**DELFT UNIVERSITY OF TECHNOLOGY**
**EEMCS FACULTY**
**Department of Applied Mathematics**

# MASTER THESIS

## INFORMATION SECURITY RETURN ON INVESTMENT MODELING

**done by second year MSc student**
**Mikhail Gurov**
**St. number 1225863**

*DELFT, 2006*

| | |
|---|---|
| **ABSTRACT** | **3** |
| **INTRODUCTION** | **4** |
| **CHAPTER 1. WHAT IS INFORMATION SECURITY** | **6** |
| **1.1. Information evolution: historical outlook ……..** | **6** |
| **1.2. Security…..……** | **6** |
| **1.3. Information Security....……** | **7** |
| **1.4. Properties of Information…….** | **10** |
| **1.5. Threats …….** | **12** |
| **1.6. Vulnerabilities……..** | **14** |
| **1.7. Countermeasures…….** | **15** |
| **1.8. Institutes which regulate Information Security ……** | **16** |
| **1.9. Investment criteria (Project Values)…..……** | **17** |
| **1.9.1. Time Value of Investments …….** | **17** |
| **1.9.2. ROI in Information Security ……** | **19** |
| **1.9.3. Expected Benefits from Investments ……..** | **20** |
| **1.10. Possible problems with evaluation …….……..** | **21** |
| **CHAPTER 2. EXPERT JUDGMENT. THE CLASSICAL MODEL** | **23** |
| **CHAPTER 3. THE MODEL FOR RETURN ON INVESTMENTS IN INFORMATION SECURITY** | **28** |
| **3.1. Problem statement……..……..** | **29** |
| **3.2. Criteria…..……** | **37** |
| **3.3. Input data…….....……** | **40** |
| **CHAPTER 4. ANALYSIS OF THE MODEL** | **45** |
| **4.1. Analysis of strategies, one year period…….……** | **45** |
| **4.2. Analysis of strategies, three years period……** | **52** |
| **4.2.1. High-high information assets…….……..** | **53** |
| **4.2.2. Medium-high information assets…….……** | **57** |
| **CHAPTER 5. CONCLUSIONS AND PERSPECTIVES** | **60** |
| **REFERENCES** | **61** |
| **APPENDIX** | **62** |
| **A1. ELICITATION PROTOCOL** | **62** |
| **A2. FORMULA FOR COMPUTING RATES** | **69** |
| **A3. DICTIONARY** | **70** |
| **A4. DATA FOR CHOOSING THE SAMPLE SIZE** | **73** |
| **A5. Expected rate of attack in the next six years depending on the chosen strategy** | **78** |
| **A7. NPV distribution depending on the chosen strategy when disclosure cost equals 10 million,, for three-year project** | **79** |
| **A8. Expected NPV, ROI and expected flows for strategies when disclosure cost equals 10 million, for five-year project** | **81** |
| | **82** |

**ABSTRACT**

**The attempts to get access to sensitive information (informational espionage) of companies occur at an increasing rate. Espionage attacks bring most of losses among all attacks aiming at access to sensitive information. This raises the question of how to invest 'wisely' in information security defense systems in order to better protect enterprise from espionage. Decisions in this area are usually made based on subjective opinion of information security specialist within a company. Due to rapid technology developments there are no useful historical data banks.**

**The project aims to model the rate of attack per year in the next year and in a year in five years from now depending on the chosen investment (defense) strategy. Incidence rates on an Return-on-Investments (ROI) model are quantified with structured expert judgment. ROI and minimum affordable cost of one security breach are computed and analyzed for 60 possible investment strategies.**

## INTRODUCTION

The attempts to get access to sensitive information of companies occur often. Three recent examples are the arrest of Pepsi-Co employees trying to sell trade secrets to Coca-Cola for 1.5 million (Atlanta, 5 July 2006, [28]) and the arrest of U.S. government consultant who gained access to the passwords of 38,000 employees, including that of FBI Director Robert Mueller (Washington, DC, 6 July 2006, [29]) and the story about hacked server of the Ohio University which was supposed to be offline (June 26, 2006, [30]). According to Computer Crime and Security Survey-2006 [25], released by FBI and Computer Security Institute, the total loss of 313 companies which were willing to give estimates, was $52,494,290 in first half of 2006. In 2005 639 respondents has reported the total loss of 130 million. 44% of respondents do not share information about computer intrusions. This shows that breaches in information security (InfoSec) systems bring significant losses and companies are not eager to share information about it. Also, this survey has shown a decrease in percentage of almost all types of detected attacks among respondents (there is a small increase in system penetration, financial fraud, misuse of public web applications and sabotage). This proves that either companies pay extensive attention to information security problems or more and more attacks become undetected. Many companies are still not properly investing into Information Security. This brings the question of how to invest "wisely" into information security systems.

This study aims to develop a model to estimate the rate of successful espionage attack per year in the next six year from now depending on the chosen investment defense strategy. Protection of any type of secret information which can be exposed due to espionage attacks is intended to be covered. Incidence rates are to be quantified with structured expert judgment. This study is an initial approach and attack rates are elicited by one expert. The efficiency of investment strategies is measured then with the return and ROI, and the minimal cost of one disclosure is calculated. For a six-year rate of attack forecasting model, the life cycle of investment project is considered to be three years. It is assumed to be middle-term project because of the assumption that a company does not invest into InfoSec until the end of the Project. At the end of Chapter 4 computations for six-year project are also analyzed. The investment into internet-based

intrusion detection has been proven to be the most efficient. This is also what Network Security Poll Results (the survey of Network Computing) has shown. In this survey, intrusion detection systems are on the first place among information security tools to be deployed in the next 12 months. The results of this study are purely illustrative and serve merely to prototype the method for estimating attack rates and quantifying return on investments.

In Chapter 1, general issues about Information Security are covered. After that, the method of combining experts' assessments is given. In Chapter 3 the Model for ROI Modeling in Information Security is presented. The analysis of both one-year model and of the model with three year life-period is done in Chapter 4. The report is concluded with perspectives and Appendices.

## CHAPTER 1. WHAT IS INFORMATION SECURITY?

In the Age of Information everybody has the intuitive feeling what Information Security is. The first thing which comes into one's mind is that it is about catching hackers who penetrate computers of big banks through the Internet to steal millions of dollars. Or it can be the fraud of documentation of a new highly technological weapon from American military research institutions in favor of Russia. Both examples are part of the story, but Information Security (InfoSec) is more than that. Before dealing with Information Security a brief look at its counterparts, Information (subsection 1.1) and Security (subsection 1.2), is presented. It will give an understanding why this field of knowledge has recently become of paramount importance and why it plays crucial role in the everyday life of a modern society. Then three commonly used properties of Information Security and its importance will be explained in subsection 1.4. After that threats and vulnerabilities (which together are the reasons for having problems with InfoSec) of enterprises are examined. Then in section 1.8 the description of main governmental institutions which influence security policies of companies is given. In order to measure efficiency of investments into InfoSec economic criteria are needed. The most important of them together with their applicability is discussed in subsection 1.9. Chapter 1 is concluded with possible problems with evaluation of efficiency of investments into InfoSec.

### 1.1. Information evolution: historical outlook

Let us have a look at transformation of society due to a few qualitative changes in the way information evolves, called *information revolutions*. We will observe how information has become the most valuable asset of the modern society.

The history of a civilization runs to at least seven thousand years. Most of this time, the society has been developing science and technology, steadily accumulating pieces of knowledge. Nonetheless, there were a few boosts caused by the way people handle information and hence called 'information revolutions'. All of them are connected with the ability of a human being to transfer information (matter, consciousness) from one information *bearer* to another, and with ability to incarnate the same piece of information on different bearers. These two properties define *portability* - the independence of information and its bearer. As we will see, the degree of portability of

information has changed greatly since the early years of humankind. It plays crucial role in the progress.

The *first information revolution* was circa 5-6 thousands of years in Mesopotamia [1]. Characters were developed then. Later it was independently discovered in China and after that in Central America by Mayas. Before using characters, the transfer of information was only verbal. This involved a few imprecise transformations of information on the way from one bearer (brain cells) to another. It follows that information was partly lost and distorted with each transfer at a time.

The *second information revolution* has happened with the development of manuscript around 1300b.c. in China, and 800 years later in Greece [1]. Now there is no necessity to know an individual personally to get pieces of knowledge from her/him. The circle of people who are able to transfer the information is enlarging, as well as the amount (size) of portable information. The distortion is still high: while copying texts by hand people tend to make mistakes and their handwriting can be difficult to read.

The next, *third information revolution* happened between 1434 and 1444 with Guttenberg's invention of the printing-press machine. It could be an exaggeration though to say that that the invention has immediately drastically influenced the transfer of knowledge by the increase of the amount of published books. There was already very powerful 'printing' industry of monasteries in Europe at that time. Each monk was able to reproduce 1200-1300 pages in one working week [1], and for a while the church could make more books than printing-press machines. Nevertheless, printing changed the way information was handled. It steadily reduced the price of books, making them affordable not only for rich and educated people, but also for the poor. There is one more important aspect. The Invention has left monks without job. The independence of bookmaking from the church has greatly changed the content of books. First, the same books as monks were copying before were published (philosophy, religion and ancient texts), but later books in languages other than Latin and on other various subjects has appeared. The invention of printing-press technology has given birth to popular literature. As a consequence of boost in knowledge, it changed the system of education. In the next decades a number of universities devoted to social sciences was opened in Europe. With the third revolution

people have become more independent in thinking and in choosing subjects they would like to know.

The *fourth information revolution* is happening these days with the development of computers and digital communication used both in business and at home. Technical operations which could take months (such as modeling the shape of airplane wings, car design with respect to crush resistance, applying to universities abroad, finding a proper employee or a trade partner) can be now performed in a few hours with help of computers, mobile connection and Internet. Availability of information has increased greatly. Any two users of the Internet interested in a particular matter can found each other by means of Internet. They can easily share information and communicate.

People can easily access enormous amounts of knowledge which is accumulated so far and contribute it. Two scientists from Berkley University estimated the size of new information appeared in the period 2000 to 2003 [2]. The research has shown the growth rate of 66% per each year! Everybody knows the famous quote of one unknown wise man – "Those who owns the information owns the world". Information has become the most valuable asset and even a special term 'information society' has appeared.

> Definition. (Wikipedia) An *information society* is a society in which the creation, distribution and manipulation of information is becoming a significant economic and cultural activity.

Information nowadays is the most valuable asset for most of businesses. No wonder that more and more crimes happen around information. Companies incur informational fraud which is very difficult to estimate, but which it is counted in billions [10]. This is why in the era of information the question of its security has gained the crucial role in business and society.

### 1.2. Security

Although the notion of Information is relatively clear, it is difficult to say what the word 'security' means exactly. There exists social security, national security, public security, data security and even food security[1]. Security means something which makes

[1] when people do not need to live with hunger or fear of starvation, [6]

you feel safe, something you can rely on. It provides stability. Security is also a financial instrument, but we do not consider it in this thesis. The right to be secure is guaranteed by constitutions of most of the countries [7]. According to [4]:

> "Security, security measures – measures taken as a precaution against theft or espionage or sabotage etc".

Security is a set of measures which keep things around go right, corresponding to the laws. It is the prime aim of the state to provide security.

### 1.3. Information Security

The search engine www.google.com, responds with about half of the links related to Information Security in response to the phrase 'security'. This speaks says for itself: people care a lot about data security nowadays. Broad definition of Information Security is given in [8]:

> *Information Security* - the concepts, techniques, technical measures, and administrative measures used to protect information assets from deliberate or inadvertent (intentional or negligent) unauthorized acquisition, damage, disclosure, manipulation, modification, loss, or use.

In some books the term *Computer Security* is used [9]. This is because computers are now the main information bearer and processor. Although forty years ago people would rather use the term Information Security, now it is almost the same. Consider a few examples:

1. You keep you credit card, photos of children and CV together in your pocket bag. It was stolen while traveling. Abusers have all data they need to threaten your family
2. Somebody intercepted your telephone call and has got the inside information
3. Your fellow gave you video cassette from Lending Service and there is a bug inside. An adversary listens to all home conversations
4. An intruder penetrated your house and stole valuable papers

5. While working at home you throw the draft of new marketing strategy of your company into the trash bin. Opponents have got it.

None of these examples is about Computer Security. We will understand Information Security in a broad sense, but mostly be focused on security of information stored and transferred by means of computer devices or related equipment, i.e. on *security of digital assets*. We are going now to examine three important properties of information security which are broadly used [10, 19].

### 1.4. Properties of Information

There are three commonly accepted and used properties of digital assets which should be preserved while residing in, possessed by and communicated through various system resources (including hardware, software, firmware, information data and telecommunications). These properties are *confidentiality*, *integrity* and *availability*.

**Confidentiality**

Imagine that a group of criminals occasionally get data about banking accounts and personal profiles of credit card holders of one of the banks. Criminals know addresses of people and their credit history. This puts card holders under risk and allows them to sue the Bank, which has failed to provide its clients with *confidentiality* of their accounts. After such a story, the bank will lose trust and get big penalties from clients and from regulating authorities. If strategic plans of an oil company have been stolen by competitors, consequences can be very serious.

*Confidentiality is preserved* when only authorized users get access to information they are supposed to access. There are several ways of providing confidentiality, including both hardware and software solutions as well as a sound security policy (called *countermeasures*). You can find some of them in table 1. The consequences of confidentiality loss can be from just embarrassing (the lost of patients' data in hospital) to terrific (disclosure of roots for nuclear fuel movement). Time component plays an important role here because many secrets are valid for a limited period of time.

| | |
|---|---|
| Hardware | Encryption<br>Authentication System<br>Data Partitioning System<br>Disk Drive Lock<br>Firewall System<br>Intrusion Detection System<br>Redundant Data Storage System<br>Theft Detection System<br>Uninterruptible Power Supply |

| | |
|---|---|
| Software | Access Control Management System<br>Anti-virus software<br>Audit Data Reduction System<br>Authentication System<br>Automated Security Policy Planning Systems<br>Content Scanning System<br>Data Partitioning System<br>Encryption Systems<br>Firewall System<br>Intrusion Detection System<br>Network Mapping System<br>Password Cracking System<br>Public Key Infrastructure (PKI) Systems<br>Risk Assessment System<br>Theft Detection System<br>Transaction Auditing Systems<br>Virtual Private Network<br>Vulnerability Scanning System |
| Policies | Data Access Restrictions |

Table 1. Measures to provide confidentiality

**Integrity**

Consider another situation. A bank has a very confidential banking system. The patch which processes data transactions of the financial year and generate reports, makes a mistake. Active is not equal to passive in the Balance Sheet. Such a report makes no sense, because information is not consistent. Bank loses time to localize the error. The Central Bank gives fines for the delay of the Balance Sheet. This shows that it is important to preserve *integrity property* of information. If hackers intrude the database and make changes, in e.g. report the data do not possess then the integrity property. If the changes are made after delivering the report and it is not used later, there are no consequences. This proves that time component can also be engaged while preserving integrity.

Any information should be consistent. It means that there should be no controversial and lost pieces of information. More or less, everybody relies upon integrity of the possessed data. Otherwise you simply cannot trust it and hence cannot use it. To draw the line, *integrity* refers to the wholeness and continued unchanged nature of information [11].

**Availability**

Consider an example. A bank has a problem: anti-virus software has not been updated in time and server gets worm-virus[2]. People are trying to access their accounts, but the system is busy with the worm and the server is not responding. Clients' businesses cannot

[2] Worm virus is self-replicated program which copies itself many times so that it conquers all the available resources

make payments and lose potential profit. They go to the court and sue the bank, which has failed to provide them with a proper service. This example shows the importance of *availability* – the last property of information (but not the least). This property means that information must be available to its users (people or processes) when desired. The reasons for loss of availability can be very different in nature – from activity of viruses (like worm) and physical destruction of server to improper planning of system and resources (hard drive is full, processor is overloaded). This property of information is also the baseline for information security: once availability is lost it can be not possible to identify if integrity and secrecy properties are still possessed. For example, if administrator detected that the system is not responding it can be due to its failure, but it is also possible that there is a malicious software which takes all processing time. What can threaten an enterprise so that it may lose some of properties of InfoSec?

We will now talk about threats and vulnerabilities which (together) can cause problems with Information Security.

### 1.5. Threats

*Threat* is a possible danger to an enterprise. There exist three broad categories of threats

- natural and physical threats – all kinds of disasters (tornado, fire, flood, power shortage, occasional failure of the equipment). You can predict and manage some of such events. For instance, you can put the second reserved (replicated) server not next to the main one, but in the neighboring district, or you can install fire alarm system. Planning actions before hand and taking preventive measures will minimize risk from this kind of threat
- unintentional threat – threat from the person with no malice. It is all kinds of wrong actions – spilling soda onto expensive equipment, deleting important files by mistake, giving wrong rights to users, etc.
- intentional threat – threat by a person with malice, either insider or outsider. If high technologies of espionage are used then the threat is intentional, other cases are less obvious to identify as intentional ones.

Who possess the threat and what is the motivation for intruders? Aims can be greatly different [11]. It can be circumstantial intrusion of user (without malice), or a random attack to company, for example, by vandals. For some individuals (hackers) it is a matter

of prestige to break through. Attacks with malice are more dangerous. This sort of action is done in order to get personal gains (by both insiders and outsiders), or simply to perform an act against society (for example, against the World Bank). Competitors can cause harm to system destructing it or slowing down the decision.

Let us examine *intentional threats* in more detail. How can a person with malice get access to data? We are mostly focused on digital assets, so it is plausible to examine the ways in which computer system can be penetrated. The attack can be either from outside or inside. The list of widely used tools and techniques to intrude the system is places here:

- port scanners – program scanning open ports of a particular server
- network scanners – tools for examining the network for holes
- packet sniffers – programs getting login and passwords
- password crackers – programs using a dictionary to find a word which match an encrypted user password
- buffer overflows – client programs which give the input to the target program longer than it is supposed to be. It contains executable code, which may be run by server (if there is no control of input length)
- Trojan horses – programs containing some useful function, but in reality just disguising some malicious activity

Once the intruder is in the system, he/she can do a few things – depending on the intentions and on the privileges he has got:

- gathering trophies – simply collecting sensitive data, with no commercial intent
- shutting down the system
- steeling sensitive data with malice
- altering data on server (creating new accounts, corrupting data)
- running malicious software (password grabbers, useless processes)
- altering software (changes in functionality, altering what users see)
- alerting users about bad security

Companies would like to know and to monitor all possible threats. To resist threats from the hyperspace, for example, there are special programs, which are responsible for network monitoring. Examples of such programs are firewalls and filtering routers: all

traffic that comes into the network is analyzed. There are different kinds of network-based attacks which can be detected via network monitoring. We give here only brief overview. More details on this subject can be found in [17, section 4.3]:

- denial-of-service (DoS) attacks. This is the attack which makes a given service not available to authorized users. It attempts to greatly reduce performance or shut down the server. This can be done, for example, by sending specially crafted packets (which uses bugs in software) to a server, or pinging it with illegally long payload.
- Probes and network mapping. The first thing to do in the attacker's list is to get (valid) IP addresses of active machines. This act is called *network mapping*. In order to penetrate the system it is important to determine specific information about individual machines within network (such as operational system). This is what probing does: it sends special pockets to computer in order to check the response. The idea behind that is simple: different OS systems react differently on different stimuli.
- Gaining access. *Password guessing* is the simplest way of trying to get access. *TCP Hijacking* is more clever attack: it takes advantage of the fact that some computers trust each other. So, if A trusts B, we can pretend to be a trusted machine B, set up the connection, but have the connection. Machine B is flooded to make sure is does not respond for any packets from machine A.

In any case, threats are not that dangerous if system is well-protected. Information system should be *vulnerable* to be at risk.

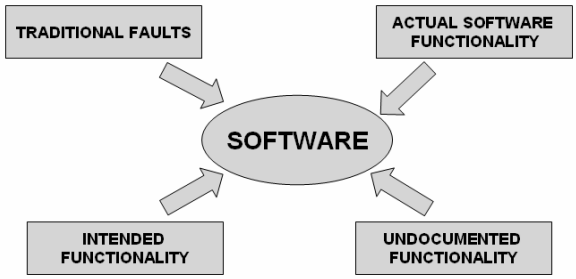
### 1.6. Vulnerabilities

*Vulnerability* is the susceptibility of an enterprise to attack. All information systems are vulnerable, the question is how much time and money it will take an intruder to get access to classified information (i.e. to successfully attack the enterprise). Besides, the longer is the time to intrusion, the more time administrators have to detect it and take protective actions. Unsafe system with no threat will exist with no problems. We give the list of possible vulnerabilities below [10]:

- physical vulnerabilities (physical access to system)
- natural vulnerabilities (susceptibility to natural disasters)

- hardware and software vulnerabilities (failures in system)
- media vulnerabilities (data bearers can be damaged or stolen)
- emanation vulnerabilities (electric signals can be intersected)
- communications vulnerabilities (the line can be tapped)
- human vulnerabilities (insider or weakly trained person)
- exploiting vulnerabilities (unexpected easy-to-discover ones)

Many intrusions from cyberspace happen because of vulnerability of software. Let us examine the reason for software failure in more detail. Most commonly it fails because of embedded defects. Nothing is perfect, and this comes up to the price. The graph below shows the main reasons of software failure [18].




Plot 1. Reasons for software failure

Threats and vulnerabilities form the opportunity for individuals to make harm to an enterprise. In order to prevent it a company should take *countermeasures*.

### 1.7. Countermeasures

*Countermeasures* (synonymous with *safeguards* and *security controls*, [19]) are activities which protect an enterprise from any malicious use of information and which reduce its susceptibility to attack. Computer Security strives to detect vulnerabilities and threats and take preventive countermeasures. We distinguish measures which protect data stored in the computer system (passwords, backups, etc), measures which take care about data being transmitted (encryption, digital signatures, authentication) and physical security (physical access, physical protection, shredders, etc.). There exists management, operational and technical controls prescribed for information system in order to preserve confidentiality, availability and integrity. A detailed classification of them (which takes about 30 pages) is given in [19]. This recommended security controls are developed by National Institute of Standards and Technology, Department of Commerce for Federal

Information Systems. We will explain the need in this kind of documents in the next section.

### 1.8. Institutes which regulate Information Security

The life of the society depends on how secure it is. In the age of information a lot depends on information systems, which store personal information, medical data, accounting data and other crucial information. In order to prevent disclosure from information systems of organizations dealing with such data, there are special governmental institutions, which set requirements for security level of companies. These institutions make research and form standards which information security systems of enterprises should satisfy to be allowed to run business. We list main of these institutions and briefly describe them below.

| Name | Description |
|---|---|
| NIST | Founded in 1901, NIST is a non-regulatory federal agency within the U.S. Commerce Department's Technology Administration. NIST's mission is to promote U.S. innovation and industrial competitiveness by advancing measurement science, standards, and technology in ways that enhance economic security and improve our quality of life.<br><br>NIST carries out its mission in four cooperative programs:<br>• the NIST Laboratories, conducting research that advances the nation's technology infrastructure and is needed by U.S. industry to continually improve products and services;<br>• the Baldrige National Quality Program, which promotes performance excellence among U.S. manufacturers, service companies, educational institutions, and health care providers; conducts outreach programs and manages the annual Malcolm Baldrige National Quality Award which recognizes performance excellence and quality achievement;<br>• the Manufacturing Extension Partnership, a nationwide network of local centers offering technical and business assistance to smaller manufacturers; and<br>• the Advanced Technology Program, which accelerates the development of innovative technologies for broad national benefit by co-funding R&D partnerships with the private sector.<br>NIST has an operating budget of about $930 million |
| NCSD | The Computer Security Division (CSD) - (893) is one of eight divisions within NIST's Information Technology Laboratory.<br>The mission of NIST's Computer Security Division is to improve information systems security by:<br>• Raising awareness of IT risks, vulnerabilities and protection requirements, particularly for new and emerging technologies;<br>• Researching, studying, and advising agencies of IT vulnerabilities and devising techniques for the cost-effective security and privacy of sensitive Federal systems;<br>• Developing standards, metrics, tests and validation programs:<br>1. to promote, measure, and validate security in systems and services<br>2. to educate consumers and<br>3. to establish minimum security requirements for Federal systems<br>• Developing guidance to increase secure IT planning, implementation, management and operation |
| Basel Committee | The Basel Committee on Banking Supervision (Comité de Bâle) is an institution created by the central bank Governors of the *Group of Ten* nations . It was created in 1975 and meets regularly four times a year.<br><br>Its membership is now composed of senior representatives of bank supervisory authorities and central banks from the G-10 countries (Belgium, Canada, France, Germany, Italy, Japan, the Netherlands, Sweden, Switzerland, the United Kingdom and the United States), and representatives from Luxembourg and Spain. It usually meets at the Bank for International Settlements in Basel, where its 12 member permanent Secretariat is located.<br><br>The Basel Committee formulates broad supervisory standards and guidelines and recommends statements of best practice in banking supervision (see bank regulation or Basel II, for example) in the expectation that member authorities and other nation's authorities will take steps to implement them through their own national systems, whether in statutory form or otherwise.<br><br>The purpose of the committee is to encourage convergence toward common approaches and standards. |

Tables 2. Main institutions forming requirements for InfoSec systems

The interested reader can find more information about it in [9, 10, 11]. In the next section, issues concerning the Model for Investments into InfoSec are approached.

