1

PyTHIA: Towards Anonymity in Authentication

Dimitris GRITZALIS[1] , Kostantinos MOULINOS[1], John ILIADIS[2], Costas LAMBRINOUDAKIS[2], Steven XARHOULACOS[2]

[1] *Dept. of Informatics, Athens University of Economics and Business*
76 Patission St., Athens GR-10434, Greece, e-mail: {dgrit,kdm} @aueb.gr

[2] *Dept. of Information and Communication Systems, University of the Aegean*
30 Voulgaroktonou St., Athens GR-11472, Greece, e-mail: {jiliad,clam,stx}@aegean.gr

Abstract There is a scale between authentication and anonymity, which is currently leaning towards the side of authentication, when it comes to e-commerce. Service providers and merchants are usually keeping track of user-related information in order to construct behavioural profiles of their customers. Service providers and merchants also correlate profiles of this kind, stemming from different sources, in order to increase their profit. This correlation is usually performed with the use of Unified Codes. Authentication, confidentiality, integrity, authentication, and non-repudiation are necessary functionalities for enabling e-commerce. Most of the currently used mechanisms that support these services do not provide anonymity. This paper presents PyTHIA, a mechanism, which is based on the use of Message Digest Algorithms and the intermediation of Trusted Third Parties in order to provide anonymity to e-commerce users who have to authenticate themselves in order to access services or buy goods from service providers and merchants respectively. With PyTHIA e-commerce users are able to authenticate without giving away any personal data and without using Unified Codes. In addition, PyTHIA ensures that service providers and merchants can effectively trace a customer in case he behaves maliciously.

Keywords Privacy Enhanced Technologies (PET), Security, Privacy, Anonymity, Certificates, Trusted Third Party, (TTP), PyTHIA

1. INTRODUCTION

Electronic Commerce (e-commerce) is expected to dominate business transactions in the future. Virtual markets and trade conducted over the Internet are anticipated to grow at an explosive rate. In 1996, Amazon.com recorded sales of less than $16 million, while in 1997 it sold $148 million worth of books [IMR98]. E-commerce eliminates the need of intermediaries, minimizes the product cost, and provides customers with worldwide market access. These are due to the wide use of data network technologies, and the evolution of the World Wide Web (Web). The Web attracted the average user to electronic business with its user-friendly interface. Despite its security problems [GRI99], the Web enabled people to interact using multimedia content.

In order to promote their sales, merchants are establishing new ways of collecting, processing and exchanging user data. Advertising is increasingly shifting towards the Web, as this communication channel fulfils promises for better targeting, more efficient response and more accurate audience measurement. During 1996, Internet advertising increased by a factor of ten from $20 million to $200 million while during 1997 it has risen to $600 million. The year 2000, $40 billion expected to be spent on Internet advertising [IMR98].

In order to measure the audience's marketing preferences and customize their product lines to specific user needs, merchants collect online personal data when a customer connects to their site. They further use advanced scientific techniques, such as data mining, to compile and analyse the data they had already collected, to form profiling databases. A user profile is a collection of personal data that uniquely identifies a person. The data collected for e-commerce purposes become critical tools in tracing potential clients' consuming patterns.

The collection and processing of personal data may lead to private and family life violation, thus discouraging the public from using new technologies. According to a Business Week/Harris poll [BUS98], lack of privacy in communications is the main reason of being off the Internet for the great majority of potential users. Users consider the lack of privacy to be a deterrent against e-commerce, even more than cost, difficulties in use and unwanted marketing messages. This situation would have a profound impact on the growth of the Internet with further consequences on the evolution of e-commerce and increase of advertising revenues [BUS98].

The antidote to online privacy infringement consists of channels that do not reveal the identity of the communicating parties. Such channels are called *anonymous channels*. Internet operation should be based on the principle of anonymity. If individuals wish to maintain the level of privacy

they enjoy in real world, they should be given the choice for anonymity in the Internet.

