CERIAS Tech Report 2001-59

Information Assurance In Agent–Based Workflow System: An Overview

Thomas Bellocci, Chwee Beng Ang, Parbati Ray, and Shimon Y. Nof
Center for Education and Research in
Information Assurance and Security
&
Department of Industrial Engineering
Purdue University, West Lafayette, IN 47907

Information Assurance In Agent-Based Workflow System: An Overview

Thomas Bellocci, Chwee Beng Ang, Parbati Ray, and Shimon Y. Nof
Center for Education and Research
in Information Assurance and Security
and
Department of Industrial Engineering, Purdue University
West Lafayette, IN 47907-1398 USA

Abstract

This research work addresses the problem of information quality in distributed information systems. A TQM-based definition of information assurance is introduced to fit the needs of inter-networked enterprises that rely on information for the fulfillment of their objectives. Information failures type and impact are investigated in distributed systems like ERP (Enterprise Resource Planning). The design and operation of autonomous agents to implement variable information assurance in workflow systems are studied.

Keywords: Information assurance, ERP, Requirements, TQM, MICSS, Industry survey, Autonomous agents, Protocols

1. INTRODUCTION

1.1. Problem background

Companies' organization and functioning have changed dramatically under the influence of information technology. Information systems have evolved from a centralized to a distributed organization. This change has enabled the development of inter-networked enterprises.

In these companies, information systems do not only support business functions but they are also integral parts of the business operation. For example, ERP systems (Enterprise Resource Planning) are essential for organizations and their supply chains. Companies completely rely on their information system for the execution and coordination of daily business operations.

Critical issues, however, are facing nowadays the enterprise information systems. Companies often have trouble obtaining valuable, timely information and exchanging correct data between different departments of the company. The distribution of information sources, and the high speed of data transfer have increased the vulnerability of companies to information failures and the impact of these failures on the performance of organizations. Incorrect information in ERP systems can have serious consequences for inter-networked companies. These problems become even more critical when a company tries to manage its supply chain. As a consequence, it is necessary to develop a definition of information assurance that fits the actual needs of inter-networked companies, and discover if there exist specific information assurance failures that are critical for the performance of a company.

Another issue is to automate the assurance practices as much as possible, as we cannot expect the workers who interact with the system to include the assurance tasks as part of their job. The objective is to design information systems to automatically apply the assurance function. The problem of how to design and operate agents to assure information in production enterprises is addressed in this article, particularly in the context of ERP systems.

1.2. Research agenda

Following this analysis of today's computing environment in inter-networked enterprises, two main research questions were formulated:

1) How to define information assurance in inter-networked enterprises? What are the requirements for information assurance, and are there critical information assurance failures?
2) How to design and operate agents to automate the information assurance process in workflow systems?

The first research question is addressed in Section 3 of this article, and the second question in Section 4.

2. LITERATURE REVIEW

The literature was reviewed in five directions: 1) Information assurance and security, 2) TQM, quality assurance and information systems, 3) Information assurance survey of requirements, 4) Security and assurance agents, and 5) Autonomous agent systems.

The topics (1) to (3) are extensively described by Bellocci and Nof (2001a), topics (4) and (5) were studied by Bellocci and Nof (2001b). The main conclusions of the literature review are presented below.

2.1. Information assurance and security

The topic of information assurance and security has been originally described mainly by computer scientists. Their approach of information management focuses on information security from internal and external threats (e.g., Longley and Shain, 1986;

Shirey, 1995; Finne, 1997; Voas, 1999). A clear distinction is made between information security and information assurance (e.g., Dobry and Schanken, 1994; Jelen and Williams, 1998). Information security is a feature of the functional components of a product or system, whereas information assurance "is a measure of confidence that the security features and architecture of an automated information system accurately mediate and enforce the security policy" (Longley and Shain, 1986).

Today's networked enterprises are, however, primarily concerned with the quality aspects of their information for the purpose of achieving their performance goals, and with reaching a global improvement in the trustworthiness and value-addition of information.

2.2. TQM, quality assurance and information systems

The approach focusing on information security from internal and external threats, described in Section 2.1., is balanced by an approach emphasizing information's accuracy, value-addition and related features.

Several articles show that companies are now seeking new approaches regarding the administration of distributed information systems. Ford Motor Co. tried to automate the maintenance of its information system (e.g., Schwartz and Zalewski, 1999). British Airways started to think about the value-addition of information in its distributed information system (e.g., Steinitz, 1998).

