An Architecture for Pseudonymous e-Commerce

Sandro Rafaeli
Lancaster University
Lancaster, UK
rafaeli@comp.lancs.ac.uk

Marc Rennhard
Swiss Federal Institute of Technology
Zurich, CH
rennhard@tik.ee.ethz.ch

Laurent Mathy
Lancaster University
Lancaster, UK
laurent@comp.lancs.ac.uk

Bernhard Plattner
Swiss Federal Institute of Technology
Zurich, CH
plattner @tik.ee.ethz.ch

David Hutchison
Lancaster University
Lancaster, UK
dh.comp.lancs.ac.uk

Abstract

Current e-commerce practice enforces a customer to disclose her identify to the e-shop. The use of credit cards makes it straightforward for an e-shop to know the real identity of its customers. Although, there are some payment systems based on untraceable tokens, they are not as widely used as credit cards. Furthermore, even without buying anything, a customer is already disclosing some information about who she may be or where she is by just connecting to the e-shop's Web server and leaving behind her IP address. The Pseudonymity System described in this document, makes it possible for a customer to buy electronic goods without having to reveal her real identity.

1 Introduction

One major goal of the IST ShopAware project is to build trust in electronic commerce to improve its acceptance. There are several technologies that can help to achieve this, e.g. by encrypting all communication between customer and merchant. Traditional stores offer a certain degree of anonymity in the sense that the customer does not have to give away his identity if he pays with cash. It is therefore desirable that online shopping offers this anonymity as well. On the other hand, anonymity may seem at odds with other security requirements such as authentication.

Generally, the process of buying electronic goods consists of three parts: (i) a customer (Bob) browses through the e-shop, looks at information about products, chooses the ones he wants, and fills them in his shopping cart; (ii) he proceeds to the checkout and provides a credit card, which is checked by the e-shop; (iii) after Bob's credit is cleared out, he has access to the products he has paid for.

Traditionally, none of these parts is anonymous. When browsing through the e-shop, IP-packets are sent from the customer's machine to the e-shop and vice versa. These packets contain the sender and receiver's IP-addresses. An eavesdropper or the e-shop can easily derive from those packets Bob's identity or at least the name of the computer he is using. When Bob has to submit his credit card, he discloses even more of his identity.

An anonymous record or transaction (Clarke, 1999) is one whose data cannot be associated with a particular individual, either from the data itself or by combining the transaction with other data. Examples for anonymous transactions are casting a vote in a ballot or a cash payment. A pseudonymous record or transaction (Clarke, 1999) is one that is identified by a pseudonym and the transaction cannot, in the normal course of events, be associated with a particular individual. This means that a transaction is pseudonymous in relation to a particular party if the transaction data contains no direct identifier for that party. But if a specific piece of additional data is available, then the transaction data can be linked to that party[1]. To be effective, a pseudonymous mechanism must involve legal and technical protections, such that the link can only be made (i.e. the index can only be accessed) under certain circumstances.

In the context of business in general, and e-business in particular, anonymity may lead to fraud. Pseudonymity is therefore a requirement for a trusted platform.

The Pseudonymity System consists of three parts: the Pseudonymity Network, the Pseudonymity Certification Authorities, which issue pseudonyms (Pseudonymous Certificates), and the Pseudonymous Transactions, which allows payment with pseudonymous credit cards.

The Pseudonymity Network (PN) enables browsing the Internet anonymously. It is not bound to e-commerce, but can be used for any browsing-activity in the Internet. It is based on a set of distributed proxies that are operated by independent institutions. It enables Web users to browse

[1] This piece of additional data could be an entry in an index that maps a party to its pseudonym.

the Web such that neither the Web server (or e-commerce site) nor any eavesdropper nor the independent operators of the proxies can find out where those users are, who they are, and where they are going.

We also provide customers with pseudonymous certificates, which are certificates not for their real identity but for self-chosen pseudonyms. On one hand, pseudonymous certificates enable customers to authenticate themselves as their pseudonym, but on the other hand, e-shops cannot derive the customers' real identities. The Pseudonymity Certification Authority (PCA), which has issued the certificate, is the only instance that knows the link between the real identity of a customer and her pseudonymous. This is needed to resolve the pseudonymity in the case the owner of a pseudonym tries to misuse the pseudonymous certificate in order to cheat. Revealing the real identity could be requested by a court order, for example.

Finally, our system is completed with the Pseudonymity Transaction. Credit cards are very popular in e-commerce and this is not likely to change very soon. We therefore want to provide credit cards for pseudonymous users. The idea is that there could be financial institutions (Pseudonymity Credit Card Providers (PCCP)) that would not require users to reveal their real identity to obtain a pseudonymous credit card. A pseudonymous certificate is needed to issue a credit card for that pseudonym. This is perfectly secure since the chain of PCAs described above can, if required, reveal the real identity of the pseudonymous user.