Deploying e-commerce infrastructures requires among others entity authentication, and confidentiality and integrity of the transmitted data. Protecting the confidentiality and integrity of data does not usually degrade the levels of privacy. Authentication, however, contrasts with anonymity. There is a scale between these two, and it is leaning towards the side of authentication when it comes for e-commerce. This is due to the fact, that strong authentication is based on the disclosure of the identity of the involved parties. On the contrary, anonymous communications do not reveal the identity of the involved parties. As a result, new technologies should evolve permitting the authentication of users while also facilitating their anonymity.

This paper presents an authentication mechanism that requires the intermediation of Trusted Third Parties (TTP), enabling Web users to authenticate themselves against the sites they visit and at the same time refrain from revealing any personal information. The mechanism averts personal data profiling and enables companies to trace the identity of a customer in case of fraud.

The paper is organized as follows: In section 2 an overview of Privacy Enhancing Technologies is presented, while in section 3 a framework is presented, the privacy mechanism should operate within. In section 4 we analyse the operation of this mechanism. Section 5 contains a discussion on the inner-workings of the mechanism and ideas for future enhancements. Finally, in section 6 some concluding remarks are provided.

2. OVERVIEW OF PRIVACY ENHANCING TECHNOLOGIES

Privacy Enhancing Technologies (PET) include those technologies developed to protect users from revealing their identity when they communicate with each other. In this section we focus on PET applied to Internet technologies.

The various PET mechanisms are strongly interrelated; many of them are based on recent technological developments and some blur the traditional distinctions between setting, implementing and enforcing privacy guidelines. The various mechanisms for the protection of privacy on global networks, according to their purpose, can be categorized as follows [OEC99].

2.1 Minimizing disclosure and collection of personal data

This category includes the following mechanisms:

- Management of cookies. Cookies comprise text files, formulated during the connection of a Web browser to a Web server via HTTP, and enabling the Web server to trace the on-line behaviour of the client.
- Anonymous re-mailers are e-mail servers permitting users to send electronic messages without revealing their identity.
- Anonymous re-webbers are proxies providing users with the ability to anonymously visit web sites.
- Anonymous payment systems. The most anonymous means of digital payment is electronic cash. Electronic cash comprises an electronic payment system that protects user anonymity and payment untraceability. In general, electronic cash schemes achieve these goals via digital signatures [LAW96].
- Digital certificates are digital tokens, issued by TTP, confirming the identity of the token holder. Digital certificates typically carry personal information. There is one category of certificates, which are used to confirm that a particular user is authorized to make a specific kind of transaction. These mechanisms do not directly reveal personal information.
- Anonymous profiles are those, which do not contain the personal identification information of a user. Each user is assigned a numerical identifier using cookies.

2.2 Informing users about on-line privacy policies

Various ways exist in order to inform users about the privacy policies adopted by web sites, including posted privacy policies, terms and conditions, and digital labels. Infrastructures exist supporting this practice. The most popular include TRUSTe, BBBOnLine, the OECD Privacy Generator, and P3P. The latter is a specification, developed by W3C [W3C99], enabling Web sites to express privacy policies in a standard format.

2.3 Providing users with options for personal data disclosure and use

Three practices belong to this category:

1. *On-line negotiation* of privacy standards through digital labels.

2. *Opting-in,* which refers to optional data fields and click-box choices commonly used by several Web sites to mark as optional several fields on the forms they use to collect personal data.
3. *Opting-out,* which refers to the ability of users to control the use of personal data they possess, either previously made known, or those being publicly available. This category includes the following mechanisms:

a) Controlling the use of personal data following the completion of collection, which refers to a common practice of several Web sites giving users the choice to change their mind and withdraw their consent to collect personal data. This is usually accomplished via e-mail.
b) Preventing the receipt of unsolicited e-mail advertising. The most popular mechanism of this category is Robinson lists, which include the names of all those people not wishing to receive electronic messages of advertising content. Legal authorities such as the national Data Protection Authorities in Europe usually dispatch the Robinson Lists to the public.
c) Opting-out of anonymous profiling which refer to the ability of users to erase collected personal data.

