Validation of 1-N OT Algorithms in Privacy-Preserving Investigations

Zbigniew Kwecka, Prof William J. Buchanan and Duncan Spiers
School of Computing, Napier University, Edinburgh, UK
z.kwecka@napier.ac.uk
w.buchanan@napier.ac.uk
d.spiers@napier.ac.uk

Abstract:
Most organisations relay on digital information systems (ISs) in day-to-day operations, and often sensitive data about employees and customers are stored in such systems. This, effectively, makes ISs enhanced surveillance measures, witch can reach further than CCTV monitoring and provide valuable resources for internal and external investigations. For privacy reasons, if a digital forensic investigation is to take place, only the investigators should know the identities of the suspects. Ideally, the investigators should not have to disclose these identities to the data holders, while the data holders, i.e. organisations whose data subjects are being investigated, should not have to disclose their full databases to investigators. The only data that should be disclosed should relate to that involving the subject – thus the need for a privacy-preserving investigation system. Several privacy-preserving algorithms have been proposed, but most of them are only of theoretical interest since empirical evaluations have rarely been undertaken. The main novelty in this paper is that it applies a 1-out-n Oblivious Transfer (1-n OT) algorithm to a new area of privacy-preserving investigations. Hence, an implementation of a straightforward privacy-preserving investigation system that can be used in real-life applications is outlined. The system uses tried and tested encryption algorithms: RSA for hiding the identity of the suspect; AES to conceal from investigators records not relating to the suspect; and commutative RSA to allow discovery of index where a suspect's data is stored in the third party records. This paper outlines an initial evaluation of the system proving that it may be successfully used in digital forensic investigations, conducted by public authorities and private organisations alike. The empirical evaluation also shows that the time required by this system to run grows in line with increasing number of records and increasing size of records, which is desirable compared to exponential growth observed in many systems that employ 1-n OT protocols .

Keywords: privacy preservation, data mining, digital forensics, digital suspect watchlist, Oblivious Transfer

1. Introduction

Most organisations relay on some form of digital ISs. These are usually used to record information about customers, employees and their interaction. However, such systems are often capable of, and used for, recording actions performed by the users, which makes these perfect surveillance measures. Thus, an organisation may record activities of system users, whether they are working in the office, or using software or equipment belonging to the organisation, while they are 'on the go' (Vlahos 2008). The majority of the population accepts monitoring as a required safeguard against threats to the organisations concerned and do not consider such method as a direct threat to their personal privacy (Gill & Spriggs 2005). At the same time, many people worry about the privacy of their digital records and the unfair use of their data (Young et al. 2006). Thus, there are two issues that need to be addressed to ensure that the privacy of law-abiding individuals is not affected by the surveillance systems in operation. The first issue relates to the fact that data obtained for legitimate reasons is often used for other purposes which could be deemed as intensifying surveillance and invasion of privacy (Ball et al. 2006). Arguably, the best course of action is to amend legislation accordingly. The second issue is linked to privacy concerns in access to the data collected, and this paper proposes a protocol to address it. In other words, it relates to internal and external investigators who request information on a third party's data subject, in such a way that traditional investigations would typically invade privacy of the data subject (Frikken & Atallah 2003). The following scenario should illustrate this clearly:

> Government security services, following a lead, investigate several potential suspects, including law-abiding Bob. Investigators need to identify Bob as a suspect in order to legally require a third party, an on-line shop where Bob made some recent purchases, to release the data. Later, Bob is removed from the list of suspects based on the evidence gathered. However, a few days later, Bob wants to use his favourite on-line shop again, but this time requests 0% finance on the purchase. His application is refused. What Bob does not know is that he was a suspect, and the on-

line shop has placed him on the list of high-risk borrowers, because of the recent enquiry from the security services. Most importantly, Bob is unable to find out why his application was refused, since disclosure of matters affecting national security and crime prevention are exempt from many provisions of the UK's Data Protection Act 1998 (DPA) (sections 28 and 29). Given also that the DPA was enacted in implement of the European Data Protection Directive 95/46/EC, similar exemptions apply across the European Economics Area.

