



Alessandro Buoni

Fraud Detection in the Banking Sector

A Multi-Agent Approach

TURKU CENTRE *for* COMPUTER SCIENCE

TUUCS Dissertations
No 150, November 2012

Fraud Detection in the Banking Sector

A Multi-Agent Approach

Alessandro Buoni

*To be presented, with the permission of the Faculty of Department of
Information Technologies of Åbo Akademi, for public criticism in
Auditorium Gamma on November 8, 2012, at 12 noon.*

Åbo Akademi University
Department of Information Technologies
Institute for Advanced Management Systems Research
Joukahainengatan 3-5A FIN-20520

2012

Supervisors

Prof. Christer Carlsson
Institute for Advanced Management Systems Research
Department of Information Technologies
Åbo Akademi University
Joukahainengatan 3-5 A FIN-20520
Åbo
Finland

Prof. Mario Fedrizzi
Department of Computer and management sciences
University of Trento
Via Inama, 5 - 38122 Trento
Italy

Reviewers

Prof. Hannu Salmela
Department of Management, Information Systems Science
Turku School of Economics
FI-20014 Turun yliopisto Turku
Finland

Dr. Andrea Molinari
Department of Computer and Management Sciences
University of Trento
Via Inama 5, I-38122 Trento
Italy

Opponent

Prof. Colin Eden
Strathclyde Business School
Department of Management Science
University of Strathclyde
199 Cathedral Street G4 0QU Glasgow
United Kingdom

ISBN 978-952-12-2801-8
ISSN 1239-1883

Abstract

Fraud is an increasing phenomenon as shown in many surveys carried out by leading international consulting companies in the last years. Despite the evolution of electronic payments and hacking techniques there is still a strong human component in fraud schemes.

Conflict of interest in particular is the main contributing factor to the success of internal fraud.

In such cases anomaly detection tools are not always the best instruments, since the fraud schemes are based on faking documents in a context dominated by lack of controls, and the perpetrators are those ones who should control possible irregularities.

In the banking sector audit team experts can count only on their experience, whistle blowing and the reports sent by their inspectors.

The Fraud Interactive Decision Expert System (FIDES), which is the core of this research, is a multi-agent system built to support auditors in evaluating suspicious behaviours and to speed up the evaluation process in order to detect or prevent fraud schemes. The system combines Think-map, Delphi method and Attack trees and it has been built around audit team experts and their needs.

The output of FIDES is an attack tree, a tree-based diagram to "systematically categorize the different ways in which a system can be attacked". Once the attack tree is built, auditors can choose the path they perceive as more suitable and decide whether or not to start the investigation.

The system is meant for use in the future to retrieve old cases in order to match them with new ones and find similarities.

The retrieving features of the system will be useful to simplify the risk management phase, since similar countermeasures adopted for past cases might be useful for present ones.

Even though FIDES has been built with the banking sector in mind, it can be applied in all those organisations, like insurance companies or public organizations, where anti-fraud activity is based on a central anti-fraud unit and a reporting system.

Sammanfattning

Bedrägerier ökar i antal och får alltmer avancerade former vilket ett flertal studier visar som genomförts och rapporterats av internationella konsultbyråer. Man kunde tro att den avgörande faktorn är den ökande digitaliseringen av betalningssystem och de alltmer avancerade programmeringslösningar som spritts bland dem som begår brott inom cyberrymden ("hackers", "crackers" och nyare kategorier av brottslingar). Likväl har det visat sig att det fortfarande finns ett stort och betydande inslag av den mänskliga komponenten i de flesta bedrägerischeman.

Intressekonflikter är en central faktor i de flesta fall då någon genomfört ett framgångsrikt bedrägeri.

De publicerade studierna visar också att automatiserade verktyg för att spåra anomalier i betalningsprocesser inte alltid är användbara när det är fråga om brottslingar som befinner sig i en sådan position i organisationen att de kan följa med hur säkerhetsrutinerna sätts upp och administreras; i vissa fall är de t.o.m. ansvariga för säkerhetssystemen. Om säkerhetsrutinerna bygger på kontroll av betalningstrafiken genom att följa upp anomalier i dokumentationen kan de enkelt sättas ur spel om t.ex. dokumenten förfalskas på rätt sätt.

Inom banksektorn är de bankens interna revisorer som har ansvar för att uppdaga interna bedrägerier. Ofta kan dessa experter inte lita till annat än sin erfarenhet, tips från misstänksamma medarbetare eller rapporter över transaktioner som haft någon form av fel.

Det finns därför ett behov av stödsystem som kunde göra internrevisorernas arbete mera systematiskt och ge dem en bättre chans att komma bedrägerier på spåret (eller snabbt kunna avfärda misstankar som riktats mot någon medarbetare i banken).

I avhandlingsarbetet har jag utvecklat ett stödsystem för internrevisorer inom banksektorn kallat FIDES [*Fraud Interactive Decision Expert System*] vilket är ett fleragentsystem som utvecklats för att hjälpa internrevisorerna med att arbeta igenom misstänkta betalningsprocesser som kan ha uppkommit genom bedrägeri eller i bästa fall komma ett pågående bedrägeri på spåren och kunna avstyra det innan det kunnat genomföras. FIDES kombinerar flera olika metoder: *think maps*, *Delphi* och *attack trees* som

samtliga utformats såatt de följer och stöder de olika faserna i internrevisorernas arbete. Ansatsen med ett stödsystem för experter kan anpassas och användas inom andra sektorer såsom försäkringssektorn och den offentliga förvaltningen.

Acknowledgements

A Ph.D process requires a lot of effort and support from other people.

It is a pleasure for me to thank these people now.

First and foremost I would like to thank my supervisor, Prof. Christer Carlsson. It has been an honor to be his Ph.D student.

He taught me how proper research must be done and his passion for his job has been a constant motivation throughout my Ph.D journey.

I would like to express my deepest gratitude to Prof. Fedrizzi for initiating me into my field of research, but most of all for his patience, moral support and constant encouragement.

I would also like to thank Jòzsef Mezei for his help, but also for his friendship and the many interesting discussions we had.

I am sincerely grateful to the reviewers of my thesis Dr. Andrea Molinari and Prof. Hannu Salmela for their constructive and encouraging comments. I would like to thank all my colleagues at IAMSR and TUCS for providing a stimulating environment for my research, especially Dr. Frank Tetard for his moral support.

Lastly, and most importantly, I wish to thank my parents and my girlfriend, without whom this thesis would not have been written.

To them I dedicate this thesis.

List of original publications

1. Buoni, A. (2010). Fraud detection: from basic techniques to a multi-agent approach, *International Conference on Management and Service Science*, 24-26 August, Wuhan, 1-4.
2. Buoni, A., Fedrizzi, M., and Mezei, J. (2010). A Delphi-Based Approach to Fraud Detection Using Attack Trees and Fuzzy Numbers, In *Proceeding of the IASK International Conferences*, 21-28.
3. Buoni, A., Fedrizzi, M., and Mezei, J. (2011). Combining attack trees and fuzzy numbers in a multi-agent approach to fraud detection, *Int. J. Electronic Business*, Vol. 9, No. 3, 186-202.
4. Buoni, A., Fedrizzi, M. (2012). Consensual dynamics and Choquet Integral in an attack tree-based fraud detection system, *Proceedings of the 4th International Conference on Agents and Artificial Intelligence*, Volume 1, Algarve (Portugal), 6-8 February, 238-288.

Contents

I	Research summary	1
1	Introduction	3
1.1	Fraud in the banking sector	4
1.2	Basel II and III accord	13
1.3	Why do people commit fraud?	15
1.4	Methodology	18
1.5	Research questions, contributions	22
1.6	Overview of the thesis	24
2	Audit team methodology in the ICT era	25
2.1	Anomaly detection tools	27
2.2	Fuzzy reasoning systems	29
2.3	Tree-based detection modelling	31
2.4	Knowledge-based architectures	32
2.5	Data-mining approach in fraud detection	36
3	FIDES and its components	39
3.1	Delphi method	39
3.2	Think-map	42
3.3	Attack-trees	48
3.4	FIDES	52
3.5	SWOT analysis of FIDES	73
4	Fraud detection processes	75
4.1	ICT based	75
4.2	Human factors based fraud	77
5	Fraud in other contexts	79
5.1	Insurance industry	79
5.2	Fraud in the European Union and public sector	83
5.3	Money laundering	88
5.4	FIDES in other contexts	92

6	Summary, Conclusions, Future research	95
6.1	Future research	97
II	Original publications	109

List of Figures

1.1	Initial Detection on Occupational Fraud (Acfé, 2008)	5
1.2	Type of Fraud Experienced During the Prior 12 Months (percentages) (KPMG, 2003)	6
1.3	Factors contributing to fraud in Organisations (Percentages) (KPMG, 2003)	6
1.4	Time at the organisation (KPMG, 2011)	7
1.5	Methods used to override controls (KPMG, 2011)	8
2.1	Patterns Generation for Zipf Analysis (Huang et al., 2008) . .	28
2.2	Patterns Generation (Huang et al., 2008)	28
2.3	An example of an FP-tree construction (Mukkamala et al., 2006)	31
2.4	Logic Attack Graph generator (Ou et al., 2006)	32
2.5	Logical IDS architecture components within each node (Sterne et al., 2005)	34
2.6	Dynamic Hierarchy scheme (Sterne et al., 2005)	35
3.1	An example of mind map ©Buzan.	42
3.2	A concept map showing the key features of concept maps (Novak and Cañas, 2008)	43
3.3	An example of expert skeleton concept map.(Novak and Cañas, 2008)	44
3.4	An ICF structure of the Mediatheque designed by Norman Foster. (Oxman, 2004)	45
3.5	Content analysis of design issues using the ICF methodology. (Oxman, 2004)	46
3.6	Content analysis of design issues using the ICF methodology. (Oxman, 2004)	46
3.7	Content analysis of design issues using the ICF methodology. (Oxman, 2004)	46
3.8	Web-Pad interface for linking concepts to the issue contextualism. (Oxman, 2004)	47

3.9	Web-Pad interface for linking concepts to the issue contextualism. (Oxman, 2004)	47
3.10	An example of an Attack Tree (Schneier, 1999)	49
3.11	Generalized Attack Tree	49
3.12	Attack Tree - Bank account compromise	51
3.13	Protection tree for electronic store	52
3.14	Fides in Buoni (2010)	53
3.15	Two different scenarios of FIDES (Buoni, 2010)	54
3.16	Ishikawa Diagram used for pruning	55
3.17	The architecture of FIDES	56
3.18	The Web-Pad interface	57
3.19	The keywords selection	57
3.20	The Think-map	58
3.21	Values of linguistic variable <i>speed</i> (Carlsson et al., 2004)	60
3.22	A fuzzy representation of the linguistic label "medium"	61
3.23	Possible representation with triangular fuzzy numbers	62
3.24	The final Attack tree	65
3.25	Another attack tree stored in the data-base	65
3.26	Scaling function f and sigmoid function f' .	67
5.1	Profile of respondents (Ernst & Young, 2011)	81
5.2	Fraud risk exposure faced by insurance company (Ernst & Young, 2011)	81
5.3	Different types of fraud affecting insurance companies (Ernst & Young, 2011)	82
5.4	Duration of assessment and instances of assessment and preliminary review completed in each calendar year (OLAF, 2011)	84
5.5	Amounts recovered from closed financial follow ups in € million in each calendar year (OLAF, 2011)	84
5.6	Fraud in public sector in UK (The cabinet Office Counter Fraud, 2011)	87
5.7	Universe of transactions	91
5.8	SARs part I and II	94

List of Tables

1.1	Median Loss Based on Presence of Anti-fraud Controls (Acfe, 2008)	5
1.2	Ex ante and ex post anti-fraud actions	10
1.3	Anti-fraud procedures	10
1.4	Anti-fraud procedures	11
1.5	Report of the inspectors	11
1.6	Rating System	12
2.1	Red flags used to measure the argument Attitude (Deshmuk and Talluru, 1997)	30

Part I

Research summary

Chapter 1

Introduction

Those who have already tried
nettles recognize silk

Cristina Donà

My research path started with money laundering and the complexity of the phenomenon was clear from the beginning.

“Money Laundering is the process by which criminals attempt to conceal the true origin and ownership of the proceeds of criminal activities¹”. If successful, money can lose its criminal identity and appear legitimate. The main difficulty in detecting money laundering schemes is due to the fact that money launderers operate both in legal and illegal business, often using conventional techniques to move money from one place to another. A typical example is when smugglers carry cash on their person to open bank accounts abroad in tax heaven. It is clear that these kinds of operations cannot be traced using artificial intelligence techniques. The lack of data concerning money laundering is another limitation on research in this area.

All these limitations and the opportunity to visit the audit team of an important European bank convinced me to focus my research on fraud detection.

The lesson one can learn from money laundering is that there is a gap between the literature and reality in creating money laundering detection tools.

In reality money launderers are not convicted by using state of the art AI tools against their principal crimes, but for tax evasion.

Being aware of these aspects is essential before studying fraud detection.

In the literature one can find a lot of evidence concerning the prevalence of human-based fraud. It is no a surprise that the main fraud scheme is

¹<http://www.fsc.gov.im/aml/>

identity fraud and basically all the fraud schemes can be considered as variations on the identity fraud theme. Hackers often check the garbage bins of employees to find codes and relevant information in the mail sent from the institution. External fraud can often be considered as internal since criminals use employees as Trojan horses to gather useful information about breaking into the system (rather than performing sophisticated cyber-attacks).

The episode that provided the main motivation for the present research was a meeting with employees of one of the most important European banks. This meeting offered the unique opportunity to interview the members of the audit-team, in particular the head of the risk management department and one member of the security department. The employees of the bank confirmed my hypothesis that identity fraud is the principal type of fraud scheme adopted and that the real cause of fraud is mainly conflict of interest. They also described their anti-fraud methodology: the main issue they emphasised was that their fraud detection strategy follows the economic imperative. This means that the bank does not attempt to prosecute fraud when the cost of prosecution is higher than the economic loss it can provoke.

Although fraudsters often use human-based techniques, fraud detection is still dominated by paper-based schemes and traditional statistical techniques. Based on the risk indicators they developed, auditors can make the decision to send inspectors to the bank's branches. The output of the inspections is a report that must be evaluated by auditors. As specified by the experts in the meeting, the key factor in detecting fraud is improvement of the interaction between inspectors and auditors. Given these inputs the challenge of my research is to build a multi-agent system to support the audit team in detecting fraud, based on their suggestions and their needs.

1.1 Fraud in the banking sector

In recent years, the amount of fraud cases has significantly increased as a consequence of the rapid development in Information and Communication Technology (ICT). Although the prevention measures adopted have also progressed, fraudsters have adapted their capabilities by developing new strategies. According to Acfe (2008), in the United States organisations lose on average 7 % of their revenues in fraudulent activities and corruption is the most common factor in 27 % of the cases.

The main industries considered in the study are: banking and financial services (15% of cases), government (12%) and healthcare (8%). Regarding how the cases have been detected, a tip is the most common way in 46 % of the cases, as shown in Fig. 1.1. The tip can be attributed to employees (57.7 %), customers (17.6 %) or vendors (12.3 %). Table 1.1 presents the percentage of reduction in fraud when controls have been implemented.

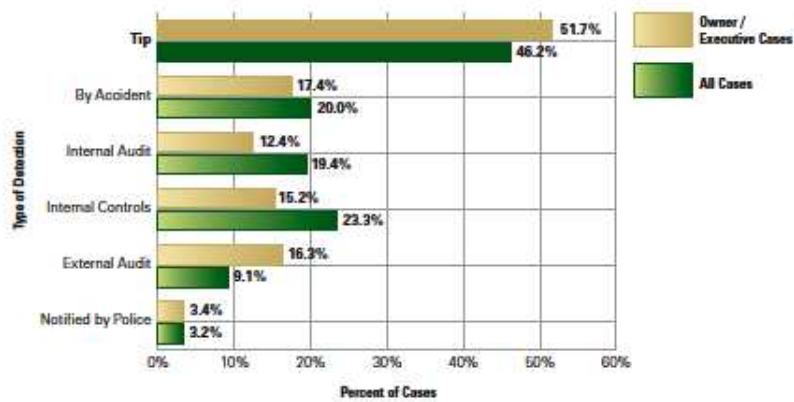


Figure 1.1: Initial Detection on Occupational Fraud (Acfe, 2008)

The two most effective controls adopted clearly suggest the importance of environmental factors in fraud. The presence of a hotline (a channel for people to describe unusual behaviours and/or suspicious activities of their colleagues) results in a significant reduction in fraud.

Control	% of cases implemented	Yes	No	% Reduction
Surprising Audit	25 %	\$ 70000	\$ 207000	66.2 %
Job rotation/ Mandatory Vacation	12.3 %	\$ 640000	\$ 164000	61.0 %
Hotline	43.5 %	\$ 100000	\$ 250000	60.0 %

Table 1.1: Median Loss Based on Presence of Anti-fraud Controls (Acfe, 2008)

As can be observed in the KPMG (2003) figures, the most frequent type of fraud experienced by auditors and fraud experts interviewed is employee fraud. (PricewaterhouseCoopers (2011) confirms the high rate of human related fraud in organisations). There is a clear prevalence of human related fraud over computer based fraud (see Fig.1.2).

Figure 1.3 demonstrates that *Collusion between employees* and third parties and *Inadequate Internal Controls* are the most common factors contributing to fraud in organisations. This indicates that employees, working

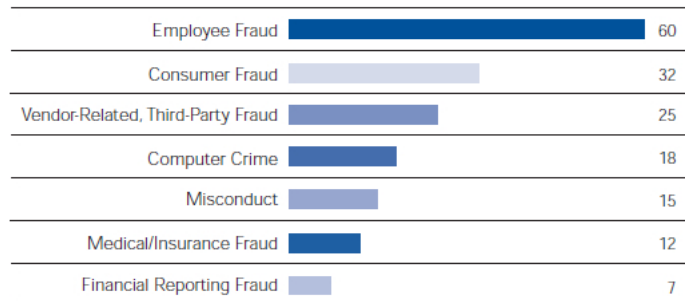


Figure 1.2: Type of Fraud Experienced During the Prior 12 Months (percentages) (KPMG, 2003)

for third parties as Trojan horses inside the institution, can mask irregularities that could be detected by internal security systems. It is interesting to observe the significant reduction of fraud cases from years 1994 and 1998 to 2003 in the category *inadequate internal controls*. This proves that improving internal control procedures can be a successful way of reducing fraud.

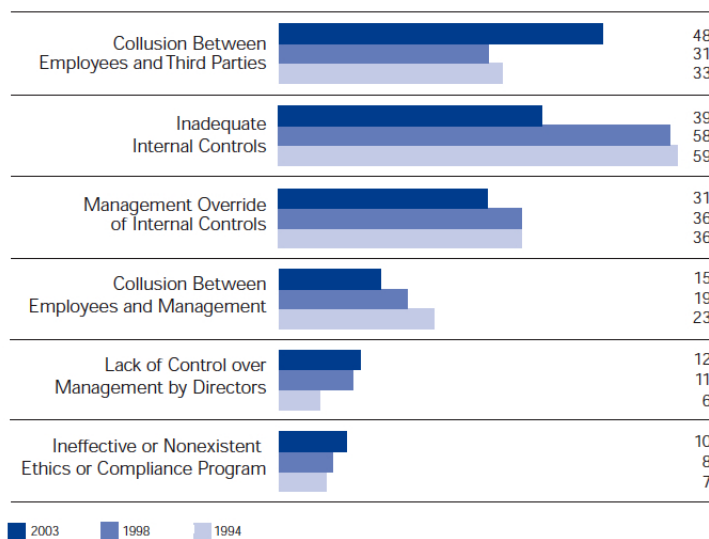


Figure 1.3: Factors contributing to fraud in Organisations (Percentages) (KPMG, 2003)

According to KPMG (2007) the profile of a fraudster is the following:

- 36-55 years old (70% of the cases)
- male (85%)

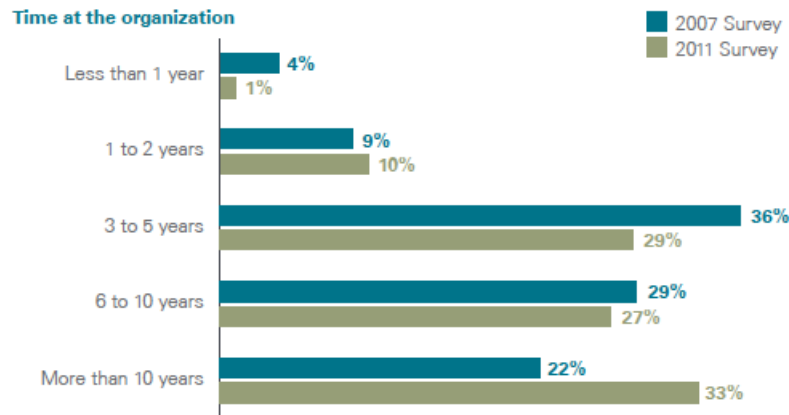


Figure 1.4: Time at the organisation (KPMG, 2011)

- acting independently (68%)
- defrauding his own employer (89%)
- someone who had been working in the institution for 2-5 years before committing fraud (36%)

The emerging picture is: a senior manager with deep knowledge of internal weaknesses and procedures, acting out of greed (73 %) and performing multiple fraudulent transactions for a relatively long period of time.

KPMG (2011) confirms the same profile and features of the typical fraudster. Additionally the survey points out that many of the fraudsters work several years in the institution before committing fraud although they are considered greedy and deceitful by nature. This implies that other factors such as financial worries, job dissatisfaction and aggressive targets, play an important role in the decision to commit fraud, even after the perpetrator acquired knowledge of the institution and gained the trust and respect of colleagues. The survey in Fig. 1.4 indeed reveals that there was an increase in the detection of fraud among long-term employees in 2011. In particular, 60% of fraudsters had worked at the company for more than 5 years before the fraud was detected, while 33 % of fraudsters had been employed there for more than 10 years.

Another strong motivation to commit fraud underlined in the survey is the need to hide losses or poor performance in order to earn bonuses. Authors observe the low number of fraud in companies that set achievable and realistic targets for their employees. The amount of fraud resulting from weak internal controls increased from 49 % in 2007 to 74 % in 2011 (see Fig. 1.5).

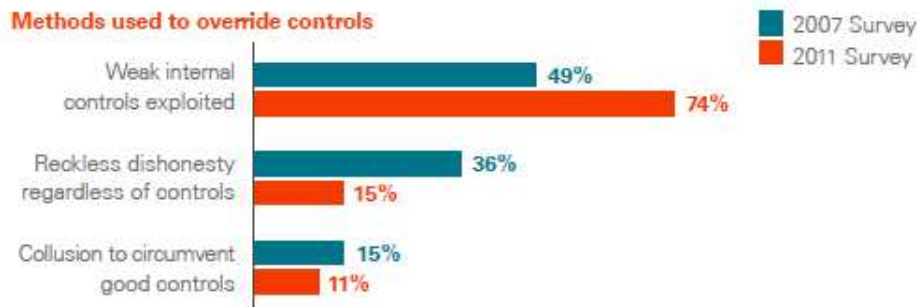


Figure 1.5: Methods used to override controls (KPMG, 2011)

In 2007, one-quarter of the fraud cases were discovered through whistle-blowers; in 2011 whistle-blower reports were used in uncovering 14 % of fraud. In 2007, 8 % of fraud were discovered by accident; this number increased to 13 % in 2011. After combining these numbers, one can observe that these informal methods of detecting fraud cover over half of the cases in the 2011 survey. These figures show how the fate of a company depends increasingly on the good conscience of employees and their motivation to act in the interest of the institution they work for.

Governments are making significant efforts to promote whistle-blowing. In the US, the Dodd-Frank Act (2010) intends to award whistle-blowers; in UK, the Public Interest Disclosure Act (1998) protects workers who are willing to cooperate, once there is good evidence of the truthfulness of their information.

Richard Powell, lead investigator of KPMG in the EMEA (Europe, Middle-East, Asia) area, notes that "Many of the fraud I've investigated in the past few years have come to light due to formal or informal whistle-blowing reports. Very few, by contrast, are discovered as a direct consequence of management, internal or external audit review".

The interviews I conducted with the members of the audit team and a member of the Computer Security Department supported these observations. The first thing they pointed out was the lack of information systems implemented in their organisation: they follow human based approaches and traditional statistical methods. The second important element they underlined is priority of economic imperative in investigating fraud cases. This tendency is also confirmed in (KPMG, 2011): in 60 % of the cases, companies recover losses in excess of \$ 25,000. This percentage rises to more than two thirds of the cases when the loss exceeded \$ 50,000.

The decision to prosecute a specific fraud case is made after an accurate cost-benefit analysis: only the fraud cases whose costs in prosecution are less than the suffered loss are prosecuted.

The analysis of the auditors involves a short description of the case, possible scenarios and identification of dangerous schemes. Experts also develop risk indicators related to different areas, to enhance their effort in cases where the possibilities of success are higher. The Computer Security Department of the Bank, follows the "know your customer" (KYC) principle, develops class risks and monitors the frequency of fraud on the total number of transactions, performs cash flow and customer profile analysis. The representative of this department pointed out that cause/effect relations between suspicious behaviours and fraudulent transactions are sometimes less important than temporal analysis in order to detect the perpetrators.

Another important aspect to consider in fraud detection is semantic analysis: the creation of a well structured scheme where the offender-crime-victim links are clear and detailed. In many cases, it can be also problematic for an audit team to interpret different suspect behaviours coming from different branches located in different European countries. In this situation, the problem is not only linguistic, but also semantic given the cultural differences between different auditors. The expert mentioned in the interview that the most common type of fraud encountered in his career is identity fraud. Other common fraud types include:

- Unauthorized transfers
- Loans to nonexistent customers
- Real estate (overpaid assessor)
- Same IP address operating in different bank accounts for a particular period of time
- Bank accounts opened with stolen cheques
- False ID (loan fraud)
- Accounting fraud
- Phishing/spyware

The expert also confirmed that the reason for the prevalence of human-based fraud is the presence of conflict of interest, which neutralises the efforts in implementing control procedures.

He also admitted the limitations of even the most advanced software in performing such complex tasks as mapping unusual behaviours.

The audit team also performs a continuous monitoring of the activities in the different areas in order to discover unusual phenomena and suspicious transactions in the normal routines of the bank.

WHEN?	WHAT?
EX-ANTE	<ul style="list-style-type: none"> • Systematic and effective mapping of the risks • Consulting on new procedures • Support on new risk areas
EX-POST	<ul style="list-style-type: none"> • Investigations in the branches • Support the evaluation of risk and fraud management

Table 1.2: Ex ante and ex post anti-fraud actions

The activities of the audit team can also be classified into two groups based on the fraud detection process: ex-ante and ex-post (see Table 1.2).

In order to prevent fraud the audit team, ex-ante, monitors all the activities in different branches of the Bank, offers consulting on new procedures to the directors of the branches and supports them once new risk areas are discovered. The ex-post activity consists of sending inspectors to the branches where fraud has been perpetrated and once they receive reports from the inspectors, they can perform the risk evaluation and fraud management procedures (see table 1.3) .

As shown in Table 1.4 there are two types of inspections: ordinary and targeted. Ordinary inspections involve: a, the control of the application of anti-fraud procedures and b, targeted inspections specifically directed towards single suspected operations, a particular customer or a specific risk area.

WHERE?	WHAT?
Procedures	Interview with the manager responsible of that procedure
AUDIT TEAM	Analysis of the reports, strategy development

Table 1.3: Anti-fraud procedures

WHERE?	WHAT?
Retail Banking	Ordinary Inspection: applications of anti-fraud procedures, reports Targeted inspections: on specific operation/customer/risk areas

Table 1.4: Anti-fraud procedures

The inspection also involves an interview with the managers of the branch in order to verify that all the procedures are followed. Finally the audit team starts to develop the anti-fraud strategy and the risk estimation based on the analysed reports (see Table 1.4). Table 1.5 shows an example of a report form used by inspectors after the inspection of a branch of the Bank.

Objective	a short description of the problem, objectives and limits of the inspections
Results	a description of the inspections and results, explanations of the causes, people involved
Responsibilities	the names of people directly involved in the operation
Conclusions	the names of people directly involved in the operation
Suggestions	a list of useful details about the inspections and personal suggestions

Table 1.5: Report of the inspectors

Table 1.6 highlights the rating system used by the audit team. Risk is calculated in 3 different areas: Credit, Investment and Finance. Priorities of the investigation are decided according to the total score obtained in each branch of the Bank: the branch with the highest score is scheduled to be inspected first.

The KPMG figures and the description of the methodology used by the audit team members indicate that human aspects play an important role in fraud. The human factor has great importance in fraud (collusion between parties, conflict of interest), but also in the detection and prevention of it: the activities of audit teams benefit from experience, intuition and the ability of the auditors to create unusual paths. Auditors have deep knowledge of the processes carried out inside the departments and they can figure

WHAT?	ADVANTAGES
Rating System	Risk Measurement in the 3 main different business areas: Credit, Investment and Finance
Continous Auditing	Analyze unusual phenomena and transaction in the 3 main different business areas

Table 1.6: Rating System

out possible schemes using their expertise and/or the information gathered through whistle-blowers. Using a cognitive approach Grazioli et al. (2006) examine the evaluation process of auditors in fraud detections, exploring the reasons for their success and failure. The authors claim that errors constitute an important issue in detecting fraud. They point out that errors are not necessarily the result of ignorance or lack of accuracy in a detection process. Errors of interpretation, for instance, can be caused by the unwillingness to criticize the opinion of a colleague or the concern of running into a false positive. In some cases, success in fraud detection is the result of interaction between different errors: the achievement of a right outcome for the wrong reasons. This suggests that auditors often follow unusual paths of reasoning. Grazioli et al. (2006) did not observe any fatal error in the evaluation of the fraud cases. Their explanation is that "fraud detection success and failure probably depends on a pattern of errors, rather than the presence of any one specific error".

Bazerman et al. (2007) point out how the interpretation of facts suffers from ambiguity. In their study, they underline how "people tend to reach self-serving conclusions whenever ambiguity surrounds a piece of evidence". Another element of bias they point out is "the threat of being fired for delivering an unfavourable audit". To remove this bias a tool to support an audit team should possess the possibility of being anonymous and "help auditors understand the unconscious errors they make and the reasons they make them". In general conflict of interest is one of the main causes of fraud in the organisations (and in particular in the banking sector), since senior managers can limit the control procedures and escape controls. Often they can themselves be in charge of these procedures. Fraud are not perpetrated by senior managers only for greed, but also to improve their social image (cars, watches, life style) or to fake the budget to gain a promotion. In a similar way, sales managers can establish a friendly relationship with naive customers, who entrust money to them. All these behaviours can hardly be detected using data mining techniques or only when it is too late, the perpetrator is gone or has already transferred the money to a secret account.

In order to prevent this kind of fraud, (when normal procedures cannot work due to the conflict of interest), auditors should focus on discovering unusual behaviours, for instance a sudden change in the lifestyle of their employees. The activities of the audit team are based on interpretation of observations, and in particular auditors have to link suspect behaviours, information and rumours with a specific loss. The risk of a wrong audit is that it can have dangerous effects on the image of the bank; a new investigation can stop the activity of a branch even for weeks. The interpretation of a fact depends on the cultural background and the experience of the auditor and on the context where he/she works. An example that can very well describe this issue is the fact that in the South of Italy most of the external fraud loss is a result of robbery. An expert auditor has to take into consideration these cultural differences in order to suggest the right countermeasures and instruments. In other words to fight robbery you need to invest in bulletproof glass and not in the latest anti-hacking software.

In light of these considerations fraud detection benefits from the behavioural intelligence of the auditors, their experience and knowledge of the context. For example the real estate market offers many opportunities for different kinds of crime, from money laundering and tax evasion to identity fraud and everything concerning mortgage fraud schemes. In order to detect a potential fraudster or money launderer it is not sufficient to have a good KYC policy: auditors need to acquire knowledge concerning the local context, rumours about business men involved in the real estate market, their relationship with politicians, their social relations, their lifestyle and the places they frequently visit, in other words "what people think and say about them".

The representation of these behaviours and decision-making dynamics constitutes a big challenge for a decision support system developer. Once the database is populated with potential and/or real fraudulent attacks auditors can easily retrieve past cases, analyse the countermeasures adopted for similar cases, play "what-if" games. Presently such tasks are performed manually by auditors. They analyse data and develop risk indicators using traditional statistical methods. A Decision Support System (DSS) for an audit team should perform more sophisticated analysis to improve the interaction between inspectors, auditors and the reporting system by taking into consideration all the aspects described in this section.

1.2 Basel II and III accord

The scandals that shook the banking community drove the Basel Committee on Bank Supervision to issue the 2001 Basel accord. This accord focuses on operational risk, which is defined as "the risk of direct or indirect loss

resulting from inadequate or failed internal processes, people and systems, and from external events”.

According to the Basel II Accord (Basel Committee, 2001), banks are encouraged to develop sophisticated methodologies to calculate operational risk, monitor bank activities, and reinforce internal control structure and auditing in order to preserve the integrity of the managerial processes. These systems also include the use of internal and external data, scenario analysis, control factors and an accurate reporting system based on key risk indicators. Operational risk includes all non-credit and market risks (Tinca, 2007), which touch on a wide range of topics such as internal and external fraud, employment practices, work safety and management risk.

Operational risk calculation

The accord suggests three methodologies for calculating operational risk:

1. The **basic indicator approach** allows a bank to use a single indicator to determine its capital charge (20% of its average annual gross income).
2. The **standardized approach**. Banks adopting this approach must calculate a capital requirement using a different risk indicator for each one of their business lines. In order to implement this approach banks must meet certain criteria such as: demonstrate to have an operational risk management in place, systematically track relevant operational risk data, regularly report to business unit management operational risk exposures, have in place an operational management system and subject their operational management process to an independent review.
3. The **Advanced Measurement Approach (AMA)**. As specified in Basel II in the AMA, banks may use their own method for assessing their exposure to operational risk, so long as it is sufficiently comprehensive and systematic. AMA gives more flexibility to the banks as long as they demonstrate to the regulators that they have adequate protection against risks. These procedures, including state of the art of anti-fraud and day-to-day reporting systems, are more expensive to implement, but they give the banks the opportunity to reduce capital reserves.

The G20 November 2010 summit in Seoul endorsed Basel III in response to the lack of regulation in the financial systems amplified by the late 2000s financial crisis. According to KPMG (2010), the Basel III accord has two main objectives:

- To strengthen global capital and liquidity regulations with the goal of promoting a more resilient banking sector; and
- To improve the banking sector's ability to absorb shocks arising from financial and economic stress.

The instruments to achieve these objectives address three main areas:

- Capital reform (including quality and quantity of capital, complete risk coverage, leverage ratio and the introduction of capital conservation buffers and a counter-cyclical capital buffer);
- Liquidity reform (short term and long term ratios); and
- Other elements relating to general improvements to the stability of the financial system.

1.3 Why do people commit fraud?

Fraudulent behaviour is often associated with personal greed and environmental factors.

In other words individuals can become fraudsters once they understand the weakness of the system and use it for their own advantage. The simplified picture emerging from this view is that of "a few rotten apples in the barrel". The mathematical model to describe this approach is based on outliers (transactions/behaviours deviating from the usual path).

This approach assumes that there is a scientific way to discriminate between "good" and "bad" acts in an organisation and that people are rationally aware of the consequences of their choices. The countermeasures often suggested to prevent and discourage these activities are devoted to improving internal controls and codes of conduct. Unfortunately, reality is more complex. An interesting point of view and a valid counterargument is offered by Kim (2005), who introduces the venality vs. banality hypothesis.

The idea is that greed (venality hypothesis) is not a sufficient motivation to explain fraud. In fact, this approach ignores how individual decisions change dramatically in an organisational setting under ideological and psychological pressures. The banality hypothesis offers an alternative and more realistic explanation of fraud. The author illustrates the banality hypothesis with the example of the tendency to listen to the superiors in an uncritical "banal way".

Miller (1986) argues that there is a propensity for people to accept definitions of actions provided by legitimate authority.

For example the order to falsify a document is perceived as right when it is ordered by the employer since there must be a reason why he has decided

to do it. Another justification for such behaviour is not having trouble with the superiors, not necessarily because people fear the revenge of their superiors, but simply to have an easy job-life.

The easiest way of avoiding trouble is to follow the job description, which means working in an efficient way according to the boss' requests.

The efficiency principle in the job environment, as a side effect, brings social irresponsibility to the actors involved in the process. If the goal of the director of a bank is to respect the budget to get a promotion and the goal of the clerk is to follow the job description of the superior in the shortest time with the minimum effort, a fraud scheme can be seen as the most efficient way to maximize the utility of the agents involved in the process.

Inside lawyers, involved in fraud, according to Suchman (1998), show an "agnostic" ethical behaviour. If the measure of quality is efficiency and therefore it can be expressed only in quantitative terms, a "good" lawyer is the one who "provides a cost-effective vehicle for his/her client's specific interest, and in doing so, he or she also facilitates the efficient functioning of the economy as a whole". It is interesting to observe that the paradigm of efficiency does not ignore the ethical issues for lack of morality, but for lack of instruments to award a goal achieved ethically.

Rosen (2002), based on empirical findings, observes that the agnostic view of law, focusing on "adding value" has become the dominant ideology for inside counsel in the twenty-first century and it is part of the rhetoric of most business management.

In this light, KPMG's recommendation of setting up realistic targets for the employees is definitely clearer.

The importance of protecting and giving incentives to whistle-blowers is becoming a common issue also on the institutional level, but what is the real opinion of whistle-blowing amongst the members of the organisations? In most cases they are considered dissidents (Near and Miceli, 1987), deviants or traitors (Greenberger et al., 1987).

A study (Milliken, 2003) offers four reasons why people are so reluctant to be whistle-blowers. The first reason is the fear of being labelled as a "troublemaker" or "tattle-tale". The second reason is the damage to the relations with colleagues. The third one is futility, the feeling that their confessions cannot have any effects on the case since they perceive themselves as outsiders in the institution. The fourth reason is the fear of getting punished or not having a promotion.

The venality hypothesis suggests that people act autonomously having a clear distinction between what is good and what is not and they are able to calculate the risk of committing fraud. The banality hypothesis shows how much more complex reality is.

During the interview granted by the head of the risk management department of an important European bank he expressed a strong agreement

with the venality vs banality hypothesis.

An important distinction he suggested was between fraud perpetrated for personal enrichment (a clerk stealing money from his/her cashier's desk) and fraud committed to increase personal reputation, for instance a bank director who fakes reports in order to obtain a promotion.

In the first case, the entity of fraud is not so relevant, considering also the insurance of the bank, but clerks also have insurance for their mistakes. In the second case, the entity is more important. The common feature in both cases is that the real reason behind the act is not greed or reputation, but the fact that these people are working for criminal organisations to obtain loans or other kinds of benefits. This means that they are victims as well.

Ten years ago an improvement in the life-style of employees was perceived by auditors as a red-flag. Nowadays however auditors are monitoring employees who are in critical financial situation, since they can be targeted by criminal organisations.

The head of risk management has also pointed out the difference between external and internal fraud and their relation. Internal fraud is characterized by low frequency (10-20 cases per year) and high damage. The amount differs from year to year. It was 4-5 million in 2011 and 20-30 million in other years in the bank.

External fraud like theft of information and a hacking attack, on the other hand is characterized by high frequency and low damage. He emphasized how external fraud are always performed with the support of at least one bank employee. In this sense there are no pure hacking attacks and for this reason, very often external fraud are classified as internal.

A tendency he noticed in the last ten years is the increase of awareness in treating the fraud problem in the Bank. In particular, the most effective controls have been performed on suspense accounts. Suspense accounts are used to temporarily place transactions which have uncertain classification. Once they have been classified or formal mistakes are corrected they are moved to the account they belong to. Suspense accounts can be easily used by a bank director to hide illegal operations or simply for faking the budget in order to show good financial performance. Monitoring these accounts is a very effective way of preventing fraud. This type of control, which can be done remotely, must be supported by employees who can directly observe suspicious activities in the job environment.

Another important red flag is the sudden change of lifestyle of employees. Often this can be the only clue to start an investigation. The head of risk management department underlined many times how direct observation of the colleagues working with the potential fraudster is fundamental in stopping him/her. Concerning the future of fraud in terms of prevention for the bank and in relation to the evolution of fraudster strategies, he believes that : a, robberies will be less and less relevant since security systems have

reached a very high level of trustworthiness, b, the technological level of the bank is superior to that of hackers, the main concern is the economic crisis as a driving force for fraud activities and cooperation with criminal organisations.

