Do I know you? User Recognition without Identification

Juho Heikkilä

<Juho.Heikkila@hut.fi>
Telecommunications Software and Multimedia Laboratory
Helsinki University of Technology

Keywords: *Recognition, privacy, anonymity, pseudonym, PKI, authorization, liability*

Abstract: *People are beginning to use the information networks for an increasingly wide scope of applications. The social and entertaining sides are important portals for non-technical people to get involved in the networked society. How do these people recognize each other and what kind of identifiers do they use for each other. How do they present themselves and how could they trust the things the other person says about himself.*

More problems are generated by the databases of registration and other information collected from the users, while security and privacy protection are not on the top of the list of importance. Fortunately in the last years, civil rights organizations have forced the issue of privacy higher on the legislator's list.

One way to protect people's privacy is to let them stay anonymous. Full anonymity, however, poses its own problems of lacking liability and continuity in relating to other users and services. For most purposes, unique identification is not required. Using strong security pseudonyms and third parties, the liability questions can be solved in such a manner that unique identification can be achieved when the demand is strong enough. In this paper I will discuss why such issues are important, present some criteria and an outline for a trust model.

Introduction

As humans, we define ourselves much in relation to others. We meet other people and relate to them, some seem interesting and we get to know them better, on the other hand, some are quite unapproachable. All these decisions, whether conscious or not, are based on the information we can gather about the other person: the looks, the voice, the attitude, gestures, to name a few. These very same attributes are the basis on which we recognize familiar people even from a crowd, or from a voice in a noisy room. In face-to-face contact making all those small decisions comes naturally, but when the environment changes, our decisions become more difficult. And most of us have a need to be able to trust the people we would call friends.

In the beginning, every community is a small one, and growth means also change. In the last decades, the network community has grown exponentially, from a relatively small group of selected researchers where it could be said that everybody knew about everybody, to a global media to which almost everyone in the most networked countries can have access to.

So how do these Jacks and Jills know what the other is really like and how do they later recognize each other? The virtual environment is very different from the real world and text based interfaces make it hard to interpret what

kind of person there is on the other side. Many of the problems people face when handling their relations online are due to the differences between the environments and the differences inducted to human behaviour by those differences.

For the purposes of this paper, I will discuss such virtual worlds where people interact with other people or services. In other words, I am excluding single user worlds, such as in many games or other Virtual Reality applications. My main emphasis is on situations where there is need to trust a previously unknown entity or recognize a known entity.

One possible solution would be to use recognition instead of identification. Recognition, as I here use it, refers to the act of acknowledging that this is a familiar person. Online environment makes it more difficult to recognize old friends as well as to know what kind of a person one is dealing with. I feel that the difference between recognition and identification is an important issue, and that services providing security through anonymity will become more important in our society.

For access control or computer aided recognition to take place, however, some sort or identification has to take place. The important questions are how to make illegitimate linking difficult yet allow tracing if liabilities are being neglected.

The emphasis of this paper is on concept rather than on implementation. The purpose is to further privacy, bring forward alternatives to strong identification and outline ideas for a Public Key Infrastructure to be used both for recognition between users and as access control to online services. As a case example, I shall use a chat environment where users can recognize old acquaintances and prove things about themselves, attributes like age and gender being the most obvious. An implementation of such an environment is to be part of my Master's Thesis.

The rest of this paper is organized as follows: Section 2 introduces names and identities from a more philosophical or humanist, rather than technical, point of view and section 3 continues to discuss privacy and problems related to it. Section 4 discusses technologies available to solve some of these problems. Then in section 5 some criteria are outlined for a solution. And in section 6 an outline solution for our example of chat environment and parties involved are discussed. Section 7 presents some further problems that remain or appear due to our solutions. Finally, section 8 concludes the paper.

2. Name and Identity

> *"We know of no people without names, no languages or cultures in which some manner of distinctions between self and other are not made." [1]*

Names are an essential part of us being human, they allow us to relate to others and build our identities. They are an essential part of understanding the difference between the self and the other. This distinction is at the very heart of us being conscious beings. [2]

Names and identities can easily be mixed up, especially in environments where unique identifiers are used and the identifier is referred to as the identity. To avoid confusion and misunderstandings caused by using the same word for both concepts, I urge the use of 'identity' for that which is the essence of an individual, and 'identifier' for that which is used to refer to that individual.

Human Identity

Forgetting the identification context, identity of a person is the definition of that person as a human. Such an identity is unique, there are no two individuals that are exactly the same, even if no apparent difference exists. Take a pair of identical twin sisters for an example, they can be so like each other that parents have difficulties differentiating them from each other. The girls, however, would always know which one is which.

