Dependability and Security in Medical Information System

Xukai Zou[1], Yuan-Shun Dai[1], Bradley Doebbeling[2], and Mingrui Qi[1]

[1] Department of Computer and Information Science
Indiana University, Purdue University, Indianapolis
IN 46202, USA
{xkzou,ydai,qim}@cs.iupui.edu
[2] School of Medicine
Indiana University, Purdue University, Indianapolis
IN 46202, USA
bdoebbel@iupui.edu

Abstract. Medical Information Systems (MIS) help medical practice and health care significantly. Security and dependability are two increasingly important factors for MIS nowadays. In one hand, people would be willing to step into the MIS age only when their privacy and integrity can be protected and guaranteed with MIS systems. On the other hand, only secure and reliable MIS systems would provide safe and solid medical and health care service to people. In this paper, we discuss some new security and reliability technologies which are necessary for and can be integrated with existing MISs and make the systems highly secure and dependable. We also present an implemented Middleware architecture which has been integrated with the existing VISTA/CPRS system in the U.S. Department of Veterans Affairs seamlessly and transparently.

Keywords: Medical Information System (MIS), Health Information Technology (HIT), Dependability, Security, Tele-medicine.

1 Introduction

The medical information system (MIS) is a typical collaborative computing application in which people such as physicians, nurses, professors, researchers, health insurance personnel, etc. share patient information (including text, images, multimedia data) and collaboratively conduct critical tasks via the networked system.
In 2004, President Bush outlined a plan to have electronic health records (EHR) for most Americans within ten years. This is a clear indication of the importance of MIS to medical practice and health care services. Medical/health data is extremely important; it can help physicians to diagnose patient's diseases, cure patient's illness, recover people's health, and save their lives. On the other hand, medical/health data is most sensitive and if used or disclosed inappropriately, it could jeopardize people's privacy and endanger their lives. The greatest concern people have with MIS is the security, privacy, and confidentiality of their personal health information [9]. People would be willing to step into the digitized health information age only when

J. Jacko (Ed.): Human-Computer Interaction, Part IV, HCII 2007, LNCS 4553, pp. 549–558, 2007.

individual's privacy/integrity can be protected and guaranteed within the MIS systems. Accessing patient data must be monitored, controlled, and granted only to authorized users. As a result, security of the MIS is very critical for those digitized health services. From the administrative/managerial point of view, governments in different countries have stipulated regulations about medical data. For example, the U.S. Department of Health and Human Services passed The Health Insurance Portability and Accountability Act (HIPAA) in 1996 [1]. However, from the technical point of view, how to enforce the security of MIS is a challenging issue. This paper aims to answer the question and provides a practical solution to it.

Dependability is also very important for MIS. For example, it could be potentially fatal if a patient's records fail to be loaded due to system failures and they are unavailable in the event of life-threatening emergency. Therefore, the MIS has very high requirements for reliability. The newly developed grid computing technology is applied to MISs and proves to be of great reliability without sacrificing performance.

In summary, the MIS needs very high levels of dependability and security. We have implemented our dependability and security mechanisms into a real MIS, in collaboration with VA (Veterans Affairs) Hospitals and DoD (Department of Defense) in the U.S. [11]. The key management services and security modules are integrated into this MIS system without modifying the existing system components including the client GUI interface, the MIS application server, and MIS databases.

Another very important collaborative service offered by the MIS is tele-medicine via a way of video conferencing and virtual reality. In this scenario, doctors, specialists, surgeons, nurses, etc. from different places can conduct consulting, diagnosis, and surgery through the Internet, just as they would gather together to have a face-to-face meeting to discuss or diagnose patients. The primary issue for this kind of tele-medicine is the transmission and processing of different types of signals such as images, video, audio, multimedia, text, documents, etc. Clearly, the security and reliability of these confidential signals, activities and services are critical. The new MIS, after integrated with our new security and reliability mechanisms, guarantees various security needs and provides secure and reliable tele-medicine service.

The remainder of the paper is organized as follows. Section 2 introduces the background of the MIS; Section 3 describes the middleware architecture for key management services that are used in the MIS; Section 4 presents how to improve the MIS reliability by integrating the Grid computing technology; and Section 5 concludes this paper and describes possible future research.