This scenario demonstrates an invasion of Bob's human and natural rights, and, in this case, the party that caused the violation are the security services, as their actions made a third-party aware that he is a suspect in an investigation. According to the DPA, organisations may provide other organisations with personal and sensitive personal information about a data subject in some exceptional circumstances (see Part IV of the DPA). Additionally, in the UK, The Regulation of Investigatory Powers Act 2000 (RIPA) and The Regulation of Investigatory Powers (Scotland) Act 2002 (RIPSA) gave even wider regulated powers to investigators. For instance, emergency services may request information on the allergies of a casualty, and of a casualty's relatives, from any institution that they suspect may have this data, and such institutions may lawfully disclose the data. Accordingly, the police and other public authorities may also request data related to their suspects, based on the same reasoning. Thus, in the above scenario, the security services and the data controlling organisation would act lawfully in accordance with the above legislation. However, their actions could seriously impinge upon Bob's natural rights and quite possibly his privacy. In the scenario it has a very particular detrimental impact upon his rights concerning his future economic relations with the data controlling organisation. This raises interesting issues about the legal remedies, if any, open to him. In similar circumstances a case could be made out that there has been a breach of Article 8 of the Council of Europe's Convention on Human Rights (now enforceable in the United Kingdom under the Human Rights Act 1998). This would be difficult to pursue for a number of reasons. Quite apart from the practical difficulty of knowing that there has been a breach of rights, how the breach has come about, who is responsible and how to prove it (what might be called "evidential difficulties"), there is also the question of the extent to which those responsible might be able to claim exemption from responsibility (which might be called "substantive difficulties"). The right of privacy under Article 8 (like most human rights) is a qualified right, meaning that a public authority is entitled to disregard the right where the interests, among others, of national security or the prevention of crime and disorder require. Such an exemption would normally exclude the possibility of Bob being able to pursue damages against the public authority. However, perhaps the correct approach is to regard the exemptions as only coming into effect where they are proportionate. If there is a way to obtain the evidence they require without invasion of privacy and other rights of the suspect and without the adverse impact the scenario predicts, then it is arguable that the public authority should take into account the rights of the suspect and so to choose the least disruptive method of obtaining the evidence they need for their purposes. It therefore could be argued that if they chose a method which invades protected rights and is likely to cause adverse impacts, then the public authority have used an exemption disproportionately and so should be obliged to recompense the suspect for the harm perpetrated by their choice of method. It is interesting to conjecture to what extent a court would entertain such a claim. Moreover, the security services may have compromised their investigation by revealing the identity of their suspect to the on-line shop and its employees.

The focus of this paper is on establishing how an appropriate level of privacy can be maintained in the processes of inter-organisational and inter-departmental personal data acquisition in such a way as to avoid both of the issues highlighted above. Thus, existing technologies that may assist public authorities and internal investigators in large organisations in gaining secure and confidential access to a suspect's data are explored. The novelty here is applying scalable 1-n OT protocols into a complete investigative system and validating its use in real-life applications through empirical evaluation.

2. Existing Technology

In the UK there is little government guidance relating to the way data access requests should be sent by public authorities, or to the method of private data transfer back to the authority. RIPA specifies that such operations should be performed according to DPA (this is regulated by the Statutory Instrument "The Regulation of Investigatory Powers (Acquisition and Disclosure of Communications Data: Code of Practice) Order 2007" S.I 2197/2007 which came into force on 1st October 2007), which requires that data is secure, but it does not provide a specific method of achieving this. However, this

paper argues that it is necessary to protect the privacy of the suspect and of the investigation, to a level that cannot be provided by the technologies used in the public domain today. The authors of this paper suggest that the only viable solution lies within the area of *multiparty computation*.