### 1.9. Investment criteria (Project Values)

In order to be able to assess the efficiency of Investments into InfoSec technologies and controls of an enterprise (which is the prime aim of the Project), we need some criteria. There are a few economic factors called *project values* which aim to help economists. A brief overview of investment criteria and its applicability for the Project is discussed below.

#### 1.9.1. Time Value of Investments

We would like to understand if we can use such economic parameters of a project as NPV, IRR and PI in Information Security. The following definitions are necessary for that.

*Cash Flow, or Net Cash Flow* (CF, or NCF) equals cash receipts minus cash payments over a given period of time [12]. CF analysis is used to maintain a healthy financial situation of an organization and is based on analyzing the counterparts of CF – *inflows* and *outflows*. In Information Security inflow is the decrease in losses caused by breaches in information system and outflow is the investment in InfoSec. Cash Flow is the baseline for computing Net Product Value (NPV) – an important parameter of any investment project.

*Discounted Cash Flow* (DCF) takes into account the time value of money and it equals (discounted) future cash flows brought to today:

$$DCF_1 = \frac{CF_1}{1+r}; DCF_2 = \frac{CF_2}{(1+r)^2}, \quad ..., \quad DCF_n = \frac{CF_n}{(1+r)^n}$$

Here $r$ is a constant interest rate and $n$ is the number of time intervals. Inflows can be money borrowed for the project and profits from the project, and outflows can be payments for the credit and expenses related to the project.

*Net Present Value* (NPV) is the sum of Discounted Cash Flows:

$$NPV = \frac{CF_1}{1+r} + \frac{CF_2}{(1+r)^2} + ... + \frac{CF_n}{(1+r)^n}$$

NPV is expressed in terms of monetary units. If it is positive, the project is profitable and investments should be made.

*Internal Rate of Return* (IRR) is based on the same principles and calculations as NPV. It is defined as the interest rate which makes discounted future cash flows equal to initial investment of the project. In other words, it is the rate $r_{IRR}$ such that:

$$NPV = \frac{CF_1}{1 + r_{IRR}} + \frac{CF_2}{(1 + r_{IRR})^2} + ... + \frac{CF_n}{(1 + r_{IRR})^n} = 0$$

IRR shows the maximum rate when project still has positive NPV. This indicator differs from industry to industry and should be bigger than the current interest rate (otherwise NPV is negative).

*Profitability Index* (PI) is a ratio between discounted benefits and discounted cost of a project. It takes positive values and measures the present value per dollar invested. PI should be bigger then 1.

NPV, IRR and PI are the main investment criteria's. They are an integral part of any business plan. One can wonder if these indicators can be used with respect to investments in Information Security. The meaning of in-flows and out-flows is to be understood in order to do that.

*In-flow* is extra money company is going to spend on Information Security. It formed not only by initial payments. There can be hidden costs involved, such as costs for extra support or expenditures caused by problems with availability because of extra preventive measures.

*Out-flow* is the revenue brought by implementation a project. It is a potential profit which an enterprise is currently losing due to the vulnerability of its information system and which an organization will get if investments are made. The value of potential profit is difficult to estimate. It will require analysis of current losses due to problems with confidentiality, integrity and availability, and the potential losses in the future. Some assessments have been already done about it. Dan Erwin of Dow Chemical, for example, explained in an interview with Richard Power of the Computer Security Institute his approach to evaluating losses after an incident occurs []. He said that if a computer system is affected, they calculate the value of a minute of a computer time. Then the costs are computed as

$$costs = \frac{VTn}{N}$$

where

$V$ – the value per minute

$T$ – duration of the disruption ni minutes

$N$ – total number of people using their computers

$n$ – number of people unable to use the system

To that number, they would add the cost of the investigation plus additional costs such as missed sales, fines for not being able to pay bills on time, or demurrage on warehouse storage or railcar usage. For insider fraud, Erwin suggests adding in the cost related to bad press, including lost goodwill and effect on stocks, plus the costs to replace people ([9], p.388). Finally, comparison of current and potential costs will give the out-flow. We see that estimation of flows is an important and difficult talk. Its calculation can be problematic because of too many uncertain and vague factors.

There are also two more parameters: *Return on Investments* and *Expected Benefits from Investments*, which are standing separately from the group of 'standard' time-value parameters NPV, IRR and PI.

### 1.9.2. ROI in Information Security

Return on Investments is a measure of corporate profitability and generally is defined as

$$ROI = \frac{profit - investment}{investment} \times 100\%$$

This parameter is well-known in economic accounting. It was first used circa 1912 by E.I. du Pont de Nemours and Company. The model developed by a 27 year old explosives sales engineer F.Donaldson Brown, which has often been referred to as the *Dupont Model*, has been used since that time. ROI is frequently used in cost-benefit analysis, which strives to answer the question: should we purchase the item/invest in a project? ROI can be easy-to-define with respect to the whole organization and overall efficiency. Although if we decide to use it as a supporting tool for taking investment decisions in Information Security, it could be problematic to directly count ROI.

The reason for that is the difficulty with defining the benefits which implementation of one or another InforSec project provides. How to measure the return

on buying extra anti-virus or a new biometric access system? If a company has suffered losses (not necessary material ones) because of the intrusion, it can be difficult to define if such an intrusion could be prevented with extra protection. It depends on the level of intruder and his/her connections with personnel from your organization. One should also distinguish financial returns (the ones which are reflected in the budget as cost reduction and increase of the revenue) and non-financial returns (customer satisfaction, reputation, higher quality, faster service, etc.). The second ones are long-term, strategic benefits which to some extend define company's survival on the competitive market. Such returns are difficult to estimate.

One thing which can be done about assessing the return on investments is to go from quantitative value of profit to probabilistic assessments of one or other characteristics of Information Security. For instance, it is possible to judge, from the experience how less likely is the probability of successful attack if we change the firewall settings from 'medium threat' to 'high threat'. If the influence of one or other technology on such probability is known, an analogue of ROI parameter where returns are in terms of probabilities can be computed. Such influence can be investigated through expert judgment questionnaires and the use of modern probabilistic models. Hence, ROI in the modified sense (its analogue) can be used as a supportive tool for making an investment decisions as a relative value. Such value separately has no sense.

### 1.9.3. Expected Benefits from Investments

ROI analogue can be approached from another point of view. In Information Security benefits are measured in terms of loss reduction. Let us define the following parameters

$E[L]$ = expected loss, calculated in monetary figures (such as dollars)

$P_O$ = the probability of occurrence of a risk prior to an investment

$P_I$ = the probability of occurrence of a risk after an investment

$V$ = the cost value of an asset

$\varepsilon$ = the exposure value of an asset

The following intuitive formula can be used to count the expected loss of a threat, which (if occurred) leads to negative consequences:

$$ExpectedLoss = Probability(Accident)*LossValue*e$$

$$E[L] = P*(V*\varepsilon)$$

ε is the amplification factor which helps to take into account not only direct value of loss *LossValue* caused by an accident, but also the exposure that occurs with the loss of the asset. This may include loss of human life, loss of opportunity, or may include all kinds of expenditures necessary to restore the situation to the initial state.

To estimate the expected benefits from investing into InfoSec, we need to understand how much we gain from the investment. Assume $E[L_O] >= E[L_I]$ and $P_O >= P_{INV.}$ Then

$$\Delta L = P_0 * L * \varepsilon - P_{INV} * L * \varepsilon = -\Delta P * L * \varepsilon$$

The higher value of $\Delta P$ is, the higher is the decrease of loss caused by making investments. The negative side of this approach is that it does not take into account the amount of investments. We need to have such an indicator that if we compare two different strategies of different size but with the same $\Delta L$, it should give higher value for a smaller investment strategy. The ratio $\frac{\Delta L}{Inv}$ satisfies such a property. This indicator (index) has similar structure as ROI. For this reason we will call it *ROI analogue for InfoSec* and write

$$ROI_{SEC} = \frac{\Delta L}{Inv}$$

This value shows what part of every invested dollar comes back in the reduction of loss.

**1.10. Possible problems with evaluation**

The human factor is one of the parameters which make ROI difficult-to-evaluate. It is well-known that 80% of the intrusions happen with help of insiders ([10], p.16). Does it mean that if we had lost some confidential materials we should hire more professional personnel managers? Or if we are not exposed to abuse it is due to talented IT managers? Assumptions of a model must be carefully defined and checked in order to get defendable results.

There are questions and problems which still wait to be solved and answered. Nevertheless, it is possible to evaluate return on investments in Information Security given certain assumptions. In the next chapter the Classical Expert Judgment Model is

presented. After that the Model for evaluating the effect of Investments in InfoSec will be defined.

## CHAPTER 2. EXPERT JUDGMENT. THE CLASSICAL MODEL

Expert judgment is used when it is impossible (or too expensive) to get direct observations of the studied quantity. For example, it is unrealistic to perform statistical experiments in order to get failure rate of energetic system, and historical data is rare. Instead, we can ask experts to express their (subjective) beliefs about the matter of interest (failure rate). The aim of combining assessments from a *pool* of experts is to achieve rational consensus. There are several methods of weighting and combining expert judgments. The scheme of performance-based weighting of experts in the *classical model* (which will be explained below) satisfies a proper scoring rule. It means that expert archives his/her maximal expected score (in the long run) only by giving true beliefs. The Classical model is approved in many expert judgment studies and for this reason is used to obtain assessments of CPTs values in the Model.

The classical model of combining expert judgments uses questionnaire. Experts are asked to give their assessments for uncertainty distributions of random variables in form of quantiles. Usually in risk analysis experts are asked to express beliefs on a studied matter with (5%, 50%, 95%) or (5%, 25, 50%, 75%, 95%) quantile assessments, but any other quantiles can be chosen as well. The 95% degree of belief can be also explained by requirements of regulating authorities. In many studies 50% quantile (medium) is used. It is important to make a clear background (definitions, assumptions, etc.) on which an uncertainty is assessed to avoid ambiguity in understanding questions. A weighted combination of the experts' assessments is called a *decision maker.* It is a normalized linear combination of the experts' assessments with respect to their weights.

Not all experts perform equally well. Assessments of 'better' experts should get a higher weight (score) in the overall linear combination. There are two types of questions. S*eed,* or *calibration* questions allow to measure how good the experts are in quantifying their uncertainty. Second, there are questions of *interest variables*, i.e. questions on uncertain quantities being assessed.

Seed questions must be chosen from the expert's research field, related to the questions of interest. Answers to them are known or will become known in the time after the project. A score measuring 'goodness' of an expert is calculated based on the answers

to seed variables as a product of two values, namely, a *calibration score* and an *information score*. The calibration score says how statistically well the expert performs. The information score measures the expert's uncertainty about the requested matter. The second type of questions concerns assessing the quantities we are interested in called *interest variables.* Questions are also called *items*. Mathematical description of the Classical model is given below. First we define information score and calibration score. After that schemas for computing weights of experts are given.

Let us assume that we have one seed question with realization $r$ and $N$ experts. Each of them give *(5%,50%,95%)* quantiles for this seed variable

$$(q_5(e), q_{50}(e), q_{95}(e)), e = 1,..., N$$

In order to compute information score, intrinsic range is to be defined. I*ntrinsic range* is the interval *(l,h),* such that

$$l = \min_{e=1..N}\{q_5(e), r\}, h = \max_{e=1..N}\{q_{95}(e), r\}$$

i.e. it is the minimum interval containing all experts' assessments and the realization of the chosen seed variable. Intrinsic range should contain the whole distribution. For this reason, *k%* overshoot is often included in range *(l,h)*. The sensivity to the choice of $k$ must be checked.

*The information score* defines the degree to which the experts' assessments are concentrated relative to user-defined *background measure.* This measure is assigned to each variable for each expert. For the uniform and for the lognormal distributions, for example, the background measure is:

$$F(x) = \frac{x - l}{h - l}, \qquad x \in (l, h)$$

$$F(x) = \frac{\ln x - \ln l}{\ln h - \ln l}, \qquad x \in (l, h)$$

If the background measure is uniform, then the expert's distribution for seed variables is uniform between quantiles 0 and 5%, 5% and 50%, etc. *The information score* of an expert for one seed question is then defined as

$$I(e,1) = \sum_{i=1}^{4} p_i \ln \frac{p_i}{r_i}, \qquad p = (0.05, 0.45, 0.45, 0.05),$$

where $r_i$ is a background measure for interval *i,* that is

$$r_1 = F(q_5(e)) - F(q_l(e)), r_2 = F(q_{50}(e)) - F(q_5(e)),$$
$$r_3 = F(q_{95}(e)) - F(q_{50}(e)), r_4 = F(q_h(e)) - F(q_{95}(e)),$$

Let us have $M$ seed questions. The *information score* of an expert is defined as the average of information scores for each of the seed variables:

$$I(e) = \sum_{i=1}^{M} I(e,i)$$

The information score shows how confident the expert is in estimations, i.e. how concentrated is the distribution. We chose experts with higher information between those with approximately equal calibration scores. The value of the information score depends on the choice of background measure and intrinsic range, but usually this dependence is negligible.

Let us consider one expert and $M$ seed questions. We would like to estimate the ability of the expert to predict realizations. If we ask him/her to give (5%,50%,95%) quantiles, then 5% of the realizations should fall in the interval (<5%), 45% in the interval (5%-50%), 45% in the interval (50%-95%) and 5% in the interval (>95%). Let us denote these realizations as $(s_1 M, s_2 M, s_3 M, s_4 M)$. To evaluate how close the empirical density of the expert $(s_1, s_2, s_3, s_4)$ is to the hypothetical one $(p_1, p_2, p_3, p_4) = (5\%,45\%,45\%,5\%)$, so-called *relative information* is used:

$$I(s, p) = \sum_{i=1}^{4} s_i \ln \frac{s_i}{p_i}$$

Its minimum value zero is archived if and only if $s=p$. *Calibration score* is based on $I(s, p)$. It is the likelihood of statistical hypothesis which is defined for each expert as [27]:

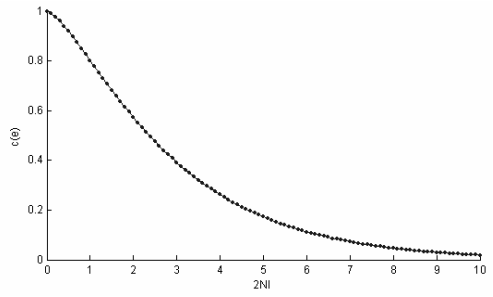
> *The realizations may be regarded as independent samples from a distribution corresponding to the expert's quantile assessments.*

We would like to test the degree to which the realizations for seed variables support this hypothesis, i.e. to check if discrepancies between the realizations and the expert's assessments have appeared by chance. It is well-known that

$$P(2M * I(s, p) \le x) \approx \chi_3^2(x)$$

The *calibration score* for expert is defined as the probability to get the relative information score worse than obtained, under assumption that his/her true distribution is $(p_1, p_2, p_3, p_4)$. The score is expressed as

$$c(e) = 1 - P(\chi_3^2 < 2NI(s, p)) = 1 - \chi_3^2(2MI(s, p))$$




Plot 5. Function c(e)

The best case is when $c(e) = 1$, because the expert's empirical distribution is just the same as the hypothetical one. Low calibration score (say less then 0.05) means that the expert's probabilities are not supported by seed variables.

Let $q(e) = (q_5(e), q_{50}(e), q_{95}(e)), e = 1,..., N$ be the assessment of *N* experts of the variable of interest. We need to combine these distributions to get the distribution of the Decision Maker:

$$q(DM) = \sum_{e=1}^{N} w_e q(e)$$

Three schemas are used for computing the weights of experts

- *Global weights* $- w_e$ equals the product of the calibration and the information score (normalized). The information score equals to the average of item information scores:

$$InfScore(e) = \sum_{i=1}^{N} InfScore_i(e)$$

- *Equal weights* $- w_e = \frac{1}{N}$
- *Item weights* – uses individual scores for each of the experts and items, and the same calibration score for any of the items within each of the experts. This means that

$$w_e = \frac{CalScore(e)\sum_{i=1}^{M} InfScore_i(e)}{\sum_{e=1}^{N} CalScore(e)\sum_{i=1}^{M} InfScore_i(e)}$$

*DM optimization* is performed to maximize the score of the decision maker after computing scores for experts. Cut-off level $\alpha$ is to be defined first. If the calibration score of an expert is lower than $\alpha$, such an expert is taken as unweighted. Scores for the rest of the experts are to be normalized and then a new DM is constructed. This DM is added to the pool of experts and its score depends on the choice of $\alpha$. The value of $\alpha$ which brings maximum to DM score within the pool of experts is the one required. This procedure can be applied for the global as well as for the item weighted Decision Makers.

A researcher should be aware of possible difficulties while performing the expert judgment with help of the Classical model. Firstly, the choice of seed questions influences the results of elicitation. Secondly, the questions should be both such precise that to avoid ambiguity of its notation by the experts, and correct from the mathematical point of view. Thirdly, there is a significant assumption that the experts' performance on interest variables can be judged on the basis of their performance on seed variables. The distributions on seed variables are assumed to be independent, i.e. the choice on any of them does not influence other choices.

## CHAPTER 3
## THE MODEL FOR RETURN ON INVESTMENT MODELING

In this Chapter, the model for quantifying the return on investment (ROI) into information security systems will be presented. A complete model involves quantifying

i. the incident rate (not necessarily constant) of unwanted events under various defense strategies
ii. the costs of various defense strategies,
iii. the costs of a security breach,
iv. the present value of future costs/potential profits

The first two parts of the problem are generic. They recur in a large population of information systems. Due to rapid technology developments and because of specific type of considered unwanted events, we do not have banks of useful historical data about its incident rates. However, we do have a robust population of knowledgeable experts who can give useful information. Due to the time restrictions and in order to verify the model, the incidence rates are first assessed by one expert and the elicitation session is not hold yet. Later, it will be quantified with *structured expert judgment,* when experts' uncertainty is treated as scientific data, and is elicited and processed according to accepted protocols. Incidence rates for a particular type of unwanted events and under certain defense strategies are considered. More detailed information about considered events and defense systems can be found in Subsection 3.1.

Defense strategies are levels of information security defense systems a company would like to invest in. Its costs may be obtained through open channels. These costs are dependent on the size of a company, number of its offices and employees. The cost of a security breach depends on the type of information asset disclosed. The Expert's answers to questionnaire, a company profile which is defined in order to estimate investment expenditures, the associated costs of defense systems and types of information assets are discussed in subsection 3.3, devoted to the input data.

In order to estimate the plausibility of making an investment, or to compare two investment strategies, criteria are needed. Criteria calculated and used in the Model are given in Subsection 3.2. We will study the time dependence of the incidence rates, and money becomes cheaper in time. This will be taken into account while comparing the strategies for one-year model and while computing the net product value (NPV) of project in the three-year model. We proceed to the description of the Model.

### 3.1. Problem statement

We focus on a defended information systems as may be found e.g. in banks, hospitals, government agencies, the military etc. The rate at which such a system will suffer successful attack will be quantified, under a variety of defense systems. The nature of the information assets is essential in determining the *costs* of a successful attack, but not for the successful attack rate itself. This can be explained by the fact that companies are aware of and do similar actions in order to keep the information assets safe. Furthermore, they are all interconnected through the Internet and telephone lines. Hence, they are dependent on their partners not to have problems with data security.

The attention is restricted to Loss-of-Confidentiality events, and Loss-of-Integrity or Unavailability are not considered. In particular, we consider the number of successful attacks resulting in unwanted data disclosure, in a given population of systems in a given year in the future, under various defense systems. Unwanted disclosures may be precipitated by agents who are either *benign* or malicious. Malicious attacks may be the result of espionage, disgruntled employee sabotage, hackers, or any other malicious activity. In terms of damage, those **attacks from espionage** constitute the dominant concern and will be the focus of this study.

Companies have already installed and running security systems. We expect that purchasing any additional security *can* improve protection of information assets. Three most common defense systems are considered for implementation:

- *Human shields:* training personal to better counter attackers
- *Barriers and enclaves:* firewalls, gaps, meshes
- *Intrusion detection*: physical and virtual

Physical intrusion detection systems are identical to the various anti-burglar systems and will not be considered here.

We distinguish the following levels of implementation for these systems:

| **Human shields** | **Barriers and enclaves** | **Intrusion Detection** |
|---|---|---|
| 0.do nothing<br>1.Training program in Information Security for all staff<br>2.(1) plus certification of key personnel | 0.Do nothing<br>1.Internet / intranet segregation<br>2.(1) plus enclaving sensitive areas with firewalls and routers<br>3.(2) plus enclaving sensitive areas with space gaps<br>4.(3) plus copper meshing (against van Eck radiation reading) | 0.Do nothing<br>1.NIDs (Network-based Intrusion Detection systems)<br>2.(1) plus HIDs (Host-based Intrusion Detection systems)<br>3.(2) plus Hybrid systems. |

Table 3. Levels of defense systems

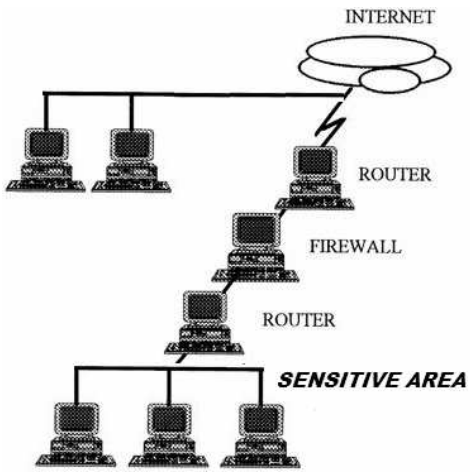
These levels are to be understood successively; thus "Internet / intranet segregation" means implementing *only* this segregation. "Segregation plus firewalls and routers" means implementing *only* firewalls and routers, in addition to segregation. This also means that any attack scenario prevented at a certain level would also be prevented at all successive levels. Below the description of defense systems levels is given.

**Human Shields**

A training program includes seminars, lectures and workshops on information security for all staff. Information security certification for all “key” personnel (i.e. who has access to system definitions) together with training program forms the second level of human shields.

**Barriers and Enclaves**

Firstly, the investment into (better) segregating inter- and intra-nets can be made. The Intranet is the collection of private computers within on organization. If these computers have access to the Internet, they can also be reached directly from it. The internet/intranet segregation includes restricting access to internal resources for users from internet. It can be done by the login/password identification, ip-validated access or giving signal/access by correct mac-address.

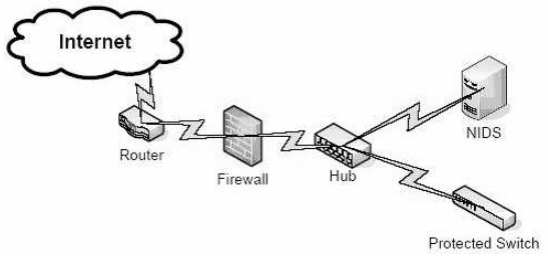


Secondly, in addition to the internet/intranet segregation, computers with sensitive information can be enclaved with firewalls and routers. All messages passing through *firewall* (which can be a piece of either hardware or software) are examined and those who do not meet the specified security criteria are blocked. A *router* is a piece of hardware which connects the LAN (local area network) to the WAN (wide-area network) of the internet. It contains the routing table, which allows filtering the traffic based on the IP addresses.

Thirdly, a spatial separation can be installed around enclaved systems so that to restrict physical access to sensitive enclaved systems. The fourth level of barriers and enclaves also includes installing copper meshing – a special copper barrier which does not allow to read electromagnetic waves emitted by devices containing sensitive information.

**Intrusion Detection Systems**

Network-based Intrusion Detection System (NIDS) is a separate independent modul which identifies intrusions by examining network traffic and monitoring multiple hosts. This system is connected to the internet through a *hub*, a network switch configured for port mirroring.




The second level besides NIDS includes implementing host-based intrusion detection system. It consists of an *agent* (a special program) on the host[3], which analyses system calls, application logs, file-system modifications and the other host activities and states in order to identify the intrusion.

