Secure printing of confidential documents for nomadic users

Luis Bengochea, Ramiro Cano

Computing Sciences Department
Universidad de Alcalá
Campus Universitario- 28871 – Alcala de Henares, Spain
luis.bengochea@uah.es - ramiro.cano@alu.uah.es

Abstract. In an economic environment more global each day, a high percent of the employees in a company spend the most part of their work time outside their offices or even outside their countries. Advances in portable devices design and arrival of communication networks capable to transmitting such data allow these nomadic employees to fulfil their tasks with the only help of a mobile phone or a PDA. However, these devices show lacks when accessing large size files, which sometimes can be needed when it is to present a report or formalize a contract at a client's head office. This problem worsens if the document contains confidential data and it cannot be hosted in an accessible server in the company nor can be sent by fax without putting its security at risk. In this paper we present a system which allows nomadic employees, using a light portable device and a secure communication channel with their company, after a search of the required document and its finding through short metadata, appropriate for being shown in the screen of this kind of devices, such as title, date and document relevant phrases, being able to request remote printing and ciphered reception with one-time key, through an open channel and for a wide range of multifunction printers supporting J2ME, avoiding the risk for the above-mentioned document to get lost.

1 Introduction

At the information era, where the concept of society is being replaced by the global village and public opinion is forged in the “blogsphere”, concepts such as ubiquitous computing become more important each day in the field of information technologies. Companies regarding still being competitive within this new scenario that grows up will need constant adaptation in order to avoid getting obsolete through the way.

Electronic components miniaturization, more advanced each day, allows us to have nowadays portable devices, capable of doing any kind of tasks, with lesser size than the taken up by a wallet in our pocket.

This evolution has morphed cassette players into digital format players with the size of a bank credit card, and weighing just a few grams; the first mobile phones into small multimedia devices capable of doing countless functions moreover those that were initially designed for; the first portable computers into laptops weighing about a kilogram, or into personal digital assistant (also known by its acronym PDA), and we could go on with many other examples.

This advance in mobile device designing, together with the simultaneous appearance of communication networks suitable for transmitting this kind of data, they allow the arising in the companies of a new type of employee -travelling or nomadic- who is able to fulfil all his tasks with the only help of a device that fits on his hand and weighs little more than a thousand grams, and who has no need of turning up in his company's head office [1].

Same as devices advance through miniaturization and functionality also does the complexity regarding the tasks they are bound for and, as a consequence, it is exponentially growing the quantity, quality and importance regarding the data being transmitted and processed by this kind of devices. Therefore, the security concerning that information is another aspect that gains special importance in this new scenario, because even when techniques and security measures have had an amazing development in the last years, also have done the malicious interest about this precious data. Fighting against this kind of criminals -crackers, phisers, spammers...- demands from computing security experts -researchers, hackers, security consultants, developers...- a constant search in new information protection mechanisms and a permanent surveillance concerning new kind of crimes related to this data.

2 Nomadic users

In this article it is presented a working scenario about a company that develops its activity in a scattered geographic field, including activities outside its origin country, and in which part of its staff it is constituted by travelling or nomadic users who spend long time without turning up in any of the company's offices, but whose needs in privacy and confidentiality with sensitive business data, are equal to the remaining employees who work inside the building with physical and logical accessing security measures implanted.

This kind of employee who fulfils his function as nomadic user carries out constant journeys for dealing with several clients of the company he works for, and often, he needs remote access to a wide variety of internal documents -rules, proceedings, manuals, reports, contracts, etc. - that are inside the company's corporate network, through a portable device -PDA, mobile phone or Blackberry- and from anywhere around the world [2].

An average size standard corporate network belonging to a standard company which have this type of employees, must be generally constituted by the next components:

- An internal network which offers service to the offices and studies in the company, divided in some subnetworks with their own Internet gateways.

- A demilitarized zone (DMZ) in which the services that must be accessible from outside the network are hosted.
- The appropriate network security elements, mainly firewall -node and network firewalls- and reactive and passive intrusion detection systems (IDS).

If there is a connection between the local area network and the DMZ -which it is not necessary-, it must have extreme security measures.

The DMZ offers various services which can be classified into two types: external and internal ones. First ones are public services that can be accessed directly by anyone connected to the Internet, such as the web server which hosts the public company webpage or a FTP file server; while the second ones can only be accessed from the DMZ itself, and are specifically aimed for being used by the company nomadic employees through a tunnelized secure virtual private network (VPN) implemented with IPsec [3].

