CERIAS Tech Report 2001-34

# Information Assurance in Networked Enterprises: Definition, Requirements, And Experimental Results

**Thomas Bellocci, Chwee Beng Ang,**
**Parbati Ray, Shimon Y. Nof**
Center for Education and Research in
Information Assurance and Security
&
School of Industrial Engineering, No. 01-05
Purdue University West Lafayette, IN 47907

# Information Assurance in Networked Enterprises: Definition, Requirements, and Lab Experiments[1]

Thomas Bellocci, Chwee Beng Ang, Parbati Ray, and Shimon Y. Nof

## ABSTRACT

With the dramatic growth of information exchanges within and between organizations, major concerns emerge about the assurance of information. Without clear knowledge of the true needs for information assurance, a company may employ local, specialized solutions that are too restrictive, or not comprehensive. On the other hand, cost-effective, variable integrity and variable security may be economically justifiable and adequate for certain situations and decisions.

Therefore, a new definition of information assurance has been developed following the TQM approach. It describes assurance as a combination of information security, integrity, and significance.

The requirements of information assurance are presented and have been justified on the basis of concrete results obtained from the lab experiments that were conducted. The experiments and results have been briefly discussed in this paper.

[1] This work was supported by sponsors of the Center for Education and Research in Information Assurance, Purdue University

**TABLE OF CONTENTS**

## 1. Introduction:

Companies are becoming increasingly dependent on their information systems. They have new requirements regarding the trustworthiness and value of their information. Therefore, it is of significant importance to develop a new approach of assuring information, not only based on security as defined by computer scientists, but also by considering the integrity, relevance, and other aspects of the quality of information displayed to the users.

In today's companies, information systems not only support business functions but are also an integral part of business operations. For example, ERP systems (Enterprise Resource Planning) are now essential for organizations and their supply chains. Incorrect information in ERP systems can have serious consequences for the inter-networked companies. [See "Experimental Results"]

In this computing environment, having a secure information system is no longer sufficient. Companies are now seeking new approaches regarding the administration of distributed information systems [1--3]. At the same time, workers need more and repeated training to operate with increasingly complex information systems; they look upon security practices as a factor in slowing them down in performing their jobs. Hence, it is necessary to automate the required assurance practices as much as possible, and to expect the information system to apply them, not the workers who interact with the system as part of their job. In other words, information assurance tasks must be handled in the background, in parallel with the users working with the system's information. The challenge is to ascertain what the true assurance requirements are for given industries, and to develop the most effective means to address these requirements.

It appears that companies can no longer be content with what traditionally has been defined as information security. A broader view of information assurance is hence needed, and a global improvement in the trustworthiness and value-addition of information must be achieved. The approach taken by this research project involves surveying the assurance requirements and developing active protocols and autonomous agents to assure information in networked enterprises, as an extension to our previous research in this direction [4, 5]. Our purpose in this paper is to explain the approach of information assurance that we have developed so far from the viewpoint of industrial engineering and information management, and justify this approach.

## 2. Literature Review:

In the literature dealing with information management, different approaches can be found. On one hand, an approach emphasizing information's accuracy, value-addition and related features [1], and on the other hand, an approach focusing on information security from internal and external threats [6--10]. The topic of information assurance has been previously defined by mainly computer scientists. According to the literature, the following definitions may be found.

### 2.1. Network Security:

Network security management is defined [9] as "supporting security policies by monitoring and controlling security services and mechanisms, distributing security information, and reporting security events." The functions associated with network security management are: controlling access to resources, retrieving and archiving security information, and managing and controlling the encryption process. It is also explained that security requires: confidentiality, integrity, authentication, access control, non-repudiation and availability.

### 2.2. Security Requirements in Distributed Systems:

Security requirements in distributed systems [10] include: identification and authentication, trusted recovery, security management, trusted path, access control, audit, availability, cryptography, data confidentiality, and data integrity.

### 2.3. Automated Information System Security:

According to [11], automated information system security implies "the totality of security safeguards needed to provide an acceptable level of protection for the system and for data handled by it."

