A Short Discussion on Privacy: Case Studies

Vahab Pournaghshband

Computer Science Department

University of California, Berkeley

vahab@cs.berkeley.edu

1 Incentive to profile

1.1 Price discrimination

Profiling customers is an incentive for the sellers to implement price discrimination. If I am a seller, I want to know as much information as possible about what you are interested in buying and how much you are willing to pay for it.

1.2 Advertising in market

Some companies like Google, Yahoo and Facebook have revenues from advertising and, as a result, their business models are based upon profiling. To make their advertising more effective, these companies have strong incentives to profile their users on what they are interested in and are likely to buy. They might not go after names but certainly they are interested in your income, demographic information, age etc.

1.2.1 Consumers won't pay for privacy

Surveys from consumers about privacy have shown that they are very concerned about their privacy. However, when they were asked to tell how much they are willing to pay for their privacy, very few responded positively.

---

[0]This note is based on David Wagner's Lectures on Computer Security at UC Berkeley.

Often companies view privacy as more of a cost issue and not a profit issue. Companies can lose a lot if they fail to preserve privacy of their customers, but will not be rewarded otherwise. For instance, an oil company will lose remarkably if it is responsible of an oil spil. On the other hand, it is not effective for an oil company to advertise that it has not been responsible for any oil spil in ten years. A few years ago, a start-up company called Zero-knowledge, Inc. released a privacy protection product called *freedom*. It claimed to protect privacy of Internet users by various ways such as anonymizing its customers' internet addresses. They sold approximately 10,000 units and as a result of such little success, they were forced to go out of business. On the other hand, some other companies like DigiCash and Paypal were successful in providing privacy protection solutions. Zero-knowledge is a good example of the fact that great technology doesn't necessarily lead to commercial success.

2 Driver's license

The magstripe on the driver's license card contains information shown on the card and possibly more. The real incentive for the magstripe on the driver license has been to help law enforcement personnel with their duties.

Problem: Other companies (e.g. bars) also can read the information stored on the magstripe for various purposes

Solution:

- Online reader: In this case, the only information stored on the magstripe could be an index that is sent to the master database to retrieve the requested information.

- Offline reader: In this case, the information could be encrypted by a private key only known to police officers (but this can lead to key management issues).

In US there is little law on privacy but there is always an exception: Video rental merchants are not permitted to reveal any information on what their customers have rented (Video Privacy Protection Act). The historical fact about this is the Judge Robert Bork's video rental history that was leaked to the public during his nomination which had serious consequences.

3   e-commerce

It is absolutely a killer for e-commerce businesses to lose sensitive information of their customers. An example of such incident is T.J.Maxx. The company lost the database of its customers, including their name, credit card information, address, and phone numbers. The server was hacked because of an insecure wireless network used, and as a result there was a lawsuit later filed against the company. TJ Maxx lost records of about 10 million customers and suffered a significant financial loss of more than $100 millions for this incident.

Problem: How to protect customers' sensitive information?

Solution? Simply not maintaining sensitive information such as credit card numbers by deleting them when the transaction is complete.

Challenge: The company may want to keep these information for various reasons. For instance, if the company wants to offer discount based on a particular customer's interest and spending. Clearly having such information is valuable. In fact, e-commerce companies normally use the credit numbers as an index to their database.

Solution? One-way hashing of credit card numbers

Solution? Multiple database
Not very plausible. All databases are compromised by attacking only the web server (Figure 1).
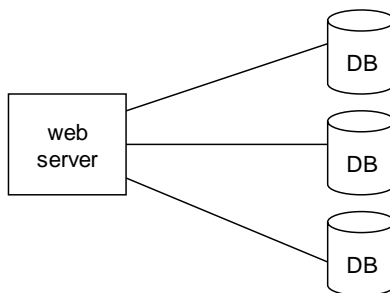

Figure 1: An e-commerce web server is connected to multiple databases.

Solution: Separate out the database containing sensitive data from non-sensitive data, and allow only one way access to this database. The server is allowed to only send queries like charge the user ID by this amount with whatever credit card number you have on the database (Figure 2).
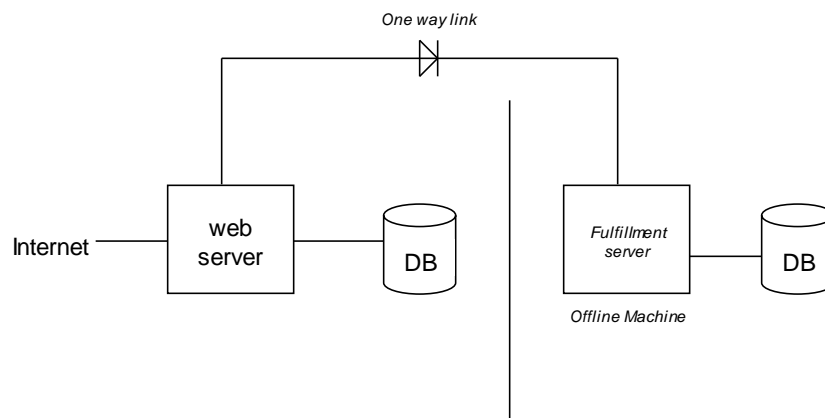

Figure 2: One-way only access to sensitive database.