2 Health Information Technology (HIT) and Medical Information System (MIS)

Information technologies have been widely used in every field of modern society including health care. However, its adoption to healthcare industry lags behind other industries by as much as ten to fifteen years [2, 9]. Medical Information System (MIS) / Health Information Technology (HIT), defined as “the application of information processing involving both computer hardware and software that deals with the storage, retrieval, sharing, and use of health care information, data, and knowledge for communication and decision making" [13], was initially used for

financial accounting of medical transactions [9]. The first EHRs were designed and deployed in the late 1960s and early 1970s. HIT/MIS tends to help reduce costs, to liberate resources to satisfy unmet needs, and to restore competitiveness [9]. In particular, when combined with the Internet, "HIT is expected to foster patient-focused care, to promote transparency in prices and performance, …" [9]. There are at least two desires which motivate the adoption of HIT/MIS: (1) increasing productivity achieved in other industries that have extensively adopted Information Technology (IT), and (2) reducing the incidence of medical errors and improving quality by relating patient outcomes to care processes, along with an increased ability to measure and pay providers for performance [9]. The primary benefit associated with the adoption of HIT/MIS is: reduced cost and improved quality of health care. Many challenges are also associated with HIT/MIS including Complexity of the health care enterprise; Magnitude of the investment and who should make it; the absence of standards; a lack of interoperability; and the danger that acceleration of using inadequate products may increase the total investment and discourage providers [9]. Many obstacles exist along with the way in developing and deploying HIT/MIS such as (1) enormous amounts of money may need to be expended or even be wasted and (2) vast expenditures must be invested in the short-term while most benefits can only be gained in the long-term [9]. All these factors can affect people's trust and confidence in MIS. But they are not people's primary concerns. What people are concerned about most is "the security, privacy, and confidentiality of their personal health information." [9]. We are putting our effort in this direction and embedding our innovative technologies into existing MISs to create the new generation MIS system.

Like many information systems in other fields, the MIS basically consists of three components (tiers): client/user interface, database, and application servers. For example, in the VISTA system (Veterans Health Information System and Technology Architecture) of the U.S. Department of Veterans Affairs, the three components are CPRS client (Computerized Patient Record System), VISTA application server, and Database (See: http://worldvista.sourceforge.net/vista/history/index.html). We implemented and integrated our security modules into VISTA to transform VISTA into a secure MIS system [11]. Apart from the provided security services such as confidentiality, integrity, authentication, fine-tuned access control, and detection of stealing and masquerading attacks, one prominent feature of our new technology is the seamless and transparent integration in the sense that the security modules are integrated without modification of any existing components and users can continue to use the system in the same way as before. In addition, different from other integration technologies which are based on Master Patient Index (MPI), the integration under our new technology can be accomplished without requirement for MPI.

3 Architecture of the Proposed Secure MIS

As discussed above, dependability and security are very important and imperative for new generation MIS systems. We have invented advanced security and dependability technologies such as the Access Control Polynomial (ACP) based key management mechanism and grid-computing based reliability services. These technologies are well

