Initial Experiences of Accessing Patient Confidential Data over the Internet using a Public Key Infrastructure

D. W. Chadwick, S. Harvey, J. New[1], A. J. Young[2],
IS Institute, University of Salford, Salford, M5 4WT, England
[1]Salford Royal Hospitals NHS Trust, Stott Lane, Salford, M6 8HD, England
[2]School of Sciences, University of Salford, Salford, M5 4WT, England

Abstract. A project to enable health care professionals (GPs, practice nurses and diabetes nurse specialists) to access, via the Internet, confidential patient data held on a secondary care (hospital) diabetes information system, has been implemented. We describe the application that we chose to distribute (a diabetes register); the security mechanisms we used to protect the data (a public key infrastructure with strong encryption and digitally signed messages, plus a firewall); the reasons for the implementation decisions we made; the validation testing that we performed and the preliminary results of the pilot implementation.

1. Introduction

The effective management of chronic disease, such as diabetes and cardiovascular disease, are increasingly dependent upon information technology. Traditionally these information systems have been developed within secondary care (hospitals), and have only been accessible from within the home institution. Many dozens (if not hundreds) of these information systems exits at hospitals around the UK. Unfortunately the majority of care is provided by primary care (GP's) and nurses, who do not have easy, real time access to the information recorded in these centralised hospital information systems. This can result in inefficient health care provision e.g. duplication of investigations.

The aim of this research was to develop a methodology to convert standalone, hospital based information systems, into highly secure distributed applications running on the Internet. This would enable geographically dispersed health care professionals to have access to the information held within the (previously centralised) information system.

After some discussions with the clinicians at our local hospital, Hope Hospital in Salford, Greater Manchester, it was decided that their Diabetic Information System (DIS) would be a useful application from which to build and test our methodology.

2. The Centralised Application

Salford is a health care district in Greater Manchester, UK, with a population of 230,510 of whom 5395 are known to have diabetes. The DIS was introduced in January 1992, and holds details about all the known diabetes patients. The DIS is used by all primary care and hospital diabetes services. Records based upon the UK Diabetes dataset [9]

are updated and verified during the annual structured preventative care review. Briefly this contains information regarding their type of diabetes, how it is treated, the presence of any diabetes related complications and biochemical indicators used to assess metabolic control.

The DIS is accessible directly via a built-in user interface, and very recently a SQL based fat client has been added to allow access via the hospital LAN. (The client is fat, as it performs authorisation decisions based on the user's login identity. By comparison a thin client would let the database perform the authorisation function.) All clinicians are given a username/password pair for login to the application, the username determines his privileges for accessing the data (i.e. hospital consultants and general practitioners can only access data relating to their own patients). As the hospital LAN is a trusted network, with no network connections to the outside world, no security is provided for the data whilst it is in transit between the client and server.

Once a year, paper printouts of each patient are produced and posted to all local GPs who then call the patients in for an annual diabetes examination. The completed pro-formas are posted back to the hospital for manual entry into the DIS, a process that can take several weeks (or months) to complete and produces additional problems. Data are double entered, once onto the pro-forma and once into the database, thus giving rise to potential transcription errors. If patients visit their GP before their annual check up, no up-to-date data is available. The pro-formas can, and do, get lost or misplaced, and the whole process is time consuming.

3. Issues in Distributing the Application

To make the application accessible over the Internet, a number of security and usability concerns were addressed. Most importantly we had to ensure the confidentiality of all patient data during transfer and ensure that only appropriate health care professionals could access the data as clinicians have a legal duty of care to ensure that patient details are kept confidential. This could not be compromised whilst the data is in transit over the Internet and so it calls for strong encryption. Secondly, the hospital DIS must be assured that the remote user is genuinely who they say they are, so that unauthorised people are not allowed to access the patient data. This calls for strong authentication of the users. Thirdly the privacy of the hospital LAN must be maintained by keeping out undesirable users and only letting in bona-fide users. This requires a firewall to protect the gateway between the hospital LAN and the Internet. Finally the user interface must be simple to use, quick to learn, and fast in performance, since health care professionals are extremely busy with very limited time to spend on learning new applications or waiting for data to arrive whilst in consultation with patients.