Largely based on the same functions as common encryption schemes, multiparty computation has a strong theoretical underpinning (Goldwasser 1997) and the structural building blocks for the creation of a solution to these privacy concerns (Frikken & Atallah 2003). Multiparty computation allows *n* different parties to engage in a protocol to enable them to compare their secret inputs, or to compute a function, without revealing these inputs. A classic case is Yao's millionaires' problem, where two millionaires seek to compare their fortunes without revealing the exact figures involved (Yao 1982). Yao provides three different solutions to the problem, giving the basis for the multiparty computation for *n* equal to 2, that is, when only two different parties are involved. He also provides a technique for scaling any 1-2 OT protocol into 1-n OT. Since Yao's hypothesis, many other multiparty computation schemes for value comparison have been defined, including loosely proposed real-life alternatives gathered in Fagin, Naor & Winkler (Fagin, Naor & Winkler 1996). Goldreich, Micali & Wigderson (Goldreich, Micali & Wigderson 1987) later introduced schemes designed for general *n*.

Two major types of protocols that may help to protect privacy, when retrieving information from large data sets, are Private Information Retrieval (PIR) and 1-n OT. In the PIR, a *chooser*, the party that requests information, may query a database in a way that the database provider (*sender*) cannot identify which row is being retrieved. A proposed solution to this problem is to provide the *chooser* with a copy of the database, since, in ordinary PIR, privacy of the records in the database is not considered a factor (Ostrovsky & Skeith III 2007). Consequently, the true aim of research in the field of PIR protocols is to achieve the lowest possible number of computations and communications during the private selection of a record. The use of ordinary PIR protocols would, therefore, solve the problem of compromising an investigation by a public authority. They would, however, negatively affect the privacy of the data subjects unrelated to the investigation, since their records would be provided to the authorities, along with the suspect records.

More advanced than PIR protocols are 1-n OT schemes, which are stronger than PIR on the basis that the *chooser* may learn only one row of the results provided by the *sender*. For this reason, 1-n OT is often called symmetric PIR. 1-n OT schemes, on their own, may be considered as having limited benefit (Schneier 1995). However, this paper argues that, when combined with other protocols, they may provide ideal tools for a privacy-preserving enquiry system.

Many oblivious transfer (OT) protocols were designed to allow the *chooser* access to a randomly selected record from *n* secrets kept by the *sender.* Such protocol is presented in Schneier 1995. This may be useful in mental games (Goldreich, Micali & Wigderson 1987) but it does not satisfy our criteria. Although some 1-n OT protocols allow the *chooser* to select the record of interest by means of an index, the *chooser* often does not know the relevant index. Returning to the scenario provided in Section 1, the security services would know the identity of the suspect, through a credit card number or an IP address, for example, but would not, however, know the index of Bob's records in the database of the on-line shop. Consequently, they would be unable to request Bob's records in a secure manner by using 1-n OT. The solution to this problem, once again, can be found in the field of multiparty computation. More precisely, multiparty computation offers a few alternative schemes for performing an asymmetric equality test (Frikken & Atallah 2003). This is a test that can be used to compare the secret inputs of two parties and provide the results to only one party. Therefore, before parties engage in 1-n OT protocol, a synchronisation phase may occur in which the *chooser* may discover the index of the suspect in the *sender's* database. Using this technique, the security services from the scenario could compare Bob's IP address to those of recent visitors to the on-line shop, in a way that none of the parties would know the IP addresses being compared. The security services would, however, know the indexes of the records being compared. Once asymmetric equality protocol has reported a match on the IP addresses being compared, the security services would know the index needed for 1-n OT enquiry. Frikken & Atallah describe a technique used for asymmetric equality testing that employ commutative encryption schemes. These schemes are characterised by the fact that a plain-text, encrypted with two different sets of keys into a cipher-text, may be decrypted in any order of applying the keys. Therefore, in a commutative encryption, where E_A and E_B are two commutative functions, the following is true:

$$E_A(E_B(\text{plaintext})) = E_B(E_A(\text{plaintext}))$$

Equation 1: Commutative functions

This paper proposes a system that uses the characteristic shown in Equation 1 during the ID synchronisation phase, as described in greater detail in Section 4.2.