Even when VPN was designed with static users in mind, it has not been necessary to make any significant changes for it in order to being able to be used in this context. [4], characterized by the neediness in ubiquity and flexibility.

When the nomadic employee needs any type of confidential document, he will establish a remote connection with the demilitarized zone's private network in the corporate network, using the aforementioned secure tunnelled VPN, which will allow him to access the internal resources of the company and make any kind of search or query over those documents he has access depending on his role in the company.

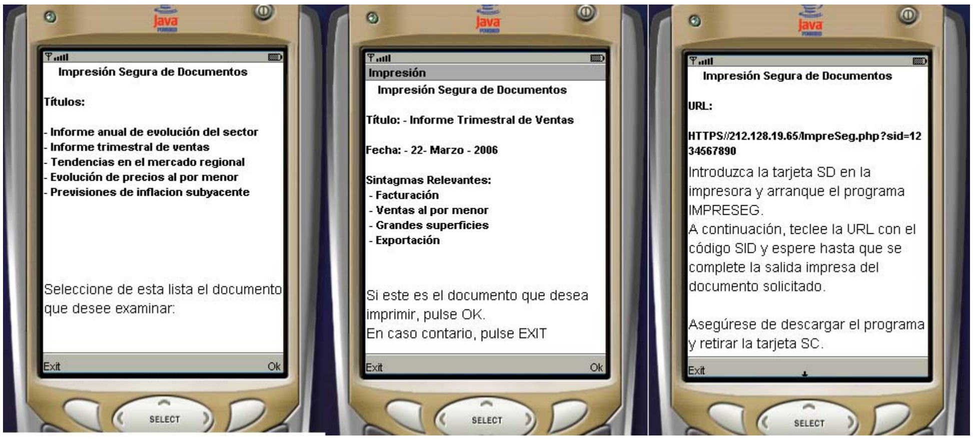
3 Applications on ubiquitous computing environments

In a company with nomadic employees it is arisen the problem about an application which has to be able to be executed using very different devices and communication means, depending on the location of the employees in each moment. Therefore, providing the ability of mobility to these applications is an important aspect in ubiquitous computing. This capacity of migrating applications in order to adapt them to different environments and contexts, in some cases with very broad differences in the available resources -such as the size of the screen, the use or not of colours, or the bandwidth available,- it constitutes even today a very important challenge, as it is stated in [5] where the authors introduce the concept of polymorphic application, in which it is preserved the functionality while its structure is changed in order to adapt it to different environments. This way, an employee can carry out the same tasks either if he is in his office's desk, or if he is at an airport's waiting room.

3.1 Accessing to documents

The content management platform used in a company include a group of processes which cover, from obtaining and classifying the contents up, to the way in which they are shown when offered to the final user. In this last point is where gets special importance the typology of portable devices [6] which the nomadic users are accessing from, characterized by their small size screens which make that design and appearance of the information both become an important part inside the services' supply chain of the content manager.

Several methods have been proposed [7] in order to being able to show a document's content within a mobile device with a small size screen, such as cutting the whole text into smaller units and show these either complete or resumed. In order to summarize a text, different techniques can be used, such as listing the more meaningful words set or extracting the most outstanding syntagms from him. In this same work [7] it is shown that resumes based in a combination of key words and important syntagms provide an excellent approximation for identifying the document which was been looked for.



Impresión Segura de Documentos

Títulos:

- Informe anual de evolución del sector
- Informe trimestral de ventas
- Tendencias en el mercado regional
- Evolución de precios al por menor
- Previsiones de inflacion subyacente

Seleccione de esta lista el documento que desee examinar:

Impresión

Impresión Segura de Documentos

Título: - Informe Trimestral de Ventas

Fecha: - 22- Marzo - 2006

Sintagmas Relevantes:
- Facturación
- Ventas al por menor
- Grandes superficies
- Exportación

Si este es el documento que desea imprimir, pulse OK.
En caso contario, pulse EXIT

Impresión Segura de Documentos

URL:

HTTPS//212.128.19.65/ImpreSeg.php?sid=1234567890

Introduzca la tarjeta SD en la impresora y arranque el programa IMPRESEG.
A continuación, teclee la URL con el código SID y espere hasta que se complete la salida impresa del documento solicitado.

Asegúrese de descargar el programa y retirar la tarjeta SC.

Fig. 1. Example on a dialog for a document secure printing. In the search phase, the titles of candidate documents are shown. In the identification phase, it is shown the metadata associated with a document and, in case of a positive identification, the server provides an URL with a session identifier which allows to request the document printing from a multifunction printer.