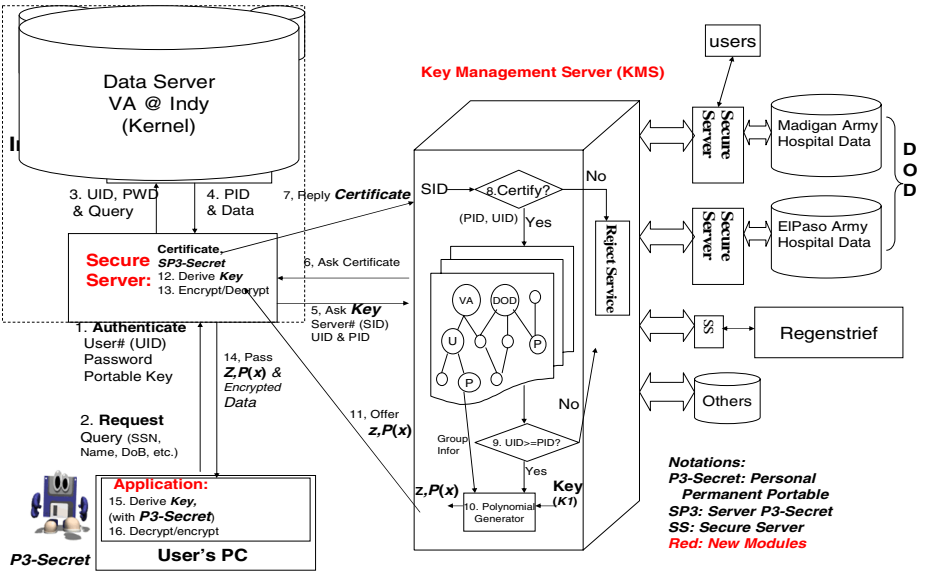

Fig. 1. ACP based Secure Medical Information System

suited for MIS systems. Due to the power of our new ACP mechanism (see Section 3.2), different functions/services required in the MIS can be implemented efficiently and flexibly. We have implemented our ACP mechanism into a real MIS. The result has been recently featured by the news [11]. We present the result and describe the architecture of the implemented secure MIS system here. Some parameters and data we have cited are either normalized or shifted so as to maintain the patients' data privacy as put forth in the agreement with the VA and DoD Sponsors, but we still retain sufficient context and clear examples for the purpose of illustration. Moreover, besides securing MIS systems, the ACP mechanism can be applied to various other collaborative computing applications that require all those security needs, such as financial systems, information forensics, government systems, distributed signal processing, etc.

3.1 System Architecture

Figure 1 shows the architecture of the ACP based secure MIS system. Three main security components are added into the existing MIS system: ACP Key Management Server (KMS), Secure Server (SS), and User Application (UA). The UA module is integrated in the existing client application, the SS is configured to stand at front of each existing database/application server and the KMS offers comprehensive key management services (It is worthy to note that the integration did not change anything in the existing systems). The three components work together to provide comprehensive security services involved in MISs. Let us begin with the simple case that a user (e.g., a doctor) accesses a patient data and how access control and confidentiality are enforced, and then dig into more interesting but complicated functions and how does the ACP mechanism support them.

When a doctor accesses a patient's data (assume that the doctor has logged in and been authenticated) (Step 1,3 in Figure 1), the request with the patient's basic information is sent to the SS (Step 2) which is standing in front of the database server. The SS passes the request to the database server and obtains the patient's record including patient ID (PID) and other required fields (Step 4). The SS then passes the user's UID and the patient's PID to the KMS (Step 5). The KMS then checks the access control hierarchy against (UID, PID). Because the hierarchy specifies which groups of users can access which groups of patient records in which granularity, the KMS can validate whether and how the user can access the patient (Step 9). If yes, the KMS generates a random key and the ACP polynomial A(x) according to the hierarchy and hides the key in P(x) (Step 10) and sends (z, P(x)) back to the SS (Step 11, See Eq.(1) and Eq.(2) in Section 3.2 for the definition and formula of A(x) and P(x)). The SS then derives the key from P(x) (Step 12) and encrypts the data (Step 13) and sends the data along with (z, P(x)) to the user (Step 14). The user can derive the key from P(x) (Step 15) and the decrypts the data (Step 16). Note: (1) the SS is treated as a (super)user and belongs to the highest node of one organization's hierarchy. (2) the ACP mechanism does not constrain data types. Patient data could be in the form of text, images, video, audio, multimedia, etc. Our security scheme can seamlessly encrypt all of them and guarantee their confidentiality and integrity.

3.2 Access Control Polynomial (ACP)

Access Control Polynomial (ACP) is the core component for key management and secure information exchange/sharing among a group of users. We briefly discuss it here. Assume that every valid user (doctor, physician, nurse, etc.) in the system is assigned a Permanent Personal Secret (PPS), denoted by SID_i for user U_i. Assume q is a large prime which forms a finite field F_q and acts as the system modulus.

Whenever there is a group of users participating in a health care service, the Key Management Server (KMS) constructs a polynomial $A(x)$ in finite field $F_q[x]$ as:

$$A(x) = \prod_{i \in \psi} (x - f(SID_i, z)) \tag{1}$$

where ψ denotes the group under consideration and SID_i are group members' PPSs assigned to the members in ψ. $f(x, y)$ is a public one-way hash function and z is a random integer from F_q. $A(x)$ is called an *Access Control Polynomial* (ACP). From Eq.(1), it is apparent that $A(x)$ is equated to 0 when x is substituted with $f(SID_i, z)$ by a valid user ψ; otherwise, $A(x)$ is a random value.