2.4 Providing access to personal data

This category includes off-line or on-line mechanisms permitting users to access personal data they have previously release. The Open Profiling Standard (OPS) is a standard for exchanging information between individuals and service providing parties. In addition, OPS supports user privacy by giving the end-user the ability to control the release of their personal data and track their exchange and usage [OEC98]. The standard specifies the following [W3C97]:

– Naming issues and rights of authorities regarding profile data.
– Varying levels of security of communicated data.
– Elementary profile operations such as *profile read* and *profile write.*

2.5 Protecting privacy through trans-border data flow contracts

This category includes all legal agreements and contracts between different countries, with respect to the protection of personal data. When studying these agreements, particular attention should be paid to the characteristics of data flow, including the nature of the data, the purpose and duration of the processing, the country of origin and destination of the flow, the data protection laws in the involved countries, and the security measures taken.

In addition, identifying the protection level "adequacy" offered by the destination country has become the most distinct debate with regard to trans-border data flow. The European Union Directive 95/46 [EUR95] and the Council of Europe Model Contract of 1992 [OEC99] have adopted the term "adequate level of protection", while OECD Guidelines state that trans-border flows may be restricted in case that no "equivalent" protection exists [EUR95].

Furthermore, one should define what the "adequate" level of protection is. For this reason, the European Union has set up a Working Party (under Articles 29 and 31 of the Directive) [EUR95]. Among other duties, this Working Party is responsible for giving the Commission an opinion on the level of data protection in the European Union Member States, as well as in third countries. In case there is no national legal framework, other means may be utilized in order to identify the adequacy of the data protection level. For example, the U.S. Federal Trade Commission follows a system of self-regulation, which established a set of data protection principles, called *Safe Harbour.* United States companies reassure their European customers that they respect individual privacy by compiling a list of companies complying with Safe Harbour principles.

2.6 Enforcing Privacy Principles

Enforcing privacy principles can be distinguished in two categories [OEC99]:

1. *Ensuring compliance with privacy standards.* Companies follow this proactive approach by reassuring their customers with regard to their compliance with national and international data protection practices and laws. In essence, data protection auditing is performed either by external or internal entities, which confirm that the examined organization actually has activated procedures and has taken measures to protect personal data, The entities that perform the audit can be internal data protection officers, third party reviewers, standards organizations, accounting firms, industrial firms, etc.
2. *Complaint resolution procedures for breaches of privacy standards.* Individuals follow this reactive approach when they believe that their personal lives have been violated. The resolution is usually made between the data subject concerning the breach and the data controller. Other means of resolution include private sector and industry bodies certification schemes, and administrative, civil and criminal proceedings.

2.7 Educating users and the private sector

Except for the entities directly involved in data protection matters, ISPs, Service Providers, and companies should promote the education of users with respect to mechanisms and practices they can use to protect their personal data.

There are, currently, several organizations that undertake this educative task, including Project OPEN (the Online Public Education Network), the U.S. Direct Marketing Association, the Centre For Democracy and Technology, the Electronic Privacy Information Centre "Call for Action" and TRUSTe, among others.

3. TOWARDS ANONYMITY IN AUTHENTICATION

Anonymous authentication is expected to contribute in the growth of e-commerce. However, there is a reverse analogous relationship between anonymity and authentication. E-commerce involves the use of on-line services and real time communication. The latter adds new challenges in protecting user anonymity while requiring the authenticated presence of users. We present a mechanism, called PyTHIA, which supports anonymous and authenticated communications. The three axes, our mechanism is based on, are the following:

1. Communication and user *anonymity* as a means to support anonymous profiling.
2. The existing *legal framework with regard to personal data protection,* which influences the deployment and release of anonymous communication.
3. *Authentication* in wide area networks, which is effectively implemented by using TTP services.

3.1 Anonymity

Anonymity is examined as a service offered and ensured by communication networks. Anonymous communication is a powerful means individuals have to ensure their privacy. One can distinguish four types of communication where the sender's physical identity is partly hidden [FRO96]:

1. *Traceable anonymity,* giving no clue about the sender's identity and leaving this information in the hands of an intermediary. Typically, the sender should trust the intermediary. Although traceable anonymity offers

the lowest security it permits the recipient of a message to trace back the identity of the sender in cases of repudiation between the involved parties.
2. *Untraceable anonymity* in which there is no way of revealing the identity of the sender.
3. *Traceable pseudonymity,* which assigns a pseudonymous (or 'nym') to the sender of message. The pseudonymous can be used to trace the real identity of the sender.
4. *Untraceable pseudonymity,* where a pseudonymous is assigned to the sender of the message as in traceable pseudonymity. However, this cannot be used in order to trace the real identity of the user.