The first two concerns we addressed using a public key infrastructure (PKI), with strong encryption and digital signatures. These would provide both the confidentiality and the strong authentication that a health care application needs. We were already using an Entrust [7] PKI for other research projects and so it was natural to choose this, but if we had not had access to Entrust, there are a number of other commercial PKI vendors and CA service providers to choose from. The hospital chose to use

FireWall-1 from Checkpoint [8] in order to adequately address the third concern about privacy of the hospital LAN.

In order to address the fourth point we had the choice between building a special purpose user friendly interface, or using an existing well known user friendly interface such as a Web browser. A special purpose interface gives the most flexibility in terms of design and capability, but at the cost of significant development effort plus time invested by the user to learn how to use the new application. The World Wide Web on the other hand is ubiquitous, and most (if not all) computer users already know how to use one of the popular browsers. With their simple click and point interfaces we felt that browsers would give us the most chance of success with primary carers, since the latter probably already know how to use browsers, and if not, it should not take much time to teach them. Furthermore minimal development effort would be needed for this, as we could design the main web page to look like the paper form that the clinicians were used to seeing. The architecture is shown in Figure 1.

Given this decision, a number of other decisions flowed from it. Firstly, we had to decide how to convert the http requests once they arrived at the web server, into SQL requests for access to the DIS. Secondly, we had to decide between using SSL[3] protocol and its X.509[2] certificates, or Entrust formatted X.509 certificates and their proprietary protocol. We decided to use CGI scripts to solve the former problem, as they are an already proven way of converting web traffic into application specific queries.

The latter decision was somewhat more difficult to make. SSL has a number of benefits, namely: it is a de-facto standard, all web servers and clients support it, and all PKI vendors and service providers can issue certificates in this format. However, SSL as implemented in 1998, had a number of severe disadvantages:

- we only had immediate access to 40 bit encryption, which was inadequate for confidential patient data. We managed to get plugins to Netscape Communicator for 128 bit encryption (from Fortify [1]), but not for Microsoft Internet Explorer (these would require a US export licence).
- SSL browsers and servers did not support the automatic retrieval of certificate revocation lists (CRLs), and this is an important factor to consider, especially from the server side.
- Trust management has to be performed by the users of the browsers and the administrator of the web server. By this we mean that deciding which root CA public keys to trust and which not to trust has to be performed by the users. As these people are typically not security specialists (and additionally in the case of GPs they simply don't have time to do this) we thought that this placed too much burden on them. It is a role that should be carried out by the security officers of the organisations.
- Certificate renewal after expiry is a manual process, and some browsers will still continue to use expired certificates.

Entrust Direct on the other hand, acts as proxies for both web clients and servers. Http requests from standard web clients are intercepted by the Entrust Direct client proxy running in the user's machine, strongly encrypted and digitally signed using the user's private key, before being sent to the web server. The Entrust Direct server proxy,