The KMS selects a random group key K for group and computes the polynomial:

$$P(x) = A(x) + K \tag{2}$$

Finally, the KMS publicizes $(z, P(x))$.

From this public information, any group member U_i can get the key by:

$$K = P(f(SID_i, z)) \tag{3}$$

Here U_i computes $f(SID_i, z)$ first and then substitutes into $P(x)$.

For any other member U_r excluded by ψ, $P(f(SID_r, z))$ yields a random value from which U_r cannot get *K*. This key management mechanism guarantees that only a user whose SID_i is included in $A(x)$ can extract the key from $P(x)$.

With this scheme, dynamic groups can be easily managed to accept and revoke users. If a new user U_t needs to be added, the KMS creates a new SID_t and assigns it to U_t. Then, the KMS includes $(x - f(SID_t, z))$ in the formation of $A(x)$ as:

$$A'(x) = \prod_{i \in \psi} (x - f(SID_i, z))(x - f(SID_t, z)) \tag{4}$$

$A'(x)$ is used to mask key *K* by computing $P'(x) = A'(x) + K$. Then $(z, P'(x))$ is sent to U_t. U_t can use SID_t to derive the key from Eq.(3).

If a current group member U_t needs to be revoked from the group, the KMS just selects a new random z' and recomputes $A'(x)$ by excluding the corresponding $(x - f(SID_t, z'))$. Then, the KMS selects a new group key K', computes $P'(x) = A'(x) + K'$, and multicasts $(z', P'(x))$. Now, the deleted user U_t cannot extract K' from $P'(x)$.

3.3 Security Functions Supported by ACP

All the security functions, i.e. Secure Group Communication (SGC)/Secure Dynamic Conferencing (SDC), Differential Access Control (DIF-AC), Hierarchical Access Control (HAC), prevail in the MIS and are supported by our ACP mechanism. The central component in the MIS system is hierarchical access relation between users (e.g., doctors, nurses, hospital clerks, insurance personnel, etc.) and patients. This is a typical HAC scenario which has been correctly implemented and enforced by our ACP mechanism. For example, in Figure 2, suppose a physician p_2 in node P_2 wants to access patient data d_8 in node P_8. The KMS will generate a random key *K*. The patient data will be encrypted using *K* before the data is sent back to the physician. The *K* will be hidden in the polynomial $P(x) = A(x) + K$ where $A(x) = (x - f(CID_8, z))(x - f(CID_5, z))(x - f(CID_4, z))(x - f(CID_2, z))$ and CID_i is the secret node ID for node *i* and it is only known to the users of node *i*. Thus, physician p_2 can obtain *K* first and then decrypt the patient data and use it.

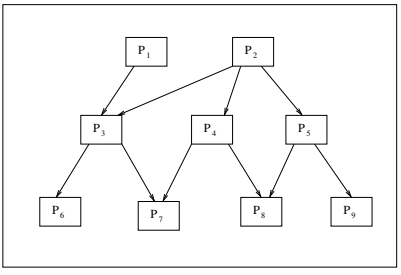

Fig. 2. An access hierarchy in MIS

The Differential Access Control (DIF-AC) also appears in MIS. One basic format of DIF-AC is that one patient has multiple doctors such as his family physician and multiple specialists who share the patient' information; on the other hand, a family physician can have multiple patients. Another format of DIF-AC is the following many-to-many relation: for any patient in a node, his data can be accessed by all the users in this node and the ancestral nodes; correspondingly, a user in a node can access the patient data of his node as well as all his descendant nodes. Yet, some other DIF-AC format is that a doctor transmits some notifications to his patients or a hospital announces some events to certain patients. No matter which kind of DIF-AC, our ACP scheme can support and enforce it correctly.

SGC/SDC scenarios occur often in MIS. (Secure Group Communication (SGC) refers to a setting in which a group of members can communicate (or share the information) among group members, in a way that outsiders are unable to understand the communication (or the information) even when they are able to intercept the communication (or steal the information). Secure Dynamic Conferencing (SDC) refers to a scenario where any random subset of the given universe of users can form a secure communication (sub)group.) One case is that a patient's data is located in several databases. The data need to be collected from all these databases and encrypted for transmission. Thus, the SSs in front of these databases form a temporary conference and the KMS generates a key for this transmission and hides the key in the P(x) which is formed from these SSs. Another typical scenario is tele-medicine. Our ACP based SGC scheme enables tele-medicine to be highly secure with great power.