In light of what has been said in this paragraph, understanding the reasons why people commit fraud is an important key in stopping them. The venality hypothesis is not only a trivial explanation, but also a methodological short-cut. On the other hand the banality hypothesis, supported by bank experts, takes into consideration social, psychological and environmental aspects and provides a better interpretative instrument for developing counter strategies to detect and prevent fraud.

In particular, the most effective countermeasures include the improvement of first level controls and the implementation of systems which are able to move the information quickly from whistle-blowers to the auditors of the risk management department. In order to facilitate this process an adequate program of protection for whistle-blowing must be implemented.

Finally, in order to make anti-fraud operators and public opinion more aware of the gravity of the problem, fraud should not be introduced only as an economic problem, but as a social one, since (as shown in this section) fraud operations might be only the tip of the iceberg of more complex activities related to criminal organisations.

1.4 Methodology

There are no moral phenomena,
only a moral interpretation of
phenomena.

F.W.Nietzsche

Since the focus of my research is to design a multi-agent system for fraud detection and a key factor in detecting fraud is the interaction between fraud experts, the methodology I followed in this research is based on design science and constructivism.

Design-science, as conceptualized by Simon (1996), supports a pragmatic research paradigm that calls for the creation of innovative artifacts to solve real-world problems. Thus, design-science research involves focus on the IT artifact with a high priority on application domain relevance.

Constructivist theories of learning state that the learner is not a passive recipient of knowledge, but has an active role in creating knowledge, based on the "learning by doing" approach. Learners construct their knowledge based on their experience and relationship with concepts (Oxman, 2004).

The fraud detection process, is seen in this thesis as a product of the interaction between agents. The construction and interpretation of reality happens through many stages. Auditors can decide to send inspectors to the branches of the Bank to verify whether the whistles they received are genuine or not. At this stage inspectors have to interpret the facts and report all the information to the auditors. Finally auditors, on the basis of the facts or simply suspicious behaviours described by the inspectors, can decide whether or not to start the evaluation process. In the case that they decide to examine the information collected by the inspectors, they have to interpret all these facts as well. Finally, they need to find an agreement on the counter strategy to adopt, which is a shared interpretation of their individual interpretations.

There is also the possibility of finding the truth (fraud) but deciding not to prosecute it. This is the case when the economic imperative has to be dealt with, which obliges the auditors not to prosecute a fraud if the costs of prosecutions are higher than the amount of the fraud itself.

Hevner et al. (2004) defines 7 guidelines for Design Science in Information Systems Research. In the following section I will discuss how my research can be anchored to these guidelines and the process of verification and evaluation of the proposed multi-agent system.

Guideline 1: Design as an Artifact

The aim of design science research is to build an IT artifact to address an organisational problem. The artifact is not independent of the users and the social context and it has to meet their needs according to the environment where they operate. The multi-agent system to be introduced in the next chapters has been built around the needs of audit team with the aim of improving interaction between inspectors and auditors. Based on this observation, we can claim that the artifact described in the thesis is very context sensitive. At the same time it can adapt to other contexts such as the insurance or public sector, where there is need for interaction between a risk management unit (which has to analyse facts and behaviours sent by external operators) and the employees who have the duty to verify and check these suspects in the field.

Guideline 2: Problem Relevance

The goal of IS research is to acquire knowledge and implement it to solve relevant business problems. Business organisations, on the other hand, are devoted to the maximization of the profits. Design science artifacts aim to solve problems and improve business performance. The relevance of a design science artifact is its ability to follow the needs of a constituent community that works to achieve certain goals.

The constituent community addressed in this thesis, is the community of experts, who can better rich their goals (detecting fraud) using the multi-agent systems developed here.

Guideline 3: Design Evaluation

A design artifact must be evaluated and rigorously demonstrated regarding its utility, quality and efficacy using specific evaluation methods. The design science process is an iterative process where the evaluation of the artifact relies on a continuous feedback from the experts in the field. The feedback provides the basis to improve the system and satisfy the requirements of the experts.

Hevner et al. (2004) suggest different methodologies for the design-science evaluation. Discussing the descriptive approach, the authors suggest how descriptive methods of evaluation should only be used for especially innovative artifacts for which other forms of evaluation may not be feasible. The thesis is mainly concerned with internal fraud which are characterized by low frequency and high damage: the phenomenon is not easy to observe and therefore an anti-fraud system cannot be tested by repeating the same experiment. For this reason other forms of evaluation than the descriptive one are not feasible. "Informed argument" and "Scenarios" are two elements of the *descriptive approach* for evaluating a design science artifact.

Informed Argument: *Use information from the knowledge base (e.g., relevant research) to build a convincing argument for utility of the artifact.*

In this dissertation relevant research in the field will be matched to prove the utility of the designed multi-agent system.

Scenarios: *Construct detailed scenarios around the artifact to demonstrate its utility.*

The description of the multi-agent system and its components, as described in the original publications, is enriched by practical examples and scenarios where the system can be applied. These scenarios are plausible as they are based on the literature and interviews with fraud experts who provided a narration of the most frequent schemes they have encountered.

Guideline 4: Research Contributions

The effectiveness of design-science research manifests itself in providing a clear contribution to the research field where the artifact is going to be applied. The artifact must provide solutions to unresolved problems and/or extend or apply existing knowledge base in new and

innovative ways. The main contributions of the multi-agent system will be underlined in the following chapters. The flexibility of the system is shown in describing possible applications outside the banking sectors.

Guideline 5: Research rigor

Evaluation and construction of artifacts require rigour. The principal aim in design-science evaluation is to determine *how well* an artifact works and not to theorize or prove *why* the artifact works.

This is an important aspect to consider for the peculiarity of fraud detection. A senior anti-fraud expert can observe in his/her career (20-30 years) only few relevant fraud. It is clear that to justify and understand the reasons why a system works would require an unacceptable lapse of time. It is more important to provide arguments and show clear contributions regarding the implementation of artifact.

Guideline 6: Design as research process

Design science is an iterative process based on the continuous interaction and feedback from the final users and/or experts in the field. Given this approach, looking for the best or optimal solution is not a realistic aim. Heuristic strategies, on the other hand, can provide realistic solutions and methodologies that can be implemented in the business world. Simon (1996), in order to express this concept uses the expression *satisficing* solution: a solution that works well for the specified class of problems. The building process of the multi-agent system follows this principle, as it offers the fraud experts a GDSS to provide realistic solutions to fraud problems.

Guideline 7: Communication of research

A key concept in design-science is that the artifact must be presented both to technology-oriented as well as management oriented audiences. Technology-oriented audiences are interested in the implementation of the system and a description of the use.

Zmud (1997) suggests that the presentation of design-research to a management oriented audience needs to put the emphasis on the knowledge required to effectively apply the artifact "within specific contexts for individual or organisational gain" (and not on the description of the artifact itself).

In this thesis both aspects have been taken into consideration in order to be suitable for both types of audience. The interaction process between inspectors and auditors is described by a mathematical model

and the role of the actors involved in the process and the knowledge they need to use the system are also underlined.

1.5 Research questions, contributions

In this section the research questions and the main contribution of the dissertation are introduced.

Research questions

Fraud is a serious economic problem with social consequences which are often graver than the economic loss itself. The literature and common understanding of fraud business are influenced by a fictional view of the phenomenon. Both fraud activity and its detection are represented as a battle between hackers and cyber-detectives supported by advanced data-mining techniques and the most advanced anti-hacking software.

RQ1: Does the (fictional) description of fraud given in the literature provide a truthful representation?

This research question deeply influenced the whole research process. The main goal of this research has been to build a multi-agent system to support an audit team in fraud detection. Who are then the actors/users of the system? It has been built around the needs of the bank's audit team and the inspectors of the bank.

RQ2: What are the needs of audit team inspectors involved in the fraud detection process and why is it important to build a system meeting these needs?

The main problem for an audit team is to build a database where fraud cases and suspicious behaviours are stored in order to be used in the future to facilitate the detection process, comparing the new cases with the old ones. In order to perform the retrieving process, experts also require a system that is able to link semantically suspect behaviour, people suspected of being involved in the fraud schemes and eventually the money stolen from a specific bank account. The two processes are carried out together since once a case is solved or analysed it can be stored in the database for the future. Internal fraud can be characterized by conflict of interest, collusion between employees and falsification of documents. These operations are often performed manually, therefore it is very difficult to detect them, but most of all it is hard to represent them in a user-friendly way. After the interviews with the employees of the bank, the lesson I learned was to improve the interaction between auditors and inspectors.

RQ3: How can a multi-agent system improve the interaction between auditors and inspectors?

Contributions

The research questions have been developed in an iterative process as a result of the continuous improvement of the main achievement of the thesis: a novel multi-agent systems for fraud detection. The opportunity to interview one of the most important European Banks helped me improve the features of FIDES. Another important support received from the bank's experts was the confirmation of the initial hypothesis, stating that there is a gap between the common representation of fraud and their real impact, features and importance.

All the hypotheses and considerations in this thesis rely on a continuous process of verification moving between the literature and the bank world. The novelty of the system is achieved by combining three different methods: Think-map, Delphi method and attack tree. In particular, the procedure of building the attack tree has been improved and mathematically well founded using fuzzy logic in the Delphi process.

According to Schneier (1999), who first introduced attack-trees, the creation of the attack tree requires the following steps:

1. Identify possible attack goals
2. Try to think of all possible attacks against each goal
3. Add them to the tree
4. Repeat this process down the tree until you are finished
5. Give the tree to someone else and have him think about the process and add any nodes he thinks of
6. Repeat as necessary, possibly over the course of several months.

The procedure suggested by Schneier does not include a decision support system to create the attack tree. These 6 steps can be useful for didactic purpose to conduct an experiment or as a general guideline, but step 6 especially is not acceptable in the real world.

Although a fraud scheme can be perpetrated over a long period of time and clues and information about a potential fraud case are few in the short run, in order to prevent fraud sometimes it might be appropriate to start the evaluation process on an incomplete tree rather than wait "until you are done". At that stage it would be too late and the fraudster could be gone already or he could have sent the money to a secret account.

Modern audit procedures follow the continuous auditing principle and auditors prefer to prevent rather than knowing everything about fraud once they are perpetrated. The key factor to prevent and detect fraud indeed

is to improve the interaction between auditors and inspectors in order to process the information as quickly as possible.

In this sense the output of FIDES, which is an attack-tree, offers the auditors an hypothesis of the attack. In this hypothesis, auditors can perform the risk analysis and decide whether to start the investigation or not. The time required for the information to reach the audit team is reduced (and consequently the response of the auditors). Even when the response is negative and/or the auditors decide that there is no evidence or economic convenience to start the investigation the case can be stored and retrieved in the future. Past cases can be useful for evaluating new cases, but also to retrieve sub-trees and elementary operations, which might look similar. The use of FIDES as an instrument to build up a structured data-base is an important and innovative feature of the system. Another important contribution of FIDES is its flexibility since it can be applied in all the contexts where an interaction between external inspectors and a central risk management unit plays a fundamental role.

1.6 Overview of the thesis

The dissertation is organized in the following way. In the second chapter the state of the art of AI methodologies developed to support auditing will be discussed. In the third chapter FIDES and its components will be described in detail. The description of FIDES and how it can perform the anti-fraud detection process will follow the storyline of the original publications. The evolution of FIDES discussed in section 3.4 respects the chronological order of the publication of the articles. The first version of FIDES in figure 3.14 is described in Buoni (2010). Figure 3.17 is the latest version of FIDES introduced in Buoni et al. (2010) and Buoni et al. (2011). The evolution of FIDES is the result of the meeting with the audit team of the bank, which provided valuable information to create a more detailed architecture compared to Buoni (2010). The fuzzy mechanism was introduced in Buoni et al. (2010) and Buoni et al. (2011). The application of cosine similarity to compare attack trees was introduced in Buoni et al. (2011). Finally, the consensual modelling of the attack tree and the Choquet-based evaluation are described as the main contributions of Buoni and Fedrizzi (2012). Chapter 4 describes ICT-based and human factors based fraud. The importance of this distinction will be underlined in order to better understand the fraud phenomenon. Chapter 5 provides an overview of fraud in other sectors (Insurance, European Union and public sector, Money Laundering). The last paragraph of the chapter illustrates how FIDES can be applied in these other sectors. Summary, Conclusions and Future research are given in the 6th and last chapter.

Chapter 2

Audit team methodology in the ICT era

The adoption of Information and Communication Technology (ICT) changed the way of exchanging information. Documents and reports can be sent in electronic format. This change is driving audit methodology to a new paradigm termed "continuous auditing". In this context, auditing becomes a process conducted in real time in order to adapt as fast as possible to the events connected to the business activity.

According to Canadian Institute of Chartered Accountants (CICA) and the American Institute of Certified Public Accountants (AICPA), continuous Auditing is defined as "a methodology that enables independent auditors (both internal and external) to provide written assurance on a subject matter using a series of auditors reports issued simultaneously with, or a short period of time after, the occurrence of events underlying the subject matter" (Searcy DeWayne and Woodroof, 2003).

The evolution of financial products increases the possibility for fraudsters to develop more sophisticated fraud scheme.

Kuhn and Sutton (2006) show that early fraud detection can be improved by integrating a continuous audit model in SAP and this practice could be a valid instrument to prevent fraud. In the literature one can find different solutions to support audit teams in fraud detection and auditing in general.

Chou et al. (2007) propose an agent-based system for continuous auditing termed as the agent-based continuous audit model (ABCAM). The main features of the system include continuous inspection features and providing daily or weekly reports to human auditors. Agents validate the data by comparing it to supporting data sources. For instance, data in ledgers and journals is compared to financial reports and documents are matched with informative documents such as check-lists and invoices. Documents and observations with errors or showing unusual behaviours can be easily analysed

by auditors.

Wang et al. (2008), aware of the limitations of traditional statistics analysis and data mining techniques, which can only detect attacks that share a common feature with at least one past observation, introduce an immune multi-agent system for network intrusion detection. The system is based on self-learning features and inspired by the biological immune system. The architecture consists of a multi-layer intelligent security system, which is able to detect intrusions and find new attacks through learning and memory features. Two main procedures characterize the system following the biological model: the detection of antigens and the generation of immune antibodies. These two features permit the detection of intrusion and create antibodies for new attacks with similar features.

Zhang et al. (2008) propose a fuzzy integrated method based on man-computer combined data and fuzzy expert evaluation using Delphi method. The evaluation is performed by experts and based on pre-defined risk factors. The security evaluation of the embedded system is a process in which the risk factors of the system are analysed and explained. The basic goal of this evaluation is to control the risk, once a satisfactory level of assurance is reached. In order to achieve consistent opinions and the consensus among the experts, the Delphi method is performed to adjust the fuzzy evaluation of each expert.

The target groups of the systems described so far are auditors and in general anti-fraud experts. The attacks are electronic based and the aim of these intrusions is to break into the security system in order to steal information and/or money. Implementing continuous auditing presents different pros and cons.

Singleton and Singleton (2005) analyse the benefits and limitations of implementing a continuous audit system. Advantages of the system include mitigating the risk, facilitating internal controls objectives and having instant access to information. The disadvantages are related to the structure of the company. Large companies can benefit more from continuous auditing since they can support the high cost required for implementing these systems; they have complex security issues that can justify such complex architecture and good infrastructure.

In the next sections detection methods will be classified into four groups: anomaly detection tools, fuzzy reasoning systems, tree-based detection modelling and knowledge-based architectures. The classification follows the complexity of the task to be performed.

Anomaly detection tools include the techniques which are able to generate alarms, acting as sentinels of the system.

Fuzzy reasoning systems rely on the evaluation of the risk performed by experts, using linguistic labels.

Tree-based detection modelling offers an appealing representation of

fraud for users (fraud experts) in order to profile potential fraudster behaviour.

Knowledge-based architectures are implemented to solve complex tasks and use a mix of computational intelligence techniques such as neural networks, genetic algorithms, fuzzy logic and heuristic rules.

In the last section data-mining approach in fraud detection and its limitations will be introduced.

2.1 Anomaly detection tools

This group of methods includes those ones that are able to generate alarms on the basis of recognition of anomalies, unusual behaviour and paths. These tools are particularly useful in managing massive amounts of data and reducing the searching space quickly. Bedford's and Zipf's laws can be considered as meta-rules. These rules can be positioned as sentinels like a first wall of defence for the systems, a sort of pre-filtering mechanism.

Bedford's law states that "the distribution of the first significant digits of numbers drawn from a wide variety of random distributions will have (asymptotically) a certain form." Huang et al. (2008) offer an interesting approach based on Zipf's Law. The basic idea of Zipf's Law is that "the product of frequency of the use of a word, f , and the rank order, r , is approximately constant" (see Fig. 2.1).

This means that a small number of keywords can characterize the content of a document. The same principle can be used in fraud detection. In Figure 2.2, the pattern generation for Zipf Analysis is described. Zipf Analysis is very suitable for detecting attacks with frequent sequential patterns. For future research Huang et al. (2008) suggest an evaluation of the fraud detection performance of Zipf's law by comparing Zipf analysis with other clustering algorithms, like K-means, Self-Organizing Maps and Digital Analysis.

Bedford's and Zipf's law can assist auditors in detecting anomalies in documents and in general in all kinds of digital analysis.

Fugate and Gattiger (2003) benchmark different computer intrusion detection methods like Mahalanobis distance and One-class Support Vector Machines.

Bolton and Hand (2002) provide an important distinction between fraud prevention and fraud detection. Fraud prevention includes several measures used to stop fraud from occurring in the first place. Examples of these measures can be personal identification numbers, security payment systems or SIM cards for mobile phones. On the other hand, fraud detection involves identifying fraud as quickly as possible once it has been perpetrated. Statistical methods are widely used for addressing this issue, combining su-

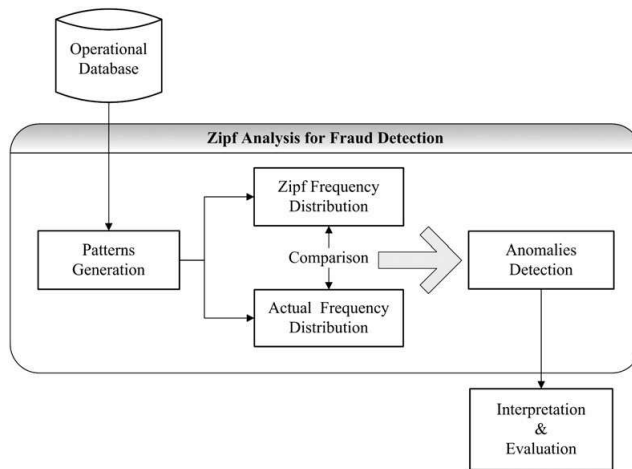


Figure 2.1: Patterns Generation for Zipf Analysis (Huang et al., 2008)

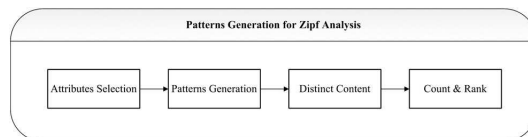


Figure 2.2: Patterns Generation (Huang et al., 2008)

pervised and unsupervised approaches. Supervised methods are introduced to classify new observations by comparing with the old ones. In unsupervised methods the goal is to find unusual and unknown patterns. The main goal of a statistical analysis is to create a suspicion score in order to rank all the suspicious activities, in a cost effective way, given the many different techniques to perpetrate fraud and the ability of the fraudsters to adapt their strategies.

Among supervised methods we can mention the traditional statistical classification approaches (Hand, 1981; Huang et al., 2008), but also more powerful methods like neural networks (Ripley, 1996; Hand, 1997; Webb, 1999).

Neural networks have been used in fraud detection for their capacity to adapt, to learn from standard behaviour and to discover new behaviour, to detect and block transactions immediately and for their ability to generalize and learn from paths. Knowledge based systems as well have been used to classify and detect suspicious transactions. The limitation of these systems is the fact that they only perform well when the behaviours cover a well structured domain, but it is well known that most of business activities are unstructured. In such situations, with interdependent and also noisy and incomplete data, neural networks can excel (KDD Cup, 1999).

Malek et al. (2008) discuss an artificial neural network (ANN) for the detection of fraud in Smart Card. The proposed fraud engine is based on a back propagation algorithm. Clusters of similar transactions are grouped into a small number of clusters. If a transaction does not fit into any cluster, it can be classified as an anomaly. The user then will be asked to authenticate himself/herself, for instance, by answering a random question. In case of a right answer the system simply allows the user to proceed, otherwise it can block him/her and suggest further investigations to the security unit.

Carpinteiro et al. (2006) propose a multilayer perceptron model for detecting attack patterns in computer networks. The input data, representing normal and attack patterns, is extracted from the files of the Third International Knowledge Discovery and Data Mining Tools Competition KDD Cup (1999). The system has been built to classify novel normal patterns and novel categories of attack patterns.

2.2 Fuzzy reasoning systems

A complex hybrid system based on different artificial intelligence techniques can be useful in recognizing hidden paths, but its power is very limited in dealing with fraud caused, for instance, by collusion between employees and third parties and lack of internal controls. In this case it might be more useful to rely on the opinions of experts, who can recognize or at least describe suspect behaviours. A team of experts or members of an anti-fraud task force could be asked to evaluate anomalies, unusual behaviours out of the ordinary routine, and produce a ranking based on specific indexes or red flags. The aggregation process and the resulting ranking based on crisp numbers could produce poor results. Fuzzy logic provides an appropriate tool to perform this task as it has the ability of dealing with imprecise information.

Deshmuk and Talluru (1997) show a rule based fuzzy reasoning system to assess the risk of management fraud using a novel measure of belief. This measure expresses the belief of the auditors in find material irregularities in the management activities. The arguments contributing to this measure are the measure of the material conditions (C) that have made the fraud possible, the motivations or the reasons (M) that pushed the fraudsters to commit the fraud and the attitude, (A) which describes the set of (un)ethical values which have inspired the fraudsters. Each argument is defined through red flags, chosen according to statistically significant factors. In Table 2.1 the argument "attitude" is defined.

Three categories are associated to each red flag: low, medium and high. Each user/expert can assign a numerical value to each red flag associated with each category/linguistic label. Based on these values, a membership

Attitude	
CO - Collusion with outsiders	MRA - Managements risk taking attitude
NCI - Need to cover up an illegal act	MD -Management dishonesty
SPA - Strong personality anomalies	ME - Management evasiveness
DRA - Disrespect for regulatory authorities	UEE - Undue emphasis on earnings
MPP - Misstatements in prior period	AA - Aggressive attitude toward financial reporting

Table 2.1: Red flags used to measure the argument Attitude (Deshmuk and Talluru, 1997)

function is built for each argument. After the three membership functions are constructed the risk of management fraud (RMF) is measured through an inference process as follows: If C is X and M is Y and A is Z then RMF is W.

A similar approach has been adopted by Bordoni and Facchinetti (2001) who developed a fuzzy expert system for insurance fraud evaluation. They reason that the main requirement of the companies is a system able to filter and split unsuspecting from suspicious claims in a automatic and fast way. The two problems can be addressed by employing a call center and fraud experts respectively. In this paper the authors illustrate the use of a Fuzzy Logic Control (FLC) model to evaluate each claim by applying an index of suspicion. The system is built using 350 rules and 69 inputs variables. The 4 output variables are the following: 1) Suspect Index: this is the main index indicating the level of suspicion of the claim 2) Element of suspect: helps the expert to detect the most suspicious areas of the claim 3) Competence: the software indicates inconsistent results 4) Body injuries: an index dealing with body injuries of the claimant.

The risk evaluation gives the opportunity to perform an accurate selection of suspect cases producing a ranking of the most important risk factors. The system, tested with the collaboration of 90 Italian insurance companies, has demonstrated effectiveness as a tool and a fast method that helps the call centers to pay unsuspecting claims immediately and filter real suspicious cases, improving the work efficiency.

2.3 Tree-based detection modelling

Trees are effective tools for systematically classifying the components of a fraudulent attack in different contexts. The graphical tree representation is appealing to users, flexible enough to be equipped with several types of information, and easily automated. Behavioural profiling consists of analysing the customer's recent activities in order to profile their normal behaviour. The following example of tree-based detection modelling refers to e-commerce fraud.

Xu et al. (2006) use the FP-tree (Frequent pattern tree) to represent relations between transactions. The aim of this research is to build a system able to detect fraud and adaptive to the dynamic behaviour patterns. Mukkamala et al (2006) present three tree-based models: FP-tree, classification regression tree (CART) and TreeNet. The FP-tree is based on a transaction database of purchased products. An initial scanning is performed in order to create an empty root. In the second stage of the process, items are placed in the tree in decreasing order of their price. Each transaction contains the class of the purchased product: daypart, grouped IP address and grouped purchased amount. Figure 2.3 shows an example of an FP-tree.

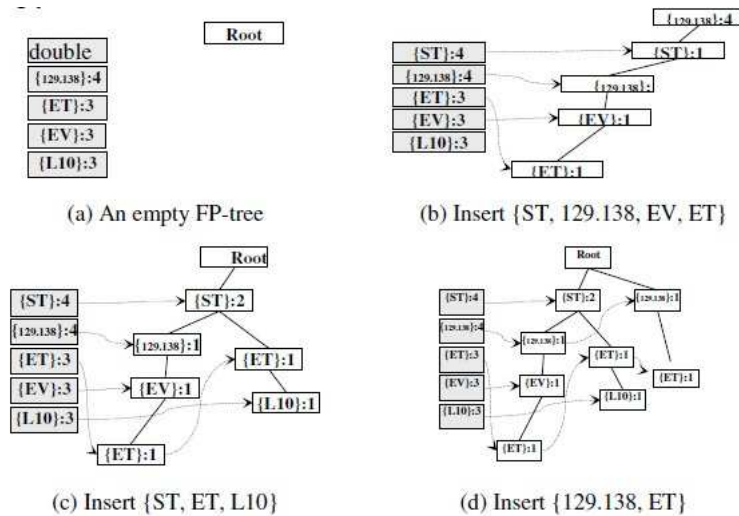


Figure 2.3: An example of an FP-tree construction (Mukkamala et al., 2006)

The external leaves of the tree suggest unusual patterns and the root nodes describe the frequency of each class. At this stage it is possible to follow all the links and obtain all the association rules. The expert/user for instance could be interested in knowing how many items have been purchased on Sunday from that IP.

CART builds classification and regression trees for predicting continuous

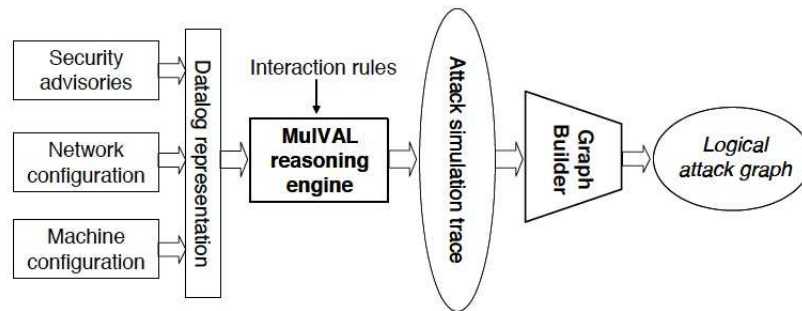


Figure 2.4: Logic Attack Graph generator (Ou et al., 2006)

dependent variables (regression) and categorical predictor variables (classification). The tree is constructed continuously until there is no significant decrease in the measure of the impurity.

TreeNet is another classification model built up through a collection of small trees with each tree improving on its predecessors through an error-correcting strategy.

Ou et al. (2006) point out the importance of taking into consideration multi-stage and multi-host attacks: an attacker could jump from one machine to another to perform different attacks. Using a logical attack, it is possible to enumerate all possible attack scenarios by depth-first traversing. In their approach they represent relations with logical expression, and generate attack graphs through automatic logic deduction, as opposed to a custom-designed graph search algorithm.

The output of a Logical Attack Graph is shown in Fig. 2.4. This attack generation tool is built upon MuIVAL, a network security analyser based on logical programming, in order to improve the security of complex network based systems.

The advantage of the logical attack graph is the improvement in understanding the causal relationship between system configuration and a successful attack.

2.4 Knowledge-based architectures

Multi-agent systems can be used to solve complex tasks involving large amounts of data using a combination of computational intelligence techniques such as neural networks, genetic algorithms, fuzzy logic and heuristic rules.

Ohsuga et al. (2001) describe three main phases in which an artificial intelligence based system is able to solve a large-scale problem: deconstruct

the problem into small sub-problems, distribute the problems to different agents, and integrate their results.

Major and Redinger (2008) develop a hybrid knowledge/statistical based system for the detection of fraud. In particular the system has been used in the healthcare fraud detection as a pre-investigative analysis tool. The proposed Electronic Fraud Detection (EFD) "integrates expert knowledge with statistical information assessment to identify cases of unusual provider behaviour. The heart of EFD is 27 behavioural heuristics, knowledge-based ways of viewing and measuring provider behaviour." The system identifies suspicious behaviours that need to be investigated. The operational cycle consists of 3 steps: 1) measure each provider's behaviour; 2) compare it with his peers; 3) refer the unusual behaviour to the Security Unit.

A machine-learning tool can also be used to create new rules and improve the detection process.

Martignoni et al. (2002) evaluate a behaviour-based malware detector. They try to correct the asymmetry present in syntax-based approaches for malware detection, defining *semantic gap* as "disconnect between a voluminous stream of low-level events and any understanding of the aggregate effect of those events". Behaviour graphs are used to describe a correlated sequence of events that have some particular semantic effect. In order to recognize complex behaviour they compose graphs hierarchically assuming that "a behaviour graph generates an event that can be used as a component within another graph." The robustness of the system is underlined by its capacity to detect different variants of malware. Given a well defined semantic structure, the system is still able to recognize variations of the same malware typology.

Flegel et al. (2008) distinguish between generic, specialized and custom-built fraud detection tools. The authors point out that most of the systems are custom-built for commercial purpose. They also argue that a different perspective is needed in detecting intrusion at host or network level. The novel concept of *syntactic gap* is introduced, which can be defined as the existing gap between real fraud, detected in the real world, and existing fraud models. An ad hoc ontology is required in order to improve the current models in terms of adaptivity and flexibility.

Sterne et al. (2005) develop a general cooperative intrusion detection architecture supporting activities in mobile ad hoc network (MANET). The model has been designed for military applications. The network discussed in this paper involves a mix of mobile systems, PDAs, laptops, which can be located on different kinds of vehicles. The presented model is an organisational model based on a dynamic hierarchy. The links in the model are dynamic since nodes can quickly activate or lose connection with their neighbours. Every node is responsible for protecting itself using its own detection system. Figure 2.5 presents an intrusion detection system (IDS) from the

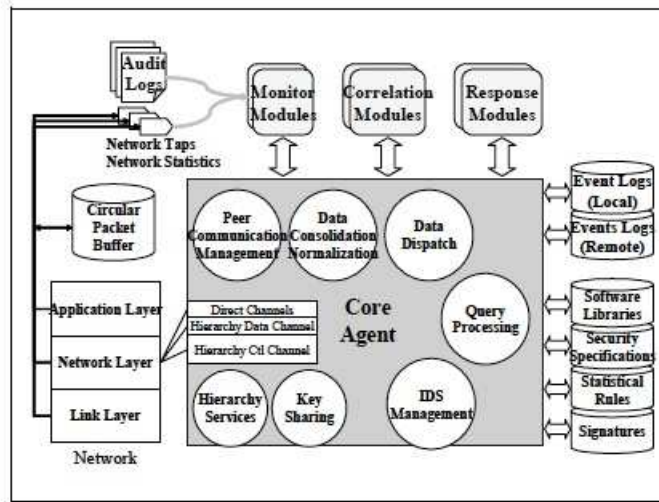


Figure 2.5: Logical IDS architecture components within each node (Sterne et al., 2005)

perspective of a single node. The grey rectangle represents the core agent of each node. Its function is to perform sensor-dependent processing and communication; circles are relevant logical processes and cylinders represent storage information units.

Figure 2.6 depicts how dynamic hierarchy works with the arrows representing hierarchical relationships. Nodes U, V, and B are children of C. In this scenario, X represents the potential attacker. The approach shown in the figure is based on link-layer monitoring and accumulation of packet counts. This means that information gathered from all neighbouring nodes is aggregated and compared. Inputs and outputs packets are monitored by neighbouring nodes. In this example, given that X is the attacker, nodes C, Z, W and B monitor the packets of X. In this way X can improve his position in the hierarchy. N, being in a lower level of the hierarchy, is in the best position to monitor X. The principle used is that intrusion detection and correlation should occur at the lowest level in the hierarchy at which the aggregated data is sufficient to enable an accurate detection or correlation decision.

Anomaly detection tools, discussed in this chapter, generate alarms and can be used by auditors to support their investigations. These systems are often introduced as exhaustive instruments to detect fraud. In the real world the number of false positives is very high and all the generated false alarms need to be evaluated and pruned by human experts.

These tools can be useful instruments to support human decisions, but they need to be constantly refined and considered for their intended purpose:

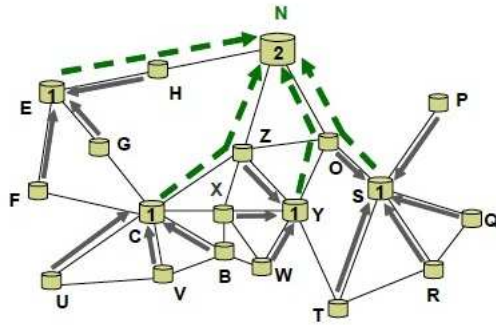


Figure 2.6: Dynamic Hierarchy scheme (Sterne et al., 2005)

algorithms that follow heuristic rules implemented by humans.

These systems must be integrated with DSS to be useful; the main concerns are the management of all these alarms and the transformation of the produced information into knowledge to detect real or potentially dangerous cases.

Another limitation of these systems is that they cannot be used to prevent human-based fraud, since human behaviour is not visible in any database. Attempts of intrusion can be detected, but the path of the fraud is visible only when it is too late and the fraud has been already perpetrated.

Fuzzy reasoning systems have the advantage of mimicking decision making dynamics and interaction between experts. In the real world, anti-fraud experts need to agree on their different evaluations on a real or potential fraud case. A fuzzy reasoning system can solve the problem of aggregating different beliefs in order to produce a shared strategy to fight against fraud.

Tree-based systems introduced in this chapter are addressed to e-commerce, but they can be used by experts for a user friendly representation of fraud and semantic oriented features. Trees can also be stored, retrieved and then matched with new cases.

Knowledge-based systems for fraud detection are more suitable for complex organisations and institutions, since they are able to support the costs of implementation and maintenance. The main limitation of these systems is the gap with the procedures applied in the real world. Audit team procedures for fraud detection are still dominated by paper-based and traditional statistical methods. Another limitation is the lack of real data about fraud schemes and the nature of fraud data itself. Fraud cases in the real world have a very low frequency and therefore it is a very difficult task to test these systems. If the goal of a knowledge-based system is to be autonomous

and self adaptive an intermediate phase is required. This phase consists in the creation of semantically encoded data-bases using DSS which are able to make the tacit knowledge of the experts explicit.

Once the knowledge is well structured new knowledge can be created. In this sense, a DSS can be seen as an intermediate step in the multi-agent system design before constructing proper knowledge-based architectures.

2.5 Data-mining approach in fraud detection

Different data-mining methods have been applied in several fields of research. The dramatic growth in data volumes - also known as the "big data" revolution - has made the analysis and pruning work very time consuming. At the same time the evolution of technology, computational power and improved computational methods have improved the capacity of analysts to interpret and manage these large and growing amounts of data. Data-mining has gradually evolved into the Knowledge Discovery Database (KDD) technology.

KDD refers to all those techniques and tools that are able to support humans in making sense of data (Fayyad et al., 1996); it appeared for the first time in 1989 (Piatetsky-Shapiro, 1991). The KDD process involves collection of data, finding paths in the data, algorithm selection and - most important - the determination of what is to be considered to be knowledge and what is not.

Phua et al. (2005) introduce a comprehensive survey of data-mining approaches in fraud detection and discuss their limitations.

The data-mining methods can utilize training/testing data with labels, but restricted to legal examples, and no labels to predict/describe the evaluation data.

1. Supervised data-mining approaches on labeled data

These methods, including supervised algorithms, examine all previous labelled transactions to mathematically determine what a standard fraudulent transaction looks like by assigning a risk score (Sherman, 2002). Wheeler and Aitken (2000) adopted a case-based (CBR) reasoning approach to analyse existing methods and techniques that produced misclassified data.

2. Hybrid data-mining approaches with labeled data

Hybrid approaches combine different techniques such as neural networks, Bayesian networks and decision trees. Chan et al. (1999) by combining naive Bayes, C4.5, CART and RIPPER as base classifiers, show that there is an improvement in terms of reducing computational costs and enhancing efficiency in credit card fraud detection.

3. Semi-supervised data-mining approaches with only legal (non-fraud) data

Murad and Pinkas (1999) edit profiles of phone calls from telecommunications accounts; common profiles are extracted and an alert is produced once a call duration and destination exceeds the threshold and standard deviation of the overall profiles.

4. Unsupervised data-mining approaches with unlabeled data

Dorransoro et al. (1997) introduced a non-linear analysis algorithm without labels to find events of credit card fraud. Since there is no history of each credit card transaction, all the transactions are segregated into different geographical locations. The authors show that the system enhanced the detection of the false positives and improved computational efficiency.

Phua et al. (2005) show that there are limitations to using the data-mining approach for fraud detection. One of the limitations is to obtain real fraud detection data. A key limitation is the emphasis on complex algorithms such as unsupervised neural networks, which contradict the empirical evidence.

Phua et al. (2005) show that less complex algorithms - such as naive Bayes and logistic regression - can produce better results. Another criticism is that few systems have been implemented but the most discouraging reality is that there has not been any empirical evaluation of commercial data-mining applications since Abbot et al. (1998).

In this thesis, data-mining techniques are used as instruments to generate alarms that can be evaluated by experts rather than being an autonomous system that can be used to detect fraud. This approach was chosen because of a number of reasons. In the real world the most dangerous fraudulent schemes are developed and run by humans and not by some kind of electronic systems. The dominant data-mining approaches paint a world dominated by hackers using sophisticated algorithms to attack the banking systems. This approach ignores conflict of interest in the world of fraud, where managers can modify documents and data to avoid controls or offer hackers all the support they need, for instance by selling usernames and passwords. These are clearly inexpensive and simple ways to commit fraud, which cannot be detected using a data-mining approach. Data-mining methods are tested on synthetic databases where the frequency of fraud usually is very high compared to the real world. Another issue is that a fraudulent operation can be perpetrated over a long period of time (5-10 years). A senior bank manager could over the course of his career find a maximum of 15-20 high impact fraudulent cases. Using a data-mining approach, the question is how much of the developed tools and methods can be extended to new cases?

Another fundamental problem is the time factor. In case of internal fraud, but also other kinds of fraud, suspicious transactions can be detected too late and when the money is already in an offshore account. Data-mining can be a valid approach to generate alarms, to block credit-cards or online transactions, but it has many limitations. In order to implement systems inspired by the KDD principles it is imperative for the fraud experts to collect a sufficient and relevant number of fraud cases, which can be done by using an interactive Decision Support System (DSS).

Chapter 3

FIDES and its components

Decision makers often operate in a surveillance mode rather than a problem-solving mode. In contrast to a theory of information that assumes that information is gathered to resolve a choice among alternatives, decision-makers scan their environments for surprises and solutions. They monitor what is going on. They characteristically do not "solve" problems; they apply rules and copy solutions from others.

James March

In this chapter the Fraud Interactive Detection Expert System (FIDES), a multi-agent system (MAS) proposed in Buoni (2010) , and in Buoni et al. (2010) and then extended in Buoni et al. (2011) and its components will be introduced. FIDES is a latin word often (and wrongly) translated as "faith" but to the Romans it meant "reliability". The target groups of the system are the auditors of the bank and inspectors. In the next paragraph the main components of FIDES will be introduced, including Delphi method, Think-map and Attack trees. In the last paragraph the anti-fraud procedure carried out by the system will be described in detail.

3.1 Delphi method

The Delphi method was used for the first time in RAND (Research AND Development), a U.S. sponsored military project, in the 1950's. The name

comes from the Greek oracle of Delphi "where necromancers foretold the future using hallucinogenic vapors and animal entrails" (Gordon, 1994).

It was designed to stimulate a debate free of the charismatic influence of the personalities involved in the process. The influence of oratory and pedagogy of the influential members of the participants is removed by the fact that extreme opinions are summarized to give equal weights to different statements. These are indeed the most important aspects of Delphi: anonymity and feedback. In this sense the Delphi method can be considered as a controlled debate. In the history of philosophy, one of the main topics that challenged philosophers is *reality*.

