Medical Informatics in a United and Healthy Europe
K.-P. Adlassnig et al. (Eds.)
IOS Press, 2009

doi:10.3233/978-1-60750-044-5-725

Authentication and Encryption in the Snow Disease Surveillance Network

Johan Gustav BELLIKA [a,b,1], Lars ILEBREKKE [b], Per Atle BAKKEVOLL [b], Håvard JOHANSEN [a], Jeremiah SCHOLL [b], Monika Alise JOHANSEN [b]
[a] *Department of Computer Science, University of Tromsø, Norway*
[b] *Norwegian Centre for Telemedicine, University Hospital of North Norway, Tromsø, Norway*

Abstract. The paper presents how authentication and encryption is implemented in the Snow disease surveillance network. Requirements for the authentication mechanism were collected from General Practitioners (GPs). The identity of each Snow user is preserved across health institutions allowing GPs to move freely between health institutions and use the system independent of location. This ability is combined with close to zero user account administration within the participating institutions. The system provides global user certificate revocation and end-to-end encryption.

Keywords. authentication, system administration, security, disease surveillance

1. Introduction

When an outbreak of a dangerous disease as meningitis is discovered in a patient population, a Disease Prevention Doctor (DPD) is responsible for raising an alert. Such alerts need to be distributed to all GPs and health personnel involved in diagnosis and treatment of communicable diseases. Alerts must be acknowledged to the DPD independent of the GPs whereabouts. After the initial alert, collecting and distributing data about the progress of the outbreak from/to all institutions involved is important to be able to prevent the disease from spreading. The Snow disease surveillance network is developed to enable extraction of anonymous epidemiological data from all kinds of health institutions, alerting of disease outbreaks and outbreak management. Users of the system can subscribe to periodical epidemiological reports, use Instant Messaging (IM) and participate in Multi User Chat (MUC) conferences dedicated to disease surveillance and local outbreak management. Users can also specify/search for epidemiological data amongst the participating health institutions. This paper presents the authentication and encryption solution implemented in the Snow disease surveillance network.

2. Method

The user requirements for the security system were discovered during a risk assessment process, collected using semi-structured individual interviews, semi-structured focus

[1] Corresponding Author: Johan Gustav Bellika, Department of Computer Science, Faculty of Science, University of Tromsø, 9037 Tromsø, Norway; E-mail: gustav@cs.uit.no.

group interviews, informal discussions with both medical and technical staff, and reviewing of electronic and paper documents. In total 13 different GPs were interviewed, with one having the role of DPD. The interviews where transcribed by a third person. Informed consent was obtained.

3. Requirements, Architecture and Design

The GPs are in general very concerned about protecting their patient's information. It is a strict requirement that no unauthorized person can get access to or be able to extract sensitive data from the EHR systems. The premise that the patient information is inaccessible for others than the patient's GP is the foundation for the patients' trust, which the GPs depend on. Based on this premise they strongly support a solution where the EHR system exports anonymous data to the Snow system. Even though the GPs are very concerned about security, they do not want to invest money in new security solutions, if this does not give them any specific benefit. With regard to mobility GPs move between different practices, and regularly appear in the accident and emergency unit (AEU) in the municipality. A local user account is normally created for each practice and AEU. These accounts function as their identification. The GPs state that they do not want to logon to the Snow system. However, satisfying this requirement is currently not possible. Secondly, the GPs prefer that the logon to the Snow system is done automatically when they logon to the EHR system. They feel that they have too many systems to deal with already, and thus express that system administration needs to be kept to a minimum. From rehearsal of pandemic outbreak situations we know that information about distribution of vaccines is very sensitive and that information to the public need to be handled by professionals on such situations. Because the outbreak management system is the natural place for administering distribution of vaccines, the information distributed within this system needs to be protected.