Anonymity has both beneficial and harmful implications in peoples' lives. For the purpose of this paper, we focus on
a) privacy protection as a means for enabling anonymous profiling,
b) avoiding impersonation,
c) avoiding fraud in on-line transactions.

3.2 Legal framework concerning data protection

Although profiling may not change the amount of actual collected data concerning a person, organizing the data into searchable form reduces the person's privacy by permitting correlations that were previously impossible. In order to limit the impact of such processing on individuals' personal lives, several data protection laws have been enacted worldwide. The most renown is the European Directive 95/46, *"On the protection of individuals with regard to the processing of personal data and on the free movement of such data"* [EUR95], which sets the prerequisites for data owners and processors for collecting, processing and exchanging personal data. The U.S. government promotes the notion of "self regulation", a set of data protection rules applying to a plurality of market sectors, the content of which has been primarily determined by members of the specific trade sector.

Special emphasis has been placed on the use of Unified Codes, in several interpretations of 95/46 Directive. For example, article 8 of the Greek National Data Protection Law (L. 2472/97) [DAT97], states that the use of Unified Codes as a means of cross-linking personal data files, belonging to different data controllers, should be prohibited. This is due to the fact that using Unified codes may result in forming personal profiles within wider communities.

3.3 Trusted Third Parties

Not all TTP services can be supported only by technological means (e.g. in the case of non-repudiation service, there should be a legal body that

recognizes digital signatures as legal evidence). In addition, functions supported by technology may sometimes fail due to errors. To cover inadequacies presented in all these cases, entities using a PKI need to be aware of the legal principles and frameworks that support their use of PKI facilities and TTP mechanisms.

4. PYTHIA

We present a prototype for a mechanism called PyTHIA (PrivacY Through Hashes In Authentication) that supports *traceable anonymity*. PyTHIA users own a cryptographic construct called Privacy-Protected Authentication Token (PPAT), issued by an appropriate authority.

We consider Trusted Third Parties (TTP) can undertake this role, in the form of a value-added service. PPAT owners can authenticate themselves against Web sites offering products or services, using this token. However, no element of their identity is disclosed. If a user later repudiates his actions, the TTP can help in adjudicating the dispute by revealing the true identity of the entity, which used a specific PPAT to authenticate itself against a site.

The mechanism uses the security infrastructure provided by TTPs and digital certificates, as a means to trace the — certified — identity of users whenever this is needed. PyTHIA users must have obtained a digital certificate from a TTP, before requesting a PPAT and using PyTHIA. Although we were considering X.509v3 certificates [ISO95] while developing the mechanism, PyTHIA can make use of other categories kinds of certificates as well.

Throughout the presentation of PyTHIA, we assume that Alice wishes to use the mechanism to protect her privacy, while authenticating herself at Bob's web site. We also assume that Alice already possesses a valid certificate $Cert_A$ from a TTP called Trent, before requesting a PPAT from that TTP.

4.1 PPAT generation

We present the basic elements a PPAT comprises of, before analysing the PPAT generation process. The first element is the output of a collision-free hash function. The input to this function must be $Cert_A$ and a pseudorandom value RV produced at the time of PPAT generation. Actually, the first element of the PPAT is the output of the hash function applied n times to the aforementioned data. Trent chooses n, and the reason behind this choice is explained in the next section where we present in detail the PPAT generation process.