Some research work has already tried to apply a TQM approach to data management. A description of data characteristics was developed from a user-centered approach by Wang (1998). He explains that any piece of information has the following quality dimensions: intrinsic quality, accessibility, contextual value, and representational value, that must be considered for Total Data Quality Management (TDQM).

2.3. Information assurance survey of requirements

Without clear knowledge of the true needs for information assurance, a company may employ local, specialized solutions that are too restrictive, or not comprehensive.

The challenge is to ascertain what the true assurance requirements are for given industries, in order to develop the most effective means to address the problem.

There are two ways of surveying requirements. On the one hand, people can use literature review and experts' knowledge to analyze their problem and see what the key issues are. On the other hand, people can conduct surveys among information systems' users to understand what the key issues are from a description of the system routine use. Several articles summarizing experts' knowledge in the field of security requirements can be found (e.g., Pfleeger, 1991; Dobry and Schanken, 1994; Shirey, 1995; King, 1997).

2.4. Security and assurance agents

Autonomous agents system is a relatively recent research area. A comprehensive review and definition of agents have been available only recently (e.g., Franklin and Graesser, 1997; Nof, 1999). Also, distinctions between agents are only starting to appear. Security agents have been among the first type of agents to be studied, for instance by Crosbie and Spafford (1995), who implemented an Intrusion Detection System. Varadharajan et al. (1998) described a security agent-based distributed authorization system for ATM machine, where a distinction can be made between "productive" software agents and "assurance" software agents. In this example, the "productive" task, money withdrawal, is still separated from the "assurance" task, namely checking agent's rights. A combination of tasks is not envisaged. In conclusion, past research investigated security and assurance agents. But it never considered a combination of security or assurance tasks with production tasks in the same agent. The issue of tasks combination in agents for assurance purpose is addressed in this article.

2.5. Autonomous agent systems

An autonomous agent system architecture able to supervise processes has been described by Kim (1996), and Kim and Nof (2000). In these papers, the authors introduce the AIMIS (Agent-based Integration Model of Information Systems). This architecture enables the monitoring and execution of processes among distributed organizations with heterogeneous information systems, using agents. This autonomous agent model was used in this research work as a model of agent-based workflow system.

3. INFORMATION ASSURANCE: DEFINITION AND REQUIREMENTS

3.1. TQM-based definition

The literature review showed the need for inter-networked companies to consider both the security and quality aspects of data for managing their operations. As a result, information assurance was defined as the combination of:

1) Information security,
2) Information integrity, and
3) Information significance.

Information security means protecting information from malicious threats and damage due to external or internal sources. *Information integrity* should be understood as permanency of the information during communications and storage. Lastly, *information significance* refers to the value that the intended user can get out of the information when s/he receives it. The broader view considers assurance from the viewpoint of "quality assurance". The broader definition is proposed as follows:

- Information assurance combines the requirements of information security, integrity and significance.
- Assuring information means having a safe information system, which guarantees that information is secure and at the same time keeps its integrity and its significance during its lifetime.
- The goal of information assurance is to provide trustworthy and significant information to users in operational, service systems that rely on the information for the fulfillment of their objectives.

3.2. Information assurance requirements

A preliminary analysis generated a list of all the requirements a company must fulfill if it wants to assure its information according to the TQM-based definition. For each category, a non-exhaustive list of measures was developed and presented by Bellocci and Nof (2001a). These comprehensive requirements were derived from the RACF parameters (Resource Access Control Facility) developed by IBM and presented by Schwartz and Zalewski (1999), and from the TDQM parameters described by Wang (1998).

3.3. MICSS lab experiments

As a step in refining the assurance requirements derived from the literature, and showing the variable needs in information assurance, experiments were conducted with an ERP software simulator called MICSS (Management Interactive Case Study Simulator) [http://www.mbe-simulations.com/, June 2001]. MICSS was used to simulate the functioning of a company with a team-oriented view. There are four views of a company, namely Marketing, Production, Purchasing and Finance. Each of these views can combine to operate the company to be profitable.

In an ERP system, a company enters an operational policy P to ease and automate some of the basic functions of the business, such as production planning. A policy can be described as a k-tuple recording the value d_i of each data item D_i composing the policy; e.g. $P = (d_1, d_2, \ldots, d_k)$. Often the company follows a baseline policy BP, recognized to provide good performance results regarding profits π and due-date-performance DDP.