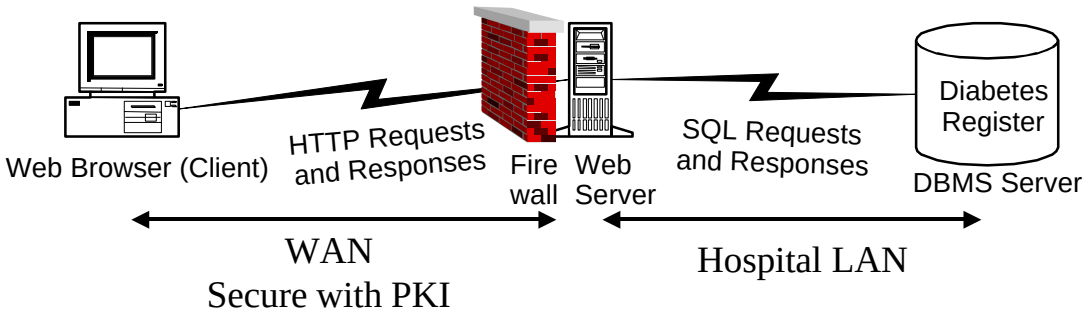

Figure 1. The Chosen Architecture

running in the hospital firewall, intercepts the traffic for the web server, decrypts it, verifies the signature, and, if the user is trusted, forwards the http request to the web server. Both the Entrust Direct client and server automatically retrieve CRLs and process them, they will not accept expired certificates. They also have automatic mechanisms for certificate renewal (without user involvement) and they have all their trust decisions made for them by the security administrator of the PKI. Whilst a Canadian export permit was required for using 128 bit encryption, this was merely a formality as Entrust had already been granted blanket approval for the UK. (Since 1998 of course, the export situation has changed somewhat, in that the US has now altered its approach and has granted blanket approval for 128 bit encryption to be exported to commercial organisations in approximately 45 countries.) Given the above advantages of Entrust Direct, especially with the security naïve set of busy users that we had, it became the natural choice for us.

4. Implementation

In order to operate a PKI, a number of components are needed. The Certification Authority (CA) is the central server that responds to certification requests from the user, and signs the certificates of the users. It also issues the revocation lists. The CA will not issue new certificates or add certificates to revocation lists without being asked to do so by an authorised party. Usually a signed message from a trusted Registration Authority (RA) is sufficient for this. The Registration Authority Agent (RAA) is the administrative client used by a RA to issue signed certification or revocation requests to the CA. In the case of Entrust, after receiving a valid certification request from the RA, the CA returns secret authorisation information to the RA, which the RA gives to the user after authenticating him/her. The user's Client

uses the secret information to establish an authenticated encrypted link with the CA, and then sends a certification request to the CA. The possession of this secret information by the user proves to the CA that the user has been authenticated by the RA, and the CA is therefore willing to issue the user certificate. All of the above are provided by Entrust as part of their basic PKI product line. All communications between the CA server and the other PKI entities are secured, but as a further precaution, we placed a Linux IPFWADM firewall between our CA server and the Internet, in order to block any improper traffic from untrusted third parties.

One additional component needed by a PKI is a directory service in which to publish the certificates and CRLs that have been issued by the CA. LDAP [4] is the Internet standard protocol for accessing directory services, and this protocol is used by the CA server to write to the directory, and by clients to retrieve certificates and CRLs from it. We used the directory server from MessagingDirect [6] to publish our certificates and CRLs.

The final components of our PKI are the Entrust Direct proxies that intercept the http traffic between the clinicians and the DIS. As stated previously, they interact with the PKI to fetch CRLs and check the signatures of incoming messages. They also digitally sign all outgoing messages. In the case of the client proxy it uses the private key of the clinician; in the case of the server proxy it uses the private key allocated to the DIS. The mode of operation is that the clinician starts the Entrust Direct client proxy instead of his usual browser, enters his password to gain access to his private key, and then the proxy starts up the browser. The user then interacts with the Web browser in the normal way, with no visible further interference from the client proxy. At the server end, the administrator of the firewall must enter the password protecting the DIS's private key every time the server proxy is started. Thus every interaction with the DIS is digitally signed and may be audited if required.

Whilst the PKI provides the strong authentication function, our CGI scripts needed to provide the authorisation function. The privileges must be administered in exactly the same way as the SQL fat client that they replace. The SQL client administered privileges by holding a table in the DIS containing the username, password and permissions of each registered user. The SQL client had super-user privileges to the DIS, and when a user logged in, it would compare the password with that stored in the DIS, and if correct, would retrieve the appropriate permissions for the user and act accordingly. We built the same functionality into our CGI scripts, but with an added enhancement. The first time a user accessed the DIS, he had to provide his DIS username and password. The CGI script checked that these were correct, and then stored the LDAP distinguished name of the user (obtained from the digital signature) in a new field in the DIS's table. On subsequent login attempts, the CGI script could simply retrieve the clinician's registered username by looking up the one equivalent to the LDAP DN in the table.