### 2.4. Assurance:

Assurance in computer security, according to [11], is a "measure of confidence that the security features and architecture of an automated information system accurately mediate and enforce the security policy." Furthermore, if the security features of this system are relied upon to handle sensitive information and restrict user access, the features must be tested to ensure that the security policy is enforced during operation. A slightly different definition is provided in [12]: "Assurance is a measure of confidence in the accuracy of a risk or security measurement".

## 3. TQM Approach of Information Assurance:

### 3.1. Definition:

As mentioned above, companies require more than information security. Wang [1] pointed out the need for companies to have information that has intrinsic, access, contextual, and representational dimensions by applying Total Quality Management to data. In our opinion, Wang's useful work can be combined with further consideration of security aspects. When information systems become the spinal cord of modern companies, these companies must have a reliable system that provides secure and useful information, and these systems have to manage security and assurance problems by themselves.

Based on our initial work, we have concluded that an information system is worthwhile for companies if it can ensure that its information is secure, keeps its integrity, and maintain its significant value for users. Therefore, we define information assurance (Figure 1) as the combination of:

1) Information security
2) Information integrity
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.
More details about these definitions are included in Table 1.

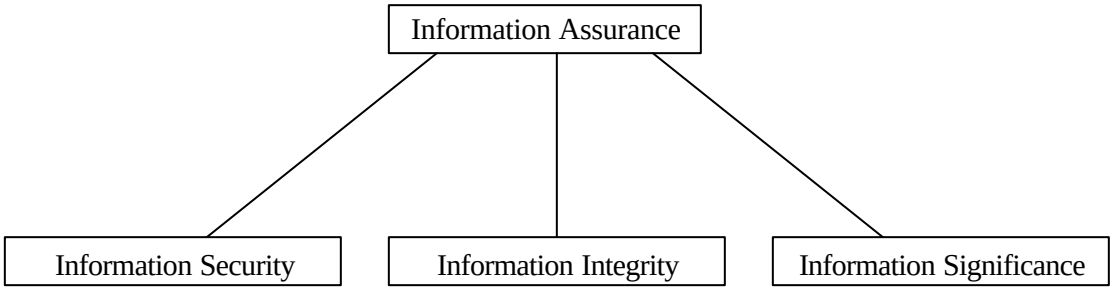



Figure 1. Broad view of Information Assurance

The definition commonly used in computer sciences, as defined above (by [11,12]) does not fit the broader view of our information assurance problem. The broader view considers assurance from the viewpoint of "quality assurance". Therefore, the definition that we will use is 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. Requirements:

Our preliminary analysis has generated a list of all the requirements that a company must fulfill if it wants to assure its information (Table 1). For each category, a non-exhaustive list of measures is shown that can guarantee that the category is fulfilled. Currently, the examples (in italics) are technical issues that may change with the state of the art.

One must understand that Table 1 points out comprehensive requirements (non italic) that must be fulfilled to assure information in networked enterprises. It can be noticed that some of the requirements specified for information security have been previously described in [2] regarding the RACF parameters.

