An Innovative Approach for the Protection of Healthcare Information Through the End-to-End Pseudo-Anonymization of End-Users

Panagiotis Gouvas, Anastasios Zafeiropoulos, Konstantinos Perakis(✉), and Thanasis Bouras

UBITECH Research, Athens, Greece
{pgouvas,azafeiropoulos,kperakis,bouras}@ubitech.eu

Abstract. Protection of data privacy and anonymity in the healthcare domain is of crucial importance and imposes many challenges since it regards a multi-fold and multidimensional process that needs to be safeguarded on multiple levels. Data protection has to be safeguarded at application and context layer, at session layer and at network layer. Taking into account the existing challenges, the scope of the current paper is to present the approach and conceptual architecture of SHIELD, an innovative methodological approach and network architecture-deployed within the framework of the FI-STAR project- targeting at the protection of healthcare information through the pseudo-anonymization of end-users. SHIELD aspires to provide value added services that could complement the service offering of the FI-STAR project in particular for the target sector of health care, and strengthen its technology basis.

Keywords: Data integrity · Data protection · Medical data management · Privacy · Pseudonymisation · Security

1 Introduction

The European Union is very sensitive with regards to the protection and integrity of personal healthcare information since processing of and access to healthcare data imposes severe legal and ethical issues, and has thus issued a plethora of guidelines and recommendations that need to be met in order to safeguard that the personal information of European patients is not forged, tampered with, or retrieved by unauthorised users in any way, as well as that the processing of information stored in clinical sites and healthcare facilities does not allow back tracking to patients, at least not without their informed consent [1, 2].

Privacy protection however is a multi-fold and multidimensional process that needs to be safeguarded on multiple levels. To comply with the aforementioned regulatory requirements, everything - the hardware, the software, the network, even the data itself - must be secured [3, 4]. As such, data protection in the healthcare domain needs to be safeguarded at (1) application and context layer, deterring phishing and impersonation attempts that may allow access to a (healthcare) information system where

R. Giaffreda et al. (Eds.): IoT360 2014, Part I, LNICST 150, pp. 210–216, 2015.
DOI: 10.1007/978-3-319-19656-5_30

personal information is stored, which an unauthorised user may retrieve and/or alter, allowing access to such a system and to the information it encloses only to authorised and verified users, (2) at session layer, deterring data flow manipulation, malicious attacks, interception and tampering attempts, that may allow the eavesdropping and/or tampering with the exchanged information during an open session between two non-authenticated parties, and allow end-users of a (healthcare) information system (e.g. a patient) to securely exchange information with this system over the internet throughout the period during which a session is active between the involved parties [5], and (3) at network layer, deterring real-time sniffing of network packets and network level traceability, protecting the end-users by preventing third-parties from monitoring the end-users' Internet connection, and retrieving the end-users' physical location which may in turn be used for malicious acts.

The scope of the current paper is to present the approach and conceptual architecture of SHIELD, an innovative methodological approach and network architecture for the protection of healthcare information through the pseudo-anonymization of end-users. The SHIELD platform approach was submitted at the Future Internet Social Technological Alignment in Healthcare (FI-STAR) Open Call for additional Project Partners and included in the FI-STAR project consortium. Towards this end, SHIELD aspires to provide value added services that could complement the service offering of the FI-STAR project in particular for the target sector of health care, and strengthen its technology basis.

2 Methodology

SHIELD will deliver a methodology, network architecture and software infrastructure by which the end-user adopts a new artificial identity, called pseudonym, pseudo-identity or pseudo-profile, provided by the proposed trusted and secure mechanism, inheriting and including all features and credentials needed by the existing Identity Management Generic Enabler of FI-WARE [6].

