Global SA&D Approach for New Information Systems Architectures Security

Snene Mehdi, Pardellas Jorge, *Database Lab, University of Geneva*

***Abstract*—Information System security is often assimilated to a set of software solutions (Firewall, data encryption,..) but rarely consider the organizational security rules as a fundamental part of the IS security policy. With the increasing use of new IS architectures (Open architecture, distributed database, web server, multi-tier application servers) security leaks become crucial and every security problem is harmful to the organization business continuity. To reduce and detect major security risks at an earlier step of the IS project, our approach is based on different knowledge exchange between end users, analyst, designers and developers. The knowledge is mainly oriented to the detection of weak signals inside the organization. In this paper, we present the different knowledge surroundings an IS project and a knowledge pattern structure that can be used for the formalization aspects of the established exchange that should be established during the IS project between the different participants**

***Index Terms*—Distributed IS, Project life cycle, Knowledge pattern, Security.**

I. Introduction

The fast evolution of Information System's architectures and of the corresponding technologies has not been followed by an adequate adaptation of SA&D methodologies for the new requirements of these new forms. The commonly used methodologies keep focused on representing the virtual world as close to the real world of the organization as possible. This is done without taking into account the resulting risks taken by using these new architectures. Indeed, the opening of the IS represents an informational danger: the data, process and results security might be violated at any time. That violation can occur during any phase of informational production, information exchange, data collection phase, process execution or the results transmission and can be either internal or external. To cope with these risks, a global security SA&D approach is needed. This approach must cover the entire SA&D process and ensure the continuity of the security policy.

Manuscript received April 10, 2006. This work was supported in part by the Swiss National research fund. This project was supported by the Swiss Federal Government unsder the SER project 03.0391-1 in the frame of the EU Network of Excellence INTEROP : Interoperability Research for Networked Enterprise applications

Dr. SNENE Mehdi is a senior researcher in the university of Geneva. He is leading research on Information systems security. Author is with the Database research lab, cui- 24 rue du general dufour 1204 Geneva Switzerland (phone: 0041 22379 7773; fax: 0041 22379 7780; e-mail: snene@ cui.unige.ch).

PARDELLAS Jorge is a Phd student in the university of Geneva. Author is with the database research lab. Pardellas@cui.unige.ch.

Basically, existing approaches such as intrusion detection, listening detection, exploits scan, etc. are commonly used to make the running IS secure. These tools do not induce a coherent and continuous process that ensures the global potentiality of the IS. Indeed, these tools are used subsequently to the deployment and are considered as a software suite which remains external to the IS. Upstream of the deployment, these tools cannot be considered or integrated within the IS life cycle because they protect the execution platform rather than the IS itself. From this postulate, two major issues remain crucial for the survival of opened IS. The first one is how to make sure that the tools guarantee the global security of the IS without exposing their survival and restricting its functionalities. The second issue concerns the efficiency of these tools applied to answer the security needs of an IS that was not conceived and developed while taking into account the secure dimension of data and process [1].

In commonly used methodologies, the differentiation between procedure and process remains implicit. Both are generally used together; however the distinction is essential in order to define the scope of security. Indeed, the security approach is different whether the procedure or the process is considered. A procedure is defined as a formalism of data treatment. It is in fact a set of steps, means and methods used in the execution of a task in order to achieve a predefined result. Hence, the security policy must be applied at the different levels composing the procedure. We define a process as an arranged succession of operations performed in order to realize a procedure in an automated manner. Existing security solutions provide answers to process security through key exchange, data encryption, etc. [2]. Different studies underline the fact that security problems are generally due to a misunderstanding of the organizational security needs or an inadequate system solution. In fact, implementing a set of security process for an IS that has been designed and developed without taking into consideration the internal security procedures, which should be integrated in it, will leads to a patching maintenance policy. Such security policy is considered as a downstream solution that reacts after vulnerabilities detection or attacks. The presented approach is based on the early detection of security leak at the different IS project life cycle and on the analyze and the integration of user informal security procedures. The main purpose of such approach is to obtain an upstream security policy that enables us to detect major vulnerabilities before IS deployment [2].