3.1. Architecture

In the current version of the Snow agent system [1, 2], each health institution that wants to provide data to the disease surveillance system needs to run a Snow server. By using the mobile agent concept (see [1, 2]), the Snow server makes it possible to extract and correlate data from multiple Electronic Health Record (EHR) systems without the need for central processing. This ability is critical for deploying a disease surveillance system in Norway, because Norwegian confidentiality laws prohibit centralization of patient information, with some exceptions. Snow enables system entities to

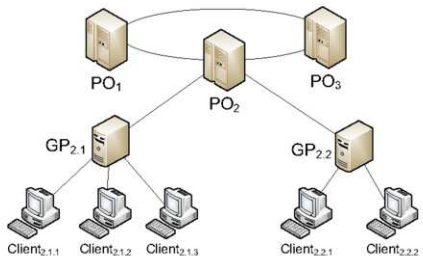

Figure 1. Organization of PO and GP servers in the Snow Disease surveillance network

communicate by sending and receiving messages using an Extensible Messaging and Presence Protocol (XMPP) [3] based routing overlay network. In this overlay structure, client workstations within each particular GP office connect to a common GP server. All GP servers within some region, like a county, connect to a common Post Office (PO) server. The PO servers of different regions connect to one another in order to facilitate global message delivery. As such, the clients, GP servers, and PO servers form a strict hierarchy, as illustrated in Figure 1.

Each Snow user is given a unique Jabber Identity (JID) in the form <health personnel number>@<domain>. The health personnel number (HPN) is a globally unique identity assigned to each health worker by the Norwegian health authorities. The domain contains the hostname of the GP office where the user connects. The JID facilitates XMPP message routing within the Snow substrate. Snow is implemented as extensions to the open-source Openfire (Jive Software http://www.igniterealtime.org/) XMPP server infrastructure.

3.2. Authentication and Encryption

Security in Snow is based on a strong notion of identity and confidentiality. It is critical that the sender of each message can be identified and that only the intended receiver of a message can read that message. To this end, Snow relies on an X.509 standard based public key infrastructure for authentication of users and servers, encryption, and message signing. Each user and server that wants access to the Snow infrastructure must first obtain an X.509 certificate from the offline Snow Certificate Authority (CA). Each user certificate binds the HPN, etc. to his public key. The CA attests this binding by signing the certificate with its private key after validating that the person holding the corresponding private key is indeed the person he claims to be. The private keys are kept secret by each user and are used for decrypting inbound messages and signing outbound messages. Outbound messages are encrypted using the public key contained in the recipient's certificate so that he is the only one that can read the message.

Each client and server stores a copy of the CA's root certificate so they may verify the authenticity of certificates passed to them. Client workstations store user certificates and encrypted versions of the private keys for every user that is allowed to access Snow on that workstation. Servers store their own certificate and private key, and a white list of other computers that may communicate with them. During TLS [4] connection establishment between a client and a server, the client presents its certificate to the server and proves access to its private key. If the certificate is valid, the client is logged into the server with the identity provided by the user certificate. If the user account is unknown, the account is automatically created based on information in the certificate.

3.3. Key and Certificate Distribution and Revocation

Users are initially given their certificate and private key via their health network e-mail accounts. These email accounts are created and maintained by the Norwegian health net authority and are also commonly used for EDI of encrypted patient information. Although we consider these e-mail accounts fairly secure, we add an extra level of protection by encrypting the private keys with a password that users receive via SMS. The users' private keys are then encrypted on the client workstations by a user provided password. When moving to another workstation, a user can bring his private

key with him on a portable storage device or, access it from his e-mail account and repeating the initial installation process.

All Snow servers are able to resolve a HPN to a user certificate, which is a part of the JID. The GP server stores the certificates of all local users. To support high availability and reliability, the POs store all certificates in its own region and all discovered certificates from other regions. When a new user certificate is discovered on a GP server, it is stored locally and pushed to the PO server. When a user wants to send a message, he first acquires the receiver's certificate by querying the GP server using the JID. Certificate requests are routed along the Snow server hierarchy (see Section 2.1) towards the PO server associated with the receiver using the domain information in the JID. Upon receiving a certificate request, a server first tries to find the matching certificate in its local store. If a match is found, the certificate is propagated back along the request's forwarding path. If not, the request is forwarded to the next PO server. If not found at this PO, then the certificate does not exist and an error is returned.

It is sometimes necessary to invalidate a certificate before its set expiry time. Typically this must be done in situations where a user leaves the organization or if a private key is stolen. To control the validity of certificates the CA publishes a Certificate Revocation List (CRL) containing the serial numbers of certificates that have been revoked by the CA. Whenever a certificate is used the CRL must be checked first. The CRL is signed by the CA to avoid denial of service attacks based on issuance of false revocations. Also, the CRL include a monotonically increasing timestamp so that an attacker cannot remove revocations from the list by replaying old CRL instances.

