The 8th Privacy Enhancing Technologies Symposium

Leuven, Belgium
July 23–25, 2008

Summarized by Rae Harbird (rbird@gmail.com)

This conference was the eighth in a series designed to promote research in privacy enhancing technologies. It marks the transition from workshop to symposium with published proceedings. The conference organizers were keen to maintain the spark and spontaneity of previous years and so kept the "talk for 3 to 5 minutes on anything you like" rump session. The program also included a new feature, HotPETs, in which invited researchers talked about their latest ideas. The social events also provided plenty of opportunity for lively discussion; the banquet featured the PET awards conferred for work that, in the view of the judges, strengthened privacy protection through technology.

PETS/WOTE JOINT SESSION

- ***Using a Touchscreen Interface in Prêt à Voter***
  David Lundin and Peter Y.A. Ryan, University of Surrey, UK

The first morning of the PETs conference was held in conjunction with the Workshop on Trustworthy Elections. In this presentation David Lundin described an extension to Prêt à Voter (PAV), an open-source, electronic voting system, designed to enable the use of a touchscreen interface. Up until now, PAV has used only printed ballot papers, but a touchscreen interface alongside traditional printed voting slips offers improved usability and increased accessibility. Recent PAV developments include the addition of a paper audit trail to support human-readable electronic verification; this has had the side effect of facilitating the use of interactive voting machines.

PAV encodes votes using a randomized candidate list, thus ensuring the secrecy of each vote and removing any bias that might occur with fixed ordering. Conceptually, the PAV ballot has two parts: One is placed in the ballot box and can be used as a Human Readable Paper Audit Trail (HRPAT); the other is used as a protected receipt, which can be printed and kept by the voter, electronically published, and used for vote counting. Both parts of the ballot paper contain an encrypted, randomized list of the election candidates known as an *onion*. The onions are related cryptographically such that it can be clearly shown that one is derived from the other, even when modified with information about the vote cast.

Briefly, the voting procedure is as follows: The voter casts a vote at a terminal, which displays a ballot form generated from a peeled (decoded) onion. The terminal prints a two-part ballot paper, each containing marks indicating the cast vote and an onion. The onion on the receipt part is modified to include information about the vote. The voter retains the receipt and, after sealing the HRPAT, takes it to a teller, at which point the teller posts the ballot (still in the envelope) into a sealed, transparent ballot box. Using the cryptographic relationship between the onions and the candidate list it can be determined that the terminal has not cheated in either printing the candidate list or recording the vote. A threshold set of decryption tellers performs the final stage, decrypting the onion representing each vote.

KEYNOTE SPEECH

- ***Analyzing PETs for Enterprise Operations***
  Stuart Shapiro and Aaron Powell, The MITRE Corporation, USA

Enterprises, as guardians of personally identifiable information (PII) in both the private and public sectors, are beginning to acknowledge the need to address the threats to privacy. There are very few PETs that can support the PII life-cycle management from collection through destruction of that information. PETs do not necessarily help unless they reflect and support the business process of the organization, and this may not involve privacy-specific technologies.

Understanding which technologies can be used to support particular business processes is not always straightforward. The authors have found it useful to categorize specific business processes and PETs in the context of an organization's requirements as a means of assisting the selection and deployment of particular technologies. In general, many PII-related business processes are common across organizations, a lesser number are specific to the type of organization, and a few pertain specifically to the organization. PETs can be categorized according to their primary function: data desensitization or anonymization, identification of PII, those that assist with policy enforcement such as network monitoring and end-point event detection, etc. Once categorization has been achieved, it is much easier to map PETs to business processes. This helps to identify higher-risk busi-

ness processes, and these can be examined in further detail with use cases delineating precisely how the technology will be employed. It is also much easier to identify potential gaps in the enterprise's processes, ensuring that the enterprise is receiving the maximum business benefit from the technologies used.

In the long term, the authors advocate a more holistic approach, introducing the concept of a Privacy-Enabled Architecture (PEA). We are familiar with using architectures as templates, encapsulating desirable properties, and it should be possible to embed privacy-enabling technologies within a system or enterprise design to achieve comprehensive privacy risk management. There is a clear analogy with service-oriented architectures, which focus on business processes and in which there is a loose coupling of services and specific technologies.

SESSION 1

- ***Perfect Matching Disclosure Attacks***
  Carmela Troncoso, Benedikt Gierlichs, Bart Preneel, and Ingrid Verbauwhede, K. U. Leuven, ESAT/SCD-COSIC, IBBT, Belgium

Anonymous network communication software is used to hide the identities of communicating parties but allow information to be gleaned from analyzing traffic data between them. Disclosure attacks can be used to uncover the relationships between those communicating and the pattern of their communications. In this presentation Carmela described the two contributions of the paper. First, the authors have developed an improved and more realistic user scenario, in which a set of users communicate with each other over an anonymous channel modeled as a threshold mix. Previous workers investigating these attacks only considered a very simple model, in which users choose their communication partners with uniform probability and the effectiveness of attacks investigated strongly relies on this model. The model in this paper relaxes many of the constraints previously imposed, introducing behaviors that are more random in nature.