Next section presents the Weak Signal concept that we

consider as the cornerstone of our approach. Then we present the different security aspects that must be taken into account while the IS is designed and developed. Finally we adapt the different aspects to the new IS architectures and we define a knowledge pattern structure that will be used for the IS project partners communication establishment.

II. RETHOUGHT IS LIFE CYCLE:

A. Weak Signal

The proposed approach is based on weak signal detection. We define a *weak signal* inside an organization as information or knowledge, which cannot be explicitly expressed during the requirements analysis by actors involved in the creation, diffusion and the use of such information or knowledge. A security weak signal is an implicit information or knowledge existing inside the organization, which can be harmful to the system security if not detected. In fact, such security vulnerability cannot be resolved during the final system life cycle steps because it requires a complete rethinking of the system.

A common case of security weak signals appears during confidential documents exchange automation. Indeed, such exchange in the real world involves different security rules based on human responsibility and vigilance. For example, in the case of a document exchange, actors ensure that copies of the document do not exist and guarantee by their presence that the right person receives the document. The exchange life cycle is closed and any leak is easily identifiable. The adaptation of this human document exchange process to an electronic process (email, ftp, etc.), only the common security rules applying to the data flow are implemented if no more requirements are specified by actors. However these security rules basically address process issues but do not consider procedure issues.

These procedure issues are constituted by the different elements surrounding the exchange, which we identify as security weak signals. For example, existing security frameworks do not guarantee that there is no other copy of the document. In the case of an email exchange, the sender and the recipient keep by default a copy of the document in their email application. This multiplication of copies greatly increases the risk of security break. This is the kind of security weak signals the proposed approach tried to identify and to circumscribe.

B. IS Life Cycle

During the users' needs analysis phase of common design methodologies, the security dimension of IS is still not well specified due to the fact that the organizational security is not considered as an autonomous procedure. Indeed, the security dimension is generally delegated to development and deployment steps where they consider the system security instead of the organizational security, which must take place at analysis and design steps. The rethinking of the different phases composing the life cycle take into consideration the security dimension from the first phase to the last one in a continuous way (see Fig 1.).

The analysis of users' needs must emphasize the risk level linked to each category of data and process. It is also primordial that the users' requirements in term of confidentiality and security and the constraints determine the system runtime. Indeed, user actions implicitly trigger security policy decisions [3], thus the activities of a user and the weak signals surrounding these activities have to be clearly defined. The design phase must take into account the technical platform and the development environment to obtain an adaptable schema that preserves the security specificities relative to the latter. It must also express the users' needs gathered during analysis phase. Finally, it has to express the different security components relative to data process and to the data itself. The development of the IS adapts the security requirements of the platform while respecting the schema specifications. In the deployment level, it's imperative to validate the technical platform along with the system constraints and the processes in order to ensure that the system answers correctly to the security criteria.

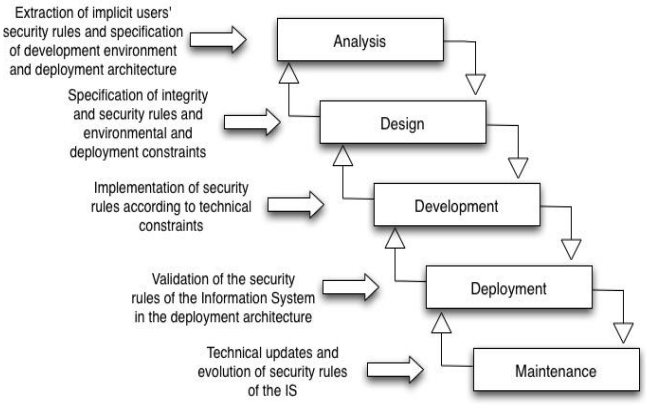


Figure1. Security dimension into Information System Project Life Cycle

The maintenance phase guarantees the global security of the system through the platform updates and the adaptation of the IS and of its security rules to modifications and new needs.