In the remaining of the paper, we define the requirements for anonymous e-commerce in section 2. We present our proposal to solve the anonymity problem using pseudonymous identities (section 3). Afterwards, the anonymity properties of the solution are analysed (section 4). We present the systems limitations in section 5. We describe other works done on the area of anonymity and pseudonymity in section 6. Section 7 gives an insight of the current status of development of the Pseudonymity System and discusses future work. Finally, we conclude our work in section 8.

2 Pseudonymity Requirements

A connection between two parties is anonymous with regard to a third instance if it is not possible for that instance to unveil more than one of the communicating parties. Note that a connection is still anonymous if one of the parties is detected. To make the term *connection anonymity* even stronger, we say that it should not be possible to determine any information about more than one of the communicating parties. By any information, we mean information that does not necessarily uncover a party's identity but that gives hints to identify the party. Connection anonymity can be achieved by disguising the communication path.

We should also consider *data confidentiality*. Even though an instance breaks the connection anonymity between two parties, the attacker should not be able to read any content of the exchanged data. We refer to this case as data confidentiality with regard to a certain instance.

Another term is *data anonymity*. Data anonymity means that the data in messages should not enable anyone to determine the identities of the communicating parties. Not even the two parties involved in the communication. This means hat Bob must not include any information about himself in the messages sent to Alice, since Alice could otherwise determine Bob's identity.

Therefore, anonymous electronic commerce should fulfil the following requirements:

Definition 1 *The e-shop should not be able to identify the customer's real identity.*

Definition 2 *It should not be possible to derive the sender's real IP-address from the source IP-address in messages sent to the e-shop.*

Definition 3 *When paying with a credit card, the e-shop should not be able to acquire knowledge about the customer's real identity from the credit card information provided.*

Definition 4 *On the other hand, the merchant should have the same guarantees about the validity of the credit card information as in non-anonymous payments.*

Definition 5 *The credit card provider should not be able to link the payment from a customer to an e-shop.*

In the next sections, we describe our Pseudonymous System that can achieves every and all of these requirements.

3 The Pseudonymity Service

The Pseudonymity Service is composed by third-party service providers and, as such, requires trust. The question is how much trust a customer is willing to have in a third-party service. One could argue that there is a single instance that offers an anonymity service for a customer to communicate with another party such that nobody besides the customer and the third-party is able to learn about the end-to-end connection (see Anonymizer (Cottrell, 1997)). However, this requires much trust in that single instance and not every customer may accept that

At least two entities should be in the middle of any pseudonymous communication: one closer to the client and the other closer to the server. We refer to these entities as Pseudonymity Entities (PE). The PE closer to the client knows about the client's location, but does not know her destination. It only knows the other PE. On the other hand, the PE closer to the server knows who the server is,

but does not know who is connecting to it since this PE only sees the PE closer to the client.

The information about the connection client-server is divided between the PEs, and these entities have to collude to recover the whole path. Each PE knows just a piece of the entire information. Furthermore, including more PEs in the chain between client and server can diminish the level of trust on every party, but one must bear in mind that such an increased trust in the system is achieved at the expense of increased operational overhead.

This pseudonymous property can be extended to other aspects of anonymity. Employing at least to PEs when generating certificates makes it possible to hide the real identity of a client without losing authentication (Definition 1). The same argument is valid for payments with credit cards. Using pseudonymous credit cards makes it possible to achieve anonymity requirements 3, 4 and 5.

Our Pseudonymity System consists of three distinct parts: the Pseudonymity Certification Authorities, which issues Pseudonymous Certificates, the Pseudonymity Network, and the Pseudonymous Transactions, which allows payment with pseudonymous credit cards. Each of these parts is composed by two or more PEs.

3.1 Pseudonymity Certification Authority

Before Bob can start browsing the Web and buying goods anonymously, he needs a pseudonym. If Bob simply wants to surf the Web, then he could choose any pseudonym. However, such a pseudonym would not help much when it comes to credit card payments, since nobody besides Bob could resolve the relation between the pseudonym and his real identity.

We strongly believe that e-commerce will use strong user authentication in the future. We therefore have designed a system that provides customers with pseudonymous certificates, which are certificates not for their real identity but for self-chosen pseudonyms. Pseudonymous certificates enable customers to authenticate themselves as their pseudonyms, so that e-shops can trust the customers without being able to derive the customers' real identities.

We introduce the Pseudonymity Certification Authority (PCA) as an instance that provides trustful pseudonymous certificates. The certificates are pseudonymous and not anonymous because the PCA knows the relation between a customer and her pseudonym. The correctness of the pseudonym-to-real name mapping can be ensured by, for example, verifying an electronic certificate endorsing the customer's real identity.

The PCA is the only instance that knows the link between the real identity of a customer and his pseudonymous. This is needed to resolve the pseudonymity in the case the owner of a pseudonym tries to misuse the pseudonymous certificate in order to cheat. Revealing the real identity could be requested by a court order, for example.