Conforming sites do not communicate with sites or users that have their certificates revoked. This implies that all TLS connections to and from the revoked entity must be considered tainted and should be torn down without prior warning. Whenever the CA updates the CRL, the new data is pushed to the POs. The POs then distribute the new CRL to the connected GPs, who in turn distributes it to the clients. Hosts that were offline during this time can poll the latest CRL from a PO before initiating other communication. Note that, in our case, pushing CRL updates is better suited than pulling them since the expected update frequency is relatively low compared to the required distribution latency. To ensure that the CRL does not grow indefinitely the CA removes revocations of expired certificates. However, because perfectly synchronized clocks between the offline CA and the Snow network cannot be assumed, a scenario where a revocation is removed before the corresponding certificate has expired on all sites must be avoided. For this we ensure that the CA keeps revocations of expired certificates in the CRL for a time equal to or larger than an expected upper bound on clock skew. In practice, this is less than a few minutes.

4. Discussion

Meeting the requirement of the GPs seems to be very hard since they want single sign-on functionality and no extra local administration of users accounts. The authentication and encryption mechanism in Snow comes a long way in achieving this goal by using certificates for authentication and user account creation on the GP servers. By making such a clear separation between the mechanisms for authentication and for authorization, it becomes possible for Snow users to move between GP offices with

little administrative cost. In particular, GP offices do not need to coordinate their local access control lists.

To protect the private key, a user provided key or password is necessary to encrypt it, which place some burden on the GPs for authenticating. However, the solution has potential for single sign-on if the client software can be integrated in the existing EHR system. A benefit of the authentication method is that it does not require work by system administrators as the information necessary for authentication on the client workstations can be added by the users themselves. The solution does require that an encrypted version of the user's private key will be stored on every client where the user certificate has been used, which increases the vulnerability of the user's private key. A more secure solution would be use of smartcards. However, GP offices have been reluctant to adopt such solutions as they incur a higher investment cost and management overhead compared to software only solutions. The important feature of the authentication used is that the GPs have the same identity and certificate across all health institutions and will be able to receive and acknowledge receipt of alert messages from the DPD.

Another beneficial administration issue is that administration of a GP office server will only need its white list updated when the PO it communicates with changes, which probably will be infrequently. The white list on a PO server will need to be changed when a new GP office is added or one of the other PO servers it communicates with changes. Revoking a PO certificate is somewhat problematic since doing so will disconnect all connected GP servers, preventing them from communicating. More alarmingly, an attacker controlling a PO might modify the software so that it does not forward updated CRLs to the GP servers. However, in the case of such an intrusion being discovered, the compromised PO servers should be powered off. Revocations can then be installed manually on the GP servers, or the sub-network might wait in a suspended state until a new PO is installed.

In conclusion, the ability to discover, monitor, and manage outbreaks of dangerous diseases in the patient population is important. However, the computational environment in which such a system is to operate in is heavily constrained by patient confidentiality considerations, firewalls, and the lack of global access control lists.

In this paper we show how the security in the Snow disease surveillance network has been designed so that it can operate within these constraints. In particular, Snow relies on a hierarchical overlay network structure to facilitate XMPP messaging between all system entities. By utilizing an X.509 based public key infrastructure, Snow users can; move between GP offices without requiring global coordination of access control lists, have the same identity and certificate across all health institutions, and be able to receive and acknowledge receipt of alert messages from the DPD.

References

[1] Bellika, J.G., Hasvold, T., Hartvigsen, G. (2007) Propagation of program control: A tool for distributed disease surveillance. *International Journal of Medical Informatics* 76(4):313–329.

[2] Bellika, J.G. et al. (2007) Properties of a federated epidemiology query system. *International Journal of Medical Informatics* 76(9):664–676.

[3] Saint-Andre, P. (2004) Extensible Messaging and Presence Protocol (XMPP): Instant Messaging and Presence. RFC 3921. http://www.ietf.org/rfc/rfc3921.txt.

[4] Dierks, T., Rescorla E. (2008) The Transport Layer Security (TLS) Protocol Version 1.2. RFC 5246. http://tools.ietf.org/rfc/rfc5246.txt.