From the perspective of an individual, that identity is Self. All other identities are something that is Other. Since we define ourselves much in relation to others, it may be seriously troubling if this separation is lost.

Human Names

Names are what people use to address different Others as well as to introduce their Self. Unlike identities, names do not need to be unique, not in the global sense. However, the namespace

needs to be large enough to avoid too many conflicts in addressing different individuals.

Names can also have a link to an identity, that what one is. A tangible example of how a name can reflect its owner can be found in the names of Native Americans. And in many cultures the coming of age was accompanied by naming, which was an important event. Even a person unfamiliar with one could get a hint of what that person was like, just from the name.

In western cultures, where such customs have disappeared, nicknames have become an important part of our lives. They are more personal than our given names. Yet, for some of us it it important to know the real names behind nicknames of our acquaintances, and, often without realizing it, people make assumptions about a person based on the name. The less is known, the more is assumed.

Unique Identifiers

Unique identifiers are something that are unique in a scope that is considered comprehensive. Inside one country, social security numbers are such. The important thing is that there are no conflicts on to which entity or account that identifier is linked to.

In a modern state with millions of people to tax and millions of bank accounts (etc.) to take care of, such identifiers ease the management and minimize mistakes. These identifiers are not designed for people to use among themselves, however. People usually have a small pool of people to recognize, conflicts using names are rare and easy to cope with. Few of us know the social security numbers of people outside the immediate family. These numbers mean little to us as people, they are not part of our identity, simply pointers to individuals.

Network Identifiers

Identifiers handled automatically by machines have to be unique, at least within the context. I mean that the entity to which a certain identifier refers, can vary in different spaces and times. As an example, there can not be two users with a same nick in the same chat room at the same time. In a different room or at a later time, however, the same nick can be used by a different individual.

Due to this demand for uniqueness, as the context is widened, the required namespace grows. This leads to network-wide names becoming hard for people to use. ICQ numbers are a good example, they are as non-descriptive as phone numbers and as hard to remember. Email-addresses at least usually have a name-like part, possibly made unique with a number, and a domain part. However, if the user has to remember whether the user part was jill99, jill-99 or jill_99, trouble is near.

These problems are one reason why in environments where users interact with each other, like chat rooms, names do not need to be unique in time. The situation is more like the real world situation with people in a room, only the amount of information required to differentiate between persons is needed. How could these users recognize each other online without needing to reveal their true identities?

3. Privacy

Merriam-Webster defines privacy as:[1]

1 a: *the quality or state of being apart from company or observation*

b: *freedom from unauthorized intrusion*

Like many animals, humans have a need to keep a territory, private space where uninvited ones are not accepted. Home is such a private place and violations are often punishable by law. What we do at home is nobody else's business and similarly our letters are private.

Orwell & Co.

In "1984"[3], George Orwell depicted a society where one third of people were watching the rest. The workings of STASI in ex-DDR can be considered similar. Yet both Orwell and STASI lived in the old analog world. Neither could automate their information gathering and management by using powerful computers, databases and networks. One last obstacle remains on the way of fully automated information gathering by listening to the information networks: knowing who is who.

[1] Dictionary definition is certainly a simplification, more complex and varying definitions can be found in various psychological and philosophical texts.

If it became easy to identify the user behind each message sent over the networks, that obstacle would be removed. Considering how essential the network is becoming in our daily lives, we should realize that identification must not become standard practice. Many feel they have nothing to hide, well, neither did the German Jews in 1930´s, but few years later even more would have died in the hands of the Nazis if identification would have been automatic.

Information Age

The world, at least the most industrialized part of it, is on the verge of the information age where information in itself will become the most important asset. The value of top companies based on intangible products has surpassed that of those based on tangible products [4].

At the same time the fundamentals of Intellectual Property Rights are being questioned as old fashioned [5] and the rights of an individual to one's personal information are being discussed in the US, where privacy protection has been minimal compared to Europe [6]. Even as the legislation is finally waking up, it can only go so far. National law enforcement can do little about databases in foreign countries and even less to information that is handed under the counter and never caught. As easily as end-users can exchange mp3-files over the network, corporations can do the same with databases. Further, both legislation and enforcement are slow processes, it would be much easier to prevent mishap than to try to fix things later.

Digital Threats

Processing and storage capabilities about double each year, at the same time as the data mining algorithms are being further developed and optimized. The capabilities of organizations databases with fast access are growing rapidly.