A Pseudonymous Certificate is a standard X.509 certificate (Housley and Polk, 1999) and is worldwide used for authentication. Therefore, no interface to e-shops is required. The only requirement for our certificates to be used is that the e-shops will have to accept PCAs as valid Certification Authorities (CA). This means including the PCA root certificate in the list of valid CAs in the e-shop database.

An e-shop does not have to have more confidence on a Pseudonymity Certification Authority than it has on a standard Certification Authority. The reasoning behind that is that the PCA generates a pseudonymous certificate based on a real certificate and since an e-shop, which would accept the real certificate, can thus accept the pseudonymous certificate. The trust e-shops have on standard CAs is based on the business model on which CAs operate. A CA lives of certifying identities, if the CA starts to issue faked or invalid certificates, its trustworthiness would be put in check and it would be out of business (Friedman et al., 2000). The same assumption can be used for PCAs, if they start mismanaging the generation of pseudonymous certificates their reliability is affected and they are out of business.

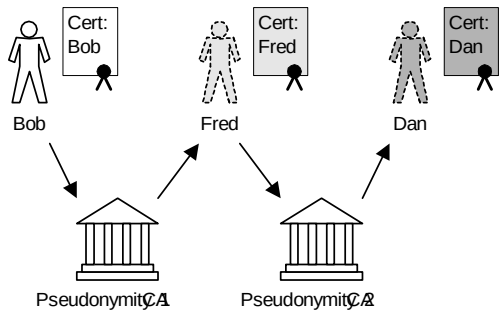

Figure 1: Pseudonymity Certification Authority.

A pseudonymous certificate can be used to obtain another pseudonymous certificate at another PCA, creating a chain of pseudonyms. Each certificate identifies a different pseudonym. This way, only all of the involved PCAs together can resolve the relation between the real identity and the last pseudonym in the chain. This greatly improves trust of the customers in the system.

For example, in (figure 1), Bob goes to PCA_1 and requests a pseudonymous certificate for a pseudonym *Fred*, and then goes to PCA_2 and, posing as *Fred*, requests a pseudonymous certificate for pseudonym *Dan*. Bob can then go to an e-shop and present himself as *Dan*. The e-shop accepts the certificate and thinks it is talking to *Dan*. The e-shop has no way to correlate *Dan* with Bob unless PCA_1 and PCA_2 collude with it. On the other hand, if the e-shop suspects that *Dan* (Bob) is trying to cheat, then the e-shop can go to a court of law and present its case. The court can order PCA_1 and PCA_2 to disclose their parts of the information and reveal that *Dan* = *Fred* = Bob, and thus Bob can be charged properly.

3.2 Pseudonymity Network

As seen in section 2, customers must not send data directly from their computer to the destination Web site since this would expose their IP address. At first glance, this might not be a big problem since most home-customers today get a temporary IP address assigned by their Internet Service Provider, which means that deriving the customer's identity from that address is not so simple (without cooperation from the customer's ISP, for instance). On the other hand, deployment of permanent access technologies (e.g. ADSL) also means that more and more home-customers will possess a permanent IP address, which is also the case for many office and university computers. Therefore, we want a solution that does not enable a destination to derive any relevant information, such as name, country, IP address, about the customer from the IP packet it receives (Definition 1). This means that the IP packets going from the customer side to the destination must be relayed by at least one intermediate proxy server. If one proxy server is used between customer and e-shop, then the e-shop's Web site sees the proxy's source address in the received IP packets and has no way to derive the customer's IP address (Definition 2).

However, there is an additional problem, namely that the proxy server knows about the end-to-end connection, hence there is no connection anonymity with regard to the proxy. It means that there must be at least two independent proxy servers between the customer and the e-shop. The knowledge is distributed among several proxies such that no single proxy knows enough to learn about the end-to-end connection between the communicating parties.

The Pseudonymity Network (PN) is a general mechanism that enables pseudonymous Web browsing. It is not bound to e-commerce, but can be used for any browsing-activity on the Internet. It is based on a set of distributed proxies that can be operated by independent institutions. It enables customers to browse the Web in a way that neither the Web server nor eavesdroppers and not even the proxies' independent operators can find out where those customers are, who they are, and where they are going.

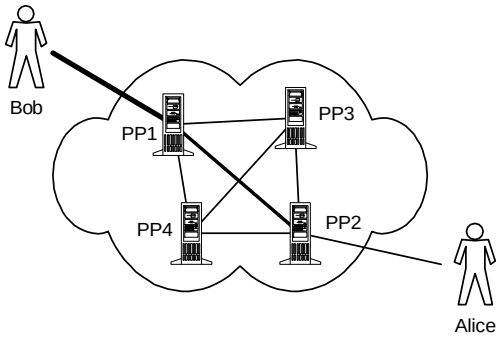

Figure 2: Pseudonymity Network.

The PN consists of two or more Pseudonymity Proxies (PP). The idea is that all communication between the customer and the e-shop is relayed by PPs. Figure 2 depicts a communication between Bob and Alice.