The proposed approach aims to elaborate an IS with a global security context that takes place at the first steps of the design

phase and continuously during the next phases. Indeed, it is primordial to determine the different characteristics of every phase to obtain a global knowledge of the system.

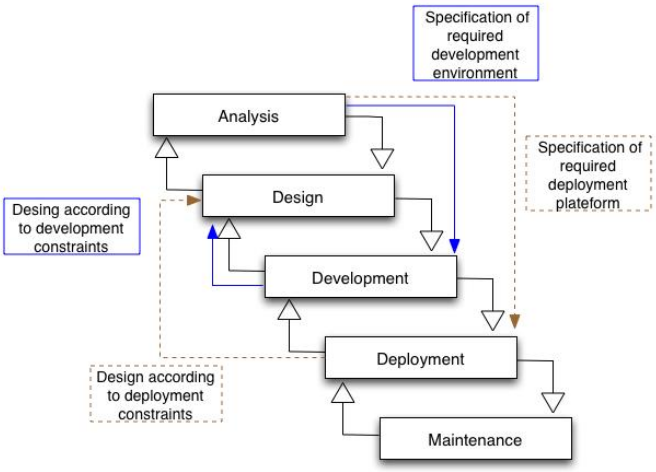
At every level of the life cycle, the responsible actor extracts the inherent local constraints. These constraints represent critical information that must be diffused to the other levels. Indeed, they represent guidelines for the elaboration of each

step [4]. This emphasizes the necessity for other actors to be aware of the global system constraints to adapt to the other phases' needs thus maintaining the good continuity of the project. In order to gather all the required characteristics at the different phases, a strict collaboration between the different actors of the project is needed. Therefore technical and conceptual constraints have to be communicated both upstream and downstream of the IS life cycle (see Fig.2).

Figure 2. Information System Project Life Cycle exchange

III. ADAPTATION TO DIS LIFE CYCLE

Among the different emerging architectures, special attention is brought on the Distributed Information System (DIS). Because of its collaborative nature, this type of systems is potentially more sensitive to threats and requires therefore a particular attention. Due to its physical and logical structure,



the Distributed Information System needs to reach a high security level to ensure the confidentiality of the transiting information. Indeed
different sites, due
encounter informat
underlines the need
with the distributior

The Distributed
distributed data ar
connected with son
commonly known a
services across p
computing environ
applications can b
available. The most
Request Broker Arc
Java Beans) from Su

The existing gap between current information system design methodologies and the distri
forced developers to operate
models to adapt them to the technology specification. Every

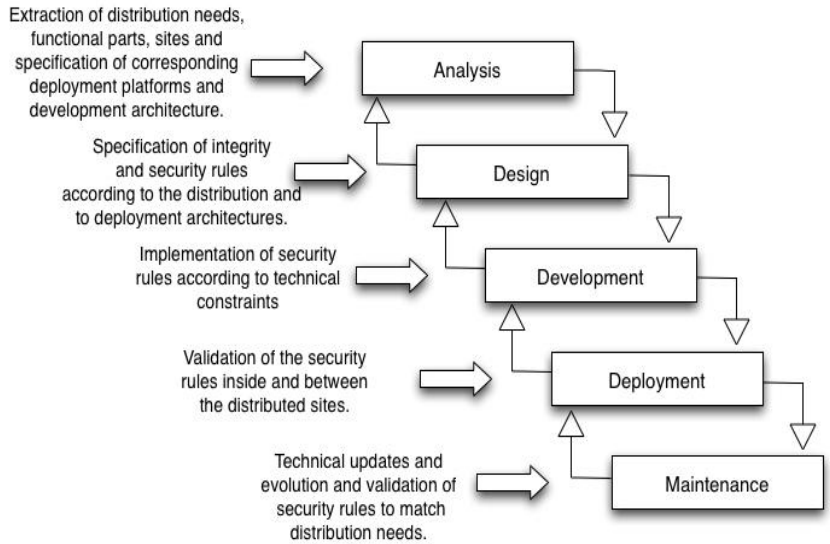


Fig3. Security requirements of a Distributed IS

middleware specification has its own implementation constraints and generally, designers do not have enough knowledge about the implementation process to realize a well-designed system that corresponds to the implementation model without major modifications [6].