What can these databases contain? Usenet articles for example, many groups are being archived and posts can be dug up years later. Companies providing services collect registration data and those selling products collect credit card numbers. Basically, anything once put online, may stay forever.

Even if those maintaining the databases are trusted to not intentionally intrude on privacy, others may intrude on them. Crackers roam all over the network, competing on what they can crack, and digital warfare is being speculated on.

Need for Unique IDs

Yes, there is a definite need for unique identifiers and identification. States want to keep record on their citizens, citizens may want to fill their tax-forms online, or seek medical advice that requires access to one's medical record. Since the network doesn't provide for any strong identification, several projects of electronic identification has been launched. Examples include the Finnish FINEID [7] and Swedish SEIS [8] projects.

However, care should be exercised on where these identification schemes are used. Sadly, to inspire willingness in citizens to pay the additional cost of an electronic ID-card, a reader, and software, the governments are supporting most projects somehow utilizing the cards. In many cases, a much more anonymous approach would be adequate, and even liability can be arranged without constantly intruding people's privacy.

Pseudonyms

People do not always actively realize how widely pseudonyms are in use today, even in the real world. Take any celebrity and you can be fairly certain that the name is not original. There are many reasons for using pseudonyms: one is having a fancy name, second could be protecting one's relatives from the merry-go-round of fame, third could be to hide from them, etc.

In real life situations, some feel awkward using a pseudonym, for some reason there seems to be a feeling of guilt, as if the only reason to appear under a false name would be to hide something criminal. Yet there is nothing criminal in pseudonyms in itself (naturally, fraud is a different thing). People can use pseudonyms for the very same reason celebrities do, to keep some things private.

4. Technologies

The very same technology that is used for identifiable identities could be used for pseudonyms as well. Thus pseudonym security technology can be just as strong. The more restricting parameter is cost. Any physical

security technology, like smart cards or other trusted computing base, comes with additional cost as well as cumbersome installation and use due to additional hardware required.

Public Key Cryptography

Public Key Cryptography (PKC) differs from symmetric cryptography in that instead of a key, a key pair is used. One key is public and the other is kept secret. Public key encryption systems can further be divided into two categories, symmetric and asymmetric. Since PKCs are also called asymmetric cryptography, it is important to not confuse these two different asymmetries. In symmetric public key cryptography, like RSA [9], both keys can reverse the application of the other, in asymmetric cryptography the reversion can only be done in one direction. Digital Signature Algorithm (DSA) is asymmetric, the private key is used to sign a document and the public key can be used to verify the signature.

Symmetric public key cryptography is practical when confidentiality is needed on top of integrity, and non-repudiation is required in communication between two parties. Since both have public keys, they have access to each others public keys. These public keys can be used to encrypt any traffic going to the key owner. And using one's own private key, a digital signature can be added to the message. The recipient can verify the sender with the sender's public key, and only the recipient can open the message since no one else has access to his private key. Using an asymmetric public key scheme would require each party to have two key pairs, one for signatures and one for encryption.

PGP [10] is a good example of using PKC for encrypting messages and distributing (public) keys with certainty added by trusted parties signing the keys (these are called 'Introducers'). Since the keys are long enough to keep collisions extremely unlikely, the public keys can be considered unique identifiers. Systems used for distributing and certifying keys are called Public Key Infrastructures (PKI).

Digital Certificates & Public Key Infrastructures

A certificate is a document that usually contains some kind of a statement and a signature or other method of verification. Letters of recommendation are one conventional example of a certificate. The Issuer certifies by a signature that the Statement about the Subject of the certificate is accurate.

Digital certificates generally contain the same basic information. The public key of the Issuer is used for identification and verifying the accompanying signature, the public key of the subject is used to define the subject and various kinds of statements are used depending on the type of certificate. Digital certificates can generally be divided into two groups:

> *Identity certificates* are used in PKIs to provide integrity into the distribution of public keys. A certification authority (CA) issues a certificate that binds the Subject's public key and the statement together. X.509 [11] certificates are usually identity certificates.
>
> *Attribute certificates* are used to make statements about the public key. In fact, Identity certificates are a subset of attribute certificates where the statement is identification. However, attribute certificates can be highly anonymous if the Subject public key is never bound to an identity. IETF's Simple Public Key Infrastructure (SPKI) [12] is more designed for attribute certification.