3. Related Research

In recent years, privacy-preserving data collection and mining have received a considerable amount of attention. Researchers have developed techniques that permit data mining within a cross-section of records rather than from the whole set (Aggarwal & Yu 2008; Agrawal & Srikant 2000; Kantarcioglu & Vaidya 2002). Some of these approaches are based on perturbations, as the one described in Agrawal & Srikant 2000, and can result in an information loss. This information loss is perceived in Aggarwal & Yu 2008 as a natural trade-off between accuracy and privacy, since the larger the number of perturbations, the greater the level of privacy. Another approach includes *k*-anonymity models first proposed in Samarati 2001, where attempts to link any gathered records to its owner creates at least *k* different entries. Also, Aggarwal & Yu 2008 provide techniques based on transforming original data sets into anonymised data sets. Overall, these well-developed solutions are suitable for data mining in relation only to statistical information rather than to individual records, where the identity of the data subject is essential – applying them in a Court of Law would raise questions of authenticity, accuracy and validity, which the above well-developed privacy-preserving solutions cannot provide.

Current research into the retrieval of individual records mainly focuses on PIR or 1-n OT protocols (Frikken & Atallah 2003; Kantarcioglu & Clifton 2003; Li et al. 2008; Naor & Pinkas 2001; Wen-Guey 2002). Possible uses of these protocols include: electronic watch-lists of suspects (Frikken & Atallah 2003); cooperative scientific computation (Du & Atallah 2001; Goldwasser & Lindell 2002); on-line auctions (Cachin 1999); and secure comparison of information (Fagin, Naor & Winkler 1996). However, most of these protocols and solutions suffer from a possible large computational overhead resulting from use of public-key algorithms (Li et al. 2008). Thus, in early solutions, the computational complexity was exponential to the number of bits used to store the private records and to the number of records (Cachin 1999). However, Naor & Pinkas (Naor & Pinkas 2001) managed to lower the number of exponential operations required below one per record, by increasing the communicational complexity. Li and others (Li et al. 2008) discussed a system that does not make use of any public-key technology. In their system, Ex-OR operations are used to provide a different approach to Yao's problem, although Impagliazzo & Rudich (Impagliazzo & Rudich 1989) argue that a system without a trapdoor function, such as public-key encryption, is unlikely to perform the required operations.

Most research into PIR and OT has focussed on perfecting previously developed schemes, with little attention paid to their □ ractical use (Li et al. 2008; Naor & Pinkas 2001; Ostrovsky & III 2007; Wen-Guey 2002). In addition, none of the PIR/OT related research provides empirical evaluation. Comparison of the different schemes is usually done on the basis of computational and communicational complexity, which, some researchers assert (e.g. Li et al. 2008) should not be directly compared. Past research has shown that the efficiency of asymmetric key operations is 0.1% of the efficiency of symmetric key operations (Li et al. 2008). If this was true, a protocol that takes O(*n*) symmetric operations, would take a similar amount of time to a protocol with O(*n* / 1000) asymmetric operations. At the same time, computational complexity, expressed in terms of the number of operations undertaken, would suggest otherwise. Consequently, it could be argued that OT and PIR schemes are most practical when they are defined and evaluated for a specific problem, rather than for a general solution (Goldwasser 1997). This was also suggested by Naor & Pinkas (Naor & Pinkas 2001), who emphasise that selection of the trade-off between computational and communicational complexity depends on the specific problem at hand. Thus, in the following sections, a basic system for solving the privacy problem during inter-organisational and inter-departmental investigation is described, as well as an empirical proof of its efficiency for the task at hand.

4. Proposed System

The system for improving data subject privacy in investigations conducted by the public authorities is built on a base of theoretically tried-and-tested ideas and common encryption protocols currently in the public domain. The following section defines these building blocks, and the way they fit together.

4.1 Querying and narrowing the scope

The modern ISs may contain large numbers of records and only a narrow subset of such records would usually be related to a person being investigated, and therefore, be of interest to investigators. To safeguard privacy, the identity of the potential suspect cannot be provided to a data holder, a *sender*. Thus, the investigator, a *chooser*, needs to narrow down the scope of an inquiry, before privacy-preserving protocols may be used efficiently.