One cannot forget about traffic analysis. Basically, any of the involved parties (the proxy servers and the e-shop) or any other party could do traffic analysis to expose the end-to-end connection. Traffic analysis means that the eavesdropper monitors the incoming and outgoing links of a proxy and tries to correlate the packets. If this is done at all intermediate proxies between customer and e-shop, it could be possible to find out about the end-to-end connection. Therefore secure channels between every entity are required.

The secure channels can be established using the Diffie-Hellman (DH) key exchange protocol (Diffie and Hellman, 1976). In the DH protocol, both parties involved in the key exchange contribute to generate a shared secret key, which is then used to encrypt the communication between the parties.

Figure 3 shows how the channels are nested in order to protect the end-to-end connection anonymity:

1 Bob establishes a secure channel C_1 with Pseudonymity Proxy 1 (PP_1) and they share key k_1. Bob then issues a request to PP_1 to connect to PP_2;

2 Bob establishes a secure channel C_2, on top of C_1, with PP_2, using PP_1 as a proxy. Bob shares k_2 with PP_2. Note that PP_1 and PP_2 already have a secure channel LC_1 between them and hence share key kl_1;

3 Bob issues a request to PP_2, which cannot be seen by PP_1 because it is encrypted with k_2 (C_2), to connect to Alice;

4 Eventually, Bob decides to buy something from Alice and wants to pay for it; Bob establishes a secure channel (C_A) to Alice (on top of C_1 and C_2), and starts the payment process (see section 3.3). Bob and Alice share k_3.

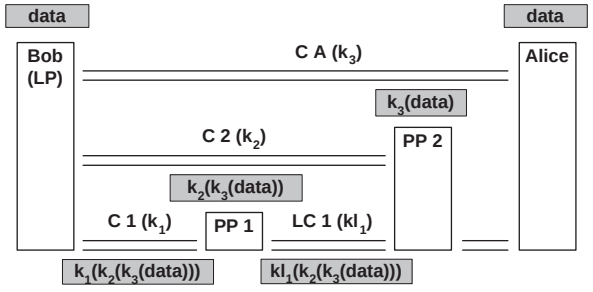

Figure 3: Layers of encryption.

Access to the PN is accomplished through a local proxy (LP), which runs on Bob's machine (or Bob's LAN) and appears as a normal Web proxy to the browser. Access to the PN is therefore transparent to the browser, and hence does not require any modifications.

Figure 3 shows Bob's LP, two PPs and Alice. There are two nested channels, one between LP and PP_2 and one between LP and Alice. When Bob sends data to Alice, his LP encrypts this data first with the key k_3, which he shares with Alice, then with the key k_2, he shares with PP_2, and then encrypts this with k_1, shared with PP_1. When PP_1 receives the data from Bob's LP, it decrypts the data with k_1, encrypts it with kl_1 and forwards the data to PP_2. PP_2 on its turn decrypts the data first with kl_1, then with k_2 and then forwards the data to Alice. Finally, Alice decrypts the data with k_3 and recovers the original message. The response from Alice follows the inverse path to Bob.

The result is that each of the three entities sees different data after removing the link encryption: PP_1 sees $[[\text{data}]_{k_2}]_{k_3}$; PP_2 sees $[\text{data}]_{k_3}$; and Alice sees the plaintext data. An external eavesdropper sees $[[[\text{data}]_{k_1}]_{k_2}]_{k_3}$ between LP and PP_1, $[[[\text{data}]_{kl_1}]_{k_2}]_{k_3}$ between PP_1 and PP_2, and $[\text{data}]_{k_3}$ between PP_2 and Alice. Since the data is differently encrypted between each pair of entities, the eavesdropper cannot correlate them.

Furthermore, authentication plays an important role in the Pseudonymity Network. In order to avoid unauthorised Pseudonymity Proxies to join the PN, the secure channels between PPs are double authenticated. PP_n and PP_{n+1} exchange certificates to prove to each other that they are who they claim to be. With double authentication we can guarantee that no malicious outsider can pose as a valid PP and join the PN.

Our PN implementation uses Secure Socket Layer (SSL) (Freier et al., 1997; Dierks and Allen, 1999) to perform the authenticated DH. Furthermore, the secure connection between Bob and PP_2 will be performed by a local proxy running on Bob's machine. After the channel is established, the local proxy gives it to Bob's browser and then Bob can access Alice's Web pages in the same way he would do without the PN.

The network of proxies hides the location of customers without limiting their access to online shops. The reason for that is PN does not need a special gateway to interface with e-shops, it uses only standard protocols, such as HTTP and SSL, to connect and talk to the online shop on behalf of customers. Thus, e-commerce Web sites do not require any technical modifications in order to accept connections from PN users.

3.2.1 Considerations

One can look at the PN and get to the conclusion that it is not a pseudonymity network; it is rather an anonymity network. We argue that we provide a certain level of anonymity throught a pseudonymous system. PN users are anonymous from the point of view of outsiders (since ousiders are not able to link origin-destination in a connection), but it is also pseudonymous because our proxies can make the link.

