A SHORT OVERVIEW OF THE SECURITY CONCEPT OF THE GERMAN ELECTRONIC HEALTH INFORMATION INFRASTRUCTURE

Stefan Knipl and Ali Sunyaev
Faculty of Management, Economics and Social Sciences, University of Cologne, Pohligstraße 1, 50969 Cologne, Germany
{*knipl, sunyaev*}*@wiso.uni-koeln.de*

Keywords: Security analysis, Healthcare telematics, Electronic health card, Information systems security, Healthcare IS security.

Abstract: Germany is about to introduce a nation-wide healthcare telematics system, intending to connect existing information systems of all stakeholders. This could result in new threats to highly sensitive medical data. In this paper we shortly describe the security concept itself and point out a few possible problems at reaching the goals of information security.

1 INTRODUCTION

Stakeholders of the German health care system are not interconnected. Since the health insurance card currently in use in Germany is not a smart card it is therefore not able to provide secure transfer of medical data, thus preventing nationwide data crosslinking. This is about to change with the introduction of a new electronic health card in Germany (Schweiger et al., 2007).
The rollout of the electronic health card is on its way, hopefully leading to reduced costs in the health system, while boosting efficiency. Critics fear that one of its biggest strengths - merging information and making it available to care providers - may become a problem. This data is extremely sensitive and must not get in the hands of unauthorized persons. This paper analyzes to which extent this has been conceptually prevented.

2 SECURITY CONCEPT

This chapter describes the essential layout of the security concept of the German HTI.

2.1 EHC, HPC and SMC-B

Electronic health card (EHC), health care professional card (HPC) and security module card of type B (SMC-B) are the essential smart cards used in the HTI. Their memory is split in two parts: one is protected, containing secret keys and valuable data, the other one being not protected especially.
Accessing data in the protected area is possible by using one of three personal identification numbers (PIN). The gematik prohibits unprotected transfer of PINs, read-out of keys or being able to create a clone of a card (gematik, 2008b, p.69). The secret keys can be decrypted by entering the right PIN.
EHC PINs often will not be needed for a longer period of time and thus could easily not remain memorized or get mixed up. Patients could therefore use a sequence of numbers that is easy to remember (and thus easy to guess) or just write down the PINs and put them next to the EHC. Astolen card may then lead to the loss of confidentiality.
If a card has to be replaced, there has to be a possibility to access the data stored in the HTI (gematik, 2009b, p.46). To enable data transcoding a secret key is divided in parts and given to trustees (gematik, 2009b, p.60). However, this concept is not yet fully defined.

2.1.1 Pseudonymization and Anonymization

Data minimization, pseudonymization and anonymization are the means of choice to save as little personal data as necessary (gematik, 2009a, p.16).
Pseudonymization is used for the EHC. There are further pseudonymized certificates, not issued on the plain name of an insured person. If, however, one has

larger amounts of data of one single pseudonym then de-pseudonymization may be possible. Since insured persons get a new pseudonym with every EHC (i.e. at least every five years) in most cases this should be no real danger. More significantly, card issuers could perform a de-pseudonymization. This can only be done by a few selected people and only if the data protection official of the card issuer is involved (gematik, 2008a, p.170f).
If data is exchanged with a health insurance then care providers always have to be anonymized (gematik, 2008b, p.31). In addition the institution of a care provider has to be anonymized when querying insured person's master data (gematik, 2009b, p.38). This way one seeks to prevent the possibility of creating profiles of insured persons.

2.1.2 Data Access, Preservation and Transferability

Access on EHC, HPC or SMC-B is only allowed using the connector. However, the connector will not be available from start, so it may be set aside (gematik, 2009b, p.39), meaning that card readers can be directly connected to the clients of care providers. Once connectors are available the protected parts of insured persons' master data can then be saved to the protected area of the EHC. Until then they may temporarily be saved to the freely readable part of the EHC (for a time period not yet defined) (gematik, 2009a, p.15), as is the current state of health insurance cards (HIC). Insured persons have the possibility to access their own data with write access to emergency records and read access to the rest. Since prescriptions are relatively often filled by proxy persons the delegation of these rights to other insured persons will be possible (gematik, 2009b, p.92).
Every access to an EHC or into the HTI will be audited. Access will not be granted if auditing fails. Replacing an EHC does not mean loss of data, since existing data stored in the HTI can be transcoded.