Many interpretations have been offered and many different epistemologies have been suggested for developing different research methods. The one proposed by D. Sam Scheele (Turrof, 2002) is the subjective or negotiated reality. This approach suits particularly well the needs of Delphi participants. In the Delphi method the reality is the product of interaction between experts and the output is the detection of illegal operations based on previous experiences and the intuition of the participants.

Rowe and Wright (1999) suggest four for a good design of the Delphi process:

1. **Anonymity of Delphi participants:** allows the participants to freely express their opinions without undue social pressures to conform to others in the group. Decisions are evaluated on their merit, rather than who has proposed the idea.
2. **Iteration:** allows the participants to refine their views in light of the progress of the group's work from round to round.
3. **Controlled feedback:** informs the participants of the other participant's perspectives, and provides the opportunity for Delphi participants to clarify or change their views.
4. **Statistical aggregation of group response:** allows for quantitative analysis and interpretation of data.

A typical Delphi process can be described in eleven steps according to Skulmoski et al. (2007) :

1) Develop the research question - The research question can be co-developed by the supervisor and the participants typically through a literature review. A pilot study can also be conducted to identify the problem and understand the relevance of the research question.

2) Design the research - In this phase, in order to help answer the research question, different kind of methods are evaluated. Normally the survey is most common way chosen to perform the Delphi process.

3) Research sample - At this stage research participants are selected. This is a critical phase since the output of Delphi is based on experts opinions. Knowledge of the problem and social and communication skills are fundamental requirements.

4) Develop Delphi round one questionnaire - In order to be sure that all the questions will be clear to all the participants, the purpose is to have a brainstorming session.

5) Delphi pilot study - Especially when there are inexperienced researchers, a pilot study may be conducted to test the comprehension of the questions and adjust them if necessary.

6) Release and analyse first round questionnaire - The questionnaires are distributed and then analysed. It is common to represent the results using mind maps in order to understand the logical connections, interactions, causes and effects of the research problem.

7) Develop second round questionnaire - On the basis of the results of the first questionnaire the questions will become more focused in order to narrow the research topic.

8) Release and analyse second round questionnaire - The questionnaire is distributed and then analysed as in the first round. Now the participants have the opportunity to verify if the responses of the first round correspond to their opinions and change or enrich them in the light of the answers of the other participants.

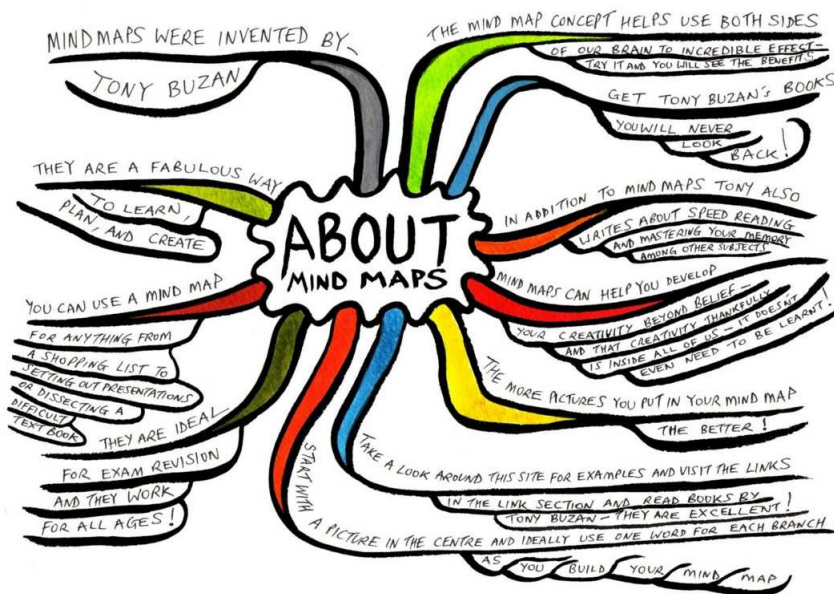
9) Develop third round questionnaire - The third round is based on the answers of the second round. New questions are added to go in deep in the research and refine the objectives.

10) Release and analyse third round questionnaire - In the third round the procedure is similar to the one used for the previous two rounds. Once again participants can change their opinions, comment and share the knowledge acquired during the process. At this stage if the consensus is reached and participants are satisfied, the process can stop.

11) Verify, Generalize and document the research results - Results are verified and a document is released to describe the research results.

The Delphi method has been used effectively in many fields of research to handle incomplete knowledge of a problem or phenomenon and to extract tacit knowledge by soliciting experts opinions.

Yao and Liu (2006) conduct a comparative study between conventional Delphi and dynamic Delphi. Dynamic Delphi is based on web technology, since survey rounds are performed online. In this version of Delphi, panellists are connected all around the world and feedback is released, only partially or fully dynamically between rounds. Their research suggests that dynamic Delphi survey may form a consensus quickly, since social-emotional exchanges may facilitate the elimination of potential misunderstandings. Delphi is best used as laying the groundwork for other methods (Franklin



© Paul Foreman <http://www.mindmapinspiration.com>

Figure 3.1: An example of mind map ©Buzan.

and Hart, 2007). This means it can be used as an instrument for setting up other methods.

3.2 Think-map

Mind mapping is an effective way to represent knowledge and share and organise new ideas, forcing users to focus on the logic construct, coherence in their reasoning and improvement of the accuracy of cause-effect relationship. According to Åhlberg (2007) the two most important names in mapping representation are Tony Buzan and Joseph Novak.

The concept of mapping was introduced by Buzan (1974). The main idea of a mind map is to organise keywords into a radiant structure that resembles a tree seen from above (Åhlberg, 2007). In a radiant structure it is possible to visualize the central idea, which can be expanded through a brainstorming process adding new branches. Figure 3.1 represents a mind map, according to the principles of Buzan (1974).

Novak and Cañas (2008) developed the idea of concept maps defined as "graphical tools for organising and representing knowledge". Concept maps (see Figure 3.2) are characterized by a hierarchical structure with the main concept at the top of the map and the less generally developed concepts placed below the main one. Another important feature of concept maps is

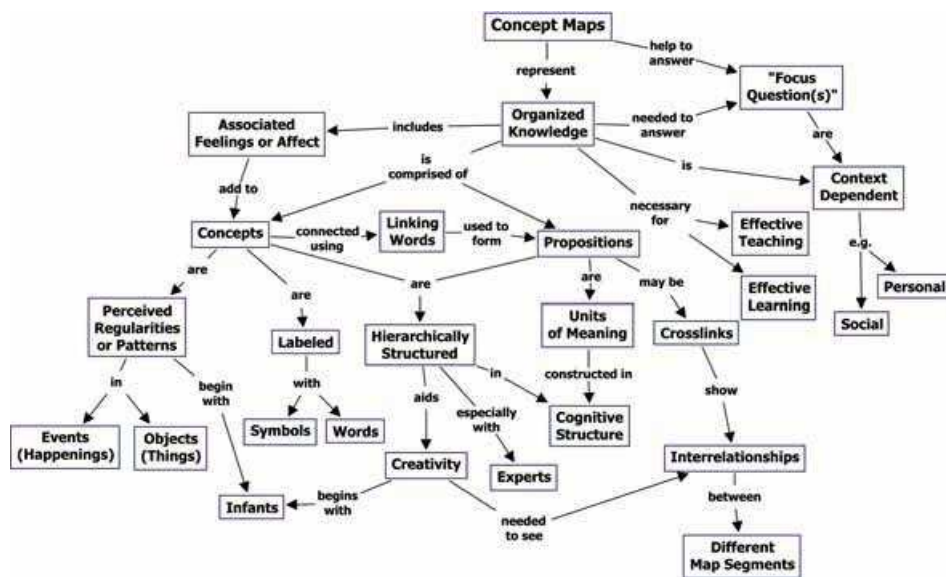


Figure 3.2: A concept map showing the key features of concept maps (Novak and Cañas, 2008)

the use of cross-links, which are concept connectors that specify the relation between labels. In Fig.3.2, cross links are "represent" and "includes" for instance.

Using concept map software like CmapTools (<http://cmap.ihmc.us/>) and with the support of a projector, it is possible to move concepts (labels) around, rebuild and expand the map as it is required. Modifications can be performed by the users in real time (synchronously) or at their convenience (asynchronously), once maps are stored in the server. Users can interact leaving comments in form of post-its or during the building of the map (Cañas et al., 2003).

Figure 3.3 depicts an example of building a concept map. A focus question is placed on the top, with the "parking lot" of listed concepts on the left, which can be moved by the users to create the map. Cross-links act as logical connectors to link concepts. In the case of particularly complex subjects an expert can build an "expert skeleton concept map" in order to offer some clues to users to complete the map.

The main contribution of concept maps (and maps in general) is to capture the tacit knowledge of the experts, forcing them to articulate their reasoning and visions.

Amongst the mind map structures found in literature, there is one which deserves more attention considering the nature of the research problem. This structure is termed as Think-map and was introduced by Oxman (2004) as

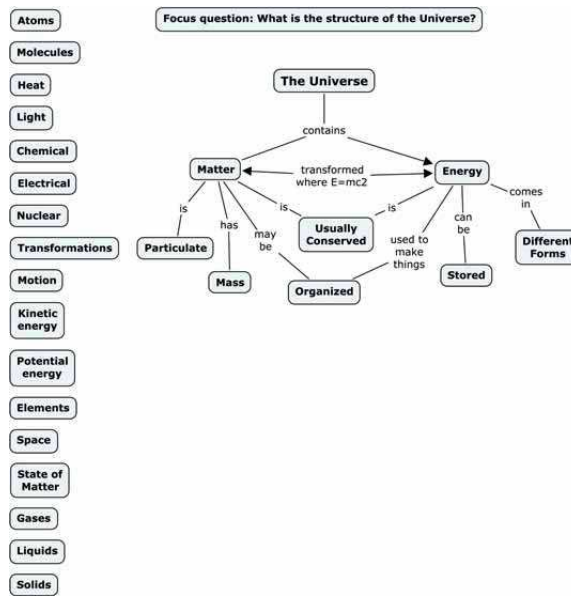


Figure 3.3: An example of expert skeleton concept map.(Novak and Cañas, 2008)

an educational instrument to support students in learning design concepts. Think-maps follow two main learning theories: constructivism and concept mapping. Constructivist theories of learning (Kolb, 1984) are built on the idea that the learner does not have a passive role in the process, but is part of the learning process, creating himself/herself and learning new concepts. Concept mapping in Think-maps acts as a tool for organizing and representing knowledge. This structure reinforces the key idea of constructivism, that is "learning by doing". The concept map structure that makes explicit conceptual mapping is called ICF (Issue-Concept-Form), which creates a network of associated concepts and acts as a "structuring ontology for the construction of conceptual networks of design concepts" (see Figure 3.4).

The ICF structure refers to the Mediatheque of Nimes (France), designed by Norman Foster. In the figure we can observe three main structures: urban attraction (issue), luminosity space (concept) and vertical space (form). For instance following the double line, we can interpret the scheme in the following way: in order to create an urban attraction (an issue), a luminosity space (a concept) was achieved by introducing a six story vertical space (a form) as a central atrium (a form). The software developed to create conceptual mapping and linkages between concepts is termed Web-Pad. The system is based on the AI methodology of Case-Based Reasoning (CBR). Web-Pad works on two levels: textual and visual. The textual level includes a case-base of design precedents that are represented textually as

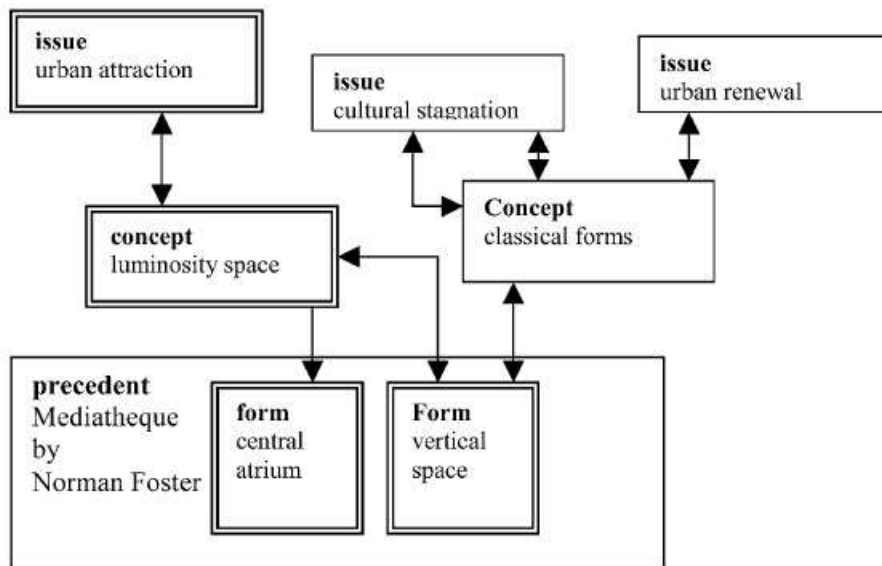


Figure 3.4: An ICF structure of the Mediatheque designed by Norman Foster. (Oxman, 2004)

a network of concepts. The visual level consists of a graphic representation of the cases. The system allows the user to perform a data-based search. As the output of the search the system returns precedents containing links to the items indicated by the user. In the CBR search the user can specify a partial description of the case and retrieve similar cases that match the entered description. To improve the system the user can express the degree of similarity between 0 and 1 comparing a new case with an old one he/she found during a previous search. In this way the system can improve its accuracy. The result of the search will be a ranking of similar cases in descending order. The following experiment was performed by a team of design students. Students were asked to create inference between concepts and significant relationships of ideas in museum design.

Figures 3.5, 3.6 and 3.7 shows the procedure to create the three main structures which form ICF: main issues, main concepts and main forms. The related keywords, associated to each label, are underlined in color.

Figure 3.8 shows the Web-Pad interface in the process of creating the issue "contextualism". Web-Pad creates "concept" and "form" labels in a similar way. The description text box shows the text that has to be analysed.

The graphic level representation as result of the issue-concept-form linking is shown in Fig.3.9.

Students in distributed environments shared their ideas and comments

When approaching the design of the Carre d'Art , we think the architect identified four main issues he believed were crucial to the planning of the building itself and to the overall development of the site as a part of the heart of the old city of Nimes.

Foster and the initiator of the project, mayor Jean Bousquet , were unanimous in the opinion that what was needed on the site was an **"anti museum"**, rather than the conventional gallery type museum .They both instinctively believed that creating an arts and media center where the citizens of Nimes could take an active part of ,would be a greater generator of attraction.

Figure 3.5: Content analysis of design issues using the ICF methodology. (Oxman, 2004)

After looking at the bigger picture, there was a need to focus in on the CONCEPTS that would eventually evolve into the physical planning of the site.

Since Bousquet and Foster had already opted for an "anti museum" , Foster thought that to achieve that he would have to design **a flexible structure** that would be able to accommodate different activities and changing needs .This , coupled with the need to work in the urban context , brought about the visual concept of the **"anti monument"** .

Figure 3.6: Content analysis of design issues using the ICF methodology. (Oxman, 2004)

The Issues identified and the planning concepts evolved, now it was time to translate all this into PHYSICAL form.

As we discussed earlier, the museum was planned as an "anti museum" and an "anti monument" through a flexible structure. It was designed to hold a varied **mix of uses** under the title of an ARTS AND MEDIA CENTER.

Since the building was situated in an existing classical urban context, there was a need to incorporate in it's design those classical qualities and forms. A visitor would meet those forms right upon entering the building, passing under the **entrance canopy** resting on steel columns, a conscious gesture to the

Figure 3.7: Content analysis of design issues using the ICF methodology. (Oxman, 2004)



Figure 3.8: Web-Pad interface for linking concepts to the issue contextualism. (Oxman, 2004)

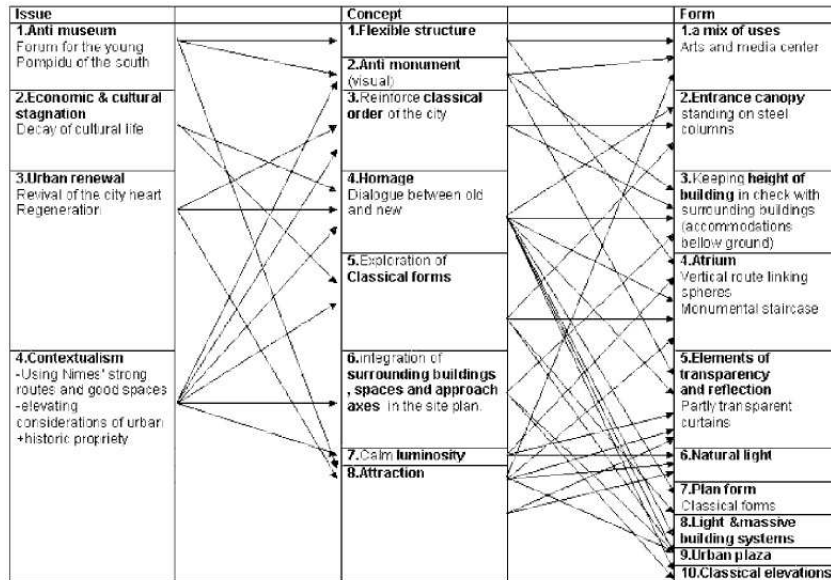


Figure 3.9: Web-Pad interface for linking concepts to the issue contextualism. (Oxman, 2004)

in real-time discussion mode. In figure 3.8, one can observe the empty field "System Message", that is used by students to interact. The goal of this task is to develop a conceptual map from the knowledge extracted from the texts. The system, once the links are created, can be browsed to perform a semantic search by selecting one label on the map with the mouse. The output is a description of the related linked concepts or an issue related to a certain form.

3.3 Attack-trees

One of the most suitable tree-based models is the attack tree. The term attack tree was introduced by Schneier (1999) to systematically categorize the different ways in which a system can be attacked. He describes an attack tree as an instrument to "provide a formal, methodical way of describing the security of systems, based on varying attacks. Basically, you represent attacks against a system in a tree structure, with the goal as the root node and different ways of achieving that goal as leaf nodes."

And and *or* operators, in the classical framework of attack trees, relate the subtask of a node. An *and* operator is used when all the features must be present to perform a particular task; in the case that at least one of them is necessary, an *or* operator will be chosen. Using this approach it is possible to analyse the contribution of any attribute to the main goal which is to finalize the attack. Another important feature is the opportunity to calculate the cost of an attack as a function of the cost of subtasks, as shown in Fig. 3.10.

At the same time it is possible to identify the path with the highest probability of success or to study the least expensive ways available for the fraudsters to perform an attack. Yager (2006) introduces the use of OWA nodes in describing the required number of children for the success of the parent. According to this approach, the contribution of the children to the success of the parents is determined by a linguistic quantifier, Q , expressing the (linguistically characterized) proportion of subtasks needed for the parent node to be accomplished. If we consider an OWA node that has n children, Q subtasks are directly responsible for the success of the parent. Q is a monotonic linguistic quantifier expressing concepts like "at most" or "at least half".

The attack tree is a powerful tool to represent an attack through nodes, analyse the contribution of a single attribute and calculate the cost of an attack (the cheapest way to perform it) or the attack with the highest probability of success.

Mauw and Oostdijk (2005) introduce a generalization of Schneiders attack trees. Figure 3.11 shows a generalized attack tree and its main elements:

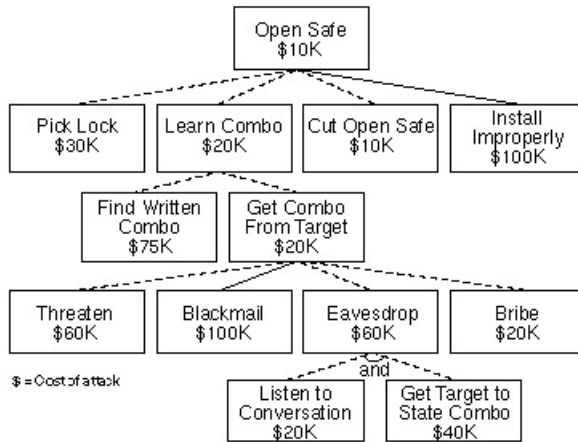


Figure 3.10: An example of an Attack Tree (Schneier, 1999)

nodes, bags and bundles. The power set of a set S , which is the set of the elementary attacks, is denoted by 2^S in Fedrizzi and Giove (2007). A bag is any subset of S containing multiple copies of the same element. For instance, if a and b are elements of S , $B = \{a, a, b\}$ is a bag. $Bag(S)$ is the set of all finite and non-empty bags from S . Given a finite set S of attack components, an attack is an element of $Bag(S)$. An attack suite is an element of $2^{(Bag(S))}$.

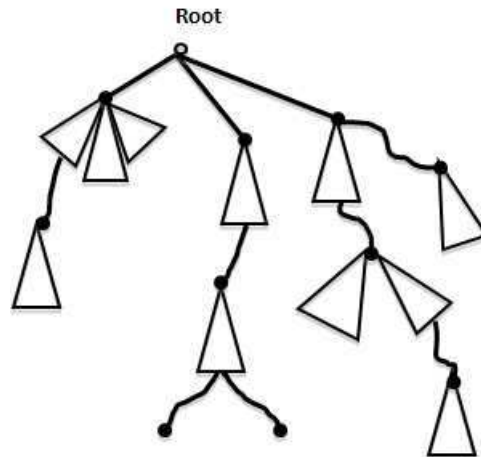


Figure 3.11: Generalized Attack Tree

In order to form an attack, we have to consider connections from a node to a bag of nodes. The set of multiple connections between a node and a bag of nodes is called a bundle. Several bundles could be associated with a given

node and the execution of any bundle of a node is sufficient to fire that node. The attack suite defined by a node can be determined by its bundles. The attack suite consists of the union of the attack suites defined by its bundles. An attack in a bundle is constructed by taking an attack from each of the nodes in the bundle and joining them together. This structure permits us to consider the possibility that a sub-attack may occur more than once. The design of the attack tree is based on a cooperative approach involving the group of auditors according to the following steps:

- a. identification of the attack components;
- b. construction of the bundles for each node;
- c. definition of the suites;
- d. qualification of the attributes.

Odubiyi and O'Brien (2006) created a framework to teach students how to develop countermeasures to the Twenty most Critical Internet Security Vulnerabilities (SANS Institute). The task was to build up an attack-tree forest taking into consideration these threads. The experiment simulated the methodology followed by companies, which try to "generalize from previously observed behaviour to recognize future behaviour, either malicious or normal."

Dimitriadis (2007) presents a case study suggesting how an attack tree should be developed based on the authentication mechanism of a major bank. In Fig.3.12 we can observe different paths to compromise a bank account.

The attack tree describes 3 main groups of attacks: UT/U (User Terminal/user) attacks, CC (Communication Channel) attacks, IBS (Internet Banking Server) attacks.

1) **UT/U attacks** - these attacks are aimed at the user equipment (smart-cards, password generators). An example is piggy-banking, which consists of installing a camera on the ATM in order to steal user's credentials.

2) **CC attacks** - these attacks target communication links. Examples of attacks are pharming (compromising domain name servers and connecting the user to a fraudulent website) and sniffing (capture information of the user).

3) **IBS attacks** - these are off-line attacks against the server that host the Internet banking application. Examples include brute force attack, bank security policy violation and web site manipulation.

Edge et al. (2007) define attack and protection trees and discuss how they can be implemented in order to perform a correct analysis in the on line banking security system. A protection tree "is constructed by identifying

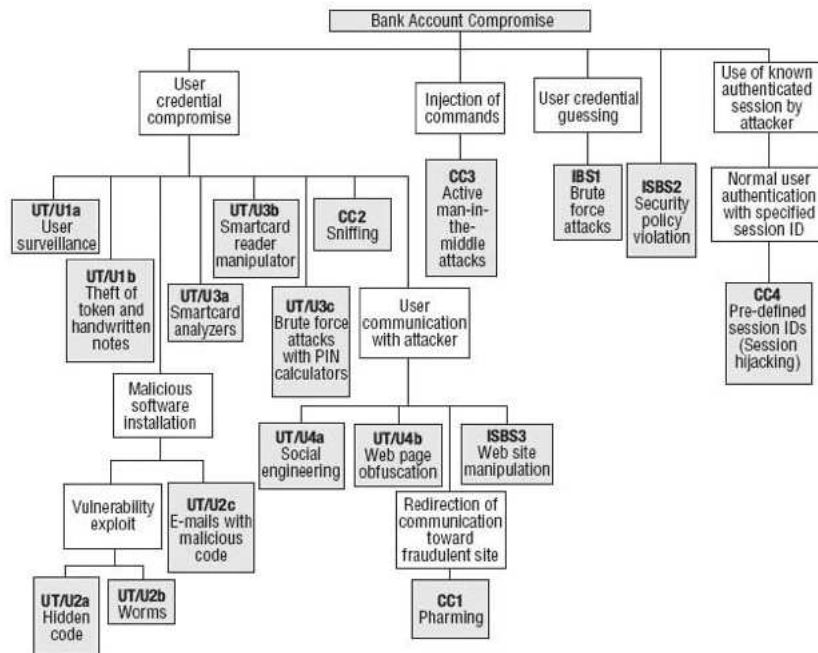


Figure 3.12: Attack Tree - Bank account compromise

protections that can mitigate an attacker’s actions in the leaf nodes of the attack tree.” It is complementary to the attack tree. Figure 3.13 shows an example of a protection tree to protect an electronic store.

Isograph (<http://www.isograph-software.com>), an attack tree software, provides the opportunity to calculate the consequences (Financial, Reputation, Safety, Political, Operational) of an attack (HIGHCOST, LOW-COST, SERIOUS, SLIGHT).

Moore et al. (2001) incorporate preconditions and postconditions in the model to improve the description of the attack tree. Preconditions include assumptions that auditors could make about the attacker and that are necessary for an attack to succeed, for instance the skills, resources, access, or knowledge that the attacker must possess, and the level of risk that he or she is willing to take. The postconditions are the consequences of the attack, for instance the knowledge gained by the attacker and changes to systems once the attack has succeeded. In this case material consequences are not as heavy as the non-material ones can be. Non-material consequences, like the loss of image or competitive advantage as a consequence of know-how misappropriation, can be hard to predict, but also to describe.

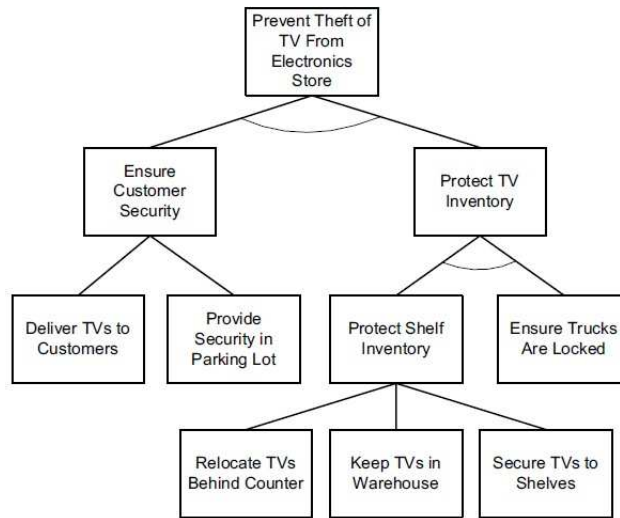


Figure 3.13: Protection tree for electronic store

3.4 FIDES

The first prototype of FIDES to support an audit team of a bank in fraud detection is shown in Figure 3.14. On the Data filtering module, we specify a set of $A_1 \dots A_k$ agents that can generate alarms.

Alarms can be produced by software agents (detection tools), suspicious behaviours indicated by whistle-blowers or the results of reports filled by inspectors of a Bank. These alarms must be evaluated by the team of experts $E_1 \dots E_M$, who interact performing the Delphi method in order to select those cases and/or behaviours that can be potentially dangerous.

The output of the Delphi process is an attack tree. In the Detection Module experts can decide which case to prosecute and then perform the risk estimation phase. Both tasks can be performed using the attack tree calculating for instance the least expensive, the most rational or the most irrational path.

Attack trees can be stored in a data base. Users in this phase can retrieve old attacks and relative paths to check the countermeasures used in the past and gain some inspiration for evaluating the new case.

Two possible kinds of scenarios where FIDES can be applied are illustrated in Fig.3.15.

In the first scenario, the left path of the figure, agents generate alarms when they discover anomalies in the data. In this case the Delphi method works in real time. A valid instrument to group cases and alarms by typologies, is the Ishikawa or fishbone diagram, which is depicted in the pruning box in Fig.3.15 and described in detail in figure 3.16; it has been adapted

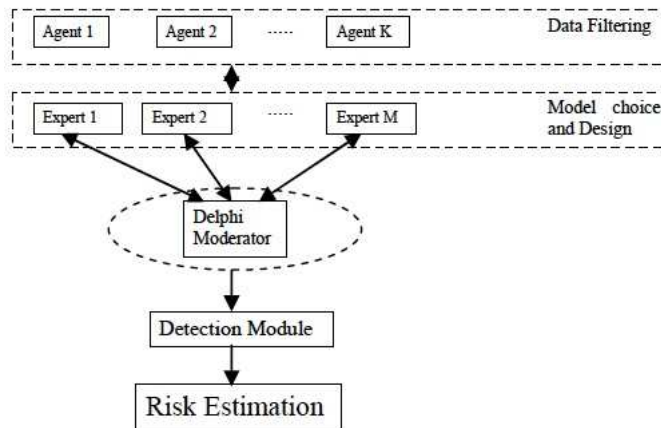


Figure 3.14: Fides in Buoni (2010)

from Balanced Scorecard Institute (1996).

The moderator has the duty to examine different alarms, group them and formulate the questions. Questions $Q_1...Q_n$ as shown in Fig.3.16 are positioned on the bones of the diagram under different typologies (common false alarms, usual behaviours, typical fraud schemes, unusual fraud schemes). Questions for instance can be formulated as "Is this one a common false alarm?". The output of the Ishikawa diagram is a list of fraud cases or behaviours after the pruning process. These good cases or behaviours are used to create the attack tree with a higher level of accuracy. Once the attack tree is built, auditors can decide whether or not to start an investigation and which path is more probable.

In the second scenario, represented in the right path of figure 3.15, the system is used in a proactive way. In this case, experts can meet to develop case scenarios to prevent future fraud cases on the basis of suspicions, rumours, information gathered, or confessions made by other people in anonymity. The output of the Delphi process is again an attack tree describing possible ways that fraudsters could choose to perform attacks on the basis of suspicions the experts have. The risk management phase consists in deciding whether to block or not certain operations or to investigate people who could be involved in illicit activities. The system works as a Group Decision Support System (GDSS) and it can be installed on mobile devices like mobile phones or tablet computers. Users can meet physically or virtually in order to respond immediately to alarms.

FIDES has been developed with the banking sector in mind, but it can be used in other situations that require a crisis management scenario (police, fire-fighters and medical personnel) and where there is the necessity to

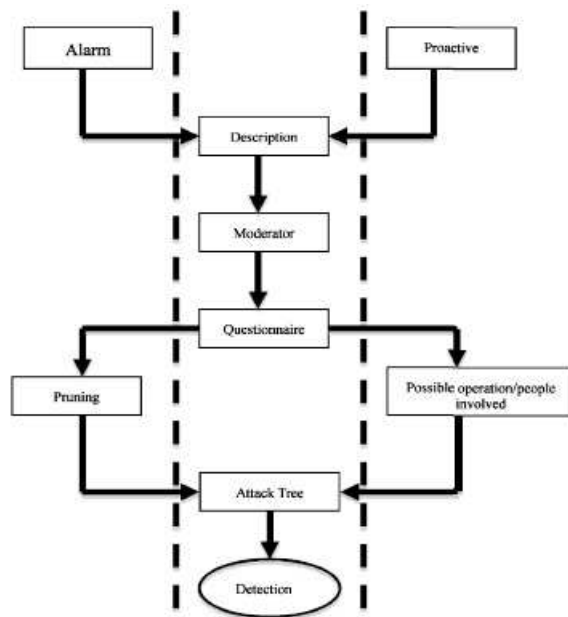


Figure 3.15: Two different scenarios of FIDES (Buoni, 2010)

coordinate complex tasks in emergency situations like natural disaster.

After the meeting with the employees of one of the most important European banks and the feedback received from them, the prototype of FIDES was improved according to their needs; in particular we worked on the interaction between inspectors and auditors which is a key feature in fraud detection. The blog platform that inspectors use to interact with each other, commenting on fraud cases, has been the starting point in improving FIDES. An inspection can literally stop the activity of the bank for a considerable period of time, causing extensive economic damage to the branch of the bank where the inspection takes place. An inspection, the creation of the report and the evaluation performed by the audit team are all time consuming activities.

A GDSS like FIDES can dramatically reduce all these procedures. The integration of the Think-map, the Delphi Method and the attack tree in FIDES is not only an instrument to improve fraud detection, but an instrument to collect fraud cases and the performed countermeasures which can be adopted in the future to solve new cases.

In Fig.3.17 the FIDES architecture and its components are introduced. The first module of FIDES is the Information Filtering Module. Inspectors check the alarms generated by the system or evaluate information and decide whether to report them or not to the audit team members. In case the

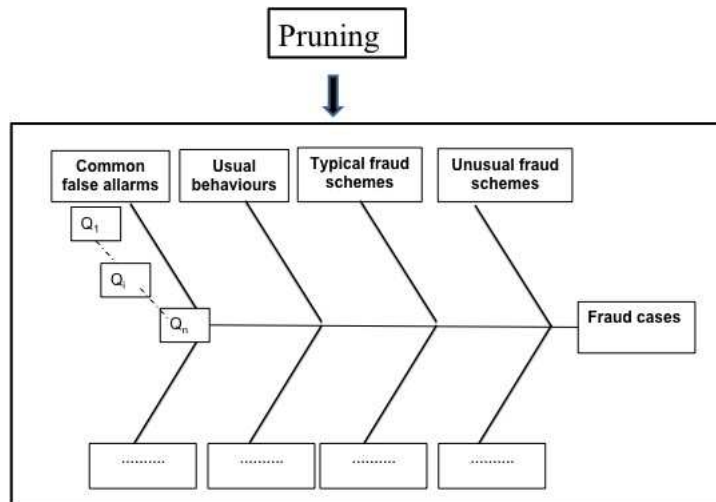


Figure 3.16: Ishikawa Diagram used for pruning

suspicious events are relevant they can decide to share them with colleagues using Web-Pad. The second module is the Attack Components Detection Module. Inspectors interact with each other by commenting on fraud cases, suspicions or the interview they just had with the director of the branch of the bank. As the result of this first phase, all this information is listed in the form of comments in a chat. As shown in Fig.3.18 messages are typed in real time in a "message box" and then they can remain in the "description box" to allow the other inspectors to read the new information. At this stage inspectors can start to underline keywords as shown in Fig.3.19 and to associate concepts to different labels in order to create a think-map.

In a think-map, the ICF structure is specified by linking Issue, Concept and Form together while in FIDES we determine an "action", "suspicious behaviour" and "suspected people", who might be responsible for the action on the basis of suspicious behaviour or alarms generated by the System. Inspectors can create new labels and link them. The result is the think-map shown in Fig.3.20.

The Think-map acts as a model to create nodes to be delivered to the auditors.

In the Delphi Module auditors connect nodes in order to create the attack tree.

The fuzzy mechanism

The construction of the attack-tree is performed through a fuzzy mechanism. The audit team performs the Delphi process focusing on the selection of nodes in order to form the attack tree. In the first phase the inspec-

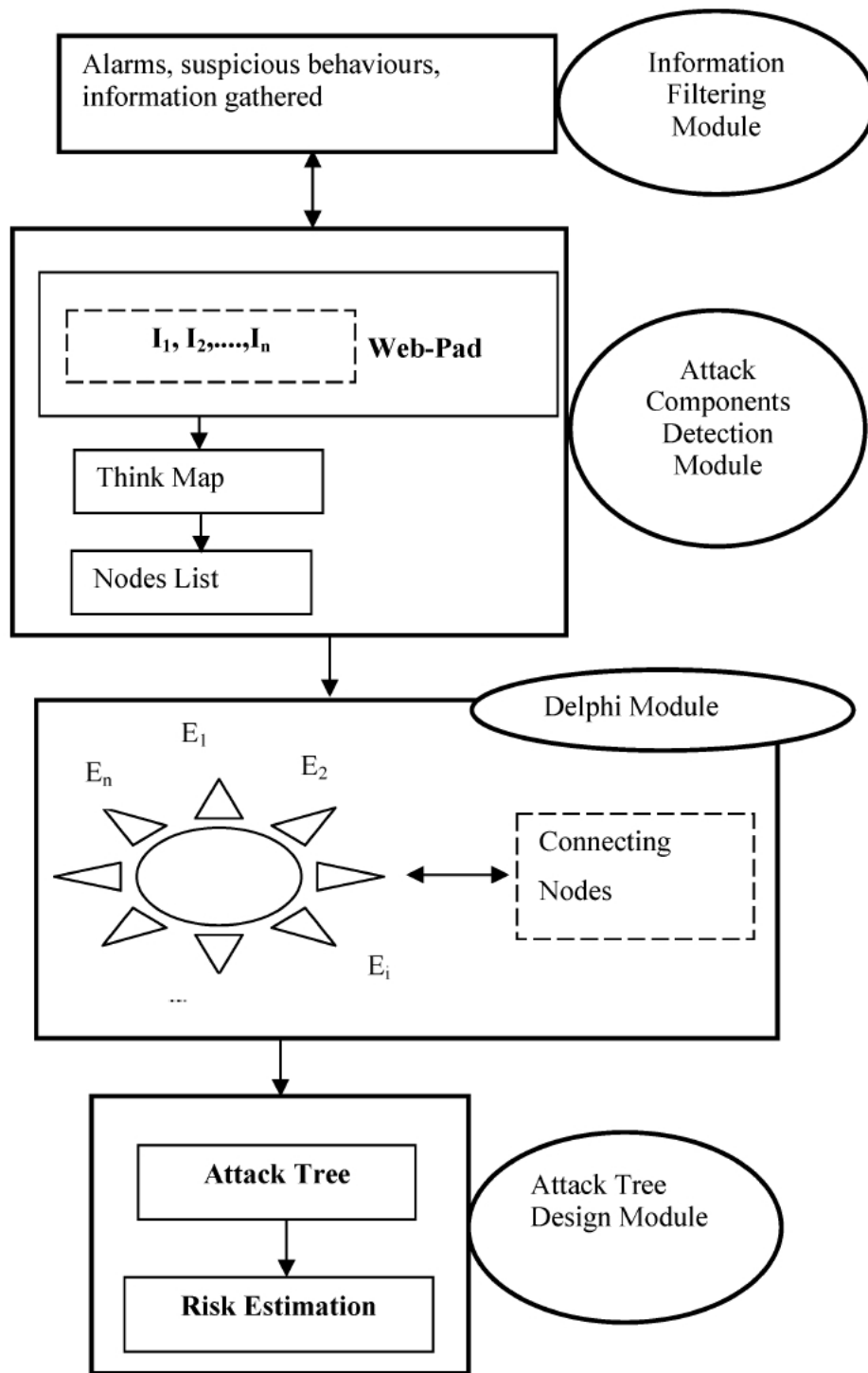


Figure 3.17: The architecture of FIDES

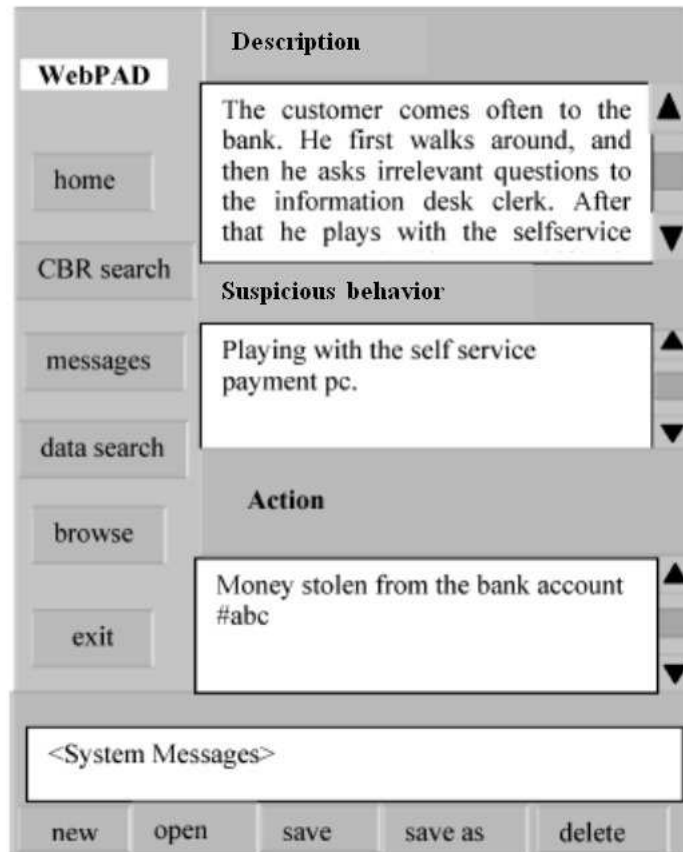


Figure 3.18: The Web-Pad interface

“The *customer* comes often to the bank. He first *walks around*, then he *asks irrelevant questions* to the information desk clerk. After that he *plays* with the *self-service payment pc.* checking around if there is someone observing him.”