As a consequence, three typical scenarios regarding information can be encountered in an ERP system. When the company whishes to input its baseline policy BP, the value d_i of a data item D_i can indeed be one of the following:

1. Correct, defined as BP
2. Correct but delayed, defined in this research as $D4(D_i)$ or $D8(D_i)$
3. Wrong, defined as $Wd(D_i)$ or $Wh(D_i)$.

A set of experiments was run with MICSS to simulate failures in information exchange and analyze the potential consequences of failures on the performance of the company. First, a Baseline Policy was built for the company. The values in the purchasing policy, marketing policy, etc… were chosen so that the company performs well. Then, information assurance failures such as Delayed information or Wrong information were introduced in the Baseline Policy to assess the consequences of low-level information assurance on the performance of the company. A complete description of the experimental process is described by Bellocci and Nof (2001a). The detailed statistical analysis is described by Bellocci et al. (2001).

Conclusions from the lab experiments:

1) The experiments show that information failures have significant impact on the performance of a company only under specific conditions.
2) Profit is very sensitive to information failures. Due Date Performance reacts more slowly and is impacted significantly only after long lasting and large errors.
3) The impact of information failure depends greatly on the Data item that is concerned by the failure. For instance, the consequences of a problem concerning Price are usually much more serious and long lasting than when the error concerns Batch Size.
4) Data items have different characteristics that make them more sensitive to a specific type of failure. For instance, a delay of 8 months has a large impact on Profit when it concerns Price, but no real impact when it concerns Batch Size.
5) A difference in the length of delay influences the performance of the company only when Price is concerned by the error.
6) The impact of an information failure depends on the error size, except when the error concerns Order Level.
7) The impact of each type of information failure on each data item is presented in Figure 3 using the maximum relative difference to the BP's performance reached in a one-year period. The specificities of each Data item are summarized in Table 1. It appears that each Data item has strong particularities regarding the performance metrics of the company that it affects, and the type of failures that it is most affected by.

Table 1. Summary of information failures' impact on Data items

Error in Data item	Impact on	Critical Failures	Length of delay	Error size
Price	Profit	Wd (positive impact) Wh, D8, D4	Important	Important
QLT	Profit	Wd	Not important	Important
	DDP	Wd	Not important	Important
Batch Size	DDP	Wh	Not important	Important
Order Level	None	None	Not important	Not important

3.4. Industry survey

On the basis of the lab experiments, an industry survey was designed to assess the information assurance requirements of the corporate world. Two questionnaires were developed. One was sent to the information system manager of a given company, and the other one to the department managers of the same company (e.g., production manager, marketing manager…). The objective of the first survey questionnaire was to understand the general approach of companies regarding information security and assurance. The second survey questionnaire was designed to study the actual information assurance problems encountered by users of the company's information system.

The design and conclusions of the industry survey are presented by Ray, et al. (2001a). The questionnaires, and the detailed analysis of companies' answers are available in Ray, et al. (2001b). The questionnaires were sent to approximately 50 companies in the United States, Europe and Asia. The analysis was based on the 9 questionnaires returned by information system managers, and the 10 questionnaires returned by department managers.

Conclusions from the industry survey:

1) Companies think that information assurance failures have significant impact on their performance.

2) From the analysis of the survey, it is inferred that companies are more concerned by information significance than information security or integrity in their information systems.
3) Further analysis showed that indeed Profit and Due Date Performance (the reputation of the company) are the parameters that are the most affected by information assurance failures.
4) At present, System Authorizations, Firewalls and Antivirus are the most popular preventive measures that companies apply. This observation shows that companies are equipped to handle information security and integrity problems, but not yet to handle information significance problems.
5) Companies introduce flexibility in their information systems mainly using user groups having access to different resources using passwords.
6) The process of assuring the data is too time-consuming for information system users. In a decision-making process, information system users spend more time on acquiring the necessary information, and arguing about its accuracy than using the data.
7) Most companies disregard the information due to the fact that they maybe from unreliable sources or the information may be inaccurate.
8) If information is missing at the time of changing the policy in their ERP system, most of the companies can wait, but not for very long. If they have to wait for longer, they go ahead and change their strategies.
9) Users have difficulties to change processes because of consequential damages due to tight integration.

