Ensuring the security and privacy of information in mobile health-care communication systems

Authors:
Ademola A. Adesina1
Kehinde K. Agbele1
Ronald Februarie1
Ademola P. Abidoye1
Henry O. Nyongesa1

Affiliation:
1Department of Computer Science, University of the Western Cape, Cape Town, South Africa

Introduction
Before the application of information and communication technology (ICT) in health-care delivery systems, some of the problems faced were the incorrect recording of diagnoses, unavailability of patient information, delays in accessing the information, space limitations for record-keeping and insufficient personnel for patient monitoring. The paradigm shift in health information technology has enabled a reduction in these hurdles and a more personalised service to be delivered. Through the acceptance of the Internet as a tool for health-care providers, medical organisations are establishing websites. In addition to being reservoirs of descriptive information about the facilities and services of the organisations, these websites allow patients global access to their medical information, such as clinical laboratory reports, appointment information, health and prevention reports, billing information and other components of their patient record, via the Internet.1,2

Acceptance of the Internet as a tool by health-care providers has not only enabled a transformation from paper-based records to electronic patient records (EPRs), but has also facilitated the use of sensor networks for remote patient monitoring, which allows for easy accessibility of medical information by health-care practitioners. For example, Intel’s Integrated Digital Hospital combines mobile point-of-care and other information technology concepts to integrate patient and administrative information into a comprehensive digital view of a patient’s health.3 The corollary for global access is that electronic use (from medical terminologies to networking protocols) must be standardised. Another consideration when using this technology to enhance health-care delivery is the need for security and privacy, so as to maintain fundamental medical ethics and social expectations. Such considerations include data access rights; where, when and how data are stored; security during transmission; data analysis rights; and governing policies.

In this paper, we examine various ways of implementing data security measures in a mobile health-care environment when data are being transmitted and where they are stored in the database repository. Data encryption, digital watermarking and steganography are various ways to protect the integrity of the data (which may be in the form of text, image, video or audio) in noisy communication channels during the transmission of patient data. Security for the database server and central monitoring system (in the case of sensor networks or telemedicine) is necessary to protect the integrity of the stored data in a mobile health-care communication system. We also present the theoretical background to the issues of privacy and data protection and discuss dynamism in health-care delivery systems, the storage of patient records and the transformation from eHealth (electronic health) to mHealth (mobile health). Some real-life scenarios regarding the privacy and security of patient records are given and recommendations for the improvement of database security and privacy are discussed before we offer our conclusions.
Theoretical background to the issues of privacy and data protection

The United Nations guidelines encourage countries to enact legislation that will accord personal information an appropriate measure of protection and also to ensure that such information is collected only for appropriate purposes and by appropriate means. In 1995, the Data Protection Directive was enacted to provide some level of protection for citizens during the free flow of personal data within the European Union. The directive stated that the flow of personal data can be only within the boundaries of the member countries that can guarantee ‘an adequate level of protection’.4

The Southern African Law Reform Commission recognises that privacy is a valuable aspect of personality. Its protection forms an element of safeguarding a person’s right to privacy and providing legal protection on personal information that is collected, stored, used or communicated by another person or institution. The meaning of information protection varies in different declarations and laws; basically it means that personal information should be dealt with according to a specific principle known as the ‘Principle of Information Protection’.5,6

The promulgation of information protection laws in South Africa has resulted in the amendment of South African legislation, most significantly the Promotion of Access to Information Act 2 of 2000, the Electronic Communication and Translations Act 25 of 2002 and the National Credit Bill of 2005.4,5 The preliminary recommendations of the Commission, as set out in the Bill, can be summarised as follows4:

1. Privacy and information protection should be regulated by a general information protection statute, with or without sector specific statutes, which will be supplemented by codes of conduct for the various sectors and will be applicable to both the public and private sector.
2. General principles of information protection should be developed and incorporated in the legislation. The Bill gives effect to eight core information protection principles: processing limitation, purpose specification, further processing limitation, information quality, openness, security safeguards, individual participation and accountability.
3. A statutory regulatory agency should be established. Provision has been made for an Independent Information Protection Commission to direct the work of the Commission in implementation of both the Protection of Personal Information Act and the Promotion of Access to Information Act of 2000. Data subjects will be under an obligation to notify the Commission of any processing of personal information before they undertake such processing.