Second, Carmela described a new attack, known as the Perfect Matching Disclosure Attack (PMDA), which can be used to discover the relationship between messages sent between the set of senders and receivers, enabling attackers to build a profile of users' behavior. PMDA is based on graph theory: communications between users are represented as a maximum weighted bipartite graph. An attacker, observing communications over several rounds, is able to represent the edges in the communications graph as a matrix of log probability values. Subsequently, linear assignment methods can be utilized to find the maximum weight bipartite graph.

The authors assessed the performance of PMDA by comparing it in simulations with the previously published Statistical Disclosure Attack. Results show that PMDA is more accurate when linking senders and receivers. In the future the authors intend to generalize the user communication model yet further by introducing behavior variance over time. They also intend to improve the efficiency of PMDA by parallelizing attacks and by parallelizing the linear assignment problem solver.

SESSION 2: ANALYSIS

- ***Metrics for Security and Performance in Low-Latency Anonymity Systems***
  Steven J. Murdoch and Robert N.M. Watson, Computer Laboratory, University of Cambridge, UK

Tor is the latest generation of Onion Routing software used for anonymous communications over the Internet. This research aims at answering the following question: How do Tor routing decisions affect the risks of communication compromise? Tor directory authorities maintain a list of nodes, with associated attributes, participating in the Tor network. Clients initiate a connection by retrieving information about a set of candidate nodes, known as a consensus directory, from a Tor directory authority. Path selection is carried out subsequently by Tor clients based upon mapping their own requirements to the available node selection algorithms and applying the algorithm to the list of candidate nodes. A user may prefer, for example, a route offering more security from compromise or may choose a route that has more bandwidth capacity.

The authors have developed a Tor path simulator that uses a consensus directory as input with a given number of the nodes within it marked as malicious. Acting as a set of Tor clients, the simulator generates paths with given characteristics (fast or stable). The security of the routes generated is measured as the probability, with respect to the cost and selection algorithm, of connection compromise, characterized by an attacker being able to control the first and last nodes in a Tor connection. The authors evaluated four path-selection algorithms, drawing from current Tor behavior and research in this area. In further experiments the network performance (connection latency) was evaluated using the set of route-selection algorithms. The authors conclude that the results are surprising: within the constraints imposed by experimentation, Tor's default bandwidth-weighted path-selection algorithm offers improved performance in terms of compromise and network latency over the supposedly more secure Tor uniform path-selection algorithms.

PANEL

- ***Data Minimisation, Proportionality, and Necessity in Privacy Systems***
  Moderator: Caspar Bowden, Microsoft UK
  Reported Panelists: Paul de Hert, Vrije Universiteit Brussel, Belgium; Eleni Kosta, K.U. Leuven, Belgium; Gordon Nardell, 39 Essex Street, UK

Caspar launched the panel discussion by introducing the panelists and briefly explaining the theme. Paul de Hert

followed, talking about problems with privacy-related legislation in Europe: There is a solid human rights framework, but it allows for many exceptions. Article 8 of the 1950 European Convention on Human Rights (ECHR) is used to protect both electronic communications and equipment; furthermore, most countries have extended these rights in their own legislation. The ECHR also contains a set of very broad exceptions; hence privacy-infringing laws are seldom challenged in court. Moreover, these exceptions are not always applied in a consistent manner. In practice, the proportionality of many new privacy-infringing technologies remains largely unchecked and they are accepted by a simple balancing of interests that always turns out in favor of the government that proposes them.

Paul believes that there are several possible remedies to redress this lack of balance. First, put technology back in context by understanding the values that are at stake. New technologies challenge current value settings; a true balancing exercise should therefore integrate our current understanding of these values and our ambitions to uphold or alter certain settings. Second, we can create new human rights. In the 2001 EU Charter on fundamental rights the right to privacy was complemented with a right to data protection. The presence of this addition forced judges to reconsider the traditional privacy balancing act. Principles such as "purpose limitation" and "consent" do now have a human rights status and their disrespect will necessarily influence a balancing act. Third, more attention needs to be paid to the concept of proportionality. German constitutional law offers a far superior concept. Next to the question "Is this law proportional?" (proportionality in a strict sense), it addresses questions such as "Is there not an alternative that better respects privacy?" (subsidiarity) and "What is left of a specific human right when the governmental initiative to deploy a certain technology gets a green light?"

Gordon Nardell followed. Like Paul de Hert, Gordon noted that Article 8 of the ECHR is a contradiction in some sense, "giving with one hand and taking away with both." Earlier this year the ECHR recognized this dilemma when it gave judgment on a case brought by Liberty and two other NGOs against the UK government. It discovered that the UK Ministry of Defence had intercepted communications that flowed through two of British Telecom's radio stations while in transit between the UK and Ireland. The data gathered was filtered by using search engines and keyword lists and used by intelligence analysts. The statutory procedure involved an application to a minister for a warrant. The problem lay with the analysis of the data after collection rather than with the interception, because a warrant authorized indiscriminate interception of huge numbers of communications, while the process of isolating individual communications, where the real interference with privacy took place, was governed by secret "arrangements," which were not detailed in the legislation. The European court upheld that legal discretion granted to the executive was unfettered and noted that most other European countries publish more information on their respective surveillance laws than the UK. The judgment raises questions over the UK's controversial Regulation of Investigatory Powers Act (2000) and there may be a revision of that law. The UK government should ensure that the use of intercepted data is just as transparent as the acquisition of that data. This ruling may also have an impact on the state's obligation to intervene in cases of private data mining.