The proposed SHIELD Platform consists of three (3) sets of software components, facilitating end-users to preserve their anonymity while interacting with the FI-STAR healthcare applications and services, without restricting their privileges and benefits raised from their actual profile:

(1) The Pseudo-anonymization Networking Infrastructure, which guarantees the anonymity of end-users and the protection of the personal data, at network and session level, incorporating two main software artifacts: (a) the Pseudonimity Network Client that operates on-top of a standard TOR-client and is responsible for anonymizing the communication at IP-level, and (b) the Opportunistic & Ephemeral Negotiator that establishes a mutual agreement between the negotiating parties, i.e. the end-user pseudonymity client and the SHIELD server, with regard to the utilization of a symmetric key for encrypting the communication among them;
(2) The Pseudo-anonymization Application Layer, which guarantees the anonymity of end-users in the course of their interaction with given applications provided

though the FI-WARE platform, including two main software components: (a) the Pseudonymizer that facilitates the generation of discrete pseudonyms and pseudo-profiles for each end user (Pseudonym Generator), ensures that no reference to these pseudo-profiles exists from any profile publicly available on social networks (Social Clearance), issues a digital certificate for each pseudo-profile (Digital Certificates Creator) that will be used for the user's interaction with the provided applications and services, and stores the associations among real users and pseudonyms in an Encrypted Database; and (b) the Virtual Proxy that acts as an intermediary for requests from end-users seeking resources from the deployed technical implementations of the FI-WARE Identity Management General Enabler (Fig. 1).

(3) The Context Aware Services, which allows the preservation of the context coherence as well as the secure logging and monitoring of all access to personal data, including (a) an Activities Monitoring service that gathers all data related with the activities of a pseudonym, (b) a Secure Logging and Audit Trail Service, and (c) an Authorized Pseudonym Resolution Service that allows authorized third-parties to resolve the association between the pseudonym and the real identity of the end-user.

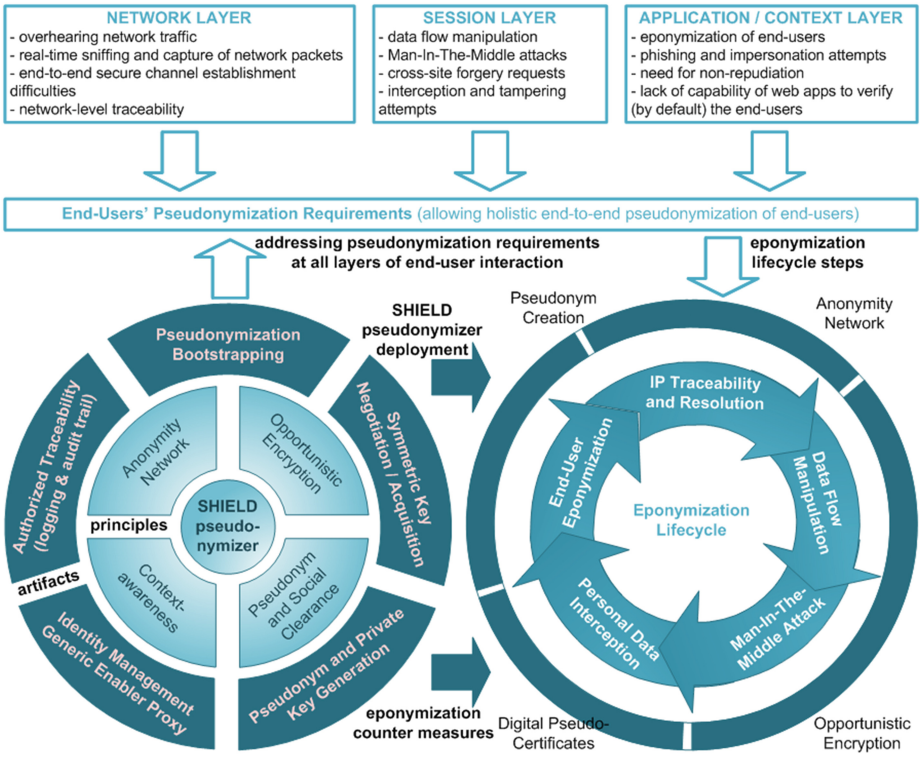

Fig. 1. SHIELD key technological concepts.