Figure 3.19: The keywords selection

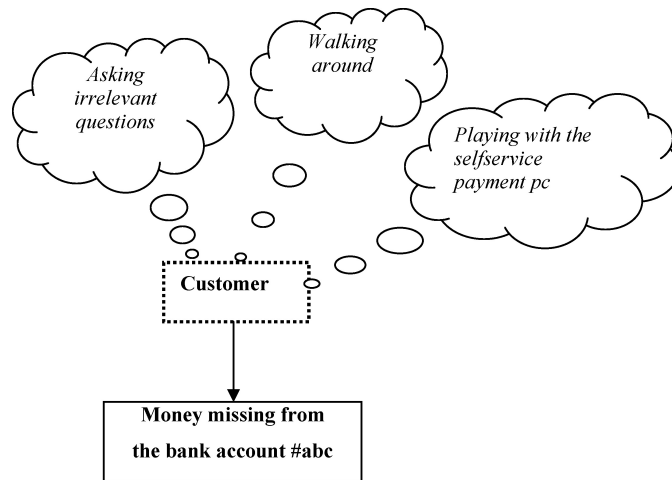


Figure 3.20: The Think-map

tors determine the possible nodes of the attack tree with the help of the think-map. The moderator, once the nodes have been delivered by the inspectors, asks the experts about the possible connection of the nodes, and aggregates the results to obtain the attack tree. The connection between nodes is determined by a fuzzy mechanism based on fuzzy sets theory which was introduced by Zadeh (1965) as an extension of the classical notion of set (Chakraborty, 2010).

In classical set theory the membership of the elements of a set is expressed in binary terms, which means that an element can belong or not to that set. A fuzzy set allows its members to have different degrees of membership, called membership function, in the interval $[0,1]$. In real world one can frequently encounter fuzziness: knowledge is imprecise, vague and uncertain. Human thinking and reasoning often involve fuzzy information. The use of systems based on classical set theory and two-valued logic is problematic in dealing with opinions expressed by experts.

A concept to deal with uncertainty is soft computing. The idea of soft computing was introduced by Zadeh (1994) as an example of a new kind of artificial intelligence to mimic the ability of humans to adopt a way of reasoning that is approximate rather than exact. Humans can tolerate and deal with imprecision and uncertainty in relation with language, sloppy handwriting, distorted speech and summarizing a text.

In hard computing the requirements are precision, certainty and rigour. In soft computing the main idea is that since precision and certainty carry a cost, it is necessary to exploit the trade-off between imprecision and uncertainty. Soft computing, according to Zadeh (1998), is not a single methodology, but rather a consortium of computing methodologies, including fuzzy

logic (FL), neurocomputing (NC), genetic computing (GC) and probabilistic computing (PC).

The contribution of FL to Soft Computing according to Zadeh is the fact that any theory can be fuzzified by replacing the concept of a crisp set with that of a fuzzy set. The advantage of using fuzzy logic is a better contact with reality, but with more computational cost compared to crisp numbers. Zadeh's main motivation for using FL is related to the concept of information granulation and how humans deal with fuzzy concepts. He stresses the fact that human concepts are fuzzy because they are the results of clumping of points drawn together by similarity. Examples of clump concepts are "middle-aged" or "partially cloudy". Linguistic values can be seen as a form of data compression. This form of data compression is termed as *granulation* (Zadeh, 1994). Linguistic labels and variables are described in the following way in Zadeh (1975) "By a linguistic variable we mean a variable whose values are words or sentences in a natural or artificial language. For example, "Age", is a linguistic variable if its values are linguistic rather than numerical i.e., young, not young, very young, quite young, old, not very old and not very young, etc., rather than 20,21,22,23."

The following description of the basic concepts of linguistic computing is based on the book by Carlsson et al. (2004) .

Definition 1. *A linguistic variable is characterized by a quintuple*

$$(x, T(x), U, G, M),$$

in which X is the name of the variable; $T(x)$ is the term set of x , that is, the set of names of linguistic values of x with each value being a fuzzy number defined on U ; G is a syntactic rule for generating the names of values of x ; and M is a semantic rule for associating with each value its meaning.

For instance, the term *speed* interpreted as linguistic variable can be the set $T(\text{speed})$, which is

$$T = \{\text{slow, moderate, fast, very slow, more or less fast, slightly slow, ...}\},$$

where each term in $T(\text{speed})$ is characterized by a fuzzy set in a universe of discourse $U = [0, 100]$.

We might interpret

- *slow* as "a speed below about 40mph"
- *moderate* as "a speed close to 55mph"
- *fast* as "a speed above about 70mph"

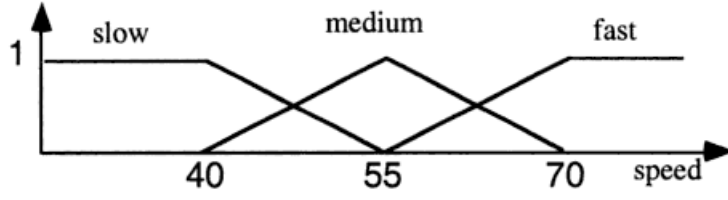


Figure 3.21: Values of linguistic variable *speed* (Carlsson et al., 2004).

These terms, represented in Fig.3.21 , can be characterized as fuzzy sets and their membership functions are

$$\begin{aligned}
 \text{slow}(v) &= \begin{cases} 1 & \text{if } v \leq 40 \\ 1 - (v - 40)/15 & \text{if } 40 \leq v \leq 55 \\ 0 & \text{otherwise} \end{cases} \\
 \text{moderate}(v) &= \begin{cases} 1 - |(v - 55)|/15 & \text{if } 40 \leq v \leq 70 \\ 0 & \text{otherwise} \end{cases} \\
 \text{fast}(v) &= \begin{cases} 1 & \text{if } v \geq 70 \\ 1 - (70 - v)/15 & \text{if } 55 \leq v \leq 70 \\ 0 & \text{otherwise} \end{cases}
 \end{aligned}$$

In our approach, the level of agreement between experts about the connection between a vertex and a leaf in the attack tree will be expressed by linguistic labels (Very Low, Low, Medium, High, Very High) as shown in Fig.3.23.

Before the description of the model we need some more basic definitions from fuzzy set theory. First, a formal definition of a fuzzy set is provided.

Definition 2. Let $X = x$ denote a collection of objects (points) denoted generically by x . Then a fuzzy set A in X is a set of ordered pairs

$$A = (x, \mu_A(x)), \quad x \in X \quad (3.1)$$

where $\mu_A(x)$ is termed the grade of membership of x in A , and $\mu_A : X \rightarrow M$ is a function from X to a space M called the membership space. When M contains only two points, 0 and 1, A is non fuzzy and its membership function becomes identical with the characteristic function of a crisp set. This means that crisp sets belong to fuzzy sets. A fuzzy number is a convex fuzzy set on the real line \mathbb{R} such that

1. $\exists x_0 \in \mathbb{R}, \mu_A(x_0) = 1$
2. μ_A is piecewise continuous

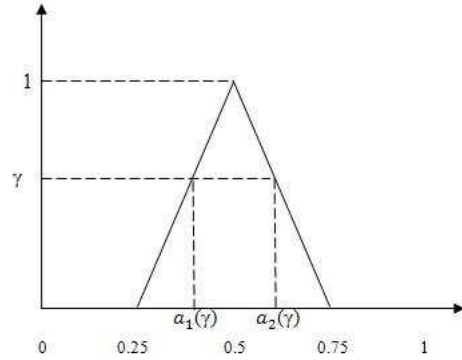


Figure 3.22: A fuzzy representation of the linguistic label "medium"

(The convexity means that all the γ -level sets are convex.) Furthermore, we call \mathcal{F} the family of all fuzzy numbers.

A γ -level set of a fuzzy set A in \mathbb{R}^m is defined by $[A]^\gamma = \{x \in \mathbb{R}^m : A(x) \geq \gamma\}$ if $\gamma > 0$ and $[A]^\gamma = \text{cl}\{x \in \mathbb{R}^m : A(x) > \gamma\}$ (the closure of the support of A) if $\gamma = 0$. Let A be a fuzzy number. Then $[A]^\gamma$ is a closed convex subset of \mathbb{R} for all $\gamma \in [0, 1]$. We use the notations

$$a_1(\gamma) = \min[A]^\gamma, \quad a_2(\gamma) = \max[A]^\gamma$$

for the left-hand side and right-hand side of the γ -cut, respectively.

Example 1. We calculate the value $a_1(\gamma)$ and $a_2(\gamma)$ of μ_{medium} represented in Fig.3.22

$$\mu_{\text{medium}}(x) = \begin{cases} 4x - 1 & 0.25 \leq x \leq 0.5 \\ -4x + 3 & 0.5 \leq x \leq 0.75 \end{cases}$$

then substituting γ in μ_{medium} we have

$$a_1(\gamma) = \frac{3 - \gamma}{4}$$

$$a_2(\gamma) = \frac{\gamma + 1}{4}$$

When we calculate the arithmetic operations on fuzzy sets (fuzzy numbers), we apply the rules of interval arithmetic. Let A and B be fuzzy numbers with the corresponding γ -cuts: $[A]^\gamma = [a_1(\gamma), a_2(\gamma)]$, $[B]^\gamma = [b_1(\gamma), b_2(\gamma)]$, then the γ -cut of the fuzzy number $A + B$ is the following:

$$[A + B]^\gamma = [a_1(\gamma) + b_1(\gamma), a_2(\gamma) + b_2(\gamma)],$$

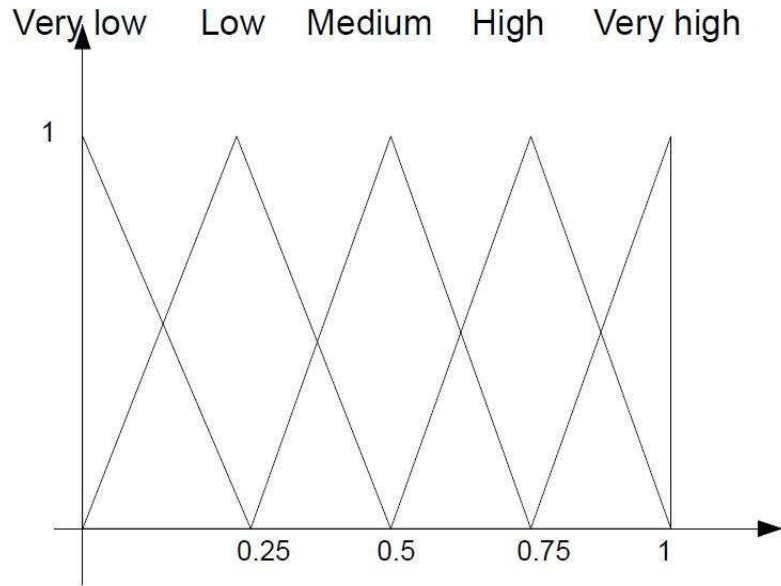


Figure 3.23: Possible representation with triangular fuzzy numbers

and the γ -cut of the fuzzy number αA , where $\alpha > 0$:

$$[\alpha A]^\gamma = [\alpha a_1(\gamma), \alpha a_2(\gamma)].$$

Example 2. When we calculate the sum of two triangular fuzzy numbers A and B the following result is obtained.

$$A = (a_1, b_1, c_1)$$

$$B = (a_2, b_2, c_2)$$

$$A + B = (a_1 + a_2, b_1 + b_2, c_1 + c_2)$$

Linguistic labels are used in the questionnaire and the labels can be presented as fuzzy numbers as shown in Fig.3.23.

We suppose that the moderator can choose which nodes are parents (V) (with descendant) and which ones are leaves (L) (without descendants, basic attack components). We obtain the following two sets

$$L = \{l_1, \dots, l_s\}, V = \{v_1, \dots, v_t\}.$$

In the first questionnaire the experts have to express their opinion in linguistic terms about statements like " $l_i \in L$ is required for $v_j \in V$ ", for every $i = 1, \dots, s, j = 1, \dots, t$.

Then experts $E = (e_1, \dots, e_N)$ are asked to determine their level of agreement on this statement based on a linguistic scale with m terms for every pair l_i, v_j . The linguistic terms in the model are represented as fuzzy numbers. In other words we have a mapping

$$\phi : T \rightarrow \mathcal{F} \quad (3.2)$$

from the set of linguistic terms into the family of fuzzy numbers.

Example 3. *One possible representation for a linguistic label is a triangular fuzzy number:*

$$A(u) = \begin{cases} 1 - \frac{a-u}{\alpha} & \text{if } a - \alpha \leq u \leq a \\ 1 - \frac{u-a}{\beta} & \text{if } a \leq u \leq a + \beta \\ 0 & \text{otherwise} \end{cases}$$

From the opinion of the experts we obtain the frequencies of the different classes. For the pair l_i, v_j we have $n_1^{ij}, \dots, n_m^{ij}$. If we denote by A_1, \dots, A_m the fuzzy numbers corresponding to the linguistic labels, we can define a new fuzzy number A_{ij} as a "weighted average", with level sets:

$$[A_{ij}]^\gamma = \left[\frac{1}{N} (n_1^{ij} a_1^1(\gamma) + \dots + (n_m^{ij} a_1^m(\gamma))), \frac{1}{N} (n_1^{ij} a_2^1(\gamma) + \dots + (n_m^{ij} a_2^m(\gamma))) \right],$$

where $[a_1^k, a_2^k]$ is the level set of A_k . This is clearly a fuzzy number with the support in the interval $[0, 1]$.

To obtain the connection degree for the pair l_i, v_j , we calculate the f -weighted possibilistic mean value of A_{ij} , defined in Carlsson and Fuller (2001).

Definition 3. *The f -weighted possibilistic mean value of $A \in \mathcal{F}$, with $[A]^\gamma = [a_1(\gamma), a_2(\gamma)]$, $\gamma \in [0, 1]$, is defined by ,*

$$E_f(A) = \int_0^1 M(U_\gamma) f(\gamma) d\gamma = \int_0^1 \frac{a_1(\gamma) + a_2(\gamma)}{2} f(\gamma) d\gamma, \quad (3.3)$$

where U_γ is a uniform probability distribution on $[A]^\gamma$ for all $\gamma \in [0, 1]$.

Example 4. *As an example one can calculate the f -weighted possibilistic mean value of μ_{medium}*

$$E_f(\mu_{\text{medium}}) = \int_0^1 \frac{\frac{1+\gamma}{4} + \frac{3-\gamma}{4}}{2} f(\gamma) d\gamma$$

$$f(\gamma) = 2\gamma$$

$$\int_0^1 \frac{1}{2} 2\gamma d\gamma = \int_0^1 \gamma d\gamma = \frac{1}{2}$$

After we have obtained these defuzzified numbers as the estimation of the connection strengths, we can determine for every attack component the ranking of the other nodes and then we can construct the adjacency matrix of the attack tree by connecting the leaves to the best ranked vertices.

Example 5. *In the simplest case we can represent the linguistic labels as fuzzy sets with the membership function:*

$$A(u) = \begin{cases} 1 & \text{if } u = c \\ 0 & \text{otherwise} \end{cases}$$

If we have 5 categories, we use the set $\{0, 0.25, 0.5, 0.75, 1\}$. The weights of the outcomes are the frequencies of the linguistic labels. If we observe the weights $n_0, n_{0.25}, n_{0.5}, n_{0.75}, n_1$, then A_{ij} is just the characteristic function of the value:

$$\sum_{i=0}^4 \frac{n_{0.25i}}{\sum_{j=0}^4 n_{0.25j}} 0.25i,$$

that is simply the sample mean value of our data. And according to the used defuzzification method, the obtained connection estimation is this sample mean.

The final result of the interaction between auditors is the attack tree, shown in Fig.3.24.

In the Attack tree design module, auditors can calculate which path of the attack tree is the most probable, the easiest and the least rational, and can view all the lines of reasoning that can help to understand whether to start or not the investigation. A useful instrument to retrieve past cases often used in information retrieval is cosine similarity. If we have two sets S_1 and S_2 then $M = |S_1 \cap S_2|$ is the number of common items between S_1 and S_2 . The cosine similarity between S_1 and S_2 can be defined as

$$\cos(S_1, S_2) = \frac{M}{\sqrt{|S_1||S_2|}}.$$

In our case S_1 is the attack tree shown in Fig.3.24, which we want to compare to another attack tree, stored in the database, S_2 shown in Fig.3.25. The similarity measure can be found using the method described in Nanopoulos and Manopoulos (2002).

The number of edges is 14 and 8 in S_1 and S_2 , respectively. Comparing the two attack trees, it can be seen that they have 5 edges in common: (transfer money to a personal account, log in), (transfer money to a personal account, hijacking bank system), (log in, guessing), (guessing, guess ID), (guessing, guess password). This means that $M = 5$ and the cosine similarity is:

$$\cos(S_1, S_2) = \frac{M}{\sqrt{|S_1||S_2|}} = \frac{5}{\sqrt{8 \times 14}} \approx 0.47$$

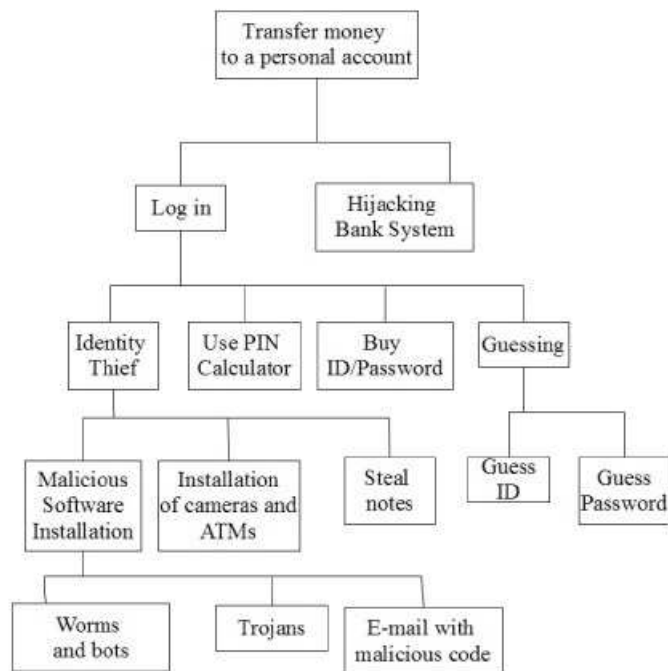


Figure 3.24: The final Attack tree

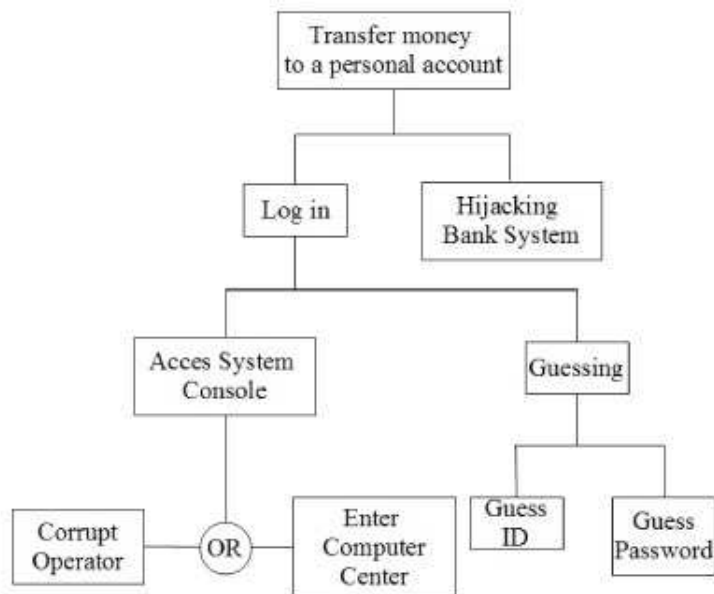


Figure 3.25: Another attack tree stored in the data-base

Consensual modelling of the attack tree

In Buoni and Fedrizzi (2012), an extension of FIDES is described. First, assuming that the opinions of experts involved in the design of the attack tree are represented by fuzzy preference relations, we introduce a dynamical consensus model aiming at finding a shared representation of the attack tree. Second, assuming that the leaf nodes of the attack tree are attributes with fuzzy number values and that the attributes are interdependent, we show how to propagate the values up the tree through an aggregation process using Choquet integral. First the consensual modelling and then the aggregation process will be described in detail.

The consensual modelling of the attack tree allows the experts to build the attack tree according to their opinions. In the attack tree design the moderator can choose which nodes are parents (V) and which ones are leaves (L), obtaining two sets

$$L = \{l_1, \dots, l_s\}, V = \{v_1, \dots, v_t\}$$

The individual preferences are represented as fuzzy preferences

$$P = L \times V.$$

If $P = \{p_1, \dots, p_m\}$ is a set of alternative connections and $E\{e_1, \dots, e_N\}$ is the set of experts, then the fuzzy preference relation of expert e_1, R_i is given by its membership function $\mu_i : P \times P \rightarrow [0, 1]$ such that

$$\begin{aligned} \mu_i(p_k, p_l) &= 1 \text{ if } p_k \text{ is definitely preferred over } p_l \\ &\in (0.5, 1), \text{ if } p_k \text{ is preferred over } p_l, \\ &= 0.5 \text{ if there is indifference between } p_k \text{ and } p_l \\ &\in (0, 0.5), \text{ if } p_l \text{ is preferred over } p_k \\ &= 0, \text{ if } p_l \text{ is definitely preferred over } p_k \\ &\text{where } i = 1 \dots, N \text{ and } k, l = 1 \dots, M. \end{aligned}$$

Each individual fuzzy preference relation R_i can be represented by a matrix

$$[r_{kl}^i], r_{kl}^i = \mu_i(a_k, a_l).$$

Each decision maker $i = 1, \dots, n$ is represented by a pair of connected nodes, a primary node (dynamic) and a secondary node (static).

The n primary nodes form a fully connected sub network and each of them encodes the individual opinion of a single decision maker, denoted by r_i .

The n secondary nodes encode the individual opinions originally declared by the decision makers, denoted by s_i , and each of them is connected only with a single primary node.

Here, for the sake of simplicity, we assume that there are only two alternatives available ($m = 2$), which means that each individual preference

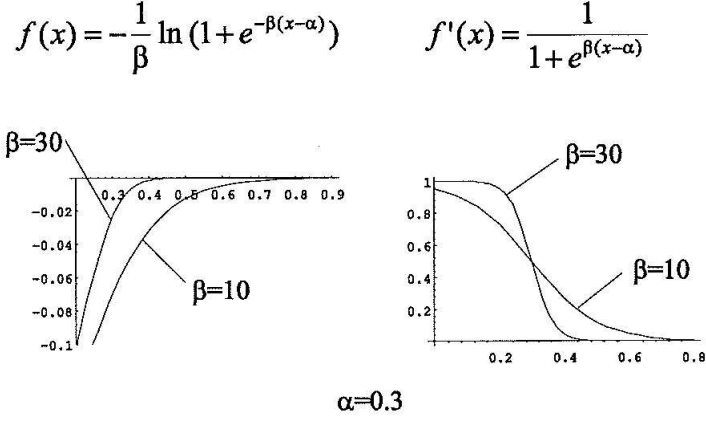


Figure 3.26: Scaling function f and sigmoid function f' .

relation R_i has only one degree of freedom, denoted by $x_i = r_{12}^i$. Accordingly the preference relation declared by expert e_i will be denoted s_i .

The iterative process of opinion transformation corresponds to the gradient dynamics of a cost function W , depending on both the present and the original network configurations, as introduced in (Fedrizzi et al., 1999) and (Fedrizzi et al., 2007).

The value of W combines a measure V of the overall dissensus in the present network configuration and a measure U of the overall change from the original network configuration.

The diffusive interaction between primary nodes i and j is mediated by the interaction coefficient $v_{ij} \in (0, 1)$, whereas the inertial interaction between primary node i and the associated node j is mediated by the interaction coefficient $u_i \in (0, 1)$,

$$v_{ij} = f'((x_i - x_j)^2) \text{ and } u_i = f'((x_i - s_j)^2)$$

The values of the interaction coefficients are given by the following derivative of scaling function (see Fig.3.26).

The value v_i as well as the weighting coefficient $v_j \in (0, 1)$ depend non-linearly on the standard Euclidean distance between the opinion x_i and x_j

$$v_i = \frac{\sum_{j \neq i} v_{ij}}{n - 1}$$

$$\bar{x}_i = \frac{\sum_{j \neq i} v_{ij} x_j}{\sum_{j \neq i} v_{ij}}$$

The individual disagreement cost $V(i)$ is given by

$$v_i = \frac{\sum_{j \neq i} v(i, j)}{n - 1}$$

where $V(ij) = f((x_i - x_j)^2)$ and the individual opinion changing cost is $U(i) = f((x_i - s_j)^2)$.

Summing over the various experts we obtain the collective disagreement cost $V = \frac{1}{4} \sum_i V(i)$ and the inertial cost $U = \frac{1}{2} \sum_i U(i)$, where 1/4 and 1/2 are conventional multiplicative factors.

Then the full cost function is $W = (1 - \lambda)V + \lambda U$ with $0 \leq \lambda \leq 1$

The consensual network dynamic acts on the individual preference x_i through the iterative process $x_i \rightarrow x' = x_i - \epsilon \frac{\partial W}{\partial x_i}$ based on the gradient term $\frac{\partial V}{\partial x_i} = v_i(x_i - \bar{x}_i)$.

We can analyse the dynamical effect of the components V and U separately. The dissensus cost V induces a non linear process of diffusion based on the gradient term

$$\frac{\partial V}{\partial x_i} = (x_i - \bar{x}_i)$$

As a result, the iterative step of the non-linear diffusion mechanism corresponds to a convex combination (with sufficiently small ϵ) between the opinion value x_i and the weighted average \bar{x}_i of the remaining preference values x_j ,

$$x'_i = (1 - \epsilon v_i)x_i + \epsilon v_i \bar{x}_i$$

The inertial cost leads to a non-linear mechanism which opposes changes from the original opinion value x_i , by means of the gradient term

$$\frac{\partial U}{\partial x_i} = u_i(x_i - s_i)$$

The full dynamics associated with function $W=(V+D)/2$ acts iteratively on each decision maker i through convex combinations of the opinion value x_i , the average opinion value \bar{x}_i , and the original opinion value s_i , in the following way

$$x'_i = (1 - \epsilon(v_i + u_i))x_i + \epsilon v_i \bar{x}_i + \epsilon u_i s_i$$

At this point the expert e_i is in dynamical equilibrium, in the sense that $x' = x_i$, if the following equation holds

$$x_i = (v_i \bar{x}_i + u_i s_i)/(v_i + u_i)$$

Choquet-based evaluation

In many applications of attack trees the main problem is to promulgate information contained in leaves to other nodes, up the tree until it reaches the root node. This process is carried out by aggregation operations occurring at the "and/or" nodes. There are different systems of aggregation that can be suitable for different cases. An extensive overview of these methods can be seen in Beliakov et al. (2007) and Grabish et al. (2009).

The most used one is the weighted average. The limitation of this method is that it is a compensative method, which means that it is not possible to appreciate the interactions among the attributes. The Ordered Weighted Aggregation (OWA) method, introduced by Yager (1988) depends on the order positions of the child nodes, but it does not consider the possible interaction between the nodes. One way of solving this limitation of OWA operators is to introduce the Choquet integral (Choquet, 1953), which can take into account the interaction between nodes, ranging from redundancy (negative interaction) to synergy (positive interaction). Choquet integral is also mathematically well founded (Klement et al., 2010) and applied in many problems in multi-criteria and multi-attribute decision models (Grabish, 1996; Grabish and Labreuche, 2010).

The main scope we want to achieve by introducing the Choquet integral is to combine the inputs in such a way that not only the importance of individual inputs, but also the importance of the coalition is considered. This means that an input, not relevant by itself, can become significant when merged with some other inputs. It is easy to show how weighted arithmetic and OWA operators are particular cases of Choquet integral with respect to additive and symmetric capacities respectively.

Consider a finite set of elements $N = \{1, 2, 3, \dots, n\}$. A (discrete) fuzzy measure μ (also called capacity) defined on N is a set function $\mu : 2^N \rightarrow [0, 1]$ satisfying:

- 1) $\mu(\emptyset) = 0, \mu(N) = 1$ (border conditions),
- 2) $S \subseteq T \implies \mu(S) \leq \mu(T), \forall S, T \subseteq N$ (monotonocity condition)

Given two coalitions $S, T \subseteq N$, with $S \cap T = \emptyset$, the fuzzy measure is said to be additive if $\mu(S \cup T) = \mu(S) + \mu(T)$, sub-additive if $\mu(S \cup T) < \mu(S) + \mu(T)$ and super-additive if $\mu(S \cup T) > \mu(S) + \mu(T)$ with respect to the two coalitions S, T .

Definition 4. Let μ be a fuzzy measure on N . The discrete Choquet integral of a function $f : N \rightarrow [0, 1]$ with respect to μ is defined by

$$\int f d\mu = C_\mu(f(1), \dots, f(n)) = \sum_{i=1}^n [f((i)) - f((i-1))] \mu(A_{(i)})$$

where (i) indicates a permutation on N so that $f((1)) \leq f((2)) \leq \dots \leq f((n))$ and $A_{(i)} = \{(i), \dots, (n)\}$. Also $f((0))=0$.

From now on we call $f(i) = x_i$ so the Choquet integral of the vector $(x_1, x_2, \dots, x_n) \in [0, 1]^n$ with respect to the fuzzy measure μ is the following

$$C_\mu(x_1, \dots, x_n) = \sum_{i=1}^n (x_{(i)} - x_{(i-1)})\mu(A_{(i)})$$

where $x_i \in (x_1, \dots, x_n)$, $x_1 \leq x_2, \dots, \leq x_n$, $A_i = \{(i), \dots, (n)\}$, and $x_0 = 0$.

The Choquet integral then is nothing else other than a linear combination of the marginal gains (differences) between the ordered criteria. An alternative way of writing the Choquet integral is

$$C_\mu(x_1, \dots, x_n) = \sum_{i=1}^n x_{(i)}(\mu(A_{(i)}) - \mu(A_{(i+1)}))$$

where $A_{(n+1)} = \emptyset$.

The Choquet integral satisfies different properties described in Grabish (1996), Marichal (1988) and Ghirardato (2000).

In this section, we will need the following essential properties (for all $x, x' \in [0, 1]^n$)

$$x_i \leq x'_i \implies C_\mu(x_1, x_2, \dots, x_n) \leq C_\mu(x'_1, x'_2, \dots, x'_n) \text{ (Monotonicity)}$$

$$x_i \leq x'_i \implies C_\mu(x_1, x_2, \dots, x_n) < C_\mu(x'_1, x'_2, \dots, x'_n) \text{ (Strict Monotonicity)}$$

$$\min(x_1, x_2, \dots, x_n) \leq C_\mu(x_1, x_2, \dots, x_n) \leq \max(x_1, x_2, \dots, x_n)$$

Choquet integral is continuous and if the fuzzy measure is additive, it coincides with the weighted average, and if every subset with the same cardinality has the same measure, it collapses into the OWA operator (Fodor et al., 1995; Grabish, 1995a).

In the Choquet-based evaluation process the problem is to promulgate the information up the tree over the nodes, taking into consideration the interaction between them, which can be redundant (negative interaction) or synergetic (positive interaction). The estimation of the attribute value is performed adopting the possibility measure, representing the numeric imprecision of attributes' values using unimodal LR fuzzy numbers, which are defined by

$$A(x) = \begin{cases} L\left(\frac{a-x}{a_1}\right) & a - a_1 \leq x \leq a, \\ R\left(\frac{x-a}{a_2}\right) & a \leq x \leq a + a_2 \\ 0 & \text{else,} \end{cases}$$

where $a \in \mathbb{R}$ is the peak of A , $\alpha = a - a_1 \leq 0$ and $\beta = a_2 - a \leq 0$ are the left and right spread, respectively, and $L, R : [0, 1] \rightarrow [0, 1]$ are two strictly continuous shape functions such that $L(0) = R(0) = 1$ and $L(1) = R(1) = 0$.

By extending the Choquet integral to a fuzzy domain several forms of information can be handled at the same time like crisp data, interval values, linguistic variables (Yang et al., 2005). The Choquet integral then is defined for a measurable interval valued function (Aumann, 1965), and then it is extended to fuzzy integrand using alpha-cuts (Grabish, 1995b). We introduce the following notations:

- I is the set of interval numbers (rectangular fuzzy numbers)
- $N = \{1, 2, \dots, n\}$ is a set of the elements
- $F : N \rightarrow I$ is an interval-valued function
- $F_L(i)$ and $F_R(i)$ are the left end point and the right end point of the interval $F(x)$, respectively
- \mathcal{F} is the set of all unimodal LR-type fuzzy numbers
- $[^L A^\alpha, ^R A^\alpha]$ is the alpha cut of the fuzzy number A
- $\Phi : N \rightarrow F$ is a unimodal LR fuzzy valued function
- \mathcal{F} , is the set of all the attack trees with unimodal LR fuzzy numbers as leaves values

Given a measurable fuzzy-valued function $\Phi(i)$ on N and a fuzzy measure μ on 2^N , the Choquet integral of $\Phi(i)$ with respect to μ is defined as

$$\int \Phi d\mu = \bigcup_{0 \leq \alpha \leq 1} \alpha \int \Phi^\alpha d\mu$$

Given a measurable interval-valued function $\Phi(i)^\alpha$ and the fuzzy measure μ on 2^N , the continuous Choquet integral of $\Phi(i)^\alpha$ with respect to μ is

$$\int \Phi^\alpha d\mu = \left[\int \Phi_L^\alpha d\mu, \int \Phi_R^\alpha d\mu \right]$$

which becomes

$$\int \Phi d\mu \bigcup_{0 \leq \alpha \leq 1} \alpha = \left[\int \Phi_L^\alpha d\mu, \int \Phi_R^\alpha d\mu \right]$$

Consider now a tree in \mathcal{F} whose leaves' values are unimodal LR fuzzy numbers. Based on the fact that the Choquet integral of unimodal LR fuzzy numbers is still an unimodal LR fuzzy number as it was proved in Bortot et al (2011), the algorithm proceeds as described below:

1. The alpha cut of each unimodal LR fuzzy number in the leaves will be considered using a suitable grid.
2. The procedure receives the extremes of the alpha-cut, and computes the aggregated value for both the lower and the upper bounds. Increasing the values of alpha in between $[0,1]$, the two computed values, form an interval included in the previous ones (for lower value of alpha).
3. Thus the obtained intervals form the alpha-cuts of the fuzzy root, i.e. the required solution.

In this chapter FIDES and its components have been introduced and described in detail. The main contribution of FIDES is the novelty of a system that combines Think-map, Delphi method and attack-tree. Based on the suggestions and feed-back obtained from the audit-team of the bank, the system mirrors real world anti-fraud procedures performed by banking experts.

FIDES improves the interaction between auditors and inspectors, shortening the time between the inspection and the final evaluation. The building of the attack tree is mathematically well founded and rigorously defined step by step.

Delphi method, used to create the attack tree, combines formal rigour and respect for human reasoning and the decision-making process. Usually hypotheses, introduced in decision-making processes, are based on the full rationality of the agents and perfect information. In FIDES uncertainty and lack of information, because of the intrinsic nature of fraud and anti-fraud procedures, are the natural conditions for the auditors to work in. The power of the system is to turn uncertainty into an opportunity to expand the set of different/possible interpretations of fraud paths. Also in the risk management phase, when the attack tree is created, the decision to choose a certain path can follow non-rational reasoning. The least-probable path can be chosen or the most expensive one, according to the intuition of the experts and the circumstances of the moment.

On the other hand, if the users are free to create unconventional decision paths, the procedure they use instead to achieve a shared solution is very well structured and methodical.

FIDES in other words combines the creativity of the decision-making process with the rationality of the procedure used in creating knowledge. FIDES is not only a DSS for fraud detection, but an instrument for creating knowledge that can be re-used in the future. In order for information to be re-used, it must be processed and stored in a well structured and semantically defined way.

In this sense a DSS like FIDES can be seen as an intermediary step in the path towards creating an autonomous knowledge-based system in the future.

3.5 SWOT analysis of FIDES

The SWOT analysis is one of the most popular tools to analyse strengths and weaknesses of a system and to offer valid suggestions to improve it. In this section a SWOT analysis will be carried out for FIDES.

Strengths

A strong success element of FIDES is that it has been built around the needs auditors have collected and expressed; FIDES was developed according to their suggestions and based on a continuous process of verification moving between the literature and the bank world. FIDES can be used as an anti-fraud system, but also in a proactive way to prevent future potential fraud; once suspect behaviors are reported to the audit team similar schemes can be detected. A third strength of FIDES is flexibility, since it can be used in all the contexts where an interaction between external inspectors and a central risk management unit plays a fundamental role.

Weaknesses (or Limitations)

There are some limitations of FIDES due to its complexity. One difficulty will be in its implementation as different tools with specialized features are problematic to integrate with each other. Another limitation is the validation of the system. Since the frequency of internal fraud is very low and sometimes up to 20 years may be needed to collect a few dozen relevant cases, a validation process in the traditional positivistic sense is not possible.

Opportunities

A system such as FIDES would permit a bank to shorten the inspection and reporting system time. Saving time is a key factor in fraud detection, which allows auditors to develop a timely counter-strategy to block potential fraudulent activities. There are substantial possibilities for improvement in

the attack tree building procedure in FIDES. This can greatly improve the detection process and consequently save time and money; thus FIDES has a great future potential. Auditors could use it as a learning tool to study old cases stored in the system and create an automatic detection system to help their work.

Threats

The decision to adopt the system by banks (or other organisations) could be an issue for the auditors since users are nowadays unwilling to use many different tools to complete a single (even complex) task.

Chapter 4

Fraud detection processes

In this chapter, a distinction is made between ICT and human-based fraud. ICT-based fraud are based on hacking attacks and characterized normally by high frequency and low damage.

Human-based fraud have low frequency with high damage and involve the intervention of human actors (for instance, in faking documents manually) and the main cause of their success is conflict of interests. Human based fraud are perpetrated not necessarily because of greed, but for instance to improve the reputation of the perpetrator (a bank director wants to show better performance to obtain a promotion). ICT and human based fraud can be combined, for instance when a hacker uses the support of a bank employee to obtain information to finalize his/her fraudulent scheme.

ICT based and human based fraud will be described in paragraph 4.2 and 4.3, respectively. In paragraph 4.3, the role of this distinction is explained.

4.1 ICT based

ICT-based fraud deals with attacks perpetrated by hackers in order to obtain information and/or to steal money. This phenomenon is growing thanks to the development of e-commerce, on-line banking and credit card usage. These methods use malware, Trojans or other schemes, aiming at cheating the users and/or stealing their identities. In the following paragraphs a few of these schemes will be discussed, focusing on e-commerce; the examples are taken from the Teach-Ict website¹.

In **online auctions**, once the buyer wins an auction, the amount of the bid is paid, but sometimes the items are not delivered. This can also happen in the case of an on-line store purchase. Another form of fraud is the offer to **work at home and make a fortune**. Initially, customers of this fake service are asked to pay a fee in order to receive a kit with all the

¹<http://www.teach-ict.com/gcse/theory/fraud/miniweb/index.htm>

information to start up the activity. When the payment is made, the kit they receive consists of only advertisements material such as pamphlets.

Other forms of fraud involve the **promise of a loan** after the payment of a fee.

Phishing is a fraud technique used by criminals to steal information such as credit card and social security numbers, user IDs and passwords to gain access to bank accounts. Criminals can set up a website that appears to be a legitimate one, for instance of a bank or an insurance company and capture user's attention by sending an email with a link that apparently points to the real website, but instead leads to the fake one. They ask the victim to send sensitive information such as credit card numbers and security access codes which can be used to withdraw money from the victim.