3.5. Conclusions

The conclusions of the lab experiments and industry survey were presented earlier. The general implications of these results for the research project are explained below:

1) The experiments showed differences in the impact of information failures between Data items. Thus, it is important to adjust the assurance tasks to the needs thanks to

variable assurance. Each company has to investigate its own specificities to define what are the most critical information assurance failures for its activity.

2) The survey showed that companies have difficulties keeping the consistency and significance of their information. It also proved that decision-makers are willing and are able to wait to get better assured information.

3) Some examples of assurance tasks that should be automated to improve decision-making processes in a company can be formulated from the literature review, the lab experiments, and the industry survey:

   a. *Dealing with information security:*

      - access authorization
      - intrusion detection
      - virus detection
      - messages encryption
      - users and data profiles management
      - critical data monitoring and history recording

   b. *Dealing with information integrity:*

      - regular back-ups
      - data decay prevention
      - communication links quality monitoring
      - communications success monitoring (termination + mapping)
      - data safety when system crashes
      - secure restarting after system crashes

   c. *Dealing with information significance:*

      - believability check (the data stored is not obviously wrong)
      - completeness check (all of the needed characteristics of a data are stored)
      - accuracy check (the value of the data is given with the required accuracy level)
      - source trustworthiness (where does the data comes from?)
      - timeliness (the data must be ready when needed)
      - representation (the data must be displayed using the correct representation)

The lab experiments showed the large variability in the impact of information failures, depending on the failure type and the data item concerned by the failure. As a consequence, variable information assurance should be introduced in information systems. The industry survey demonstrated that information significance is the true concern of information systems users in inter-networked companies. It also helped developing a list of examples of assurance tasks that should be automated in distributed information systems. The next step of this research work is to design and evaluate agent models and variable assurance protocols to automate these assurance tasks.

4. AGENTS AND PROTOCOLS FOR INFORMATION ASSURANCE

In an ERP system, autonomous agents can perform production-related tasks or assurance-related tasks. In the frame of this research, two dimensions of autonomous agent systems are investigated:

a. the conditional execution of assurance tasks, and
b. the agents used to perform the tasks.

The first problem is referred to as “variable assurance” problem, and the second one is called “task combination” problem. The first section of this chapter describes the justification for variable assurance and presents the basis for its implementation. The second section focuses on task combination in agents, and introduces models for assurance in autonomous agent systems.

4.1. Variable assurance

The MICSS lab experiments and the Industry survey demonstrated the differences of impact of information failures on transactions, and motivated the need to adjust the assurance tasks to each requests. The implementation of variable assurance requires two steps: 1) Evaluating the importance of performing assurance tasks for a given transaction, and 2) Deciding if the assurance tasks should be performed according to this importance level.

4.1.1. Risk assessment

The decision of whether or not to perform assurance tasks for a given production request needs to be supported by two separated information gathering activities:

a. The Request Analysis
b. The Context Analysis.

The purpose of the Request analysis is to gather the request's characteristics to tailor an assurance process to he request needs, based on the analysis of the critical information assurance failures showed by Bellocci and Nof (2001a). The purpose of the Context Analysis is to gather information about the system to adjust the assurance processes to the status of the system. A risk assessment model was developed by Bellocci and Nof (2001b) to enable the implementation of variable information assurance in agent-based workflow systems, based on the AIMIS model described by Kim and Nof (2000).

4.1.2. Variable assurance protocols

Three different variable assurance protocols were designed to support the implementation of the variable assurance model. The results of the MICSS lab experiments showed the importance to consider two different request characteristics:

(1) The assurance needs of the request, that can be assimilated to the need of assurance features (trustworthiness, completeness, integrity…) for instance due to the information sender location, or receiver identity, and
(2) The priority of the request, that corresponds with the need to receive the information on time.

Based on this conclusion, three different Variable Assurance Protocols with different logic were designed:

a. VAP0 assures all the requests,
b. VAP1 assures requests based on their assurance needs, and
c. VAP2 assures requests based on their assurance needs and priority level,

These protocols are described by Bellocci and Nof (2001b). They are modeled and analyzed later, in Sections 4.3 and 4.4.

4.2. Assurance models

In the frame of this research, two categories of agents were considered: 1) Dedicated agents, A_D, and 2) Polyvalent agents, A_{AP}. There are two types of dedicated agents: 1) Assurance dedicated agents, A_A, that can only perform Assurance Tasks, T_A, and 2) Production dedicated agents, A_P, that can only perform Production Tasks, T_P. A polyvalent agent, A_{AP}, is able to execute both a production task and the associated assurance task. The execution of these two tasks in a row by the same agent is referred to as an assurance-production task, and noted T_{AP}. A polyvalent agent can also execute single production tasks, T_P, by skipping the assurance part of its code.