Table 1 – Requirements to assure information

| **Information Assurance** | | |
|---|---|---|
| **Information Security requires:** | **Information Integrity requires:** | **Information Significance requires:** |
| **Protection against external threats:**<br>• *Anti-virus, hacking watch*<br>• *Firewalls*<br>• *Encryption, personalization*<br>• *System authorizations (login + password)*<br><br>**Access profiles management:**<br>Profiles and attributes definition:<br>• *Users groups*<br>• *Class authorizations*<br>• *Attribute of groups*<br>Profiles and attribute maintenance:<br>• *No user with non-standard password intervals*<br>• *No userids that have never been used, or inactive users*<br><br>**Data logging:**<br>• *Global access table entries*<br>• *Started task table entries*<br>• *Class descriptor table entries*<br>• *Dataset name table entries*<br>• *Range table entries*<br>• *Inbuilt audit trails*<br><br>**Data management:**<br>• *Definition of sensitive dataset profiles*<br>• *Definition of general resources profiles* | **Data integrity:**<br>• *Preventing data decay*<br>• *Preventing accidental loss of data*<br>• *Updating and maintenance*<br><br>**Communications integrity:**<br>• *Assuring quality of communications links*<br>• *Recovering from transmission failures*<br>• *Ensuring that the data of receiver and sender map correctly*<br><br>**System recovery:**<br>• *Restarting the system after it crashes*<br>• *Reverting to stable state after system interruption* | **Intrinsic value of information:**<br>• *Accuracy*<br>• *Objectivity*<br>• *Believability*<br><br>**Contextual value of information:**<br>• *Relevancy*<br>• *Value-added*<br>• *Timeliness*<br>• *Completeness*<br>• *Correct amount of data*<br><br>**Representational value of information:**<br>• *Interpretability*<br>• *Ease of understanding*<br>• *Concise representation*<br>• *Consistent representation* |

## 4. Experimental Results:

Following our list of requirements, it was concluded that there are three possible situations for communications in an ERP system. The user can either get correct information, correct but delayed, or wrong information.
This theoretical analysis has been supported by some experiments.

As a step in refining the assurance requirements survey and showing the variable needs in information assurance, experiments have been conducted with an ERP software simulator-trainer called MICSS (Management Interactive Case Study Simulator) [16]. MICSS was developed to simulate the functioning of a company with a team-oriented view.
A set of experiments, using this software, was conducted to simulate failures in information exchange and the potential consequences of subsequent failures on the company.

### 4.1. Problem:

It has been discovered that we can encounter 3 scenarios regarding information in an ERP system. A data item can indeed be correct, correct but delayed, or wrong.
Hence, we decided to study the influence of the following parameters on these scenarios: dataset (type of data affected by information failure), length of delay and error size (difference between the correct data and the wrong value).

First a class experiment involving the undergraduate students of course IE332 was conducted. This provided us with a large amount of data that was analyzed [17]. The measures were not fully reliable to carry out a deep statistical analysis. Nevertheless, it showed interesting trends that encouraged us in organizing our own experiment, where we could master all the parameters. The results of our team lab experiment are presented in the following paragraphs.

### 4.2. Methodology:

MICSS

MICSS (Management Interactive Case Study Simulator) [16] is an ERP simulator that has been developed to simulate the functioning of a company with a team-oriented view.
MICSS has four views of a company, namely Marketing, Production, Purchasing and Finance. Each of these views has certain policies, which combine in an optimal way in order to be profitable for the company. However often the four departments of the company are unable to communicate properly and this creates discrepancies in the policies developed and hence, in information assurance.
MICCS enables us to simulate the functioning of a company through one year. We divided this period of one year into 6 periods of 2 months.

Design of Experiment

We have decided to study 4 factors in this experiment.
Factor 1:
*Dataset*; with 4 levels: Prices, QLT (Quoted Lead Time), Batch Size, and Order Levels.
Factor 2:
*Failure type*; with 2 levels: "wrong information", and "delayed information"
Factor 3 (nested in "wrong information"):
*Error size*; with 2 levels "value doubled", and "value halved".
Factor 4 (nested in "delayed information"):
*Length of delay*; with 2 levels "4 months", and "8 months".

So, we finally had 17 scenarios to simulate:

List of all the scenarios:

- *Correct information:*

(1) Baseline policy

- *Wrong information:*

(2) QLT doubled
(3) Prices doubled
(4) Batch Size doubled
(5) Order Level doubled
(6) QLT divided by 2
(7) Prices divided by 2
(8) Batch Size divided by 2
(9) Order Level divided by 2

- *Delayed information:*

(10) QLT delayed 4 months
(11) Prices delayed 4 months
(12) Batch Size delayed 4 months
(13) Order Level delayed 4 months
(14) QLT delayed 8 months
(15) Prices delayed 8 months
(16) Batch Size delayed 8 months
(17) Order Level delayed 8 months

Metrics

The Profit and the Due Date Performance (DDP) were recorded at the end of each period of 2 months. Profit represents how the whole company is performing, and the DDP gives an idea of how well the company is organized.
For each scenario 10 runs per year were conducted in order to have a statistical overview of the results.