Depending on the level of privacy required, the scope of an enquiry should be narrowed to a number of records/identities between 100 and 1000, which this paper identifies as *selected records*. This may be achieved using techniques known from data mining. Thus, assuming that the *chooser* is a public authority that is investigating the data subject, it should be able to provide some basic information to narrow down the records to be processed. For example, the time of the last transaction performed by the suspect may be known by the *chooser*. This, in turn, can be used to limit the scope of the privacy-preserving enquiry to a set of data subjects who performed a transaction at around the same time as the *suspect.* The *sender* may be queried about its services usage statistics before the enquiry, in order to make sure the scope is not narrowed too much. For example, in a small on-line shop, the enquiry may be made about records of all customers that performed a transaction on a given date, while a larger e-commerce site may be queried about records of customers who performed a transaction in a much smaller time window, such as for less than one hour.

4.2 ID synchronisation

After the enquiry has been narrowed down, the *chooser* must check that the *suspect's* data is among the *selected records*. The core of this system is a 1-n OT protocol that requires the *chooser* to know the index of the *suspect's* data within the *selected records*, otherwise the protocol will not work. Thus, a commutative encryption is used for this purpose in the following way:

(1) The *chooser* encrypts the *suspect's* ID with key E_{ch} and sends it to the *sender*.
(2) The *sender* encrypts the received input, $E_{ch}\ (ID_{suspect})$, with its own key E_s, and sends $E_s(E_{ch}(ID_{suspect}))$ back, together with all IDs in the *selected records* encrypted one-by-one, $E_s(ID_1), \ldots, E_s(ID_n)$, where *n* is the number of the *selected records*.
(3) The *chooser* encrypts $E_s(ID_1), \ldots, E_s(ID_n)$, the encrypted IDs received, with its own key, to produce $E_{ch}\ (E_s(ID_1)), \ldots, E_{ch}\ (E_s(ID_n))$ and compares resulting cipher-texts to the ID of the *suspect* encrypted by the both parties, $E_s(E_{ch}(ID_{suspect}))$.

If any of the *selected records* refers to the *suspect*, its cipher-text will match $E_s(E_{ch}(ID_{suspect}))$, and the *chooser* will recognise the index, *i*, of the *suspect's* data within *selected records*. At the same time, the *sender* will not be able to recognise the ID of the *suspect,* as long as the encryption used holds.

4.3 OT Protocol

ID synchronisation provides an index that is needed for the correct operation of 1-n protocol. This system uses a straightforward 1-n OT protocol, which is based on the 1-out-of-2 Oblivious Transfer (1-2 OT) scheme described in Schneier (1995). The resulting protocol operates as follows:

(1) The *sender* generates *n* sets of *public/private keys* pairs, and sends all *public keys* to the *chooser*, preserving the order in which they have been sent.
(2) The *chooser* generates a key with a private encryption algorithm, such as AES, later called *AES key*. It then uses the i^{th} *public key* received from the *sender* in Step 1 to encrypt the *AES key* and send it to the *sender*.
(3) The *sender* does not know which *public key* has been used to encode the *AES key*, or which record has been selected, thus protecting the privacy of the *suspect*. The *sender* can then decode cipher-text received in Step 2 using all *private keys* generated in Step 1, whilst preserving the order in which they have been decrypted. In this way *n* potential AES keys are created. Only the i^{th} one is the proper *AES key*; the other outputs are random sets of bits, which cannot be distinguished from ordinary *AES keys.*
(4) The *sender* encrypts all records using appropriate keys decrypted in Step 3. Thus, the first record in *selected records* is encrypted with an *AES key* decrypted using the first private key generated in Step 1. Consequently the i^{th} record, which includes data about the *suspect,* is encrypted using the *AES key* generated by the *chooser* in Step 2, sent to the *sender* encrypted by the i^{th} *public*

key, and then decrypted using i^{th} private key. This way the i^{th} record will be encrypted using the proper *AES key*.

(5) The *chooser* gets *n* encrypted records, but using the *AES key* it is able to decrypt only the i^{th} record. Other records are unreadable to the *chooser* provided that the false keys generated in Step 3, and used to encrypt these records in Step 4, are not broken.