PKIs can further be divided by the topology used in certification. Whereas X.509 is hierarchic structure where each node is linked to each other in a certification tree (see figure 1 left side), non-hierarchic schemes allow independent certification, where anyone can be a certifier (see figure 1 right side). PGP and its trust model are non-hierarchic, anyone can sign the public keys and thus become an Introducer. However, only the ownership of keys, i.e. identity, can be certified. SPKI allows for both hierarchic structures and non-hierarchic ones. On top of that, being an open attribute certificate infrastructure, it hold any type of a Statement, so application specific certificates are easy to create.

The entity x in the PGP model might run into difficulties if he cannot find anyone to trust. In the X.509 model each node below the root has a single parent, which has to be trusted. And since

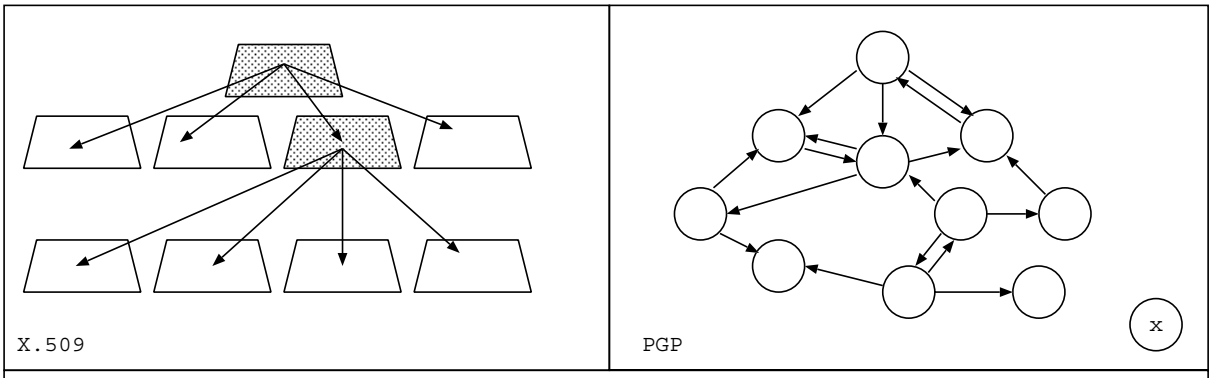

Figure 1. Certification paths in basic X.509 and PGP models

the root certifies each node on the level below it, trust is implied to the whole tree. Such is not always desirable either. However, if x is willing to trust some well-known entity, a Trusted Third Party, he could trust the statements made by that party like depicted in figure 2. However, x would still have no trust to the separate group on the right side in figure 2. Further, x could now decide, how far below the TTP the trust put to it might extend [13].

Trusted Third Parties

A Trusted Third Party (TTP) is an entity that is seen as trustworthy by the entity to whom the certificate is presented. In the certification business, the Certification Authority (CA) is a trusted third party, an entity that makes a statement about a possibly unknown entity, and based on that statement actions can be taken, like granting access to a resource.

Especially in an environment where anyone can make statements, the important question about Trusted Third Parties is who to trust. In a purely peer-to-peer certification network, like PGP, the question is further emphasized. If the user has no idea of who has signed the certificate, not much trust can be put on it. Yet anyone making a statement is generally a "trusted" third party. Therefore, as in any real world statement, it becomes important to define trust in different parties. Known people are always more trustworthy than unknown people and well-known organizations, like population register center, can be considered highly reliable for certain kinds of information. In our chat identity example, the population register center would be ideal certifier for things like gender and age.

When dealing with Trusted Third Parties, the trust usually needs to go both ways. The TTP need to trust the user to keep the private key private and the user needs to trust the TTP to not leak the given information to outsiders. Total anonymity is often impossible since identification brings liability and liability is often required for trust. For a Trusted Third Party to be able to certify things like age and gender, identification should be required. A TTP that certifies attributes that it is not certain about is not worth the trust. When dealing with matters of greater liability than age or gender, as is case in a payment service, the liability aspect is further emphasized. An abuser needs to be brought to justice and if there is no way to link the pseudonym to a real identity, not many merchants are likely to take such a payment seriously. A merchant needs to be fairly certain