Eleni Kosta discussed the European Union's data retention directive of 2006. The purpose of the directive is to harmonize the obligations of Internet Service Providers (ISPs) and telecom operators with respect to the retention of certain data. The directive states that information such as the source, destination, and type and date of communications should be harvested but not the content. Privacy analysts argue that this is not as clear-cut as it might seem. For example, an email address or source information can reveal more information about the participants than is strictly necessary. Data collected must only be provided to competent national authorities, but there is some question surrounding which entities this covers. In the case where a provider shares information inappropriately (i.e., to an authority that is not entitled to receive the data), it will be liable for any resultant damage. The directive has been challenged in front of the European Court of Justice (ECJ): Ireland has asked for an annulment of the legislation on the grounds that it has not been adopted on a proper legal basis. In April of this year, over forty NGOs signed an amicus curiae brief asking the ECJ to annul the EU directive on data retention. They pointed out that, apart from the formal grounds put forward by Ireland, the directive is illegal on material grounds, mainly for infringement of the right to privacy (Article 8, ECHR). In Germany the piece of legislation that transposes the data retention directive into national law has already been challenged in front of the Constitutional Court, which has still not published its decision. However, the Court adopted an interim ruling, stating that data retention as such is not unconstitutional. However, the law implementing the data retention directive does not provide sufficient guarantees concerning the access to the data and the crimes for which data retention may be used. Data retention may only be used for serious crimes and when a judicial warrant is present and therefore the relevant article must be revised.

SESSION 3: ATTACKS

- ***Chattering Laptops***
  Tuomas Aura and Michael Roe, Microsoft Research, Cambridge, UK; Janne Lindqvist, Helsinki University of Technology, Finland; Anish Mohamed, Royal Holloway, University of London, UK

Janne Lindqvist reported the results of an investigation into the information that could be gleaned from wireless-enabled laptop computers by an attacker snooping the packets

broadcast by the operating system in order to bootstrap networked system services such as network connection (DHCP), network address resolution (DNS), and network file systems (e.g., NFS). Most operating systems will attempt to initiate these kinds of services as soon as they are switched on and will periodically retry even if failure occurs. Inspection of these preliminary protocol interactions shows that a lot of information is revealed which may enable identification of the service providers (domain name, service details, and server name or IP address) and users (user name, email address, and real name). This information may invite unwanted attention for the user or expose the user's computer to further hacking attacks.

The authors propose a partial solution to this dilemma, namely, policy-controlled use of Network Location Awareness (NLA). Some operating systems allow the user to configure and select one of a set of network profiles. For example, Windows Vista implements the NLA service, which will identify the network without user intervention. NLA creates a fingerprint of the access network based on a set of parameters associated with that network. In the case of authenticated networks the fingerprint might be the security parameters; in all other instances it could be the gateway MAC address. A cryptographic hash is generated from the fingerprint and this is used to identify the network when the user is interacting with it. The next step is to disable and enable service discovery protocols depending on the observed network fingerprint. Some useful default policies were described: NetBIOS should be disabled and enabled separately for individual networks, the default DNS suffix should be disabled outside the domain network, and network file shares and printers should be probed only in the network for which they were originally configured.

HOTPETS

- ***PRAIS—PRivacy impact Assessment for Information Sharing***
  Rae Harbird, Mohamed Ahmed, and Anthony Finkelstein, University College London, UK; Andrew Burroughs, Coram, UK; Elaine McKinney, Logica, UK

The UK government is promoting multi-agency information sharing as a key component of new work practices for those providing services to children and families. Despite the plethora of guidance available, staff do not always feel confident to share what they know. This research project has involved a short-term collaboration between computer scientists and child-protection experts. Together they have developed a prototype decision support tool, known as PRAIS (PRivacy impact Assessment for Information Sharing) in the domain of children's social care. The PRAIS system is designed to assist in the decision-making process, not to replace it. Users are not bound to follow the information-sharing actions advocated by PRAIS and staff will be aware that all information-sharing decisions are taken at their own discretion, based on, among other things, their assessment of risk to the individuals involved. PRAIS can also be used as a training tool to help professionals learn experientially about the issues in managing personal information.

PRAIS has been engineered as an expert system, comprising a user interface, a knowledge base containing privacy-related facts and rules, and an inference engine, which can interpret the knowledge base and draw conclusions. The user interface is a Web application representing some of the common work procedures in a social worker's day-to-day tasks that may involve information sharing. The rules are compliant with the UK Data Protection Act and have been reviewed by the Information Commissioner's Office. Project participants are working toward securing a new partner with whom they can develop an operational version of PRAIS and evaluate its efficacy in a realistic environment.