resident within each tamper-free terminal or it can be read by the terminal from a publicly readable file.

2. After individual i inserts his or her smartcard into the terminal, the terminal must provide the smartcard with the public key of the institution j. The tamper-proof chip within the smartcard of individual i then computes the common key $K_{i,j}$ as:

$$K_{i,j} = \begin{cases} h(h^{\delta_1}(X_1) \,||\, h^{\delta_M}(X_M) \,||\, i \,||\, j), & i \leq j, \\ h(h^{\delta_1}(X_1) \,||\, h^{\delta_M}(X_M) \,||\, j \,||\, i), & i > j, \end{cases} \tag{1}$$

where $\delta_m = \max(\alpha_m, \beta_m)$, $m = 1, \ldots, M$ and $||$ denotes concatenation.

Note that the tamper-proof chip can easily compute $h^{\delta_m}(X_m) = h^{|\delta_m - \alpha_m|}(Y_m)$ $(m = 1, \ldots, M)$ and thus $K_{i,j}$ because it has available the values $Y_m = h^{\alpha_m}(X_m)$ residing in its internal memory. The tamper-proof chip of the terminal at the institution performs symmetric procedures, and thus obtains the same key $K_{j,i} = K_{i,j}$ [3].

In order to aid our subsequent discussions we will simplify eq. (1) into

$$K_{i,j} = \begin{cases} h(X \,||\, i \,||\, j), & i \leq j, \\ h(X \,||\, j \,||\, i), & i > j. \end{cases} \tag{2}$$

As before, the k-bit value X is chosen randomly by the Trusted Authority where k should be sufficiently large, say $k \geq 100$, in order to prevent it from an exhaustive search attack [3]. The value X is kept secret by the Trusted Authority, and during the enrolment of individuals the Trusted Authority also injects a copy of X into the chip of the smartcard belonging to the individual and into that belonging to the local authority at each institution. Hence, in fact, the value X is common to all parties in the system.

3.2 Anonymity

In order to provide anonymity to individuals within the system, the Trusted Authority must create distinct pseudonyms for each user with respect to each of the institutions. In order to do this each individual i must enrol in-person to the Trusted Authority and provide it with some identification information P_i. The Trusted Authority uniformly chooses a unique identity I_i and a unique secret value S_i, and associates them with P_i. It is the duty of the Trusted Authority to keep the values (P_i, I_i, S_i, SC_i) secure, where SC_i is the unique chip number built into the tamper-proof chip of the individual's smartcard. The same procedure is also observed for the local authority of each institution.

Assuming that each institution has been assigned a publicly known identity B_j, the Trusted Authority creates the pseudonym $I_{i,j}$ of the individual with identity I_i with respect to B_j as:

$$I_{i,j} = h(S_{\mathrm{TA}} \,||\, I_i \,||\, B_j).$$

Here we assume that the encoding scheme for the identities of individuals I_i and institutions B_j is uniform. The key S_{TA} is maintained as secret by the Trusted Authority. The secret value S_i and the pseudonyms for an individual are then inserted into that individual's chip. Similarly, each institution is given the respective pseudonym that the individual will present to the institution.

Another secret parameter injected into the tamper-proof chips of both the individual $I_{i,j}$ and the institution B_j is a database key $D_{i,j}$, uniformly chosen by the Trusted Authority. This database key will be used to create other keys which are further used to control access to the database. Thus, for example, these created keys can be used

to hide passwords of individuals, to encipher the access matrix or to encipher the data in the database. In this paper we will use them to encipher stored data, although it is clear to the reader that other modes of their usage are possible.

3.3 Data storage

Within each institution B_j data in the database concerning individual $I_{i,j}$ must be stored in such a way that only the individual and the institution (i.e. its staff) can view the data. Assuming $R_{i,j}$ represents the data of individual $I_{i,j}$ at institution B_j, a key $K_{R_{i,j}}$ must be uniformly chosen by the local authority within the institution to be applied in order to hide data $R_{i,j}$. Bearing in mind that a secure DBMS running above a secure operating system is crucial for overall system security, there are a number of ways in which data can be stored in a manner that will make it accessible to users only through the key $K_{R_{i,j}}$. One simple method through which data can be protected from unwanted disclosures is by way of direct encipherment using the above key $K_{R_{i,j}}$ (see [10]).