Also, the set of key terms and important syntagms associated to each document can be previously saved as metadata associated to itself [8], which could allow the applications that access the documents to use them directly for being shown during the search dialogue in a specific document.

A decisive factor for the effectiveness in the univocal identification of a document using outstanding syntagms extracted from the text is their quality. In our project we have used KEA algorithm adapted for Spanish language, which uses a Naïve-Bayes classifier built from a set of learning documents and provides appropriate results [9] for our purpose. In the fig. 1 it is shown how to reach the document identifier for the document desired for printing and the SID key provided by the server.

3.2 Document printing preparation

Once the document for printing is identified, a petition is sent to the server, which will unleash the generation of a one-time session identifier (SID) which will be associated to the document and will have an expiry associated (TTL) in the database. The system for generating that session identifier lies in the using of a pseudorandom number generator (PRNG) [10] y [11] widely used in cryptographic and secure data erasing systems in computing applications.

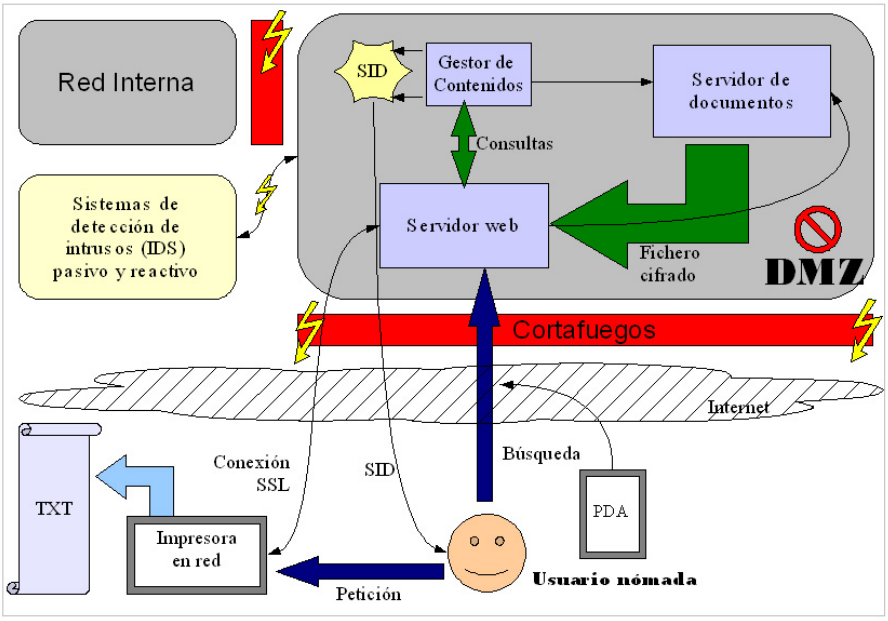


Fig. 2. Diagram of the document secure printing system.

At the same time, the document server will get ready a copy of the document, in PDF ciphered format. That copy is sent to the web server in the company's demilitarized zone (DMZ) so that it can be accessed from the remote printer (see fig 2).

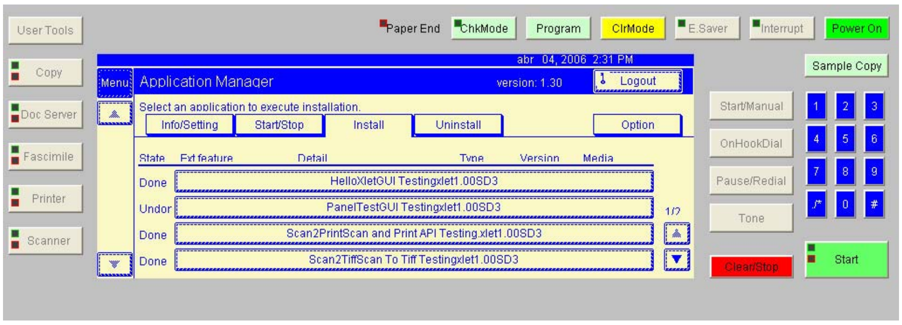
4 Secure printing

The use of multifunction printers -with the ability of document scanning, fax sending and receiving, etc.- with a standard TCP/IP connection, such as Internet, is something usual nowadays in any company. In many cases, they have at their disposal a Java virtual machine environment capable of executing applications developed with Java 2 Micro Edition (J2ME).