Accordingly to Clarke's definition, no matter how hard one tries, it can never be possible to *trace back* an anonymous transaction, and this holds true irrespective of how much information one has access to (note that this extra information can come from anywhere). Furthermore, "a pseudonymous record or transaction is one that cannot, in the normal course of events, be associated with a particular individual" and we go on saying that a transaction is pseudonymous if its data does not contain anything to link it to an individual, but that with additional info, that the transaction and the individual could be linked.

Note that a "transaction" on the network will usually involve two IP addresses: the client's and the server's. Now, PPs simply add a chain of IP addresses in any transaction, in a way that nobody, "in the normal course of events", knows the *entry* IP address (the client's) and the *exit* IP address (server's), and hence providing anonymity. However, each PP along the line has the opportunity to remember its *ingress* IP addresses (the IP address that connected to its server side) and its *egress* IP addresses (the IP address its client side connects to). Whether the proxies choose to remember this information or not by logging it, or even to collude or not, is irrelevant: there is an opportunity to "resolve" the chain of IP addresses and therefore link the client's address with the web server's address. The result is always the same: there is an opportunity to complete the chain back to the browser's IP address, and from there, one may get the user's identity.

Furthermore, we are trying to provide anonymity for privacy purposes, not for criminal purposes. If the user is not doing anything illigal and is just tryng to be unknown, then she would certainly have no restrictions with the logging. On the other hand, if a malicious-to-be user wants to use our system for bad-doing then she would think twice before using our system (she cannot be sure that the logs will be lost).

Moreover, for a system to be pseudonymous it does not mean it has to enforce that there must be always a mapping between pseudonym and real identity. What we want to make clear here is that although there is a possibility of tracing users we do not intend to provide everlasting traceability.

3.3 Pseudonymous Transaction

Finally, our system is completed with the Pseudonymous Transaction. Credit cards are very popular in e-commerce and this is not likely to change very soon. We therefore want to provide credit cards for pseudonymous customers (Definition 3). The requirement for such a payment system is that the customer should pay with her own credit card. Of course, she must not pay the e-shop directly but some pseudonymity entity. From the e-shop's point-of-view, it is the PE that pays in behalf of the anonymous customer. Obviously, PE must not be used alone because none involved parties are allowed to know everything about the end-to-end connection between the customer and the e-shop. This means that there must be at least two independent PEs between the customer and e-

shop during the payment process.

The idea is that there could be financial institutions, such as Pseudonymity Credit Card Providers (PCCP), that would not require customers to reveal their real identity to obtain a pseudonymous credit card. Just a pseudonymous certificate would be required to issue a credit card for that pseudonym. This is as secure as the verification of the customer's real identity by the first PCA in the pseudonymous certificate chain. As the chain of PCAs described above can, if required, reveal the real identity of the pseudonymous customer, for PCAs issuing pseudonymous certificates based on other certificates, the system is therefore as secure as these certificates.

Consider now that a customer goes to a first PCCP, provides his real identity, his real credit card, and the first pseudonym in the chain, and then the PCCP issues a pseudonymous credit card. All purchases made with the pseudonymous credit card are charged to the real credit card. The customer can go to a second PCCA, identify himself with the first pseudonym and his first pseudonymous credit card to obtain a pseudonymous credit card for the second pseudonym. This means that there is again a chain of PCCP, and only all of them together can break the pseudonymity.

If the customer goes now to the e-shop and provides a pseudonymous credit card, then the e-shop accepts it, because it can be validated as a normal credit card (Definition 4) at the last PCCA in the chain. The credit cards are validated backwards through the whole chain, which means that the payment is only approved if the real customer and his credit card are credit-worthy. This has the advantage that there is no financial risk for the PCCPs, since validation always goes back to the real customer's credit card (see figure 4).

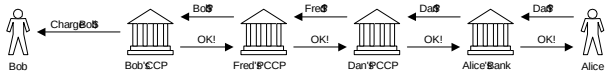

Figure 4: Pseudonymous Transaction.

The pseudonymity is maintained due to the fact that Alice's bank contacts Dan's PCCP for clearing the purchase out. Dan's PCCP does not know who Dan really is. At the other end of the chain, Fred's PCCP contacts Bob's CCP for clearing the transaction out, hence Bob's CCP believes that the merchant involved in the transaction is Fred's PCCP. Also, Fred's PCCP does not know who the real merchant is because it has been contacted by Dan's PCCP.

A pseudonymous credit card can be accepted as a normal credit card and, from the e-shop's point of view, the authorisation of the pseudonymous credit card transaction would be done exactly as it is done today with a normal credit card: through the e-shop's bank.