The last level includes implementing a Hybrid Intrusion Detection System, which combines approaches to intrusion detection. Host agent data is combined with network information to form a comprehensive view of the network. This level is the most comprehensive and costly.

We would like to assess the uncertainty on the rate of disclosure due to successful espionage attack in a population of comparable information systems, under these three defense strategies. Hence, each strategy is associated with three numbers stating for the chosen level of implementation. Strategy "do nothing" about all three defense systems is

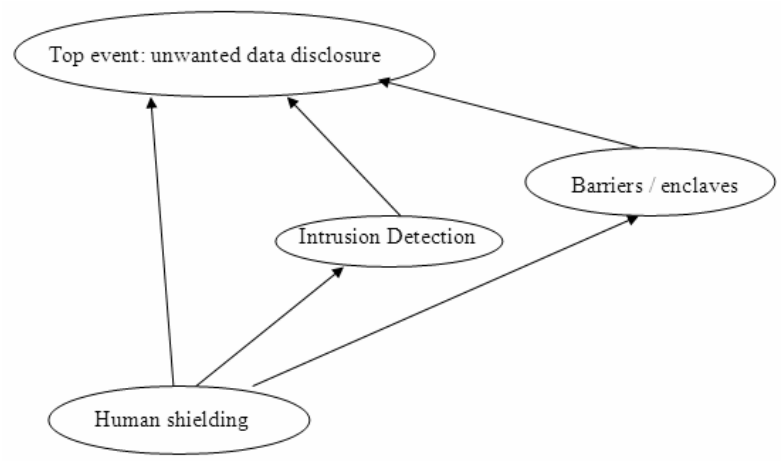
[3] host is any device in the network which has a unique name by which a network attaches this device

called the *base*, or *zero* strategy. We assume that, given a human shielding posture, there are no significant interactions between the barrier / enclave and intrusion detection strategies. That is, the effect of a barrier strategy is neither amplified nor diminished by the implementation of an intrusion detection strategy[4]. Human shielding, on the other hand, may interact with the other strategies. Whereas training of personnel is primarily designed as a defense against human ingression, it may be expected to heighten security awareness in general and thus enhance the effectiveness of other strategies. We anticipate possible interactions between human shielding and:

-space gap barriers

- Host-Based IDs

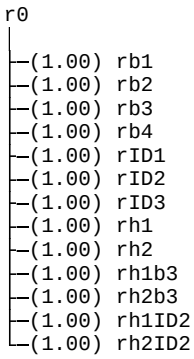
The model may be graphically represented as




Plot 6. Graphical representation of the Model

Any attack scenario prevented if no extra defense systems are installed, is also prevented by successive levels. If more attacks are prevented (stopped) at the base level (i.e. rate decreases), the more attacks are prevented at higher levels of defense systems (i.e. with any of the strategies). There can not be the situation when the decrease in the attack rate for zero-strategy corresponds to the increase in the attack rate for non-zero strategy. This means that corresponding rates are highly correlated with the rank correlation 1. Hence, the dependence structure can be chosen as:

[4] Suppose the DO NOTHING rate is $3\times10^{-4}$ events per unit time. Suppose the rate under strategy 1.3.2.2 is $2\times10^{-4}$, and that under 1.3.3.3 the rate is $1.5\times10^{-4}$. Then the rate for implementing BOTH 1.3.2.2 and 1.2.3.3 is $3\times10^{-4}\times (2/3)*(1.5/3) = 10^{-4}$. This reflects the fact attacks prevented under a weaker strategy are also prevented under a stronger strategy, and the effects of Barrier and Detection strategies are (conditionally) independent.




Plot 7. Dependence structure

In order to define the influence of levels of defense systems to the distribution of successful attack rate, sampling is performed in Unicorn. The diagonal band copula is used to join random variables. Marginal distributions for components of dependence tree are necessary for that. We assess them in the following way.

A population of one million comparable information systems under a given defense strategy is considered, and questions of the following form shall be asked:

*In a population of one million comparable Information systems with defense strategy X, how many unwanted data disclosures will occur NEXT year as a result of malicious espionage attack?*

| | | | | |
|---|---|---|---|---|
| _____ | _____ | _____ | _____ | _____ |
| 5% | 25% | 50% | 75% | 95% |

*How many unwanted data disclosures will occur in one year starting FIVE years from now as a result of malicious espionage attack (assuming no further changes to the system)?*

| | | | | |
|---|---|---|---|---|
| _____ | _____ | _____ | _____ | _____ |
| 5% | 25% | 50% | 75% | 95% |

The full questionnaire is given in the Appendix. Distributions of successful attack rate (resulted in unwanted data disclosure, per million of comparable systems, per one year) for the following strategies has been elicited:

| N Quest | Human Shielding (hum) | Barriers (bar) | Intrusion Detection (id) | Rate of Attack | Used in function N |
|---|---|---|---|---|---|
| 13 | No (0) | No (0) | No | R0 | all |
| 15 | No | Inter/intra (1) | No | rb1 | 1 |

| | | | | | |
|---|---|---|---|---|---|
| 17 | No | Firewall (2) | No | rb2 | 1 |
| 19 | No | Space gap (3) | No | rb3 | 1,3 |
| 21 | Training (1) | Space Gap | No | rh1b3 | 3 |
| 23 | Certification (2) | Space Gap | No | rh2b3 | 3 |
| 25 | No | Copper Mesh (4) | No | rb4 | 1,3 |
| 27 | No | No | NIDs (1) | rID1 | 2 |
| 29 | No | No | HIDs (2) | rID2 | 2,5 |
| 31 | Training | No | HIDs | rh1ID2 | 5 |
| 33 | Certification | No | HIDs | Rh2ID2 | 5 |
| 35 | No | No | Hybrid (3) | rID3 | 2,5 |
| 37 | Training | No | No | rh1 | 4 |
| 39 | Certification | No | No | rh2 | 4 |

Table 4. Strategies from questionnaire

The rate of successful attack for any of 60 strategies can be computed from these 14 rates and dependence structure. The formula for that can be found in Appendix and more explanation is given below.

**Dependence and independence in the Model**

There are two kinds of dependences in the Model. First, *"strategy<->rate" dependence* is the dependence between the defense strategy (which is associated with three values $hum \in \{0,1,2\}, bar \in \{0.1.2.3.4\}, id \in \{0,1,2,3\}$) and the corresponding rate of attack. The independence in this sense means that the change in strategy has no influence over the change of the rate. Any of the strategies are highly correlated to "base" strategy $hum = 0, bar = 0, id = 0$ with rate $r0$, because any attack scenario prevented at certain level is also prevented by the successive levels.

Second, there is *"strategy<->strategy" independence*, which is preserved when there is no significant interactions between strategies with respect to how it affects the rate of attack.

Suppose that, given the level of human shielding *hum*, the barrier strategy (its effect) *bar* and the intrusion detection strategy *id* are independent. Denote the attack rate of the strategy $A = (hum = 0, bar = i, id = 0)$ as $r_{hum=0,bar=i,id=0}$ and the attack rate of the strategy $B = (hum = 0, bar = 0, id = j)$ as $r_{hum=0,bar=0,id=j}$. Then from the independence and the successive nature of strategies, the rate of implementing both is:

$$r_{hum=0,bar=i,id=j} = r0 \times \frac{r_{hum=0,bar=i,id=0}}{r0} \times \frac{r_{hum=0,bar=0,id=j}}{r0}$$

The effect of changing strategy *bar* has no influence over the effect that strategy *id* has on the overall rate. If we do not anticipate possible interactions between *hum* and *bar*, *hum* and *id*, and if strategy *hum=0* is changed to *hum=k*, then:

$$r_{hum=k,bar=i,id=j} = r0 \times \frac{r_{hum=k,bar=0,id=0}}{r0} \times \frac{r_{hum=0,bar=i,id=0}}{r0} \times \frac{r_{hum=0,bar=0,id=j}}{r0}$$

In this case the change of the base rate caused by the strategy

$$(hum = 0, bar = 0, id = 0) \rightarrow (hum = k, bar = i, id = 0)$$

is expressed with a ratio

$$\frac{r_{hum=k,bar=0,id=0}}{r0} \times \frac{r_{hum=0,bar=i,id=0}}{r0}$$

If interaction between *hum* and *bar* is expected, all rates $\{r_{hum=k,bar=i,id=0}\}_{k=0:2,i=0:4}$ are required in order to compute the change in the attack rate for corresponding strategies.

The independence of *bar* and *id* given *hum* means that:

$$\frac{r_{hum,bar=i,id}}{r_{hum,bar=i-1,id}} = \frac{r_{hum,bar=i,id=0}}{r_{hum,bar=i-1,id=0}}$$

If, in addition, *hum* and *bar*, *hum* and *id* are independent, then this ratio is the same for any *hum* and

$$\frac{r_{hum,bar=i,id}}{r_{hum,bar=i-1,id}} = \frac{r_{hum,bar=i,id=0}}{r_{hum,bar=i-1,id=0}} = \frac{r_{0,bar=i,id=0}}{r_{0,bar=i-1,id=0}}$$

The "strategy<->strategy" independence means the same proportional increase/decrease of rate with the change in level of one type of strategy given any level of another type of strategy. If this increase/decrease is dependent on the condition, two types of strategies are dependent.

**An example**

We can expect the interaction between certain levels of strategies. The attack rate cannot be computed from the effects that strategies (say *hum* and *bar*) have on base rate *r0* alone and is to be requested from the experts. Consider the example.

Let us anticipate the interaction between human shielding strategy *hum* and the intrusion detection strategy *id=2*. Let us compute the rate for strategy $(hum = 1, bar = 0, id = 2)$. Dependence means that we cannot compute its rate as:

$$r_{hum=1,bar=0,id=2} = r0 \times \frac{r_{hum=1,bar=0,id=0}}{r0} \times \frac{r_{hum=0,bar=0,id=2}}{r0}$$

or as

$$r_{hum=1,bar=0,id=2} = r0 \times \frac{r_{hum=1,bar=0,id=1}}{r0} \times \frac{r_{hum=0,bar=0,id=2}}{r_{hum=0,bar=0,id=1}}$$

Assume that an expert gives assessment $r_{hum=1,bar=0,id=2} = rh1id2$. How to compute the rate for the strategy $(hum=1, bar=0, id=3)$ from it? If there is no interaction between *hum* and *id=3*, the change in rate because of

$$(hum=1, bar=0, id=2) \rightarrow (hum=1, bar=0, id=3)$$

is the same as for

$$(hum=0, bar=0, id=2) \rightarrow (hum=0, bar=0, id=3)$$

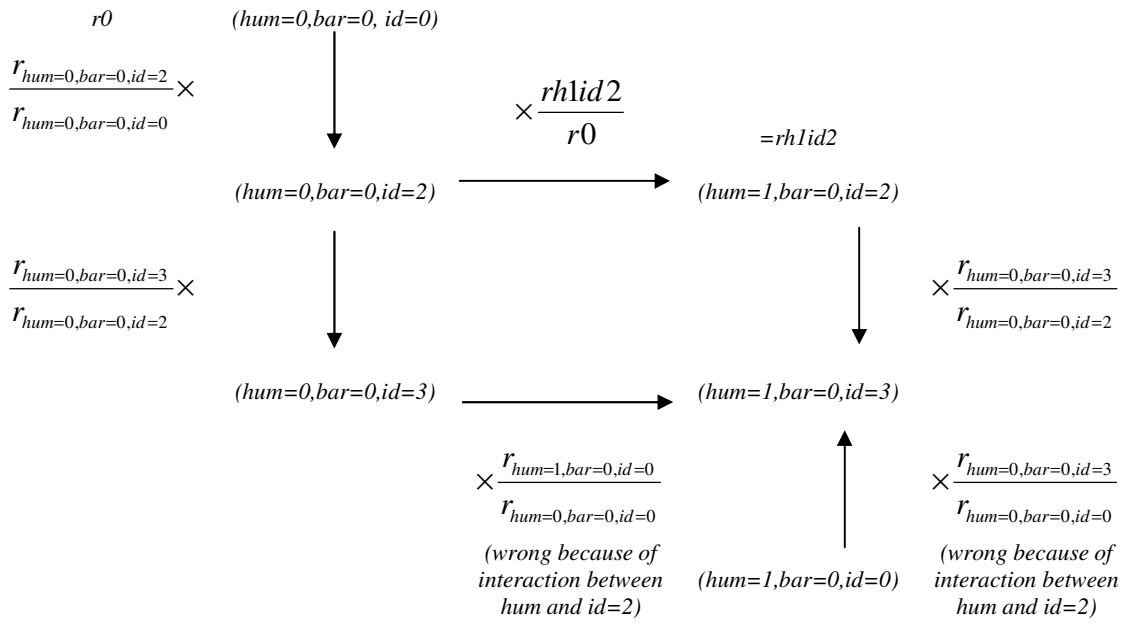
and

$$r_{hum=1,bar=0,id=3} = r_{hum=1,bar=0,id=2} \times \frac{r_{hum=0,bar=0,id=3}}{r_{hum=0,bar=0,id=2}}$$

If there is an interaction between *hum* and *id=3, then*

$$\frac{r_{hum=1,bar=0,id=3}}{r_{hum=1,bar=0,id=2}} \neq \frac{r_{hum=0,bar=0,id=3}}{r_{hum=0,bar=0,id=2}}$$

and expert **has to be asked about** $r_{hum=1,bar=0,id=3}$. The ratios changing the base rate r0 (no interaction between *hum* and *id=3* is assumed) can be represented as:




### 3.2. Criteria

Some criteria for judging the plausibility of the investment are needed. Suppose there are three strategies with rates and investment costs $(R_{ZERO}, P_{ZERO}=0)$, $(R_A, P_A)$ and $(R_D, P_D)$. How to identify if the decrease in rate 'worth' extra costs?

The ratio

$$-\frac{R_A - R_{ZERO}}{P_A - P_{ZERO}}, -\frac{R_D - R_{ZERO}}{P_D - P_{ZERO}} \qquad (1)$$

give the decrease (increase) of rate with every extra unit of investment *P*, when ZERO strategy (no investments) is changed to A or D. This way allows to order strategies, but it does not answer the question if the strategy is plausible (profitable). To define it, the price of one unwanted data disclosure is needed.

Let a company suffer (on average) loss $D_1$ with each disclosure. $D_1$ is the factor which works for the decision of making an investment. If it is high, the prevention of even several successful attacks will make significant profit. The Investment is more significant source of losses than disclosures, if secrets worth little.

The investment into defense systems is defensible if the income on it is higher than the investment. For the strategy A it is true if:

$$\mathrm{Re}turn = Income - Investment = (R_{ZERO} - R_A)D_1 - P_A > 0$$

Or if we compare strategies A and D, D is better than A if:

$$\mathrm{Re}turn(D) - \mathrm{Re}turn(A) = (R_A - R_D)D_1 - (C_D - C_A) > 0$$

After rearranging terms:

$$(0 - R_D) \times D_1 - (C_D - 0) > (0 - R_A) \times D_1 - (C_A - 0)$$

Now add $R_{ZERO} \times D_1$ to both parts:

$$(R_{ZERO} - R_D) \times D_1 - (C_D - 0) > (R_{ZERO} - R_A) \times D_1 - (C_A - 0) \quad (2)$$

Left part shows how much more profitable is D relative to the zero strategy and left part compares A to ZERO. With the help of this equation, all strategies are ordered by its profitability. However, the investment in one million can make 1 dollar profit. Then it is better to go to the bank and to have guaranteed incomes of a certain rate r. Investment managers are interested in the return on each unit of currency. For this reason, both parts should be divided by the corresponding investments. This criterion is called the *Return-on-Investment* (ROI) and was presented in Chapter 1. In case of strategies A and D the inequality is:

$$\frac{(R_{ZERO} - R_D) \times D_1 - P_D}{P_D} > \frac{(R_{ZERO} - R_A) \times D_1 - P_A}{P_A}$$

or

$$ROI_D > ROI_A \quad (3)$$

If this inequality is true, investment in D is more efficient. Given the value of investment, strategy with higher ROI is more preferable. However, ROI should be also higher than the bank rate *r*. Minimum value -1 is achieved when the price of one disclosure is zero or if the investment does not influence the rate of attack. It means that every invested dollar is lost.

The choice of criteria depends on the aim of making an investment

- If reducing the rate of attack is the primary goal, no criteria are needed. The highest investments give the highest reduction in the successful attack rate.
- If the company would like to maximize return (expressed in the reduction of rate) on each dollar invested in defense systems, first criteria is suitable

- If the company is interested in comparing return on investment of several projects, the second and the third criteria should be used

The price of one unwanted data disclosure is very business-specific and is assumed not to be known. Nevertheless, it is possible to take this price into account. Solving the following equation for a given investment $P_A$

$$ROI = \frac{(R_{ZERO} - R_A) \times D_1 - P_A}{P_A} > r$$

relative to $D_1$ we get:

$$D_1 > \frac{(1+r)P_A}{R_{ZERO} - R_A} = D_{MIN} \tag{4}$$

Value $D_{MIN}$ is the minimum price per one exposure at which there is economic sense of implementing a project A. If a project brings no reduction in rate, $D_{MIN}$ equals infinity, saying that there is no price of exposure at which a project is reasonable. If $R_A = 0$, the possible minimum value of $D_{MIN}$ is obtained. It means that all attacks are prevented and it gives maximum possible potential profit $R_{ZERO}D_{MIN}$.

Suppose that the rate in the next year from now is $R^1_{ZERO}$ and in the year starting five years from now it equals to $R^5_{ZERO}$. Suppose that after the investments, the rate in the next year is $R^1_A$ and in the year starting five years from now $R^5_A$. Assume also that the price of one exposure D stays the same in the next six years and no extra investments besides $P_A$ are made in next five years. We forecast rates in between linearly:

$$R^i_{ZERO} = R^1_{ZERO} + \frac{i}{5} \times (R^5_{ZERO} - R^1_{ZERO})$$

$$R^i_A = R^1_A + \frac{i}{5} \times (R^5_A - R^1_A)$$

Now the Net Present Value (NPV) of the project can be calculated from the table.

| | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Inflow (IF) | - | $R^1_{ZERO}D - R^1_A D$ | $R^2_{ZERO}D - R^2_A D$ | $R^3_{ZERO}D - R^3_A D$ | $R^4_{ZERO}D - R^4_A D$ | $R^5_{ZERO}D - R^5_A D$ |
| Outflow (OF) | $P_A$ | - | - | - | - | - |

Table 3. Inflow and outflow of the five year project

$$NPV(A) = -P_A + \sum_{1=1}^{5} \frac{(R^i_{ZERO} - R^i_A)D}{(1+r)^i}$$

$r$ is usually equal to refinance rate of the Central Bank. The project should have positive NPV to be attractive to investors.

We compute the minimal price at which the investment becomes reasonable from the following formula (investment is expected to pay itself and to bring profit in three years):

$$\frac{D_1(R^1_{ZERO} - R^1_A)}{1+r} + \frac{D_1(R^2_{ZERO} - R^2_A)}{(1+r)^2} + \frac{D_1(R^3_{ZERO} - R^3_A)}{(1+r)^3} - P_A \geq 0$$

$$D \geq D_{MIN} = \frac{P_A}{\frac{R^1_{ZERO} - R^1_A}{1+r} + \frac{R^2_{ZERO} - R^2_A}{(1+r)^2} + \frac{R^3_{ZERO} - R^3_A}{(1+r)^3}}$$

### 3.3. Input data

Expert's distributions are given in table 6.

| | | 0.05 | 0.25 | 0.5 | 0.75 | 0.95 | Price of strategy, USD |
|---|---|---|---|---|---|---|---|
| 13 | nothing at all | 1 | 500 | 750 | 900 | 1000 | 0 |
| Barriers and Enclaves | | | | | | | |
| 15 | Inter/Intra segregation | 1 | 5 | 10 | 15 | 20 | 15000 |
| 17 | I/I with firewalls | 1 | 2 | 3 | 4 | 5 | 90000 |
| 19 | I/I, F/w, and space gap | 0.01 | 1 | 2 | 3 | 4 | 135000 |
| 21 | I/I, F/w, space, w/ trng | 0.01 | 0.5 | 1 | 1.5 | 2 | 142000 |
| 23 | I/I, F/w, space, w/ cert | 0.01 | 0.5 | 1 | 1.25 | 1.5 | 162000 |
| 25 | I/I, F/w, space, w/ meshing | 0.01 | 0.25 | 0.5 | 0.75 | 1 | 210000 |
| Intrusion Detection Systems | | | | | | | |
| 27 | NIDs only | 1 | 5 | 10 | 15 | 20 | 5000 |
| 29 | NIDS plus HIDS | 1 | 3 | 7 | 12 | 15 | 35000 |
| 31 | NIDS + HIDS w/ trng | 1 | 3 | 7 | 12 | 15 | 42000 |
| 33 | NIDS + HIDS w/ certification | 1 | 3 | 5 | 8 | 10 | 62000 |
| 35 | Hybrid IDS | 1 | 2 | 3 | 4 | 5 | 85000 |
| Human Shields | | | | | | | |
| 37 | Training alone | 1 | 500 | 750 | 900 | 1000 | 7000 |
| 39 | Certification alone | 1 | 250 | 500 | 750 | 900 | 27000 |

Table 6. Expert distributions for number of attacks per on million of comparable information systems, in the next year

From this table, training alone has no effect on the number of successful espionage attacks, and the certification has the least influence of all other strategies. Also, implementing internet/intranet segregation has the same effect as implementing network-based intrusion detection systems, although the segregation is more expensive. The next table shows how the attack rate distributions are changed in five years. The rate significantly decreases for any strategy. Expert explains it with the predictions of a new type of machines which are much less vulnerable than today's ones.

| | | 0.05 | 0.25 | 0.5 | 0.75 | 0.95 |
|---|---|---|---|---|---|---|
| | | | | | | |
| 13 | nothing at all | 1 | 50 | 75 | 100 | 150 |
| Barriers and Enclaves | | | | | | |
| 15 | Inter/Intra segregation | 1 | 2 | 5 | 8 | 10 |
| 17 | I/I with firewalls | 0.01 | 1 | 2 | 3 | 4 |
| 19 | I/I, F/w, and space gap | 0.01 | 0.45 | 1 | 1.5 | 2 |
| 21 | I/I, F/w, space, w/ trng | 0.01 | 0.45 | 1 | 1.5 | 2 |
| 23 | I/I, F/w, space, w/ cert | 0.01 | 0.45 | 1 | 1.25 | 1.5 |
| 25 | I/I, F/w, space, w/ meshing | 0.01 | 0.228 | 0.5 | 0.75 | 1 |
| Intrusion Detection Systems | | | | | | |
| 27 | NIDs only | 1 | 3 | 7 | 12 | 15 |
| 29 | NIDS plus HIDS | 1 | 2 | 5 | 8 | 10 |
| 31 | NIDS + HIDS w/ trng | 1 | 2 | 5 | 8 | 10 |
| 33 | NIDS + HIDS w/ certification | 1 | 2 | 4 | 7 | 9 |
| 35 | Hybrid IDS | 1 | 2 | 3 | 4 | 5 |
| Human Shields | | | | | | |
| 37 | Training alone | 1 | 50 | 75 | 100 | 150 |
| 39 | Certification alone | 1 | 40 | 65 | 80 | 100 |

Table 7. Expert distributions for number of attacks per on million of comparable information systems, occur in one year starting five year from now

**Defining the cost of one successful disclosure**

According to our expert (Julie Ryan), setting a cost of an attack is problematic, unless a type of compromised data is specified. Once this is done, it is possible to guesstimate costs. This issue will be examined in more detail. The way the cost of an attack is valued in Information Security is via the equation

$$P = c * i$$

where *c* is the *cost of the asset* and *i* is the impact of the compromise (or an *exposure factor*). Thus, the cost value of a successful attack would vary according to the importance or exposure factor of the information. Consider a few examples (provided by expert):

1) The asset cost of identity information is minimal -- perhaps $10. The exposure factor is enormous, depending on the identity being compromised. For sake of argument, assume the identity of a middle class American with a job, a wife, and 1.2 kids. The cost of a successful attack could run into the hundreds of thousands of dollars not counting the years that it takes to clear up such an issue.

2) The asset cost of the plans for the F-117 stealth fighter is approximately $2 Billion. The exposure factor of revealing the secret aspects of design include not only the costs of lives lost and the costs of airplanes lost, but also the cost associated with designing and manufacturing a replacement airplane

3) The asset cost of a trade secret, say one that required 5 years of research and development, may be a couple million dollars. The exposure factor includes not only the costs of development but also the potential profits, market share acquisition, and opportunity costs.