2.2 Public Key Infrastructures

A public key infrastructure (PKI) is a trustworthy unit helping to check certificate validity and issuing new certificates. That way requests can only be conducted by registered components and can be verified, preventing man in the middle attacks and replay attacks. Mutual checks happen on all levels in the HTI. The smart cards that are used can perform such a check with a few restrictions even without availability of the HTI by using Card Verifiable Certificates (CVCs).
Certificates that are necessary for authentication must contain the role of certificate owners (gematik, 2009b, p.309). Otherwise a role assignment register would be needed. The PKIs in the HTI provide the possibility to verify X.509 certificates and CVCs (gematik, 2009b, p.49). The PKI in the HTI is divided in two parts: one for infrastructure, and one for persons and institutions (gematik, 2009b, p.141). It will not be possible to use services in the HTI by fraud with a fake certificate. Additionally a security alarm would be executed by a failed authentication (gematik, 2008b, p.127), allowing counter measures.
Requesting certificate state is done by using the Online Certificate Status Protocol (OCSP).In exceptional cases certificate revocation lists (CRLs) are allowed (gematik, 2009b, p.142). OCSP does not use blacklisting. It returns the state in the CA data base, making a revocation immediately active. OCSP answers may be cached, but there is no severe threat that a revoked certificate will be seen as valid by mistake, since they are rarely used in such short intervals (the cache normally lasts 12 to 24 hours) (gematik, 2009b, p.303f). It has to be considered as critical that non-availability of a revocation list of integrity must not lead to the result of all objects being considered as revoked as a precaution (gematik, 2008b, p.78). If the PKI is available, but no data of revoked objects can be accessed then all certificates had to be accepted if there were no contradictory data cached. Availability would otherwise not be given, but at least a security alert should be executed.
CVCs allow smart cards to independently verify certificates. They are used for mutual authentication of EHCs, HPCs and SMC-As/-Bs (gematik, 2008b, p.68). Since the public certificate of the issuing CA is known (gematik, 2008b, p.82), validity verification is directly possible. However it is not possible to verify if a CVC has been revoked. Corresponding X.509 certificates are verified when using services in the HTI. A revoked card may therefore not be misused if attackers can not also disrupt the availability of the HTI at the same time.

2.3 Data Transfer and Data Storage

Connection between card reader and connector is encrypted and mutual authentication is required (gematik, 2008b, p.51). Since part of the connection to professional services lie in the internet it is open for attacks. Therefore connectors have to be approved (gematik, 2008a, p.238) and they possess an integrated VPN client, allowing a connection using IPSec and mutual authentication (gematik, 2008b, p.121). At the other end of the connection a firewall should be passed before reaching a VPN concentrator, which may also include a packet filter (gematik,

2008b, p.120). Furthermore a VPN concentrator may not contain any backdoors in its software (gematik, 2008b, p.121).
Components in the care provider region should verify integrity and authenticity of the services they contacted (gematik, 2008a, p.155). Regarding security this should be obligatory.
It is not assumed that the inner regions of the HTI can not be reached by external attackers or are free of internal attackers, so transfer of unencrypted personal data has been prohibited (gematik, 2008a, p.153). Additionally plain text protocols, such as FTP, are not allowed to use (gematik, 2008b, p.98), as well as is the case for operating WLANs (gematik, 2008e, p.20). Between regions demilitarized zones (DMZ) should be established (gematik, 2008b, p.98). Furthermore intrusion prevention systems (IPS) may be used (gematik, 2008b, p.98). Employed firewalls may only open necessary ports and only allow connections as far as they are needed for providing services (gematik, 2008a, p.124). Additionally measures have to be taken to prevent spoofing (gematik, 2008b, p.97f). Unique message identities have to be assured to prevent replay attacks (gematik, 2009b, p.193).
XML signature and encryption are used (gematik, 2009b, p.108) to provide consistent end-to-end authentication and encryption on application layer.
Storage of data in the HTI is achieved in form of data objects encrypted with a symmetric encryption method. It is necessary to access an "object ticket" containing the symmetric key to the data objects. Object tickets contain a list of entitled persons, who may access this symmetric key. Insured people themselves are always entitled and able to entitle others.
It is important to prevent accesses without insured person's approval (gematik, 2009b, p.43), since this would mean compromising confidentiality.

2.4 NTP Service

It is necessary that all service in HTI got nearly the same system time to prevent replay attacks and CVCs with expired validity date to be seen as valid. The NTP service is periodically called by all systems in all regions answering with the current time. In total there have to be at least four servers to recognize a server returning a different wrong time to each client, respectively three servers to detect a server responding with a time shifted by a constant factor ("false ticker"). There must be at least four servers in the HTI and each data center must include at least three servers to still be able to detect false tickers if other data centers should not be available (gematik, 2008a, p.139f). Additionally all components accept no time changes of more than 1000 seconds. Since time may not be changed by entire hours or days, such an attack may fairly easily be noticed and will not be useful for exploitation.

2.5 Broker

A connector must contact a broker before it is granted access to professional services. The broker contains of the auditing service, the trusted service and the service directory service.
The auditing service is logging all attempts of accesses to realize legal certainty. The broker does not forward requests as long as auditing has not successfully been completed (gematik, 2008b, p.115f). Directly contradicting is the specification that "non-accessibility respectively non-availability of the auditing service MUST NOT lead to a complete stagnancy of the telematics infrastructure" (gematik, 2008b, p.114), especially since the auditing service has to be established in a redundant way (gematik, 2008b, p.114).
The Service Directory service (SDS) lists all services available in the HTI and their URLs. Items are added after compatibility testing and approval by the gematik (gematik, 2009b, p.196f). If attackers would be able to add an entry this would have no effect, if they could not also add an entry in the DNS service. However there would be consequences if attackers could change the URL of an existing entry. Without being able to change DNS entries requests could not be redirected to a fake server, though it could affect availability of a service. Changes in the SDS are able after successful authentication (gematik, 2007, p.32f). If it can be assured that employees handle their X.509 certificates carefully damage potential in this area is low. It should be considered if the four-eyes principle could be applied here since otherwise availability of essential parts of the HTI may be lowered.