3.3.1 Considerations

The most important aspect of a payment process is its correctness. When a non-anonymous customer purchases some goods, the risk lies on the customer's credit card provider. The e-shop contacts the provider to check if the credit card is valid and that there is balance available in that card to cover the requested transaction amount, i.e. it is checked whether the customer is creditworthy or not. Based on figure 4, we see that the shop asks Dan's PCCP to validate the payment and Dan's PCCP asks Fred's PCCP to validate the payment. Fred's PCCP knows Bob and his credit card, hence it contacts Bob's CCP to validate the payment. If checking is not successfully validated, then none of the other validation requests will be validated. This implies that if Bob is not creditworthy, then Fred's PCCP, Dan's PCCP, and Alice will not accept the payment as valid. Therefore the payment fails. Note that this is exactly the same as for non-anonymous customers: if the customer is not creditworthy, the payment fails.

When Fred's PCCP is convinced that Bob is creditworthy, it replies to Dan's PCCP validating Fred's transaction. Likewise, Dan's PCCP replies to the Alice's bank validating Dan's transaction. Note that there is no financial risk for Fred's PCCP, Dan's PCCP and Alice. Fred's PCCP only validates the request from Dan's PCCP after it is convinced that it will receive the money from Bob (via Bob's CCP). Dan's PCCP only validates the request after it is convinced that it will receive the money from Fred's PCCP. Alice only delivers the goods after it is convinced that she will receive the money from Dan's PCCP.

4 Analysis

In this section we discuss why a certain customer Bob is anonymous for all entities in the PS and external eavesdroppers.

In figure 5, Bob connects to Alice through the PN and introduces himself to her as Dan. Bob uses Dan's certificate he had acquired at PCA_2 (2). Note that PCA_2 does not know that Bob is Dan, PCA_2 believes that Fred is Dan. Yet, PCA_1, which has issued Fred's certificate (1), only knows that Bob is Fred and has no clue that Fred is Dan.

Alice sees Bob's connection coming from PP_2 (4), therefore she does not know where Dan (Bob) is. PP_2 also does not know where Bob is, because Bob connected to PP_2 via PP_1 (3). Moreover, PP_1 does not know where Bob is browsing, because the destination address that PP_1 sees is PP_2's.

If Bob decides to buy something, he pays using Dan's credit card. Alice gets the credit card, but there is no way for her to link that card to Bob. The authorisation of the payment is done via Alice's bank, Dan's PCCP (7), Fred's PCCP (6), and Bob's CCP (5). Bob is charged by his CCP via his usual monthly bill (8). The bill contains a transaction made with Fred's PCCP. There is no link to Alice.

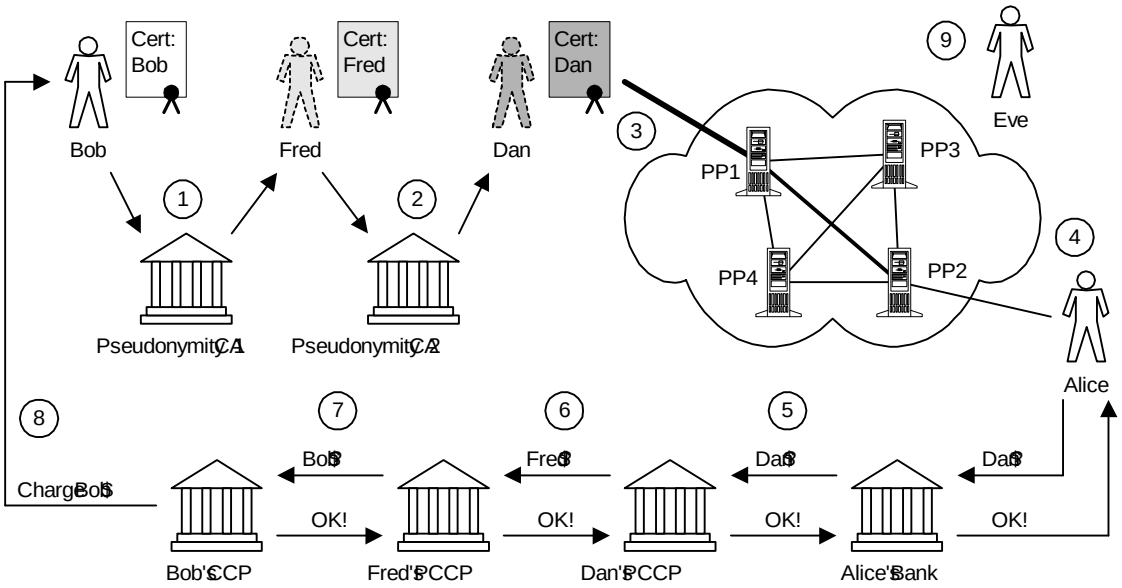

Figure 5: The Pseudonymity System.

Again, none of the involved parties can learn anything about the end-to-end connection from Bob to Alice.

Eve (9), who was trying to eavesdrop on Bob's activities, does not know Bob's pseudonyms or where he was browsing. Bob contacted PCA_1 and PCA_2 using secure channels, hence Eve could not read which Bob's choices for his pseudonyms were. When Eve was trying to find out where Bob was browsing, the only thing she could see was that Bob was connecting to PP_1. Bob contacted PP_1 using a secure channel, thus when Bob requested a connection to PP_2, Eve did not see it. When PP_2 connected to Alice, Eve had no way to know Bob requested that connection.