Hence, a specific type of compromise should be chosen and used as a point of comparison for calculations on ROI. The following table helps in choosing the price of one disclosure. Further analysis has shown that investments into defense systems can be profitable only for the last two types of information assets.

| Asset cost | Exposure factor | Price |
|---|---|---|
| Low | Low | 100 |
| Medium | Low | 10 000 |
| High | Low | 100 000 |
| Low | Medium | 500 000 |
| Medium | Medium | 1 000 000 |
| High | Medium | 10 000 000 |
| Low | High | 1 000 000 |
| **Medium** | **High** | **10 000 000** |
| **High** | **High** | **100 000 000** |

Table 8. Prices for unwanted data disclosure depending on the type of secret

**Defining the price of investment in defense systems**

Naturally, the cost of investment into defense systems is dependent upon the size of the company and the number of offices. Thus, we have to put assumptions on the company we are making this analysis for. A company with 100 employees in one central office with three sensitive enclaves (finance, R&D, and operations) is chosen. Ten employers out of all personnel are "key" for purposes of certification. The following prices are estimated for such company profile:

| Defense systems, levels of implementation | Price, USD | Final price |
|---|---|---|
| **Human shields** | | |
| 0. do nothing | 0 | 0 |
| 1. Training program for staff in Info sec risk | 7000 per 100 employees | 7,000 |
| 2. (2) plus certification of key personnel | 70000 per 35 employees | 27,000 (7K for training plus 20K for cert) |
| **Barriers and enclaves** | | |
| 0. Do nothing | 0 | 0 |
| 1. Internet / intranet segregation, | 15000 | 15,000 |
| 2. (2) plus enclaving sensitive areas with firewalls and routers, | 30000 per enclave | 90,000 |
| 3. (3) plus enclaving sensitive areas with space gaps | 45000 per enclave | 135,000 |
| 4. (4) plus copper meshing (against van Eck radiation reading) | $25,000 per termpested room plus costs of previous features | 210,000 |
| **Intrusion Detection** | | |
| 0. Do nothing | 0 | 0 |
| 1. NIDs (Network-based Intrusion Detection systems) | 5000 per NIDS | 5,000 |
| 2. (2) plus HIDs (Host-based Intrusion detection systems) | 30000 per 100 machines for the HIDS alone | 35,000 |
| 3. (3) plus Hybrid systems | 50000 | 85,000 |

Table 9. Prices for defense systems

**The choice of the sample size**

The model gives the growth of the expected attack rate for the strategies with fixed *bar* and *id*, when there is no interaction between *hum&bar* (i.e. space gaps), *hum&id* (i.e. host-based IDS) and when *hum=0* is changed to *hum=1*. This happens due to the sampling error, because the influence of *hum=0* and *hum=1* on the rate is the same (expected rate should not change). This brings the question of the choice of the sample size. The difference is the highest for the strategy *(bar,id)=(0,0),* because the range of its counterparts is the widest (i.e. the highest nominator divided by the smallest nominator).

We compute the absolute difference between the expected rate of attack for strategies (0,0,0) and (1,0,0), depending on the number of samples and for the seed values 1,2 and 3. Results are presented in the table 10:

| Sample size, ~ per strategy | Seed=1<br>E(rate 0,0,0);<br>E(rate 1,0,0);<br>Abs. difference | Seed=2<br>E(rate 0,0,0);<br>E(rate 1,0,0);<br>Abs. difference | Seed=3<br>E(rate 0,0,0);<br>E(rate 1,0,0);<br>Abs. difference | Max absolute difference for a given sample size |
|---|---|---|---|---|
| 1,000 | 6.5091e-004;<br>6.6039e-004;<br>0.0948-004 | 6.5807e-004;<br>6.6151e-004;<br>0.0344-004 | 6.5530e-004;<br>6.5616e-004;<br>0.0086-004 | 0.0948-004 |
| 2,000 | 6.5800e-004;<br>6.5593e-004;<br>0.0207-004 | 6.5803e-004;<br>6.5730e-004;<br>0.0073-004 | 6.4631e-004;<br>6.5878e-004;<br>0.1247-004 | 0.1247-004 |
| 4,000 | 6.5874e-004;<br>6.5683e-004;<br>0.0191-004 | 6.5374e-004;<br>6.5976e-004;<br>0.0602-004 | 6.4719e-004;<br>6.5888e-004;<br>0.1169-004 | 0.1169-004 |
| 8,000 | 6.5456e-004;<br>6.5817e-004;<br>0.0361-004 | 6.5252e-004;<br>6.5605e-004;<br>0.0353-004 | 6.5449e-004;<br>6.5411e-004;<br>0.0038-004 | 0.0361-004 |
| 32,000 | 6.5361e-004;<br>6.5822e-004;<br>0.0461-004 | 6.5379e-004;<br>6.5567e-004;<br>0.0188-004 | 6.5544e-004;<br>6.5812e-004;<br>0.0268-004 | 0.0461-004 |
| 100,000 | 6.5368e-004;<br>6.5603e-004;<br>0.0235-004 | 6.5611e-004;<br>6.5539e-004;<br>0.0072 | 6.5436e-004;<br>6.5733e-004;<br>0.0297 | 0.0297-004 |
| 200,000 | 6.5486e-004;<br>6.5486e-004;<br>0.0000-004 | 6.5575e-004;<br>6.5574e-004;<br>0.0001-004 | 6.5585e-004<br>6.5585e-004;<br>0.0000-004 | 0.0001-004 |

Table 10. The difference between the expected values depending on sample size

We see that the difference between the expected rates becomes smaller with the increasing sample size. However, it also increases the amount of computations. We choose 60000 samples per strategy as the optimal one with respect to the time needed for the sampling and to the accuracy. Totally it makes about 1.2Gb of samples per 60 strategies. Now we are ready to analyze the model. This is done in the next Chapter.

## CHAPTER 4. ANALYSIS OF THE MODEL

In this Chapter it will be shown how the Model can be used for analyzing investments into defense systems. Firstly, the one year model is considered. An investor expects in it that the investment will pay itself back and bring profit in the one year from now. The Model has shown investing into internet-based intrusion detection systems (IDS) to be the best strategy with respect to return and ROI. It will be investigated that the three year model also approves this strategy as the best one with respect to NPV of the project.

### 4.1. Analysis of strategies, one year pay-back period

Let us check if any of sixty strategies can bring profit in case of high asset costs and high exposure factor. The cost of disclosure and the reduction in the expected attack rate are the factors which generate income through prevented disclosures. Given a strategy, if the return is negative for the highest price, it will be also negative for any lower cost of disclosure. If there are profitable strategies for the highest price, we will go to the lower levels until no strategy is profitable.

The model says that if a company would like to be 95% sure that the investment into defense systems will pay for itself, the disclosure cost of “high-high” information assets (100 million for every disclosure) is not high enough to justify investments. All 5% quantiles of the return are negative. According to the expert, the left end of the attack rate for zero is close to zero. Together with the assumption that all prevented disclosures are prevented at all successive levels, it makes the sample with small base rate. It means that only few disclosures can be prevented. Hence, income in such cases will not justify the investments. We have more than 5% of such cases. This is why the Model shows that if the company would like to be 95% sure that the investment will pay itself back, it is more profitable to pay for every disclosure than to invest into defense systems.

If a company would like to be 75% confident[5] that the investments into defense systems will be compensated, there exist favorable strategies (highlighted in "return, 25% quantile" column):

| Human shields | Barriers | IDS | COST, $ | RETURN 25% QUANTILE, $/year | ROI, $/$year | RETURN 50% QUANTILE, S/year | EXPECTED RETURN, $/year | EXPECTED ROI, $/$year |
|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 1 | 5000 | 4.45E+04 | 8.90E+00 | 6.90E+04 | 5.96E+04 | 1.19E+01 |
| 1 | 0 | 1 | 12000 | 3.75E+04 | 3.13E+00 | 6.20E+04 | 5.24E+04 | 4.37E+00 |
| 2 | 0 | 1 | 32000 | 1.78E+04 | 5.55E-01 | 4.23E+04 | 3.28E+04 | 1.03E+00 |
| 0 | 0 | 2 | 35000 | 1.47E+04 | 4.20E-01 | 3.93E+04 | 2.99E+04 | 8.54E-01 |
| 1 | 0 | 2 | 42000 | 7.70E+03 | 1.83E-01 | 3.23E+04 | 2.27E+04 | 5.41E-01 |
| 2 | 0 | 2 | 62000 | -1.23E+04 | -1.98E-01 | 1.25E+04 | 3.09E+03 | 4.98E-02 |
| 0 | 0 | 3 | 85000 | -3.52E+04 | -4.14E-01 | -1.03E+04 | -1.97E+04 | -2.31E-01 |
| 1 | 0 | 3 | 92000 | -4.22E+04 | -4.59E-01 | -1.73E+04 | -2.68E+04 | -2.92E-01 |
| 2 | 0 | 3 | 112000 | -6.22E+04 | -5.55E-01 | -3.72E+04 | -4.66E+04 | -4.16E-01 |
| 0 | 1 | 0 | 15000 | 3.45E+04 | 2.30E+00 | 5.90E+04 | 4.96E+04 | 3.31E+00 |
| 1 | 1 | 0 | 22000 | 2.75E+04 | 1.25E+00 | 5.20E+04 | 4.24E+04 | 1.93E+00 |
| 2 | 1 | 0 | 42000 | 7.75E+03 | 1.85E-01 | 3.23E+04 | 2.28E+04 | 5.44E-01 |
| 0 | 1 | 1 | 20000 | 3.00E+04 | 1.50E+00 | 5.50E+04 | 4.56E+04 | 2.28E+00 |
| 1 | 1 | 1 | 27000 | 2.30E+04 | 8.52E-01 | 4.80E+04 | 3.84E+04 | 1.42E+00 |
| 2 | 1 | 1 | 47000 | 3.00E+03 | 6.38E-02 | 2.80E+04 | 1.86E+04 | 3.96E-01 |
| 0 | 1 | 2 | 50000 | -3.00E+00 | -6.00E-05 | 2.50E+04 | 1.56E+04 | 3.12E-01 |
| 1 | 1 | 2 | 57000 | -7.00E+03 | -1.23E-01 | 1.80E+04 | 8.44E+03 | 1.48E-01 |

Table 11. Strategies with positive expected return

The expected return is lower than its 50% quantile. This means that the left tail of return distribution is longer. There is a chance that the return will be much less than its expected value, though value in more than half of the cases it is bigger than its expected (column "Expected Return").

Some unprofitable strategies are also presented in table 11 for the sake of demonstrating if the investments into human shields given barriers and intrusion detection systems should be made. As we see, for espionage attacks the investment into training/certification does not increase the return given *(bar,id)*. This effect can be seen and explained from expert's distributions. From rate assessments, there is almost no influence of the investment into human shields on distribution of successful attack rate. This influence appears only on higher levels of Barriers (and Enclaves) and an intrusion detection system. High levels of investments are required then, and the decrease in rate of attack is not justified with the increase in investments. Also, given a barrier/intrusion detection strategy, investments into certification of key personnel do not make the return

[5] 75% quantile is used in many expert judgment studies, for this reason we use it here. Any other quantile can also be chosen.

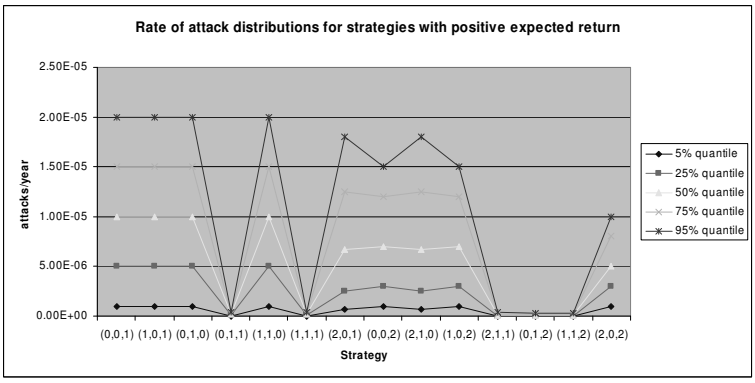
on investment negative, but it reduces it. This effect can be also explained by already trained personnel (according to the FBI survey [25], more than half of companies consider training to be important) and by the specific type of attacks considered (espionage).

From table 11, ROI decreases with investing in human shields, given levels of barriers and intrusion detection. It does not mean though that the company gets less return with the decrease in risk of secret disclosure. Let us take a look at the following rate distributions:

| Strategy (hum,bar,id) | Expected attacks/year | Attacks/year, 5% quantile | Attacks/year, 25% quantile | Attacks/year, 50% quantile | Attacks/year, 75% quantile | Attacks/year, 95% quantile | COST, $ |
|---|---|---|---|---|---|---|---|
| (0,0,1) | 1.02E-05 | 1.00E-06 | 5.00E-06 | 1.00E-05 | 1.50E-05 | 2.00E-05 | 5000 |
| (1,0,1) | 1.02E-05 | 1.00E-06 | 5.00E-06 | 1.00E-05 | 1.50E-05 | 2.00E-05 | 12000 |
| (0,1,0) | 1.02E-05 | 1.00E-06 | 5.00E-06 | 1.00E-05 | 1.50E-05 | 2.00E-05 | 15000 |
| (0,1,1) | 1.94E-07 | 3.38E-08 | 6.55E-08 | 1.55E-07 | 2.78E-07 | 4.22E-07 | 20000 |
| (1,1,0) | 1.02E-05 | 1.00E-06 | 5.00E-06 | 1.00E-05 | 1.50E-05 | 2.00E-05 | 22000 |
| (1,1,1) | 1.94E-07 | 3.37E-08 | 6.55E-08 | 1.55E-07 | 2.78E-07 | 4.22E-07 | 27000 |
| (2,0,1) | 7.84E-06 | 7.12E-07 | 2.50E-06 | 6.67E-06 | 1.25E-05 | 1.80E-05 | 32000 |
| (0,0,2) | 7.55E-06 | 1.00E-06 | 3.00E-06 | 7.00E-06 | 1.20E-05 | 1.50E-05 | 35000 |
| (2,1,0) | 7.84E-06 | 7.12E-07 | 2.50E-06 | 6.67E-06 | 1.25E-05 | 1.80E-05 | 42000 |
| (1,0,2) | 7.53E-06 | 1.00E-06 | 3.00E-06 | 7.00E-06 | 1.20E-05 | 1.50E-05 | 42000 |
| (2,1,1) | 1.59E-07 | 1.71E-08 | 3.80E-08 | 1.09E-07 | 2.36E-07 | 3.86E-07 | 47000 |
| (0,1,2) | 1.53E-07 | 2.39E-08 | 4.63E-08 | 1.13E-07 | 2.19E-07 | 3.21E-07 | 50000 |
| (1,1,2) | 1.52E-07 | 2.39E-08 | 4.38E-08 | 1.13E-07 | 2.19E-07 | 3.21E-07 | 57000 |
| (2,0,2) | 5.38E-06 | 1.00E-06 | 3.00E-06 | 5.00E-06 | 8.00E-06 | 1.00E-05 | |

Table 12. Successful attack rate distributions for strategies with positive return

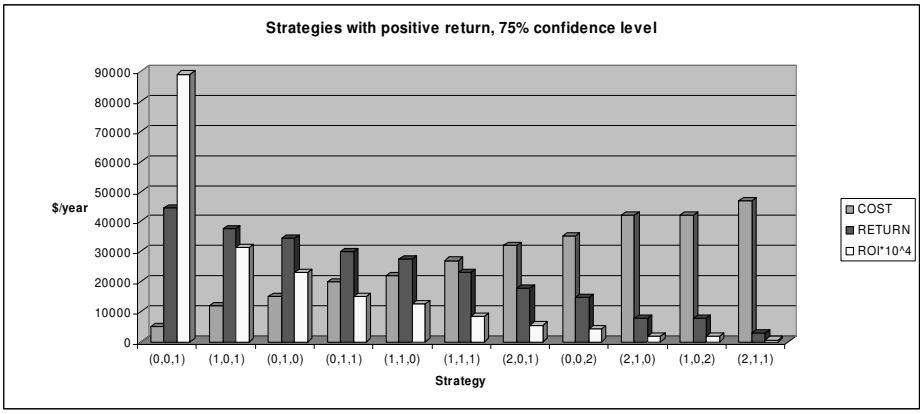
or, graphically,




Plot 7. Rate distributions for strategies with positive return

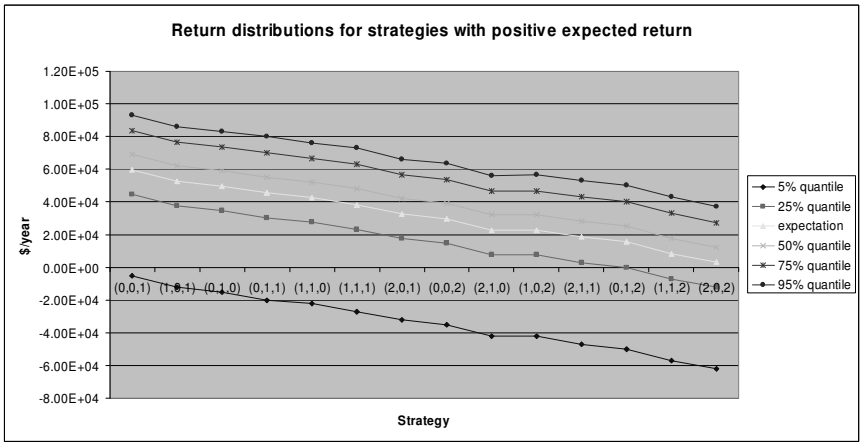
Consider the second strategy (1,0,1). It includes internet-based IDS and personnel training in Infosec risks. From the experts' answers, the distribution of successful attack

rate when no defense systems are implemented is just the same as when personnel have had training in InfoSec risks. The rate distribution (and hence profit) is the same, but expenditures increase. Hence, the decrease in return and ROI (plot 8) for the second strategy (and also the third) is caused only by the increasing expenditures. Investment strategies (1,0,1) and (0,1,0) are not recommended, if there is possibility to implement strategy (0,0,1). Also, for any given barrier-intrusion detection strategy from table 11, the investment into training personnel alone makes absolutely no sense, but together with certification it gives extra reduction in the expected attack rate, though this reduction is not economically plausible. All profitable strategies and their return distributions are given in the graph below.



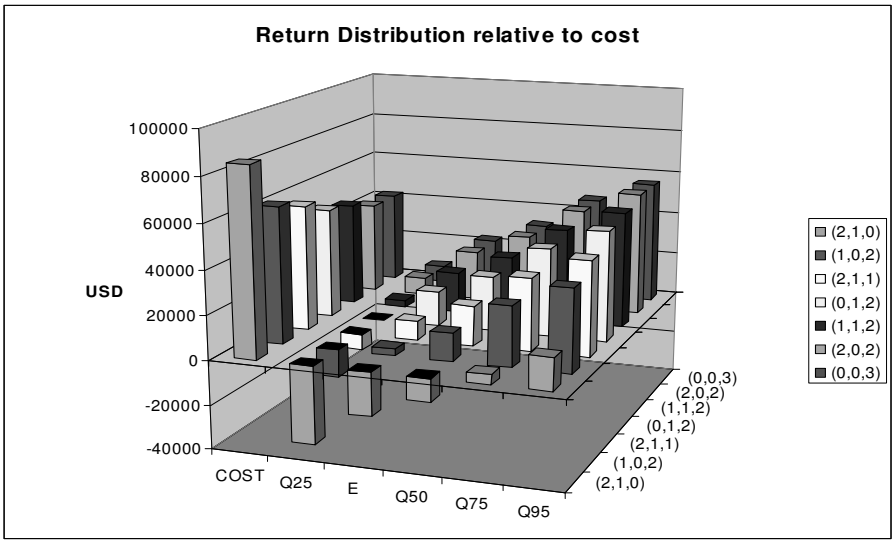

Plot 8. Strategies with positive return, 75% confidence level

As we can see from the Plot 8, the strategy of investing into network-based intrusion detection systems gives the highest return and ROI. Every dollar invested in it brings almost 9 dollars back. This is incredibly profitable. Plot 9 shows that the chance to cause the loss is around 5 and 10% for this strategy.



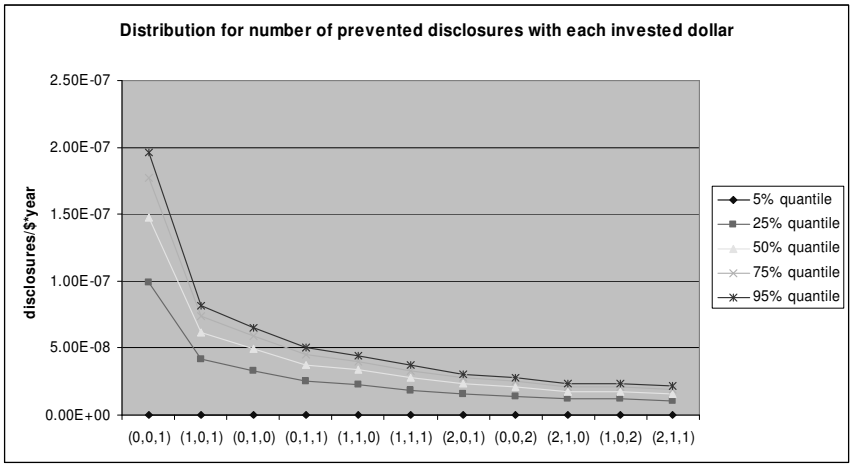

Plot 9. Return distributions for strategies with positive return

Plot 9 shows the return distribution for different strategies. We can read from it, for example, that strategy (0,1,2) has non-negative return with 75% confidence level and that the return is at least $40,000 with 25% confidence level. The lower is the confidence level, the higher the return is. Together with the investments, it gives the following picture:




Plot 10. Strategies on the edge of changing to ones with positive expected return

The next plot says that the higher is the level of defense systems (and hence the price), the less disclosures are prevented with each invested dollar, i.e. extra prevented disclosures (if there are ones) cost more.




Plot 11. Number of prevented disclosures with each dollar invested

Setting the cost per disclosure one level lower (medium-high information assets, 10 millions per disclosure) finalizes the search for profitable strategies. It gives no profitable strategies for 95% and 75% confidence levels. The expected return (and return for 50% confidence level) is positive only for the strategy of investing into internet-based intrusion detection systems (0,0,1). Its ROI is 29.2%, which is still very high. Exact statistics for this strategy is given in the table below:

| Percentile | Return, $/year | Corresponding ROI, $/Syear | Rate of attack, attacks/year |
|---|---|---|---|
| 5 | -5.00E+03 | -1 | 1.00E-06 |
| 25 | -5.00E+01 | -0.01 | 5.00E-06 |
| 50 | 2.40E+03 | 0.48 | 1.00E-05 |
| 75 | 3.85E+03 | 0.77 | 1.50E-05 |
| 95 | 4.80E+03 | 0.96 | 4.80E+03 |
| Expectation | 1.46E+03 | 0.292 | 1.02E-05 |

Table 13. Statistics for strategy (0,0,1), medium-high information assets

The result is interpretable. Loss-of-Secrecy events which happen due to espionage are

| | **Cost, 5% quantile** | **Cost, 25% quantile** | **Cost, 50% quantile** | **Cost, 75% quantile** | **Cost, 95% quantile** |
|---|---|---|---|---|---|
| (0,0,1) | 5.10E+06 | 5.65E+06 | 6.76E+06 | **1.01E+07** | 1.00E+15 |
| (1,0,1) | 1.22E+07 | 1.36E+07 | 1.62E+07 | 2.42E+07 | 1.00E+15 |
| (0,1,0) | 1.53E+07 | 1.69E+07 | 2.03E+07 | 3.03E+07 | 1.00E+15 |
| (0,1,1) | 2.00E+07 | 2.22E+07 | 2.67E+07 | 4.00E+07 | 1.00E+15 |
| (1,1,0) | 2.24E+07 | 2.49E+07 | 2.97E+07 | 4.44E+07 | 1.00E+15 |
| (1,1,1) | 2.70E+07 | 3.00E+07 | 3.60E+07 | 5.40E+07 | 1.00E+15 |
| (2,0,1) | 3.26E+07 | 3.61E+07 | 4.30E+07 | 6.43E+07 | 1.00E+15 |
| (0,0,2) | 3.55E+07 | 3.94E+07 | 4.71E+07 | 7.04E+07 | 1.00E+15 |
| (2,1,0) | 4.28E+07 | 4.73E+07 | 5.65E+07 | 8.44E+07 | 1.00E+15 |
| (1,0,2) | 4.26E+07 | 4.73E+07 | 5.65E+07 | 8.45E+07 | 1.00E+15 |
| (2,1,1) | 4.70E+07 | 5.22E+07 | 6.27E+07 | 9.40E+07 | 1.00E+15 |

| | | | | | |
|---|---|---|---|---|---|
| (0,1,2) | 5.00E+07 | 5.56E+07 | 6.67E+07 | 1.00E+08 | 1.00E+15 |
| (1,1,2) | 5.70E+07 | 6.33E+07 | 7.60E+07 | 1.14E+08 | 1.00E+15 |
| (2,0,2) | 6.26E+07 | 6.95E+07 | 8.32E+07 | 1.25E+08 | 1.00E+15 |

Table 14. Distribution for minimal cost per disclosure (return=0)

rare. Hence, cheap secrets can not pay for themselves. Either the price of disclosure or its probability should be high to justify the investment.