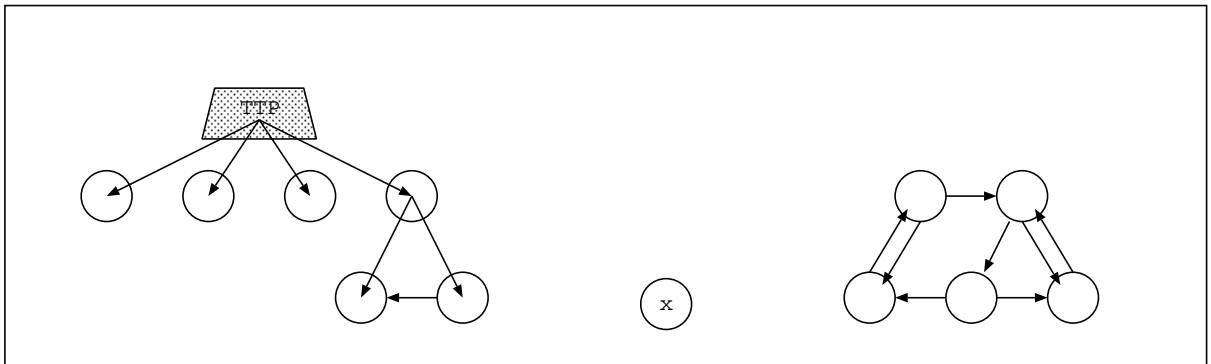

Figure 2. A model with a generally trusted third party and an untrusted group.

that he gets money for his goods or services.

In this attribute certification no single party can, or should, be the ultimate source of information. In the information age, no one should have monopoly over information. Businesses should also realize that privacy protection is not just a legal complication, but it is likely to become an important business opportunity. Different parties can certify different kinds of information and multiple certifiers help keep the user's personal information unlinked. Competition in the certification business should keep the prices acceptable and the information reliable, at least for the successful brands. [14]

Smart cards

Smart cards are small computers, usually a single chip embedded onto a plastic card, like a credit card or a mobile phone SIM-card. Smart cards are usually considered to be trustworthy as a computing base, since attacking them is difficult. Since smart cards can be designed to not allow the private key to leave the card, all operations are to be kept on the card, and even the legal user of the card should not be able to access the key directly. Also, unlike magnetic cards, smart cards are considered impossible to copy. This makes smart cards a likely platform for the electronic identity card systems.

Yet smart cards, like any real devices, are not perfect. The smart card requires an external terminal to interact with the user. The commands and the access code are given to this terminal to mediate to the card. However, there is no guarantee that a malicious device does what the user expects [15]. It might sign a different document from what it shows on the screen, it could store the access code for future reference, or it could even make multiple signatures instead of one. Also, there are cases, where, under laboratory conditions, secret information has been extracted from a card [16].

For the case at hand, smart cards also present the problem of cost. Any physical device has a price, and for smart cards there is the price for the card, and the price for a reader. A mobile user soon runs into the problem of finding a trusted terminal with a reader. If we yet consider that the security of the card is based on that information is kept on the card, backing it up would defeat this purpose. Further, the memory space of the cards is very limited, 16 KB being usual high-end to date. Therefore, smart cards are not a likely choice for storing pseudonyms, of which there may be plenty. However, they could be used as a relatively secure notebook for holding a key or passphrase used to decrypt a file containing the pseudonym keys.

5. Criteria

What kind of requirements would be put on a system providing safe interaction between people, as well as between people and businesses. Since there is no considerable extra cost for extra security, there is no need to keep the people-to-people connection any less secure than the people-to-business. Therefore, I feel that these cases can be considered fairly similar.

1. Anonymous to new Acquaintaces

Anonymity here means that the person's real identity can not be directly seen or derived from the pseudonym. In most cases no reason exists for the users to always shout their real identity at everything they see.

2. Traceability after Crime or Misuse

Should the pseudonym be used for something bearing legal responsibility, the users need to be traceable. For this trusted third parties are needed. And I feel it is important to emphasize that the third party has to be trustworthy to all parties. The users need to be able to trust that their true identity is not traced when it should not, and the law enforcement needs to trust that the third party can link the pseudonym and the true identity accurately.

3. Easy to Create new Identifiers

To allow people to use different identifiers on each of the services used, it is important that they can be easily created on the fly. Also, should an identifier be exposed, it is vital to be able to get rid of it and get a new one. In some services, like anonymous payment, getting new identifiers could be an important part of maintaining anonymity [17].

4. Secure Against Forging

The new identifiers should at least be more secure than conventional username-password pairs. And providing any financial action is to be performed, forging identities should be far more

difficult. A level of security provided by electronic identity smart cards should be seeked.

5. Usable for Access Control

As the new identifiers should be more secure than old username-password pairs, they could be used as a replacement for these. A simple challenge-reply test with changing challenge would also prevent anyone from overhearing the password. This would also allow one to prove the ownership of accounts.