Key logging software is another fraud technique. Fraudsters send an attachment in a form of a photograph, but once it is opened, software is installed on the computer in order to register all the keys that have been typed. Finally, a record with all the keystrokes is transferred to the hacker's computer. After receiving the data, the hacker starts to use the password to log in to a website previously visited by the victim (for example, on-line bank account).

Identity theft is the most common way to create fraud. Many of the fraud schemes do not require the use of brute force to break into security systems but aim to steal sensitive information from the victims employing different "creative" methods.

In UK, the estimated number of identity fraud victims is more than 100,000, at a cost of £1.3 billion annually.

The most common type of identity fraud is credit card fraud. ²

The simplest form of identity fraud indeed can take place at a restaurant, when people forget the original copy of their credit card receipts, which contains the credit card number. These copies can be used to make purchases on the net. Personal information can be acquired in many ways: memorized by clerks and waiters, removed from different bills, stolen from employee files or a hospital record. Employers can be bribed to provide personal information of their employees and clerks can put skimmers on credit card machines in order to record personal data.

It is important to observe that in the ICT fraud, involving electronic payments, there is also a strong human component. Identity theft is the main goal of fraudsters' activity in order to use the information to log into the banking systems.

This aspect contradicts the traditional rhetoric of investments in security, which states that the improvement in technology in terms of computational power is the main focus. Investment in the latest security systems is some-

²<http://money.howstuffworks.com/identity-theft1.htm>

thing that cannot be avoided, due to the fact that hackers can develop newer and more sophisticated methods, but in order for institutions and private companies to avoid wastage of resources the decision to invest in sophisticated ICT countermeasures has to deal with the discussed trivial methods that are largely applied in real world, otherwise the organisation may run the risk of investing in the perception of security instead of security itself. In the next paragraph human based fraud and its relation to ICT-based ones will be discussed and investigated from the perspective of the architecture of FIDES.

4.2 Human factors based fraud

In the previous paragraph, the strong human component in ICT-based fraud have been demonstrated, since the easiest way to gain access to a payment system is to acquire username and password by using different techniques.

The main difference between human and ICT-based fraud is the key-factor: in human-based fraud it is conflict of interest. In the case of a customer fraud in a bank, a director of a branch who has friendly relationships with his/her customers is a likely scenario. Customers, who have limited experience in finance or place too much trust in the director or the finance expert of the bank may decide to instruct the "expert" to invest the money on the basis of the "do what you think is right" principle. This excess of trust, of course, implies that customers do not pay close attention to transactions in their account, leaving the manager free to perpetrate his fraud scheme. Another crucial difference is the scope of the fraud: in the case of ICT-based fraud, the focus is personal enrichment or secret information as industrial espionage, but for human based fraud it can be reputation, the possibility of obtaining a promotion, or as the banality hypothesis suggests just a short-cut to be competitive in the market. This can be the case also in tax fraud, where specialised accountants or inside lawyers are handsomely paid to find all the possible ways to, in the best case scenario, elude taxes or, in the worst one, evade them. The real danger related to human based fraud is the involvement of employees in criminal organisations. The reasons for this involvement can be various: to obtain loan because of their problematic situation; they are paid by criminals to act as informers and Trojan horses inside that institution (bank, public office); they operate under constant threat from the criminal organisations. This is the main reason why more focus should be on human based fraud, not only for the economic impact, but for detecting members of big criminal organisations and their illegal business. These fraud schemes are often just the tip of the iceberg of more complex activities.

In FIDES, electronic systems that are able to generate alarms in case of

intrusion, can be used as sentinels in the Information Filtering Module. All these alarms, which can be associated with suspicious behaviours reported by whistle-blowers, have to be evaluated by inspectors, who have to decide whether or not to start the evaluation process. Already in the filtering phase, inspectors, being aware of the links between electronic attacks and the human support that is needed to finalize them (faking documents, identity theft), can use their experience to find connections between the digital and behavioural side of fraud.

In this sense FIDES endorses the idea of technology as a support for human decisions and not an instrument of truth and total control managed by an autonomous system.

Chapter 5

Fraud in other contexts

In this chapter fraud in other contexts will be introduced such as in the insurance and the public sector. The first paragraph will be dedicated to the Insurance industry. The second one will focus on fraud in the European Union and public sector. In the third paragraph, money laundering is highlighted. Even though money laundering is not the main topic of this thesis, it deserves a general view because of the strong relationship to fraud and its importance in the banking sector. Finally, we will explore the potential of FIDES in these different contexts.

5.1 Insurance industry

The Insurance industry is one of most critical sectors affected by fraud. The Insurance industry includes different areas; California Department of Insurance (2012) provides a list of the most common insurance areas and typologies of fraud.

- **Automobile Property.** In car insurance there are several ways of committing fraud. There is the possibility of faking damages in order to obtain a bigger refund or inflating the damage by making a deal on the bill with the body repairer. Another common technique is vehicle arson (setting fire usually on an old vehicle in order to obtain from the insurance a higher sum of money than one would get if the car was sold in the market (Insurance fraud hotline, 2012)). This type of fraud also poses a danger to the surrounding properties and people.
- **Medical.** It is hard to estimate the level of Healthcare fraud in USA. Healthcare fraud expert and Harvard University Professor Malcolm K.Sparrow estimates the loss as hundreds of billions of dollars per year (Aldrich, 2010). The most common healthcare fraud schemes include: inflated billing by any medical facility, doctor, dentist, chiropractor,

laboratory, etc; falsified billing (pharmacist can falsify billing in order to distribute drugs without medical prescription).

- **Life.** In this case the fraud involves identity fraud schemes, since the aim is to give the false identity of a dead person or fake information in the death report.
- **Workers' Compensation.** This category includes situations where employers commit illegal acts against employees (such as having uninsured employees). Misclassification of the worker typology in order to obtain higher compensation from the insurance company is a typical example. Another example is to classify a roofer as a clerk; the same goal can be achieved by misrepresenting the payroll showing an under-reported wage or injuries history.
- **Fire.** This type of fraud includes all those attempts to burn down commercial activities, private houses, in order to receive compensation. Another way is to exaggerate the loss in the claim form.
- **Property.** Property fraud takes place when damage to a property is fabricated in order to claim compensation. It can be done on private houses or offices in order to obtain a refund. Fabricating theft, claiming that the object was inside a property (car, house) is another scheme adopted to obtain a refund from the insurance company.
- **Healthcare.** An example of healthcare fraud is identity theft when identity is stolen in order to secure healthcare benefits. There are cases where patients are recruited to request medical procedures without necessarily using these services afterwards. Fraud can be perpetrated also by pharmacies when they inflate or falsify bills or codes. Disability is another area where fraud can be committed: patients continue to receive benefits even after the temporary disability is over.

In insurance fraud, India represents an interesting case. It is one of fastest growing economies amongst the BRIC countries and as a consequence, insurance fraud is on the rise as well.

In the Ernest and Young report (Ernst & Young, 2011) authors point out how the results and patterns are very similar to the ones observed in the UK, in particular in the area of claims.

The survey, as shown in Figure 5.1 is based on 51 respondents including CXO (Chief Executive Officers 29%) , agents (23%), internal audit & head compliance officers (22%), MD/Director (16%) and middle management (10%). 40% of the respondents declared that in the institution where they work they do not have an anti-fraud department. 40% of the respondents agreed that there has been a rising of insurance fraud during the last



Figure 5.1: Profile of respondents (Ernst & Young, 2011)

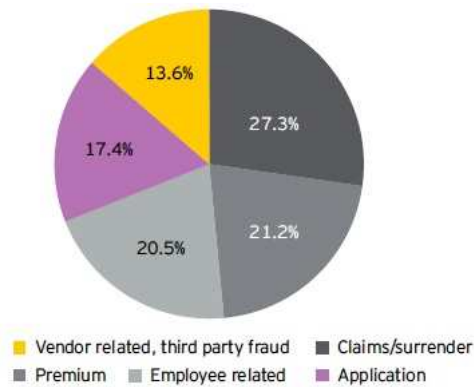


Figure 5.2: Fraud risk exposure faced by insurance company (Ernst & Young, 2011)

year. The survey identifies three broad fraud categories: policy holder and claims fraud, intermediary fraud and internal fraud.

The first category includes fraud against the insurer by a policyholder and/or other parties in the purchase and/or execution of an insurance product. The second category includes the fraud perpetrated by intermediaries against insurers and/or policyholders. The third category is the one where employees in collusion with other employees commit fraud against the insurance company. The five risk areas of Insurance companies are shown in Fig. 5.2.

The three major risk areas are the following ones: claims/surrender (27%), Premium (21.2%) and Employee related (20.5 %). An example of claim fraud is travel abroad for surgery without disclosing it; a premium

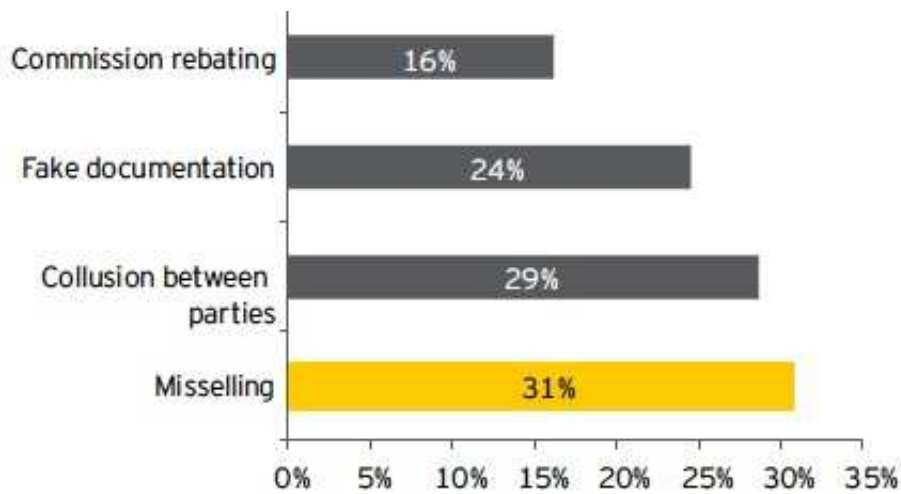


Figure 5.3: Different types of fraud affecting insurance companies (Ernst & Young, 2011)

related fraud is commission rebating. "Rebating is defined as giving a customer something of monetary value in exchange for making a purchase. This is usually conceived of as cash discounts, but can include expensive gifts, free trips or concert tickets, prizes, anything of significant value¹". According to 50% of respondents in the Ernest and Young report, the areas that require more regulation are claims and surrender. 27% of respondents declared that applications constitute the second largest risk exposure for fraud that requires more regulation. In this area there is concern about money laundering activities. The authors (Ernst & Young, 2011) suggest that institutions should invest more in the Know Your Customer (KYC) strategy. Some of these recommendations include a centralized KYC database and the use of intelligent systems to monitor and screen the negative list against third party databases.

As shown in Fig.5.3, misselling (fraudulent misrepresentation of material information) constitutes the most important cause of fraud for the respondents. The second cause is collusion between parties in 29% of the cases and fake documentation is the third one with 24% of the cases. Examples of fake documentation fraud in the life insurance business include submitting fake medical test reports and age proof falsified by agents. More than 80% of the respondents estimates that fraud can increase the costs for the insurer by 1% to 5%.

¹http://www.ehow.com/about_6111979_insurance-rebating_.html

Many indicators in this report confirm that there is room for implementing IS to support existing anti-fraud units; 40% of respondents indeed claimed that they do not have a dedicated anti-fraud department in their organisation; 43% of the respondents use manual red flags as a means to detect fraud in their organisation and a systematic screening is done for less than 25% of the key vendors and employees in their organisation. Once again the survey indicates that the most used mechanism to detect fraud is whistle-blowing followed by internal audit.

According to the report insurance fraud can be classified in two categories: hard and soft. Hard fraud occurs when people report a false injury for an accident. Soft fraud occurs when people either lie to their insurance companies or hide certain information for financial gain.

Insurance sectors can benefit from the ICT methodology techniques described in chapter 2. Arson fraud on vehicles and houses are usually assigned to experts who have to find clues in order to prove the fraudulent nature of the fire. Since the evaluation process still lacks in scientific methodologies a fuzzy reasoning experts system using objective criteria could help experts calibrate their opinions and reach a common consensus. Inflated medical bills can be detected using statistical methods based on outlier detection or neural networks. Knowledge based systems can only be applied with high maintenance cost and large amounts of data: these systems can provide solutions for big companies with complex structures and resources.

5.2 Fraud in the European Union and public sector

The European institution responsible for protecting the European Union from fraud, corruption and any other illegal activity is known as OLAF (Office européen de Lutte Antifraude). OLAF (2011) investigates EU members and staff, protects the reputation of European institutions and supports the EU in implementing anti-fraud prevention and detection procedures. OLAF also conducts external investigations in Member States of the EU. The huge amount of information about suspected cases, whistle-blowing and potential fraud activities is managed by a web-based tool, the Fraud Notification System (FNS). Using the website of OLAF² it is possible to report information in 4 steps:

- Submit an initial questionnaire
- Attach document if possible
- Create a password-protected account

²Available at http://ec.europa.eu/anti_fraud/index_en.html

	2006	2007	2008	2009	2010
Assessments completed	462	543	645	740	691
Average duration (months)	5,2	6,2	6,2	7,1	7,4
Prima facie non-case after preliminary review	300	259	243	267	197
Total (assessments + preliminary review)	762	802	888	1007	888

Figure 5.4: Duration of assessment and instances of assessment and preliminary review completed in each calendar year (OLAF, 2011)

Major sector	Co-ordination Case	Criminal Assistance Case	External Investigation Case	Internal Investigation Case	Mutual Assistance Case	Total	Assessment
Agriculture (expenditure + revenue)	37	44	29	0	7	117	101
Cigarettes	16	9	0	0	0	25	4
Customs	17	1	13	0	6	37	8
Direct Expenditure	0	2	43	0	0	45	58
EU Institutions + EU Bodies	0	7	24	108	0	139	94
External Aid	0	3	74	0	0	77	82
Structural Funds	2	5	46	0	0	53	201
Total	72	71	229	108	13	493	548

Figure 5.5: Amounts recovered from closed financial follow ups in € million in each calendar year (OLAF, 2011)

- Communicate with an investigator via a "blind" mailbox into which both parties can drop off messages

In 88% of the cases most of the informants fall into 3 main categories: the general public, the European Commission and member State authorities. OLAF, once it verifies the trustworthiness and the fulfilment of certain conditions, offers protection to whistle-blowers.

In the report authors underline the importance of anonymity in the FNS and the cooperation between OLAF operators, national authorities and Member States, which has to put in place good quality control instruments and implement anti-fraud measures. The report specifies that OLAF is an organisation driven by the "learning by doing" principle. Based on this principle, OLAF uses the experience of its experts to examine new cases in the light of the mistakes and improvements of previous cases. In the first phase of OLAF methodology, evaluators perform an initial scan of the reports before recommending the most serious cases to the Director General, who can decide to open the case or not.

Before the introduction of this methodology (2004), the average time for the assessment of a case was 10.6 months (in 2002). By analysing Fig. 5.4 one can appreciate the benefits of the introduction of this procedure in 2006 and the following years. Years 2009 and 2010 show a significant increase, but the numbers are still considerably lower than in 2002.

In Fig. 5.5, the amounts recovered in different sectors in 2010 are listed: the structural fund sector (€32.9 million), followed by agriculture (€11.9 million) and direct expenditure (10.6 €). Ongoing cases have resulted so far in a further €351.2 million.

We report four fraud cases in EU as they appear in Open Europe (2008).

Belgian city spends €12 million on junkets and dinners: A secret account containing non-declared funds of the Belgian city of Charleroi was discovered in 2007. The account had been used to illegally put away EU funds that the town had received - amounting to some €12 million. The account was apparently used to fund a whole range of junkets and dinners. For instance, a delegation of members of the ruling Socialist Party (PS) used the money to go to Belarus on a hunting trip. The city's Secretary, and current President of the Walloon Parliament, Bernard Bermils Jos Happart was among the travellers. Happart defended himself saying, "If I'm invited, I don't ask where the money has come from". Bizarrely, the party was also handed an illegal \$ 3,000 cash donation from the Embassy.

€50 million to "ghost farmers" : The EU paid out approximately 50 million euros during the period 2001-2004 to farmers in southern Italy, for buying and selling surpluses of citrus fruits under the EU's Common Agricultural Policy. However it was later revealed that the farmers, buyers and even the fruit did not actually exist.

Bulgarian fraudsters steal €9.6 million of EU funds : Bulgarian "businessmen" Mario Nikolov and Ljudmil Stojkov, who have close ties with the Bulgarian President, siphoned off €9.6 million from the EU's Agricultural and Rural Development SAPARD program. Nikolov and Stojkov used false documents on numerous occasions to import used meat processing and packing machines, which they presented as brand new, in turn allowing them to purchase the machines with SAPARD funds. Stojkov has also been charged separately with money laundering.

MEPs "misuse" £100m worth of staff allowances : Senior MEPs and EU officials tried to hush up an internal audit that found severe problems and endemic misuse of funds worth at least £98.4 million a year, more than £125,000 for each of the 785 Euro-MPs. Many MEPs were found to be diverting office payments to 'service providers', which were supposed to be accountants, professionals or companies delivering administrative services. But in many cases the whole allowance was paid to a single individual or MEP's member of staff.

Another sensitive target for fraud is the public Sector. The cabinet Office Counter Fraud (2011) report estimates the amount of fraud in the public sector in UK to be £21 billion, that is 55% of the nation's total fraud loss.

In October 2010, the Government established the Counter Fraud Task Force. In January 2011, in order to support the Task force in fighting fraud, the Cabinet Office created a network of Counter Fraud Champions (CFC)

in every department. The Task Force has four priorities:

- **Collaboration:** all departments in the public sector must work together sharing information on fraudsters and perform data analysis.
- **Assessment of Risk and measurement of losses:** losses must be reported via quarterly data summary.
- **Prevention:** prevention is a priority and where most of resources and investment should go.
- **Zero tolerance:** the public sector does not follow the economic imperative and there is no acceptable level of fraud.

As shown in Fig.5.6 tax fraud is the area with the biggest loss, including £7 billion from tax evasion, £3 billion from lost taxes in the hidden economy and £5 billion lost in criminal attacks. Another important loss is in Procurement Fraud. Typically this kind of fraud consists of collusion between suppliers using falsified or duplicated receipts. Another £515m is lost to grant fraud each year. In this case there are applications from fictitious organisations or individuals for public funds. It includes cases where funds are not used for the purpose of the application once they are obtained.

Amongst the recommendations and suggestions to improve fraud detection and prevention, authors mention the establishment of a repository of fraud related information and supporting material. They also encourage the promotion of an e-learning tool to train personnel and improve their awareness of fraud.

Large organisations such as the EU and governmental institutions are particularly suitable for implementing complex knowledge based systems, since they are rich in human and economic resources and they have to deal with heterogeneous data. Almost all the techniques could be applied in organisations like OLAF or the EU by crossing different databases. Given the huge amount of data in these organisations, there is high potential for implementing a proactive approach to create case-based scenarios dealing with different parameters. EU fraud cases or tax evasion in the public sector can be detected using traditional statistical techniques, looking for outliers deviating from the usual behaviours and analysing discrepancies by matching different databases (owning a luxury car or a boat with low personal income).

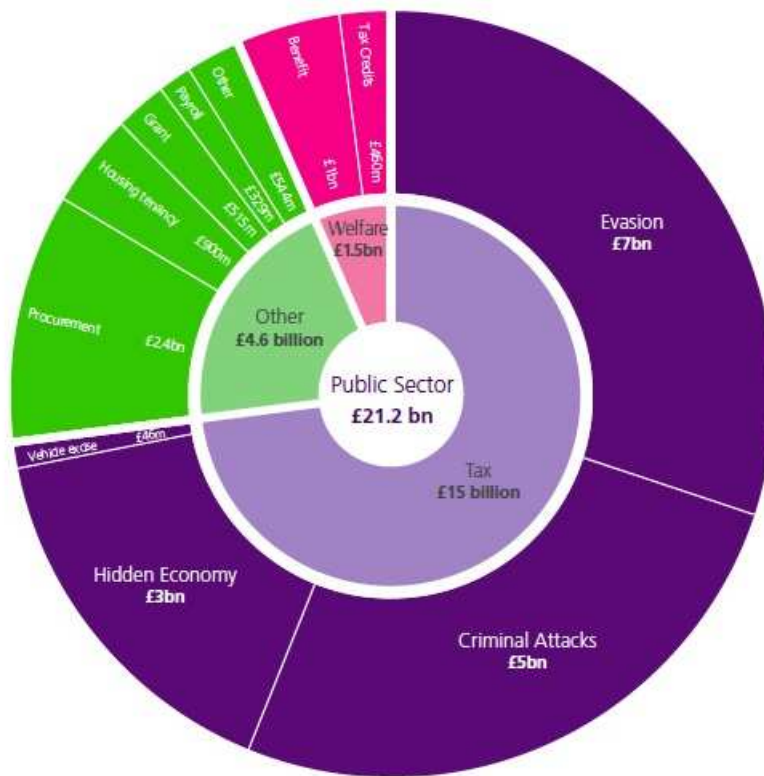


Figure 5.6: Fraud in public sector in UK (The cabinet Office Counter Fraud, 2011)

5.3 Money laundering

Money laundering is defined by the Cambridge Advanced Learner's Dictionary³ as "the crime of moving money that has been obtained illegally through banks and other businesses to make it seem as if the money has been obtained legally." Money laundering business is the main instrument of organised crime to legitimise their money coming mostly from drug and arms trafficking, terrorism, racketeering and prostitution. The necessity of laundering money is related to the danger of storing large amounts of cash. The main difference between money laundering and fraud is the purpose of the two activities: in fraud, the aim is to take advantage in economic terms or improve your reputation; in money laundering the purpose is to hide the origin of illicit business. Once the money is laundered, the launderer can manage a perfectly genuine business. Revenues from this legal business can be reinvested in the same business or they can be used to buy illegal products and start the cycle again. To perform a money laundering scheme, criminals frequently use fraud techniques as well. An example is identity fraud, often used to hide identity in money laundering schemes.

OECD (2006) provides various example of identity fraud schemes used in money laundering and tax evasion with a list of the most vulnerable industries. Tax evasion is another example of fraud activity that can be used to hide the origin of dirty money.

Spreutels and Grijseeels (2000) point out that the most significant difference between tax evasion and money laundering is that in tax evasion funds are moved to a single location while in the case of money laundering funds are moved to several offshore locations. The authors suggest that tax evasion for money laundering purposes implies a higher level of sophistication compared to traditional tax evasion. On the other hand people pay taxes on legal activities to legitimate the origin of their income coming from illegal activities.

This can be done by using the money to run a legal business, investing in the real estate market, or simply adding the cash into the till cash machine at the counter, pretending that it is part of business, (owning a restaurant for instance), as suggested by (Brewer, 2007).

Money laundering processes can be divided into three main phases as described in the Financial Action Task Force (FATF, 2005):

Placement. In the placement stage the launderer introduces his illegal profits into the financial system (e.g. depositing the money with a bank or an insurance company).

Layering. In the second stage the launderer separates the criminal proceeds from their source by the creation of layers of transactions designed

³Available at <http://dictionary.cambridge.org/dictionary/british/money-laundering>

to disguise the audit trail and provide the appearance of legitimacy. He/she will process the funds in possibly several transactions through the financial services sector by purchasing various financial instruments. The purpose of the transactions is to break the (paper) trail between funds and their origin.

Integration. In the final stage of integration the criminal proceeds are treated as legitimate. If layering has succeeded, integration places the criminal proceeds back into the economy in such a way that they appear to be legitimate funds or assets.

The following **high risk areas** are specified in FATF (2005) :

- Insurance Industry
- PEPs (politically exposed persons)
- Wire transfers (Complex wire transfer schemes). Currently, there are a limited number of indicators to help identify potential terrorist wire transfers.
- NPOs (non-profit organisations) ;

High risk products and services (Office of the Comptroller of the Currency, 2002)

- Electronic Banking
- Private Banking Relationships;

High-Risk Customers

- Non-bank Financial Institutions (NBFI)
- Non-governmental organisations (charitable organisations)
- Offshore corporations.

FATF members have noticed an increase in schemes involving gatekeepers. Gatekeepers provide a range of services, such as providing advice, preparing legal documentation and carrying out certain types of financial transaction. One typical example, for instance, can be the purchase of real estate.

In Reuter and Truman (2004) a list of the most common money laundering techniques is provided:

- **Smurfing:** this involves breaking down cash deposits into amounts below the reporting threshold of \$ 10,000. Couriers (smurfs) are used to make deposits in several banks.

- **Informal value transfer systems (hawalas):** *hawalas* is an Arabic word for a particular underground banking system, where the *hawaladar* can personally carry cash from country A to country B, taking care of all the service, from placement to integration.
- **Wire electronic funds transfers:** funds can be transferred by smurfing to a principal collecting account, often located abroad in an offshore financial center.
- **Legitimate business ownership:** "dirty" money can simply be added to the revenues of legal (cash intensive) activities (bar, restaurants).
- **Shell corporation:** these are set up, usually offshore, complete with bank accounts; money can reside there in the layering phase.
- **Real estate transactions:** properties can be bought and sold under a false name or for a shell corporation.
- **Overvaluing imports:** if an imported product is overvalued, the foreign exporter receives an inflated value for the product, and wealth is shifted from the domestic importer to the foreign exporter.
- **Undervaluing exports:** the money launderer converts his illegal money into products by purchasing products for cash at the market price. The products are then exported to a foreign colluding importer at below market prices. The foreign importer receives the undervalued exports and resells them in the market at the real prices that reflect their true value.

The success of these two last schemes is a consequence of lack of controls by the governments in import-export activities.

- **Casinos:** chips are bought with cash. After a certain period, since casinos have establishments in different countries, they can be moved to other casinos and the customers can receive cash to purchase goods.
- **Self-money laundering:** the laundering of one's own money.

Figure 5.7 represents the universe of transactions. By analysing the scheme it is possible to appreciate the interaction between the illegal and the legal world and the complexity of the problem. We can divide transactions into four main groups:

- **Legal/Usual:** in everyday life we withdraw money, with a certain frequency (once a week, for instance).

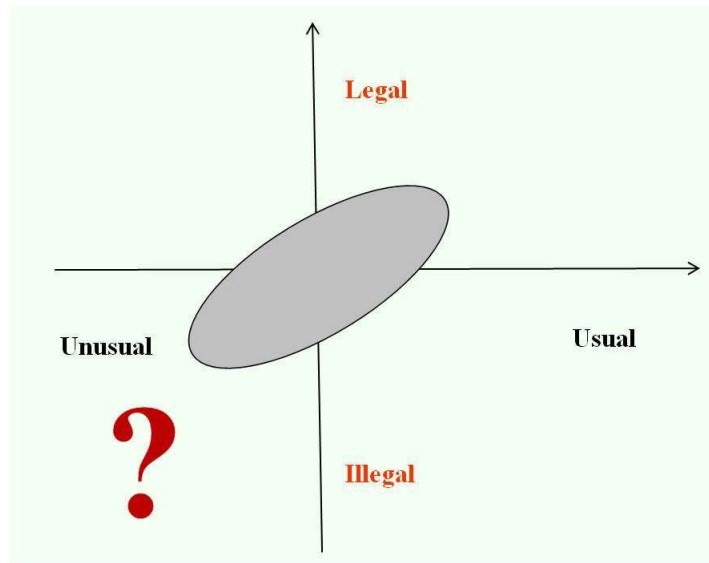


Figure 5.7: Universe of transactions

- **Legal/Unusual:** sometimes the number of our withdrawals can be unusual in terms of the amount or frequency, according to our usual behaviour. In that case the bank can call to verify if we are still in possession of the ATM card or credit card.
- **Illegal/Usual:** these are typical transactions and procedures that banks can easily identify as illegal according to their experience.
- **Illegal/Unusual:** these ones are the most difficult to detect because they are not typical or common or are still unknown.

The main problem with money laundering is that the transactions involved can belong to all these four categories, creating a sort of grey zone where legal and illegal activities are mixed, as a result of complicated schemes. The money launderers' activities consist of moving money from the legal to the illegal world and what seems to be legal business actually hides an illegal activity. It is almost impossible to detect these small transactions and at the same time connect them with unusual behaviours. How can we define "suspicious behaviour"? In practice the real challenge is not to separate all the single transactions, but to find a meaningful path between transactions and associate them with suspicious behaviour like a sudden change of life style or typical money laundering red flags. Suspicious behaviours related to bank activities include: reluctance to provide personal information, customers with multiple bank accounts, frequent deposit and withdrawals without any business activity to justify these operations, activity within high

risk areas and with non cooperative countries listed by FAFT⁴. A general view on money laundering, a list of red flags and suspicious behaviours can be found in the Federal Financial Institutions Examination Council manual (FFICE, 2010).

The OECD (2006) report shows how the real estate sector tax evasion can hide money laundering schemes. Three ways of concealing ownership are used: a) onshore acquisitions through off-shore companies and/or through a complex structure of ownership; b) unreported acquisition of properties overseas; and c) use of nominees. In the same report the role of fraud detection schemes in uncovering ML schemes is specified, since there are no specific techniques to identify the tax payer or money launderer in the real estate sector. Different techniques, like risk analysis, risk profiling and case selection are integrated with modern tools based on data mining and statistical analysis.

5.4 FIDES in other contexts

In this chapter a typology of crimes related to fraud in Money Laundering is introduced with fraud schemes in the European Union, Insurance and public sector. All these different organisations base their anti-fraud strategy on the analysis of reports, performed by an anti-fraud central unit. This unit has to analyse the reports, set priorities, remove false alarms and take actions on the basis of a report or suspects, filed by inspectors, agents, operators or normal citizens. All these features suggest that FIDES can be implemented to improve the interaction between agents, developing a case-based reasoning approach in storing different cases in order to detect and prevent fraud.

OLAF authorities introduced the Fraud Notification System (FNS) to submit information related to fraud cases, suspects and suspicious behaviours. They underline the importance of anonymity in their system in a similar way as it was emphasized in the description of FIDES. All the information collected in the questionnaires is stored in a dedicated server and used as a source of intelligence/evidence (OLAF, 2007). This data could be the basis for OLAF evaluators to use as the input for FIDES. In this case the interaction takes place between OLAF evaluators and the Director General. On the basis of the information submitted in the FNS and acquired during the personal investigations, the most dangerous cases can be taken into consideration and the FIDES processes can be activated.

In the UK public sector the Cabinet Office encourages collaboration between departments and different sectors of public administration and insists on investing in prevention by creating a repository for information and e-

⁴Available at http://www.fatf-gafi.org/document/4/0,3343,en_32250379_32236992_33916420_1_1_1_1,00.html

learning tools to improve employees awareness of fraud. These needs can be accomplished by a MAS like FIDES since it has the capability to improve the interaction between users and to build maps and attack trees, where relevant information can be stored.

In money laundering, banks and money transfers companies have the duty to fill in a Suspicious Activity Report (SAR), in case customers show suspicious behaviour. In Fig. 5.8, the Suspicious Activity Report (SAR) is shown⁵, including part I and part II. In part I the fields are dedicated to the personal data of the suspect. In part II the officer responsible for filling out the report has to add information about the activity.


Particularly interesting is part VI of SAR, where there is a space to fill with a text based description of the case. Members of the institutions have to include all the information that was not requested in the previous section of the report, but it can be useful to have a complete description of the case. An automatic text retrieval of SAR documents and text summarization can be used as input for FIDES to generate think-maps first and then attack trees. In think-maps the red-flags and risk areas can be underlined in order to support inspectors in building the attack tree.

Insurance companies have a report system based on hot-lines and on-line reports. All these systems can be integrated in FIDES to create an attack tree catalogue to support anti-fraud experts in fighting and preventing fraud.

The arson expert Lentini (Lentini, 2011) points out the importance of determining the cause of fire and recognizing arson patterns. In the '80s and '90s fraud detection in connection with fire was performed based on "anecdotal evidence at best and witchcraft at worst". This caused the conviction of many innocent people. Presently most of the fire investigators agree on using scientific methods to determine fraud, but still nowadays some of them neither understand nor follow the scientific method. FIDES in this context could be used also not only as a fraud or case-based reasoning detection system, but also as a learning tool for fire investigators. The cases stored in the database as think-maps or attack trees could be used to train them and expand their knowledge.

Millner and Duhl (2011) suggest that insurance investigators should use search engines and social networks to gather information about claimants' profiles. They report cases of people sharing information about their illegal behaviours or their intention to commit fraud. The use of social networks can be the new frontier for investigation. Fraud investigation can be supported by screening the status of people using social networks or forums of discussion and all the information can be retrieved and represented in a structured way in order to detect or prevent possible fraud.

⁵Available at <http://www.docstoc.com/docs/10131012/FinCEN-Form-109>

FinCEN Form 109 March 31, 2007 Previous editions will not be accepted after September 30, 2007 (Formerly Form TD F 90-22.56)	Suspicious Activity Report by Money Services Business Please type or print. Always complete entire report. Items marked with an asterisk * are considered critical. (See instructions.)	 OMB No. 1506-0015
1 <input type="checkbox"/> Check this box only if amending or correcting a prior report (see item 1 instructions) 1a <input type="checkbox"/> Check this box if this is a recurring report		
Part I Subject Information 2 <input type="checkbox"/> Multiple subjects (see item instructions)		
3 Subject type (check only one box) a <input type="checkbox"/> Purchaser/sender b <input type="checkbox"/> Payee/receiver c <input type="checkbox"/> Both a & b z <input type="checkbox"/> Other		
*4 Individual's last name or entity's full name		*5 First name
6 Middle initial		
*7 Address		
*8 City	*9 State	*10 ZIP Code
*11 Country Code (If not US)		
*12 Government issued identification (if available)		
a <input type="checkbox"/> Driver's license/state I.D. b <input type="checkbox"/> Passport c <input type="checkbox"/> Alien registration z <input type="checkbox"/> Other _____ e Number _____ f Issuing state/country _____		
*13 SSN/ITIN (individual) or EIN (entity)	*14 Date of birth	15 Telephone number
Part II Suspicious Activity Information		
*16 Date or date range of suspicious activity		*17 Total amount involved in suspicious activity a <input type="checkbox"/> Amount unknown
From MM/DD/YYYY To MM/DD/YYYY		\$ _____,00
*18 Category of suspicious activity (check all that apply)		
a <input type="checkbox"/> Money laundering b <input type="checkbox"/> Structuring c <input type="checkbox"/> Terrorist financing z <input type="checkbox"/> Other (specify) _____		
*19 Financial services involved in the suspicious activity and character of the suspicious activity, including unusual use (check all that apply).		
a <input type="checkbox"/> Money order b <input type="checkbox"/> Traveler's check c <input type="checkbox"/> Money transfer z <input type="checkbox"/> Other _____ e <input type="checkbox"/> Currency exchange		
Check all of the following that apply		
(1) <input type="checkbox"/> Alters transaction to avoid completing funds transfer record or money order or traveler's check record (\$3,000 or more) (5) <input type="checkbox"/> Individual(s) using multiple or false identification documents (2) <input type="checkbox"/> Alters transaction to avoid filing CTR form (more than \$10,000) (6) <input type="checkbox"/> Two or more individuals using the similar/same identification (3) <input type="checkbox"/> Comes in frequently and purchases less than \$3,000 (7) <input type="checkbox"/> Two or more individuals working together (4) <input type="checkbox"/> Changes spelling or arrangement of name (8) <input type="checkbox"/> Same individual(s) using multiple locations over a short time period (9) <input type="checkbox"/> Offers a bribe in the form of a tip/gratuity (10) <input type="checkbox"/> Exchanges small bills for large bills or vice versa		
If mailing, send each completed SAR report to: Enterprise Computing Center - Detroit Attn: SAR-MSB P.O. Box 33117 Detroit, MI 48232-0980		A free secure e-filing system is available to file this report. Go to http://bsaeiling.fincen.treas.gov for more information and to register.

Catalog No. 49340J

(Rev. 3/07)

Figure 5.8: SARs part I and II

Chapter 6

Summary, Conclusions, Future research

This research started with money laundering, a very complex phenomenon that is very difficult to detect, since money laundering goes back and forth from legal to illegal business. The activity of smuggling is often perpetrated by human actors who carry on their persons large sums of money in order to open bank accounts abroad in tax havens.

These limitations encountered while studying this subject and the opportunity I had to meet the audit team of an important European bank led me to change my research topic to fraud detection. Lessons learned from studying the money laundering phenomenon, however, were invaluable in understanding the gap that exists between the description of this subject and the reality.

The reality of financial crime is often not as fascinating as cyber-thrillers portray. Identity fraud is indeed the most common type of technique used by fraudsters.

The instruments used for stealing personal information are not only based on hacking systems such as Trojan horses or malware, but one can also find less sophisticated ones.

Internal fraud in the banking sector, as in other kinds of institutions such as insurance companies or governmental ones, is characterized by conflict of interest.

In order to execute a fraudulent scheme one needs the complicity of the management, trust of the co-workers and a good lawyer.

The distorted representation of fraud includes not only the prevalence of electronic hacking systems over less sophisticated techniques, but also the motivation that leads people to commit financial crimes.

An interpretation key for understanding this concept is the venality vs. banality hypothesis introduced by Kim (2005). According to the venality

hypothesis, which is commonly adopted in literature and in the business world, fraudsters are driven by greed and they are aware of the consequences of their actions. This interpretation drives organisations to limit the countermeasures to a list of good intentions such as improving control procedures. These good intentions ignore conflict of interest, or, analysing it from a cynical point of view, they might be the product of it. Based on the banality hypothesis the main reason that drives people to commit fraud is the banal tendency to follow the boss' orders. In an efficiency-oriented society people have to adhere to the job description.

In most cases orders are executed without perceiving the reasons why things are done, thus their level of integrity is not an issue at all for the employee. In the venality hypothesis, there is a simplification of human morality. In other words there is a common understanding about who the "good" and the "bad" employees are.

Based on the banality hypothesis, a good lawyer or a good accountant is the one who helps the company achieve its goals and be more competitive on the market.

The banality hypothesis is not simply a cynical way of studying fraud, but a valid theoretical base for developing effective anti-fraud techniques. The banality hypothesis suggests that rather than just having a list of good intentions it would be more efficient to develop anti-fraud systems based on anonymity and activate programs to protect whistle-blowers.

By adopting this principle, the audit team will benefit in developing more efficient anti-fraud countermeasures. An audit team also has to be able to improve the reporting systems and map these behaviours in a user-friendly way in order to prevent fraud schemes. A major bank has many branches even in different countries; this means that inspectors from different countries describe and report their suspicions according to their cultural background. Their description could be perceived by the audit-team in a distorted way. The audit team of a bank requires a system able to unify risk criteria and to adapt them to the context. This can be performed by improving the interaction between inspectors and auditors in order to produce a shared representation of the fraud schemes and a common agreement on the counter-strategy to adopt. The audit-team activity deals with the information received and with the interpretation of the facts and suspicions.

The experience that auditors have accumulated during their careers, and their ability to integrate missing information with similarities from past cases, is essential in fraud detection.

FIDES is designed according to these premises: audit-team's needs and capabilities. Modelling these behavioural aspects, improving interaction between auditors and inspectors and creating a knowledge-based system to retrieve information about past cases, are the main features of FIDES. Auditors can benefit from using FIDES since they can improve the interaction

with the inspectors and obtain semi-structured information to start their evaluation process.

Inspectors can use the system as a platform to communicate with the audit team and force themselves to improve the accuracy of their reports. A multi-agent system such as FIDES can work in all those contexts where there is need for interaction between a central anti-fraud unit and inspectors/operators who have to report irregularities back to this main anti-fraud unit.

Other sectors where FIDES can be applied include insurance companies, public organisations and governmental institutions, but also in other risk management contexts: police, fire-fighters and medical personnel who need to coordinate complex tasks in emergency situations such as terrorist attacks, natural disasters or multiple car accidents.

6.1 Future research

The head of the risk management department of the bank I interviewed expressed his ideas about the future of fraud and its detection. The most important predictions concerning the future of fraud are the following:

- robberies will not be issues anymore for banks, since the security systems will be increasingly effective
- the use of advanced anti-hacking technology will easily block external attacks
- fraud could increasingly become a necessity for employees as a consequence of the financial crisis
- the connections of fraud activities with criminal organisations might increase

Based on these predictions, we can claim that the potential of FIDES is that of a system that can fight crime in general and not only fraud.

In the future, once the data-base is populated with a significant number of fraud cases, auditors can produce a text description of the fraud cases and then these interpretations could be matched with the information stored in the data-base. Using this procedure, the different interpretations of the auditors could be summarized quickly and the attack tree or a possible rough version of it can be produced directly and refined in the Delphi process.