4.4 Prototype

The feasibility of the above system was assessed with the use of a prototype, built in C# .NET and the Legion of the Bouncy Castle cryptography API (The Legion of the Bouncy Castle, 2007). To simplify operation and testing, the current prototype does not connect to a database, but instead uses files as records. Taking into consideration that IPv4 addresses are 32-bits in length, IPv6 addresses are 128-bits long, and a credit card number may be written using 54 bits, 256-bit identifier fields are used in the prototype. Such a solution provided sufficient data for the initial evaluation of the system, which did not, at this stage, include additional levels of complexity.

The first step is to narrow down the scope of the enquiry. For this operation, guidelines may be issued on the cooperation between the public authorities and data holders. However, due to large variations in ISs used by different organisations, this step cannot be automated. Thus, narrowing down the scope is not a part of the software prototype.

The ID synchronisation of the prototype uses a modification of the RSA scheme, sometimes referred to as SRA. This modification takes account of the fact that RSA is commutative for keys generated with common modulus n, and secure as long as the exponent is kept private (Chevalier, et al. 2005). During ID synchronisation, the *chooser's* key protects the identity of the *suspect*. Thus, different commutative encryption schemes, such as Pohlig-Helman protocol, may be used depending on the security requirements.

Finally, the 1-n OT protocol employed in the prototype uses RSA with 1024-bit keys and AES with 256-bit keys, which are currently standard in the encryption field. Breaking any or all of the *n* RSA keys at this stage will not provide any measurable benefits to a potential perpetrator. However, if the private encryption scheme keys can be broken, the public authority, or an adversary (if the communication channel is not secured), could decrypt all *selected records*. Thus, it is important to choose a well-tested encryption scheme with an appropriate length of key.

5. Initial Evaluation

Usually protocols are evaluated on the basis of the number of operations and level of communication required. The ID synchronisation phase requires a maximum of $O(2(n+1))$ computations, but only two rounds of communication are needed. Likewise, the computational complexity of the 1-n OT used is large, and is affected by two major operations: symmetric key exchange; and processing of the *selected results*. Both of these operations require $O(n+1)$ each; however, only three rounds of communications are required, in total. Thus, the complete system requires only five rounds of communication, which simplifies the exchange of information between two parties. Consequently, real-time data exchange between the *chooser* and the *sender* is not necessary, so asynchronous transmissions may be used. This also permits the manual exchange of data, such as via post or secure email.

Figure 1 illustrates that the processing time is almost linear to the number of records in the set. Thus, processing time per record of a set-size can be established. For example, when running both the *chooser* and the *sender* processes on an ordinary workstation (RAM: 1GB, CPU: 1.6GHz), the total processing time is around 0.6 seconds per 1MB record. Increasing the physical size of records also affects the processing time in a linear manner (please see Figure 2 for details). More importantly, during the tests, CPU usage on the test workstations did not exceed 55%, meaning the impact of the investigative system on the underlying platform is limited.

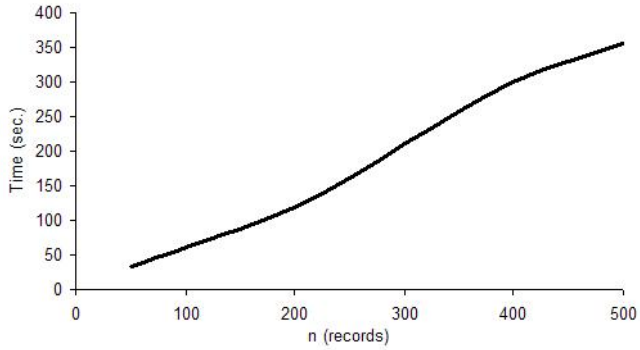

Figure 1: Linear increase in processing time as number of records in the set, *n*, increases