This encipherment key $K_{R_{i,j}}$ must also reside in an enciphered form under a key which is available to the individual. This key-enciphering key is calculated by the staff's terminal at the institution as:

$$KE_{i,j} = \begin{cases} h(X \,||\, D_{i,j} \,||\, I_{i,j} \,||\, B_j), & I_{i,j} \leq B_j, \\ h(X \,||\, D_{i,j} \,||\, B_j \,||\, I_{i,j}), & I_{i,j} > B_j. \end{cases} \quad (3)$$

In addition, for each entry in the database belonging to the individual $I_{i,j}$ a signature or checksum [11, 12] must be created to prevent undetected changes to the data without the individual's consent. This is achieved by using the values S_i and S_j which are in the tamper-proof chips of the individual and the institution's local authority, respectively.

One simple way to create such signatures or checksums is as follows. When the individual is requested by the institution to verify and approve the data $R_{i,j}$ about the individual to be committed into the database, tamper-proof chips of the respective parties must generate certain parameters as input to some secure signature function *sig*. Thus, the chip belonging to the individual $I_{i,j}$ creates $t_{i,j} = h(R_{i,j} \,||\, S_i \,||\, B_j)$, while the chip belonging to the institution's local authority creates $t_{j,i} = h(R_{i,j} \,||\, S_j \,||\, I_{i,j})$ (note that here $t_{i,j} \neq t_{j,i}$). The two terminals onto which the individual and local authority are connected obtain the respective values from the chip within the respective (inserted) smartcards, and then the terminals exchange $t_{i,j}$ and $t_{j,i}$ over the secure channel established previously using the session key.

After receiving $t_{i,j}$ from the individual's terminal, the institution's terminal then computes the signature for the individual's entry. That is, the entry for individual $I_{i,j}$ within the database of institution B_j is

$$\{K_{R_{i,j}}\}_{KE_{i,j}}, \{R_{i,j}\}K_{R_{i,j}}, sig(R_{i,j}, t_{i,j}, t_{j,i}),$$

where the symbol '$\{\}_K$' denotes encipherment using key K.

The institution's terminal then sends this complete entry (including the signature) and $R_{i,j}$ to the individual's terminal which re-computes the signature (if needed, the entry and the signature can also be sent by the individual's terminal to a lawyer who represents the individual). If both signatures are identical, the individual's terminal sends an acknowledgement to the terminal at the institution. Both terminals then erase the values $t_{i,j}$ and $t_{j,i}$. In this manner, neither the individual nor the institution can modify the data illegally, since neither $t_{i,j}$ nor $t_{j,i}$ are ever directly available to the individual or the institution's staff, respectively.

Here we have illustrated a simple method for the storage of data in an institution's database, focusing on the ease of access to the data from remote tamper-free terminals. Other secure data storage methods can be devised, and the reader is directed

to [10–12] for a more comprehensive discussion on this particular issue.

3.4 Verifiability

When an individual $I_{i,j}$ wishes to view his or her data $R_{i,j}$ held at an institution B_j the individual must use his or her smartcard with a tamper-free terminal:

1. The individual then selects via the tamper-free terminal the identity of the institution B_j that holds the data the individual wishes to view.

2. After inserting his or her smartcard into the tamper-free terminal, the terminal provides the smartcard — and thus the chip within — with the identity B_j. The individual's terminal must also indicate to the institution's tamper-free terminal that a session is being requested. The institution's terminal then looks up the identity $I_{i,j}$ of the individual.

3. The individual's chip then computes $K_{i,j}$, while the chip within the institution's terminal computes $K_{j,i}$. As before, $K_{i,j} = K_{j,i}$.

4. The individual's terminal (or the institution's terminal) generates a session key K_s. The session key K_s is then exchanged by way of enciphering it with $K_{i,j} = K_{j,i}$.

5. The institution's terminal then instructs the database system to return the entry $\{K_{R_{i,j}}\}_{KE_{i,j}}$, $\{R_{i,j}\}_{KR_{i,j}}$, $sig(R_{i,j}, t_{i,j}, t_{j,i})$ for individual $I_{i,j}$. This entry is enciphered using the session key K_s and the result is dispatched to the individual's terminal.

6. The individual's terminal deciphers the entry using the session key K_s, and the key-enciphering key $KE_{i,j}$ is recreated following eq. (3). The individual's terminal then recovers $K_{R_{i,j}}$ and uses it to decipher and present to the individual the data $R_{i,j}$. The integrity of the data may also be verified by way of recreating the signature in the manner previously discussed. This would involve the institution's terminal re-computing $t_{j,i}$ and sending it to the individual's terminal via a secure channel.