2.1 SHIELD Pseudo-Anonymization Networking Infrastructure

The SHIELD Pseudo-anonymization Networking Infrastructure will be based on the establishment and operation of a TOR network combined with the development of the SHIELD pseudonymity network client [7]. The pseudo-anonymization process begins with the installation and initiation of a Pseudonimity Network Client at the typical Web browser of the end-user. The Pseudonimity Network Client constitutes a JAVA-based client component that is able to be installed and executed on all major operating systems and encapsulates all the pseudonym acquisition business logic. This acquisition business logic is realized through a concrete interaction with the SHIELD Server. The network client will be used by the end-users of the provided services in order to join the TOR network and handle all the networking communication activities of the client, ensuring its anonymity in the IP network.

By joining the SHIELD TOR network, connections among network nodes will be based on the dynamic creation of virtual tunnels. Network flows within the virtual tunnels will follow random pathways through several relays. The path from the source to the destination node will be created on the fly and each relay node will be aware of only the predecessor relay. Network flows data will be encrypted multiple times while each relay node will be able to decrypt the "layer" of encryption that permits him to be aware of the successor relay. The negotiation of the per network flow set of encryption keys for each hop will be done by the SHIELD network client. Thus, the complete end-to-end path among the SHIELD end-users and the application servers will be unknown to all the relays in the network. Since the routing scheme followed for the establishment of communication among the SHIELD nodes of the TOR network will not follow any specific rules, traffic analysis from malicious users will be very hard to be achieved [8].

By ensuring the anonymity of the SHIELD users at network level, the SHIELD Opportunistic & Ephemeral Negotiator will be used for establishing a mutual agreement between the end-user clients and the SHIELD server. Through the exchange of a mutually-agreed symmetric key between the end-user and the SHIELD server through a 1-way SSL secure communication channel, end-to-end encryption between the sender and the recipient is supported. The key agreement takes place on top of a completely un-trusted network (although TOR provides anonymity, the risk of packet overhearing from a third-party still exists). In order to achieve a shared key agreement the principles of ephemeral Diffie–Hellman keys will be employed [9].

2.2 SHIELD Pseudo-Anonymization Application Layer

The Pseudo-anonymization Application Layer builds upon the functionality provided by the operation of the Pseudo-anonymization Networking Infrastructure and is responsible for the generation of discrete pseudonyms and pseudo-profiles and the deployment of the Virtual Proxy for the interconnection to the FI-WARE Identity Management General Enabler. Since a shared-key is already selected based on the Diffie–Hellman key exchange method, a secure channel based on 2-way SSL is going to be established over which the communication of the Pseudonimity Network Client with the SHIELD Pseudonymizer will be realized.

In order to support pseydonymization of end-users, the generation of the discrete pseudonyms and pseudo-profiles for each end user is required. These pseudonyms and profiles consist the unique identifiers of the users that are used later on for the generation of digital certificates. In order to create these pseudonyms, a combination of proper parameters has to be defined, ensuring that the correlation that arises does not correspond to any profile that is publicly available on existing social or other type of community networks. By creating such pseudonyms by the SHIELD Pseudonymizer, it is ensured that the de-pseudonymization cannot lead to identification of the end-user. Untraceable pseudonymization is validated through the Social Clearance Component of the SHIELD Server that will be based on existing state of the art tools, such as Maltego.

In case the evaluation of the produced pseudonym is successful, the SHIELD PKI can be used in order to create, manage and distribute a digital certificate for the pseudonymized user. The digital certificate is communicated and installed to the end users' PC/browser while it is also used in order to store the associations among the real users and pseudonyms in an Encrypted Database.

Upon the successful installation of the certificate, the end user is able to access the SHIELD server. Since this server acts as an intermediary Virtual Proxy for requests from end-users seeking resources from the deployed technical implementations of the FI-WARE Identity Management General Enabler, access is provided to the FI-WARE Identity Management (IM) services. User identification is achieved through the two-way SSL communication and the produced pseudonym of the end user.

One critical issue that has to be tackled regarding the Virtual-Proxying of FI-WARE Identity Management services is the fact that many authentication and authorization implementations check the end-user's IP. Since in our approach the IP is continuously changing (because of the TOR anonymity), the Pseudonimity Network Client will be responsible to configure the user's browser in order to use a predefined HTTP proxy, which is bundled with the SHIELD Server and provides the end-user with a predefined validated IP address.