Let us take a look at the minimal disclosure cost distribution (at which the strategies with positive return becomes profitable, table 14). Let the cost per disclosure is its 75% quantile for strategy (0,0,1) (highlighted in table 14). This means that in the 75% cases the cost, at which investment into internet-based intrusion detection system becomes profitable, is less than $10,102,000. Hence, in the 75% cases the return should be non-negative. We put this cost into the model and run it to compute statistics. The Model shows that in this case the 25% quantile of the return is around 2 dollars and the expected return is $1,530. Hence, computations in the model are consistent.

The cost of disclosure at which the expected return becomes so that *ROI=r* can be found.

Denote:

$P_A$ - price of strategy A=(0,0,1)

$D_1$ - price of one disclosure

r - refinance rate, set to 0.07

$R_{ZERO}$ - rate of attack if no extra defense systems is purchased

$R_A$ - rate of attach after implementing A

Then

$$\text{Return} = Income - Investment = (R_{ZERO} - R_A)D_1 - P_A \geq rP_A$$

$$(1+r)P_A = D_{MIN}(ER_{ZERO} - ER_A)$$

$$D \geq D_{MIN} = \frac{(1+0.07)*5000}{(0.00065634 - 0.000010189)} = 8.2798e+006$$

Setting the cost of disclosure to $8.3 million, the following statistics is obtained:

| Percentile | Return, $/year | Corresponding ROI, $/$year |
|---|---|---|
| 5 | -5.00E+03 | -1 |
| 25 | -8.92E+02 | -0.1784 |
| 50 | 1.14E+03 | 0.228 |

| | | |
|---|---|---|
| 75 | 2.35E+03 | 0.47 |
| 95 | 3.13E+03 | 0.626 |
| Mean | 362.95 | 0.052 |

Table 15. Statistics for disclosure cost which brings minimum positive expected return

We see that the price $8.3 million is the approximate cost at which the expected return equals the income from investment to the bank (instead of investing into defense systems). If the price of one successful attack resulted in unwanted data disclosure is less than this price, then purchasing any of defense systems is not plausible. The next step is to extend the model for a few years. Three-year model is considered in the next Subsection.

### 4.2. Analysis of strategies, three year period

The software developed for the Project allows computing statistics when the life cycle of the investment into defense systems is between one and six years. It uses the expert's assessments of successful attack rate for the year starting in five years from now for forecasting. To check if software in the second model (we are going to analyze three year investment project) works fine, the life cycle of the Project is set to one year first. We expect to get the same results as we have already got and described in the previous subsection. Fixing barrier level to "DO NOTHING", ID level to "Network-Based IDS" and the cost of one disclosure to $10 million, the following output is obtained:

| | |
|---|---|
| ########################################<br>## STRATEGY BAR=0, ID=1 ##<br>########################################<br>----------HUM=0--------HUM=1---------HUM=2<br>**MEAN1=1.0135e-005; 1.0171e-005; 7.7963e-006**<br>MEAN2=9.6088e-006; 9.6432e-006; 7.4026e-006<br>MEAN3=9.0828e-006; 9.1153e-006; 7.0090e-006<br>MEAN4=8.5569e-006; 8.5874e-006; 6.6153e-006<br>MEAN5=8.0309e-006; 8.0595e-006; 6.2216e-006<br>MEAN6=7.5049e-006; 7.5316e-006; 5.8280e-006<br>E_RET=**1.0129e+003;-5.9705e+003;-2.5957e+004**<br>INVES=5.0000e+003; 1.2000e+004; 3.2000e+004<br>PRC75=1.0808e+007; 2.5939e+007; 6.8824e+007<br>SIZE=6.5865e+004; 6.6075e+004; 6.8047e+004 | ########################################<br>## STRATEGY BAR=0, ID=1 ##<br>########################################<br>----------HUM=0--------HUM=1---------HUM=2<br>**ERATE=1.0189e-005; 1.0164e-005; 7.8402e-006**<br>DEV=6.0219e-006; 6.0359e-006; 5.7070e-006<br>Q5=1.0000e-006; 1.0000e-006; 7.1238e-007<br>Q25=5.0000e-006; 5.0000e-006; 2.5000e-006<br>Q50=1.0000e-005; 1.0000e-005; 6.6667e-006<br>Q75=1.5000e-005; 1.5000e-005; 1.2500e-005<br>Q95=2.0000e-005; 2.0000e-005; 1.8000e-005<br>INVES=5.0000e+003; 1.2000e+004; 3.2000e+004<br>E_RET=**1.4614e+003;-5.5558e+003;-2.5516e+004**<br>SIZE=6.6009e+004; 6.6244e+004; 6.7736e+004 |
| Output of the second model | Output of the first model |

Table 16. Output for two models

Datasets for two models are different, so the expected attack rates are not exactly the same. There is noticeable difference in expected return, which can be explained. For first model, we count return as (expected rate of zero strategy is approximately 6.5634e-004):

$$E(\mathrm{Re}\,turn) = (ER_{ZERO} - ER_A)D_1 - P_A \geq rP_A$$

$$E(\mathrm{Re}\,turn) \approx (656.3 - 10.2) * 10 - 5000 = 1461$$

For the second model (refinance rate[6] is assumed to be 0.07),

$$E(\mathrm{Re}\,turn) = \sum_i E\left(\frac{D_1(R^i_{ZERO} - R^i_A)}{(1+r)^i}\right) - P_A \geq 0$$

or

$$E(\mathrm{Re}\,turn) \approx \frac{(656.3 - 10.1) * 10}{1.07} - 5000 \approx 1039$$

[6] *Refinance rate* is set by Central Bank and states for the rate at which banks are allowed to borrow money from Central Bank. It is often used for computing criteria of investment projects as a discount rate.

Hence, the difference is as follows: in first model, the time value of money is not taken into account while computing the return. ROI has the meaning of how much profit each dollar invested now will bring in one year. For example, if a bank is suggesting deposit at rate (or ROI) 0.09, it means than one dollar invested now will bring 0.09 dollar (besides giving back the deposit) in one year. For this reason, we compare the return with value $rP_A$ (the same as demanding $ROI>r$). In the second model, the future inflows are discounted to the present value of money and return is required to be positive. In fact, this return is the *Net Present Value* of the investment strategy. ROI in the second model shows how much does each invested dollar will bring back in today's value of money. For example, if bank gives $1.09 back in one year for depositing one dollar now, ROI=0.09 (in the context of first model). If refinance rate is 0.07 per year, then ROI in the second model is (1.09-1)/[1*(1+0.07)]. The requirements for plausibility of the strategy in both models (with one-year project life period) are the similar in both models:

$$ROI = \frac{(ER_{ZERO} - ER_A)D_1 - P_A}{P_A} \geq r$$

$$\text{Re}turn = (ER_{ZERO} - ER_A)D_1 - P_A \geq rP_A$$

$$NPV = \frac{(ER_{ZERO} - ER_A)D_1}{1+r} - P_A \geq 0$$

Also, in the second model, expected rates for the sixth year correspond to expert's assessments from table 7. Thus, this model gives sensible results. We analyze it for the three year lifetime period.

**4.2.1. High-high information assets**

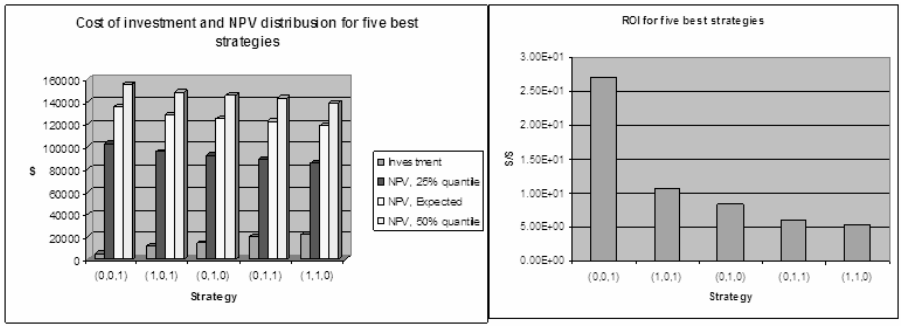
Firstly, consider the "high-high" type of information assets. Totally 33 strategies have positive expected NPV at the disclosure price of 100 million. We analyze five strategies with the highest and five strategies with the lowest positive NPV and discuss their plausibility. The following criteria are taken into consideration:

- 25% quantile and the expected NPV (so that to be 75% sure that the NPV is higher than this value)
- The expected ROI
- The rate of attack depending on chosen strategy

These values for all 33 strategies with positive NPV are given in the Appendix.

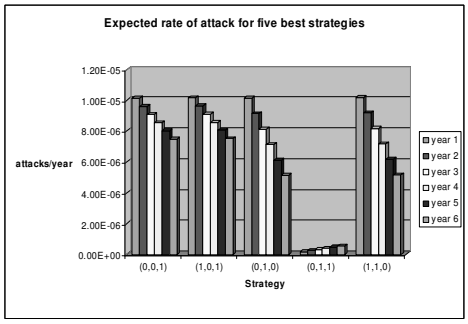
**Five "best" strategies**

The NPV distribution and expected ROI are given on Plot 12 and 13.




Plot 12, 13. NPV and ROI for five best strategies

The strategy with the highest NPV corresponds to the smallest investment. This could be different if such an investment would not give a sufficient reduction in the rate of attack. All five strategies have positive 25% quantile, which means that the NPV is positive in more than 75% of the cases. Let us take a look at expected successful attack rates:




Plot X13. Expected attack rates for five "best" strategies

We see that only the investment in internet/intranet segregation together with internet based IDS (fourth strategy) gives a good decrease in the rate of attack comparing to the strategy with highest NPV. Hence, if a company would like to better reduce number of attacks, such a strategy is a good choice. Strategies other than first and fourth one are not recommended: they give almost no reduction in rate, but are more expansive than the first strategy.

The reader should not be surprised with the increasing expected rate of successful attack. Consider the example. Suppose, in the first year and in the year starting in five years from now, two independent strategies A and B give the following decrease in the rate of attack (per million of comparable information systems):

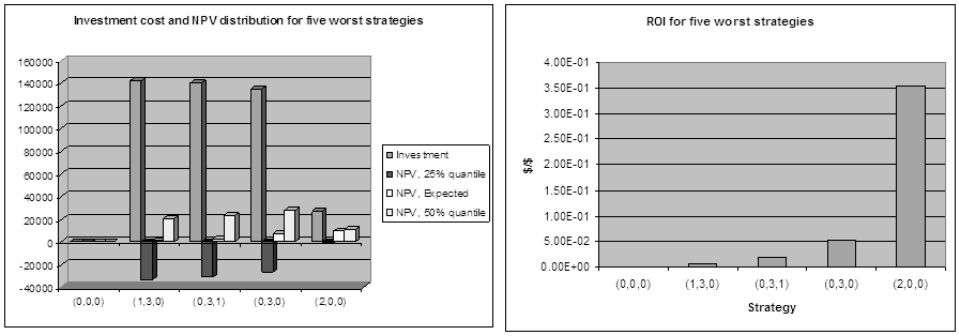
| | Zero | A | B |
|---|---|---|---|
| 1st year | 750 | 10 | 10 |
| 6th year | 75 | 7 | 5 |

Then the rate of implementing A and B is:

| | 1st year | 6th year |
|---|---|---|
| A and B | $750 \times \frac{10}{750} \times \frac{10}{750} = 0.133$ | $75 \times \frac{7}{75} \times \frac{5}{75} = 0.466$ |

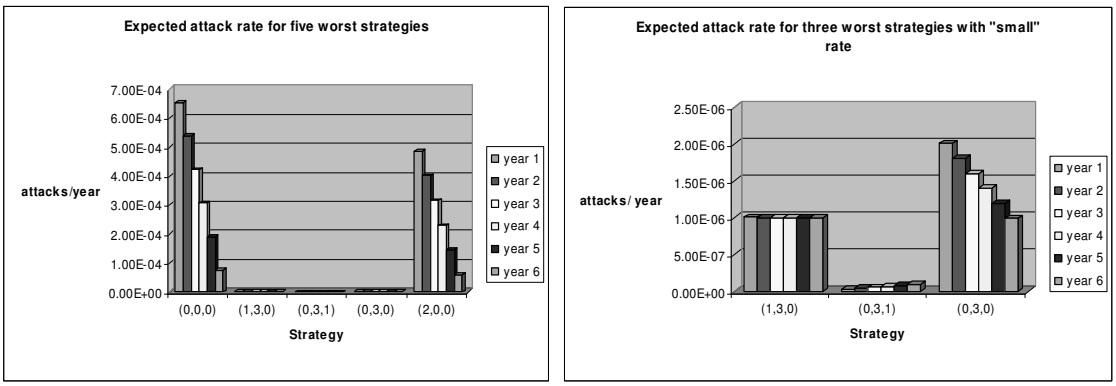
In the first year, the influence of A and B on the base rate is much higher. Thus, even if the attack rate decreases in time for A and for B separately, it can grow if both A and B are implemented. Five strategies with the highest NPV are also five strategies with the highest ROI.

**Five "worst" strategies**



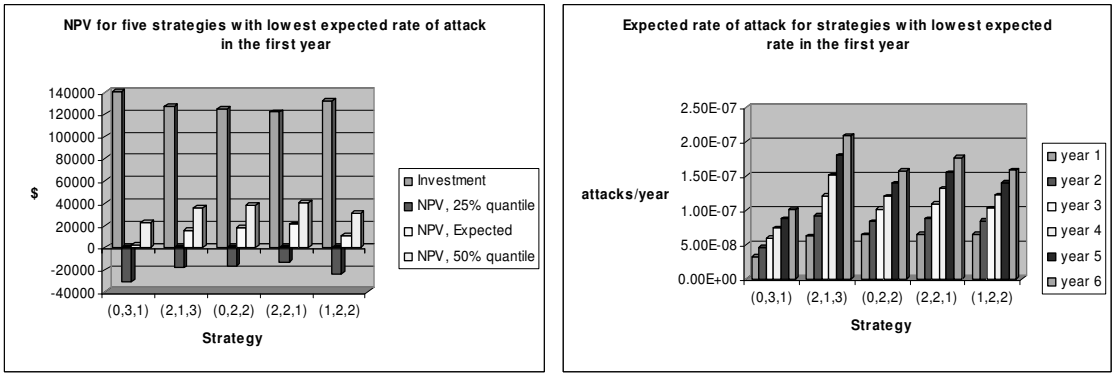

Plot 15, 16. Expected attack rates for five "worst" suitable strategies

The strategy with zero expected NPV depicted on plot 15 is zero strategy. If nothing is invested into defense systems, then there is no income from the "improvement". It can be seen from plots 15 and 16 that ROI does not necessary decrease with growing investments. Investing into training with certification of key personnel is much cheaper than purchasing the third level of barriers and enclaves, but it has the highest ROI among this five strategies. Also, only investing into training with certification of key personnel guarantees positive NPV in the 75% cases among five "worst" strategies.




Plot 17, 18. Expected attack rates for five "best" strategies

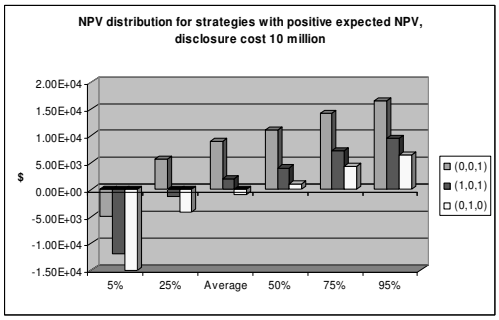
It can be seen from plots 17 and 18 that the strategy is sensible not only if it gives a significant reduction in the rate of attack. The amount of investments and its effect are both important. The rate of attack in the next six years can either stay constant (strategy (1,3,0)), growing (strategy (0,3,1)) or decreasing (the rest). If the company decides to choose a strategy with small NPV, it means that this company may not be driven by economic values. Some data (personal details or army secrets) may worth people's lives. For this reason, strategy of investing into space gaps and personnel training out of the five worst strategies is recommended. On the other hand, in order to find the strategy giving the highest decrease of expected attack rate it is better to look though all strategies with positive NPV. Five investment decisions which give the highest decrease in expected rate of attack in the first year are given on plots 19 and 20.




Plot 19,20. Five strategies with the highest decrease of expected attack rate

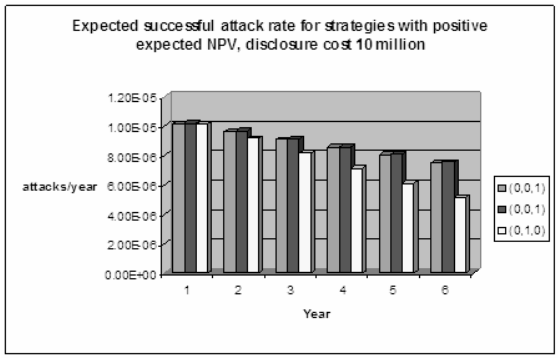
### 4.2.2. Medium-high information assets

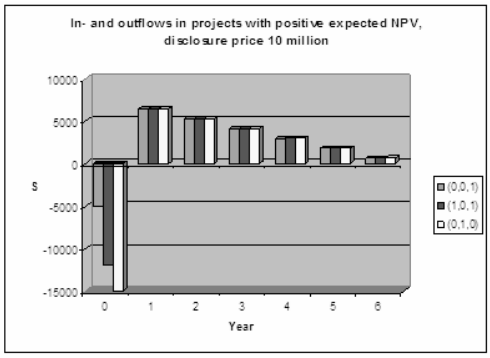
Consider "medium-high" and "high-medium" types of information assets. The assessed cost of one unwanted data disclosure is set to 10 million for them.




Plot 21. NPV distribution for three strategies

There are two strategies with positive expected NPV (Plot 21). We also add the strategy with minimum negative expected NPV to the analysis. On Plot 23 expected inflows and outflows for these strategies are given. In year 0 the company buys defense systems (outflows) and in the next six years it gets the income from it, represented by the reduction in potential loss. Extra yearly expenditures caused by the use of extra defense systems are not taken into consideration. Hence, in year from 1 to 6 we get only inflows from the project. As we see from plot 23, these inflows for strategies are approximately the same. For strategy (1,0,1) this happens because training has no influence on effect of internet-based IDS. It may seem to be a mistake that its expected rate of attack is decreasing faster for third strategy than for the first two ones, while their inflows are all similar. This happens because relatively to the base rate, all three strategies give the significant reduction in expected attack rate of approximately one order. Difference between rates is noticeable, but all of them are small comparing to the base rate. This is why the inflows for all three projects are similar.




Plot 22, 23. Expected flows and attack rates for three strategies

In the long run, internet/intranet segregation performs better than Network-based IDS. For this reason, the income from investing into (0,1,0) on the sixth year is *a bit* bigger than from two other strategies. The conclusion is obvious: if the effect is approximately the same, there is no sense in paying more. Investing into internet-based IDS is the "best" (most effective with respect to NPV) strategy. We do not need ROI value to decide if the strategy is more profitable than investing into a bank, because all flows in NPV are already discounted. Nevertheless, this parameter may help estimating how better the investment is. ROI corresponding to the expected NPV for the 'best' project, depending on disclosure cost is given in the table 17.

| Disclosure cost | NPV, 25% quantile | Expected NPV | Expected ROI |
|---|---|---|---|
| 100 000 000 | 102450 | 134990 | 26.998 |
| 10 000 000 | 5745 | 8999 | 1.7998 |
| 4 155 000 | -535.47 | 816.6 | 0.16332 |
| 3 572 000 | -1161.9 | 0.45878 | 0.000091755 |

Table 17. Expected NPV and ROI for strategy (0,0,1)

For the strategies which give positive expected NPV at the disclosure cost of $10 million, let us find the minimal cost per disclosure such that the expected NPV becomes around zero, i.e. the cost at which the project on the average becomes profitable.

$$\sum_i E\left(\frac{D_{MIN}(R^i_{ZERO} - R^i_A)}{(1+r)^i}\right) - P_A = 0$$

$$D_{MIN} = \frac{P_A}{\sum_{i=1}^{3}\left(\frac{ER^i_{ZERO} - ER^i_A}{(1+r)^i}\right)}$$

The expected yearly rates are given in the following table

| Strategy | Expected attack rate, attacks/year, in the *i*-th year from now | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| (0,0,0) | 6.54E-04 | 5.38E-04 | 4.22E-04 | 3.07E-04 | 1.91E-04 | 7.52E-05 |
| (0,0,1) | 1.01E-05 | 9.61E-06 | 9.08E-06 | 8.56E-06 | 8.03E-06 | 7.50E-06 |
| (1,0,1) | 1.02E-05 | 9.64E-06 | 9.12E-06 | 8.59E-06 | 8.06E-06 | 7.53E-06 |
| (0,1,0) | 1.01E-05 | 9.14E-06 | 8.14E-06 | 7.14E-06 | 6.13E-06 | 5.13E-06 |

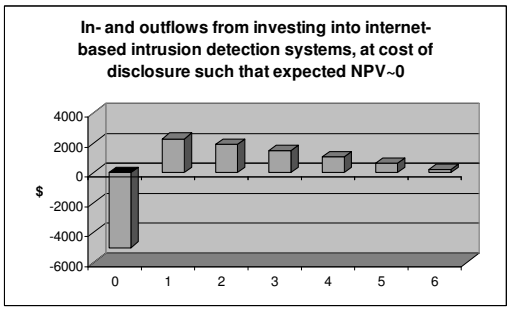
Table 18. Expected yearly attack rates for three strategies

From table 18 (it is important not to forget that the length of the project is set to three years),

| Human shields | Barriers | IDS | MINIMAL COST, $/disclosure |
|---|---|---|---|
| 0 | 0 | 1 | 3.57176E+06 |
| 1 | 0 | 1 | 8.57277E+06 |
| 0 | 1 | 0 | 1.07063E+07 |

Table 19. Minimal cost of one disclosure for three strategies

The expected flows for the strategy (0,0,1) at disclosure cost of $3.572 million are:




Plot 24. Expected flows for strategy (0,0,1) at disclosure price 3.572 million

We would like to take more strategies into consideration while deciding about investing into defense systems. For this reason, the "life period" of the project is extended to six years and the Model at the disclosure cost of 10 million is recalculated. Results are presented in the table 20:

| Human shields | Barriers | IDS | Investment, $ | Expected NPV, $ | Expected ROI, $/$ |
|---|---|---|---|---|---|
| 0 | 0 | 1 | 5000 | 1.30E+04 | 2.61E+00 |
| 1 | 0 | 1 | 12000 | 6.08E+03 | 5.07E-01 |
| 0 | 1 | 0 | 15000 | 3.10E+03 | 2.07E-01 |
| 0 | 0 | 0 | 0 | 0.00E+00 | 0.00E+00 |
| 0 | 1 | 1 | 20000 | -1.5E+03 | -7.7E-2 |

Table 20. Five strategies with highest NPV, six-year period

Obviously, investing into Intrusion-based IDS is the “best” strategy with respect to the NPV and the ROI again. In the next Chapter we draw the conclusions and discuss possible perspectives of the Project.

## CHAPTER 5. CONCLUSIONS AND PERSPECTIVES

The model has approved to be a simple tool of comparing strategies of making investments into information security defense systems. It allows to easily computing such parameters of investment project as return, ROI and NPV. Also, distributions for attack rates in the next six years has been built and studied. Making investments into defense system has appeared to be plausible only for “high-high”, “medium-high” and “high-medium” information assets, when the cost of successful attack resulted in unwanted data disclosure is high. The most efficient strategy is investing into internet-based intrusion detection systems. This strategy has appeared to have very high ROI, which means that investing into defense systems can be much more profitable than investing into financial markets. In addition, minimum cost per attack which makes a strategy profitable bas been assessed. Within investing into internet-based intrusion detection systems, secrets with disclosure cost from 8.3 million dollars for one year project and from 3.6 million for three year project keeps expected return positive and investment is more profitable than making a deposit at 7% rate.