6. Provider Independence

The identifiers must not come from a single source, any such source would find it easy to link them all together. Since we might want some of the identifiers to be next to anonymous and non-tracable, and ones used in financial transactions would require traceability, there must be different methods of providing tracability. Secondly, it would be difficult for any single entity to be certain of all the different kinds of attributes that might be required at different services.

7. Peer-to-Peer

If the users in a chat room only want to make certain that they recognize each other (from any imposter) the next time, there is no need for a third party provider. Using shared secrets could naturally suffice, but many people might feel awkvard to come up with strange phrases to recognize their acquaintaces. A technology based recognition service might be more comfortable.

Also, people often want to form groups among themselves without external control. And associations might want to issue their own certificates, like electronic membership cards.

8. Proof without Identification

In real world people meet and see things about the other people they meet. They trust their eyes to tell things about them. In cyberspace this is not so, all that is seen is text on a screen or a computer rendered presentation of what the other person wants to show. In most cases this doesn't cause serious problems, but if we consider the case where people are looking for new acquaintances, any misinformation can lead to at least disappointment. If users could prove certain attributes about themselves, like gender, age, location (to a chosen accuracy), this would help keep out the imposters.

9. Low cost

Services on the Internet have been mostly free to end users, and there is no reason should change. The most successful way of spreading something around seems to be offering the basic package for free, but without warranty or support. Additionally, from a human right and equality point of view I feel that the infrastructure providing security and privacy should be everyone's right. Support and extra services, which also create more expenses to the provider, may need to come with a price tag attached.

10. Accessible on different Terminals

Users might want to—or need to—use multiple computers for online access. It would be highly cumbersome to need to carry the identifier key pairs around on a dedicated device, an encrypted floppy disk or the like, therefore it would be nice to be able to retrieve the keys from the network itself. However, this puts extra requirements on the secure storage.

11. Secure Storage

The access to the private keys used for the identifiers must be highly secure. Even on a single user computer not continuously connected. Especially in case such a store would be accessible from the network, it should be encrypted strongly enough to stand long term scrutiny, even by organizations dedicated to breaking encryption.

6. Example Environment

Finally, let us take a quick look at what kind of parties are involved in our example case of a chat environment. Users who have never met before do not and should not need to trust each other before hand. Thus there is no need for a PGP style Introducer to certify the identity keys. The users may exchange public keys freely so that they can recognize each other later. For discussion purposes, however, normal, non-registered nicks are a natural choice. It would in fact be nice if the system is usable even without the recognition infrastructure. Our case users are Alice and Bob, who both come from behind

an ISP, that may or may not present security problems.

If Alice wants to prove to Bob attributes about herself, a Trusted Third Party is needed. Generally both Alice and Bob need to trust the TTP that certified the attributes. Alice needs to trust it because she seeks certification from it and she might need to identify herself to the TTP and does not want the world to know who she is. And Bob needs to trust the TTP to have adequately checked the information on Alice to be accurate. It may even be necessary to have different TTPs certify different things, especially since basically anyone can certify things and only parties considered trustworthy in a specific field should be trusted in that field. For example, an email-provider may be considered trustworthy to prove ownership of addresses on it's accounts, but may not be trustworthy to certify age or gender, because most web-based providers accept without checking any attributes the user gives at registration time.

The trust requirements for the service itself are different. It does not need to know anything about the users, unless the service wants to limit its use to certain groups and therefore wants the users to present certificates to it. However, since all the communication between the users is going through the service, it could check all pseudonym queries and interchange them to some of its own ones, thus performing a Man-In-the-Middle attack. If no third party certificates are passed between the users, the service could fool them in recognition. Passing third party certificates reduces the chance of such a fraud to minimal. This is one reason, however, why the service provider, or any party too closely associated with it, is not to be trusted as a TTP, at least, not as the only identifying TTP.

For mobility, the best solution would be a future smart card with enough capacity for multiple applications, one of which would be to store the key pairs and certificates. Alternatively it could encrypt the file, fetch it from a network location, decrypt it and use the keys to sign material provided by the external terminal. Of course, as long as the cards do not have direct human interfaces, the problems presented by a non-trusted terminal remain. If PDAs can be kept secure from Trojans etc., they might present a platform that is more secure than a PC, has more memory and processing capacity than a smart card, and comes with a display and input.

7. Further Problems

In the sections above, outlines of solutions to some of the problems have been discussed. However, some problems remain and some new ones appear.

Liability of the Anonymous

Anonymity always creates problems, as well as solves some others. People who can remain anonymous, feel they can do about anything without fear of being caught or even blamed for it. A prime example is spamming, most ISPs have forbidden this act of sending thousands or millions of copies of unwanted email. With the cover of anonymity and forged sender address, spammers still roam the net.