Wrong information scenarios

A data of the baseline policy is modified (double or half) and MICSS is run for 2 months. Then the data is corrected and MICSS is run for intervals of 2 months to reach the end of the year.

Delayed information scenarios

A data of the baseline policy is modified (data-25%, because it is a realistic value that can be encountered in the functioning of the company). Then MICSS is run for 4 or 8 months, for intervals of 2 months, depending on the length of the delay we were simulating. Then the data is corrected and MICSS is run for intervals of 2 months to reach the end of the year.

Statistical Analysis

The hypothesis of the experiment was that the profits and DDP of the company in the case of delayed and wrong information would be different from the case of correct information.

$H_0$ = Performance (Profit or DDP) in the case of information failure (delayed or wrong information) is similar to the performance of the correct information.

$H_1$ = they are significantly different.

$\alpha = 0.05$ (a 95% confidence interval to prove the hypothesis.)

if p val $\leq 0.05$,we can conclude with 95% confidence that we reject the null hypothesis $H_0$

To verify the above hypothesis, the data was analyzed using single factor ANOVA, an analysis tool in EXCEL.

### 4.3. Results:

The observations haven't been analyzed like a nested design. We didn't need all the information given by a nested design analysis. For simplicity and time saving, we have used single ANOVAs to compare each time two different scenarios.

For each dataset, the following comparisons are presented in [18]:
Dataset delayed 4 months / Baseline policy (for profit and DDP).
Dataset delayed 8 months / Baseline policy (for profit and DDP).
Dataset wrong half / Baseline policy (for profit and DDP).
Dataset wrong double / Baseline policy (for profit and DDP).

The datasets are presented in this order: Prices, QLT, Batch Size, Order Level.
Then, the influence of the length of the time delay and of the difference between the wrong and correct data are presented.

Summary of the graphs that can be found in [18]:

*Prices*

Fig.A1 - Dataset delayed 4 months / Baseline policy (for profit and DDP).
Fig.A2 - Dataset delayed 8 months / Baseline policy (for profit and DDP).

Fig.A3 - Dataset wrong half / Baseline policy (for profit and DDP).
Fig.A4 - Dataset wrong double / Baseline policy (for profit and DDP).

*QLT*

Fig.A5 - Dataset delayed 4 months / Baseline policy (for profit and DDP).
Fig.A6 - Dataset delayed 8 months / Baseline policy (for profit and DDP).
Fig.A7 - Dataset wrong half / Baseline policy (for profit and DDP).
Fig.A8 - Dataset wrong double / Baseline policy (for profit and DDP).

*Batch Size*

Fig.A9 - Dataset delayed 4 months / Baseline policy (for profit and DDP).
Fig.A10 - Dataset delayed 8 months / Baseline policy (for profit and DDP).
Fig.A11 - Dataset wrong half / Baseline policy (for profit and DDP).
Fig.A12 - Dataset wrong double / Baseline policy (for profit and DDP).

*Order Level*

Fig.A13 - Dataset delayed 4 months / Baseline policy (for profit and DDP).
Fig.A14 - Dataset delayed 8 months / Baseline policy (for profit and DDP).
Fig.A15 - Dataset wrong half / Baseline policy (for profit and DDP).
Fig.A16 - Dataset wrong double / Baseline policy (for profit and DDP).

*Dataset delayed 4 months / Dataset delayed 8 months*

Fig.A17 - Prices
Fig.A18 - QLT
Fig.A19 - Batch Size
Fig.A20 - Order Level

*Dataset wrong half / Dataset wrong double*

Fig.A21 - Prices
Fig.A22 - QLT
Fig.A23 - Batch Size
Fig.A24 - Order Level

Notations:

"D" means: The two scenarios give significantly different results.
"D –" means that the performance with information failure, for profit or DDP, is worse than with the baseline policy.
"D +" means that the performance with information failure, for profit or DDP, is better than with the baseline policy.
"S" means: The two scenarios give significantly similar results.