4. Remarks and conclusion

In this paper we have extended the notions embodied in the Clipper Chip concept towards achieving anonymous and verifiable databases. The Trusted Authority creates a pseudonym for an individual corresponding to each institution that holds data about the individual. Hence, an individual has a different pseudonym when dealing with each institution. For each individual, the Trusted Authority issues a tamper-proof smartcard containing that individuals's set of pseudonyms and other cryptographic parameters. The database at each institution is assumed to be managed by a trusted DBMS which can be used by staff members at the site only through a number of tamper-free terminals. These tamper-free terminals represent the only valid access points to the database. A number of tamper-free terminals are also provided at the site for use by individuals in the public, while remote tamper-free terminals may also be connected, provided a secure channel can be created between the remote tamper-free terminals and the local tamper-free terminals. When an individual wishes to submit data to an institution or to verify existing data held by an institution, he or she must interact via a tamper-free terminal which establishes a connection with another tamper-free terminal located at the institution. Communications between these two terminals must be via a session key which is selected by either terminal and is transferred securely to the other. The session key is then discarded by both terminals after the session is over.

The security of the scheme relies on the tamper-resistance of the chips and the randomness of the one-way hash function. To reduce the risk of abusing stolen chips, authentication of a chip's owner should be conducted by such means as user password [3]. In its current stage the scheme does not pretend to cover all possible points of

attack, and clearly it does not provide a balanced burden of trust between an individual and an institution. After all, it is the institution that maintains the database containing the individual's private information. In practice it is difficult to prevent an institution from creating an informal and separate 'black list' database containing 'off-the-record' information upon which in reality it bases its decisions concerning a given individual. Other security measures are also required to prevent staff members of an institution from sharing data illegally with other institutions (e.g. manually copying onto a removable hard disk). The scheme in this paper represents a first step towards providing a practical mechanism in the face of an emerging new technology.

Acknowledgements

We thank Yuliang Zheng for bringing the ILM protocol to our attention. We also thank Azad Jiwa for comments and support. This work was supported in part by the Australian Research Council under the reference numbers A49130102, A9030136, A49131885 and A49232172.

References

[1] J. Brandt, I.B. Damgard and P. Landrock, Anonymous and verifiable registration in databases, in C.G. Gunther (ed.), *Advances in Cryptology—Proc. EUROCRYPT '88 (Lecture Notes in Computer Science No. 330)*, Springer-Verlag, 1988, pp. 167–176.

[2] T. Leighton and S. Micali, Secret-key agreement without public-key cryptography, in D.R. Stinson (ed.), *Advances in Cryptology—Proc. Crypto '93* (Vol. 773 of *Lecture Notes in Computer Science*), Springer-Verlag, 1993, pp. 456–479.

[3] Y. Zheng, Amending Leighton and Micali's key distribution protocol, *Technical Report Preprint 93-17*, Centre for Computer Security Research, Computer Science Department, University of Wollongong, Sept. 1993.

[4] E. Smith, *The Australia Card: the story of its defeat*, Sun, South Melbourne, 1989.

[5] White House Press Release, The Clipper Chip initiative, *IEEE Inf. Theory Soc. Newsl.*, 43(2) (1993) 21–23.

[6] DOC/NIST, A proposed Federal information processing standard for an escrowed encryption standard (EES), *Fed. Reg.*, 58(145) (1993).

[7] J.K. Omura, A computer dial access system based on public-key techniques, *IEEE Commun. Mag.*, 25(7) (1987) 73–79.

[8] P.J. Lee, Secure user access control for public networks, in J. Seberry and J. Pieprzyk (eds), *Advances in Cryptology—Proc. AUSCRYPT '90 (Sydney) (Lecture Notes in Computer Science No. 453)*, Springer-Verlag, Jan. 1990, pp. 46–57.

[9] Y. Zheng, On key agreement protocols based on tamper-proof hardware, *Inf. Process. Lett.*, 53 (1995) 49–54.

[10] D.E. Denning, Field encryption and authentication, in D. Chaum (ed.), *Advances in Cryptology: Proc. CRYPTO '83*, Santa Barbara, CA, pp. 231–247, Plenum Press, New York, 1983.

[11] D.E. Denning, Cryptographic checksums for multilevel data security, in *Proc. 1984 IEEE Symposium on Security and Privacy*, Oakland, CA), IEEE Computer Society, April 1984, pp. 52–61.

[12] D.E. Denning, Commutative filters for reducing inference threats in multilevel database systems, in *Proc. 1985 IEEE Symposium on Security and Privacy*, Oakland, CA, IEEE Computer Scoeity, April 1985, pp. 134–146.