Table 1 summarises the required collusions in order to break the pseudonymity; pseudonymous connection and data confidentiality. By breaking the pseudonymity we mean to unveil Bob's real identity. By pseudonymous connection, we mean to recover where Bob was when he used the PN. Finally, by breaking data privacy, we mean to know what Bob bought from Alice. We conclude that only the complete set of PPs, PCAs and PCCPs colluding can break the security of the Pseudonymity System.

Another considerations were given throughout the paper.

5 System Limitations

Although our Pseudonymity System provides reliable pseudonymous Web browsing, where no single entity has the total knowledge of the end-points, it suffers from some limitations. Java applets and java scripts pose as a serious problem when a secure channel between customer and merchant is used.

Whilst a customer is browsing using an open channel (not encrypted), the local proxy can remove HTTP headers and java references from the HTML files. However, when a secure channel is established between customer and e-shop (usually during checkout) the local proxy can no longer filter it. It means that some information about the customer can leak to the server. For further security, the customer should not enable Java applets and java scripts at least during checkout time.

Furthermore, Felten and Schneider described a timing attack (Felten and Schneider, 2000) that can be used against browser caching. The attack is performed with the attacker measuring the time the target's browser takes to fetch a specific resource. If the target has already gotten the specific resource, then its browser has it cached and hence the time to fetch the resource is smaller than to fetch it over the Internet for the first time. The attacker can use different techniques to force the target's browser to fetch the specific resource. It seems that this kind of attack can be successful against all known anonymity systems.

Recently, Yahoo was legally required to prevent French citizens from accessing auctions of Nazi material. It seems that French user could easily infringe their law by using the Pseudonymity Network to access those auctions. Yahoo would not be able to identify French users using our system. Not even the PS providers would be able to thwart this situation. Ultimately, Yahoo, or other concern sites would have to block completely access from users using such anonymity system.

Finally, our system is initially limited to electronic goods, such as printable books or digital libraries access. The reason for that is that for material goods, a delivery address is required and it breaks completely the customer anonymity. Although, the customer could have access to an anonymous mailbox where the purchases could be delivered to, some merchants may not be willing to accept

Table 1: Collusions between Pseudonymity Proxies (PP_i), Pseudonymity Certification Authorities (PCA_i) and Pseudonymous Credit Card Providers ($PCCP_{name}$) and Alice that can break pseudonymity aspects.

Colluding parties	Break pseudonymity	Break pseudonymity connection	Break data privacy
(PP_1 and PCA_2) or (PP_2 and PCA_1)	–	–	–
PP_1 and PP_2	–	√	–
PP_1, PP_2 Alice	–	√	√
PP_1, PCA_2 and Alice	–	–	√
PCA_1 and PCA_2	√	–	–
PP_1, PP_2, PCA_1 and PCA_2	√	√	–
$PCCP_{Dan}$ and $PCCP_{Fred}$	√	–	√
PP_1, PP_2, PCA_1, PCA_2, $PCCP_{Dan}$ and $PCCP_{Fred}$	√	√	√

it and require a real address. We can visualise a solution using a pseudonymous delivery system employing a chain of pseudonymity delivery companies, but the price burden imposed by such solution would make it prohibitively expensive for customers.

6 Related Work

The Anonymizer (Cottrell, 1997) is a tool for anonymising Web communications. The Anonymizer is a Web site that acts as a proxy for Web requests. A user's request for an URL is first sent to the Anonymizer, which then gets the desired Web page from the end server and sends it back to the user. The Anonymizer is vulnerable to traffic analysis because the data to and from the Anonymizer are not encrypted.

The Freedom network (ZeroKnowledgeSystems, 2000) operates by dynamically building anonymous connections within a network of Anonymous Internet Proxies (AIPs). During connection setup, the client determines the path (i.e. which AIPs to use) and generates an initial packet that contains decryption information (keys) for each AIP along the way to a destination server. During data communication, the client sends fixed sized packets through the network. The packets are first encrypted with the key of the destination, then this is encrypted with the key of the last AIP and so on and finally, it is encrypted with the key of the first AIP. The packet is sent to the first proxy, which decrypts the packet (i.e. strips off a layer from the onion) with the key received during setup. Each AIP along the way strips off the top layer and forwards the packet until it arrives at the destination. Since one layer of the onion is stripped off by decryption, incoming and outgoing packets cannot be correlated. This makes Freedom network resistant against traffic analysis. Our PN is very similar to the Freedom network, but they diverge in the technology employed. The Freedom network is built using a homemade protocol that still needs to be proven safe, while we employ the well-known SSL to create the secure channels.