In fact, methodologies for the design of information systems have usually paid modest consideration to distribution and communication characteristic of systems. However, this situation has improved: information systems have noticeably grown in dimension and range, organizations want many of their separate systems to be integrated and consequently the number and the variety of geographically distributed users of these systems have become wide. The distributed systems community has produced methodologies for the design of distributed systems [7]. However, these methodologies pay little attention to the information aspects of distributed information system; instead their strength is in the distribution and communication aspects of these systems. The information aspect still not well specified in these methodologies despite its importance for the distributed information system survival [8].

Firstly, in the analysis level, we extract the users' needs that the DIS has to answer. These different needs are categorized into different functional parts that will be distributed among the distribution sites. The subdivision of the system into sub-systems inherently creates data exchange inter or extra-sites. Thus, the DIS must be secured from internal and external threats. Indeed, the subdivision into subsystems generates the need to focus into confidentiality and access restrictions to data flow. The separation into functional parts has to be followed by security rules that ensure confidentiality between these parts.

In the analysis level, the deployment platforms must be
have its own architecture.
ave their own security
g for the global security
owing level. At the design
and the DIS schema are
development environment
nade by the developer, the
their security constraints.
specify the distribution
data flow between the
tical decisions related to
bution are made. These
ty of the system. At the
chema is applied while
ints, using the platforms
g the different sites for
el, it is imperative to verify
that the communication between the sites respects both design
ns and that the security
bution. These conditions
ensure the DIS integrity, correctness and both local and global

security. At the maintenance step, any modification in the DIS has to be followed carefully to ensure the needed continuity of the system hence avoiding any data or security violation. That's why the schema must be adapted and rethought through changes. Moreover, the deployment platform has to be up-to-date in order to prevent any attack through exploits or known bugs (see Fig.3).

IV. KNOWLEDGE PATTERN

We define a set of interaction pattern between the various participants in the life cycle of a distributed information system. These patterns initially treat the exploited knowledge by each one of these participants. Then, these patterns will be the intermediary of communication to establish a form of formal communication between them to allow a better comprehension of the system and thus to ensure a better continuity between the various levels [9]. These patterns are divided into three categories corresponding to three knowledge interaction levels [10]. The first is the designers' knowledge needed by developers to capture the design environment and to well understand the purposed conceptual diagram. For example, in some special case designers need to fix some critical data on some specific sites to ensure its availabilities in case of network shutdown. Developers usually
adapt the purposed conceptual diagram to the technical environment to increase the distributed information system performance and can, if not noticed, move these data from the specified site to another one without informing designers [11]. The second category is the developers' knowledge pattern that regroups the different implementation parameters and the technical environment specificities. This pattern is useful for designers and assists them in adapting their purposed conceptual diagram to the technical environment. For example, some application servers used to develop distributed information system (Tomcat, J2EE, …) have their own security framework. Any security approach used to design the distributed information system must fit into this framework. Finally we purpose a knowledge overlap pattern that encapsulates the common knowledge between the designers and the developers. An example of this knowledge is to exception treatment. The technical behavior of each used technical environment is different from others and its integration in the conceptual diagram must be discussed and approved by both developers and designers.