However, the results will be more rational if assessments from more than one expert are used. Other experts may be disagreeing with the assessments used in the Project. Rational consensus would be archived though combining experts’ assessments with the Classical model. Besides, the model requires big amount of samples, and sample size has been taken with respect to both time of calculations and preciseness. It will take more time to compute the Model if preciseness is the priority. Also, such assumptions as the constant cost of successful attack resulted into unwanted data disclosure, constant discounting rate, absence of extra expenditures caused by investing into defense systems and absence of investments in the following five years can be released.

## REFERENCES


1. Peter F. Drucker. Managing in the Next Society. Truman Talley Books, 2002. ISBN: 0750656247
2. Hal Varian and Peter Lyman. “How Much Information?". Report. http://www.berkeley.edu
3. V.I. Rabinovich “Mendeleev” http://www.pseudology.org/Rabinovich_VI
4. http://www.wordreference.com
5. http://www.sei.cmu.edu/str/indexes/glossary/security.html
6. http://www.medterms.com/script/main/art.asp?articlekey=32945
7. The US Constitution online, http://www.usconstitution.net
8. McDaniel, George, ed. *IBM Dictionary of Computing.* New York, NY: McGraw-Hill, Inc., 1994
9. Dorothy E.Denning. Information Warfare and Security. Addison-Wesley,1999
10. Deborah Russell, G.T. Gangemi Sr. Computer Security Basics. O’Reilly % Associates, Inc., 1991
11. Randall K. Nichols, Daniel J. Ryan & Julie J.C.H.Ryan. Defending Your Digital Assets. RSA Press, 2000
12. /www.investorwords.com
13. http://www.riskglossary.com
14. www.nist.gov
15. http://csrc.nist.gov
16. Information Security Handbook. Ed. Harald F. Tipton, Micki Krause. Auerbach, 2000
17. David J.Marchette. Computer Intrusion Detection and Network Monitoring. A Statistical Viewpoint. Springer, 2001
18. James A. Whittaker, Herbert H. Thompson. How to break software security. Pearson Educstion, 2004
19. Ron Ross en al. Information Security. NIST Special Edition 800-53. February 2005
20. GRi While Paper. How to detect hackers on your web server. www.gfi.com
21. Tim Bedford, Roger Cooke. Probabilistic Risk Analysis. Foundation and Methods. Cambridge University Press, 2001
22. Dorota Kurowicka and Roger Cooke. Uncertainty Analysis with High Dimensional Dependence Modelling. Whiley, 2006
23. Cooke R.M., “Experts in Uncertainty”. Oxford University Press 1991
24. The Economics of Informatino Security Investment. Lawrence A. Gordon and Martin P. Loeb, University of Madrid
25. CSI/FBI Computer Crime and Security Surver 2006. **http://i.cmpnet.com/gocsi/db_area/pdfs/fbi/FBI2006.pdf**
26. Network Security Poll Results, Network Computing, 2000, **img.cmpnet.com/nc/1105/graphics/f22.pdf**
27. 3 arrested in Coca-Cola trade secret scheme, **CNNMoney.com, July 5, 2006** **http://money.cnn.com/2006/07/05/news/companies/coke_pepsi/**
28. **Hacker Invades FBI Computers,** WASHINGTON, D.C., July 6, 2006, CBS, **http://www.cbsnews.com/stories/2006/07/06/national/main1779905.shtml**
29. Report: One hacked OU server should have been offline, 26 June 2006, By Jim Phillips, Athens NEWS Senior Writer, http://seclists.org/lists/isn/2006/Jun/0115.html

# APPENDIX A. Elicitation Protocol

### *Calibration variables*

1. If today you are using a 128-bit cryptographic key, how many bits would you need to use in six years in order to maintain the same level of security as you enjoy today (assuming no transformative leaps in technology or mathematical breakthroughs in cryptanalysis)?

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

2. What percentage of attacks on a system is successful?

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

3. What percentage of successful attacks are detected?

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

4. What percentage of successful attacks are detected AND reported to authorities by managers responsible for the security of the system?

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

5. How long would it take a brute force attack by a single general purpose computer (PC) on a 56-bit cryptographic key to recover the key? (hours)

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

6. How long would it take a brute force attack by a distributed network of computers on a 56-bit cryptographic key to recover the key? (hours)

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

7. What percentage of announced vulnerabilities in computer systems and networks are easy to exploit, requiring only moderate computer skills or readily available tools?

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

8. How many essential security patches did Microsoft release for Windows XP in 2005?

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

9. What is the regular patch release rate (patches per year) of Microsoft?

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

10. How many vulnerabilities were reported by CERT in 2005?

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

11. What has been the annual percentage increase in reported vulnerabilities to CERT from 2000 to 2005? (In other words, on average by what percentage did the number change from year to year during that time frame)

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

12. By the end of 2005, how many products had been evaluated at the highest evaluation level of the CCEVS?

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

#### Variables of Interest

We adopt the following abbreviations for defense strategies

**NO**: No Human shielding, no barricades or enclaves, no intrusion detection

**H. Training** Institute recurring seminars, lectures, workshops etc. on info sec. for all staff

**H. Certification:** Require Info sec certification for all staff with access to system definitions.

**B. inter/intra:** Segregate the inter- and intra-nets.
**B. Firewall** Install firewalls and routers
**B. Space Gap:** Install a spatial separation around encalved systems
**B. Mesh:** Install copper meshing around sensitive enclaved systems.

**ID. NIDs** Network-based Intrusion Detection systems
**ID,HIDs** Host-based Intrusion Detection systems
**ID.Hybrid**

**Each question refers to a hypothetical population of *one million comparable information systems*. Please give your 5, 50, 95%-tiles for the number of unwanted data integrity breaches due to espionage attack, for NEXT year, and for one year, starting FIVE years hence, assuming no further changes to the system, under the stated defense strategy.**

**In each case, ONLY the strategy mentioned is implemented.**

13. **NO** *defense strategy, NEXT year*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

14. *in one year starting FIVE years from now*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

15. **B.inter/intra***, defense strategy, NEXT year*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

16. *in one year starting FIVE years from now*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

17. ***B.firewall*** *defense strategy, NEXT year*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

18. *in one year starting FIVE years from now*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

19. ***B.Space Gap*** *defense strategy, NEXT year*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

20. *in one year starting FIVE years from now*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

21. ***B.Space Gap*** **WITH H.Training** *defense strategy, NEXT year*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

22. *in one year starting FIVE years from now*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

23. ***B.Space Gap*** **WITH H.Certification** *defense strategy, NEXT year*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

24. *in one year starting FIVE years from now*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

25. ***B.mesh*** *defense strategy, NEXT year*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

26. *in one year starting FIVE years from now*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

27. ***ID.NIDs*** *defense strategy, NEXT year*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

28. *in one year starting FIVE years from now*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

29. ***ID.HIDs*** *defense strategy, NEXT year*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

30. *in one year starting FIVE years from now*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

31. ***ID.HIDs*** **WITH H. Training** *defense strategy, NEXT year*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

32. *in one year starting FIVE years from now*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

33. ***ID.HIDs WITH H.certification*** *defense strategy, NEXT year*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

34. *in one year starting FIVE years from now*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

35. ***ID.Hybrid*** *defense strategy, NEXT year*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

36. *in one year starting FIVE years from now*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

37. ***H.Training*** *defense strategy, NEXT year*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

38. *in one year starting FIVE years from now*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

39. ***H.certification*** *defense strategy, NEXT year*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

40. *in one year starting FIVE years from now*

| _____ | _____ | _____ |
|---|---|---|
| 5% | 50% | 95% |

## A2. FORMULA FOR COMPUTING RATES

Denote:
Hum – level of Human Shielding defense system
Bar - level of Barriers and Enclaves defense system
Id - level of Intrusion Detection defense system

Hum=0

| Bar (0 .. 4) | ID (0…3) | Formula for the rate of attack | Terms |
|---|---|---|---|
| bar=i | id=j | $r0 \times \frac{rbi}{r0} \times \frac{rIDj}{r0}$, where $rb0 = r0, rID0 = r0$ | h0brate<br>h0IDrate |

Hum=1

| Bar (0 .. 4) | ID (0…3) | Formula for the rate of attack | Terms |
|---|---|---|---|
| bar<3 | id<2 | $r0 \times \frac{rh1}{r0} \times (\frac{rbi}{r0} \times \frac{rIDj}{r0})$, where $rb0 = r0, rID0 = r0$ | b0ID0hrate<br>h12extra1 |
| bar=3 | id<2 | $r0 \times \frac{rh1b3}{r0} \times (\frac{rIDj}{r0})$ | h12brate<br>h12extra1 |
| bar=4 | id<2 | $[r0 \times \frac{rh1b3}{r0} \times (\frac{rIDj}{r0})] \times \frac{rb4}{rb3}$ | h12brate<br>h12extra1 |
| bar<3 | id=2 | $r0 \times \frac{rh1ID2}{r0} \times (\frac{rbi}{r0})$ | h12IDrate<br>h12extra3 |
| bar<3 | id=3 | $[r0 \times \frac{rh1ID2}{r0} \times (\frac{rbi}{r0})] \times \frac{rID3}{rID2}$ | h12IDrate<br>h12extra3 |
| bar=3 | id=2 | $r0 \times \frac{rh1b3}{r0} \times \frac{rh1ID2}{r0}$ | h12brate<br>h12IDrate |
| bar=4 | id=2 | $r0 \times [\frac{rh1b3}{r0} \times \frac{rb4}{rb3}] \times \frac{rh1ID2}{r0}$ | h12brate<br>h12IDrate |
| bar=3 | id=3 | $r0 \times \frac{rh1b3}{r0} \times [\frac{rh1ID2}{r0} \times \frac{rID3}{rID2}]$ | h12brate<br>h12IDrate |
| bar=4 | id=3 | $r0 \times [\frac{rh1b3}{r0} \times \frac{rb4}{rb3}] \times [\frac{rh1ID2}{r0} \times \frac{rID3}{rID2}]$ | h12brate<br>h12IDrate |

Hum=2

| Bar (0 .. 4) | ID (0…3) | Formula for the rate of attack | Terms |
|---|---|---|---|
| bar<3 | id<2 | $r0 \times \frac{rh2}{r0} \times \frac{rbi}{r0} \times \frac{rIDj}{r0}$, where $rb0 = r0, rID0 = r0$ | b0ID0hrate<br>h12extra1 |

| | | | |
|---|---|---|---|
| bar=3 | id<2 | $r0\times\frac{rh2b3}{r0}\times\frac{rIDj}{r0}$ | h12brate<br>h12extra1 |
| bar=4 | id<2 | $[r0\times\frac{rh2b3}{r0}\times\frac{rIDj}{r0}]\times\frac{rb4}{rb3}$ | h12brate<br>h12extra1 |
| bar<3 | id=2 | $r0\times\frac{rh2ID2}{r0}\times\frac{rbi}{r0}$ | h12IDrate<br>h12extra3 |
| bar<3 | id=3 | $[r0\times\frac{rh2ID2}{r0}\times\frac{rbi}{r0}]\times\frac{rID3}{rID2}$ | h12IDrate<br>h12extra3 |
| bar=3 | id=2 | $r0\times\frac{rh2b3}{r0}\times\frac{rh2ID2}{r0}$ | h12brate<br>h12IDrate |
| bar=4 | id=2 | $r0\times[\frac{rh2b3}{r0}\times\frac{rb4}{rb3}]\times\frac{rh2ID2}{r0}$ | h12brate<br>h12IDrate |
| bar=3 | id=3 | $r0\times\frac{rh2b3}{r0}\times[\frac{rh2ID2}{r0}\times\frac{rID3}{rID2}]$ | h12brate<br>h12IDrate |
| bar=4 | id=3 | $r0\times[\frac{rh2b3}{r0}\times\frac{rb4}{rb3}]\times[\frac{rh2ID2}{r0}\times\frac{rID3}{rID2}]$ | h12brate<br>h12IDrate |

## A3. DICTIONARY (taken from the Web)

### Security

1. Freedom from risk or danger; safety.
2. Freedom from doubt, anxiety, or fear; confidence.
3. Something that gives or assures safety, as:
    a. A group or department of private guards: *Call building security if a visitor acts suspicious.*
    b. Measures adopted by a government to prevent espionage, sabotage, or attack.
    c. Measures adopted, as by a business or homeowner, to prevent a crime such as burglary or assault: *Security was lax at the firm's smaller plant.*
    d. Measures adopted to prevent escape: *Security in the prison is very tight.*
4. Something deposited or given as assurance of the fulfillment of an obligation; a pledge.
5. One who undertakes to fulfill the obligation of another; a surety.
6. A document indicating ownership or creditorship; a stock certificate or bond.

### Internet

The internet is a network of all the computers all over the world that communicate with each other by using TCP/IP protocols. The internet includes several services, such as email, file transfer, chat and the World Wide Web.

### Intranet

(*Private Network*) An intranet is the collection of private computer networks within an organization. Intranets generally use standard network technologies like Ethernet and TCP/IP. An organization's intranet often enjoys Internet access but is firewalled so that its computers cannot be reached directly from the public Internet. (An extranet opens "holes" in the firewall for select outsiders.)

**Router**

A router is a piece of hardware that connects one or more networks together. A router is technically a "layer three gateway". This means that it connects networks the same way that gateways do and it also operates at the network layer (three) of the OSI model. In a home network, and Internet Protocol (IP) router is normally used. IP routers such as a DSL or cable modem router will connect to the LAN (local area network) to the WAN (wide-area network) of the internet. The routing table in most routers enables it to filter traffic based on the IP addresses. The router automatically updates the routing table from a Web browser interface.

**Firewall**

A system designed to prevent unauthorized access to or from a private network. Firewalls can be implemented in both hardware and software, or a combination of both. Firewalls are frequently used to prevent unauthorized Internet users from accessing private networks connected to the Internet, especially *intranets*. All messages entering or leaving the intranet pass through the firewall, which examines each message and blocks those that do not meet the specified security criteria.

There are several types of firewall techniques:

- **Packet filter:** Looks at each packet entering or leaving the network and accepts or rejects it based on user-defined rules. Packet filtering is fairly effective and transparent to users, but it is difficult to configure. In addition, it is susceptible to IP spoofing.
- **Application gateway:** Applies security mechanisms to specific applications, such as FTP and Telnet servers. This is very effective, but can impose a performance degradation.
- **Circuit-level gateway:** Applies security mechanisms when a TCP or UDP connection is established. Once the connection has been made, packets can flow between the hosts without further checking.
- **Proxy server:** Intercepts all messages entering and leaving the network. The proxy server effectively hides the true network addresses.

In practice, many firewalls use two or more of these techniques in concert.

A firewall is considered a first line of defense in protecting private information. For greater security, data can be encrypted

**Intrusion Detection System**

An Intrusion Detection System (or IDS) generally detects unwanted manipulations to systems. There are a lot of different types of IDS, some of them are described here. The manipulations may take the form of attacks by skilled malicious hackers, or script kiddies using automated tools.

An Intrusion Detection System is required to detect all types of malicious network traffic and computer usage that can't be detected by a conventional firewall. This includes network attacks against vulnerable services, data driven attacks on applications, host based attacks such as privilege escalation, unauthorized logins and access to sensitive files, and malware (viruses, trojan horses, and worms).

In a **network-based system**, or NIDS, the sensors are located at choke points in the network to be monitored, often in the DMZ (demilitarized zone) or at network borders. The sensor captures all network traffic flows and analyzes the content of individual packets for malicious traffic. In systems, PIDS and APIDS are used to monitor the transport and protocols illegal or inappropriate traffic or constricts of language (say SQL). In a host-based system, the sensor usually consists of a software agent which monitors all activity of the host on which it is installed. Hybrids of these two types of system also exist.

A **Network Intrusion Detection System** is an independent platform which identifies intrusions by examining network traffic and monitors multiple hosts. Network Intrusion Detection Systems gain access to network traffic by connecting to a hub, network switch configured for port mirroring, or network tap. An example of a NIDS is Snort.

A **Protocol-based Intrusion Detection** System consists of a system or agent that would typically sit at the front end of a server, monitoring and analyzing the communication protocol between a connected device (a user/PC or system). For a web server this would typically monitor the HTTPS protocol stream and understand the HTTP protocol relative to the web server/system it is trying to protect. Where HTTPS is in use then this system would need to reside in the "shim" or interface between where HTTPS is un-encrypted and immediately prior to it entering the Web presentation layer.

An **Application Protocol-based Intrusion Detection System** consists of a system or agent that would typically sit within a group of servers, monitoring and analyzing the communication on application specific protocols. For example; in a web server with database this would monitor the SQL protocol specific to the middleware/business-login as it transacts with the database.

A **Host-based Intrusion Detection System** consists of an agent on a host which identifies intrusions by analyzing system calls, application logs, file-system modifications (binaries, password files, capability/acl databases) and other host activities and state.

A **Hybrid Intrusion Detection System** combines one or more approaches. Host agent data is combined with network information to form a comprehensive view of the network. An example of a Hybrid IDS is Prelude.