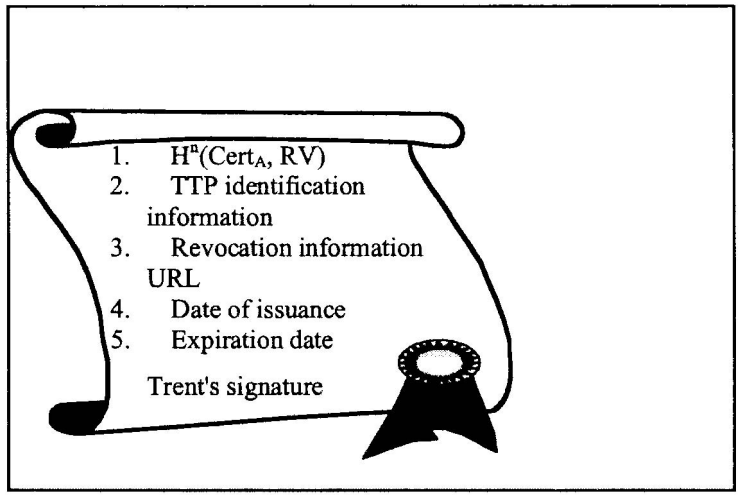

Figure 1: Privacy-Protected Authentication Token

The second element of the PPAT is identification information of Trent. The third element of the PPAT is a Uniform Resource Locator [LEE98] pointing to Trent's PPAT revocation status service. PPATs get revoked when the respective user certificates are revoked. The fourth and the fifth element refer to the date and time of issuance of the PPAT, as well as its expiration date and time. This must be equal to the expiration date of the respective digital certificate Alice has obtained from Trent.

Alice initiates the PPAT generation process by requesting a PPAT from Trent. Trent requests from Alice to authenticate herself using the certificate Trent has issued for Alice at a previous time. Trent computes the time period between the expiration date of $Cert_A$ and the current date. Trent proceeds with expressing the aforementioned time period in a predefined time unit (for the sake of simplicity we will be using *hours* as a specific time unit for our example). Having computed the amount n of hours contained in the aforementioned time period, Trent computes $H^n(Cert_A, RV)$.

Finally, Trent gathers the output of the aforementioned hash function, the information contained in the second and third field of the PPAT, the current time (fourth field) and the expiration date of $Cert_A$ (fifth PPAT field) and digitally signs them, using a private key reserved for that purpose only. The resulting construct is the PPAT of Alice ($PPAT_A$).

Trent stores $PPAT_A$ in his protected database, along with a link to (or a copy of) $Cert_A$, enabling him to quickly identify the owner of $PPAT_A$, whenever this is needed. Trent communicates to Alice $H(Cert_A, RV)$, that is the output of the hash function applied once on $Cert_A$ and RV. Trent also communicates to Alice the $PPAT_A$ itself, the number n and the RV. Alice stores this information at her protected, local repository.

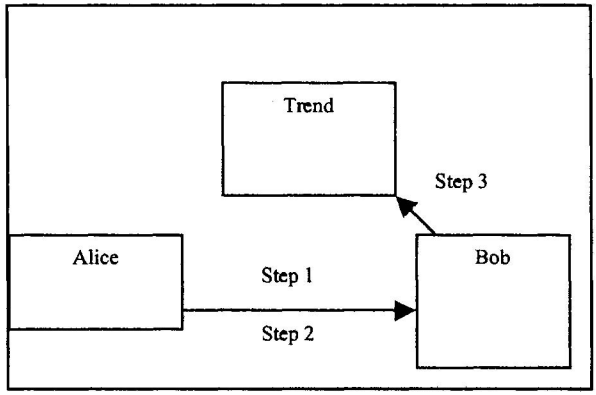

Figure 2: PPAT Authentication

4.2 Using PPAT to authenticate

Alice visits Bob and performs an action, which requires Alice's authentication lest she repudiates this action at a later stage. Alice communicates to Bob (Step 1) the $PPAT_A$. Although $PPAT_A$ does not disclose any personal information of Alice, it identifies Alice as a specific entity, carrying this unique identification badge and registered with the TTP that issued the $PPAT_A$. Alice must proceed with calculating the amount of hour k that has passed since the time the $PPAT_A$ was issued. Alice sends (Step 2) to Bob H^{n-k}, by recursively applying the hash function H n-k-1 times to the value $H(Cert^A, RV)$. Bob calculates k as well and verifies that the first element of the received $PPAT_A$ derives by applying k times the hash function H to the value H^{n-k} he has received from Alice. Alice is authenticated, since only Alice (and the TTP) could produce H^{n-k} at that time.