Example:

Two examples of the results obtained from our team lab experiment are presented (Fig. 2 and 3). The entire analysis can be found in [18].

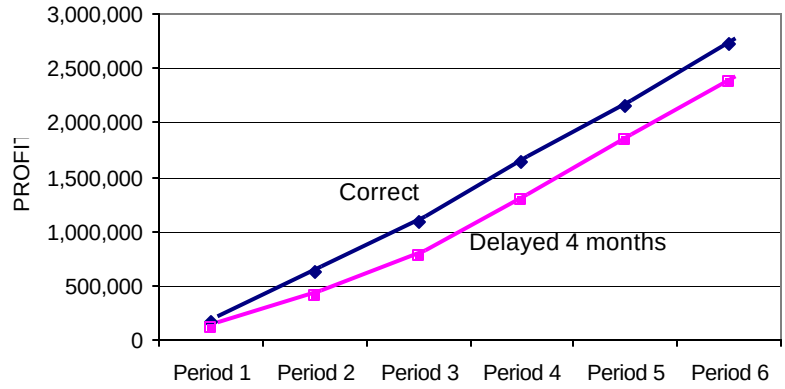
Fig. 2 – Prices; Dataset delayed 4 months / Baseline policy (for profit and DDP).

**Prices Delayed 4 months (10 runs)**

**Baseline policy (10 runs)**

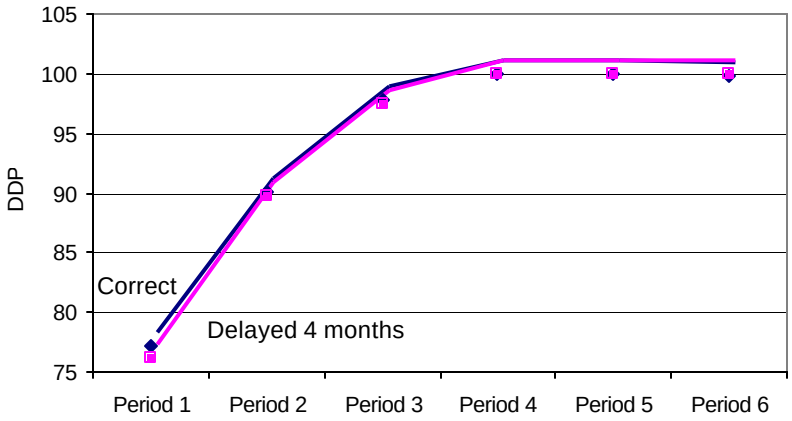
Metric: Profit

| Period 1 | Period 2 | Period 3 | Period 4 | Period 5 | Period 6 |
|---|---|---|---|---|---|
| D - | D - | D - | D - | D - | D - |