Following these observations, three Assurance Models were proposed depending on the agents available to execute the Assurance and Production Tasks, namely: 1) The Separated Model, M_{Sep}, 2) the Combined Model, M_{Com}, and 3) the Mixed Model, M_{Mix}. The nature of the agents involved in each model is summarized in Table 2.

Table 2. Summary of Assurance Models

	Type of agents involved	
Model	**A_D**	**A_{AP}**
M_{Sep}	Yes	No
M_{Com}	No	Yes
M_{Mix}	Yes	Yes

4.3. Design of Experiment

A simulation model was built and described by Bellocci and Nof (2001b) to analyze the impact of the assurance model and variable assurance protocol in implementing variable information assurance in an agent-based workflow system.

Two metrics were used to assess the performance of an autonomous agent system S:

(1) The processing time of requests, θ (S)

(2) The assurance exit level of requests, η(S),

Four characteristic parameters of an agent-based workflow system were investigated:

(1) Variable assurance protocol, symbolized V, with three levels:

 a. V0 = VAP0 (Total assurance)

 b. V1 = VAP1 (Needs-based assurance)

 c. V2 = VAP2 (Needs- and priority-based assurance)

(2) Assurance model, symbolized M, with three levels:

 a. M1 = M_{Sep} (Separated model)

 b. M2 = M_{Com} (Combined model)

 c. M3 = M_{Mix} (Mixed model)

(3) Assurance policy level, symbolized L, with three levels:

 a. L1 = 300 A.U. (Low requirements)

 b. L2 = 500 A.U. (Medium requirements)

 c. L3 = 700 A.U. (High requirements)

(4) Total number of agents, symbolized N, with three levels:

 a. N1 = 10 agents (Low quantity)

 b. N2 = 15 agents (Medium quantity)

 c. N3 = 20 agents (High quantity)

Based on this design of experiment, 81 different treatments were simulated. For each treatment, two simulation runs were executed with different random numbers. During a run, the processing time and exit assurance level of the first 500 executed requests were recorded. The stationary state is reached after 20 to 50 requests, depending on the treatment. As the transient regime ends relatively quickly, the first requests were kept in the pool of 500 requests used for the analysis.

The following experimental research questions were investigated:

1) Experimental Question 1:

 What are the significant parameters for the processing time of requests?

2) Experimental Question 2:

What are the significant parameters for the exit assurance level of requests?

3) Experimental Question 3:

What is the best variable assurance protocol overall?

4) Experimental Question 4:

What is the best assurance model overall?

5) Experimental Question 5:

What is the best combination of Variable Assurance Protocol and Assurance Model given an assurance policy level and a number of agents?

4.4. Results

The answers to the experimental research questions are presented below. Detailed answers and justifications are presented by Bellocci (2001).

1) What are the significant parameters for the processing time of requests?

According to the ANOVA results, all of the four parameters V, M, L, N and their interactions are significant with a confidence level of 95%. Hence:

$$\text{(E1)} \qquad \theta(S) = f(V,\ M,\ V{*}M,\ L,\ V{*}L,\ M{*}L,\ V{*}M{*}L,\ N\ ,V{*}N,\ M{*}N,\ V{*}M{*}N,\ L{*}N,\ V{*}L{*}N,\ M{*}L{*}N,\ V{*}M{*}L{*}N)$$

As a conclusion, Variable Assurance Protocols and Assurance Models have a significant impact on the processing time of requests.

2) What are the significant parameters for the exit assurance level of requests?

According to the ANOVA results, only some of the parameters have significant impact on the exit assurance level of the request with a confidence level of 95%. In fact:

$$\text{(E2)} \qquad \eta(S) = f(V, M, V{*}M, L, V{*}L, M{*}L, V{*}M{*}L)$$

As a conclusion, Variable Assurance Protocols and Assurance Models have a significant impact on the exit assurance level of requests.

3) What is the best variable assurance protocol overall?