At this stage also the information retrieval features of the system will be refined to support the evaluation process of the audit team. Retrieving old cases to be compared with new ones can be useful also in discovering new schemes related to old attacks. Fraud schemes can be perpetrated in a

long window of time by the same author. In this case the system could alert auditors about similarities, indicating that the same criminal is behind that scheme and offering strong evidence to incriminate the fraudster.

In a similar way, the comparison of geographical attributes could provide an indication that the same criminal organisation is operating in a certain area.

In general, the matching process could suggest that the style of the actual fraud scheme is similar to an old one offering a valid support for the investigation. Retrieving the countermeasures or strategies adopted in the past for similar cases could be useful to quickly develop a counter-strategy. Inspectors could also benefit from retrieving the information to prune their suspicions in the case they are analysing. In some cases the retrieved data can reveal that a certain behaviour is just routine and not an element of a potential fraud case. Inspectors could also check profiles of fraudsters stored in the database to obtain useful information to write their reports.

Another important future improvement of FIDES would be its integration with different types of data-bases in order to edit criminal profiles.

In FIDES the Risk management phase consists of choosing the most suitable path amongst the possible ones offered by the attack tree. The choice of the path can be supported by different aggregation methods and reasoning procedures. The choice and development of these methods could be a crucial improvement to FIDES. The risk management phase can benefit from valid and well-structured methods to obtain not only the shortest, but also the least rational path and other types of combinations that auditors perceive as important.

As suggested by Millner and Duhl (2011), the new frontier of fraud investigation lies in retrieving information from social networks. Hackers could use social networks to share their deeds in their profiles or communicate using a cryptic language. The integration of FIDES with the Web and social media is certainly a feature that can be developed in the future. The creation of a more accurate fraud ontology using a semantic search engine could improve the quality of the information and expand the detection capabilities of the system.

In the distant future we can hypothesize fraud detection systems driven by virtual experts: virtual agents programmed with different personalities or, to use a more specific term, with different ontologies. Virtual auditing would be the term to define the environment where FIDES could operate in the future. Virtual auditors could be programmed with different personalities, experience of past fraud cases and different risk aversions in order to perform the audit process simulating human capabilities and the decision-making process.

References

- Abbott, D., Matkovsky, P. and Elder, J. (1998). An Evaluation of High-End Data Mining Tools for Fraud Detection. *Proc. of IEEE SMC98*.
- Åhlberg, M. (2007), History of Graphic Tools Presenting Concepts and Propositions, Available at <http://www.reflectingeducation.net/index.php?journal=reflecting&page=article&op=downloadSuppFile&path\%5B\%5D=49\&path\%5B\%5D=6>
- Acfé, (2008), Report to the Nation on occupational fraud & abuse, Available at <http://www.acfe.com/documents/2008-rtnn.pdf>.
- Aldrich, N. (2008), Medicare Fraud Estimates: A Moving Target?, Available at http://www.smpresource.org/Content/NavigationMenu/AboutSMPs/MedicareFraudEstimatesAMovingTarget/Medicare_Fraud_Estimates.pdf
- Aumann, R. J. (1965). Integrals of set-valued functions, *Journal of Mathematical Analysis with Applications*, 12, 1-12, 1965.
- Balanced Scorecard Institute, (1996), Cause and effect diagram, Available at <http://www.balancedscorecard.org/Portals/0/PDF/c-ediag.pdf>.
- Basel Committee on Bank Supervision (2001). The New Basel Capital Accord.
- Bazerman, M. Lowenstein, G. and Moore, D., (2007), Why good accountants do bad audits, *Harvard Business Review*, 3-8.
- Beliakov G., Pradera A., and Calvo T. (2007). Aggregation Functions: A Guide to Practitioners, Springer-Verlag, Heidelberg.
- Bolton, Richard J. and Hand, D. J., (2002), Statistical Fraud Detection: A Review, *Statistical Science*, vol.17, No.3, 235-255.
- Bordoni S. and Facchinetti, G. (2001) Insurance Fraud Evaluation : a fuzzy expert system, *IEEE International Fuzzy Systems Conference*, Volume: 3, 1491-1494.
- Bortot, S., Fedrizzi, M. Giove, S. (2011). "Modelling fraud detection by attack trees and Choquet integral", *Advances in fuzzy sets and systems*, v. 9, n. 2, 137-165
- Brewer, P., (2007), How to Launder Money, Wisebread, Available at <http://www.wisebread.com/how-to-laundry-money>.

- Buoni, A. (2010) Fraud detection: from basic techniques to a multiagent-approach, *International Conference on Management and Service Science*, 24-26 August, Wuhan, pp.14.
- Buoni, A. Fedrizzi, M.(2012),Consensual dynamics and Choquet Integral in an attack tree-based fraud detection system,*Proceedings of the 4th International Conference on Agents and Artificial Intelligence*, Volume 1,283-288, Algarve (Portugal)6-8 February.
- Buoni, A. Fedrizzi, M. and Mezei, J. (2010). A Delphi-Based Approach to Fraud Detection Using Attack Trees and Fuzzy Numbers, *In Proceeding of the IASK International Conferences*, 21-28.
- Buoni, A., Fedrizzi, M. Mezei, J. (2011), Combining attack trees and fuzzy numbers in a multi-agent approach to fraud detection,*Int. J.Electronic Business*, Vol. 9, No. 3, 186-202.
- Buzan, T.,(1974) *Use Your Head*, BBC Books, London.
- California Department of Insurance,(2012). Fraud: What is Insurance Fraud? Available at <http://www.insurance.ca.gov/0300-fraud/0100-fraud-division-overview/0100-what-is-insurance-fraud/#Medical>
- Cañas, A. J., Hill, G., Lott, J.and Suri, N. (2003). Permissions and access control in CmapTools (Technical Report No. IHMC CmapTools 2003-03). Pensacola, FL: Institute for Human and Machine Cognition.
- Carlsson, C. Fedrizzi, M. Fuller, R. (2004). Fuzzy logic in management, Dordrecht, Boston, New York: Kluwer Academic.
- Carlsson, C. and Fuller, R. (2001),On possibilistic mean value and variance of fuzzy numbers, *Fuzzy Sets and Systems*, Vol. 122, 315-326.
- Carpinteiro, O.A.S. Netto, R. S., Lima, I., Zambroni de Souza, A.C., Moreira, E. M. and Pinheiro, C.A.M.,(2006).A Neural Model in Intrusion Detection Systems, S.Kollias et al. (Eds.): ICANN 2006, Part II, LNCS 4132, 856-862.
- Chakraborty, R.C. (2010),Fuzzy sets and theory:Soft Computing Course lecture. Available at http://www.myreaders.info/06_Fuzzy_Set_Theory.pdf
- Chan, P., Fan, W., Prodromidis, A. and Stolfo, S. (1999). Distributed Data Mining in Credit Card Fraud Detection.*IEEE Intelligent Systems* 14: 67-74.

- Choquet, G., Theory of capacities, *Annales de l'Institut Fourier*, 5, 131-295, 1953.
- Chou, C.L-Y Du, T. V Lai. , S., (2007), Continuous auditing with a multi-agent system, *Decision Support Systems*,42 (4) 2274-2292.
- Deshmukh, A.and Talluru, T.L.N.,(1997), A Rule Bases Fuzzy Reasoning System for Assessing the Risk of Management Fraud, *Intelligent Systems in Accounting, Finance & Management*, Volume 7 Issue 4,223-241.
- Dimitriadis, K.,(2007), Analyzing the Security of Internet Banking Authentication Mechanism, *Information System Control Journal*, Voloume 3, 1-8.
- Dorronsororo, J., Ginel, F., Sanchez, C. and Cruz, C. (1997). Neural Fraud Detection in Credit Card Operations. *IEEE Transactions on Neural Networks* 8(4): 827-834.
- Edge, K., Raines, R., Grimaila, M., Baldwin, R., Bennington, R., & Reuter, C. (2007). The Use of Attack and Protection Trees to Analyze Security for an Online Banking System. *2007 40th Annual Hawaii International Conference on System Sciences HICSS07*, 144b-144b. Ieee. Retrieved from <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4076665>
- Ernst & Young,(2011),Fraud in insurance on rise Survey 2010-2011. Available at [http://www.ey.com/Publication/vwLUAssets/Fraud_in_insurance_on_rise/\\\$FILE/Fraud_in_insurance.pdf](http://www.ey.com/Publication/vwLUAssets/Fraud_in_insurance_on_rise/\$FILE/Fraud_in_insurance.pdf)
- FATF,(2005), Money laundering & Terrorist financing typologies 2004-2005. Available at <http://www.fatf-gafi.org/dataoecd/16/8/35003256.pdf>.
- FATF,(2004), Report on money laundering typologies 2003-2004.Available at <http://www.fatf-gafi.org/dataoecd/19/11/33624379.pdf>
- Fayyad U., Piatetsky-Shapiro G., and Smyth P., (1996). From Data Mining to Knowledge Discovery in Databases, *AI Magazine*, vol.17, number 3.
- Fedrizzi, M., Fedrizzi, M., and Marques Pereira, R.A. (1999). Soft consensus and network dynamics in group decision making. *Intl. Journal of Intelligent Systems*, 14, 63-77.
- Fedrizzi, M., Fedrizzi, M., and Marques Pereira, R.A. (2007). Consensus modelling in group decision making: a dynamical approach based on fuzzy preferences.*New Mathematics and Natural Computation*, 3, 219-237.
- Fedrizzi,M. Giove,S.,(2007) Multi-expert fraud risk assessment through OWA-based attack trees, FIDIPRO-IAMSR Workshop, Turku (FI).

- FFICE (2010), Bank Secrecy Act/Anti-Money Laundering Examination Manual.
- Flegel, U., Vayssire, J. and Bitz, G., (2008), Fraud Detection from a Business Perspective: Future Directions and Challenges, Available at <http://drops.dagstuhl.de/opus/volltexte/2008/1795/>.
- Fodor J, Marichal J. L., and Roubens M., Characterization of the ordered weighted averaging operators, *IEEE Trans. on Fuzzy Systems*, 3, (2), 236-240, 1995.
- Franklin, K.K. and Hart, J.K., (2007), Idea Generation and Exploration: Benefits and Limitations of the Policy Delphi Research Method, *Innovative Higher Education*, volume 31 (4) 237-246.
- Fugate M. and Gattiker, J., (2003), Computer intrusion detection with classification and anomaly detection, using SVMs, *International Journal of Pattern Recognition and Artificial Intelligence*, Vol.17, No. 3 (2003) 441-458.
- Ghirardato P., Le Breton M., Choquet Rationality, (2000). *Journal of Economic Theory*, 90, (2), 277-285.
- Gordon, T.J., (1994), The Delphi Method, in *Futures Research Methodology*, AC/UNU Millenium Project, Washington, AC/UNU.
- Grabisch, M., Labreuche, (2010). A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid, *Annals of Operations Research*, 175, 247-286.
- Grabisch, M., Marichal, J.-L., Mesiar, R., and Pap, E., (2009). *Aggregation Functions*, Cambridge University Press.
- Grabisch M., Nguyen H. T., Walker E. A. (1995b). *Fundamentals of Uncertainty Calculi, with Applications to Fuzzy Inference*. Kluwer, Boston, MA.
- Grabisch M., (1995a). On equivalence classes of fuzzy connectives - The case of fuzzy integrals, *IEEE Trans. on Fuzzy Systems* 3, (1), 96-109.
- Grabisch M., (1996). The applications of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research*, 89, (3), 445-456.
- Grazioli, S., Johnson P. E. and Karim, J. (2006), A cognitive approach to fraud detection. Available at <http://ssrn.com/abstract=920222>.

- Greenberger, D.B., Miceli, M.P., Cohen, D.J., (1987), Oppositionists and Group Norms: The Reciprocal Influence of Whistle-blowers and Co-workers, 6 *J. Bus. Ethics*.
- Hand, D.J. (1997), Construction and Assessment of Classification Rules, Chichester: Wiley.
- Hand, D.J., (1981), Discrimination and Classification, Chichester: Wiley.
- Hevner, A.R., March, S.T. Park, J. and Ram, S. (2004). Design science in information systems research. *MIS Q.* 28, 1 (March), 75-105.
- Huang, S.-M., Yen, D.-C. Yang, L.-W., Hua, J.-S., (2008), An investigation of Zips law for fraud detection, *Decision support system*, 46, 70-83.
- Insurance fraud hotline, (2012), Arson, Available at <http://www.insurancefraudhotline.com.au/About-Insurance-Fraud/Arson.aspx>.
- Kim, Sung Hui, The Banality of Fraud: Re-Situating the Inside Counsel as Gatekeeper. *Fordham Law Review*, Vol. 74, No. 983, 2005. Available at SSRN: <http://ssrn.com/abstract=876125>
- KDD Cup 1999 Data, Available at <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html>.
- Klement, P., Mesiar, R., and Pap, E., (2010). A universal integral as common frame for Choquet and Sugeno integral. *IEEE Transactions on Fuzzy Systems*, 18 (1), 178-187.
- Kolb, D A Experiential Learning: Experience as the Source of Learning and Development Prentice-Hall, New Jersey (1984)
- KPMG (2003), Available at http://www.kworld.com/aci/docs/surveys/Fraud%20Survey_040855_R5.pdf.
- KPMG (2007), Available at <http://www.kpmg.co.uk/pubs/ProfileofaFraudsterSurvey%28web%29.pdf>.
- KPMG (2010), Basel 3, Pressure is building?, Available at <http://www.kpmg.com/BH/en/Documents/Basel%203-%20Pressure%20is%20building%E2%80%A6.pdf>
- KPMG (2011), Who is the typical fraudster? Available at <http://www.kpmg.com/US/en/IssuesAndInsights/ArticlesPublications/Documents/who-is-the-typical-fraudster.PDF>

- Kuhn, J. Randel, Jr., and Sutton,S.G., (2006). Learning from Worldcom: Implications for Fraud Detection through Continuous Assurance. *Journal of Emerging Technologies in Accounting*, Vol. 3, 61-80.
- Lentini,J.J.,(2011) Arson probes:Instinct giving way to modern science, *The Journal of Insurance Fraud in America*,8-14.
- Major, J. A. and Riedinger, D. R., (2002), EDF: a hybrid knowledge/statistical-based system for the detection of fraud, *The Journal of Risk and Insurance*, Vol. 69, No.3, 309-324.
- Malek, W. W. Z., Mayes, K. and Markantonakis, K.,(2008),Fraud Detection and Prevention in Smart Card Based Environment Using Artificial Intelligence, G. Grimaud and F.-X. Standaert (Eds.): CARDIS 2008, LNCS 5189, 118-132.
- Marichal J. L. (1988) , Aggregation operators for multi-criteria decision aid, Ph.D. Thesis, University of Lige, Belgium.
- Martignoni, L., Stinson,E., Fredrikson, M., Jha, S. and Mitchell, J.C. (2002), A layered architecture for detecting malicious behaviors, *The Journal of Risk and Insurance*, Vol.69, No.3, 309-324.
- Mauw,S. and Oostdijk,M.,(2005), Foundation of Attack Trees,*Information security and cryptology:ICISC 2005, 8th international conference*, Seoul, Korea, December 1-2,186-198.
- Miller, A.G., (1986), The Obedience Experiments: A Case Study of Controversy in the Social Sciences, Praeger Publishers.
- Milliken,F.J., Morrison, E.W.,Hewlin,P.F.,(2003), An Exploratory Study of Employee Silence: Issues that Employees Don't Communicate Upward and Why, 40 *J. Mgmt. Stud.*
- Millner,J.S. and Duhl,G.M.,(2011), Legal Issues: Using social networking to uncover fraud.*Journal of Insurance Fraud in America*.Vol.2.Number 1,3-7.Eds.Coalition Against Insurance Fraud
- Moore, A.P., Ellison, R.J., Linger, R.C, (2001), Attack Modelling for Information Security and Survivability, <http://www.itsec.gov.cn/docs/20090507164628959972.pdf>.
- Morse,J. and Bower,A. (2002), The Party Crasher, Time, Dec. 30, at 52.
- Mukkamala, S., Zu, D.and Sung, A. H.,(2006), Intrusion detection based on Behaviour Mininig and Machine Learning Techniques, *Advances in Applied Artificial Intelligence*, 619-628.

- Murad, U. and Pinkas, G., (1999). Unsupervised Profiling for Identifying Superimposed Fraud. *Proc. of PKDD99*.
- Nanopoulos, A. and Manolopoulos, Y. (2002) Efficient similarity search for market basket data, *The VLDB Journal*, Vol. 11, 138-152.
- Near, J. P. and Miceli, M. P., (1987), Whistle-Blowers in Organizations: Dissidents or Reformers?, *9 Res. Org. Behav.*
- Novak, J. and Cañas, A. (2006) The Theory Underlying Concept Maps and How to Construct Them, Technical Report IHMC Cmaptools. Available at <http://cmap.ihmc.us/publications/researchpapers/theorycmaps/theoryunderlyingconceptmaps.htm>
- Odubiyi, J. B. and C. W. O'Brien, (2006), Information security attack tree modeling, *In proceedings of Seventh Workshop on Education in Computer Security (WECS7)*, Monterey, California, 04-06 January, 29-37.
- OECD (Organization for Economic CO-operation and Development), (2006), Report on Identity Fraud: Tax evasion and money laundering vulnerabilities.
- Office of the Comptroller of the Currency, (2002), Money Laundering: A Banker's Guide to Avoiding Problems.
- Ohsuga, S. (2001), How Can AI Systems Deal with Large and Complex Problems?, *Presented at IJPRAI*, 493-525.
- OLAF, Annual report 2011, European anti-fraud office.
- Olaf, (2007), Privacy statement for fraud notification system. Available at http://ec.europa.eu/dgs/olaf/data/docpst/fraud_notif.pdf
- Open Europe, (2008), 100 example of EU fraud and waste. Available at <http://www.openeurope.org.uk/research/top100waste.pdf>
- Ou, X. Boyer, W. F., McQueen, M. A., (2006), A scalable approach to graph attack generation, *Conference on Computer and Communications Security Proceedings of the 13th ACM conference on Computer and communications security*, Alexandria, Virginia, USA, Software and network exploits, 336-345.
- Oxman, R. 2004, Think-maps: teaching design thinking in design education, *Design Studies*, Vol. 25, Number 1.
- Piatetsky-Shapiro, G. (1991). Knowledge Discovery in Real Databases: A Report on the IJCAI-89 Workshop. *AI Magazine* 11(5): 68-70.

- Phua, C., Lee V., Smith K. and Gayler R., (2005). A Comprehensive Survey of Data Mining-based Fraud Detection Research, *Artificial Intelligence Review*.
- Fighting fraud in government (2011). Available at http://www.pwc.com/en_UK/gx/psrc/pdf/fighting-fraud-in-government.pdf
- Reuter, P. and Truman, E.M., (2004), Chasing Dirty Money: The Fight Against Money Laundering. Available at http://www.piie.com/publications/chapters_preview/381/3iie3705.pdf
- Ripley, B. D. (1996), Pattern recognition and neural networks, Cambridge University Press.
- Rosen, R.E., (2002), "We're All Consultants Now": How Change in Client Organizational Strategies Influences Change in the Organization of Corporate Legal Services, *Ariz. L. Rev.* 44.
- Rowe, G., Wright, G., (1999) The Delphi technique as a forecasting tool: issues and analysis, *International Journal of Forecasting* 15, 353-375.
- SANS Institute, The Twenty Most Critical Internet Security Vulnerabilities (Updated) - The Experts Consensus, <http://www.sans.org/top20/>
- Schneier, B., (1999), Attack Trees: Modeling security threats, *Dr. Dobbs's Journal*, December, 1-9. Available at <http://www.schneier.com/paper-attacktrees-ddj-ft.html>
- Searcy DeWayne, L. and Woodroof, J.B. (2003). Continuous Auditing: Leveraging Technology. *The CPA Journal*, 1-4.
- Sherman, E. (2002). Fighting Web Fraud. *Newsweek* June 10.
- Simon, H.A. (1996). The Sciences of the Artificial, Third edition. *MIT Press*, Cambridge, MA.
- Singleton, T. and Singleton, A.J. (2005), Auditing Headaches? Relieve Them with CAR. *The Journal of Corporate Accounting & Finance*, 17-27.
- Skulmoski, G. J. Hartman, F. T. and Krahn, J., (2007). The Delphi Method for Graduate Research. *Journal of Information Technology Education* Volume 6.
- Spreutels, J. and Grijseels, C., (2000), Interaction between money laundering and tax evasion, Belgian and international measures in the fight against money laundering. Available at http://www.ctif-cfi.be/website/images/EN/pub_art/s9T10088.pdf.

- Sterne , D., Balasubramanyam, P., Carman, D., Wilson, B., Talpade, R., Ko, C., Balupari, R., Tseng, C.-Y., Bowen, T., Levitt, K. and Rowe, J., (2005), A General Cooperative Intrusion Detection Architecture for MANETs, IWIA, *Proceedings of the Third IEEE International Workshop on Information Assurance*, 57-70.
- Suchman,M.C. (1998) *Working Without a Net: The Sociology of Legal Ethics in Corporate Litigation*, 67 Fordham L. Rev. 837-849.
- The cabinet Office Counter Fraud,(2011), Eliminating Public Sector Fraud. Available at http://www.scarborough.nhs.uk/controlpanel/shoppics/pdfs/Governmentpublicationeliminatingpublicsectorfraud_1.pdf
- Tinca, A. (2007). "The Operational Risk in the Outlook of the Basel II Accord Implementation," *Theoretical and Applied Economics*, Asociatia Generala a Economistilor din Romania - AGER, vol. 5(5(510)), 31-34.
- Turof, M. and Linstone, H.A.(2002), *The Delphi Method: Techniques and Applications*, Available at <http://is.njit.edu/pubs/delphibook/ch2a.html>.
- Wang, D.G., Li, T., Liu, S.J.L., Liang, G. and Zhao, K., (2008), An immune multi-agent system for network intrusion, *Proceedings of the Third International Symposium on Intelligence Computation and Applications* (ISICA 2008), 1921 December, Wuhan, China, LNCS 5370, Springer-Verlag, Berlin Heidelberg, 436-445.
- Webb, A.R. , (1999). *Statistical Pattern Recognition*, London: Arnold.
- Xu, J. , Sung,A. H. and Liu, Q. (2006), Tree Based Behavior Monitoring for Adaptive Fraud Detection, *Proceeding of 18th International Conference on Pattern Recognition* (ICPR'06), Volume ,1208-1211.
- Yager, R. R.,(1988), On ordered weighted averaging aggregation operators in multicriteria decision making, *IEEE Trans. On Systems, Man and Cybernetics*, 18, 183-190.
- Yager, R. R., (2006), OWA trees and their role in security modeling using attack trees, *Information Sciences* 176,2933-2959.
- Yang R., Wang Z., Heng P. A., and Leung K. S. (2005),. Fuzzy numbers and fuzzification of the Choquet integral, *Fuzzy Sets and Systems*, 153, 95-113.
- Yao,J.T. and Liu,W.N.,(2006) Web-based dynamic Delphi: a New Survey instrument, *Proc. SPIE*, Vol. 6241, 62410I; DOI:10.1117/12.666849

- Zadeh, Lotfi A., (1994). Fuzzy logic, neural networks, and soft computing. *Communication of ACM*, 37, 3 (March 1994), 77-84.
- Zadeh, L.H., (1965). Fuzzy sets. *Information and Control*. 8: 338-353.
- Zadeh, L., Roles of Soft Computing and Fuzzy Logic in the Conception, Design and Development of Information/Intelligent Systems, *Computational Intelligence: Soft Computing and Fuzzy-Neuro Integration with Applications*, Volume 162, 1-9.
- Zadeh, L., A. (1975), The concept of a linguistic variable and its application to approximate reasoning, *I. Inf. Sci.*, 8(3): 199-249.
- Zhang, L.S., Zhou, N. and Wu, J.X., (2008), The fuzzy integrated evaluation of embedded system security, *International Conference on Embedded Software and Systems (ICCESS 2008)*, 2931 July, Sichuan, China, 157-162.
- Zmud, R. (1997), Editors Comments, *MIS Quarterly* (21:2), June, 21-22.

Part II

Original publications

The contribution of the author to the original publications

1. Single author.
2. Main author. Built the model with the co-authors. Wrote most of the paper (section 1, 2 and 4 completely).
3. Main author. Built the model with the co-authors. Wrote most of the paper (section 1,2 and 4 completely).
4. Joint author. Wrote section 1 and section 2.

Paper 1

Buoni, A. (2010). Fraud detection: from basic techniques to a multi-agent approach, *International Conference on Management and Service Science*, 24-26 August, Wuhan, 1-4.

© 2010 International Conference on Management and Service Science.
Reprinted with the permission from IEEE.

Fraud detection: From basic techniques to a multi-agent approach

Alessandro Buoni
Institute for Advanced Management
System Research, Turku Center for
Computer Science
Joukahainengatan 3-5 B, Åbo,
Finland
abuoni@abo.fi

Abstract—According to KPMG figures, fraud represents a serious economical problem, which has been studied in different ways due to the fact that fraudsters are benefiting from the fast development of ICT and are developing their techniques. In this paper, after summarizing different fraud detection methods and tools proposed in the literature like meta-rules and tree-based detection, we will introduce a multi-agent system, called FIDES, which integrates the computational power of data mining tools and attack trees with experts' judgments negotiated through a Delphi-based system. Two scenarios are described: in the first one FIDES, supported by cause-effect diagrams, is used to classify alarms generated by the system to help the experts to focus on the real dangerous ones; in the second one FIDES is used in a proactive way in order to block or prevent human based frauds.

Keywords—*attack tree, decision support system, Delphi method, fraud detection, multi-agent systems.*

I. INTRODUCTION

In recent years, the development of IC technologies has resulted in increased cases of frauds. Although the prevention measures adopted have also progressed, fraudsters have proved to have adapting capabilities by developing new strategies. In light of what is described in the KPMG's material and statistics this paper review sets out to explain why systems based on the evaluation of experts could have better chance to detect fraudulent behavior than those ones based on search algorithm and automatic detection of transactions.

According to Basel 2 Accord [15] banks are encouraged to develop sophisticated methodologies to calculate the operational risk, monitor the bank activities, reinforce internal control structure and auditing to preserve the integrity of the managerial processes. These systems include also the use of internal and external data, scenario analysis and control factors and an accurate reporting system based on key risk indicators.

Someone may think that fraud is the results of complex operation driven by expert hackers using the state of the art of technology. In this case it would be more logical to orient detection and prevention to the development of algorithms based on sophisticated mathematical and statistical models.

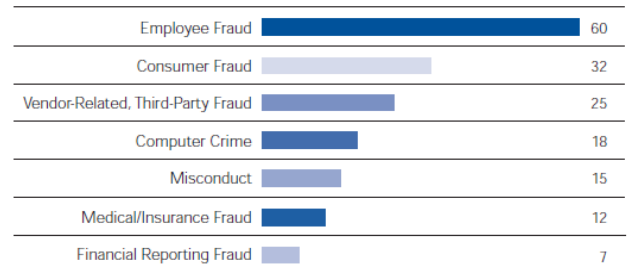


Figure 1. Types of fraud (percentage) [7].

The picture emerging from the KPMG survey [6; 7; 8] would rather suggest an approach based on the contribution of experts sharing their experience in the field and evaluating all the possible fraud strategies. One reason more to support this is that “fraudsters were mainly detected by whistle blowers or management reviews (accumulated 46 percent)” [9]. Since most of the frauds are caused by humans, as shown in Fig.1, and detected by whistle blowers, the scenario we think about is the one where a team of fraud experts (they can be senior managers or members of an audit team) through a brainstorming process can suggest how to choose the suitable fraud detection techniques and figure out their implementation in a multi-agent system (MAS) architecture.

One of the most effective approaches to manage experts' judgment is Delphi method, “an iterative process to collect and distill the anonymous judgment of experts” [13]. The choice of this approach is justified by the fact that Delphi is a good system of forcing experts to analyze a phenomenon and come up to conclusions which would be extremely hard

to reach if we had to take into account of their opinions separately.

Delphi method was used the first time in the 50s for a U.S. sponsored military project. The most important features of Delphi are anonymity and feedback. This means that it can be considered as a controlled debate. The typical Delphi process is based on questionnaires driven by a moderator. The role of the moderator is to distribute and analyze the questionnaires forcing the experts after each round to improve the accuracy and quality of their answers. The novelty of our work consists in the introduction of FIDES (Fraud Interactive Decision Expert System), a MAS based on a Delphi DSS (Decision Support System), designed to support an auditing committee in fraud detection. The MAS, described in section 3, has three main components: Data Filtering, Modeling Choice and Design and the Detection module. The output generated by Delphi is a generalized attack tree, which will be the main detection instrument, particularly useful since it is a schematic representation of experts' judgments.

II. DETECTION TOOLS

In this section we will describe the main detection tools based on meta-rules and tree-based architectures.

A. Anomaly detection tools

In this group of methods we can include those ones able to generate alarms on the basis of recognition of anomalies, unusual behavior and paths. These tools are particularly useful to manage massive amount of data and reduce quickly the searching space. Bedford's and Zipf's laws can be considered as kind of a meta-rule. These rules can be positioned as sentinels like a first wall defense of the systems, a sort of pre-filtering mechanism. Bedford's law states that "the distribution of the first significant digits of numbers drawn from a wide variety of random distributions will have (asymptotically) a certain form."

Huang et al. [4] offer an interesting approach based on Zipf's Law. The basic idea of Zipf's Law is that "the product of frequency of the use of a word, f , and the rank order, r , is approximately constant". This means that a small number of keywords can characterize a document's content. The same principle can be used with fraud detection. The assumptions the authors do is that, in a given dataset, if any patterns 'frequency doesn't follow the Zipf's law, then it might imply that there are some anomalies existing.

Zipf Analysis is very suitable for detecting attacks that have frequent sequential patterns. For future research they suggest an evaluation of the fraud detection performance of Zipf's law by comparing Zipf analysis with other clustering algorithms, like K-means, Kohonen and Digital Analysis.

Bolton and Hand [1] make an important distinction between fraud prevention and fraud detection.

Fraud prevention includes several measures used to stop fraud from occurring in the first place. Examples of these

measures can be personal identification numbers, security payment systems and SIM cards for mobile phones.

On the other hand, fraud detection involves identifying fraud as quickly as possible once it has been perpetrated and this is the main focus of this paper. Among supervised methods we mention the traditional statistical classification ones [3; 11], but also more powerful ones like neural networks [2; 12; 16].

Knowledge based systems as well have been used to classify and detect suspicious transactions. The limit of these systems is that they work very well when the behaviors cover a well structured domain, but it is well known that most of business activities are unstructured. In such situations, where we have interdependent and also noisy and incomplete data, neural networks can excel [5].

B. Tree-based detection modelling

Trees are effective tools to systematically classify the components of a fraudulent attack to any given system. The graphical tree representation is appealing to users, flexible enough to be equipped with several types of information and easy to be automated.

Amongst tree-based detection modeling, we consider attack tree the most suitable and appropriate for the purpose of this paper.

The term attack tree was introduced by Schneier [14] to systematically categorize the different ways in which a system can be attacked, and then extended in [10]. He describes an attack tree as an instrument to "provide a formal, methodical way of describing the security of systems, based on varying attacks. Basically, you represent attacks against a system in a tree structure, with the goal as the root node and different ways of achieving that goal as leaf nodes" [14].

"And" and "or" operators, in the classical framework of attack tree, relate the subtask of a node. An "and" operator is used when all the features must be present to perform a particular task and when at least one of them is necessary it will be chosen an "or" operator.

Using this approach it is possible to analyze the contribution of the any attribute to the achievement the goal that is to finalize the attack. Another important feature is the opportunity to calculate the cost of an attack as a function of the cost of subtasks. At the same time it is possible to find out the path with highest probability of success or studying the least expensive ways chosen by the fraudsters to perform an attack.

Yager [17] introduces the use of OWA nodes in describing the required number of children nodes for success of the parent nodes. The idea expressed is that the contribution of the children to the success of the parents is determined by a linguistic quantifier Q expressing the proportion (linguistically characterized) of subtasks needed for the parent node to be accomplished.

If we consider an OWA node that has n children, Q subtasks are directly responsible of the success of the parent. Q is a monotonic linguistic quantifier expressing concepts like “as most” or “at least half”.

III. FIDES (FRAUD INTERACTIVE DETECTION EXPERT SYSTEM)

This section introduces a MAS called FIDES to support an auditing committee in the fraud detection process.

The MAS will be described in its three main components (Data Filtering, Modeling Choice and Design and the Detection module) and the human computer interaction aspects of the system. In FIDES as shown in Fig.2 there are 3 components.

The first includes a set of filtering agents and it is the Data filtering module. The second one is the Model Choice and Design and the third one is the Detection module.

The role of agents, which belong to the Data filtering module, is to generate alarms. These agents may contain meta-rules as Zipf’s or Bedford’s laws or other tools like neural networks. In case of alarm, experts can join together to come up with the best strategy that is developed using the Delphi method. The output of Delphi Moderator is a document which describes all the features of the attack which can then be represented by an attack tree. Once experts have settled the tree, they can evaluate the risk in the Risk Estimation module and then decide the measures to adopt, for instance stopping a transaction or fire/sue people involved in the fraudulent operation.

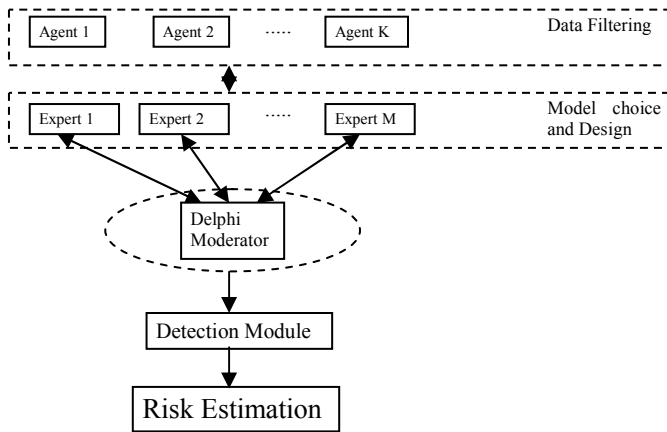


Figure 2. FIDES architecture

All the trees can be stored in a data base. In this way, experts using a case-based reasoning approach in the future will be able to recognize similar paths and strategies discovered in the past.

We introduce two different scenarios where FIDES can be used, as shown in Fig.3.

In the first scenario the situation we are figuring out is the one where agents generate alarms because they find anomalies in the data. All these alarms need to be grouped by categories in order to be understood by the moderator who is responsible of managing the Delphi process and distributing the questionnaires.

The aim of the questionnaire in this case is to reduce the search space in the first round and focus on the most dangerous operations in the following ones.

A useful instrument that can help the moderator for this purpose is the Ishikawa or fishbone diagram, which is contained in the Pruning box of Fig.3.

The first questionnaire can be used to fill the diagram with questions that will allow filtering and removing non-dangerous operations.

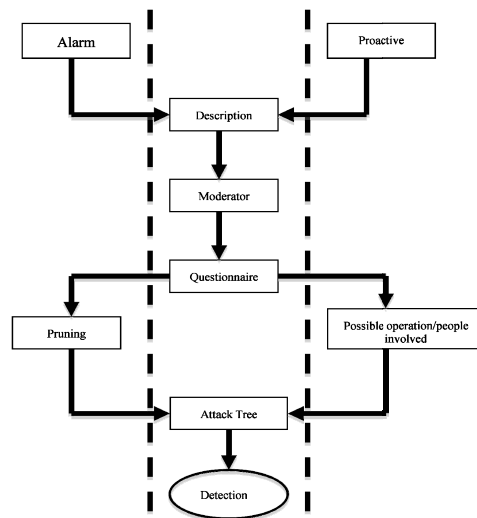


Figure 3. Two different scenarios of FIDES.

Questions can be grouped in relation to specific categories in order to facilitate the pruning. Categories can be chosen through a brainstorming process driven by the moderator, following the principles of Delphi. The output of the diagram is a list of significant fraud cases, which can be formalized then with one or more attack trees.

In the second scenario the alarm is not generated by the agents, but the system is used in a proactive way. Experts, for instance, could connect some unusual behaviors, small bureaucratic irregularities, customers complaints related to a manager or an employee matching his/her behavior with his/her psychological profile and investigate more on him/her, preventing the worse or discovering something previously unknown.

The moderator, in this case, has to force the experts to guess possible schemes based on their suspects, rumors, information gathered, or confessions made by other people in anonymity about suspect behaviors of other colleagues. The output of Delphi is again an attack tree which is a

representation of the path followed by the fraudster or the possible consequence (his/her goal) of certain suspect behaviors which could hide an operation in progress or his/her intention to carry it out.

In this sense this is the proactive scenario. In the first one experts has to filter and interpret the alarms, in the second one using abductive reasoning they try to prevent or block an operation which can't be seen through the data, because it is human driven.

Experts can meet physically or virtually using a decision support system tool that might be installed for instance on their mobile phone. This would be considered a crisis management scenario, where experts can interact and take decisions quickly. The system can be used by the members of a bank committee or different authorities involved in a crisis management context as police, firefighters and medical personnel that needs to coordinate complex tasks in emergency situation as a computer hacking attack, a terroristic attack, a natural disaster or a multiple car accident. The infrastructure of the system will be based on the Internet Protocol (IP), in order to use the state of the art of mobile technology (GSM and GPRS, 3G, WLAN) and perform different tasks. The system will follow an observe-decide-act cycle which means that each action can trigger a new one, creating a chain of events which can be visualized in order to understand all the links and focus on the main targets. The system can work as a distributed approach interconnecting different actors, but this doesn't prevent from having a central command station able to communicate and send orders to the different actors.

IV. CONCLUSIONS AND FUTURE WORK

Several methods have been applied in fraud detection. Anomaly detection tools showed their power to generate alarms in relation to past behaviors, but they can't do so much with respect to new behaviors or frauds perpetrated by humans, which are the majority according to KPMG's figures. In this case systems based on experts' judgments can be more effective. A MAS called FIDES is introduced to integrate the computational power of OR and AI tools with the experts' experience.

The management system used to extract expert's tacit knowledge is the Delphi method. The output of Delphi is an attack tree that can be used to detect illegal operations after they have been perpetrated or stored in a library in a proactive way.

On the basis of this MAS future research will be directed towards the development of FIDES, in particular in adapting the system to a bank auditing committee's features and needs. Typical knowledge management issues as group consensus achievement, semantic problems, GDSS (group

decision support systems), risk management, group emotional intelligence, information retrieval and the definition of linguistic variables using fuzzy logic, can all be future research topics.

A mobile phone application for crisis management situations could also be designed. This would be the Mobile Module of FIDES. The scenario we think about is the one where experts have to take a quick decision under an imminent attack or when a suspect operation is happening and they can't physically meet in the office or they can't have access to their pc.

REFERENCES

- [1] J.R. Bolton and D.J. Hand, *Statistical Fraud Detection: A Review*, *Statistical Science*, vol.17, No.3, 235-255, 2002.
- [2] D.J. Hand, *Construction and Assessment of Classification Rules*. Chichester: Wiley, 1997.
- [3] D.J. Hand, *Discrimination and Classification*. Chichester: Wiley, 1981.
- [4] Shi-Ming Huang, David C.Yen, Luen-Wei Yang, Jing-Shiuan Hua, *An investigation of Zip's law for fraud detection*, *Decision support system* 46, 70-83, 2008.
- [5] I. Jagielska and J. Jaworski, *Neural Network for Predicting the Performance of Credit Card Accounts*, *Computational Economics* 9: 77-82, 1996.
- [6] Kpmg forensic, *Fraud Risk Management Developing a Strategy for Prevention, Detection and Response*.
- [7] Kpmg forensic, *Fraud Survey 2003*.
- [8] Kpmg forensic, *Integrity Survey 2005-2006*.
- [9] Kpmg forensic, *Profile of Fraudster Survey 2007*.
- [10] S. Mauw and M. Oostdijk, *Foundation of Attack Trees*, *Information security and cryptology : ICISC 2005*, 8th international conference, Seoul, Korea, 186-198, December 1-2, 2005.
- [11] G.J. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition*. New York: John Wiley and Sons, 1992.
- [12] B. D. Ripley, *Pattern recognition and neural networks*. Cambridge University Press, 1996
- [13] G. Rowe , G. Wright, *The Delphi technique as a forecasting tool: issues and analysis*, *International Journal of Forecasting* 15, 353-375, 1999.
- [14] B. Schneier, *Attack Trees: Modeling security threats*, *Dr. Dobb's Journal* 1-9, December 1999.
- [15] *The New Basel Capital Accord* , April 2003.
- [16] A.R.Webb, *Statistical Pattern Recognition*, London: Arnold, 1999.
- [17] R. R. Yager, *OWA trees and their role in security modeling using attack trees*, *Information Sciences* 176, 2933-2959, 2006.