Metric: DDP

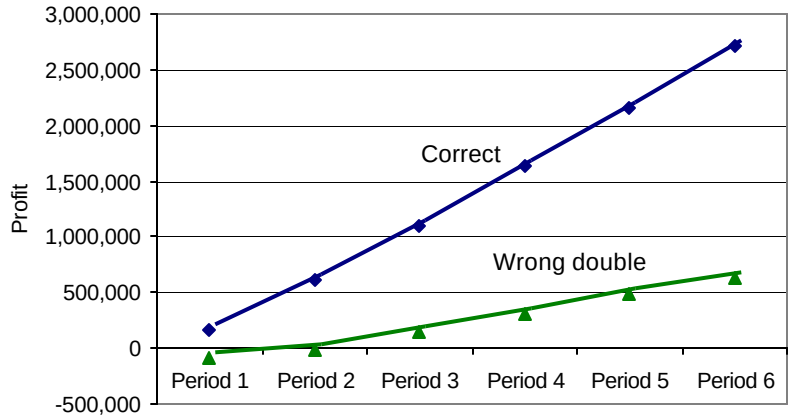
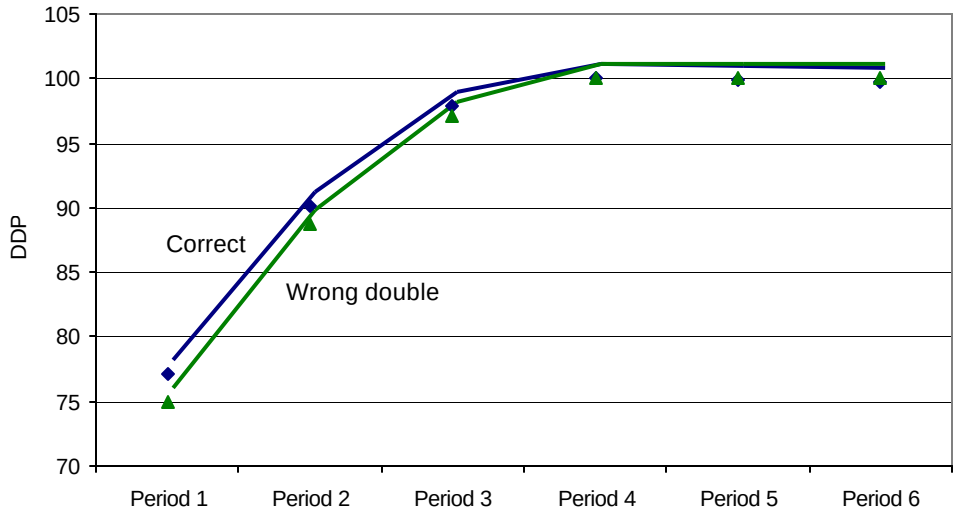
| Period 1 | Period 2 | Period 3 | Period 4 | Period 5 | Period 6 |
|---|---|---|---|---|---|
| D - | S | S | S | S | S |




Observations:

- For profit: during the 4 months of delay, the performance was worse. Then when the information was corrected (return to the baseline policy) the company followed the same trend as of the correct scenario, but the gap due to the delay could not be filled.
- For DDP: There were slight consequences that could be easily removed when the information was corrected.

Fig. 3 - QLT; Dataset wrong double / Baseline policy (for profit and DDP).

| | Baseline policy (10 runs) | | | | | |
|---|---|---|---|---|---|---|
| QLT Wrong double (10 runs) | Metric: Profit | | | | | |
| | **Period 1** | **Period 2** | **Period 3** | **Period 4** | **Period 5** | **Period 6** |
| | D - | D - | D - | D - | D - | D - |
| |  | | | | | |
| | Metric: DDP | | | | | |
| | **Period 1** | **Period 2** | **Period 3** | **Period 4** | **Period 5** | **Period 6** |
| | D - | D - | D - | S | S | S |
| |  | | | | | |





Observations:

- For profit: After returning to the baseline policy, the company did not follow the same trend as of the correct information. The slope is smaller. There were long-lasting consequences.
- For DDP: There were major consequences that lasted even after returning to the baseline policy. But finally the gap was filled.

### 4.4. Summary:

Table 2 summarizes for each dataset:

- Which information failure scenario had the largest impact on the functioning of the company ("1" means greatest impact).
- Which metric was the most affected by a failure in each dataset.
- Whether or not the length of delay had an influence on the results.
- Whether or not the error size had an influence on the results.

A complete analysis and graphical representation of these results can be found in [18].