Decision-makers relying on information to complete their tasks are particularly interested in the proportion $\tau(S)$ of trusted requests that exit the system S. A request is called "trusted" if its exit assurance level is higher than the assurance policy level of the company. A Student-Newman-Keuls range test was used to rank the variable assurance protocols with a confidence level of 95%. Two protocols provide interesting results. Both VAP1 and VAP2 offer a reduction of processing time compared to VAP0. The needs-based protocol VAP1 does not decrease the proportion of trusted requests. The needs- and priority based protocol VAP2 allows a larger reduction of processing time than VAP1, but also implies a diminution of the proportion of trusted requests.

4) What is the best Assurance Model overall?

A Student-Newman-Keuls range test was used to rank the Assurance Models with a confidence level of 95%. The results show that the Combined assurance model M_{Com} performs better than the Separated and Mixed models. It is the fastest model in requests processing, and provides the largest proportion of trusted requests.

5) What is the best combination of Variable Assurance Protocol and Assurance Model given a combination of Assurance Policy Level and Number of Agents?

Selecting a combination of Variable Assurance protocol and Assurance Model is a trade-off between low processing time of requests, and high proportion of trusted request. Two treatments can compete with the Total Assurance protocol: VAP1*M_{Com} and VAP2*M_{Com}. Their performances compared to Total Assurance are summarized in Tables 3 and 4, regarding the mean processing time of requests, $\theta(S)$, and the proportion of trusted requests, $\tau(S)$.

Table 3. Comparison of VAP1*M_{Com} performance to Total Assurance

VAP1*M_{Com} versus Best VAP0		**Assurance Policy Level**		
		300 A.U. (low requirements)	**500 A.U.** (medium requirements)	**700 A.U.** (high requirements)
Total Number of	**10** (low)	$\theta(S)$: -50% $\tau(S)$: -0%	$\theta(S)$: -60% $\tau(S)$: -0%	$\theta(S)$: -34% $\tau(S)$: -0%

Agents	**15** (medium)	$\theta(S)$: -72% $\tau(S)$: -0%	$\theta(S)$: -45% $\tau(S)$: -0%	$\theta(S)$: -19% $\tau(S)$: -0%
	20 (large)	$\theta(S)$: -70% $\tau(S)$: -0%	$\theta(S)$: -44% $\tau(S)$: -0%	$\theta(S)$: -20% $\tau(S)$: -0%

Table 4. Comparison of VAP2*M_{Com} performance to Total Assurance

VAP1*M_{Com} versus Best VAP0		**Assurance Policy Level**		
		300 A.U. (low requirements)	**500 A.U.** (medium requirements)	**700 A.U.** (high requirements)
Total Number of Agents	**10** (low)	$\theta(S)$: -69% $\tau(S)$: -16%	$\theta(S)$: -79% $\tau(S)$: -38%	$\theta(S)$: -55% $\tau(S)$: -65%
	15 (medium)	$\theta(S)$: -82% $\tau(S)$: -15%	$\theta(S)$: -72% $\tau(S)$: -43%	$\theta(S)$: -46% $\tau(S)$: -59%
	20 (large)	$\theta(S)$: -80% $\tau(S)$: -15%	$\theta(S)$: -71% $\tau(S)$: -37%	$\theta(S)$: -44% $\tau(S)$: -64%

Both VAP1*M_{Com} and VAP2*M_{Com} are interesting alternatives to Total assurance. VAP2, however, implies a diminution of the proportion of trusted requests. When the company's assurance requirements are low (i.e., L = 300 A.U.), the reduction of $\tau(S)$ is limited. The conclusions about which combination of Variable Assurance Protocol and Assurance Model to choose for a given combination of L and N are summarized in Table 5.

Table 5. Best combination of Variable Assurance and Assurance Model depending on the Assurance Policy Level, and Number of Agents

		Assurance Policy Level		
		300 A.U. (low requirements)	**500 A.U.** (medium requirements)	**700 A.U.** (high requirements)
Total Number of Agents	**10** (low)	VAP2*M_{Com} (for time) VAP1* M_{Com} (for assurance)	VAP1* M_{Com}	VAP1* M_{Com}
	15 (medium)	VAP2*M_{Com} (for time) VAP1* M_{Com} (for assurance)	VAP1* M_{Com}	VAP1* M_{Com}
	20 (large)	VAP2*M_{Com} (for time) VAP1* M_{Com} (for assurance)	VAP1* M_{Com}	VAP1* M_{Com}

4.5. Validation of experiment

The simulation experiments need to be compared to known results to be validated. The influence of the parameters has been investigated independently. It appears that the processing time increases when the number of agents decreases. When the assurance policy level increases, the processing time increases because the number of assurance tasks to be performed increases. These simple observations validate the correct behavior of the simulation model from a practical point of view.