Problem: How to update information on the fulfillment server?

Solution: Having a trusted employee store the updates on a CD and then implement them to the database or using a trusted software program to do that.

Alternative solution: get use of a throttle between the database and the web server to limit the number of sensitive queries per time unit. For example, allow only 1000 sensitive queries per day, and depending on the database it might take over a month to pull off all the sensitive information, and hopefully it will be noticed by then (Figure 3).
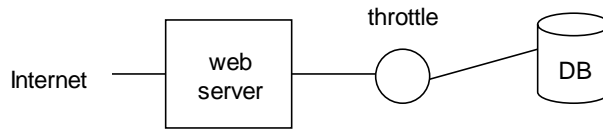
Figure 3: Deploying throttle between the database and the web server.

What do banks require from merchants?

Banks require the merchants to keep the information of transactions for several months for various reasons (e.g. disputed transactions). Thus, they are not allowed to just delete the transaction information once it is complete. To do this, banks require the merchants to protect credit card numbers using the PCI compliance.

One seemingly plausible solution in this case is to keep only a transaction ID or code to eliminate the need for keeping the full transaction information including the credit card number.

4   FastTrack

4.1   How does the system work?

A radio antenna at the toll that pings (normally with a beeping sound as a communication indicator) and a transponder installed on the car that replies to it instantly with its ID.

4.2   Problems with the current system

The protocol is simplistic and no cryptography is involved. When the transmitter pings it, it can set the *beep?* to NO. This way the driver wouldn't notice that it is being pinged and replies accordingly. The risk is for a third party to get access to the database that kept the FastTrack information. Aside from the "Big Brother" issue of government keeping track of where people go, there was actually a divorce lawsuit that the lawyer used this information to prove that the husband had another affair.

There are, in fact, other occasions that FastTrack transponder is forced to reply unnoticeably by setting the beeping sound to NO. One is for the traffic control purposes and realizing how congested the traffic is; but it is

not used to catch violation like speeding even though it is capable of catching it. So, because of this overly simplistic model that the FastTrack system has adopted, the system is vulnerable to attacks by third parties.

Question: So how can we assure that only Caltrans can effectively communicate with the transponders?

Solution? A solution would be for Caltrans to have a public-key p_k and have the transponder to reply the encrypted ID: $(ID)_{p_k}$. But this is also vulnerable to replay attack where the attacker can transmit $(ID)_{p_k}$ and fake to be another person.

Solution: Encrypt the timestamp, t along with the ID: $(ID, t)_{p_k}$.

5 e-passport

US have recently adopted e-passports that contain an RFID tag that carry digital information such as a digital photo and possibly thumbprint in near future.

A major key management issue involved here is that passport control officer might need to know the private-keys and whether US government wants to share its private-keys to other countries' governments.

Problem: A third party pings the e-passport which can make us vulnerable to:

- Stealing identity

- Figuring out I am an American (by pinging it by a fake readers), which might be bad for various reasons!

Solution? Bar code might be a better solution since someone in 5 feet away can not read it or communicate with it.

Solution? Maybe implement something that it is active to respond *only* when the passport is open.

Solution? The scheme below:

$$\text{e-passport} \quad \xrightarrow{\quad nonce \quad} \atop \xleftarrow{\quad\quad} \quad \text{reader}$$
$$(nonce, ID)_{p_k}$$
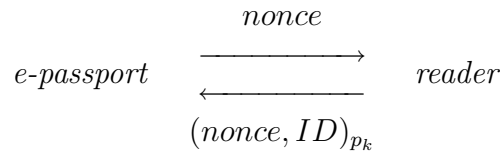
For this scheme to work, you do not want the response to be a deterministic function of the nonce, $\tilde{n}$, sent by the reader and ID, so the e-passport must produce a new randomness r and encrypt it along $\tilde{n}$ and ID.

$$\text{e-passport} \quad \xrightarrow{\quad \tilde{n} \quad} \atop \xleftarrow{\quad\quad} \quad \text{reader}$$
$$(\tilde{n}, r, ID)_{p_k}$$
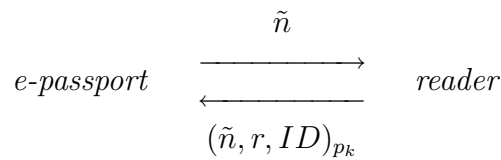
6 Libraries

Some libraries found RFID beneficial and have adopted this technology. Main advantage for libraries is self-check-out primarily for cost reduction benefit gained by adopting it. The RFID tag keeps information o f the book and the circulation information.

Problem: It puts the privacy of its patrons at risk since third parties might be able to read the RFID tags and infer, for instance, what books a particular patron has checked out.

Solution? Store a random 64-bit string as an index to the library database that matches with the desired information.

Problem: Tracking issues (the book I carry becomes a tracking device)

Characteristics of the solution: the scheme must let only the library to read the RFID tag and not any third party. In fact, it happens to be a cryptographic solution to the tracking problem.