Finally, Bob has to send his identity (Step 3), $PPAT_A$, and H^{n-k} to Trent *or have this information time-stamped by an independent Time-stamping Authority (TSA).*

If Alice repudiates her actions at a later stage, Bob communicates the aforementioned timestamp to Trent, or requests from Trent to search his protected repository and locate the information Bob had sent him at the time Alice visited Bob. Since the exact time this information was made available to Bob could be verified and this information could be produced at that time only by Alice, therefore Alice cannot repudiate having visited Bob then.

However, Alice could claim having performed different actions at Bob's site, at that time. Bob has no means to prove that Alice had performed indeed the actions he claims she had. We discuss possible extensions to the mechanism to support this, in later sections.

4.3 PPAT revocation

Bob can verify the revocation status of $PPAT_A$, by querying the appropriate TTP service (the URL for this service is the third element in $PPAT_A$). Bob must send to this service the $PPAT_A$ and the service will check the status of $Cert_A$ and return that to Bob. The status of $PPAT_A$ always depends on the status of $Cert_A$.

5. DISCUSSION

PyTHIA is an authentication mechanism that proactively protects the privacy of personal data belonging to the authenticating entities. PyTHIA does not address privacy issues related to the underlying communication protocols and mechanisms used at a transaction, like the mechanisms presented in section 2.1 do. However, PyTHIA could be used in conjunction with some of those mechanisms, in order to decrease the leak of personal data due to the underlying communication mechanisms.

PyTHIA users do not need to trust that the entities they communicate with (and authenticate against) shall not attempt to collect their personal data, or that they follow any specific policy regarding privacy. The privacy mechanisms presented in sections 2.2 and 2.3 depend on that kind of trust, and primarily on the trust, users place on the authorities that audit the privacy policy - and its implementation throughout the business functions - of businesses.

Furthermore, PyTHIA users do not need to control the amount of personal data they give away, nor do they need to use mechanisms to retract personal data they had given away at a previous time. PyTHIA does not release any personal data at all, therefore it should not be required to provide mechanisms for data subjects to access the personal data 2.4 a company has collected for them.

PyTHIA could release, indirectly, personal data. In detail, personal data could be released through inter-business data mining. Future work on PyTHIA may provide solutions to this problem, as well. However, preventing inter-business data mining can also be achieved by using PyTHIA only in environments where privacy regulatory frameworks (as those described in section2.5) and voluntary compliance schemes (as those described in sections 2.2 and 2.6) apply. The technical measures by themselves could prove to be inadequate, either due to misuse from the data subjects themselves, or due to deliberate attacks by entities that attempt to violate the privacy of the aforementioned data subjects.

Technical measures should be enforced with related regulatory frameworks, and wide dissemination of information both on the technical measures and on the legal frameworks towards users. User awareness on privacy matters should be encouraged by authorities who regulate the protection of personal data, and should be promoted by entities that can successfully push information to end-users, such as ISPs, renown companies and organizations targeted to informing the public on privacy matters (also see section 2.7).

PyTHIA does not provide a mechanism for protecting the confidentiality or the integrity either of the exchanged transactional information, or of the exchanged information concerning the mechanism itself. Other mechanisms (e.g. SSL [FRE96] without client-side authentication) could be used in order to protect the confidentiality and integrity of information exchanged between Alice, Bob and Trent.

While investigating PyTHIA we have came up with various ways for providing Alice with the necessary information to authenticate herself against Bob. We have seen that the use of public key encryption could facilitate this task, in certain ways. However, we opted out of using public key encryption and we chose to use hash functions only, for a specific reason. If public key encryption was uses, then in some scenarios a private key compromise would potentially reveal Alice's personal information to all the Web sites she had visited up to that time. Since personal data can be considered highly sensitive or confidential, depending on the place and time of their use by Alice or data collectors, we preferred to opt out of using public key encryption.