In our case, we suppose that the nomadic user have, in the place he moved to, one of these devices and therefore he will be able to install into it a small program previously developed which provides him means for establishing SSL connections with a remote web server, launching petitions into it, and negotiating both the possible responses and file receptions.

The program can be installed in the printer either from a portable memory device which the employee carries (for example, a CS, SD or other similar type of memory cards), or being downloaded and installed straight from the Internet if the printer implements such function.

After installing and executing the program (fig 3), the employee will have to enter the document identification URL provided during the search phase, so that it unleashes the petition to the company remote web server, where it will be yet ready the ciphered copy of the document.



Select an application to execute installation.

State Ext feature Detail Type Version Media

Done HelloXletGUI Testingxlet1.00SD3

Undor PanelTestGUI Testingxlet1.00SD3

Done Scan2PrintScan and Print API Testing.xlet1.00SD3

Done Scan2TiffScan To Tiff Testingxlet1.00SD3

Fig. 3. Control panel of a multifunction printer with the installation of java developed programs ability.

After receiving the petition, the server will check the provided identifier validity -both its existence and relevance- and the existence of the ciphered document. If an error with the document petition identifier happens, which could be considered as a malicious access attempt on a document, it must be planned the option of launching an alert into the passive IDS for it to be logged, or even into the reactive IDS for it to fulfil any defensive action against the attack, such as blocking in the firewall the access over the IP that made the petition for a time greater than the document session identifier expiration one.

If no error happens during the process, the server labels the session identifier as used -which is not same as expired- and logs all data concerning the document transfer: employee name, IP from where the petition was made, public key of the remote node (printer), date and time of the transference.

Finally, the PDF file is sent from the web server in a completely normal way, since SSL has been inserted in a transparent way in the communication between printer and server. The printer operated by the nomadic employee will so receive the document, close the communication channel and print it.

The program in charge of the confidential document petition will proceed, after its printing, with the secure erasing of received file in its temporal storage, which will preferably be the memory card, in order to avoid any risk, even minimum, in subsequent printing. The secure erasing is carried out by data rewriting techniques with patterns and pseudorandom data using the Guttmann method [12].

5 Security in communications for an ubiquitous office

Often, documents printed in strange offices by nomadic employees, confidential or strategic by nature, can compromise the business, which justifies the establishing of strong security measures which avoid those documents to get lost or theft and, eventually, end up in someone not authorized's hands.

It is possible that in any point within the connection between the printer and the web server hosted in the corporate network's DMZ exists a potential attacker with the purpose of carrying on some sort of attack or spy act over the connection in order to obtain the document it is going to be transferred. In a standard situation, with a HTTP or FTP transference, with the only help of a simple network sniffer and through simple techniques as ARP cache poisoning, a person inside the appropriate subnetwork in the company the nomadic employee is displaced, could have enough means to obtain a copy of such document.

Even in this situation and inside the DMZ, which hosts the valuable internal resources that would be only accessible through physical or VPN connection with the internal network, the systems providing such connectivity -usually network switches- can offer additional protection means in order to detect or even react over this kind of spy techniques and threats against the network security, being able to act through varied countermeasures from sending administrative alerts into the network administrators working places, short text messages SMS into their mobile phones, e-mails... up to deactivating the communication port which link the attacker computer cable with the whole rest of the network, leaving him outside it and frustrating at the same time any attempt in granting continuity to the attack. Even in case the confidence in the physic network security over used devices in lower communication model layers is very strong, it is convenient to activate the alerts they implement. If it is awful being theft, much worse is not being aware of it.

Moreover the physic intrinsic security of the used network devices, the program installed into the printer must assure that transmitted data are wrapped into some kind of “safe boxes” which in case of being opened without the appropriate authorization offer nonsense data. In order to being able to be opened, they require the establishing of a secure link between both emitter and receiver of the transmission, so that the communication gets strengthened further than the physic layer.

This is, in short, which both SSL and public/private key asymmetric cryptographic system provide us. In a hand, the connection must be accepted as valid and, in the other, transmitted data will not be visible by the only fact of being obtained. In the SSL connection, it is not transmitted any data until the transfer channel is considered secure. So, the connection go through different phases before being considered: in “Hello” phase, where the multifunction printer sends to the web server its public cypher key -which resides in the loaded program-, and its cryptographic preferences. In the “key exchange” phase, the server sends to the printer its public key once the printer's one is processed and accepted. In the “session key generation” phase, the server generates a symmetric cypher key for transmitting ciphered data during all the session and at last the “fin” phase, where once key exchange is finished, it does begin the ciphered communication channel. From this precise moment, the connection is considered as secure.