2.3 Context Aware Services

In addition to the services that are directly provided from existing implementations within FI-WARE, access will be given to an additional set of context-aware services. Part of these services regard personalized services targeted at the end-user associated with a specific pseudonym and include access to personal data, view of statistics as well as monitoring of the daily end-user activities. Another set of context-aware services will be provided to authorized third-parties that will be able to resolve the established association between the created pseudonyms and the end users.

3 Discussion

SHIELD suggests a new network and software architecture targeting at the provision of high quality pseudonymized context-aware services. SHIELD will provide a holistic framework for the provision of pseudonymized services in the healthcare domain,

based on: (1) the design of the networking architecture that guarantees the anonymity of end-users and the protection of their personal data, (2) the dynamic creation of unique pseudonyms for the SHIELD end users based on the design and development of the SHIELD Pseudonymizer and (3) the deployment and operation of a PKI and issuance of digital certificates per pseudonymized end user.

By combining and integrating the abovementioned technologies, and coupling them with properly designed testing pilots, business planning and targeted dissemination, SHIELD will provide end-users with value-added services including: (1) preventing non-authorized third-parties to trace the IP or the physical location of the user, to intercept the personal data and the real identity of the end-user during his first interaction with the platform, and to trace back the association of the pseudonym to the end-user; (2) preserving application and context coherence implementing secure logging, audit trail and monitoring of all access to healthcare applications and services,; (3) preserving the non-repudiation security requirement allowing authorized parties to resolve the association between the pseudonym and the real identity.

SHIELD architecture and software paradigm can be used for the provision of advanced pseudonymized services within the FI-STAR and FI-WARE platforms, thus widening the opportunities to deploy new innovative services in the healthcare domain by exploiting the provided extensions in these platforms.

Acknowledgments. SHIELD is funded within the context of FI-STAR from the European Union's FP7 under grant agreement No. 318389.

References

1. EU Directive 95/46/EC of the European Parliament and the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Official J. EC L **281**, 31–50 (1995)
2. Callens, S.: The EU legal framework on e-health. In: Mossialos, E., Permanand, G., Baeten, R., Hervey, T.K. (eds.) Health Systems Governance in Europe: The Role of EU Law and Policy, pp. 561–588. Cambridge University Press, Cambridge (2010). ISBN 978-0-521-76138-3
3. Riedl, B.; Grascher, V.: Assuring integrity and confidentiality for pseudonymized health data. In: International Conference on Electrical Engineering/Electronics Computer Telecommunications and Information Technology, pp. 473–477 (2010)
4. Neubauer, T.; Kolb, M.: Technologies for the pseudonymization of medical data: a legal evaluation. In: 4th International Conference on Systems, ICONS 2009, pp. 7–12 (2009)
5. Jovanovic, D.; Mladenovic, D.; Blagojevic, D.: Implementation of ZRTP protocol for protection multimedia session. In: 19th Telecommunications Forum (TELFOR), pp. 246–249 (2011)
6. FIWARE Architecture Description, Open Specification, Security, Identity Management Generic Enabler. http://forge.fi-ware.eu/plugins/mediawiki/wiki/fiware/index.php/FIWARE.ArchitectureDescription.Identity_Management_Generic_Enabler
7. Chaabane, A.; Manils, P.; Kaafar, M.-A.: Digging into anonymous traffic: a deep analysis of the Tor anonymizing network. In: 4th International Conference on Network and System Security (NSS), pp. 167–174 (2010). doi:10.1109/NSS.2010.47

8. Kelly, D., Raines, R., Baldwin, R., Grimaila, M., Mullins, B.: Exploring extant and emerging issues in anonymous networks: a taxonomy and survey of protocols and metrics. IEEE Comm. Surv. Tutorials **14**(2), 579–606 (2012)
9. Diffie, W., Hellman, M.E.: New directions in cryptography. IEEE Trans. Inf. Theor. **22**(6), 644–654 (1976). doi:10.1109/TIT.1976.1055638