6. FUTURE WORK

Alice is using the PPAT to identify herself to the Web sites she is visiting. The PPAT does not contain any personal data therefore no such data is leaked to these Web sites. However, if two or more Web sites collude into cross-referencing the PPAT they have collected from their visitors, then anonymous user profiles could be constructed. PyTHIA could be improved to deal with this threat. Alice could request and obtain more than one PPAT at a time from Trent, each one containing a different pseudorandom value RV. If Alice obtains r PPAT from Trent, then she will be able to visit at most r Web sites, excluding any possibility for those sites to cross-reference their visitor databases and construct a user profile on Alice. This presupposes that Alice will be using a different PPAT for each Web site she visits and that she will use no PPAT twice. However, this scenario can be quite unrealistic, since the number r of Web sites Alice visits (and to which she has to

authenticate herself) could be rather high. Issuing a high number of PPAT would result in high computational burden for Trent and high communication burden between Trent and Alice.

There is a balance between the level of privacy Alice wishes and the computational and communication burden this entails (see Figure 3). Furthermore, managing a high number of PPAT may become difficult for Alice, since she will have to track the use of her PPAT, in order to ensure that a specific PPAT is not used twice or at least is not used in too many Web sites.

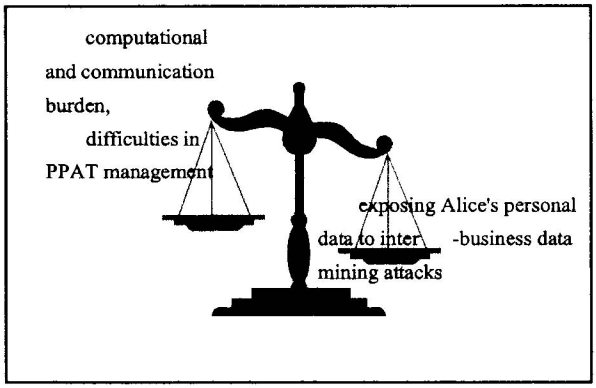

Figure 3: Consequences of managing numerous PPAT

Managing numerous PPAT could be facilitated if each PPAT is issued by Trent with a short, different - and potentially partially overlapping - validity period. Therefore, Alice must request a high enough number of PPAT in order to protect her privacy from inter-business data mining attacks, and at the same time minimize the consequences a very high number of PPAT requests would incur.

Another issue that has to be studied further in PyTHIA is to minimize the effects of a potential compromise of that part of the TTP that offers PyTHIA services. If Mallory succeeds in obtaining unauthorized access to the PyTHIA database, then Mallory would obtain personal data regarding all entities that have obtained PPAT from that TTP. All Mallory has to do is locate the PPAT of the entity, and retrieve the respective digital certificate. Trent could employ a mechanism to stall Mallory from discovering the aforementioned information and provide the time to deal reactively with the successful unauthorised access (block the access Mallory obtained to the database, or even monitor Mallory's activities and notify the PPAT whose identities have been revealed).

In order to stall Mallory, Trent could refrain from storing the PPAT themselves to the database, at PPAT generation time. Trent could store instead only the produced RVs in the database, and not link each RV to the corresponding PPAT and digital certificates.

$$H^{n}(Cert_i, RV_j), \forall\, i \in [1..\mathit{NumberofIssuedCertificates}] \text{ and } \forall\, j \in [1..\mathit{NumberofRandomValuesinDatabase}]$$

Equation 1: Mallory attempts to discover personal data for a PPAT owner, after having obtained unauthorized access to the PyTHIA database

This would increase much the time it would take for Mallory to discover the identity of a specific entity, since Mallory would have to retrieve the whole list of RVs, produce all the hashes described by Equation 1.

However, Trent would also have to perform all these computations whenever he would have to locate a specific digital certificate, based on a PPAT (e.g. when checking the revocation status of that PPAT). If the RV was stored in the PPAT, encrypted under Trent's private key, then Trent could immediately locate a digital certificate, based on the information provided by a PPAT, and at the same time if Mallory managed to obtain unauthorized access to the PyTHIA database, she would have to perform all the aforementioned computations.