## A4. DATA FOR THE CHOICE OF SAMPLE SIZE

```
----------Model2_seed2.sae-------------
########################################
##         STRATEGY BAR=0, ID=0       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =6.60630627e+002; 6.53810554e+002; 4.91493822e+002
  DEV=3.04795902e+002; 3.09222971e+002; 2.89405040e+002
   Q5=1.00000000e+000; 1.00000000e+000; 1.00000000e+000
  Q25=5.10000000e+002; 5.00000000e+002; 2.50000000e+002
  Q50=7.56000000e+002; 7.50000000e+002; 5.10000000e+002
  Q75=9.05000000e+002; 9.10000000e+002; 7.51875000e+002
  Q95=1.00000000e+003; 1.00000000e+003; 9.00000000e+002
PRICE=0.00000000e+000; 7.00000000e+003; 7.00000000e+004
 SIZE=3.37200000e+003; 3.30600000e+003; 3.34100000e+003

########################################
##         STRATEGY BAR=0, ID=1       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =1.00068585e+001; 1.00541167e+001; 7.71756834e+000
  DEV=6.05372846e+000; 6.09765967e+000; 5.71015249e+000
   Q5=1.00000000e+000; 8.01000000e-001; 7.12380000e-001
  Q25=4.60000000e+000; 4.60000000e+000; 2.39950000e+000
  Q50=9.80000000e+000; 1.00000000e+001; 6.48920000e+000
  Q75=1.50000000e+001; 1.50000000e+001; 1.25000000e+001
  Q95=1.97500000e+001; 2.00000000e+001; 1.77150000e+001
PRICE=5.00000000e+003; 1.20000000e+004; 7.50000000e+004
 SIZE=3.29800000e+003; 3.29800000e+003; 3.39800000e+003

########################################
##         STRATEGY BAR=0, ID=2       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =7.51519514e+000; 7.55451002e+000; 5.34306663e+000
  DEV=4.70488907e+000; 4.74227078e+000; 2.88674699e+000
   Q5=1.10000000e+000; 1.00000000e+000; 1.00000000e+000
  Q25=3.16000000e+000; 3.00000000e+000; 3.00000000e+000
  Q50=6.84000000e+000; 7.00000000e+000; 5.00000000e+000
  Q75=1.21500000e+001; 1.20000000e+001; 7.88000000e+000
  Q95=1.50000000e+001; 1.50000000e+001; 9.90000000e+000
PRICE=3.00000000e+004; 3.70000000e+004; 1.00000000e+005
 SIZE=3.25400000e+003; 3.31800000e+003; 3.43300000e+003

########################################
##         STRATEGY BAR=0, ID=3       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =2.97128312e+000; 3.03041374e+000; 2.21628120e+000
  DEV=1.27961454e+000; 1.28301171e+000; 7.17505174e-001
```

```
   Q5=1.00000000e+000; 1.00000000e+000; 1.00000000e+000
  Q25=2.00000000e+000; 2.04000000e+000; 1.96847500e+000
  Q50=3.00000000e+000; 3.04000000e+000; 2.14290000e+000
  Q75=3.96000000e+000; 4.05000000e+000; 2.70000000e+000
  Q95=5.00000000e+000; 5.00000000e+000; 3.33330000e+000
PRICE=5.00000000e+004; 5.70000000e+004; 1.20000000e+005
 SIZE=3.34400000e+003; 3.21700000e+003; 3.38100000e+003

########################################
##         STRATEGY BAR=1, ID=0       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =1.02092449e+001; 1.01431468e+001; 7.74822897e+000
  DEV=6.07505298e+000; 5.98375333e+000; 5.61042150e+000
   Q5=1.00000000e+000; 1.00000000e+000; 7.12380000e-001
  Q25=5.00000000e+000; 5.00000000e+000; 2.50000000e+000
  Q50=1.00000000e+001; 1.00000000e+001; 6.66670000e+000
  Q75=1.50000000e+001; 1.50000000e+001; 1.25000000e+001
  Q95=2.00000000e+001; 2.00000000e+001; 1.77150000e+001
PRICE=1.50000000e+004; 2.20000000e+004; 8.50000000e+004
 SIZE=3.26400000e+003; 3.29800000e+003; 3.29900000e+003

########################################
##         STRATEGY BAR=1, ID=1       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =1.97535840e-001; 1.91922590e-001; 1.57240601e-001
  DEV=1.70133944e-001; 1.60596901e-001; 1.61788454e-001
   Q5=3.36990000e-002; 3.36990000e-002; 1.69000000e-002
  Q25=6.54550000e-002; 6.22960000e-002; 3.57020000e-002
  Q50=1.55130000e-001; 1.55130000e-001; 1.05130000e-001
  Q75=2.85470000e-001; 2.78260000e-001; 2.28890000e-001
  Q95=4.28840000e-001; 4.21630000e-001; 3.85950000e-001
PRICE=2.00000000e+004; 2.70000000e+004; 9.00000000e+004
 SIZE=3.34700000e+003; 3.32100000e+003; 3.42500000e+003

########################################
##         STRATEGY BAR=1, ID=2       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =1.55091008e-001; 1.54846416e-001; 1.14343500e-001
  DEV=1.52715084e-001; 1.56364283e-001; 1.34490117e-001
   Q5=2.38660000e-002; 2.39520000e-002; 2.38660000e-002
  Q25=4.38430000e-002; 4.38430000e-002; 3.85290000e-002
  Q50=1.16840000e-001; 1.16840000e-001; 7.89740000e-002
  Q75=2.23990000e-001; 2.19130000e-001; 1.49320000e-001
  Q95=3.21230000e-001; 3.21230000e-001; 2.13620000e-001
PRICE=4.50000000e+004; 5.20000000e+004; 1.15000000e+005
 SIZE=3.22500000e+003; 3.29900000e+003; 3.32700000e+003

########################################
##         STRATEGY BAR=1, ID=3       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =7.88991975e-002; 7.71217002e-002; 5.82832382e-002
  DEV=1.37565507e-001; 1.36622260e-001; 1.22449718e-001
   Q5=1.82540000e-002; 1.82540000e-002; 1.82540000e-002
```

```
  Q25=2.64000000e-002; 2.56000000e-002; 2.27890000e-002
  Q50=4.82610000e-002; 4.72120000e-002; 3.22050000e-002
  Q75=7.79280000e-002; 7.62900000e-002; 5.08600000e-002
  Q95=2.04000000e-001; 2.04000000e-001; 7.05000000e-002
PRICE=6.50000000e+004; 7.20000000e+004; 1.35000000e+005
 SIZE=3.26600000e+003; 3.30600000e+003; 3.42100000e+003

########################################
##         STRATEGY BAR=2, ID=0       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =3.01741091e+000; 2.95797553e+000; 2.22095912e+000
  DEV=1.29433214e+000; 1.30699564e+000; 1.30088923e+000
   Q5=1.00000000e+000; 1.00000000e+000; 5.81440000e-001
  Q25=2.04000000e+000; 1.95000000e+000; 1.00000000e+000
  Q50=3.04000000e+000; 2.96000000e+000; 2.00000000e+000
  Q75=4.05000000e+000; 3.96000000e+000; 3.33330000e+000
  Q95=5.08000000e+000; 5.00000000e+000; 4.50000000e+000
PRICE=3.00000000e+004; 3.70000000e+004; 1.00000000e+005
 SIZE=3.30000000e+003; 3.33000000e+003; 3.36800000e+003

########################################
##         STRATEGY BAR=2, ID=1       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =7.59280562e-002; 7.69296868e-002; 6.22066793e-002
  DEV=1.38523285e-001; 1.40483202e-001; 1.29435088e-001
   Q5=1.81946000e-002; 1.82540000e-002; 9.13280000e-003
  Q25=2.56000000e-002; 2.56000000e-002; 1.43710000e-002
  Q50=4.51280000e-002; 4.61680000e-002; 3.18210000e-002
  Q75=7.14340000e-002; 7.30430000e-002; 6.35640000e-002
  Q95=1.07160000e-001; 2.04000000e-001; 9.54120000e-002
PRICE=3.50000000e+004; 4.20000000e+004; 1.05000000e+005
 SIZE=3.25900000e+003; 3.34900000e+003; 3.37500000e+003

########################################
##         STRATEGY BAR=2, ID=2       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =6.37051345e-002; 6.75930793e-002; 5.47325006e-002
  DEV=1.39997551e-001; 1.46610079e-001; 1.36888417e-001
   Q5=1.17470000e-002; 1.17170000e-002; 1.17470000e-002
  Q25=1.66670000e-002; 1.71200000e-002; 1.55200000e-002
  Q50=3.28210000e-002; 3.38020000e-002; 2.54090000e-002
  Q75=5.64670000e-002; 5.85810000e-002; 4.04760000e-002
  Q95=8.02400000e-002; 2.04000000e-001; 5.47500000e-002
PRICE=6.00000000e+004; 6.70000000e+004; 1.30000000e+005
 SIZE=3.38300000e+003; 3.30200000e+003; 3.54200000e+003

########################################
##         STRATEGY BAR=2, ID=3       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =4.47983986e-002; 4.37916561e-002; 3.72619864e-002
  DEV=1.43047290e-001; 1.41892243e-001; 1.32791140e-001
   Q5=8.16000000e-003; 8.09600000e-003; 7.89230000e-003
  Q25=9.92000000e-003; 1.00800000e-002; 8.25010000e-003
```

```
  Q50=1.40410000e-002; 1.42850000e-002; 1.00930000e-002
  Q75=2.04170000e-002; 2.05960000e-002; 1.42100000e-002
  Q95=2.04000000e-001; 5.47500000e-002; 5.47500000e-002
PRICE=8.00000000e+004; 8.70000000e+004; 1.50000000e+005
 SIZE=3.28000000e+003; 3.38700000e+003; 3.44100000e+003

########################################
##         STRATEGY BAR=3, ID=0       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =2.03472582e+000; 1.00565056e+000; 8.76602436e-001
  DEV=1.23757949e+000; 6.12128910e-001; 4.66515107e-001
   Q5=1.00000000e-002; 1.00000000e-002; 2.83750000e-002
  Q25=1.04000000e+000; 5.15000000e-001; 5.20000000e-001
  Q50=2.00000000e+000; 1.00000000e+000; 1.00000000e+000
  Q75=3.10000000e+000; 1.48000000e+000; 1.25000000e+000
  Q95=4.00000000e+000; 2.00000000e+000; 1.50000000e+000
PRICE=4.50000000e+004; 5.20000000e+004; 1.15000000e+005
 SIZE=3.30800000e+003; 3.38900000e+003; 3.32500000e+003

########################################
##         STRATEGY BAR=3, ID=1       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =3.26754034e-002; 1.67003527e-002; 1.43302510e-002
  DEV=2.42038734e-002; 1.20939671e-002; 9.20079466e-003
   Q5=4.48180000e-003; 2.29200000e-003; 2.10820000e-003
  Q25=1.06030000e-002; 6.54550000e-003; 6.54550000e-003
  Q50=2.66670000e-002; 1.33330000e-002; 1.33330000e-002
  Q75=5.00000000e-002; 2.50000000e-002; 2.08330000e-002
  Q95=8.00000000e-002; 4.00000000e-002; 3.00000000e-002
PRICE=5.00000000e+004; 5.70000000e+004; 1.20000000e+005
 SIZE=3.32700000e+003; 3.30700000e+003; 3.47200000e+003

########################################
##         STRATEGY BAR=3, ID=2       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =2.48438371e-002; 1.27034188e-002; 7.93374004e-003
  DEV=1.89607836e-002; 9.52454242e-003; 5.14592966e-003
   Q5=3.45130000e-003; 1.60684000e-003; 1.66690000e-003
  Q25=7.81930000e-003; 4.14550000e-003; 3.70910000e-003
  Q50=1.86670000e-002; 9.71430000e-003; 7.53850000e-003
  Q75=4.00000000e-002; 2.04740000e-002; 1.14890000e-002
  Q95=6.00000000e-002; 3.00000000e-002; 1.52680000e-002
PRICE=7.50000000e+004; 8.20000000e+004; 1.45000000e+005
 SIZE=3.33300000e+003; 3.29600000e+003; 3.33900000e+003

########################################
##         STRATEGY BAR=3, ID=3       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =9.57484919e-003; 4.92287748e-003; 3.42666416e-003
  DEV=5.42741403e-003; 2.79198851e-003; 3.28464383e-003
   Q5=2.64830000e-003; 1.35430000e-003; 1.39530000e-003
  Q25=4.96000000e-003; 2.32000000e-003; 2.17940000e-003
  Q50=8.40730000e-003; 4.51280000e-003; 3.04800000e-003
```

```
  Q75=1.33330000e-002; 7.00000000e-003; 3.95650000e-003
  Q95=1.98604000e-002; 1.00000000e-002; 6.00000000e-003
PRICE=9.50000000e+004; 1.02000000e+005; 1.65000000e+005
 SIZE=3.33800000e+003; 3.32600000e+003; 3.43200000e+003

########################################
##         STRATEGY BAR=4, ID=0       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =5.04827309e-001; 2.46820295e-001; 2.19746144e-001
  DEV=3.07851117e-001; 1.52501155e-001; 1.14864417e-001
   Q5=1.00000000e-002; 1.00000000e-002; 1.27560000e-002
  Q25=2.50000000e-001; 1.19060000e-001; 1.25000000e-001
  Q50=5.00000000e-001; 2.40000000e-001; 2.50000000e-001
  Q75=7.50000000e-001; 3.65000000e-001; 3.10000000e-001
  Q95=1.00000000e+000; 5.00000000e-001; 3.75000000e-001
PRICE=1.00000000e+006; 1.00700000e+006; 1.07000000e+006
 SIZE=3.30300000e+003; 3.31300000e+003; 3.35800000e+003

########################################
##         STRATEGY BAR=4, ID=1       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =8.63744645e-003; 3.50551478e-003; 3.29113264e-003
  DEV=5.92662265e-003; 2.54936762e-003; 4.39457536e-003
   Q5=1.19700000e-003; 4.63310000e-004; 4.45230000e-004
  Q25=3.43210000e-003; 1.15640000e-003; 1.03630000e-003
  Q50=7.53490000e-003; 2.92110000e-003; 2.76460000e-003
  Q75=1.28480000e-002; 5.59970000e-003; 4.59480000e-003
  Q95=2.00000000e-002; 7.91480000e-003; 6.11710000e-003
PRICE=1.00500000e+006; 1.01200000e+006; 1.07500000e+006
 SIZE=3.23500000e+003; 3.32800000e+003; 3.29000000e+003

########################################
##         STRATEGY BAR=4, ID=2       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =6.53309504e-003; 3.45257073e-003; 2.53985663e-003
  DEV=4.64956688e-003; 2.55141933e-003; 4.15243504e-003
   Q5=9.10020000e-004; 4.63310000e-004; 4.63310000e-004
  Q25=2.07270000e-003; 1.03630000e-003; 9.63210000e-004
  Q50=6.00000000e-003; 2.92110000e-003; 1.92850000e-003
  Q75=1.00000000e-002; 5.47830000e-003; 3.06320000e-003
  Q95=1.50000000e-002; 7.91480000e-003; 6.00000000e-003
PRICE=1.03000000e+006; 1.03700000e+006; 1.10000000e+006
 SIZE=3.31500000e+003; 3.30100000e+003; 3.39000000e+003

########################################
##         STRATEGY BAR=4, ID=3       ##
########################################
------------HUM=0----------HUM=1----------HUM=2
MEAN =2.71102266e-003; 1.55945434e-003; 1.40407593e-003
  DEV=1.84560505e-003; 1.69102301e-003; 3.99873318e-003
   Q5=7.34770000e-004; 3.67920000e-004; 3.67920000e-004
  Q25=1.24000000e-003; 6.20000000e-004; 5.44850000e-004
  Q50=2.25640000e-003; 1.12820000e-003; 7.70200000e-004
  Q75=3.73310000e-003; 1.82610000e-003; 1.02100000e-003
```

```
  Q95=5.37600000e-003; 5.25000000e-003; 6.00000000e-003
PRICE=1.05000000e+006; 1.05700000e+006; 1.12000000e+006

 SIZE=3.26000000e+003; 3.25900000e+003; 3.38200000e+003
```

**A5. Expected rate of attack in the next six years depending on the chosen strategy**

| HUM | BAR | ID | INV | E_RATE1 | E_RATE2 | E_RATE3 | E_RATE4 | E_RATE5 | E_RATE6 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 7000 | 6.55E-04 | 5.39E-04 | 4.23E-04 | 3.07E-04 | 1.91E-04 | 7.55E-05 |
| 0 | 0 | 0 | 0 | 6.54E-04 | 5.38E-04 | 4.22E-04 | 3.07E-04 | 1.91E-04 | 7.52E-05 |
| 2 | 0 | 0 | 27000 | 4.86E-04 | 4.01E-04 | 3.15E-04 | 2.30E-04 | 1.44E-04 | 5.85E-05 |
| 1 | 1 | 0 | 22000 | 1.02E-05 | 9.19E-06 | 8.18E-06 | 7.17E-06 | 6.16E-06 | 5.16E-06 |
| 1 | 0 | 1 | 12000 | 1.02E-05 | 9.64E-06 | 9.12E-06 | 8.59E-06 | 8.06E-06 | 7.53E-06 |
| 0 | 1 | 0 | 15000 | 1.01E-05 | 9.14E-06 | 8.14E-06 | 7.14E-06 | 6.13E-06 | 5.13E-06 |
| 0 | 0 | 1 | 5000 | 1.01E-05 | 9.61E-06 | 9.08E-06 | 8.56E-06 | 8.03E-06 | 7.50E-06 |
| 2 | 1 | 0 | 42000 | 7.81E-06 | 7.05E-06 | 6.29E-06 | 5.52E-06 | 4.76E-06 | 4.00E-06 |
| 2 | 0 | 1 | 32000 | 7.80E-06 | 7.40E-06 | 7.01E-06 | 6.62E-06 | 6.22E-06 | 5.83E-06 |
| 1 | 0 | 2 | 42000 | 7.55E-06 | 7.07E-06 | 6.59E-06 | 6.11E-06 | 5.63E-06 | 5.16E-06 |
| 0 | 0 | 2 | 35000 | 7.51E-06 | 7.03E-06 | 6.56E-06 | 6.08E-06 | 5.61E-06 | 5.13E-06 |
| 2 | 0 | 2 | 62000 | 5.37E-06 | 5.20E-06 | 5.03E-06 | 4.86E-06 | 4.69E-06 | 4.52E-06 |
| 1 | 2 | 0 | 97000 | 2.99E-06 | 2.80E-06 | 2.61E-06 | 2.42E-06 | 2.23E-06 | 2.04E-06 |
| 1 | 0 | 3 | 92000 | 2.99E-06 | 2.99E-06 | 2.99E-06 | 2.99E-06 | 2.99E-06 | 2.99E-06 |
| 0 | 0 | 3 | 85000 | 2.98E-06 | 2.98E-06 | 2.98E-06 | 2.98E-06 | 2.98E-06 | 2.98E-06 |
| 0 | 2 | 0 | 90000 | 2.98E-06 | 2.79E-06 | 2.60E-06 | 2.41E-06 | 2.22E-06 | 2.02E-06 |
| 2 | 2 | 0 | 117000 | 2.23E-06 | 2.10E-06 | 1.97E-06 | 1.84E-06 | 1.71E-06 | 1.58E-06 |
| 2 | 0 | 3 | 112000 | 2.20E-06 | 2.29E-06 | 2.38E-06 | 2.46E-06 | 2.55E-06 | 2.64E-06 |
| 0 | 3 | 0 | 135000 | 2.01E-06 | 1.80E-06 | 1.60E-06 | 1.40E-06 | 1.19E-06 | 9.92E-07 |
| 1 | 3 | 0 | 142000 | 1.01E-06 | 1.01E-06 | 1.00E-06 | 1.00E-06 | 9.99E-07 | 9.97E-07 |
| 2 | 3 | 0 | 162000 | 8.74E-07 | 8.72E-07 | 8.70E-07 | 8.67E-07 | 8.65E-07 | 8.63E-07 |
| 0 | 4 | 0 | 210000 | 5.02E-07 | 5.01E-07 | 5.00E-07 | 4.99E-07 | 4.98E-07 | 4.97E-07 |
| 1 | 4 | 0 | 217000 | 2.53E-07 | 3.02E-07 | 3.52E-07 | 4.01E-07 | 4.50E-07 | 5.00E-07 |
| 2 | 4 | 0 | 237000 | 2.19E-07 | 2.62E-07 | 3.05E-07 | 3.47E-07 | 3.90E-07 | 4.33E-07 |
| 1 | 1 | 1 | 27000 | 1.94E-07 | 2.66E-07 | 3.38E-07 | 4.10E-07 | 4.82E-07 | 5.54E-07 |
| 0 | 1 | 1 | 20000 | 1.93E-07 | 2.64E-07 | 3.36E-07 | 4.08E-07 | 4.79E-07 | 5.51E-07 |
| 2 | 1 | 1 | 47000 | 1.58E-07 | 2.13E-07 | 2.68E-07 | 3.23E-07 | 3.79E-07 | 4.34E-07 |
| 1 | 1 | 2 | 57000 | 1.53E-07 | 2.00E-07 | 2.47E-07 | 2.94E-07 | 3.41E-07 | 3.88E-07 |
| 0 | 1 | 2 | 50000 | 1.51E-07 | 1.98E-07 | 2.45E-07 | 2.92E-07 | 3.39E-07 | 3.86E-07 |
| 2 | 1 | 2 | 77000 | 1.15E-07 | 1.60E-07 | 2.06E-07 | 2.51E-07 | 2.97E-07 | 3.43E-07 |
| 1 | 2 | 1 | 102000 | 7.74E-08 | 1.07E-07 | 1.36E-07 | 1.65E-07 | 1.95E-07 | 2.24E-07 |
| 1 | 1 | 3 | 107000 | 7.74E-08 | 1.09E-07 | 1.40E-07 | 1.72E-07 | 2.03E-07 | 2.35E-07 |
| 0 | 2 | 1 | 95000 | 7.65E-08 | 1.06E-07 | 1.35E-07 | 1.64E-07 | 1.93E-07 | 2.22E-07 |
| 0 | 1 | 3 | 100000 | 7.65E-08 | 1.08E-07 | 1.39E-07 | 1.71E-07 | 2.02E-07 | 2.34E-07 |
| 1 | 2 | 2 | 132000 | 6.55E-08 | 8.42E-08 | 1.03E-07 | 1.22E-07 | 1.40E-07 | 1.59E-07 |
| 2 | 2 | 1 | 122000 | 6.49E-08 | 8.73E-08 | 1.10E-07 | 1.32E-07 | 1.55E-07 | 1.77E-07 |
| 0 | 2 | 2 | 125000 | 6.46E-08 | 8.32E-08 | 1.02E-07 | 1.20E-07 | 1.39E-07 | 1.58E-07 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 1 | 3 | 127000 | 6.29E-08 | 9.22E-08 | 1.21E-07 | 1.51E-07 | 1.80E-07 | 2.09E-07 |
| 2 | 2 | 2 | 152000 | 5.41E-08 | 7.15E-08 | 8.88E-08 | 1.06E-07 | 1.24E-07 | 1.41E-07 |
| 1 | 2 | 3 | 182000 | 4.48E-08 | 5.57E-08 | 6.65E-08 | 7.74E-08 | 8.83E-08 | 9.91E-08 |
| 0 | 2 | 3 | 175000 | 4.40E-08 | 5.48E-08 | 6.56E-08 | 7.64E-08 | 8.72E-08 | 9.80E-08 |
| 2 | 2 | 3 | 202000 | 3.98E-08 | 4.96E-08 | 5.94E-08 | 6.92E-08 | 7.90E-08 | 8.88E-08 |
| 0 | 3 | 1 | 140000 | 3.26E-08 | 4.62E-08 | 5.98E-08 | 7.34E-08 | 8.70E-08 | 1.01E-07 |
| 0 | 3 | 2 | 170000 | 2.45E-08 | 3.33E-08 | 4.21E-08 | 5.10E-08 | 5.98E-08 | 6.86E-08 |
| 1 | 3 | 1 | 147000 | 1.67E-08 | 3.36E-08 | 5.05E-08 | 6.74E-08 | 8.43E-08 | 1.01E-07 |
| 2 | 3 | 1 | 167000 | 1.44E-08 | 2.89E-08 | 4.34E-08 | 5.79E-08 | 7.23E-08 | 8.68E-08 |
| 1 | 3 | 2 | 177000 | 1.25E-08 | 2.38E-08 | 3.51E-08 | 4.64E-08 | 5.77E-08 | 6.89E-08 |
| 0 | 3 | 3 | 220000 | 9.51E-09 | 1.54E-08 | 2.12E-08 | 2.71E-08 | 3.30E-08 | 3.88E-08 |
| 0 | 4 | 1 | 215000 | 8.49E-09 | 1.69E-08 | 2.53E-08 | 3.38E-08 | 4.22E-08 | 5.06E-08 |
| 2 | 3 | 2 | 197000 | 7.85E-09 | 1.67E-08 | 2.55E-08 | 3.43E-08 | 4.31E-08 | 5.19E-08 |
| 0 | 4 | 2 | 245000 | 6.44E-09 | 1.21E-08 | 1.77E-08 | 2.33E-08 | 2.90E-08 | 3.46E-08 |
| 1 | 3 | 3 | 227000 | 5.00E-09 | 1.18E-08 | 1.86E-08 | 2.54E-08 | 3.22E-08 | 3.90E-08 |
| 2 | 3 | 3 | 247000 | 3.49E-09 | 8.79E-09 | 1.41E-08 | 1.94E-08 | 2.47E-08 | 3.00E-08 |
| 1 | 4 | 1 | 222000 | 3.46E-09 | 9.73E-09 | 1.60E-08 | 2.23E-08 | 2.85E-08 | 3.48E-08 |
| 1 | 4 | 2 | 252000 | 3.46E-09 | 9.73E-09 | 1.60E-08 | 2.23E-08 | 2.85E-08 | 3.48E-08 |
| 2 | 4 | 1 | 242000 | 3.20E-09 | 8.56E-09 | 1.39E-08 | 1.93E-08 | 2.46E-08 | 3.00E-08 |
| 0 | 4 | 3 | 295000 | 2.70E-09 | 6.10E-09 | 9.51E-09 | 1.29E-08 | 1.63E-08 | 1.97E-08 |
| 2 | 4 | 2 | 272000 | 2.45E-09 | 7.20E-09 | 1.20E-08 | 1.67E-08 | 2.15E-08 | 2.62E-08 |
| 1 | 4 | 3 | 302000 | 1.58E-09 | 5.22E-09 | 8.86E-09 | 1.25E-08 | 1.61E-08 | 1.98E-08 |
| 2 | 4 | 3 | 322000 | 1.36E-09 | 4.14E-09 | 6.93E-09 | 9.71E-09 | 1.25E-08 | 1.53E-08 |

**A6. Expected NPV, ROI and expected flows for strategies when disclosure cost equals 10 million, for three-year project**

| Strategy | E_NPV | E_ROI | E_INC1 | E_INC2 | E_INC3 | E_INC4 | E_INC5 | E_INC6 |
|---|---|---|---|---|---|---|---|---|
| (0,0,1) | 9.00E+03 | 1.80E+00 | 6.43E+03 | 5.28E+03 | 4.13E+03 | 2.98E+03 | 1.83E+03 | 6.77E+02 |
| (1,0,1) | 2.04E+03 | 1.70E-01 | 6.45E+03 | 5.30E+03 | 4.14E+03 | 2.99E+03 | 1.83E+03 | 6.80E+02 |
| (0,0,0) | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 |
| (0,1,0) | -9.72E+02 | -6.48E-02 | 6.44E+03 | 5.29E+03 | 4.15E+03 | 3.00E+03 | 1.85E+03 | 7.02E+02 |
| (0,1,1) | -5.74E+03 | -2.87E-01 | 6.54E+03 | 5.38E+03 | 4.22E+03 | 3.07E+03 | 1.91E+03 | 7.47E+02 |
| (1,1,0) | -7.93E+03 | -3.60E-01 | 6.46E+03 | 5.31E+03 | 4.16E+03 | 3.01E+03 | 1.86E+03 | 7.05E+02 |
| (1,1,1) | -1.27E+04 | -4.70E-01 | 6.56E+03 | 5.40E+03 | 4.24E+03 | 3.08E+03 | 1.91E+03 | 7.51E+02 |
| (2,0,1) | -1.79E+04 | -5.60E-01 | 6.47E+03 | 5.31E+03 | 4.16E+03 | 3.00E+03 | 1.85E+03 | 6.95E+02 |
| (0,0,2) | -2.09E+04 | -5.98E-01 | 6.47E+03 | 5.32E+03 | 4.16E+03 | 3.01E+03 | 1.86E+03 | 7.02E+02 |
| (1,0,2) | -2.79E+04 | -6.64E-01 | 6.49E+03 | 5.33E+03 | 4.18E+03 | 3.02E+03 | 1.86E+03 | 7.05E+02 |
| (2,1,0) | -2.79E+04 | -6.64E-01 | 6.47E+03 | 5.32E+03 | 4.17E+03 | 3.02E+03 | 1.87E+03 | 7.15E+02 |
| (2,1,1) | -3.27E+04 | -6.96E-01 | 6.55E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.50E+02 |
| (0,1,2) | -3.57E+04 | -7.15E-01 | 6.54E+03 | 5.38E+03 | 4.22E+03 | 3.07E+03 | 1.91E+03 | 7.49E+02 |
| (1,1,2) | -4.27E+04 | -7.49E-01 | 6.56E+03 | 5.40E+03 | 4.24E+03 | 3.08E+03 | 1.91E+03 | 7.53E+02 |
| (2,0,2) | -4.78E+04 | -7.72E-01 | 6.50E+03 | 5.34E+03 | 4.18E+03 | 3.03E+03 | 1.87E+03 | 7.09E+02 |
| (2,1,2) | -6.27E+04 | -8.14E-01 | 6.55E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.51E+02 |
| (0,0,3) | -7.08E+04 | -8.33E-01 | 6.51E+03 | 5.36E+03 | 4.20E+03 | 3.04E+03 | 1.88E+03 | 7.23E+02 |
| (0,2,0) | -7.58E+04 | -8.42E-01 | 6.51E+03 | 5.36E+03 | 4.20E+03 | 3.04E+03 | 1.89E+03 | 7.33E+02 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (1,0,3) | -7.78E+04 | -8.45E-01 | 6.53E+03 | 5.37E+03 | 4.21E+03 | 3.05E+03 | 1.89E+03 | 7.27E+02 |
| (0,2,1) | -8.07E+04 | -8.50E-01 | 6.54E+03 | 5.38E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.51E+02 |
| (1,2,0) | -8.28E+04 | -8.53E-01 | 6.53E+03 | 5.37E+03 | 4.21E+03 | 3.06E+03 | 1.90E+03 | 7.36E+02 |
| (0,1,3) | -8.57E+04 | -8.57E-01 | 6.54E+03 | 5.38E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.51E+02 |
| (1,2,1) | -8.77E+04 | -8.60E-01 | 6.56E+03 | 5.40E+03 | 4.24E+03 | 3.08E+03 | 1.92E+03 | 7.54E+02 |
| (2,0,0) | -2.33E+04 | -8.65E-01 | 1.68E+03 | 1.38E+03 | 1.08E+03 | 7.73E+02 | 4.70E+02 | 1.68E+02 |
| (1,1,3) | -9.27E+04 | -8.66E-01 | 6.56E+03 | 5.40E+03 | 4.24E+03 | 3.08E+03 | 1.92E+03 | 7.54E+02 |
| (2,0,3) | -9.78E+04 | -8.73E-01 | 6.53E+03 | 5.37E+03 | 4.21E+03 | 3.05E+03 | 1.89E+03 | 7.28E+02 |
| (2,2,0) | -1.03E+05 | -8.78E-01 | 6.53E+03 | 5.37E+03 | 4.21E+03 | 3.06E+03 | 1.90E+03 | 7.39E+02 |
| (2,2,1) | -1.08E+05 | -8.83E-01 | 6.55E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.53E+02 |
| (0,2,2) | -1.11E+05 | -8.86E-01 | 6.54E+03 | 5.38E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.51E+02 |
| (2,1,3) | -1.13E+05 | -8.88E-01 | 6.55E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.52E+02 |
| (1,2,2) | -1.18E+05 | -8.92E-01 | 6.56E+03 | 5.40E+03 | 4.24E+03 | 3.08E+03 | 1.92E+03 | 7.55E+02 |
| (0,3,0) | -1.21E+05 | -8.95E-01 | 6.52E+03 | 5.37E+03 | 4.21E+03 | 3.06E+03 | 1.90E+03 | 7.43E+02 |
| (0,3,1) | -1.26E+05 | -8.98E-01 | 6.54E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.52E+02 |
| (1,3,0) | -1.28E+05 | -8.99E-01 | 6.55E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.47E+02 |
| (1,3,1) | -1.33E+05 | -9.03E-01 | 6.56E+03 | 5.40E+03 | 4.24E+03 | 3.08E+03 | 1.92E+03 | 7.56E+02 |
| (2,2,2) | -1.38E+05 | -9.06E-01 | 6.55E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.53E+02 |
| (2,3,0) | -1.48E+05 | -9.12E-01 | 6.54E+03 | 5.38E+03 | 4.22E+03 | 3.07E+03 | 1.91E+03 | 7.46E+02 |
| (2,3,1) | -1.53E+05 | -9.14E-01 | 6.55E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.54E+02 |
| (0,3,2) | -1.56E+05 | -9.16E-01 | 6.54E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.52E+02 |
| (0,2,3) | -1.61E+05 | -9.18E-01 | 6.54E+03 | 5.38E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.52E+02 |
| (1,3,2) | -1.63E+05 | -9.19E-01 | 6.56E+03 | 5.40E+03 | 4.24E+03 | 3.08E+03 | 1.92E+03 | 7.56E+02 |
| (1,2,3) | -1.68E+05 | -9.21E-01 | 6.56E+03 | 5.40E+03 | 4.24E+03 | 3.08E+03 | 1.92E+03 | 7.56E+02 |
| (2,3,2) | -1.83E+05 | -9.27E-01 | 6.55E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.54E+02 |
| (2,2,3) | -1.88E+05 | -9.29E-01 | 6.55E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.54E+02 |
| (0,4,0) | -1.96E+05 | -9.32E-01 | 6.54E+03 | 5.38E+03 | 4.22E+03 | 3.06E+03 | 1.91E+03 | 7.48E+02 |
| (0,4,1) | -2.01E+05 | -9.34E-01 | 6.54E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.52E+02 |
| (1,4,0) | -2.03E+05 | -9.34E-01 | 6.56E+03 | 5.40E+03 | 4.24E+03 | 3.08E+03 | 1.91E+03 | 7.52E+02 |
| (0,3,3) | -2.06E+05 | -9.35E-01 | 6.54E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.53E+02 |
| (1,4,1) | -2.08E+05 | -9.36E-01 | 6.56E+03 | 5.40E+03 | 4.24E+03 | 3.08E+03 | 1.92E+03 | 7.56E+02 |
| (1,3,3) | -2.13E+05 | -9.37E-01 | 6.56E+03 | 5.40E+03 | 4.24E+03 | 3.08E+03 | 1.92E+03 | 7.56E+02 |
| (2,4,0) | -2.23E+05 | -9.40E-01 | 6.55E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.50E+02 |
| (2,4,1) | -2.28E+05 | -9.41E-01 | 6.55E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.54E+02 |
| (0,4,2) | -2.31E+05 | -9.42E-01 | 6.54E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.53E+02 |
| (2,3,3) | -2.33E+05 | -9.42E-01 | 6.55E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.54E+02 |
| (1,4,2) | -2.38E+05 | -9.43E-01 | 6.56E+03 | 5.40E+03 | 4.24E+03 | 3.08E+03 | 1.92E+03 | 7.56E+02 |
| (2,4,2) | -2.58E+05 | -9.47E-01 | 6.55E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.54E+02 |
| (0,4,3) | -2.81E+05 | -9.52E-01 | 6.54E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.53E+02 |
| (1,4,3) | -2.88E+05 | -9.53E-01 | 6.56E+03 | 5.40E+03 | 4.24E+03 | 3.08E+03 | 1.92E+03 | 7.57E+02 |
| (2,4,3) | -3.08E+05 | -9.56E-01 | 6.55E+03 | 5.39E+03 | 4.23E+03 | 3.07E+03 | 1.91E+03 | 7.54E+02 |
| (1,0,0) | -7.00E+03 | -1.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 |