The insights coming from the analysis of the survey can be used to validate our experiments from the corporate viewpoint. For instance, managers explain in the survey that no company ever reaches a proportion of trusted requests equal to 100%. Also, when the assurance policy level of the company increases fewer requests can meet the requirements, and the proportion of trusted requests decrease. These observations validate the behavior of the simulation model.

4.6. Conclusions

The analysis of the experiments showed that flexibility can be introduced in assurance tasks execution without reducing the confidence level of data. The Total assurance protocol VAP0 provides the best exit assurance level of requests, but can overshoot the assurance level required by the company's assurance policy. The requests executed using protocol VAP1 (needs-based assurance) exit the system with a significantly smaller assurance level than with VAP0, at a confidence level of 95%. Nevertheless, the proportion of trusted requests is similar with VAP1 and VAP0, and the processing time with VAP1 is significantly smaller than with VAP0, at a confidence level of 95%. As a consequence, flexibility in execution of assurance tasks can be introduced in agent-based workflow system using protocol VAP1, which allows the system to reach similar confidence level to total assurance and save significant processing time.

The results of the experiments showed that in the case where assurance tasks are serialized with production tasks, the best assurance model is the one involving only polyvalent agents, M_{Com}. Compared to M_{Sep} and M_{Mix}, this model reaches indeed the smallest processing time and highest proportion of trusted requests for any assurance policy or number of agents.

The best combination of variable assurance protocol and assurance model depends essentially on the company's assurance policy level. When the requirements are medium or high, VAP1 combined with M_{Com} allows a significant reduction of the processing time compared to Total Assurance without reducing the proportion of trusted requests. When the requirements are low, companies can decide between using VAP1* M_{Com}, that reduces the processing time without decreasing the confidence level, and VAP2*M_{Com}, that implies a larger processing time reduction than VAP1 with however a decrease of 15% in the proportion of trusted requests. In this case, information system managers have to decide what is the best trade-off for the functioning of the company.

5. SUMMARY AND DISCUSSION

The development of inter-networked enterprises created a new computing environment in which information assurance is critical. This article investigated the information assurance needs of today's companies. For this purpose, the literature dealing with information assurance was reviewed. A new definition of information assurance was also introduced following the TQM approach to better fit the needs of inter-networked enterprises. A list of requirements for information assurance was developed. The critical aspects of information assurance failures in ERP systems were also investigated using MICSS lab experiments and an industry survey.

The design and operation of autonomous agents to assure information in ERP systems of inter-networked enterprises were investigated. A variable information assurance implementation model was proposed based on the AIMIS model, and a risk assessment procedure was applied. The protocols and models needed to support variable assurance were introduced and their performance is assessed. Experimentation showed the possibility to reduce the processing time of requests without decreasing the proportion of trusted requests, compared to a systematic total assurance approach. They also demonstrated that in the case where assurance tasks are serialized with production tasks, the best assurance model is the one involving only polyvalent agents.

The following directions can be recommended for future research:

(1) It has been assumed in the simulation models that the assurance policy level was fixed over time. The influence of assurance policy level variation over time could be investigated.

(2) In this research work, it has been assumed that the entry assurance level of the requests could be modeled by a normal distribution. A possible direction for future research would be to study if another law could fit better the actual distribution of requests' entry assurance level, and assess the performance of Variable Assurance Protocols and Assurance Models with this new distribution.

(3) The simulation models focused on the sequence of optional assurance tasks followed by production tasks. In such a case, the Combined assurance model appeared to be the most advantageous. The development of assurance protocols to distinguish between the processing of requests that need parallel assurance tasks or the ones that need serial assurance tasks could be investigated.

(4) The variable assurance approach presented in this research work showed that significant resources can be saved by adjusting the assurance tasks to the request and the context. However, additional resources can be saved if the assurance tasks that are performed on concurrent requests are taken into account, as they increase indirectly the assurance level of the given request. Negotiation-based variable assurance protocols could be investigated to solve this research problem.

6. ACKNOWLEDGMENTS

This research was supported by the CERIAS (Center for Education and Research in Information Assurance and Security) at Purdue University.