Dynamism in health-care delivery systems

Health-care providers have explored information technology opportunities to reduce the overall costs of health-care delivery without compromising the quality of health-care service.5,8 This dynamism has brought about scenarios in which health care is decentralised and distributed, and responsibility is shared among different health-care providers to render optimal medical, psychological and social help to patients.7 These scenarios make it important for a practitioner to be able to search the medical records of a patient and establish the history of their ailment, as well as previous diagnoses and treatment in order to provide current treatment.

Monitoring the health of a patient in a remote area is achievable through the use of a mobile device (e.g. a cellular phone), local server and remote patient monitoring system. A periodic report from the sensors is sent to the system server by means of wireless communication, such as Bluetooth within the patient’s house, which is connected to the central monitoring station. A final response (the appropriate treatment or interpretation of the sensor signals) is received from the central monitoring station if the signals are beyond what the local server can interpret. Internet connection facilitates the link between the two ends – the patient’s environment and the central monitoring station.

With the sophistication of health-care facilities improving, the likelihood is that fewer personnel will cater to more patients, whilst still delivering prompt services and efficient patient care by making judicious use of limited time, which thus allows for more time for clinical activities. Building an ‘ecosystem’ that relates different aspects of the hospital management system could bring about precise results, minimise costs and improve the efficient management of a facility.

Figure 1 represents the proposed dynamism in health-care delivery system architecture as a block diagram of a generic form of a health-care delivery information system in which the patient is the centre point that all departments or units concentrate their services on. It is a form of ecosystem that gives priority to the patient over all other facets of the hospital management system. There is a need for information flow in this particular system, for example, the laboratory information system should make available laboratory results to the surgery unit before any surgical procedure can commence, the pharmacy dispensing a drug is dependent on the recommendation of the orthopaedic department, and radiology will present the report of their findings to the orthopaedic unit. In this system, the various forms of information have to be made available for the proper treatment of the patient.

The health-care delivery information system is an interwoven relationship. It is clear from Figure 1 that interaction between the patient and different departments is inevitable. There is a need to ensure the security of the transmission between patients and departments within the domain because vast amounts of data are constantly being generated electronically.

Other benefits of hospital-care delivery systems include:

- enabling hospitals and skilled professionals to render better services to their patients
- improving the quality of patient care in all areas of the hospital system from the laboratory to the pharmacy, and even bed management systems
increasing professionalism, such as physician and nurse productivity
• reducing the time spent by staff in filling out forms, thus freeing resources for more critical tasks
• improving the quality of care, procedures and service to patients
• controlling the costs incurred by diagnosis-related groups.

Patient records

A patient record may be defined as ‘any relevant record made by a health-care practitioner at the time of or subsequent to a consultation and or examination or the application of health management’10. A patient record contains information about the health of an identifiable individual recorded by health-care professionals, either personally or at their direction.11
The patient record documents the trend of medical activities over a particular period of time, including the treatments prescribed for an ailment. An electronic medical record is the record of the medical information of a patient for a specific enterprise, such as a hospital, whereas the EPR contains all the health-care-related information on one person, that is, the integration of the patient’s health information from diverse and disparate systems, as is practised in a distributed environment.12

Patient records can be kept in paper or electronic form. Paper-based records require significantly more storage space than digital records. Patient records should be kept for a certain number of years and such retention incurs a storage cost. Paper-based records also require collation, especially when parts are stored in different locations, whereas electronic records do not. Another problem associated with paper-based records is that of poor legibility, which may result in serious medical error. The interpretation of standard medical jargon and the standardisation of abbreviations are unreliable in paper-based records, whereas these issues are automatically addressed in electronic records because of the standardisation of forms, terminologies and abbreviations used for the input of data electronically.