Paper 2

Buoni, A., Fedrizzi, M., and Mezei, J. (2010). A Delphi-Based Approach to Fraud Detection Using Attack Trees and Fuzzy Numbers, In *Proceeding of the IASK International Conferences*, 21-28.

© 2010 Reprinted with the permission from IASK - International Association for the Scientific Knowledge.

A Delphi-based approach to fraud detection using attack trees and fuzzy numbers

Alessandro Buoni, Mario Fedrizzi, Jozsef Mezei

Abstract — . The fraud surveys carried out in the last five years by leading international consulting companies demonstrate that fraud is an increasing phenomenon depending most of all on behavioral aspects. Therefore, when addressing fraud detection processes the adoption of traditional statistical techniques comes out to be not as adequate as those based on the evaluations of experts working in a multiagent framework. In this paper we introduce a multiagent system called Fraud Interactive Decision Expert System (FIDES), which puts more emphasis on the evaluation of behavioral aspects of fraud detection according to the judgments expressed by two groups of experts, inspectors and auditors respectively. FIDES combines think-maps, attack trees and fuzzy numbers under a Delphi-based team work support system and offers to the users a suitable way to better understand and manage fraud schemes.

Index Terms — Fraud detection, Think-maps, Attack trees, Delphi method, Fuzzy numbers

1 INTRODUCTION

David J. Hand [14] points out how institutions in persecuting fraud follow the economic imperative, meaning it doesn't worth spending \$200m to stop \$20m fraud. Participants in his study estimate that U.S. organizations lose 5% of their annual revenues to fraud. This means, applied to 2006 U.S. GDP, approximately \$652 billion in fraud losses. According to [16], there is a prominent increase of fraud by individuals. Company managers, employees and customers together have been responsible for £300m of fraud in 2008, three times the value of year 2007.

In [15] it is shown that the most effective countermeasures for fraud are those ones developed by internal audit using clues given by employee whistleblowers as shown in Fig.1. The survey has been conducted on executives of U.S companies who answered the following question: Through which source do you believe your organization would be most likely to uncover fraud or misconduct?

The embarrassment of admitting to mainly

follow an economic imperative in persecuting fraud is coherent with the choice of preferring internal resources on external ones, but this is not only the main reason. It is also related to the awareness that an internal audit team has a better knowledge of their organization, the weakness of internal procedures and the personnel.

	%
Internal Audit	47
Employee whistleblowers	20
Line managers	13
External Auditors	9
Customers or suppliers	4
Government regulators or law enforcement	3
Other means	2

Fig.1 Fraud countermeasures [15].

According to the Advanced Measurement Approaches (AMA), introduced in Basel II accord [2], banks are encouraged to develop sophisticated methodologies to calculate the operational risk, monitor the bank activities, reinforce internal control structure and auditing to preserve the integrity of the managerial processes. These systems include also the use of internal and external data, scenario analysis and control factors and an accurate reporting system based on key risk indicators.

In the banking sector there is a prevalence

- *Alessandro Buoni, IAMSR, Turku Centre for Computer Science, Joukahaisenkatu 3-5 B, 20520 Turku, Finland, email: abuoni@abo.fi.*
- *Mario Fedrizzi Department of Computer and Management Science, University of Trento, Via Inama 5, 38122 Trento, Italy, email: mario.fedrizzi@unitn.it.*
- *Jozsef Mezei, IAMSR, Turku Centre for Computer Science Joukahaisenkatu 3-5 B 20520 Turku, Finland, email: jmezei@abo.fi.*

of human fraud. One of the causes encouraging internal fraud is conflict of interests, which limits the effectiveness of the control procedures.

Frauds are perpetrated not only for greed but also for career, as a way that managers use to augment their reputation, for instance faking their achievements and showing the improvements in their life style (cars, watches). Many frauds are also the results of the abuse of trust given by their customers. These people typically establish a friendly relationship with their investment consultant entrusting them their money blindly.

Commonly, the biggest problem of an audit team is to interpret and summarize all these behavioral aspects in order to come up with effective solutions to prevent fraud before they happen or to detect quickly the type of fraud when perpetrated. To this end, audit teams collect information about past behaviors in order to provide a formal representation of the most common typologies of fraud. This way, a repository of domain expert represented by standardized fraudulent attacks can be created and reused for, e.g., playing "what-if" games with potential countermeasures or identifying the nature of new attacks.

Since before addressing the design of our system we had the chance to meet several experts in charge of diverse risk management activities inside one of the largest European banking group, our work has been inspired by the information collected during the meetings. One of the main challenges for banks is to unify risk criteria in the different countries where their branches are located, perform a realistic temporal analysis and establish cause/effect relationship in a rather short time combining objective information with the subjective judgments expressed by experts.

Auditors, who are the fraud experts, can use their experience to remove false alarms, but also to detect those crimes which cannot be detected electronically because they are the results of untraceable human behaviors most often perpetrated inside the departments of the bank. The evaluation process of auditors in fraud detection has been examined, e. g., by Bazerman et al. [1] and by Grazioli et al. [12], exploiting the reasons of their success and failure, and studying the impact of ambiguity, analyzing a quite extended sample of case studies.

Several authors have demonstrated that a multiagent approach is particularly suitable to address fraud detection when behavioral aspects play a key role, see for instance Chou et al. [5], Wang et al. [24], and Zhang et al. [25].

We believe that the multiagent system we are going to introduce in this paper, combining think-maps, attack trees, and fuzzy numbers under a Delphi-based team work support system, do offer to the agents involved (inspectors and auditors) an innovative and suitable way to better understand and manage fraud schemes.

The multiagent architecture of the Fraud Interactive Decision Expert System (FIDES) [3] will permit them to open up and share their knowledge and then link all the clues in a coherent scheme. The learning by doing approach of think-maps is a good means for inspectors to formally represent their knowledge linguistically expressed, to reconsider their opinions and correct their statements in the light of the comments received by their colleagues.

The paper has been organized as follows. In the second section we introduce the main components of FIDES, i.e. think-maps, attack trees and Delphi method. In the third section we describe the fuzzy mechanism that is used for representing and aggregating the linguistic information provided by the experts of the audit team when addressing the design of the attack tree. In the fourth section the whole structure of the system is figured out combined with the description of a work session. In the last paragraph some conclusions are drawn and future line of research are sketched out.

2 THINK-MAPS, ATTACK TREES, DELPHI METHOD

Scope of this paragraph is to provide a synthetic description of the main components of FIDES, focusing the description on those features characterizing the collection and representation of knowledge involved in the analysis and detection of fraudsters' attacks.

2.1 Think-maps

The concept of mind mapping was introduced for first by Tony Buzan in 1974 [4]. The idea was to organize keywords into a radiant structure that looks like a tree seen from above [26]. A radiant structure permits to put in the middle of the paper the central idea (goal) and through a brainstorming process add around it the branches of the map. Novak and Cañas developed the idea of concept maps (see [19]), which they defined as "graphical tools for organizing and representing knowledge".

Oxman introduces the idea of think-maps

([20]), which means that conceptual mapping of designing ideas can be constructed in larger structures where it is possible to organized knowledge acquired by the learner and make it explicit. The software they develop to create these maps is called "Web-Pad".

The theoretical basis of think-maps is constructivism. Constructivist theories of learning state that the learner in not a passive recipient of knowledge, but it has an active role in creating knowledge, based on the "learning by doing" approach. Learners construct their knowledge based on their experience and relationship with concepts. Think-maps have a formal representation called ICF (Issue-Concept-Form), which acts as "a structuring ontology for the construction of conceptual networks of design concepts"([20]). To show the process we will adapt a case of study ([18]), under the hypothesis that this operation could be anticipated by suspicious behaviours, which are common to other internal fraud cases ([6]).

In our case inspectors will use Web-Pad (Fig.3) to comment suspicious behaviours and/or activities they observed or information acquired by whistleblowers or insert their reports in order to share them with all the other colleagues.

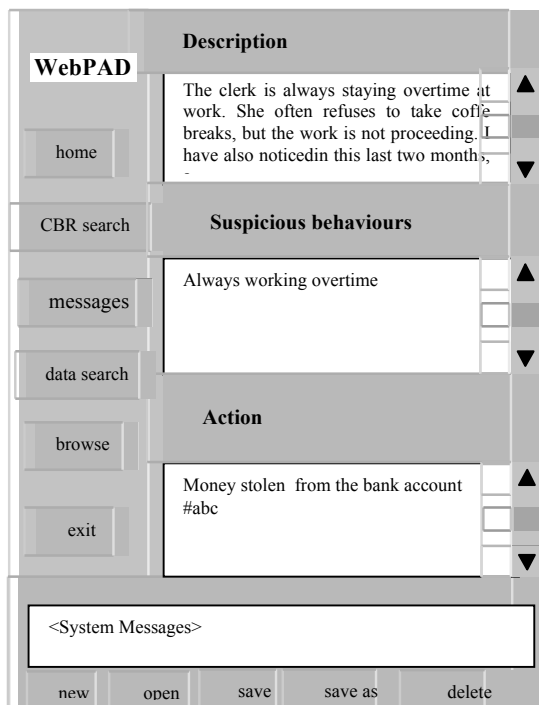


Figure 2. Web-Pad interface ([20]).

Since the software works in a real time discussion environment their comments will appear in the text box named "Description" as

a common chat discussion shape. At this stage inspectors will have the possibility to arrange the text in order to have a clear description of the case.

Inspectors at this point will start to underline keywords, as shown in Fig.3, which will be associated to 2 main labels "suspicious behaviours" and "action", which are related to one or more suspected.

The clerk is always staying overtime at work. She often refuses to take coffee breaks, but the work is not proceeding. I have also noticed in this two months, a big change in her lifestyle. She is dressing new clothes and wearing an expansive bag. She also changed her car.

Figure 3: The keywords selection.

At the end of this process the think-map shown in Fig.4 is obtained. In the map we can see the main labels, which have been used to visualize a fraud operation. In the cloud callouts are represented the activities related to "suspicious behaviour", label previously created with Web-Pad, associated to three different suspected persons. There is a clerk, or a hacker or a customer who has the habit to play with the self-service payment pc of the bank. In the middle of the figure we have the suspected and on the bottom the fraudulent action, which is the amount of money missing in a specific bank account.

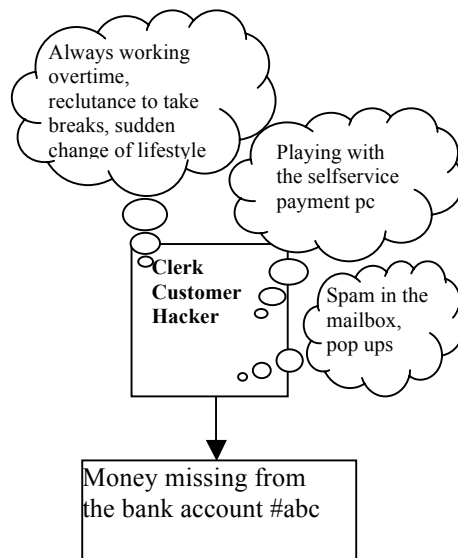


Figure 3: The keywords selection.

Having this think-map as a descriptive model, inspectors will start to create nodes, which will be sent to the auditors to be connected through the Delphi process to build

the attack tree. Web-Pad basically works on two levels. A graphic level where think-maps can be visualized and a text level where different operations can be stored and retrieved for the future, supporting the inspectors when they have to build the think-map. For instance inspectors could be interested to have a list of the cases associated to a particular behaviour and so on. In the data retrieval mode the system can bring up precedents that inspectors consider similar, according to their subjective judgement. Users can express the level of similarity between two different cases as a number between 0 and 1. Once the database is populated with a significant number of cases, it will be possible to retrieve and visualize them in descendant order, according to their level of similarity, ready to be examined.

2.2 Attack tree

The attack tree, introduced by Schneier ([23]) is a tree-based diagram to "systematically categorize the different ways in which a system can be attacked".

Nodes are the elementary attacks and the root node is the goal of the attack. Children of a node are refinements of this goal, while leaves are attacks which cannot be further on refined. The process of creating an attack tree starts identifying the possible attack goals, where each goal forms a separate tree, but there is also the possibility that different trees share nodes and subtrees.

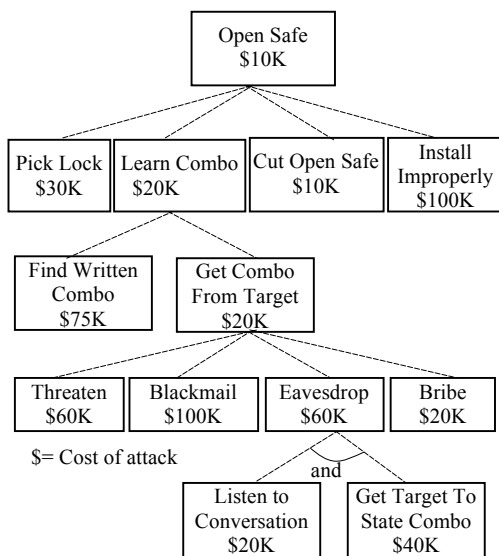


Figure 5. An example of attack tree ([23]).

Accordingly, modeling an attack tree is a matter of associating a logical AND and a

logical OR with each node, and therefore encouraging a structured representation of events and of the ways they are connected.

This supports the discovering of the most likely avenues of approach for an attack making easier the deployment of the most effective countermeasures.

Even though the Schneier's attack trees (illustrated in Fig.5) have been considered from their first appearance a convenient tool to systematically categorize the different modes in which an attack can be carried out, nonetheless their network structured has been criticized for its simplicity e for the lack of well sounded theoretical foundations.

Mauw and Oostdijk ([17]) arguing that Schneier's approach to attack trees is semantically not well sounded, provide a generalization based on the observation that an attack tree describes an attack suite and that a node can be connected to a multi-set of nodes (bundle) and may contain several bundles. But, since our paper is more focused on the way on which the team of experts carry out, in a Delphi-based context, the consensual construction of the tree, than on the complexity of its structure, the nature of the tree is irrelevant and therefore we will adopt the Schneier's approach.

2.3 Delphi method

Delphi method, introduced for the first time in the 50s for a U.S. sponsored military project ([13]), is a systematic, interactive and iterative method which relies on a team of experts, aiming at discussing and structuring the solution of a given problem. The experts are asked to answer questionnaires in two or more rounds, and after each round a moderator provides an anonymous summary of the experts' analysis from the previous round as well as some explanations of their judgments. The moderator encourages experts to reuse their earlier opinions in light of the outcomes of the analysis provided by the other experts of the team. The process is stopped according to a pre-defined criterion and some average measure of the outcomes of the final round determine the output of the process.

Rowe and Right suggest four key features for a good design of Delphi [21]:

1. Anonymity of Delphi experts: allows the experts to freely express their opinions without undue social pressures to conform from others in the team. Decisions are evaluated on their merit, rather than who has proposed the idea.
2. Iteration: allows the experts to refine

their views in light of the progress of the team's work from round to round.

3. Controlled feedback: informs the experts of the other experts' perspectives, and provides the opportunity for Delphi experts to clarify or change their views.

4. Statistical aggregation of team response: allows for a quantitative analysis and interpretation of output information.

In our system, Delphi will be used as a method to connect different nodes delivered by the inspectors. The role of the moderator will be to ask the experts their opinion about strength of the links connecting different nodes.

The aggregation process will be described in details in section 3. In the end of the Delphi process, the output will be an attack tree.

3 THE FUZZY MECHANISM

The audit team performs the Delphi process aiming to select the nodes and connect them in order to design the attack tree. In the first phase the inspectors determine the possible nodes of the attack tree with the help of the think-map. Then the moderator will ask the experts about the possible connection of the nodes, and aggregate the results to obtain the attack tree. In this paragraph will describe the fuzzy mechanism which helps the experts to form the attack tree. In the literature, a number of different fuzzy approaches to the analysis of negotiation processes in multiagent decision making have been proposed, and for an extended overview the interested reader could see, e.g., [9] and [11]. Before the description of the model we need some basic definitions from fuzzy set theory.

Definition 1. Let $X = x$ denote a collection of objects (points) denoted generically by x . Then a fuzzy set A in X is a set of ordered pairs

$$A = (x, \mu_A(x)), x \in X \quad (1)$$

where $\mu_A(x)$ is termed the grade of membership of x in A , and $\mu_A : X \rightarrow M$ is a function from X to a space M called the membership space. When M contains only two points, 0 and 1, A is non fuzzy and its membership function becomes identical with the characteristic function of a crisp set. This means that crisp sets belong to fuzzy sets. A fuzzy number is a convex fuzzy set on the real line such that

1. $\exists x_0 \in A, \mu_A(x_0) = 1,$
2. μ_A is piecewise continuous.

(The convexity means that all the γ -level sets are convex. Furthermore, we call F the family of all fuzzy numbers).

A γ -level set of a fuzzy set A is defined by $[A]^\gamma = \{x \in A: \mu_A(x) \geq \gamma\}$ if $\gamma > 0$ and $[A]^\gamma = cl\{x \in A: \mu_A(x) > \gamma\}$ (the closure of the support of A) if $\gamma = 0$. Let A be a fuzzy number. Then $[A]^\gamma$ is a closed convex subset of R for all $\gamma \in [0,1]$. We use the notations

$$a_1(\gamma) = \min[A]^\gamma, a_2(\gamma) = \max[A]^\gamma$$

for the left-hand side and right-hand side of the γ -cut, respectively.

When we calculate the arithmetic operations on fuzzy sets (fuzzy numbers), we apply the rules of interval arithmetic. Let A and B be fuzzy numbers with the corresponding γ -cuts: $[A]^\gamma = [a_1(\gamma), a_2(\gamma)]$, $[B]^\gamma = [b_1(\gamma), b_2(\gamma)]$ then the γ -cut of the fuzzy number $A + B$ is the following:

$$[A + B]^\gamma = [a_1(\gamma) + b_1(\gamma), a_2(\gamma) + b_2(\gamma)]$$

and the γ -cut of the fuzzy number αA , where $\alpha > 0$:

$$[\alpha A]^\gamma = [\alpha a_1(\gamma), \alpha a_2(\gamma)]$$

We will use linguistic labels in the questionnaire and we represent the labels as fuzzy numbers. We suppose that the moderator can choose which nodes are parents (V) (with descendant) and which ones are leaves (L) (without descendants, basic attack components). We obtain two sets:

$$L = \{l_1, \dots, l_s\}, V = \{v_1, \dots, v_t\}.$$

In the first questionnaire the experts have to express their opinion in linguistic terms about statements like " $l_i \in L$ is required for $v_j \in V$ ", for every $i = 1, \dots, s, j = 1, \dots, t$.

Then experts $E = \{e_1, \dots, e_N\}$ are asked to determine their level of agreement on the statements based on a linguistic scale with m terms for every pair (l_i, v_j) . The linguistic terms in the model are represented as fuzzy numbers. In other words we have a mapping

$$\Phi : T \rightarrow F$$

from the set of linguistic terms into the family of fuzzy numbers.

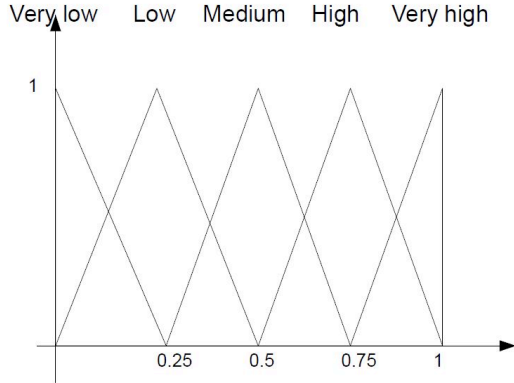


Fig.6. Possible representation with triangular fuzzy numbers

Example 1. One possible representation for a linguistic label is a triangular fuzzy number:

$$A(u) = \begin{cases} 1 - \frac{a-u}{\alpha}, & a - \alpha \leq u \leq a \\ 1 - \frac{u-a}{\beta} & a \leq u \leq a + \beta \\ 0 & \text{otherwise.} \end{cases}$$

From the opinion of the experts we obtain the frequencies of the different classes. For the pair (l_i, v_j) we have $n_1^{ij}, \dots, n_m^{ij}$. If we denote by A_1, \dots, A_m the fuzzy numbers corresponding to the linguistic labels, we can define a new fuzzy number A_{ij} as a "weighted average", with level sets:

$$[A_{ij}]^\gamma = \left[\frac{1}{N} \sum_{k=1}^m n_k^{ij} a_1^k(\gamma), \frac{1}{N} \sum_{k=1}^m n_k^{ij} a_2^k(\gamma), \right]$$

where $[a_1^k, a_2^k]$ is the level set of A_k . This is clearly a fuzzy number with the support in the interval $[0,1]$.

To obtain the connection degree for the pair (l_i, v_j) , we calculate the f -weighted possibilistic mean value of A_{ij} , defined in [10].

Definition 2. The f -weighted possibilistic mean value of $A \in F$, with γ -level sets $[A]^\gamma = [a_1(\gamma), a_2(\gamma)]$, $\gamma \in [0,1]$, is defined by:

$$E_f(A) = \int_0^1 M(U_\gamma) f(\gamma) d\gamma = \int_0^1 \frac{a_1(\gamma) + a_2(\gamma)}{2} f(\gamma) d\gamma \tag{2}$$

where U_γ is a uniform probability distribution on $[A]^\gamma$ for all $\gamma \in [0,1]$.

After we have obtained these defuzzified numbers as the estimation of the connection strengths, we can determine for every attack component the ranking of the other nodes, then we can construct the adjacency matrix of the attack tree by connecting the leaves to the best ranked vertices.

Example 2. In the simplest case we can represent the linguistic labels as a fuzzy set with the membership function

$$A(u) = \begin{cases} 1 & u = c \\ 0 & \text{otherwise} \end{cases}$$

If we have 5 categories, we use the set $\{0, 0.25, 0.5, 0.75, 1\}$. The weights of the outcomes are the frequencies of the linguistic labels. If we observe the weights $n_0, n_{0.25}, n_{0.5}, n_{0.75}, n_1$, then A_{ij} is just the characteristic function of the value

$$\frac{1}{\sum_{j=0}^4 n_{0.25^j}} \sum_{i=0}^4 n_{0.25^i} * 0.25^i,$$

what is simply the sample mean value of our data. And according to the used defuzzification method, the obtained connection estimation is this sample mean.

4 THE ARCHITECTURE OF FIDES

In this section we will describe the architecture of FIDES (Fig.7), showing how the main components are interrelated and the role played by the inspectors and auditors, the experts which have to decode the alarms generated by software agents and to detect and describe the suspicious human behaviors (see, e.g. Sanchez et al. [22] and Edge et al. [7]).

FIDES indeed, has been built on the base of the suggestions we collected interviewing the

managers of risk management department of a leading European bank and thus the multiagent system has been designed according to their opinions. In a fraud detection process an audit team have to deal with numerical data and alarms, produced by software agents, but also documents, reports and information gathered by different actors (managers, whistleblowers, anonymous informers) and then summarized by inspectors. Therefore, the key factor in detecting fraud behaviors is to improve the interaction between inspectors and auditors.

Alarms, generated by software agents and/or suspicious behaviors noticed personally by inspectors are filtered using the Web-Pad software. The Information Filtering Module is nothing else that a preliminary session where inspectors can decide which alarms and behaviors to take into considerations to perform all the process. Once the case has been well detailed as shown in Fig.2, inspectors start to underline the keywords in order to create the think-map as described in paragraph 2.

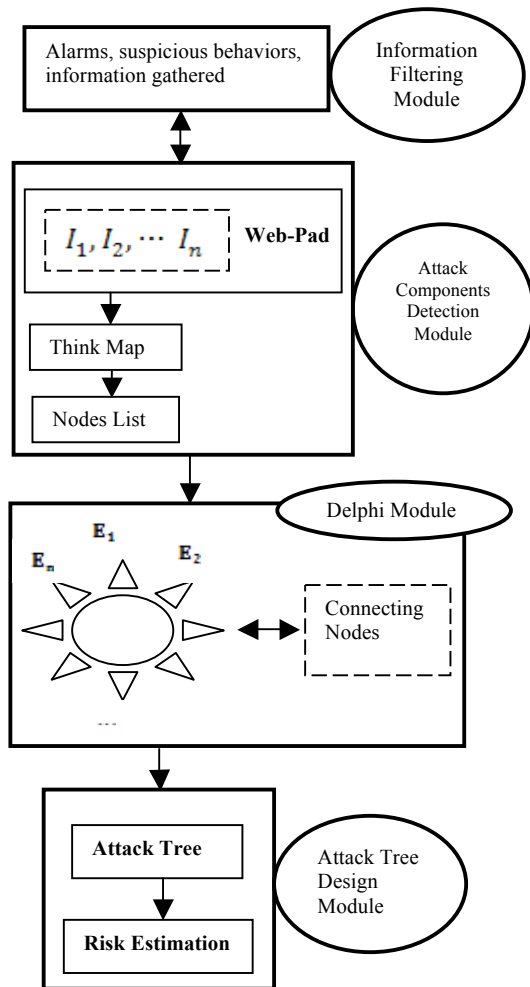


Figure 7: The architecture of the system.

This is the Attack Components Detection Module whose output is the creation of nodes to be sent to the auditors.

Since think-map represents a hypothesis of correspondence between behavior, suspected fraudsters and fraudulent actions, inspectors create nodes to try to offer an explanation of how this conceptual framework can be associated to concrete actions performed by the fraudster.

Nodes are elementary attacks expressed by labels which define all the possible steps of the fraudulent action.

At this stage auditors activate the Delphi Module, driven by the moderator, as explained in paragraph 3, to the end of connecting the nodes to form the attack tree, taking into account the strength of the links between nodes. In the Attack Tree Design Module, experts can estimate the risk and develop the strategy to persecute the fraudster. The estimation process can be performed calculating the most probable or the least expensive path.

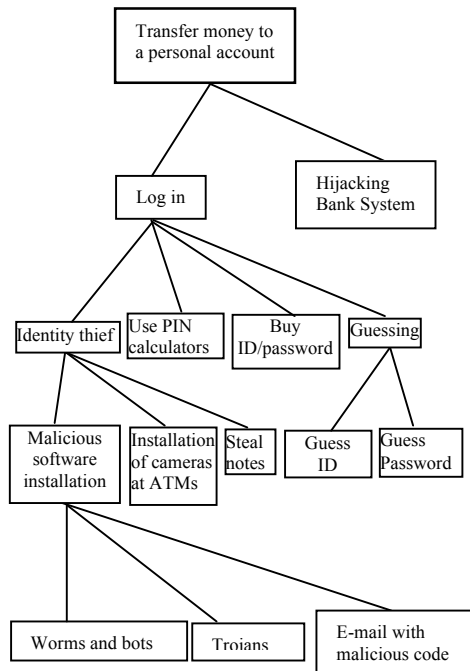


Figure 8: Final attack tree.

Continuing with the example described in paragraph 2.1, in Fig.8 is shown the final attack tree as result of Delphi process where auditors connect the nodes delivered by the inspectors. In this example the goal of the attacker is to transfer money to his/her personal account. The tree shows different ways he/she can achieve this goal. The easiest but the least successful path is to guess ID and password. The most difficult one

is to hijack the bank system. The other paths show identity thief sub-attacks based on malicious software installation or the use of cameras installed at ATMs.

5 CONCLUSION

The most of the fraud surveys carried out in the last five years by leading international consulting companies have shown that behavioral aspects play a central role, and therefore the biggest problem to be addressed by inspectors and auditors is to interpret the signals and summarize the information coming from several, sometimes conflicting, sources. The successful solution of this kind of problems depends to a large extent by a proper combination of critical analysis, knowledge-based actions, whistleblowers' messages interpretation, involving groups of interacting experts. This paper describes the FIDES system intended for the management of fraud detection situations where inspectors and auditors are collaborating at distance according to a two phase detection process. Think-map and Delphi method can offer to the inspectors and auditors respectively an effective environment to explicit their knowledge, select the most likely fraud attack components, and finally structuring them in an attack tree. Attack tree-based representation of fraudulent processes has the advantage of offering a clear visualization of the attack which permits to the auditors to highlight the most probable attack paths.

Future work will be focused on two aspects. Firstly, after introducing the cost or impact of an attack starting from a set of attributes attached to the nodes of the attack tree, the consensus reaching process of the group of audit experts will be studied taking care as well of the representation of the uncertainty involved. Secondly, assuming that the formally expressed representation of the information related to the attacks (attack trees) is stored in references sources (catalogue of attack trees), an intelligent searching module will be introduced in FIDES for retrieving information from the catalogue. To wit, our aim is to introduce an ontology-based description of the attack trees to provide a formal basis for sharing knowledge and for reusing it during the attack tree design process.

REFERENCES

- [1] M. Bazerman, G. Lowenstein and D. Moore, "Why good accountants do bad audits," *Harvard Business Review*, 3-8, 2007.
- [2] Basel II, Basel Committee on Banking Supervision, 2006.
- [3] A. Buoni, Fraud detection: From basic techniques to a multi-agent approach. 2010 International Conference on Management and Service Science, Wuhan, August 24-26, 2010 (accepted).
- [4] T. Buzan, Use your head. London: BBC Books, 1974.
- [5] C.L-Y Chou, T. Du, S. V. Lai., "Continuous auditing with a multi-agent system," *Decision Support Systems*, 42 (4) 2274-2292, 2007.
- [6] CIFAS, "Fraud indicators - some warning signs". http://www.cifas.org.uk/default.asp?edit_id=57957, 2003.
- [7] M.E. Edge, R.P.F. Sampaio, Choudhary M., "Towards a proactive fraud management framework for financial data streams," *Proceedings, of the Third IEEE International Symposium on Dependable, Autonomic and Secure Computing*, 55-64, 2007.
- [8] D. Dubois, and H. Prade, "Operations on fuzzy numbers," *International Journal of Systems Science*, 9 (6), 613-626, 1978.
- [9] C.Carlsson, M Fedrizzi, R. Fuller, Fuzzy logic in management, Dordrecht ; Boston ; New York, Kluwer Academic, 2004.
- [10] C. Carlsson, and R. Fuller, "On possibilistic mean value and variance of fuzzy numbers," *Fuzzy Sets and Systems*, 122, 315-326, 2001
- [11] M.Fedrizzi, Kacprzyk J., Nurmi H., Consensus under fuzziness, Dordrecht ; Boston ; New York, Kluwer Academic, 1997.
- [12] S. Grazioli, P. E. Johnson and J. Karim, "A cognitive approach to fraud detection".<http://ssrn.com/abstract=920222>, 2006.
- [13] T. J. Gordon, "The Delphi method, in futures research methodology," *AC/UNU Millenium Project*, Washington, AC/UNU, 1994.
- [14] D. J. Hand, "Statistical techniques for fraud detection and evaluation." http://langtech.jrc.ec.europa.eu/mmdss2007/htdocs/Presentations/Docs/MMDSS_Hand_PUBLIC.pdf. 2007.
- [15] KPMG Fraud survey, 2009.
- [16] KPMG Forensic fraud barometer, 2009.
- [17] S. Mauw and M. Oostdijk, "Foundation of attack trees," *Information security and cryptology: ICISC 2005, 8th international conference. Seoul, Korea*, 186-198, 2005.
- [18] NHS Business Service Authority, "Case study-employee fraud". http://www.forensic-computing.nhs.uk/resources/ghost_employees.pdf. 2003.
- [19] J Novak and A. Cañas, "The theory underlying concept maps and how to construct them," Technical Report IHMC Cmaptools, 2006.
- [20] R. Oxman, "Think-maps: teaching design thinking in design education," *Design Studies*. Vol. 25, Number 1, 2004.
- [21] G. Rowe, and Wright, G., "The Delphi technique as a forecasting tool: issues and analysis," *International Journal of Forecasting*, 15, 353-375, 1999.
- [22] D. Sanchez, M.A Vila, L. Cerda, J.M. Serrano, "Association rules applied to credit card fraud detection," *Expert Systems with Applications*, 36, 3630-3640, 2009.
- [23] B. Schneier, "Attack trees". *Dr. Dobb's Journal December 1999*, available at <http://www.schneier.com/paper-attacktrees-djft.html>.

- [24] D.G. Wang, T. Li, S.J.L Liu, G. Liang, and K. Zhao, "An immune multiagent system for network intrusion," *In ISICA 2008, LNCS 5370*, L. Kang et al. (Eds.), Springer-Verlag Berlin Heidelberg, 436-445, 2008.
- [25] L.S. Zhang, N. Zhou and Wu J.X., "The fuzzy integrated evaluation of embedded system security," *International Conference on Embedded Software and Systems*, 157-162, 2008.
- [26] M. Åhlberg, "History of graphic tools presenting concepts and propositions", <http://www.reflectingeducation.net/index.php?journal=reflecting&page=article&op=downloadSuppFile&path%5B%5D=49&path%5B%5D=6>, 2007.

A.Buoni is a Phd student at IAMSR, Åbo Akademi University, Finland. His research interests are in fraud detection, multiagent systems and knowledge management.

M.Fedrizzi is a Professor at Department of Computer and Management Sciences, University of Trento, Italy. His research interests are in decision modelling under uncertainty, fuzzy logic and soft computing.

J.Mezei is a Phd student at IAMSR, Åbo Akademi University, Finland. His research interests are in fuzzy logic and optimization.

Paper 3

Buoni, A., Fedrizzi, M., and Mezei, J. (2011). Combining attack trees and fuzzy numbers in a multi-agent approach to fraud detection, *Int. J. Electronic Business*, Vol. 9, No. 3, 186-202.

© 2011 Reprinted with the permission from Inderscience Publishers.

Combining attack trees and fuzzy numbers in a multi-agent approach to fraud detection

Alessandro Buoni*

IAMSR,
Turku Centre for Computer Science,
Joukahaisenkatu 3-5 B, 20520 Turku, Finland
E-mail: abuoni@abo.fi
*Corresponding author

Mario Fedrizzi

Department of Computer and Management Sciences,
University of Trento,
Via Inama 5, 38122 Trento, Italy
E-mail: mario.fedrizzi@unitn.it

József Mezei

IAMSR,
Turku Centre for Computer Science,
Joukahaisenkatu 3-5 B, 20520 Turku, Finland
E-mail: jmezei@abo.fi

Abstract: The fraud surveys carried out in the last five years by leading international consulting companies demonstrate that fraud is an increasing phenomenon depending most of all on behavioural aspects. In this paper, we introduce a multi-agent system called Fraud Interactive Decision Expert System (FIDES), which puts more emphasis on the evaluation of behavioural aspects of fraud detection according to the judgements expressed by experts (inspectors and auditors). FIDES combines think-maps, attack trees and fuzzy numbers under a Delphi-based team work support system and offers to the users a suitable way to better understand and manage fraud schemes.

Keywords: fraud detection; think-maps; attack trees; Delphi method; fuzzy numbers.

Reference to this paper should be made as follows: Buoni, A., Fedrizzi, M. and Mezei, J. (2011) 'Combining attack trees and fuzzy numbers in a multi-agent approach to fraud detection', *Int. J. Electronic Business*, Vol.

Biographical notes: Alessandro Buoni received his Master degree in 2005 from the Faculty of Economics of University of Trento. He is presently is a Doctoral student at the Institute of Advance Management System and Research (IAMSR) of Åbo Akademi University, Finland. His research interests are in fraud detection, multi-agent systems and knowledge management.

Mario Fedrizzi has been working since 1986 as Full Professor of Mathematical Methods for Economics at the Department of Computer and Management Sciences, University of Trento, Italy. His fields of interest are utility and risk theory, group decision making, fuzzy decision analysis, and soft computing. He has authored more than 150 papers, appeared in international proceedings and peer reviewed international journals. He is member of the editorial boards of *International Journal of Uncertainty Fuzziness and Knowledge-Based Systems*, *International Journal of General Systems*, *Group Decisions and Negotiation*, and of two other international scientific journals.

József Mezei received his Master Degree in 2008 from the Faculty of Mathematics of the Eötvös Loránd University, Hungary. He is presently a Doctoral student at the Turku Centre for Computer Science and a Research Fellow for the Institute of Advanced Management Systems Research of Åbo Akademi University, Finland. He has authored more than 20 papers, appeared in international proceedings and peer reviewed international journals. His main research interests are in fuzzy logic and optimisation.

1 Introduction

Hand (2007) points out how institutions in persecuting fraud follow the economic imperative, meaning it does not worth spending \$200 m to stop \$20 m fraud. Participants in his study estimate that US organisations lose 5% of their annual revenues to fraud. This means, applied to 2006 US GDP, approximately \$652 billion in fraud losses. According to KPMG (KPMG Forensic fraud Barometer, 2009), there is a prominent increase in fraud by individuals. Company managers, employees and customers together have been responsible for £300 m of fraud in 2008, three times the value of the year 2007.

In KPMG, Fraud Survey (2009) is shown that the most effective countermeasures for fraud are those ones developed by internal audit using clues given by employee whistleblowers as shown in Table 1. The survey has been conducted on executives of US companies who answered the following question: Through which source do you believe your organisation would be most likely to uncover fraud or misconduct?

Table 1 Fraud countermeasures

Internal audit	47%
Employee whistleblowers	20%
Line managers	13%
External auditors	9%
Customers or suppliers	4%
Government regulators or law enforcement	3%
Other means	2%

Source: KPMG Forensic fraud barometer (2009)

The embarrassment of admitting to mainly follow an economic imperative in persecuting fraud is coherent with the choice of preferring internal resources on external ones,

but this is not only the main reason. It is also related to the awareness that an internal audit team has a better knowledge of their organisation, the weakness of internal procedures and the personnel.

According to the Advanced Measurement Approaches (AMAs), introduced in Basel II accord (Basel Committee, 2006), banks are encouraged to develop sophisticated methodologies to calculate the operational risk, monitor the bank activities and reinforce internal control structure and auditing to preserve the integrity of the managerial processes. These systems include also the use of internal and external data, scenario analysis and control factors and an accurate reporting system based on key risk indicators.

In the banking sector, there is a prevalence of human fraud. One of the causes encouraging internal fraud is conflict of interests, which limits the effectiveness of the control procedures.

Commonly, the biggest problem of an audit team is to interpret and summarise all these behavioural aspects to come up with effective solutions to prevent fraud before they happen or to detect quickly the type of fraud when perpetrated. To this end, audit teams collect information about past behaviours to provide a formal representation of the most common typologies of fraud. This way, a repository of domain expert represented by standardised fraudulent attacks can be created and reused for, e.g., playing ‘what-if’ games with potential countermeasures or identifying the nature of new attacks.

Auditors, who are the fraud experts, can use their experience not only to remove false alarms, but also to detect those crimes that cannot be detected electronically because they are the results of untraceable human behaviours most often perpetrated inside the departments of the bank. The evaluation process of auditors in fraud detection has been examined, e.g., by Bazerman et al. (2002) and Grazioli et al. (2006), exploiting the reasons of their success and failure, and studying the impact of ambiguity, analysing a quite extended sample of case studies.

Several authors have demonstrated that a multi-agent approach is particularly suitable to address fraud detection when behavioural aspects play a key role, see for instance Chou et al. (2007), Wang et al. (2008) and Zhang et al. (2008).

We believe that the multi-agent system we are going to introduce in this paper, combining think-maps, attack trees and fuzzy numbers under a Delphi-based team work support system, do offer to the agents involved (inspectors and auditors) an innovative and suitable way to better understand and manage fraud schemes.

The multi-agent architecture of the FIDES (Buoni, 2010) will permit them to open up and share their knowledge and then link all the clues in a coherent scheme. The learning by doing approach of think-maps is a good means for inspectors to formally represent their knowledge linguistically expressed, to reconsider their opinions and correct their statements in the light of the comments received by their colleagues.

The paper has been organised as follows. In Section 2, we introduce the main components of FIDES, i.e., think-maps, attack trees and Delphi method. In Section 3, we describe the fuzzy mechanism that is used for representing and aggregating the linguistic information provided by the experts of the audit team when addressing the design of the attack tree. In Section 4, the whole structure of the system is figured out combined with the description of a work session. In the last section, some conclusions are drawn and future lines of research are sketched out.

2 Think-maps, attack trees, Delphi method

The scope of this section is to provide a synthetic description of the main components of FIDES, focusing the description on those features characterising the collection and representation of knowledge involved in the analysis and detection of fraudsters' attacks.