The biggest problem we had to overcome was how to store the DIS super-user name and password securely. As each CGI script needs access to these, we could not expect the firewall administrator to have to type them in each time. Therefore we had to rely on obscurity rather than security, and hide the username and password somewhere where the scripts could find them (details withheld for obvious reasons!). A future enhancement is to build a Session Manager that is a permanently running service that

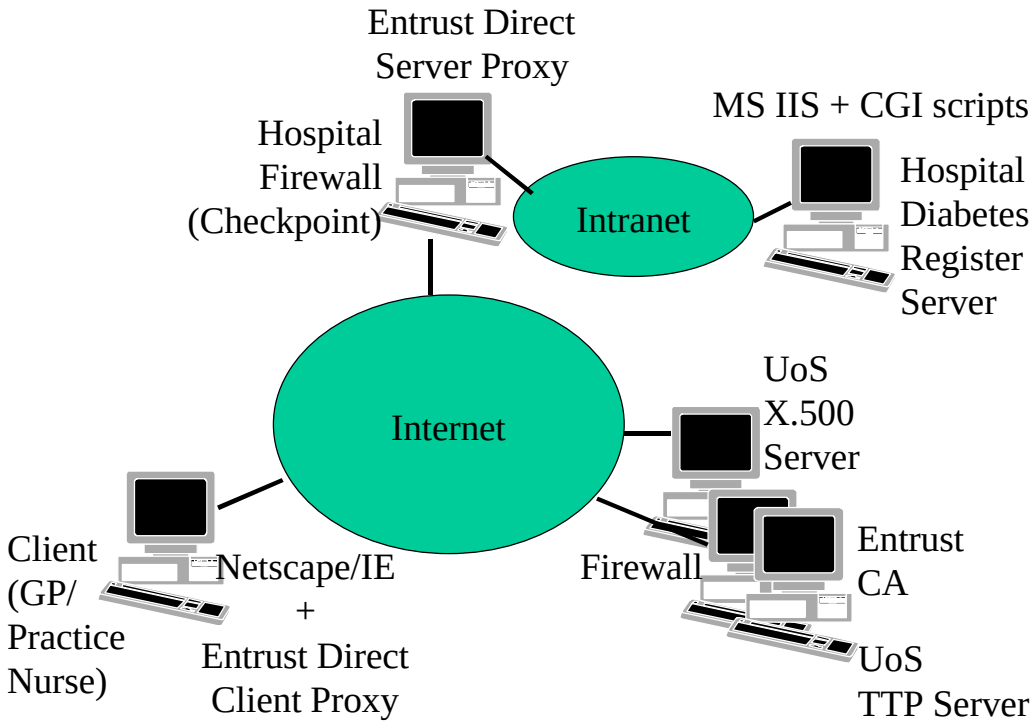

Figure 2 System Components

sits between the CGI scripts and the DIS. This will have to have the super-user name and password entered at start up time by the firewall administrator.

The complete system with all its components is shown in Figure 2.

5. Validation Testing

Once the system was built, members of the project team and closely associated colleagues in the university and hospital (7 testers in all) performed a series of validation tests on the system. These were carried out prior to the implementation with the GP pilot users, as we wanted to be sure that the system was user friendly, reliable, and fast enough in performance before we gave it to busy GPs in their working environment. Each database access test lasted between twenty and thirty minutes, to simulate what a clinician in the field would probably do under normal clinical situations. The following areas were examined under timed conditions:

- Logging onto the Application securely
- Searching for a patient's record and retrieving the details (4 times per testing session)
- Updating one patient's data with new information and retrieving it
- Closing the Application down and logging off
- Attempting to log on in an insecure manner (for performance comparison purposes)