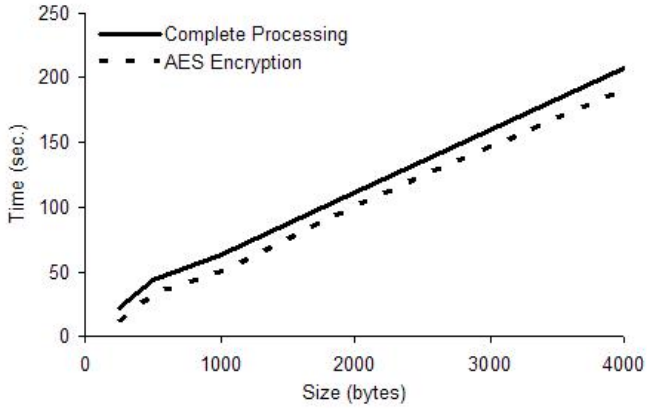

Figure 2: Linear increase in processing time as record size increases.

The most costly operation in the whole system is the preparation phase. On the testing platform, generation of 100 RSA key pairs took around 160 seconds for 1024-bit keys, and around 85 seconds for 786-bit keys. This appears to be the only downfall of the system, however, the RSA keys may be prepared beforehand by the *sender* and stored in a repository until they are required. Thus, this aspect is not considered to be a major disruptive issue.

6. Conclusion

Modern research in multiparty computation is rich in theoretical solutions related to privacy concerns. Some of the solutions are mature enough to assist public authorities and internal investigators in safeguarding the privacy of the data subjects during inter-organisational and inter-departmental investigations. Arguably, before such solutions can be legally used in the inter-operational domain, some amendments to the current data protection laws will be required. The reason for this is the fact that 1-n OT protocols, apart from the requested (i^{th}) record, pass all other records to the requesting party. From a technical point-of-view, these records are encrypted in a way that renders them unusable. However, from a legal perspective, these records would be considered as processed data and as being sent to the requesting party unlawfully – that is, outwith the permitted exemptions allowed by the DPA.

The results of the evaluation clearly show that use of privacy-preserving systems during investigations is possible without causing delay to the investigations and with a negligible impact on the level of processing required by the third party providing the records. Operations requiring the input of such a third party, can be performed quickly on a standard developer's or administrator's workstation. This paper has proposed a system that would successfully protect Bob's privacy in the scenario described in Section 1. Thus, the security services would obtain the data that would clear Bob from suspicion, without identifying him as a suspect to the on-line shop. The whole operation would take around 6 minutes if the records processed by the system could be narrowed to down to *n* equal 500.

This paper provides a description of a simple, but complete system that may be used for both inter-organisational and inter-departmental investigations. The simplicity of the design allows full comprehension of the operations performed, and understanding of the system's potential. The transparency of the design is increased by the fact that well tested, widely available private and public encryption algorithms are used. Although C# was used for prototyping, a variety of different programming languages may be used to implement the real-life system. Moreover, two parties should be able to engage in protocol using different implementations of the system, since its specification will be in the public-domain. This system, apart from the small number of communication rounds required, has the advantage of good scalability, since the processing time is linear to the number of records processed and to the size of records. Thus, this paper could be regarded as a step forward in safeguarding the privacy of individuals during investigations. With such system in place, there is no need to compromise the privacy of suspects due to time restrictions. The time required is small compared to manual data handling methods, used currently to obtain a suspect's data from third parties. Further planned research will involve making the above system more efficient without affecting its complexity, followed by more detailed evaluation using synthetic and real data.

7. Acknowledgements

The authors would like to thank Professor Rao Bhamidimarri, Dean of Faculty of Engineering and Computing, for making funding available for research, as well as Bob Rankin, Head of School, for assigning grants to this research program.

8. References

Aggarwal, C. C. and Yu, P. S. (2008) "On static and dynamic methods for condensation-based privacy-preserving data mining", *ACM Trans. Database Syst.*, Vol. 33, No. 1, pp 1-39

Agrawal, R. and Srikant, R. (2000) "Privacy-preserving data mining", *SIGMOD Rec.*, Vol. 29, No. 2, pp 439-450

Ball, K., Lyon, D., Wood, D.M., Norris, C. & Raab, C. (2006) *A Report on the Surveillance Society*, Surveillance Studies Network.