REFERENCES

Bellocci, T., Planning Variable Information Assurance in Agent-Based Workflow Systems, *Master's Thesis, School of Industrial Engineering Purdue University*, December 2001.

Bellocci, T., and Nof, S.Y., Context of Information Assurance in Inter-Networked Enterprises, *CERIAS Technical Report 2001-57, Research Memorandum No. 01-18, School of Industrial Engineering Purdue University*, December 2001a.

Bellocci, T., and Nof, S.Y., Agents and Protocols for Variable Information Assurance in Workflow Systems, *CERIAS Technical Report 2001-58, Research Memorandum No. 01-19, School of Industrial Engineering Purdue University*, December 2001b.

Bellocci, T., Ray, P., and Nof, S.Y., Information Assurance in Networked Enterprises: MICSS Lab Experiments, Results and Analysis, *CERIAS Technical Report 2001-35, Research Memorandum No. 01-06, School of Industrial Engineering, Purdue University*, January 2001.

Crosbie, M., and Spafford, E., Active Defense of a Computer System using Autonomous Agents, *COAST Technical Report 1995-008, Purdue University*, 1995.

Dobry, R., and Schanken, M., Security Concerns for Distributed Systems, *Annual Computer Security Applications Conference*, 1994, 12-20.

Finne, T., What are the Information Security Risks in Decision Support Systems and Data Warehousing, *Computers & Security*, v 16, n 3, 1997, 197-204.

Franklin, S., and Graesser, A., Is it an Agent, or just a Program? A Taxonomy for Autonomous Agents, *Intelligent Agents III, Muller, J.P., Wooldridge, M.J., and Jennings, N.R., (ed.)*, 1997, 21-35.

Jelen, G., and Williams, J., A Practical Approach to Measuring Assurance, *14th Annual Computer Security Applications Conference*, Phoenix, AZ, Dec 1998.

Kim, C.O., DAF-Net and Multi-Agent Based Integration Approach for Heterogeneous CIM Information Systems, *Purdue University, School of Industrial Engineering, Ph.D. Dissertation*, August 1996.

Kim, C.O., and Nof, S.Y., Investigation of PVM for the Emulation and Simulation of a Distributed CIM Workflow System, *International Journal of Computer Integrated Manufacturing*, v 13, n 5, 2000, 401-409.

King, C., Intranet Applications Security Checklist, *Computer Security Journal*, v 13, n 1, 1997, 47-54.

Longley, D., and Shain, M., Data & Computer Security – Dictionary of Standards Concepts and Terms, *Stockton Press*, 1986.

MICSS (Management Interactive Case Study Simulator) [http://www.mbe-simulations.com, June 2001].

Nof, S.Y., Intelligent, Collaborative Agents, *Research Memorandum No. 99-08*, *School of Industrial Engineering, Purdue University*, March 1999.

Pfleeger, S.L., A Framework for Security Requirements, *Computers & Security*, v 10, 1991, 515-523.

Ray, P., Bellocci, T., and Nof, S.Y., Information Assurance in Networked Enterprises: MICSS Class Experiments and Industry Survey Conclusions, *CERIAS Technical Report 2001-37*, *Research Memorandum No. 01-08*, *School of Industrial Engineering Purdue University*, June 2001a.

Ray, P., Bellocci, T., and Nof, S.Y., Information Assurance in Networked Enterprises: MICSS Class Experiments and Industry Survey Analysis, *CERIAS Technical Report 2001-38*, *Research Memorandum No. 01-09*, *School of Industrial Engineering Purdue University*, June 2001b.

Schwartz, A.P., and Zalewski, M.A., Assuring Data Security Integrity at Ford Motor Company*, Information Systems Security*, 1999, 18-26.

Shirey, R., Security Requirements for Network Management Data, *Computer Standards & Interfaces*, v 17 n 4, September 1995, 321-331.

Steinitz, D., Information Security Management at British Airways: Implementing a Strategic Security Program, *15th World Conference on Computer Security*, November 1998.

Varadharajan, V., Kumar, N., and Mu, Y., Security Agent Based Distributed Authorization: An Approach, *21st NISSC Proceedings*, Crystal City, VA, October 1998.

Voas, J., Protecting Against What? The Achilles Heel of Information Assurance, *IEEE Software*, January 1999, 28-29.

Wang, R.Y., Total Data Quality Management, *Communication of the ACM*, v 41, n 2, February 1998.