**A7. NPV distribution depending on the chosen strategy when disclosure cost equals 10 million,, for three-year project**

| HUM | BAR | ID | INV | Q5_NPV | Q25_NPV | E_NPV | Q50_NPV | Q75_NPV | Q95_NPV |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 1 | 5000 | -5.00E+03 | 5.75E+03 | 9.00E+03 | 1.11E+04 | 1.42E+04 | 1.65E+04 |
| 1 | 0 | 1 | 12000 | -1.20E+04 | -1.26E+03 | 2.04E+03 | 4.05E+03 | 7.23E+03 | 9.48E+03 |
| 0 | 0 | 0 | 0 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 |
| 0 | 1 | 0 | 15000 | -1.50E+04 | -4.25E+03 | -9.72E+02 | 1.06E+03 | 4.25E+03 | 6.51E+03 |
| 0 | 1 | 1 | 20000 | -2.00E+04 | -9.14E+03 | -5.74E+03 | -3.71E+03 | -4.01E+02 | 1.97E+03 |
| 1 | 1 | 0 | 22000 | -2.20E+04 | -1.13E+04 | -7.93E+03 | -5.94E+03 | -2.75E+03 | -4.92E+02 |
| 1 | 1 | 1 | 27000 | -2.70E+04 | -1.61E+04 | -1.27E+04 | -1.07E+04 | -7.40E+03 | -5.03E+03 |
| 2 | 0 | 1 | 32000 | -3.20E+04 | -2.12E+04 | -1.79E+04 | -1.59E+04 | -1.27E+04 | -1.04E+04 |
| 0 | 0 | 2 | 35000 | -3.50E+04 | -2.42E+04 | -2.09E+04 | -1.89E+04 | -1.57E+04 | -1.34E+04 |
| 1 | 0 | 2 | 42000 | -4.20E+04 | -3.12E+04 | -2.79E+04 | -2.59E+04 | -2.27E+04 | -2.04E+04 |
| 2 | 1 | 0 | 42000 | -4.20E+04 | -3.12E+04 | -2.79E+04 | -2.59E+04 | -2.27E+04 | -2.04E+04 |
| 2 | 1 | 1 | 47000 | -4.70E+04 | -3.61E+04 | -3.27E+04 | -3.07E+04 | -2.74E+04 | -2.50E+04 |
| 0 | 1 | 2 | 50000 | -5.00E+04 | -3.91E+04 | -3.57E+04 | -3.37E+04 | -3.04E+04 | -2.80E+04 |
| 1 | 1 | 2 | 57000 | -5.70E+04 | -4.61E+04 | -4.27E+04 | -4.07E+04 | -3.74E+04 | -3.50E+04 |
| 2 | 0 | 2 | 62000 | -6.20E+04 | -5.12E+04 | -4.78E+04 | -4.58E+04 | -4.26E+04 | -4.03E+04 |
| 2 | 1 | 2 | 77000 | -7.70E+04 | -6.61E+04 | -6.27E+04 | -6.07E+04 | -5.74E+04 | -5.50E+04 |
| 0 | 0 | 3 | 85000 | -8.50E+04 | -7.42E+04 | -7.08E+04 | -6.88E+04 | -6.55E+04 | -6.31E+04 |
| 0 | 2 | 0 | 90000 | -9.00E+04 | -7.92E+04 | -7.58E+04 | -7.38E+04 | -7.05E+04 | -6.81E+04 |
| 1 | 0 | 3 | 92000 | -9.20E+04 | -8.12E+04 | -7.78E+04 | -7.58E+04 | -7.25E+04 | -7.01E+04 |
| 0 | 2 | 1 | 95000 | -9.50E+04 | -8.41E+04 | -8.07E+04 | -7.87E+04 | -7.54E+04 | -7.30E+04 |
| 1 | 2 | 0 | 97000 | -9.70E+04 | -8.62E+04 | -8.28E+04 | -8.08E+04 | -7.75E+04 | -7.51E+04 |
| 0 | 1 | 3 | 100000 | -1.00E+05 | -8.91E+04 | -8.57E+04 | -8.37E+04 | -8.04E+04 | -7.80E+04 |
| 1 | 2 | 1 | 102000 | -1.02E+05 | -9.11E+04 | -8.77E+04 | -8.57E+04 | -8.24E+04 | -8.00E+04 |
| 2 | 0 | 0 | 27000 | -2.70E+04 | -2.44E+04 | -2.33E+04 | -2.32E+04 | -2.16E+04 | -2.16E+04 |
| 1 | 1 | 3 | 107000 | -1.07E+05 | -9.61E+04 | -9.27E+04 | -9.07E+04 | -8.74E+04 | -8.50E+04 |
| 2 | 0 | 3 | 112000 | -1.12E+05 | -1.01E+05 | -9.78E+04 | -9.58E+04 | -9.25E+04 | -9.01E+04 |
| 2 | 2 | 0 | 117000 | -1.17E+05 | -1.06E+05 | -1.03E+05 | -1.01E+05 | -9.75E+04 | -9.51E+04 |
| 2 | 2 | 1 | 122000 | -1.22E+05 | -1.11E+05 | -1.08E+05 | -1.06E+05 | -1.02E+05 | -1.00E+05 |
| 0 | 2 | 2 | 125000 | -1.25E+05 | -1.14E+05 | -1.11E+05 | -1.09E+05 | -1.05E+05 | -1.03E+05 |
| 2 | 1 | 3 | 127000 | -1.27E+05 | -1.16E+05 | -1.13E+05 | -1.11E+05 | -1.07E+05 | -1.05E+05 |
| 1 | 2 | 2 | 132000 | -1.32E+05 | -1.21E+05 | -1.18E+05 | -1.16E+05 | -1.12E+05 | -1.10E+05 |
| 0 | 3 | 0 | 135000 | -1.35E+05 | -1.24E+05 | -1.21E+05 | -1.19E+05 | -1.15E+05 | -1.13E+05 |
| 0 | 3 | 1 | 140000 | -1.40E+05 | -1.29E+05 | -1.26E+05 | -1.24E+05 | -1.20E+05 | -1.18E+05 |
| 1 | 3 | 0 | 142000 | -1.42E+05 | -1.31E+05 | -1.28E+05 | -1.26E+05 | -1.22E+05 | -1.20E+05 |
| 1 | 3 | 1 | 147000 | -1.47E+05 | -1.36E+05 | -1.33E+05 | -1.31E+05 | -1.27E+05 | -1.25E+05 |
| 2 | 2 | 2 | 152000 | -1.52E+05 | -1.41E+05 | -1.38E+05 | -1.36E+05 | -1.32E+05 | -1.30E+05 |
| 2 | 3 | 0 | 162000 | -1.62E+05 | -1.51E+05 | -1.48E+05 | -1.46E+05 | -1.42E+05 | -1.40E+05 |
| 2 | 3 | 1 | 167000 | -1.67E+05 | -1.56E+05 | -1.53E+05 | -1.51E+05 | -1.47E+05 | -1.45E+05 |
| 0 | 3 | 2 | 170000 | -1.70E+05 | -1.59E+05 | -1.56E+05 | -1.54E+05 | -1.50E+05 | -1.48E+05 |
| 0 | 2 | 3 | 175000 | -1.75E+05 | -1.64E+05 | -1.61E+05 | -1.59E+05 | -1.55E+05 | -1.53E+05 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 3 | 2 | 177000 | -1.77E+05 | -1.66E+05 | -1.63E+05 | -1.61E+05 | -1.57E+05 | -1.55E+05 |
| 1 | 2 | 3 | 182000 | -1.82E+05 | -1.71E+05 | -1.68E+05 | -1.66E+05 | -1.62E+05 | -1.60E+05 |
| 2 | 3 | 2 | 197000 | -1.97E+05 | -1.86E+05 | -1.83E+05 | -1.81E+05 | -1.77E+05 | -1.75E+05 |
| 2 | 2 | 3 | 202000 | -2.02E+05 | -1.91E+05 | -1.88E+05 | -1.86E+05 | -1.82E+05 | -1.80E+05 |
| 0 | 4 | 0 | 210000 | -2.10E+05 | -1.99E+05 | -1.96E+05 | -1.94E+05 | -1.90E+05 | -1.88E+05 |
| 0 | 4 | 1 | 215000 | -2.15E+05 | -2.04E+05 | -2.01E+05 | -1.99E+05 | -1.95E+05 | -1.93E+05 |
| 1 | 4 | 0 | 217000 | -2.17E+05 | -2.06E+05 | -2.03E+05 | -2.01E+05 | -1.97E+05 | -1.95E+05 |
| 0 | 3 | 3 | 220000 | -2.20E+05 | -2.09E+05 | -2.06E+05 | -2.04E+05 | -2.00E+05 | -1.98E+05 |
| 1 | 4 | 1 | 222000 | -2.22E+05 | -2.11E+05 | -2.08E+05 | -2.06E+05 | -2.02E+05 | -2.00E+05 |
| 1 | 3 | 3 | 227000 | -2.27E+05 | -2.16E+05 | -2.13E+05 | -2.11E+05 | -2.07E+05 | -2.05E+05 |
| 2 | 4 | 0 | 237000 | -2.37E+05 | -2.26E+05 | -2.23E+05 | -2.21E+05 | -2.17E+05 | -2.15E+05 |
| 2 | 4 | 1 | 242000 | -2.42E+05 | -2.31E+05 | -2.28E+05 | -2.26E+05 | -2.22E+05 | -2.20E+05 |
| 0 | 4 | 2 | 245000 | -2.45E+05 | -2.34E+05 | -2.31E+05 | -2.29E+05 | -2.25E+05 | -2.23E+05 |
| 2 | 3 | 3 | 247000 | -2.47E+05 | -2.36E+05 | -2.33E+05 | -2.31E+05 | -2.27E+05 | -2.25E+05 |
| 1 | 4 | 2 | 252000 | -2.52E+05 | -2.41E+05 | -2.38E+05 | -2.36E+05 | -2.32E+05 | -2.30E+05 |
| 2 | 4 | 2 | 272000 | -2.72E+05 | -2.61E+05 | -2.58E+05 | -2.56E+05 | -2.52E+05 | -2.50E+05 |
| 0 | 4 | 3 | 295000 | -2.95E+05 | -2.84E+05 | -2.81E+05 | -2.79E+05 | -2.75E+05 | -2.73E+05 |
| 1 | 4 | 3 | 302000 | -3.02E+05 | -2.91E+05 | -2.88E+05 | -2.86E+05 | -2.82E+05 | -2.80E+05 |
| 2 | 4 | 3 | 322000 | -3.22E+05 | -3.11E+05 | -3.08E+05 | -3.06E+05 | -3.02E+05 | -3.00E+05 |
| 1 | 0 | 0 | 7000 | -7.00E+03 | -7.00E+03 | -7.00E+03 | -7.00E+03 | -7.00E+03 | -7.00E+03 |

**A8. Expected NPV, ROI and expected flows for strategies when disclosure cost equals 10 million, for five-year project**

| HUM | BAR | ID | INV | Q5_NPV | Q25_NPV | E_NPV | Q50_NPV | E_ROI |
|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 1 | 5000 | -5.00E+03 | 8.76E+03 | 1.30E+04 | 1.55E+04 | 2.61E+00 |
| 1 | 0 | 1 | 12000 | -1.20E+04 | 1.76E+03 | 6.08E+03 | 8.52E+03 | 5.07E-01 |
| 0 | 1 | 0 | 15000 | -1.50E+04 | -1.22E+03 | 3.10E+03 | 5.56E+03 | 2.07E-01 |
| 0 | 0 | 0 | 0 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 |
| 0 | 1 | 1 | 20000 | -2.00E+04 | -6.05E+03 | -1.54E+03 | 9.15E+02 | -7.71E-02 |
| 1 | 1 | 0 | 22000 | -2.20E+04 | -8.22E+03 | -3.84E+03 | -1.44E+03 | -1.74E-01 |
| 1 | 0 | 0 | 7000 | -7.00E+03 | -7.00E+03 | -7.00E+03 | -7.00E+03 | -1.00E+00 |
| 1 | 1 | 1 | 27000 | -2.70E+04 | -1.31E+04 | -8.48E+03 | -6.09E+03 | -3.14E-01 |
| 2 | 0 | 1 | 32000 | -3.20E+04 | -1.82E+04 | -1.39E+04 | -1.14E+04 | -4.33E-01 |
| 0 | 0 | 2 | 35000 | -3.50E+04 | -2.12E+04 | -1.68E+04 | -1.44E+04 | -4.81E-01 |
| 2 | 0 | 0 | 27000 | -2.70E+04 | -2.31E+04 | -2.23E+04 | -2.21E+04 | -8.26E-01 |
| 1 | 0 | 2 | 42000 | -4.20E+04 | -2.82E+04 | -2.38E+04 | -2.14E+04 | -5.66E-01 |
| 2 | 1 | 0 | 42000 | -4.20E+04 | -2.81E+04 | -2.38E+04 | -2.13E+04 | -5.66E-01 |
| 2 | 1 | 1 | 47000 | -4.70E+04 | -3.31E+04 | -2.85E+04 | -2.61E+04 | -6.07E-01 |
| 0 | 1 | 2 | 50000 | -5.00E+04 | -3.61E+04 | -3.15E+04 | -2.91E+04 | -6.31E-01 |
| 1 | 1 | 2 | 57000 | -5.70E+04 | -4.31E+04 | -3.85E+04 | -3.61E+04 | -6.75E-01 |
| 2 | 0 | 2 | 62000 | -6.20E+04 | -4.82E+04 | -4.37E+04 | -4.13E+04 | -7.05E-01 |
| 2 | 1 | 2 | 77000 | -7.70E+04 | -6.31E+04 | -5.85E+04 | -5.61E+04 | -7.60E-01 |
| 0 | 0 | 3 | 85000 | -8.50E+04 | -7.11E+04 | -6.67E+04 | -6.42E+04 | -7.84E-01 |
| 0 | 2 | 0 | 90000 | -9.00E+04 | -7.61E+04 | -7.16E+04 | -6.92E+04 | -7.96E-01 |
| 1 | 0 | 3 | 92000 | -9.20E+04 | -7.81E+04 | -7.36E+04 | -7.12E+04 | -8.00E-01 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0 | 2 | 1 | 95000 | -9.50E+04 | -8.11E+04 | -7.65E+04 | -7.41E+04 | -8.06E-01 |
| 1 | 2 | 0 | 97000 | -9.70E+04 | -8.31E+04 | -7.86E+04 | -7.62E+04 | -8.10E-01 |
| 0 | 1 | 3 | 100000 | -1.00E+05 | -8.61E+04 | -8.15E+04 | -7.91E+04 | -8.15E-01 |
| 1 | 2 | 1 | 102000 | -1.02E+05 | -8.81E+04 | -8.35E+04 | -8.11E+04 | -8.18E-01 |
| 1 | 1 | 3 | 107000 | -1.07E+05 | -9.31E+04 | -8.85E+04 | -8.61E+04 | -8.27E-01 |
| 2 | 0 | 3 | 112000 | -1.12E+05 | -9.81E+04 | -9.36E+04 | -9.12E+04 | -8.36E-01 |
| 2 | 2 | 0 | 117000 | -1.17E+05 | -1.03E+05 | -9.86E+04 | -9.62E+04 | -8.43E-01 |
| 2 | 2 | 1 | 122000 | -1.22E+05 | -1.08E+05 | -1.04E+05 | -1.01E+05 | -8.48E-01 |
| 0 | 2 | 2 | 125000 | -1.25E+05 | -1.11E+05 | -1.07E+05 | -1.04E+05 | -8.52E-01 |
| 2 | 1 | 3 | 127000 | -1.27E+05 | -1.13E+05 | -1.09E+05 | -1.06E+05 | -8.54E-01 |
| 1 | 2 | 2 | 132000 | -1.32E+05 | -1.18E+05 | -1.13E+05 | -1.11E+05 | -8.60E-01 |
| 0 | 3 | 0 | 135000 | -1.35E+05 | -1.21E+05 | -1.17E+05 | -1.14E+05 | -8.64E-01 |
| 0 | 3 | 1 | 140000 | -1.40E+05 | -1.26E+05 | -1.22E+05 | -1.19E+05 | -8.68E-01 |
| 1 | 3 | 0 | 142000 | -1.42E+05 | -1.28E+05 | -1.24E+05 | -1.21E+05 | -8.70E-01 |
| 1 | 3 | 1 | 147000 | -1.47E+05 | -1.33E+05 | -1.28E+05 | -1.26E+05 | -8.74E-01 |
| 2 | 2 | 2 | 152000 | -1.52E+05 | -1.38E+05 | -1.34E+05 | -1.31E+05 | -8.78E-01 |
| 2 | 3 | 0 | 162000 | -1.62E+05 | -1.48E+05 | -1.44E+05 | -1.41E+05 | -8.86E-01 |
| 2 | 3 | 1 | 167000 | -1.67E+05 | -1.53E+05 | -1.49E+05 | -1.46E+05 | -8.89E-01 |
| 0 | 3 | 2 | 170000 | -1.70E+05 | -1.56E+05 | -1.52E+05 | -1.49E+05 | -8.91E-01 |
| 0 | 2 | 3 | 175000 | -1.75E+05 | -1.61E+05 | -1.57E+05 | -1.54E+05 | -8.94E-01 |
| 1 | 3 | 2 | 177000 | -1.77E+05 | -1.63E+05 | -1.58E+05 | -1.56E+05 | -8.95E-01 |
| 1 | 2 | 3 | 182000 | -1.82E+05 | -1.68E+05 | -1.63E+05 | -1.61E+05 | -8.98E-01 |
| 2 | 3 | 2 | 197000 | -1.97E+05 | -1.83E+05 | -1.79E+05 | -1.76E+05 | -9.06E-01 |
| 2 | 2 | 3 | 202000 | -2.02E+05 | -1.88E+05 | -1.84E+05 | -1.81E+05 | -9.08E-01 |
| 0 | 4 | 0 | 210000 | -2.10E+05 | -1.96E+05 | -1.92E+05 | -1.89E+05 | -9.12E-01 |
| 0 | 4 | 1 | 215000 | -2.15E+05 | -2.01E+05 | -1.97E+05 | -1.94E+05 | -9.14E-01 |
| 1 | 4 | 0 | 217000 | -2.17E+05 | -2.03E+05 | -1.98E+05 | -1.96E+05 | -9.15E-01 |
| 0 | 3 | 3 | 220000 | -2.20E+05 | -2.06E+05 | -2.02E+05 | -1.99E+05 | -9.16E-01 |
| 1 | 4 | 1 | 222000 | -2.22E+05 | -2.08E+05 | -2.03E+05 | -2.01E+05 | -9.17E-01 |
| 1 | 3 | 3 | 227000 | -2.27E+05 | -2.13E+05 | -2.08E+05 | -2.06E+05 | -9.18E-01 |
| 2 | 4 | 0 | 237000 | -2.37E+05 | -2.23E+05 | -2.19E+05 | -2.16E+05 | -9.22E-01 |
| 2 | 4 | 1 | 242000 | -2.42E+05 | -2.28E+05 | -2.24E+05 | -2.21E+05 | -9.24E-01 |
| 0 | 4 | 2 | 245000 | -2.45E+05 | -2.31E+05 | -2.27E+05 | -2.24E+05 | -9.25E-01 |
| 2 | 3 | 3 | 247000 | -2.47E+05 | -2.33E+05 | -2.29E+05 | -2.26E+05 | -9.25E-01 |
| 1 | 4 | 2 | 252000 | -2.52E+05 | -2.38E+05 | -2.33E+05 | -2.31E+05 | -9.26E-01 |
| 2 | 4 | 2 | 272000 | -2.72E+05 | -2.58E+05 | -2.54E+05 | -2.51E+05 | -9.32E-01 |
| 0 | 4 | 3 | 295000 | -2.95E+05 | -2.81E+05 | -2.77E+05 | -2.74E+05 | -9.37E-01 |
| 1 | 4 | 3 | 302000 | -3.02E+05 | -2.88E+05 | -2.83E+05 | -2.81E+05 | -9.39E-01 |
| 2 | 4 | 3 | 322000 | -3.22E+05 | -3.08E+05 | -3.04E+05 | -3.01E+05 | -9.43E-01 |