Algorithms that SSL use for implementing its ciphering system have proved along time being strong and effective for this task. For asymmetric cypher they are implemented RSA, DH-DSS and Fortezza algorithms, although RSA [14] is the most used. For symmetric ciphering they are used AES, IDEA, RC2, RC4, DES and 3DES algorithms, being AES the most frequent and strong.

For the one-way hash function system [15] the chosen algorithms are SHA-1 and MD5. Both are widely used, although in security field -including SSL- the first one is the most used. Even when recently it has been published a cryptanalytic attack against SHA-1 system which makes possible to reduce a brute force attack complexity under the expected using a standard birthday attack, it is still being considered as secure nowadays.

6 Conclusions

Using a system such as the one described in this work, a nomadic employee who fulfils part of his job moved into other companies' offices and needs to obtain a printed copy of any confidential document relating his company, can make use of a portable device and a VPN connection in order to locate the document and request a temporal key for accessing through the Internet a ciphered copy of it. In any multifunction printer, with J2ME support, available at the office where is located, he can install a small program which function is request, print and erase the document from the printer's memory in a way that grants that such document cannot end in the hands of any person without the appropriate authorization, neither accidentally or spitefully.

References

1. Kleinrock, L.; "Nomadic computing and smart spaces". Internet Computing, IEEE. Volume 4, Issue 1, Jan.-Feb. 2000 Page(s):52 – 53
2. Kleinrock, L.; "On some principles of nomadic computing and multi-access communications". Communications Magazine, IEEE. Volume 38, Issue 7, July 2000 Page(s):46 - 50
3. Cisco Systems. “Deploying IPsec Virtual Private Networks”. Solutions Guide. 2002. http://www.cisco.com/warp/public/cc/pd/iosw/prodlit/depip_wp.pdf.
4. Kamouskos,S.; "Supporting nomadic users within virtual private networks". Service Portability and Virtual Customer Environments, 2000 IEEE. 1 Dec. 2000 Page(s):128 - 133.
5. Ranganathan,A.;Chetan,S.;Campbell,R.; "Mobile polymorphic applications in ubiquitous computing environments". International Conference on Mobile and Ubiquitous Systems: Networking and Services, 2004. MOBIQUITOUS 2004. 22-26 Aug. 2004 Page(s):402-411.
6. Emilio Javier Torres Mateos, Antonio Orgaz Climent, Héctor Hernanz Barrero. “Gestión de contenidos para portales móviles UMTS”. Comunicaciones de Telefónica I+D. Número 24. Enero 2002

7. Buyukkokten, Orkut; Garcia-Molina, Hector; Paepcke, Andreas. "Seeing the Whole in Parts: Text Summarization for Web Browsing on Handheld Devices". 10th International WWW Conference. 11 pags. February 2001.

8. Bengochea, Luis. "Asignación automática de metadatos a colecciones de documentos empresariales en bibliotecas digitales". Conferencia IADIS Ibero-Americana WWW/Internet 2005. Lisboa 18-19 Oct.2005. ISBN: 972-8924-03-8

9. S. Jones, M. Jones, S. Deo, "Using Keyphrases as Search Result Surrogates on Small Screen Devices", Personal and Ubiquitous Computing, vol. 8, no. 1, pp. 55-68, 2004.

10. Peter Hellekalek. “A Concise Introduction to Random Number Generation”. Lecture Talk at the Santa's Crypto conference, Prague, December 7, 2004.
http://random.mat.sbg.ac.at/~peter/slides_Prag2004_printversion.pdf

11. Mohan Atreya. “Pseudo Random Number Generators (PRNGs)”. RSA Security. White paper. http://www.rsasecurity.com/products/bsafe/overview/Article4-PRNG.pdf

12. Peter Gutmann. "Secure Deletion of Data from Magnetic and Solid-State Memory". University of Auckland. Sixth USENIX Security Symposium Proceedings, San Jose, California, July 22-25, 1996.

13. R.L. Rivest, A. Shamir, and L.M. Adleman. “A method for obtaining digital signatures and public-key cryptosystems”. Communications of the ACM, 21(2)” - Febrero de 1978.

14. RSA Laboratories. “Public-Key Cryptography Standards (PKCS) #1: RSA Cryptography Specifications Version 2.1” - ftp://ftp.rfc-editor.org/in-notes/rfc3447.txt

15. Preneel, B. “Analysis and Design of Cryptographic Hash Functions” - Ph.D. Thesis, Katholieke University Leuven, 1993.