Table 2 – Summary of the team experiment results.

| Dataset | Prices | QLT | Batch Size | Order Level |
|---|---|---|---|---|
| **Impact ranking** | 1. Wrong double<br>2. Wrong half<br>3. Delayed 8 months<br>4. Delayed 4 months | 1. Wrong double<br>2. Wrong half<br>3. Delayed 8 months<br>4. Delayed 4 months | 1. Wrong half<br><br>Then similar for:<br>Wrong double<br>Delayed 8 months<br>Delayed 4 months | Similar for all the scenarios |
| **Metrics sensitivity** | 1. Profit<br>2. DDP | Similar for profit and DDP | 1. DDP<br>2. Profit | Similar for profit and DDP |
| **Length of delay** | Important | Not important | Not important | Not important |
| **Error size** | Important | Important | Important | Not important |

### 4.5. Impact graphs:

Impact graphs summarize the impact of each information failure type by dataset (Fig. 4.a and 4.b). The relative differences:

a. (Profit with information failure – Profit with baseline policy) / (Profit with baseline policy)
b. and: (DDP with information failure – DDP with baseline policy) / (DDP with baseline policy)

are represented respectively in Fig. 4.a and 4.b.
These differences are shown using levels: [> 70%; 35 to 70%; 5 to 35%; +/- 5%; -5 to –35%; -35 to –70%; < -70%]

The following notations are used in Fig. 4.a and 4.b:
D4: scenario with information delayed 4 months
D8: scenario with information delayed 8 months
Wh: scenario with information wrong half
Wd: scenario with information wrong double