As explained above, three knowledge categories are distinguishable in a distributed information system project. These zones are called Knowledge zone. In fact, every zone is characterized by its own knowledge that is needed by other partner to successfully achieve the project. To put this knowledge under a formalized form comprehensible by different project partners who do not share the same technical and knowledge language, we define it as knowledge pattern form. These patterns will contain and express the different needed knowledge contributing to ensure a normal continuity for the project. Three different pattern categories are identified.

The first pattern category is the distributed information system design knowledge pattern. This pattern conceptualizes the diverse knowledge used to design the system and to produce it under the form of schema. As this schema is the result of a transformation of different information collected from end users, this transformation, as any other transformation, causes some data loss. Two domains are concerned by this data loss: the conceptual domain and the process domain. The first domain represents the different data that make up the system. They are represented under the form of a design schema. But, some of the collected data cannot be represented under this form due to their particularity. Nevertheless, these data can be important for developers for a better understanding of the given schema. For example, on a database table, designers can specify a user-address attribute under many fields, which are street-number, street, and town. This specification can result from a user need, such statistical use or for a further evolution. At the development step, as these needs are not clearly expressed in the design schema, developer for better performance and to decrease the answering time can unify the different fields in a unique one that represents the same data. In this case, the obtained information system will provide users with the same data then the wanted information system. But at the time when the end users will try to get the statistical information which concern only a part of the global address such as the street, this information will be unavailable and the information system will be unable to satisfy the users' requirements. The second domain is the process domain that represents the different information system functionalities. These functionalities are represented under the form of functions and transactions. At the design step, designers usually do not specify the runtime manner and site needed for each transaction. However, this information can be vital for some critical functions. The term critical does not indicate their system aspect but the organizational one. In fact, if we suppose that the statistical functionalities mentioned in the example above will be used for each new data entry, and if we take the case of an international interim worker company, the location of this functionality runtime is crucial to determine the system answering time and its quality of services.

The second pattern category is the distributed information system development knowledge pattern. This pattern regroups structural and technical information related to the development languages and the runtime platforms. Such information is generally extracted from the different specifications of the used technologies. This information is expressed through the developed system but still hard to get at the design step. In fact, designers usually ignore the technical constraint that can be transgressed by their produced design schema. This obliges the developers to adapt the design schema to their used technologies and by the way to carry out some modifications that can be harmful to the system objectives. The different

knowledge that must be considered by these patterns concern: security, performance and availability.

The security knowledge is of two kinds. The first is the data security framework defined and requested by the technical platform. Indeed, according to the used platform different protocols exist and must be respected at the design step. This special requirement must be expressed at the beginning of the design step by a clear and formal specification. Different forms can be used to express this specification, but the most comprehensible one is to provide the designer with a conceptual schema that must be integrated in the system schema. For example, the Tomcat server obliges the designer to use a unique security framework defined by the Tomcat specification. The security data that are: User, Login, Password and Session must be defined as fields of a unique table called security. Otherwise, even the data are protected by other control checks and constraints; the transition from the login step to the session step cannot be done. In fact, if the requested data are expressed under another form than the one that the platform specifies, they cannot be treated.

The second security knowledge form is the system security Knowledge. In fact, any distributed information system needs to be protected from external or internal attacks. For this purpose, developers use different technologies and define different lockers on the system. This set of technical information must be transmitted to designers by indicating the security constraints that will be applied to the design schema at the development step. For example, using distributed databases must be accompanied by a data flow restrictions. This means that the critical data will be used just on a defined site and cannot be called by non-authorized sites. If designers distribute the different process without taking in account this fact, the security can be transgressed and the critical data are doomed to be in danger.

The third pattern category is the distributed information system knowledge overlap pattern. By overlap we designate the overlapping zone between the design and the development step. This knowledge results from the mixture of different kinds of information kinds related to both domains. It is composed mainly of three different components that are: data location, site relation and integrity constraint.

The data location concerns the distribution manner of the data. Indeed, a part of the data cannot be distributed in a classical way (fragmentation and allocation [12]). Some data is needed on specific site because of its importance to the good running of this site, or due to the end-user request. This information results both, from designers and developers. It implies that the design and implementation constraints have to be resolved jointly. The first pattern component's aim is to regroup these different locations' information in a formal way to facilitate the resolution of their related different restrictions.