2.6 DNS Service

The DNS service answers requests with the IP address corresponding to a given mnemonic address. DNS service is built in a redundant way, arranged in multiple levels. Precautions should be taken so that a synchronisation between different levels may not be able for unauthorized persons (gematik, 2008b, p.92f). It is therefore an attempt to achieve security by obscurity. Security of a system may not rely on this lack of knowledge.
To prevent attackers from redirecting requests away from valid services, DNSSEC will be used (gematik, 2008f, p.45), guaranteeing authenticity and integrity

of DNS transactions, although facilitating denial of service attacks because of its high calculation effort. Thus the number of DNS requests is permanently monitored. If a threshold is exceeded a security alert is executed (gematik, 2008b, p.87).

2.7 Connector and Extended Trusted Viewer

Connectors have a device specific identity in form of a X.509 certificate. Only VPN concentrators will be accepted as remote points, forwarding a connection to the broker. It has to be assured that the key of a connector's SMC-B may not be read out (gematik, 2008b, p.54) to assure connector integrity. A connector must be able to reliably detect spoofing of wrong system times. Besides, keys that are not trustworthy may not be adopted and data that needs to be protected may not be transferred unencrypted (gematik, 2008b, p.54f). Lastly, it also may not be possible to sign data, that a signer did not want to sign (gematik, 2008b, p.55f). This will be done by an Extended Trusted Viewer (xTV). The necessary X.509 identity of an xTV can also be deposited in the connector. If a signature should be performed it shows the certificate of the signer and the document to be signed, else the result of certificate validation (gematik, 2009c, p.38). An xTV can be realized by a software solution as well as by suitable hardware.

2.8 Cryptographic Specifications

Generally medical data may exclusively be stored and transferred in encrypted state (gematik, 2008a, p.95). Cryptographically the principle of forward secrecy is continuously implemented. If a key of lower level of hierarchy is broken this initially has no effect for keys on a higher level (gematik, 2008d, p.65). All keys that are not part of the lowest level of hierarchy and that are valid for at least a year have to be generated by truly random processes (gematik, 2008d, p.67), since even a complex generation of random numbers may hold the danger of attackers being able to reconstruct the process. Once key validity expires it has to be deleted including all its copies (gematik, 2008d, p.76), since otherwise old keys' protection requirements would be lowered, but it would still pose a threat to the goal of confidentiality.

3 CONCLUDING REMARKS

In this work we conducted a literature based analysis of the security concepts of the German electronic health card project. However, security always is a process and never a state. Thus a reliable and mathematically verifiable total fulfillment of all security goals can not be guaranteed (Wright et al., 2008). The HTI and its components are therefore designed for constant adjustment and enhancement. Ultimately the security concept of the gematik seems to be well elaborated in most parts. However this is not yet a conclusion for implementations in use, since these have not been tested in the frame of this paper.

REFERENCES

gematik (2007): Spezifikation Infrastrukturkomponenten: Registrierungsdienst (SDS) - Lokalisierungsdienst Stufe 2. Version 1.1.0, 2007.

gematik (2008a): Übergreifendes Sicherheitskonzept der Telematikinfrastruktur. Version 2.4.0, 2008.

gematik (2008b): Übergreifendes Sicherheitskonzept der Telematikinfrastruktur - Anhang B - Sicherheitsanforderungen. Version 2.4.0, 2008.

gematik (2008c): Übergreifendes Sicherheitskonzept der Telematikinfrastruktur - Anhang E - Kryptographiekonzept. Version 2.4.0, 2008.

gematik (2008d): Übergreifendes Sicherheitskonzept der Telematikinfrastruktur - Anhang F - PIN/PUK-Policy. Version 2.4.0, 2008.

gematik (2008e): Übergreifendes Sicherheitskonzept der Telematikinfrastruktur - Anhang G - Sicherheitsanforderungen Betrieb. Version 2.4.0, 2008.

gematik (2008f): Verwendung kryptographischer Algorithmen in der Telematikinfrastruktur. Version 1.4.0, 2008.

gematik (2009a): Fachkonzept Versichertenstammdatenmanagement (VSDM). Version 2.9.0, o.O., 2009.

gematik (2009b): Gesamtarchitektur. Version 1.7.0, o.O., 2009.

gematik (2009c): Konnektorspezifikation. Version 3.0.0, o.O, 2009.

Schweiger, A., Sunyaev, A., Leimeister, J. M., Krcmar, H. (2007) Information Systems and Healthcare XX: Toward Seamless Healthcare with Software Agents. Communications of the Association for Information Systems, 19, 692-709.

Wright, D., Friedewald, M., Schreurs, W., Verlinden, M., Gutwirth, S., Punie, Y., Maghiros, I., Vildjiounaite, E., Alahuhta, P. (2008) The Illusion of Security. Communications of the ACM, 51, 56-63.