EPRs take the current paper-based documents and convert them to a digital format so that they are available in an electronic form. When an EPR is initiated, information is gathered from a patient’s record at a specific location and the information is then shared via the Internet with the authorised health-care practitioners who have the right to access the database. The records include various types of data, such as physician’s notes, magnetic resonance images and clinical laboratory results. Using EPRs allows real-time access to health-care records, irrespective of the physical location of the user. Physicians, nurses, insurance companies and patients can access the records via the Internet. In addition, EPRs can be more easily backed up than paper-based records, which prevents the possible loss of data.1 Accessing an EPR is easy because it is stored in the database, which confines it within a particular location. Patient information is exchanged across the server via the Internet or other interfaces designed for presenting the records. EPRs can also be continuously updated, irrespective of the location of the health-care practitioner. The ability to exchange records between different EPR systems would facilitate the coordination of health-care delivery in non-affiliated health-care facilities. Another advantage of the EPR is that data from an electronic system can be used anonymously for statistical reporting in matters such as quality improvement, resource management and public health communicable disease surveys.13,14

A major disadvantage of the EPR is its connection to the Internet, which makes it vulnerable to ‘hacking’ or unauthorised access. Eavesdropping and skimming can also occur when sensitive data for remote patient monitoring are transmitted wirelessly. In contrast, patients need to appear in person at a health-care facility in order to be monitored or to access paper-based medical records in the traditional health-care system, which restricts the number of personnel that have access to their information. There is therefore a greater challenge in ensuring data security and the integrity of the EPR compared to traditional health-care systems.

Transformation from eHealth to mHealth

There is no consensus on the definition of eHealth (or electronic health) as it goes beyond the use of the ICT ecosystem. eHealth is defined by Eysenbach15 as ‘an intersection of medical informatics, public health and business, referring to health services and information delivered or enhanced through the Internet and related technologies’. eHealth involves patients and stakeholders that deliver good-quality health-related services at low cost.16 In addition, the term ‘eHealth’ reflects
Increasing development in mobile technology has positively influenced mobile health (mHealth) delivery services. mHealth is a new term for health-care practitioners who use a mobile phone, a voice recorder, telehealth services, patient monitoring devices, personal digital assistants (PDAs) and other mobile devices in their practice. mHealth forms part of an increasing movement towards citizen-centred health-care delivery. mHealth involves new technology, policies, devices, systems and standards for communication between patients and health-care providers, integration of applications and communication-enhanced disease management programmes, collaboration and care coordination systems and much more. Expansion in telecommunication networks and the use of smarter handsets has transformed weak health systems and assisted in combating health challenges ranging from maternal and child illnesses and mortality to chronic and infectious diseases. Mobile technology has enabled remote and isolated communities to communicate in real time in a way that was not possible before. mHealth takes into account a patient’s literacy, clinician and staff education, a strategy for wireless connectivity, an inventory of existing medical applications and creates a management system for incoming emails and text messages. Communication-enhanced health-care through mobile technology is a paradigm for the future, but it may be inhibited by the costly infrastructural requirements of sophisticated technology, a lack of availability of highly skilled operators, unreliability of Internet connectivity and the amount of training and re-training of available personnel. Despite these limitations, mHealth is expected to bring a revolutionary change to health-care delivery systems because of the exploding field of mobile digital tools like PDAs, enterprise digital assistants, tablet computers, smartphones and sensor gadgets.

mHealth requires mobile devices and other mobile technologies that are not needed by eHealth systems. The scaling up of technologies for implementation of mHealth may be a most promising investment in developing countries, especially in Africa, because of the support that can be provided to health workers in remote locations. The technology for mHealth can reach people anywhere and at any time, because it is continually expanding with sophisticated 3G networks and mobile broadband.

Privacy and security concerns

Before now, health-care information system software vendors and health-care providers have had the philosophy, ‘make it work first, then think about the security later’, for most of the technology installed. The revolutionary emergence of the personal computer in 1980, coupled with inevitable systems of networking, has brought about today’s data security. But technological changes in both the computer world and the health-care delivery information system have made securing these systems a priority in order to protect the confidentiality of information.

As advantageous as the technological aspects of an EPR are for health-care delivery systems, the benefits need to be balanced against the privacy and security concerns of the patients. Information needs to be captured, stored and maintained in a database such that the integrity and confidentiality of the information are guaranteed. Patients are entitled to be informed of their conditions, which necessitates ready access to all relevant health-care information. The integrity of the stored data can be compromised deliberately or through carelessness on the part of the personnel. For instance, data may be manipulated to gain advantage in an insurance policy or claim, or for the benefit of gaining employment.