Cachin, C. (1999) *Efficient private bidding and auctions with an oblivious third party*, 6th ACM conference on Computer and communications security - CCS '99, Singapore, pp 120 - 127

Chevalier, Y., Kusters, R., Rusinowitch, M. and Turuani, M. (2005) "Deciding the Security of Protocols with Commuting Public Key Encryption", *Electronic Notes in Theoretical Computer Science*, Vol. 125, No. 1, pp 55-66

Du, W. & Atallah, M.J. (2001) 'Privacy-Preserving Cooperative Scientific Computations', paper presented to the *Proceedings of the 14th IEEE workshop on Computer Security Foundations*.

Fagin, R., Naor, M. and Winkler, P. (1996) "Comparing information without leaking it", *Commun. ACM*, Vol. 39, No. 5, pp 77-85

Frikken, K.B. & Atallah, M.J. (2003) 'Privacy preserving electronic surveillance', paper presented to the *Proceedings of the 2003 ACM workshop on Privacy in the electronic society*, Washington, DC.

Gill, M. and Spriggs, A. (2005). *Assessing the impact of CCTV*: Home Office Research, Development and Statistics Directorate

Goldreich, O., Micali, S. & Wigderson, A. (1987) 'How to play ANY mental game', paper presented to the *Proceedings of the nineteenth annual ACM conference on Theory of computing*, New York, New York, United States.

Goldwasser, S. (1997) 'Multi party computations: past and present', paper presented to the *Proceedings of the sixteenth annual ACM symposium on Principles of distributed computing*, Santa Barbara, California, United States.

Goldwasser, S. & Lindell, Y. (2002) 'Secure Computation without Agreement', paper presented to the *Proceedings of the 16th International Conference on Distributed Computing*.

Impagliazzo, R. & Rudich, S. (1989) 'Limits on the provable consequences of one-way permutations', paper presented to the *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, Seattle, Washington, United States.

Kantarcioglu, M. & Clifton, C. (2003) 'Assuring privacy when big brother is watching', paper presented to the *Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery*, San Diego, California.

Kantarcioglu, M. & Vaidya, J. (2002) 'An architecture for privacy-preserving mining of client information', paper presented to the *Proceedings of the IEEE international conference on Privacy, security and data mining - Volume 14*, Maebashi City, Japan.

Li, S., Wang, D., Dai, Y. and Luo, P. (2008) "Symmetric cryptographic solution to Yao's millionaires' problem and an evaluation of secure multiparty computations", *Inf. Sci.*, Vol. 178, No. 1, pp 244-255

Naor, M. & Pinkas, B. (2001) 'Efficient oblivious transfer protocols', paper presented to the *Proceedings of the twelfth annual ACM-SIAM symposium on Discrete algorithms*, Washington, D.C., United States.
Ostrovsky, R. and Skeith III, W. E. (2007) *A Survey of Single-Database PIR: Techniques and Applications*, PKC
Samarati, P. (2001) "Protecting Respondents' Identities in Microdata Release", *IEEE Trans. on Knowl. and Data Eng.*, Vol. 13, No. 6, pp 1010-1027
Schneier, B. (1995) *Applied Cryptography: Protocols, Algorithms, and Source Code in C*, John Wiley & Sons, Inc.
The Legion of the Bouncy Castle (2007) "Bouncy Castle C# AP", [online], The Legion of the Bouncy Castle, http://www.bouncycastle.org/csharp/
Vlahos, J. (2008) "Surveillance Society: New High-Tech Cameras Are Watching You", *Popular Mechanics*, Vol. No. 01/2008
Wen-Guey, T. (2002) 'Efficient 1-Out-n Oblivious Transfer Schemes', paper presented to the *Proceedings of the 5th International Workshop on Practice and Theory in Public Key Cryptosystems: Public Key Cryptography*.
Yao, A. (1982) "Protocols for Secure Computation", *23th FOC$*, Vol. No. pp 160-164
Young, B. C., Kathleen, E. C., Joshua, S. K. and Meredith, M. S. (2006) "Challenges Associated with Privacy in Health Care Industry: Implementation of HIPAA and the Security Rules", *J. Med. Syst.*, Vol. 30, No. 1, pp 57-64