In addition we logged the availability of the underlying hardware, software and networking infrastructure throughout the testing period. We measured the network availability by issuing Pings at timed intervals to various stations on the university

LAN, the university's Internet gateway, University College London (a well known reliable site on the Internet), Demon Internet (to test LINX connectivity for dialup users), the hospital gateway, and the hospital's LAN. We measured data integrity by asking the testers to Email a copy of their retrieved record to a researcher, who compared this bit-wise with a copy he had previously downloaded from the diabetic register. Finally we measured the time it took to install each user with the Entrust PKI. We revoked one user to see if this worked satisfactorily, and we reinstalled another user who pretended to forget his password to his private key.

5.1. Test Results

We decided that 15 minutes was a reasonable time in which to install a new user with the client components of the Entrust PKI system. Out of the 8 users who volunteered to perform the tests, 1 was already installed prior to the testing, but we only managed to install 3 of the remaining 7 users within the allotted 15 minutes. 3 more were successfully installed within 24 hours of starting the installation, but we could not install the final user due to her PC's limited capabilities and the interactions with her Internet Service Provider (which limited outgoing TCP/IP connections). Clearly client installation is not as straightforward as it should be, even given that we were working in a very heterogeneous environment of PCs (different memories, CPUs, manufacturers, operating systems, configurations etc.).

The 7 installed users completed 32 series of tests. This was less than we had planned, but never the less we felt it was sufficient to perform a meaningful analysis of the results.

The system proved to be very reliable throughout the testing period, with most components scoring 100% availability. Only the Entrust Direct server was temporarily unreliable, at 66% availability (it crashed at 4pm every day). When we changed the frequency of CRL publication from 4 hours to 24 hours, the problem disappeared. Network availability, at 93%, was also poor, but unfortunately this was out of our control. The poor performance was almost entirely accounted for by failures in the University of Salford LAN, and not in the Internet.

We were particularly encouraged by the time taken to learn to use the system (less than 8 minutes) plus the performance once the application was installed and the connection made to the database (between 2 and 32 seconds to search for and retrieve a patient record). However, the time taken to launch the application and make the connection to the database (up to 2 minutes) was longer than we would have hoped for. Certainly we felt that this time would be prohibitive for a GP during a patient consultation.

Data integrity measured 99% and not 100% as expected. However, the 1% failure was found to be due to a documented bug in Windows NT. Of the 72 sets of results received from the testers, all were identical except for one where the dates were in American format rather than European format. A workaround has since been applied to the application.

User revocation and user re-installation both worked correctly with no hickups.

Our overall conclusions from the validation testing were that the application was easy to use, performed well after launching, but that application installation is problematical and application launching is too time consuming.

6. Piloting with the Users

At the time of writing the piloting has not completed, although installation has completed and the piloting is over half way through. The results of the installation saw a repeat of the difficulties that we found during the validation testing, only this time the environments were even more heterogeneous, with a combination of different PCs, some connected by LANs and some stand alone, using a variety of different ISPs. Consequently we experienced an even greater number of unforeseen technical problems during installation.

The original plan was to pilot the application with 12 doctors and practice nurses located in 4 different surgeries. Surgery A is split over 2 physical locations and has 2 GPs and a practice nurse that work in both locations. Surgery B has 4 GPs and 2 practice nurses (but only 1 of the latter agreed to participate in the pilot). Surgery C has 1 GP and 1 practice nurse, and surgery D has a diabetic specialist in a tertiary care unit.

Surgery A had no suitable PC at either location, and so the project lent them a suitably configured PC and modem. This was placed in one of their locations on a worktop away from the GPs desks, and so was not readily accessible during consultations. Installation proved difficult due to the fact that their ISP, one of the numerous free ones in the UK, demanded to know the calling telephone number before giving them complete Internet access (i.e. access to all protocols and port numbers. We were using non-standard ports for our LDAP service, and Entrust uses non-standard protocols and ports). As most GPs are ex-directory, they usually withhold their telephone number from outgoing calls. We had to try 10 different free ISPs before we found one that did not demand to know the calling number during login. Halfway through the trial our logging indicated that this surgery had not used the system once. When queried about this they admitted that they needed the PC to be on a mobile trolley rather than on a worktop, so that it could be moved to where the GPs were working. They would also like to have a printer available to them so that they could print out the patient's details. These were provided for the second half of the trial.