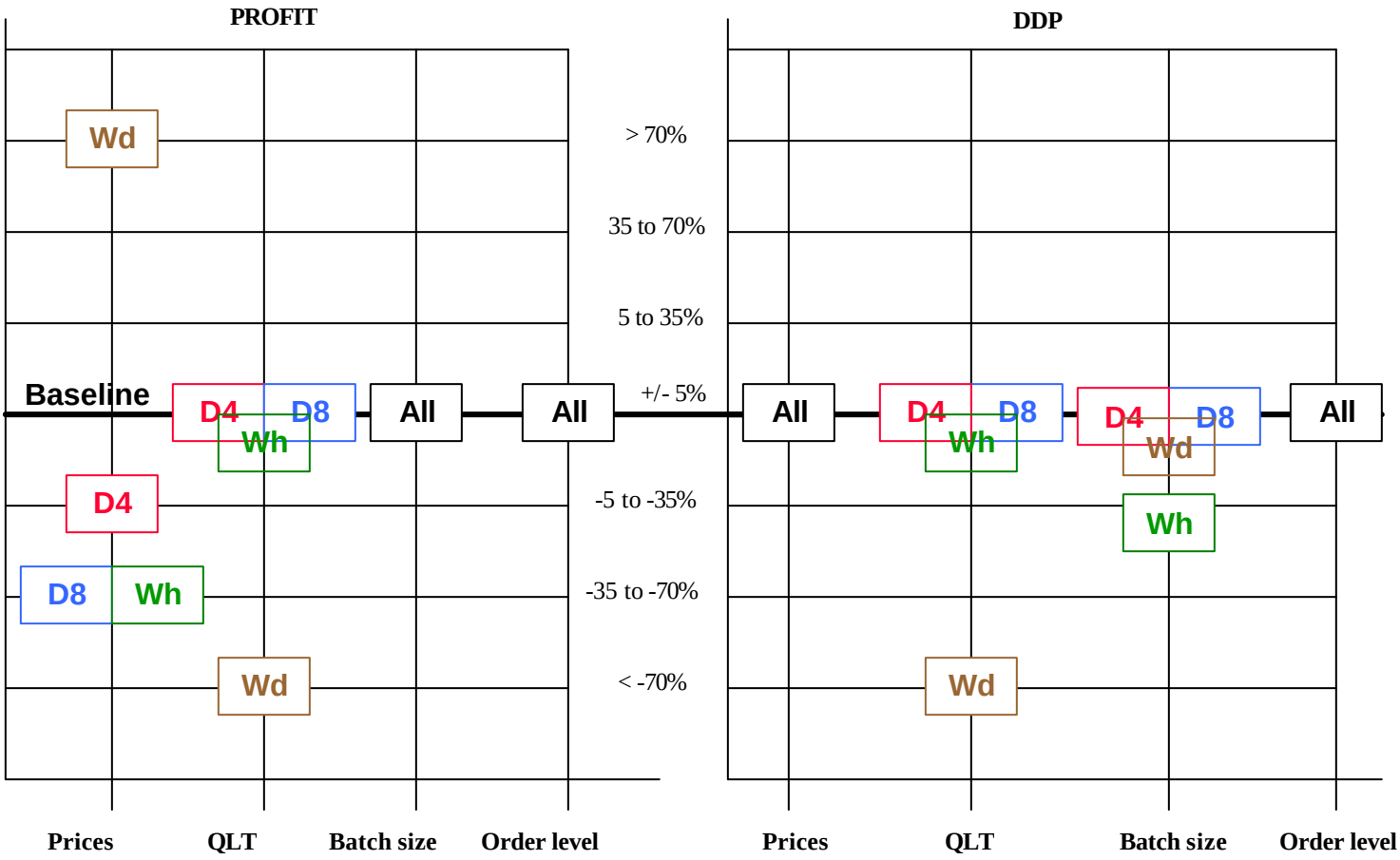



Fig. 4.a. – Failure Impact on Profit

Fig. 4.b. – Failure Impact on DDP

### 4.6. Conclusions:

1/ Some datasets are more sensitive than other. For example the consequences of a problem concerning Prices are much more serious and long lasting than when it concerns QLT. We can rank the datasets that have been tested by decreasing sensitivity: Prices, QLT, Batch Size, Order Level.

2/ Datasets have different characteristics that make them more sensitive to a specific type of information failure. For example, a delay of 8 months has a large impact on Profit when it concerns Prices, but no real impact when it concerns QLT.

3/ Profit is very sensitive to information failures. DDP react more slowly and need long lasting and large errors to be modified.

4/ The importance of information failure has been proved.

5/ The importance of the length of delay, and of the error size has been proved.

6/ We have seen that different scenarios can have very different consequences. A targeted security solution can then be designed to prevent the most serious cases first.

## REFERENCES


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

[2] Schwartz A.P. and Zalewski M.A., Assuring Data Security Integrity at Ford Motor Company, *Information Systems Security*

[3] Steinitz D., Information Security Management at British Airways. Implementing a Strategic Security Program, *15th World Conf. on Computer Security*, Nov. 1998

[4] Huang, C.Y., and Nof, S.Y., Formation of Autonomous Agent Networks for Manufacturing Systems, *Int. J. Production Research*, v.38, n.3, 2000, p.607-624

[5] Nof, S.Y., Tools and Models of e-Work, *Proc. Vth Int. Conf. On Simulation and AI*, Mexico City, February 2000.

[6] Voas J., Protecting Against What? The Achilles Heel of Information Assurance, *IEEE Software*, Jan.-Feb. 1999

[7] Finne T., What are the Information Security Risks in Decision Support Systems and Data Warehousing, *Computers & Security*, 16(3), 1997, p.197-204

[8] Ciechanowicz Z., Risk Analysis: Requirements, Conflicts and Problems, *Computers & Security*, 16(3), 1997, p.223-232

[9] Shirey R., Security Requirements for Network Management Data, *Computer Standards & Interfaces*, v.17 n.4, Sep 1995, p.321-331

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

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

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

[13] Fox B. and LaMacchia B., "Cooperative security": a model for the new enterprise, *7th Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises*, IEEE, Los Alamitos, CA, 1998, p. 314-319

[14] King C., Intranet application security checklist, *Computer Security Journal*, v.13, 1997, p. 47-53

[15] Holbein R., Teufel S., Morger O. and Bauknecht K., A comprehensive need-to-know access control system and its application for medical information systems, *13th International Conference on Information Security*, IFIP, 1997, p. 403-414

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

[17] 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 Memo School of Industrial Engineering No. 01-08*, Purdue University, June 2001

[18] 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 Memo School of Industrial Engineering No. 01-06*, Purdue University, January 2001

[19] Ang, C.B. and Nof, S.Y., Design issues for information assurance with agents: Coordination protocols and role combination in agents, *CERIAS Technical Report 2001-36, Research Memo School of Industrial Engineering No. 01-07*, Purdue University, February 2001