Securing data in a distributed environment over the mobile network has been a greater challenge than doing so in a centralised system, although the total failure of a central system is more costly than one or more elements failing in a distributed system. The mobility of data in the process of distribution results in decentralisation and the spread of data security concerns. Whereas, abuse of privilege is more prominent in a centralised system, especially if a person authorised to use the system illicitly gains access to restricted security codes or measures.

There are three basic elements of data security to be considered: confidentiality, integrity and availability. To establish a level of confidence in the data, health-care organisations must process all confidential data so that it is not disclosed to those to whom it should not be, whether the disclosure is accidental or malicious. There have been several instances in which information about a patient’s health has been ‘leaked’; such unauthorised disclosure, whether it be the health information of a public figure or a private citizen, can ruin a person’s career, affect their insurability or destroy their life. For example, the Sunday Times released the medical report of Dr Manto Tshabalala-Msimang to the public in an article entitled ‘Manto: A drunk and a thief’, published on 19 August 2007. A controversy erupted and the newspaper was sued for divulging medical information. Surprisingly, instead of querying the protection of a patient’s medical information against unauthorised disclosure by hospital staff, the attention was focused on whether the Sunday Times had permission to publish such information. Thus technology should be used and policies should be put in place to protect the confidentiality of electronic health-care information.

Data integrity is not always associated with security, particularly in the eyes of the general public. Protecting the integrity of data means ensuring that the recorded information is correct and is not in any way corrupted. A corrupted patient record is a serious problem and could lead to the death of a patient. The third and the last major aspect of the data security concept is system availability. Computer
The protected transmission of confidential information is a serious matter within the health-care system. Experts have warned against the transmission of highly confidential information, such as diagnostic test results, to avoid the possibility that a patient’s privacy can be breached. Patients have a right to the confidentiality and privacy of their medical treatment. Ethical and legal guidelines state that health workers must keep all patient information confidential unless the patients’ consent is sought and that such information cannot be divulged to a third or unauthorised person.\textsuperscript{24,25}

The case of Mr McGeary is an example of an infringement of patient rights. Mr McGeary needed an HIV test to apply for a life insurance policy. His doctor performed the HIV test, which was positive. His doctor told two other people (a doctor and a dentist) of the result. Other people then learned of Mr McGeary’s HIV status from these two people. Legal action was instituted against the doctor for breaching Mr McGeary’s legal and ethical rights to confidentiality.\textsuperscript{26,27}

The revised guideline\textsuperscript{28} of the Health Professions Council of South Africa – a body that regulates the activities of health professionals practising in South Africa – is as follows:

No practitioner may divulge verbally or in writing any information which ought not be divulged regarding the ailments of a patient except with the express consent of the patient or in the case of a minor, with the express consent of his guardian, or in the case of a deceased patient, with the consent of his next-of-kin or the executor of his estate.

Data security methods, like cryptography, digital watermarking and steganography, employed in the transmission of health information under a secured noisy channel could be the panacea or better alternative for ensuring the confidentiality of the health information. These methods are discussed within the context of protecting health information transmitted using ever-growing mHealth technology.

**Encrypting**

Encrypting prevents a third person from understanding patient information if it is intercepted. A patient’s record can be digitally scrambled such that only authorised people who possess the ‘key’ to the encryption can transform the data to its original form. Encryption can be symmetric or asymmetric.

Symmetric encryption systems provide a two-way channel for their users: A and B share a secret key and they can both encrypt information to send to the other as well decrypt information in the reverse manner. Authentication is genuine as long as the message received was not fabricated by someone other than the declared sender. The only challenge in this scheme is how the secret key is sent to the recipient and key distribution can be difficult, especially if there is a need for another user. In general, \( n \) users who want to communicate in pairs will need \( n(n – 1)/2 \) keys. What this means is that the number of keys needed increases at a rate proportional to the square of the number of users.

Asymmetric encryption systems involve each user having two keys that are unique to them – a public key and a private key. A trusted third party is used to facilitate secure interactions between the two parties. The user may send the public key freely because each key is used for only half of the process. That is, one key decrypts the encryption made by the other and vice versa. Only the corresponding private key (presumably it is kept private) can decrypt what has been encrypted with the public key.\textsuperscript{29}

Table 1 shows a comparison between symmetric and asymmetric encryption systems in terms of their transformational speed, diffusion of information, propagation of error and insertion of symbols.