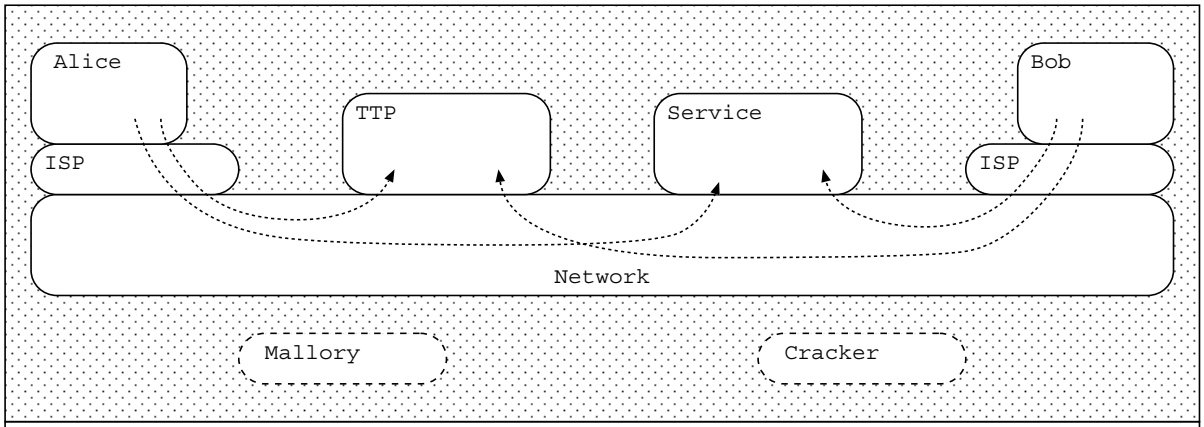

Figure 3. Parties involved when Alice and Bob interact through the Service using third party certificates. Mallory and Cracker represent possible threats of somebody listening to the traffic and somebody breaking into any of the computers involved. The arrows represent required trust.

When to reveal True Identities

One reason for third party certification is to bring liability. Trusted third parties that identify the user before certification, can be used to make the link between the pseudonym and the real person, should the need be great enough. As can be seen, one difficulty is where to draw the line, somewhere between liability and privacy protection.

In cases of proven crime, the line is clear. And even in such cases, the information should only be revealed to necessary parties. However, drawing the line even for those suspect of crime, is more difficult. If law enforcement is authorized to gather a complete database of pseudonyms, 'just in case it is needed', this again defeats the purpose. Even if we do not doubt the misuse of information by law enforcement, such a database in wrong hands is a powerful tool. Such wrong hands could include for example an invading power. Even if we trust the powers of today, we can not trust all possible powers to be. Also, databases can be attacked without invading physical space.

Publicity of Private Keys

On a personal level, abuse of anonymous pseudonyms is also possible. Private keys that are not truly private, can be used to transfer the identifier from person to person. Therefore, certifying something about somebody requires trust in that somebody, trust that the identity is not distributed. Trust in a TTP is based on the premises that its certificates are accurate. As an example, if a girl applies for a gender and age certificate for a new key and then gives the private key and the certificates to a male friend, none of the information may be correct. Should the fraud is discovered, it certainly is in the interest of the TTP to identify the girl for retribution. But what happens if the girl is not at fault, instead her identity was stolen and used. How could she prove that?

Stolen IDs

Stolen and lost identities are certain to appear. What can be done to minimize the damage? Certificate revocation and online checking both bring complexity to the infrastructure. By keeping the length of validity of the certificates short is one way of cutting out false certificates. But what about identities, should these even have a duration? If yes, keeping their validity short certainly helps revoke them as well as to provide anonymity, but it also makes them less practical for recognition.

Negative Recognition

When people are allowed to create dozens of identifiers for themselves, there is no way to be sure that this completely unrecognizable new identity does not in fact belong to an old acquaintance, perhaps a mischievous one. Not if we want to keep the identities separate, a service by a TTP could of course be used to query if this new identity belongs to a same person as any of these other keys. However, this would require one to give a list of one's acquaintances to the TTP, as well as turn the TTP into a party that binds pseudonyms together. It is noteworthy that the TTP can in fact do that already, but this is against the idea. Further, such service would only tell if that pseudonym is certified by the same TTP.