Apart from its supports to SGC/SDC, DIF-AC, and HAC, the ACP mechanism also provides several important advantages for MIS such as seamless integration of different data sources, fine-tuned access control in flexible granularity, and portable login and accessing to patient data from anywhere.

3.4 Tele-Medicine Service

Tele-medicine is the utmost goal of MIS systems. In tele-medicine, physicians and specialists from geographically distant locations collaborate to investigate some common disease such as SARS, diagnose a patient illness, and participate in or instruct a tough surgery. Most of them need to share the real-time image signals/streams and audio/video information. The size of data transmitted through the Internet could be huge (such as images). Therefore, the communication channels

would be unstable due to the long distance. The above ACP based security mechanism is generic and independent from data content, formats, or volumes; therefore, it can be applied to secure tele-medicine. As for reliability of tele-medicine, we will introduce our new grid-computing based reliability mechanism which can provide comprehensive reliability services for MIS in Section 4. Here we propose using an effective technology to improve the reliability for data transmission: checkpoint. During the transmission, checkpoints are set up. At each checkpoint, the sender and receiver exchange signals to make sure the data that have been received before this point are correct. The signal sent by the receiver can include the check information. The sender will compare the signal from the receiver with the real values. If matched, the information before the checkpoint has been correctly received. Thus, the data being transmitted between two checkpoints is temporarily cached in the memory. Otherwise, some errors occurred during the communication, and the cached information between the two checkpoints will be resent. When it passes the check, the temporary cached data can be deleted and the cache storage can be used for the next section of data. As a result, this scheme can reach a very high reliability and is very efficient compared to the prior scheme which includes resending the large file again when the receiver cannot read it (perhaps just due to the loss/error of a small section of information). In addition, the overhead of this scheme is small (<1%) in time.

Apart from the security and reliability guaranty for tele-medicine, we propose to enhance the Tele-medicine service including the following functions: remote diagnosis involving multiple doctors' collaboration, real-time and mobile communication and data sharing as dynamic conferencing, and faster, smoother multimedia stream (e.g. visualization) by the Grid Computing and P2P technology.

4 Dependable MIS Based on Grid Computing Technology

Reliability is extremely important for information systems, including MIS. Traditional techniques of using duplication can enhance the capability of fault-tolerance and thus, reliability, but can potentially affect the performance. Tai et al. [12] mentioned that there exists tradeoff between performance and reliability. When more components are duplicated for completing the same subtask in order to reach a high reliability, the performance can be decreased due to the lack of full load sharing. On the contrary, if a task is fully divided into disjoint subtasks shared by the resources without duplication, any single failure can cause the incompletion of the task. Grid computing [6] is a newly developed technology for complex systems with large-scale resource sharing, wide-area communication, and multi-institutional collaboration etc, see e.g. [3, 5, 7, 8, 10]. Many experts believe that grid technologies will offer a second chance to fulfill the promises of the Internet. A novel Grid computing architecture proposed in [4] offers a new reliability technique, instead of the prior duplication, and thus performance and reliability do not negatively affect each other but both are improved.

The new services based on grid computing technology include the caching and coherence service, backup and recovery service, fault-tolerant service, and self-healing service [4]. We briefly describe each of them below.

Caching and coherence service. By building a caching module using a portion of disk in the client machine to temporarily store the recently requested results, one can

in fact retrieve most of the data directly from its cache on the local disk when the user requests the data again. By this means, the caching service can alleviate the bottleneck of network transmission and achieve high performance. One problem with caching is data consistency, which means that the data in one's local cache may be out-of-date and inconsistent with the data in the server or other caches. Coherence service with update request buffer will solve this problem [4].

Backup and recovery service. Traditionally, the backup of a database was to duplicate at least one redundant databases. This will double cost. The new grid based backup service can significantly save money while maintaining the same availability as the duplicated database backup scheme. The idea is as follows. During subscription to the Resource Management Service (RMS), a data source will be asked to reserve a portion of its disk space for the RMS to use, and as a return, same portion of space for backuping its data will be free of charge. Then the data sources can mutually backup each other's data. (Note: in order to prevent the data of one data source A from the access of data source B when it is backed up in B, encryption and key management mechanisms can be applied.) This way, once a data source was damaged totally, its data can be completely recovered from the backups in other data sources.