Encrypting patient information before transmission can help to protect the information, although anyone who obtains the key can access the data. The key to the success of encryption is to limit the number of personnel who have the key to encrypt and decrypt the data, and to determine the most appropriate length of the key.\textsuperscript{29,30}

**Digital watermarking**

Digital watermarking of data provides a means to protect information in cases where access control to the information may be compromised. It is the art of embedding data (as a watermark) into a multimedia object, such that the watermark can be detected or extracted later without impairing the object. Watermarks are often inserted into images that can be detected when the image is compared

<table>
<thead>
<tr>
<th>Encryption type</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symmetric (stream encryption algorithm)</td>
<td>Transformational speed is high because the symbol is encrypted without regard for any other plain text symbols – each symbol is encrypted as soon as it is read. Encryption algorithm is the factor that determines the time to encrypt a symbol, but not the time it takes to receive the plain text. Low error propagation: an error in the encryption process affects only that character, because each symbol is separately encoded.</td>
<td>Diffusion is low: each symbol is encrypted separately. The symbol’s information is contained in only one symbol of the cipher text. Malicious insertion and modification: the symbols are separately enciphered, which allows the code to be compared with a similar or previous message and allows a counterfeit or new message that may look genuine to be transmitted in place of the original.</td>
</tr>
<tr>
<td>Asymmetric (block encryption algorithm)</td>
<td>High diffusion: information from the plain text is diffused into several cipher text symbols. One cipher text block may depend on several plain text letters. Difficulty in symbol insertion: enciphering is done based on blocks of symbols therefore it is rather difficult to insert a single symbol into one block, otherwise the length of the block will be incorrect.</td>
<td>Slow encryption: all plain text symbols will have to be received before the encryption process commences. High error propagation: if an error occurs in the block, it will spread across the block and affect the block transformation.</td>
</tr>
</tbody>
</table>
with the original. Watermarks used for copyright protection are designed to identify both the source of the image as well as its authorised users. Public key encryption, such as the RSA algorithm (invented by Ronald L. Rivest, Adi Shamir and Leonard Adleman in 1977), does not completely prevent unauthorised copying because of the ease with which images can be reproduced from previously published documents. All ‘watermarked’ documents and images must be extracted before they can be read and disseminated. Chang-Tsun et al.\textsuperscript{39} presented a role-based access control framework using data hiding techniques for combating security threats in a picture archiving and communication system. Access to the databases and the information contained in the pictures, in this case mammograms, was controlled through the issuance of a stego key and a watermarking key. Wilson et al.\textsuperscript{3} suggested using a steganographic filing system for storing and accessing information that is distributed over different medical records and in different locations. This filing system was designed with the intention of concealing the existence of the files and authorised users were required to be aware of the existence of the file, then supply a file name and associated password to access the desired file. This method uses initialisation of the file system with several randomly generated cover files. A newly created file is embedded within a single cover file or a subset of cover files.

**Steganography**

Steganography is the ancient art and science of hiding information by embedding messages within other, seemingly innocent-looking messages. The word steganography is derived from a work by Trithemius (1462–1516) entitled *Steganographia* – a Greek word meaning ‘cover writing’. Steganography (hiding the message being communicated) differs from cryptography (obscuring the meaning of the message). The communication medium is referred to as the cover object, the ‘stego’ object is the embedded message and together they form the stegosystem. A stego key keeps the operation secure and stego objects cannot be extracted from cover objects within the stegosystem without the stego key.\textsuperscript{32,33} Concealment of secret messages within a natural language has been in existence as early as the 16th century. However, the increase in digital information transmission and distribution has resulted in the spread of steganography from ordinary text to multimedia transmission. An example of such communication is the null cipher.