The site relation component is the set of different information concerning the different existing relations between the sites. These relations can be of different kinds. The first relation category is the organizational relation. In fact, some needs require special connection between the different organization sites. For example, some distributed information systems with critical sites are completely replicated due to their information importance. In such case, each site must inform the others of its correct management. It has also to maintain its backup site correctly and continuously. The second relation category is technical. Indeed, some technical constraints need some specific relations between the different existing sites. For example, using the J2EE platform requires the establishment of a special permanent connection between the different application containers in order to guarantee their coherences.

The integrity constraint component represents the set of the different sorts of information needed to design and implement these constraints correctly between the different sites of the distributed information system. In this thesis, we have a special focus on this component due to its major importance and impact on the system good running. The different data existing around the integrity constraints are necessary for a good implementation of these constraints. In fact, as these different constraints regroup and result from different system environments such as the organizational, the security and the technical, it is necessary to classify these data in order to obtain the best possible level of system integrity.

V. Conclusion

The set of Knowledge Patterns extracted from a distributed information system project are a group of proven reusable assets that can be used to increase the speed of developing and deploying distributed applications. These patterns have to help and to identify the interaction and processes of selecting and runtime topology. They will provide enterprise developers with a set of guidelines for building information application, including performance, technology options, application design, development and security. These patterns will aim to reduce the existing gap between information designers and the developers by providing them with a unified interaction language. This language will enrich the design step by new concepts, which help developers manage the distribution step while respecting the project goals. They will also provide designers with different information summarizing the technical environment with its constraints. Such information is important due to the modification that has to be done on the design schema to be adapted it to the technical platform. Finally these patterns will define a formal way of communication between the different participants of the project. It will be useful specifically in the overlap domain case.

References

[1] Snene M., "Knowledge patterns of distributed information systems- the case of distribution design and implementation based on integrity constraints optimisation", Ph.D thesis, N576, Geneva university, 2004.

[2] Snene M., Pardellas J., Leonard M., “Information system architectures: where are we?”, Proceedings of the ICTTA Conf, IEEE Press, Syria, 2004.
[3] Snene M., Secure design and implementation of distributed and interoperable IS based on overlap knowledge pattern, IBEC, Hammamet, Tunisia 2005.
[4] Snene M., Leonard M., “Distributed Framework for real time web basedcollaboration:M7TOOL CASE”, Proceedings of AICCSA, IEEE Press, Tunisia, 2003.
[5] A. Ekberg, Enabling technologies for web centric applications, PhD thesis, Lund institute of technology, November 1999.
[6] R. Hirschfeld, Three tiers distributed architecture, Proceedings PloP 96, Allerton Park, IL, 1996.
[7] J.A. Casal, J.A. Garda, R.G Vazquez, S.R. Yarrez, A practical experience in analysis and design of distributed information systems, I+D Computation, Vol.1, No.1, July 2002.
[8] H.G. Sol, R.L. Crosslin, Dynamic modelling of information systems II, North Holland, Amsterdam,1992.
[9] P.I Rivera-Vega, R. Karlapalem, M. Ra, A mixed fragmentation approach for inintial distributed database design, Proceedings of International conference on data engineering, IEEE, 1990.
[10] K. Hui, Knowledge Fusion and Constraint Solving in a Distributed Environment, PhD Thesis, University of Aberdeen, Kings College, Aberdeen, 2000.
[11] R. Varadarajan, P.I. Rivera-Vega, S.B. Navathe, Data redistribution scheduling in fully connected networks, Proceedings of 27th Annual Alberton conference on communication, Control and Computing, 1989.
[12] M.T. Özsu, P. Valduriez, Principles of distributed database systems, Prentice Hall Edt, New Jersey, 1999.