Social Binding

Finally, the information a user shares about oneself can lead to identification as well. However, there is nothing the technology can do about this. The technology can not say what information can or can not be shared and with whom. The goal is to keep the control of identification on the person rather than the system, to eliminate systematic identification but provide tools for people seeking other people. Asimov's laws for robots led to robots not allowing people to do anything, just to protect them from harm.

Identification through IP address

Further, a tracking problem not directly linked with this paper, but important for users' privacy. Even for a user trying to hide his or her identity by hiding personal information, the network itself can pose additional problems. Even in the case where no personal information is directly given, the address from where the surfing is done falls into the hands of the service provider. Internet Protocol (IP), the protocol used to send information from computer to computer. It is a connectionless, packet based protocol, therefore, to receive any feedback from the service, a return address must be provided. Sometimes, usually for users behind a more fixed connection, the return address of the user's computer is always the same, and sometimes as in case of dial-up connections, it may change from session

to session. Dial-up could therefore actually be considered a privacy enhancement. There are a few candidates trying to solve the IP-address problem, including technologies like Onion Routing [22], LPWA [19], Crowds [20], Anonymizer [21] and Freedom [22].

8. Conclusions

The field is certainly wide for a one man's quest. New problems appear from behind many a corner, and some solutions to one problem seem to create a different one. The important thing is to realize that the world is indeed changing and the networks are no longer the playground for nerds only. Cultural differences will also become more important as more and more different people get connected. These people may see privacy differently from the western point of view (which in itself is a generalization). Further research in the field is required and interdisciplinary studies are becoming more important as the needs of non-technical people are to be addressed. We cannot simply stand by and wait to see where the world is going.

References

[1] Social Theory and Politics of Identity, Calhoun, G., 1994, Oxford: Blackwell

[2] Lectures on handling social and work stress by Matti Ylikoski, National Public Health Institute

[3] 1984, George Orwell, 1948

[4] Jukka Kemppinen on the Seminar on Manuel Castell's "The Information Age", autumn 1999.

[5] Elektronisen kaupankäynnin kysymyksiä, Jukka Kemppinen, 22.1.2000

[6] Privacy as Intellectual Property, Pamela Samuelson, Stan. L. Rev. (draft; forthcoming 2000). http://www.sims.berkeley.edu/~pam/papers/privasip_draft.doc

[7] http://www.vaestorekisterikeskus.fi/fineid.htm [referenced: 7.8.2000]

[8] http://www.seis.se/ [referenced: 7.8.2000]

[9] A Method for Obtaining Digital Signatures and Public-Key Cryptosystems, Ronald L. Rivest, Adi Shamir, and Leonard M. Adleman. Communications of the ACM 21,2 (Feb. 1978), 120-126

[10] PGP Users Guide, Phil Zimmermann, 1994, ftp.pgpi.org/pub/pgp/2.x/doc/

[11] ITU-T Recommendation X.509, http://km.lanl.gov/docs/97x509final.doc

[12] SPKI Certificate Theory IETF SPKI Working Group, C. Ellison, B. Frantz, B. Lampson, R. Rivest, B. Thomas, T. Ylönen. September 1999. http://www.ietf.org/rfc/rfc2693.txt

[13] Modelling a Public-Key Infrastructure, Ueli Maurer, Proc. 1996 European Symposium on Research in Computer Security.

[14] Discussions with the finnish data security ombudsman, Reijo Aarnio, May 2000

[15] Providing authentication to messages signed with a smart card in hostile environments, Stabell-Kulo, Arild, Myrvang, Proceedings of USENIX Workshop on Smartcard Technology, May 1999

[16] Investigations of Power Analysis Attacks on Smartcards, Messerges, T., Dabbish, E., Sloan, R., Proceedings of USENIX Workshop on Smartcard Technology, May 1999

[17] SPKI based solution to anonymous payment and transaction authorization, Heikkilä, J. & Laukka, M., Proceedings for Nordsec '99.

[18] Onion Routing, Goldschlag, D., Reed, M. and Syverson, P., Communications of the ACM, vol. 42, no. 2, Feb. 1999.

[19] Consistent Yet Anonymous Web Access With LPWA, Gabber, E., Gibbons, P., Kristol, D., Matias, Y. and Mayer, A., Communications of the ACM, vol. 42, no. 2, Feb. 1999.

[20] Anonymous Web Transactions with Crowds, Reiter, M. & Rubin, A., Communications of the ACM, vol. 42, no. 2, Feb. 1999.

[21] www.anonymizer.com [referenced: 7.8.2000]

[22] www.freedom.net [referenced: 7.8.2000]