A system called Crowds is presented by Reiter (Reiter and Rubin, 2000). Crowds provides a protocol to retrieve Web contents anonymously. The protocol works by collecting Web users in a crowd that performs Web transactions on behalf of its members. When a user requests an URL, this request is forwarded randomly to another member in the crowd. Whenever a crowd member receives a request from another member, it makes a random choice to either forward the request to another crowd member (chosen randomly) or submit this request to the end server to which the request was destined. The reply from the server uses the same way back. If he crowd is large enough, then neither the other members nor the server nor any eavesdropper outside the crowd can tell which member in the crowd initiated the request and the system provides anonymity in the sense that any crowd member could have requested the Web page. The system has its weaknesses against local eavesdroppers (that monitor the traffic within the crowd) and collaborating members.

The Lucent Personalized Web Assistant (LPWA) (Gabber et al., 1997) provides its users aliases where each alias consists of an alias username, alias password and alias e-mail address. The LPWA acts as a proxy and whenever the user has to submit username, password or the e-mail address, he uses predefined two character escape sequences (\u, \p or \@), and LPWA replaces them with the appropriate alias. The advantage of LPWA is that it provides the user a simple and effective way to generate and use pseudonyms.

7 Current Status and Future Work

We have implemented the Pseudonymity Certification Authority server and two of them are running at the moment for demonstration purposes: one at Lancaster and the other at ETH. Users owing a Verisign certificate can request a pseudonymous certificate from both sites. The prototype has been developed using OpenSSL (OpenSSL-ProjectTeam, 1999) on Linux platform.

We finished designing the Pseudonymity Network (there is a technical report (Rennhard et al., 2001) describing the architecture) and started implementing the Pseudonymity Proxies. By the time this paper is published we should have them up and running. We plan to have a PP running at Lancaster and another at ETH. After

they have been installed on both sites, we plan to start a thorough testing of performance and shortly after that to make the network available for users worldwide.

The PP's implementation is being done in Java. We decline the performance of the application on behalf of fast development and portability. We understand that rewriting the application in a more efficient language, such as *C*, will be just a matter of time once we have finished our tests.

By the middle of 2001, we should have a working demo of the system, which will include also pseudonymous transactions using pseudonymous credit cards.

8 Conclusions

We have presented an architecture aimed at improving trust in e-commerce systems. Based on the use of a Pseudonymity Service, we have shown a solution for providing an e-commerce experience emulating the "anonymity" that can be achieved in traditional shops. Our system provides pseudonymous credit card payment, and does not require any modification to the browsers, server base and financial networks. A simple interface to the pseudonymous system that will allow material goods delivery, although not discussed here, is also being considered.

The Pseudonymity Service gives to a real user a pseudonymous identity that can be used consistently in the Internet and therefore allow e-commerce sites to perform important customer management functions. Our solution is suitable for e-commerce, but is not restricted to it. It can be used for any Internet-activity where pseudonymity is desired.

Acknowledgement

The work presented here was done within ShopAware - a research project funded by the European Union in the Framework V IST Programme. Marc Rennhard would like to thank also the Swiss Federal Office for Education and Science for his sponsorship.

References

Clarke, R. (1999). Identified, Anonymous and Pseudonymous Transactions: The Spectrum of Choice. In *Proc. User Identification & Privacy Protection Conference, Stockholm, Sweden*.

Cottrell, L. (1997). Anonymizer.com. Web site at http://www.anonymizer.com/.

Dierks, T. and Allen, C. (1999). The TLS Protocol Version 1.0. RFC 2246.

Diffie, W. and Hellman, M. E. (1976). New Directions in Cryptography. *IEEE Transactions on Information Theory*, IT-22(6):644–654.

Felten, E. and Schneider, M. (2000). Timing Attacks on Web Privacy. In jajodia, S. and Samarati, P., editors, *7th ACM Conference in Computer and Communication Security 2000*, pages 25–32.

Freier, A. O., Karlton, P., and Kocher, P. C. (1997). The SSL Protocol Version 3.0. tls-ssl-version3-00.txt, expires May 1997.

Friedman, B., Kahn, P. H., and Howe, D. C. (2000). Trust Online. *Communications of the ACM*, 43(12):34–40.

Gabber, E., Gibbons, P., Matias, Y., and Mayer, A. (1997). How to Make Personalized Web Browsing Simple, Secure, and Anonymous. In *Proceedings of Financial Cryptography 97*. Springer-Verlag.

Housley, R. and Polk, W. (1999). Internet X.509 Public Key Infrastructure. RFC 2528.

OpenSSLProjectTeam (1999). OpenSSL Project. The Open Source toolkit for SSL/TLS. Web site at http://www.openssl.org/.

Reiter, M. K. and Rubin, A. D. (2000). Crowds: Anonymity for Web Transactions. *ACM TISSEC*.

Rennhard, M., Rafaeli, S., and Mathy, L. (2001). The Pseudonymity Proxy Architecture. Technical Report MPG-01-02, Computing Department, Lancaster University, Lancaster, UK.

ZeroKnowledgeSystems (2000). Freedom.com. Web site at http://www.freedom.net/.