Fault-tolerant service. Fault-tolerance means that a user can successfully get the requested data even though some faults or failures may occur when he requests the service. The grid-based information system can provide a strong capability to tolerate faults without additional cost or hardware through the following steps: fetch from one's own local cache, visit to its data source, go to backup data source, search other local caches, and resort to self-healing.

Self-healing service. The self-healing mechanism will make use of the computation resources in the Grid system (such as the idle machines). The following three services, when integrated together, realize the intelligent self-healing service: Fault-Detection service, Diagnosis service, and Curing service. Fault-detection service will be automatically run in the grid system and detect any system fault or failure. Once faults or failures are detected, the diagnosis service is triggered. It will monitor (specific) regions and nodes and analyze the causes for the combination of symptoms such as the over-used memory, the over-slow CPU, the virus infections, the network traffic jams etc. After diagnosis, the curing service is activated. It will match some possible prescriptions based on the diagnosed results and assign actuators to heal the problem according to the prescriptions.

The MIS system is very sensitive to reliability. An unreliable or faulty MIS system could be a disaster and may damage people's lives. Fortunately, the MIS system we are designing includes the reliability model and modules as its coherent components. In particular, the above described highly robust reliability services when integrated into the MIS system, will greatly increase the dependability of the MIS system.

5 Conclusion

The secure and dependable medical information system will benefit everyone. Security and reliability come as two important and safety-critical concerns in

developing any mission-critical information systems like medical information system. We, collaborating with the Veterans Affairs hospitals (VA) and Indiana University

Medical School (IUMS), have implemented a preliminary secure medical information system [9]. The security and reliability model and the seamless integration capability of the model have been proven via this system. We are planning to extend it into the larger scope of MIS systems such as Regional Health Information Organization (RHIO) and finally the nation-wide health information system.

This paper presented the core concepts and techniques of security and reliability for MIS system. Along with further development of the MIS systems, new modules and methods will be designed and presented.

Acknowledgments. The authors appreciate the support and VISTA platform provided from the U.S. Department of Veterans Affairs Medical Centers and IUMS.

References

[1] Health Insurance Portability and Accountability Act of 1996 (HIPAA). U.S. DHHS (1996)
[2] Is it the cure? Economist (May 2003)
[3] Berman, F., Wolski, R., Casanova, H., Cirne, W., Dail, H., Faerman, M., Figueira, S., Hayes, J., Obertelli, G., Schopf, J., Shao, G., Smallen, S., Spring, N., Su, A., Zagorodnov, D.: Adaptive computing on the grid using apples. IEEE Transactions on Parallel and Distributed Systems 14(14), 369–382 (2003)
[4] Dai, Y.S., Zou, X., Guo, Y.: Novel grid services for data access with high performance, fault tolerance and self-healing. Inter. J. of High Performance Computing and Networking (In press 2007)
[5] Das, S.K., Harvey, D.J., Biswas, R.: Parallel processing of adaptive meshes with load balancing. IEEE Transactions on Parallel and Distributed Systems 12(12), 1269–1280 (2001)
[6] Foster, I., Kesselman, C.: The grid 2: Blueprint for a new computing infrastructure. Morgan Kaufmann, San Francisco (2003)
[7] Foster, I., Kesselman, C., Nick, J.M., Tuecke, S.: Grid services for distributed system integration. Computer 35(6), 37–46 (2002)
[8] Foster, I., Kesselman, C., Tuecke, S.: The anatomy of the grid: Enabling scalable virtual organizations. Inter. J. of High Performance Computing Applications 15(2), 200–222 (2001)
[9] Goldschmidt, P.G.: HIT and MIS: implications of health information technology and medical information systems. Comm. of ACM 48(10), 68–74 (2005)
[10] Kumar, A.: An efficient supergrid protocol for high availability and load balancing. Transactions on Computers 49(10), 1126–1133 (2000)
[11] Schneider, R.: IUPUI computer scientists develop revolutionary medical information system (2006) http://www.iupui.edu/news/releases/060222_med_info_system.htm
[12] Tai, A., Meyer, J., Avizienis, A.: Performability enhancement of fault-tolerant software. IEEE Transactions on Reliability 42(2), 227–237 (1993)
[13] Thompson, T., Brailer, D.: Health it strategic framework. U.S. DHHS (2004)