Surgery B has a LAN installed, with PCs on every GPs desk and a central server which automatically dials the Internet when needed. Though this dialup can take a minute to connect, it means that the Internet is always there when needed and no explicit action is needed on the part of the GPs. However, despite making several visits to the surgery, after one month we had only installed the software with 2 GPs and the practice nurse. The problems were all of a technical nature, but were compounded by the limited time available to the GPs. Therefore if the installation did not proceed as planned within the allotted time we usually had to book another appointment. Problems arose from the LAN configuration, and this made installation difficult. One of the GPs had volunteered to use a smart card to hold his private key, but after an hour of trying to install the card and reader he decided to use software based keys instead. (A report of some of the difficulties we experienced with smart

cards is given in [5].) However, after installation our log files showed that they were using the system.

In surgery C the GP uses the Internet frequently, and we were specifically asked not to alter any of his Internet settings. However, the GP was using an ISP provided dialer, rather than Windows Dial-Up Networking, and this forced the web browser to use that ISPs default proxy settings. These settings stopped Entrust/Direct from working. Fortunately the GP did have another ISP already installed, and so we had to instruct him to use that ISP, but even this caused him some disruption. The GP also agreed to pilot the use of a smart card, and this increased the installation time to 1.5 hours. However he had problems subsequently when using the smart card, and eventually reverted to using software based keys.

Surgery D had no suitable PC, and so the project lent the specialist a laptop and a modem and set her up with an ISP. She was only able to dial the ISP from one of her 3 consulting rooms as (a) the telephone connection in one room barred external calls starting with a zero and (b) the telephone connection in another room had a broken dialling tone which the modem did not recognise (we had to dial the “9” for an outside line manually to get a normal dialling tone!). However, once installed, this user uses the application very frequently from the third consulting room (several times a day in fact).

7. Conclusions

We have shown that it is possible to provide highly secure remote access to a hospital information system via the Internet, using commercial products and tailor made CGI scripts. By using a web browser interface the system is extremely easy to learn to use by both doctors, specialists and practice nurses. The response time is also adequate. However the time taken to the launch the application is an inhibiting factor, especially when using dial-up access via an ISP. The largest problem we faced was installing the users as members of the public key infrastructure. Installation was plagued with unforeseen technical difficulties, and the smart cards proved so time consuming to install and problematical to use that all our users stopped using them during the trials.

8. Acknowledgements

This research was funded by: the European Commission IV Framework Programme Trusthealth 2 (Contract No: HC 4023) and ICE-CAR (Contract No: RE 4006) projects and the UK EPSRC under grant number GR/L60548. The authors would also like to thank Entrust Technologies for making their security software available to the University on preferential terms.

References

[1] see www.fortify.net
[2] ITU-T Recommendation X.509: "The Directory - Authentication Framework". 1997.
[3] Frier, A., Karlton, P., Kocher, P. "The SSL 3.0 Protocol", Netscape Communications Corp., Nov 18, 1996.
[4] Yeong, W., Howes, T., Kille, S. “Lightweight Directory Access Protocol”. RFC1777, March 1995

[5] Chadwick, D.W. “Smart Cards aren't always the Smart Choice”, IEEE Computer, December 1999
[6] see www.messagingdirect.com
[7] see www.entrust.com
[8] see www.checkpoint.com
[9] Vaughan NJ, Home PD. The UK Diabetes Dataset: a standard for information exchange. Diabetes Audit Working Group of the Research Unit of the Royal College of Physicians. British Diabetic Association. *Diabet Med* 1995; 12: 717-22.