The null cipher applies a series of characters and words intended to confuse a hacker. The communication appears as nonsense, but can be decoded to a meaningful message. This is an ancient form of encrypted communication in which a message is surrounded by a large number of redundant characters (known as null ciphers). This form of communication is, in fact, known to have been used by the German army during World War II. The following is an example of a null cipher form of steganography: ‘Apparently neutral’ protest is thoroughly discounted and ignored. Isman hard hit. Blockade issue affects pretext for embargo on by-products, ejecting suets and vegetable oils.’ Decoding this message by extracting the second letter in each word reveals the message: ‘Pershing sails from NY June 1.’

The drawback of this form of steganography is that the message sender is forced to make a text cover according to a preset procedure, hence defeating the purpose of steganography. Also, applying a ‘brute force’ approach to decoding will reveal the message.

**Database security**

Database security refers to security within the server, excluding data transmission across the network. An advantage of database technology is the ability to perform data mining – a technique that involves the use of analytical tools to study corporate data in order to increase the efficiency of the organisation. Data mining allows for information sharing with other organisations.\textsuperscript{34} However, information sharing has security implications and so restricting access to the database is essential. Restricting access can be achieved by a multilevel security database, for which access is controlled by policies that are enforced and limit the sharing of information to only those who are authorised.\textsuperscript{35} Undesired data mining is resolved by getting an integral part of the data mining with some guiding rules so as to make data manipulations difficult for an unauthorised user;\textsuperscript{36} applying such rules will reinforce the security of the database.

Prevention of unauthorised data mining can be achieved by:

- Limiting access to the database. Eliminating grouping of the database structure by higher-order digits or ‘unique identifiers’. For example, grouping can be done with reasonable reliability by location, sex or age.
- Augmenting the data without altering its usefulness, if we have pre-knowledge of the way the data are to be used. Misleading data can then be added as they will only be retrieved by inappropriate queries.
- Auditing the database, which will discourage legitimate users against the indiscriminate misuse of their privilege. Although this approach does not enforce control, it does detect misuse by legitimate users.\textsuperscript{34}

Existing solutions to the problem of database security are:

- Role-based access control: Because of the complexity of security in a multi-user environment,\textsuperscript{37} control of a database is restricted by the degree of the user’s involvement in the patient’s treatment. For instance, a medical insurer will not be given the same access to the patient’s treatment as the physician; likewise the physician will not have access to the financial matters of the patient’s care, whereas the insurer will.
- Encryption: Encryption is used to ensure security of the data and help in protecting against eavesdropping and skimming. It includes both software and hardware and it is always better to use both forms to ensure the greatest degree of security.
- Authentication assurance or mechanisms: This solution works by confirming that data are being received from the person or entity claimed.\textsuperscript{37} Authentication algorithms (such as passwords, digital signatures and challenge response authentication protocols) play a major role in this security measure being successful.
Recommendations for improvement on data security and privacy

Medical data in the new technological dispensation can be secured within the database server and also during transmission by doing the following:

- Defining clear attributes for role-based access as the systems are put into place.
- Developing policies to protect the patient’s right to privacy with regard to their medical data.
- Defining the extent of medical data transmitted via the Internet from patients’ homes to the central monitoring station, and whether patients have partial or full control of their data.
- Specifying within data mining rules and technological measures who has the right to analyse the data. As EPRs are becoming widespread, more health organisations will have databases that store patient information in a common computerised format, allowing the sharing of data over the communication network; hence the administration of the particular part of the data in circulation has to be secured or restricted.

Conclusion

Methods of protecting electronic health data have been discussed and weaknesses in real-world applications have been highlighted. Many of the existing data security techniques are not yet robust enough to prevent detection and removal of embedded data. Notably, the quality of the media should not noticeably be degraded upon addition of a watermark; watermarks should be undetectable even in the presence of the payload of the message (or message content), multiple watermarks in a payload should not interfere with each other, watermarks should survive ‘hacking’ attacks and, most importantly, digital watermarks should not degrade the payload message. Hence, it is suggested that implementation of digital watermarking should be complemented with data encryption mechanisms to improve the assurance and integrity of the data stored, retrieved or transmitted across electronic devices. It is vital that both patients and healthcare workers have confidence in the confidentiality and integrity of the information and data, and the security of the transmission channels.

References

34. Clifton C, Marks D. Security and privacy: implications of data mining. Workshop presented at: ACM SIGMOD Workshop on Data Mining and Knowledge Discovery; 1996 June 2–3; Montreal, Canada.