Another improvement for PyTHIA concerns preventing Bob from claiming that Alice had visited him at an earlier time, than she really did. The present status of PyTHIA requires Bob to timestamp the authentication information he has received from Alice in order to prevent Bob from falsely claiming that Alice visited his site at an earlier point in time.

However, PyTHIA would be more efficient if Bob did not have to timestamp the aforementioned information. Solutions that would replace the need for Bob to communicate online with Trent or a TSA must be studied. We believe that these solutions could consist of including time-related information in the hashes produced by Alice, and making use of new technologies concerning digital signatures like forward-secure signatures [BEL99] or other cryptographic schemes.

7. CONCLUSIONS

We presented the prototype of a proactive mechanism for traceable anonymity. PyTHIA prevents any leak of personal data of a subject, when the subject is authenticated. PyTHIA can be used in conjunction with others, in order to provide a multilevel, integrated solution to the problem of privacy protection.

Furthermore, improvements to PyTHIA could prevent inter-business data mining, resulting in the construction of anonymous user profiles. There is still need for improvement in the suggested mechanism; the most important aspects that will be dealt with in the future are mentioned in section 6.

No PET mechanism by itself is sufficient for protecting privacy. Privacy clearly needs to be studied from a technical point of view. However, the technical mechanisms that protect privacy should be supported by an appropriate underlying legal infrastructure. Besides that, user awareness is a major issue. Until we achieve a satisfying degree of privacy-literacy, the privacy mechanisms and the legal infrastructures will not be able to operate efficiently.

REFERENCES

[BEL99] M. Bellare, S. Miner, "A forward-secure digital signature scheme", *Lecture Notes in Computer Science,* Vol. 166, M. Wiener (Ed.), Springer-Verlag, 1999.

[BUS98] Business Week, "A little net privacy please", 16 March 1998 (available at http://www.businessweek.com

[DAT97] *Data Protection Law* (Law 2472/97), 10 April 1997, Greece.

[EUR95] Directive 95/46, " On the protection of individuals with regard to the processing of personal data and on the free movement of such data", The European Parliament and the Council of the European Union, 24 October 1995.

[FRO96] A. Froomkin, "Flood Control on the Information Ocean: Living with Anonymity, Digital cash and Distributed Databases", *Univ. of Pittsburgh Journal of Law and Commerce* (also available at http://www.law.miami.edu/~froomkin/articles/oceanno.htm), 1996.

[GRI99] Gritzalis S., Aggelis G., Spinellis D., "Architectures for Secure Portable Executable Content", *Internet Research Journal,* Vol. 9, No. 1, 1999.

[IMR98] IMRG Ltd., *Electronic Commerce in Europe: An action plan for the marketplace,* White Paper, July 1998.

[ISO95] ISO/IEC 9594-8 (1994), *Open Systems Interconnection - The Directory: Authentication Framework.*

[FRE96] Freier A., Karlton P., Kocher P., SSL ver. 3.0, Netscape Communications Corp., 1996.

[LAW96] L. Law, S. Sabett, J. Solinas, "How to make a mint: The cryptography of anonymous electronic cash", National Security Office of Information Security Research and Technology, 18 June 1996.

[LEE98] Berners-Lee T., Fielding R., Masinter L., Uniform Resource Identifiers (URI): Generic Syntax, August 1998 (available at http://www.ietf.org/rfc/rfc2396.txt).

[W3C97] P. Hensley, M. Metral, U. Shardanand, D. Converse, M. Myers, *Proposal for an Open Profiling Standard,* W3C, 2 June 1997.

[W3C99] L. Cranor, M. Langheinrich, M. Marchiori, J. Reagle, "The Platform for Privacy Preferences Specification", 2 November 1999 (available at http://www.w3.org/TR).

[OEC98] OECD, "Implementing the OECD Privacy Guidelines in the electronic environment: Focus on the Internet", DSTI/ICCP/REG(97)6/Final, 27 May 1998.

[OEC99] OECD, "Inventory of instruments and mechanisms contributing to the implementation and enforcement of the OECD privacy guidelines on global networks", DSTI/ICCP/REG(98) 12/Final, 19 May 1999.