2.1 Think-maps

The concept of mind mapping was introduced for first by Buzan (1974). The idea was to organise keywords into a radiant structure that looks like a tree seen from above (Åhlberg, 2007). A radiant structure permits to put in the middle of the paper the central idea (goal) and through a brainstorming process add around it the branches of the map. Novak and Cañas developed the idea of concept maps (see Novak and Cañas, 2006), which they defined as “graphical tools for organising and representing knowledge”.

Oxman (2004) introduces the idea of think-maps, which means that conceptual mapping of designing ideas can be constructed in larger structures where it is possible to organise knowledge acquired by the learner and make it explicit. The software they develop to create these maps is called ‘Web-Pad’. The theoretical basis of think-maps is constructivism. Constructivist theories of learning state that the learner is not a passive recipient of knowledge, but it has an active role in creating knowledge, based on the “learning by doing” approach.

Learners construct their knowledge based on their experience and relationship with concepts. Think-maps have a formal representation called ICF (Issue-Concept-Form), which acts as “a structuring ontology for the construction of conceptual networks of design concepts” (Oxman, 2004).

Internal fraud cases can be anticipated by particularly suspicious behaviours like working overtime, reluctance to take breaks and sudden change of lifestyle. In this section, we will describe a case to show the process and how inspectors would model it.

In our case, inspectors will use Web-Pad (Figure 1) to comment suspicious behaviours or activities they observed or information acquired by whistleblowers or insert their reports to share them with all the other colleagues. Since the software works in a real-time discussion environment, their comments will appear in the text box named ‘Description’ as a common chat discussion shape. At this stage, inspectors will have the possibility to arrange the text to have a clear description of the case.

At this point, inspectors will start to underline keywords as it can be seen in the following example:

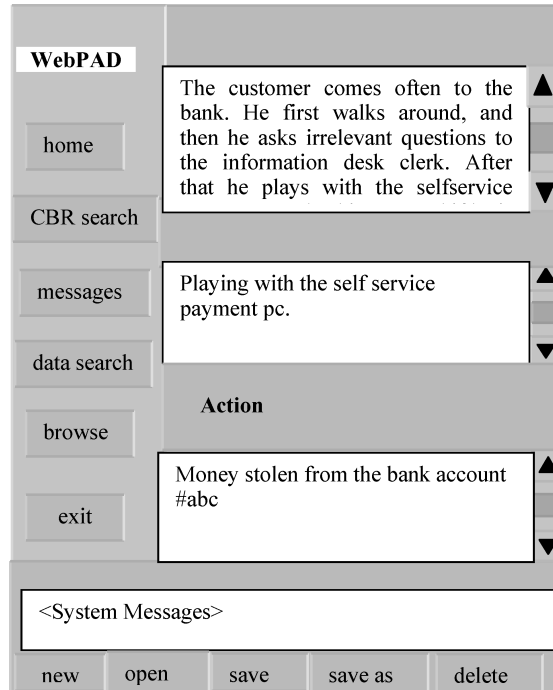
“The customer comes often to the bank. He first walks around, then he asks irrelevant questions to the information desk clerk. After that he plays with the self-service payment pc, checking around if there is someone observing him.”

Then, keywords, which are considered important to interpret the case, will be associated to two main labels ‘suspicious behaviours’ and ‘action’, which are related to one or more suspected.

At the end of this process, the think-map shown in Figure 2 is obtained. In the map, we can see the main labels, which have been used to visualise a fraud operation. The cloud callouts represent the activities related to ‘suspicious behaviour’, whose label

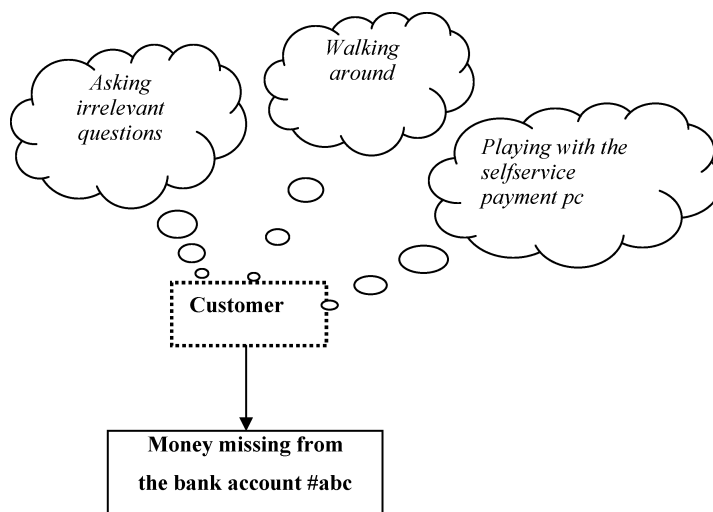
was previously created with Web-Pad, associated to three different suspected persons. There is a customer who has the habit to play with the self-service payment pc of the bank. In the middle of the figure, we have the suspected and on the bottom the fraudulent action, which is the amount of money missing in a specific bank account.

Figure 1 Web-Pad interface



Source: Oxman (2004)

Figure 2 The think-map



Having this think-map as a descriptive model, inspectors will start to create nodes, which will be sent to the auditors to be connected through the Delphi process to build the attack tree. Web-Pad basically works on two levels: a graphic level where think-maps can be visualised and a text level where different operations can be stored and retrieved for the future, supporting the inspectors when they have to build the think-map.

For instance, inspectors could be interested to have a list of the cases associated to a particular behaviour and so on. In the data retrieval mode, the system can bring up precedents that inspectors consider similar, according to their subjective judgement.

Users can express the level of similarity between two different cases as a number between 0 and 1. Once the database is populated with a significant number of cases, it will be possible to retrieve and visualise them in descendant order, according to their level of similarity, ready to be examined.

2.2 *Attack tree*

The attack tree, introduced by Schneier (1999), is a tree-based diagram to “systematically categorise the different ways in which a system can be attacked”. Nodes are the elementary attacks and the root node is the goal of the attack. Children of a node are refinements of this goal, whereas leaves are attacks, which cannot be further on refined. The process of creating an attack tree starts by identifying the possible attack goals, where each goal forms a separate tree, but there is also the possibility that different trees share nodes and subtrees.

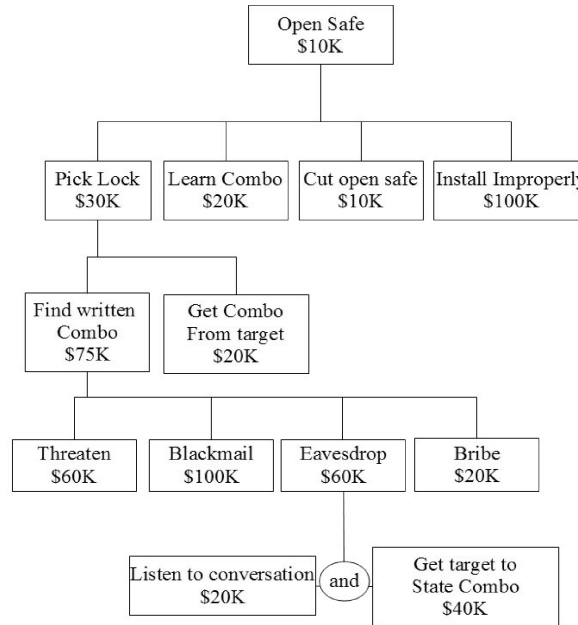
Accordingly, modelling an attack tree is a matter of associating a logical AND and a logical OR with each node, and therefore encouraging a structured representation of events and of the ways they are connected.

This supports the discovering of the most likely avenues of approach for an attack making easier the deployment of the most effective countermeasures. Even though Schneier’s attack trees (illustrated in Figure 3) have been considered from their first appearance, a convenient tool to systematically categorise the different modes in which an attack can be carried out, nonetheless their network structure has been criticised for its simplicity and for the lack of well-sounded theoretical foundations.

Mauw and Oostdijk (2005), arguing that Schneier’s approach to attack trees is semantically not well sounded, provide a generalisation based on the observation that an attack tree describes an attack suite and that a node can be connected to a multi-set of nodes (bundle) and may contain several bundles.

But, since our paper is more focused on the way on which the team of experts carry out, in a Delphi-based context, the consensual construction of the tree, than on the complexity of its structure, the nature of the tree is irrelevant and therefore we will adopt Schneier’s approach.

Niitsoo (2010) also points out how it is important to develop attack tree models that take into consideration not only whether the attack is possible or not, which cannot tell so much about not only the likelihood of the attack, but also incentives and possibilities available to the fraudster to try to analyse his or her behaviour.

Figure 3 An example of attack tree

\$=Cost of the attack

Source: Schneier (1999)

2.3 Delphi method

Delphi method, introduced for the first time in the 1950s for a US-sponsored military project (Gordon, 1994), is a systematic, interactive and iterative method, which relies on a team of experts, aiming at discussing and structuring the solution of a given problem. The experts are asked to answer questionnaires in two or more rounds, and after each round a moderator provides an anonymous summary of the experts' analysis from the previous round as well as some explanations of their judgements. The moderator encourages experts to reuse their earlier opinions in light of the outcomes of the analysis provided by the other experts of the team. The process is stopped according to a pre-defined criterion and some average measures of the outcomes of the final round determine the output of the process.

Delphi method has also been used recently in many different fields of research like R&D to explore the barrier factors to the adoption of mobile service (Steinert, 2009), security evaluation of embedded systems (Zhang et al., 2008), road safety (Ma et al., 2011) and clinical nursing (McElhinney, 2010).

Rowe and Wright (1999) suggest four key features for a good design of Delphi:

- *Anonymity of Delphi experts*: Allows the experts to freely express their opinions without undue social pressures to conform from others in the team. Decisions are evaluated on their merit, rather than who has proposed the idea.
- *Iteration*: Allows the experts to refine their views in light of the progress of the team's work from round to round.

- *Controlled feedback*: Informs the experts of the other experts' perspectives, and provides the opportunity for Delphi experts to clarify or change their views.
- *Statistical aggregation of team response*: Allows for a quantitative analysis and interpretation of output information.

In our system, Delphi will be used as a method for finding the agreement on the connections between different nodes figured out by the inspectors. The role of the moderator will be to ask the experts their opinion about strength of the links connecting different nodes.

The aggregation process will be described in detail in Section 3. In the end of the Delphi process, the output will be an attack tree.

3 The fuzzy mechanism

The audit team performs the Delphi process aiming to select the nodes and connect them to design the attack tree. In the first phase, the inspectors determine the possible nodes of the attack tree with the help of the think-map. Then, the moderator will ask the experts about the possible connection of the nodes, and aggregate the results to obtain the attack tree. In this section, we will describe the fuzzy mechanism, which helps the experts to form the attack tree. In the literature, a number of different fuzzy approaches to the analysis of negotiation processes in multi-agent decision making have been proposed, and for an extended overview the interested reader could see, e.g., Carlsson et al. (2004) and Fedrizzi et al. (1997). Before the description of the model, we need some basic definitions from fuzzy set theory.

Definition 1 (A fuzzy number (Dubois and Prade, 1978)): Let $X = x$ denote a collection of objects (points) denoted generically by x . Then, a fuzzy set A in X is a set of ordered pairs

$$A = (x, \mu_A(x)), \quad x \in X \quad (1)$$

where $\mu_A(x)$ is termed the grade of membership of x in A , and $\mu_A: X \rightarrow M$ is a function from X to a space M called the membership space. When M contains only two points, 0 and 1, A is non-fuzzy and its membership function becomes identical with the characteristic function of a crisp set. This means that crisp sets belong to fuzzy sets. A fuzzy number is a convex fuzzy set on the real line such that

- 1 $\exists x_0 \in A, \mu_A(x_0) = 1$,
- 2 μ_A is piecewise continuous.

(The convexity means that all the γ level sets are convex. Furthermore, we call F the family of all fuzzy numbers).

A γ level set of a fuzzy set A is defined by $[A]^\gamma = \{x \in A: \mu_A(x) \geq \gamma\}$ if $\gamma > 0$ and $[A]^\gamma = cl\{x \in A: \mu_A(x) > \gamma\}$ (the closure of the support of A) if $\gamma = 0$. Let A be a fuzzy number. Then, $[A]^\gamma$ is a closed convex subset of R for all $\gamma \in [0, 1]$.

We use the notations

$$a_1(\gamma) = \min[A]^\gamma, \quad a_2(\gamma) = \max[A]^\gamma$$

for the left-hand side and right-hand side of the γ -cut, respectively. When we calculate the arithmetic operations on fuzzy sets (fuzzy numbers), we apply the rules of interval arithmetic. Let A and B be fuzzy numbers with the corresponding γ -cuts: $[A]^\gamma = [a_1(\gamma), a_2(\gamma)]$, $[B]^\gamma = [b_1(\gamma), b_2(\gamma)]$, then the γ -cut of the fuzzy number $A + B$ is the following:

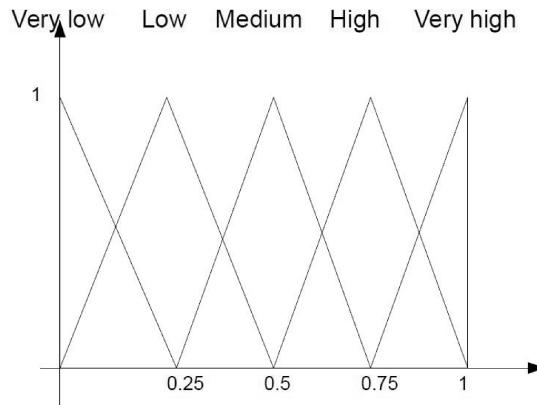
$$[A + B]^\gamma = [a_1(\gamma) + b_1(\gamma), a_2(\gamma) + b_2(\gamma)],$$

and the γ -cut of the fuzzy number αA , where $\alpha > 0$: $[\alpha A]^\gamma = [\alpha a_1(\gamma), \alpha a_2(\gamma)]$.

We will use linguistic labels in the questionnaire and we represent the labels as fuzzy numbers (see Figure 4 for a possible representation). We suppose that the moderator can choose which nodes are parents (V) (with descendant) and which ones are leaves (L) (without descendants, basic attack components). We obtain two sets:

$$L = \{l_1, \dots, l_s\}, V = \{v_1, \dots, v_t\}.$$

Figure 4 Possible representation with triangular fuzzy numbers



In the first questionnaire, the experts have to express their opinion in linguistic terms about statements like “ $l_i \in L$ is required for $v_j \in V$ ”, for every $i = 1, \dots, s$, $j = 1, \dots, t$.

Then, experts $E = \{e_1, \dots, e_N\}$ are asked to determine their level of agreement on the statements based on a linguistic scale with m terms for every pair (l_i, v_j) .

The linguistic terms in the model are represented as fuzzy numbers. In other words, we have a mapping $\Phi: T \rightarrow F$ from the set of linguistic terms into the family of fuzzy numbers.

Example 1: One possible representation for a linguistic label is a triangular fuzzy number:

$$A(u) = \begin{cases} 1 - \frac{a-u}{\alpha}, & a - \alpha \leq u \leq a \\ 1 - \frac{u-a}{\beta} & a \leq u \leq a + \beta. \\ 0 & \text{otherwise.} \end{cases}$$

From the opinion of the experts, we obtain the frequencies of the different classes.

For the pair (l_i, v_j) , we have $n_1^{ij}, \dots, n_m^{ij}$. If we denote by A_1, \dots, A_m the fuzzy numbers corresponding to the linguistic labels, we can define a new fuzzy number A_{ij} as a ‘weighted average’, with level sets:

$$[A_{ij}]^\gamma = \left[\frac{1}{N} \sum_{k=1}^m n_k^{ij} a_1^k(\gamma), \frac{1}{N} \sum_{k=1}^m n_k^{ij} a_2^k(\gamma) \right],$$

where $[a_1^k, a_2^k]$ is the level set of A_k . This is clearly a fuzzy number with the support in the interval $[0,1]$.

To obtain the connection degree for the pair (l_i, v_j) , we calculate the f -weighted possibilistic mean value of A_{ij} , defined in Carlsson and Fuller (2001).

Definition 2: The f -weighted possibilistic mean value of $A \in F$, with γ -level sets $[A]^\gamma = [a_1(\gamma), a_2(\gamma)]$, $\gamma \in [0,1]$, is defined by:

$$E_f(A) = \int_0^1 M(U_\gamma) f(\gamma) d\gamma = \int_0^1 \frac{a_1(\gamma) + a_2(\gamma)}{2} f(\gamma) d\gamma \tag{2}$$

where U_γ is a uniform probability distribution on $[A]^\gamma$ for all $\gamma \in [0,1]$.

After we have obtained these defuzzified numbers as the estimation of the connection strengths, we can determine for every attack component the ranking of the other nodes, then we can construct the adjacency matrix of the attack tree by connecting the leaves to the best ranked vertices.

Example 2: In the simplest case, we can represent the linguistic labels as a fuzzy set with the membership function

$$A(u) = \begin{cases} 1 & u = c \\ 0 & \text{otherwise} \end{cases}$$

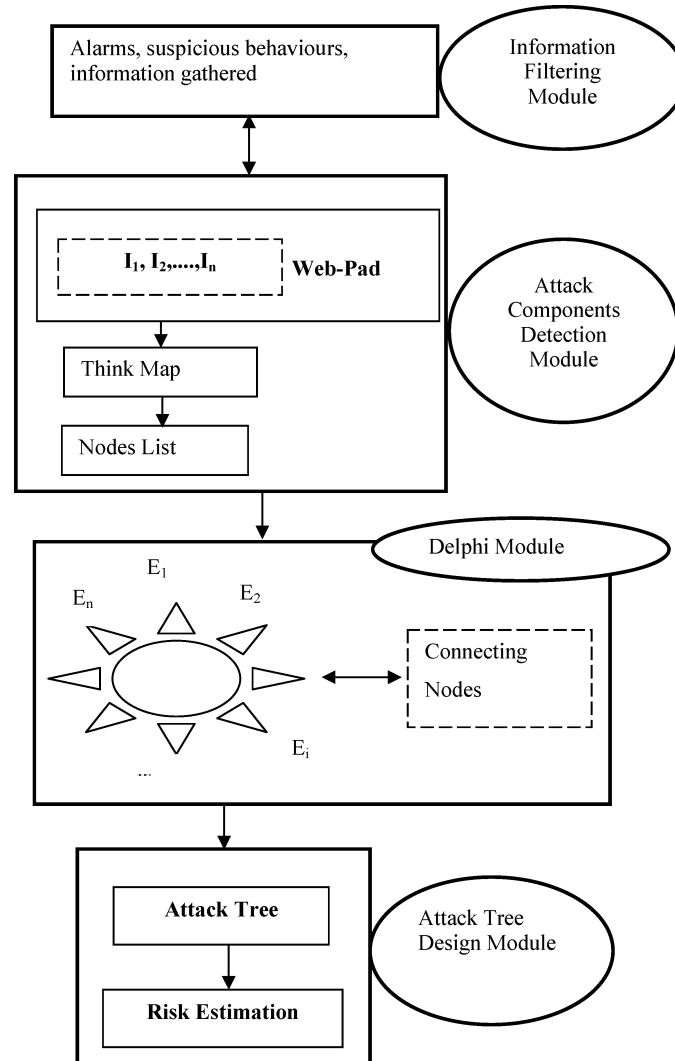
If we have five categories, we use the set $\{0, 0.25, 0.5, 0.75, 1\}$. The weights of the outcomes are the frequencies of the linguistic labels. If we observe the weights $n_0, n_{0.25}, n_{0.5}, n_{0.75}, n_1$, then A_{ij} is just the characteristic function of the value

$$\frac{1}{\sum_{j=0}^4 n_{0.25^*i}} \sum_{i=0}^4 n_{0.25^*i} * 0.25i,$$

what is simply the sample mean value of our data. And, according to the used defuzzification method, the obtained connection estimation is this sample mean.

4 The architecture of FIDES

In this section, we will describe the architecture of FIDES (Figure 5), showing how the main components are interrelated and the role played by the inspectors and auditors, the experts, who have to decode the alarms generated by software agents and to detect and describe the suspicious human behaviours (see, e.g., Sanchez et al., 2009; Edge et al., 2007).

Figure 5 The architecture of the system

FIDES, indeed, has been built on the base of the suggestions we collected interviewing the managers of risk management department of a leading European bank and thus the multi-agent system has been designed according to their opinions.

In a fraud detection process, an audit team has to deal with not only numerical data and alarms, produced by software agents, but also documents, reports and information gathered by different actors (managers, whistleblowers and anonymous informers) and then summarised by inspectors.

Therefore, the key factor in detecting fraud behaviours is to improve the interaction between inspectors and auditors. Alarms, generated by software agents or suspicious behaviours noticed personally by inspectors, are filtered using the Web-Pad software. The Information Filtering Module is nothing else that a preliminary session where inspectors can decide which alarms and behaviours to take into considerations to perform

all the process. Once the case has been well detailed as shown in Figure 1, inspectors start to underline the keywords to create the think-map as described in Section 2.

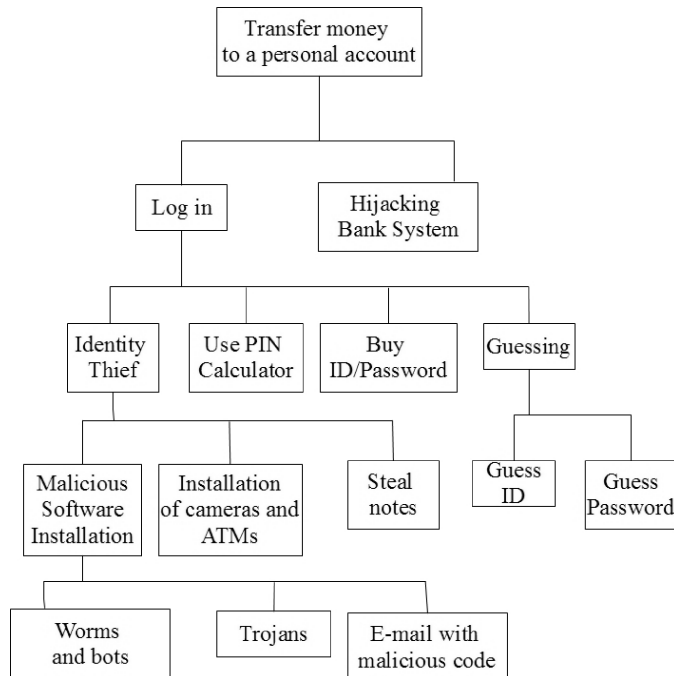
This is the Attack Components Detection Module whose output is the set of nodes to be sent to the auditors.

Since think-map represents a hypothesis of correspondence between behaviour, suspected fraudsters and fraudulent actions, inspectors create nodes to try to offer an explanation of how this conceptual framework can be associated to concrete actions performed by the fraudster. Nodes are elementary attacks expressed by labels, which define all the possible steps of the fraudulent action.

At this stage, auditors activate the Delphi Module, driven by the moderator, as explained in Section 3, to the end of connecting the nodes to form the attack tree, taking into account the strength of the links between nodes. In the Attack Tree Design Module, experts can estimate the risk and develop the strategy to persecute the fraudster. The estimation process can be performed calculating the most probable or the least expensive path.

Continuing with the example described in Section 2.1, in Figure 6 is shown the final attack tree as a result of Delphi process where auditors connect the nodes delivered by the inspectors. In this example, the goal of the attacker is to transfer money to his or her personal account. The tree shows different ways he or she can achieve this goal. The easiest but the least successful path is to guess ID and password. The most difficult one is to hijack the bank system. The other paths show identity thief sub-attacks based on malicious software installation or the use of cameras installed at ATMs.

Figure 6 Final attack tree



To select the most probable path, in the risk evaluation phase, experts can have the opportunity to compare the actual attack tree with the ones created in the past and use the information that could be useful to develop a counterstrategy for the new attack. Our assumption is that the identified attack of similar trees can be useful in finding the real path in the present tree. Moreover, by comparing the background of the cases we can obtain useful information for example the possible amount of fraud, the fraudster or his or her strategy. A typical situation could be that the same person is performing the same attack, which can be discovered by identifying similar attack trees in a specific range of time.

An attack tree represents the set of potential fraud schemes. After we constructed the tree, we would like to determine the real attack, which takes place in the present. To do this, we need a database containing all the attack trees constructed in the past, then comparing the newly created tree with the ones in the database and based on the similar trees and their outcomes, we will be able to choose the most probable attack.

A graph (in our special case a tree) can be represented by its adjacency matrix. The adjacency matrix can be used to find the similarity value between attack trees but the problem is that a tree with n vertices has $n - 1$ edges so the matrix is very sparse. Moreover, since we need comparable matrices for determining the similarity, we have to consider the matrix with all the possible keywords for every tree, and if the number of keywords is k , even in the best case we will have approximately $k(k - 2)$ number of 0s in the matrix. Our aim is to find trees with similar sets of edges. It is important to note that every vertex of the attack tree (like every decision tree) can be associated with the distance value from the root node. Thanks to the structure of the attack tree, we are able to represent the edges of the tree as ordered pairs where the first element is the vertex with smaller distance from the root (the goal of the attack). The set of these pairs is different for two attack trees and if we know this set we can construct the tree. Then, we find the similarity of these sets (the number of pairs in a set is equal to the number of vertices in the tree minus 1).

To obtain the similarity of two sets, we will employ an index, which is frequently used in information retrieval, the cosine similarity. Before the definition of the index, we need to introduce a notation: if we have two sets, S_1 and S_2 then $M = |S_1 \cap S_2|$ is the number of common items between S_1 and S_2 . The cosine similarity of two sets can be defined as

$$\cos(S_1, S_2) = \frac{M}{\sqrt{|S_1||S_2|}} \quad (3)$$

where $|S_i|$ is the number of items in S_i . In our case, the sets consist of the edges of the trees, so if a tree has n vertices, then the corresponding set will contain $n - 1$ elements. Finding this similarity measure can be done effectively for example by using the method described in Nanopoulos and Manolopoulos (2002).

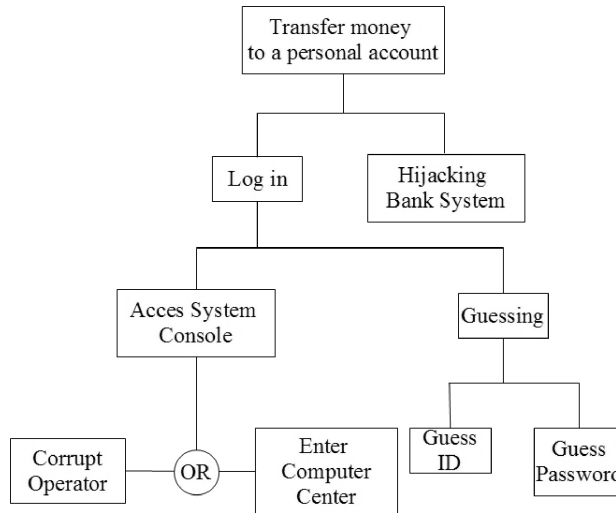
Example 3: If we look at the two constructed attack trees in Figure 6 (S_1) and Figure 7 (S_2), we can calculate the similarity by using the formula (3).

The number of edges is 14 and 8 in S_1 and S_2 , respectively. If we compare the sets, we can see that the two trees have 5 edges in common: (transfer money to a personal account, log in), (transfer money to a personal account, hijacking bank system),

(log in, guessing), (guessing, guess ID), (guessing, guess password). This means that $M = 5$, and the cosine similarity:

$$\cos(S_1, S_2) = \frac{M}{\sqrt{|S_1| |S_2|}} = \frac{5}{\sqrt{8 \times 14}} \approx 0.47.$$

Figure 7 Possible attack tree from the database



5 Conclusions

Most of the fraud surveys published in the last five years by leading international consulting companies have shown that behavioural aspects play a central role, and therefore the biggest problem to be addressed by auditors and inspectors is to interpret the signals and summarise the information coming from several, sometimes conflicting, sources. The successful solution of this kind of problems depends to a large extent on a proper combination of critical analysis, knowledge-based actions, whistleblowers' messages interpretation, involving groups of interacting experts.

Since before addressing the design of our system, we had the chance to meet several experts in charge of diverse risk management activities inside one of the largest European banking groups, our work has been inspired by the information collected during the meetings. Banks have a huge amount of data and experience concerning fraud cases, but they do not use it in an efficient way, since most of the knowledge is unstructured. One of the main challenges of the bank is to unify fraud risk assessment processes in the different countries where its branches are located, perform a realistic temporal analysis and establish cause/effect relationship in a rather short time combining objective information with the subjective judgements expressed by experts.

Treasuring the knowledge and the opinions collected during the meetings, we designed the multi-agent system FIDES, intended for the management of fraud detection situations where inspectors and auditors are collaborating according to a two-phase

detection process. In the first phase, think-maps and Delphi method offer to the inspectors and auditors, respectively, an effective environment to explicit their knowledge, select the most likely fraud attack components, and finally structuring them in an attack tree. In the second phase, the attack tree structure is definitively settled introducing a representation of uncertainty involved based on fuzzy numbers. Attack tree-based representation of fraudulent processes has the advantage of offering a clear visualisation of the attack, which permits to the auditors to highlight the most probable attack paths, after introducing the cost or impact of an attack starting from a set of attributes attached to the nodes of the attack tree.

On the basis of the suggestions and comments collected during the work sessions with the members of the audit team of the bank, we started to develop a first prototype of FIDES. The prototype is limited to two modules, i.e., the Attack Components Detection and Delphi ones. A test has been conducted with the collaboration of the inspectors of the bank to verify the usability and understand the limitations of the system. During the experiments, a list of different cases was presented as a short description and possible consequences (typically a loss from a bank account) will be provided to the users (inspectors) to create think-maps and then nodes as described in the previous section.

A second group of experts will carry out the process of linking the nodes, aiming at designing the final attack tree. The group process will be driven by a moderator, selected among the members of the same group, according to the Delphi method, and the consensual dynamics based on software implemented in Fedrizzi et al. (2008).

After we have sent few screenshots of the first draft of the prototype, we asked the users to give us a short feedback highlighting positive and negative aspects about what they observed and suggest further developments. The first reaction was a positive evaluation of the improvement of the interaction dynamics between inspectors and auditors. Another feature, which has been appreciated, is the anonymity, which allowed the users to operate with more freedom, without the pressure of performing a wrong evaluation of the case.

Regarding negative aspects, users pointed out possible problems in the interpretation of the fraud cases once the system was used by people with different cultural backgrounds and languages, a problem arising due to the international profile of the banking group. A limitation would be the possibility to manage big fraud operation on international scale, where different auditors and inspectors from different contexts might find difficulties in understanding each other. In particular in building the think-map, the perception and interpretation of the facts might encounter problems on a semantic level, like underestimate or overestimate the same behaviour or action performed in different contexts.

Concerning future improvements of FIDES, first of all, assuming that the formally expressed representation of the information related to the attacks (attack trees) is stored in references sources (catalogue of attack trees), an intelligent searching module will be introduced for retrieving information from the catalogue. To wit, our aim is to propose an ontology-based description of the attack trees to provide a formal basis for sharing knowledge and for reusing it during the attack tree design process.

Another possible improvement would be a mobile version of FIDES, particularly useful when inspections are carried out in the different branches of the bank spread to a large geographic area, and aiming at homogenising the information coming from different information sources.

References

- Åhlberg, M. (2007) *History of Graphic Tools Presenting Concepts and Propositions*, <http://www.reflectingeducation.net/index.php?journal=reflecting&page=article&op=downloadSuppFile&path%5B%5D=49&path%5B%5D=6>
- Basel Committee (2006) Basel II, 2006, *Basel Committee on Banking Supervision*, <http://www.bis.org/publ/bcbs128.htm>
- Bazerman, M., Lowenstein, G. and Moore, D. (2002) 'Why good accountants do bad audits', *Harvard Business Review*, Vol. 80, pp.3–8.
- Buoni, A. (2010) 'Fraud detection: from basic techniques to a multi-agent approach', *International Conference on Management and Service Science*, 24–26 August, Wuhan, pp.1–4.
- Buzan, T. (1974) *Use Your Head*, BBC Books, London.
- Carlsson, C. and Fuller, R. (2001) 'On possibilistic mean value and variance of fuzzy numbers', *Fuzzy Sets and Systems*, Vol. 122, pp.315–326.
- Carlsson, C., Fedrizzi, M. and Fuller, R. (2004) *Fuzzy Logic in Management*, Kluwer Academic, Dordrecht, Boston, New York.
- Chou, C.L-Y., Du, T. and Lai, S.V. (2007) 'Continuous auditing with a multi-agent system', *Decision Support Systems*, Vol. 42, No. 4, pp.2274–2292.
- Dubois, D. and Prade, H. (1978) 'Operations on fuzzy numbers', *International Journal of Systems Science*, Vol. 9, No. 6, pp.613–626.
- Edge, M.E., Sampaio, R.P.F. and Choudhary, M. (2007) 'Towards a proactive fraud management framework for financial data streams', *Proceedings of the Third IEEE International Symposium on Dependable, Autonomic and Secure Computing*, 25–26 September, Columbia, MD, pp.163–171.
- Fedrizzi, M., Fedrizzi, M., Marques Pereira, R. and Brunelli, M. (2008) 'Consensual dynamics in group decision making with triangular fuzzy numbers', *Proceedings of the 41st Annual Hawaii International Conference on System Sciences*, IEEE Computer Society, Waikoloa, Hawaii, pp.70–78.
- Fedrizzi, M., Kacprzyk, J. and Nurmi, H. (1997) *Consensus under Fuzziness*, Kluwer Academic, Dordrecht, Boston, New York.
- Gordon, T.J. (1994) *The Delphi Method in Futures Research Methodology*, AC/UNU Millennium Project, AC/UNU, Washington.
- Grazioli, S., Johnson, P.E. and Karim, J. (2006) *A Cognitive Approach to Fraud Detection*, <http://ssrn.com/abstract=920222>
- Hand, D.J. (2007) *Statistical Techniques for Fraud Detection and Evaluation*, http://langtech.jrc.ec.europa.eu/mmdss2007/htdocs/Presentations/Docs/MMDSS_Hand_PUBLIC.pdf
- KPMG Forensic fraud barometer (2009) http://www.yhff.co.uk/Fraud%20Barometer%20-%20Feb%202009%20_2_.pdf
- KPMG Fraud survey (2009) <http://www.kpmginstitutes.com/aci/insights/2009/pdf/kpmg-fraud-survey-2009.pdf>
- Ma, Z., Shao, C., Mac, S. and Ye, Z. (2011) 'Constructing road safety performance indicators using fuzzy Delphi method and grey Delphi method', *Expert Systems with Applications*, Vol. 38, No. 3, pp.1509–1514.
- Mauw, S. and Oostdijk, M. (2005) 'Foundation of attack trees', *Proceedings of the 8th Annual International Conference on Information Security and Cryptology*, 1–2 December, Seoul, Korea, LNCS 3935, Springer-Verlag, Berlin Heidelberg, pp.186–198.
- McElhinney, E. (2010) 'Factors which influence nurse practitioners ability to carry out physical examination skills in the clinical area after a degree level module – an electronic Delphi study', *Journal of Clinical Nursing*, Vol. 19, pp.21, 22, 3177–3187.

- Nanopoulos, A. and Manolopoulos, Y. (2002) 'Efficient similarity search for market basket data', *The VLDB Journal*, Vol. 11, pp.138–152.
- Niitsoo, M. (2010) 'Optimal adversary behavior for the serial model of financial attack trees', *Proceedings of the 5th International Workshop on Security*, 22–24 November, Kobe, Japan, LNCS 6434, Springer-Verlag, Berlin Heidelberg, pp.354–370.
- Novak, J. and Cañas, A. (2006) *The Theory Underlying Concept Maps and How to Construct Them*, Technical Report IHMC Cmaptools.
- Oxman, R. (2004) 'Think-maps: teaching design thinking in design education', *Design Studies*, Vol. 25, No. 1, pp.63–91.
- Rowe, G. and Wright, G. (1999) 'The Delphi technique as a forecasting tool: issues and analysis', *International Journal of Forecasting*, Vol. 15, pp.353–375.
- Sanchez, D., Vila, M.A., Cerda, L. and Serrano, J.M. (2009) 'Association rules applied to credit card fraud detection', *Expert Systems with Applications*, Vol. 36, pp.3630–3640.
- Schneier, B. (1999) 'Attack trees', *Dr. Dobbs's Journal*, December, Available at <http://www.schneier.com/paper-attacktrees-ddjft.html>
- Steinert, M. (2009) 'A dissensus based online Delphi approach: an explorative research tool', *Technological Forecasting & Social Change*, Vol. 76, pp.291–300.
- Wang, D.G., Li, T., Liu, S.J.L., Liang, G. and Zhao, K. (2008) 'An immune multi-agent system for network intrusion', *Proceedings of the Third International Symposium on Intelligence Computation and Applications (ISICA 2008)*, 19–21 December, Wuhan, China, LNCS 5370, Springer-Verlag, Berlin Heidelberg, pp.436–445.
- Zhang, L.S., Zhou, N. and Wu, J.X. (2008) 'The fuzzy integrated evaluation of embedded system security', *International Conference on Embedded Software and Systems (ICCESS 2008)*, 29–31 July, Sichuan, China, pp.157–162.

Paper 4

Buoni, A., Fedrizzi, M. (2012). Consensual dynamics and Choquet Integral in an attack tree-based fraud detection system, *Proceedings of the 4th International Conference on Agents and Artificial Intelligence*, Volume 1, Algarve (Portugal), 6-8 February, 238-288.

© 2012 Reprinted with the permission from SciTePress Digital Library Online.

CONSENSUAL DYNAMICS AND CHOQUET INTEGRAL IN AN ATTACK TREE-BASED FRAUD DETECTION SYSTEM

Alessandro Buoni¹, Mario Fedrizzi²

¹*IAMSR, Turku Centre for Computer Science, Joukahaisenkatu 3-5 B, 20520 Turku, Finland*

²*Department of Computer and Management Sciences, University of Trento
abuoni@abo.fi, mario.fedrizzi@unitn.it*

Keywords: fraud detection, attack tree, consensus, Choquet integral

Abstract: In this paper we extend two modules of the multi-agent system FIDES (Fraud Interactive Detection Expert System) previously introduced in Buoni et al. (2011), and involving the attack tree representation of fraudulent attacks. First, assuming that the opinions of experts involved in the design of the attack tree are represented by fuzzy preference relations, we introduce a dynamical consensus model aiming at finding a shared representation of the attack tree. Second, assuming that the leaf nodes of the attack tree are attribute fuzzy numbers valued and that the attributes are interdependent, we show how to propagate the values up the tree through an aggregation process based on Choquet integral.

1 INTRODUCTION

KPMG Fraud survey (KPMG 2009), conducted on executives of U.S. companies, shows that the most important sources to detect fraud are internal audit (47%) and employee whistle blowers (20%).

An audit team (experts) have to deal with both numerical data and unusual behaviours, create different scenario, develop risk indicators to detect and prevent fraud.

In order to achieve this goal, experts analyse information about the past fraud cases, as collected by inspectors along the processes. Reuse this information and deal with this huge amount of data is a typical knowledge management problem.

A system to support the work of experts has to take into considerations the complexity of managing this kind of information, affected by imprecision, uncertainty, behavioural aspects, and false alarms.

Moreover, one critical issue to address is to aggregate the judgments of the single experts in order to extract useful knowledge in a structured way, develop countermeasures to detect frauds in real time, activate effective strategies to prevent and adapt them when new unusual schemes happen.

Several authors have demonstrated that a multi-agent approach is particularly suitable to address fraud detection when behavioural aspects play a key

role, see for instance Chou et al. (2007), Wang et al. (2009), and Zhang et al. (2008).

Accordingly, in Buoni et al. (2011) we introduced FIDES (Fraud Interactive Detection Expert System), a multi-agent system combining think-maps, attack trees, and fuzzy numbers under a Delphi-based team work support framework, to offer to the experts an innovative and suitable way to better understand and manage fraud schemes.

The system has been developed in cooperation with a group of analysts coming from the risk management department of a leading European bank.

The most critical issue to address in FIDES is to perform the Delphi process aiming to select the nodes and connect them in order to design the attack tree that is used to systematically categorize the different ways in which a system can be attacked.

In this paper, at first we extend the Delphi module of FIDES introducing a dynamical consensus model based on individual fuzzy preferences representing the opinions of experts.

Secondly, assuming that to the leaves of the attack tree fuzzy attribute values are associated, we propose an innovative approach for aggregating these values based on Choquet integral.

The paper is organized as follows. The second section is devoted to a short description of FIDES. In the third section we introduce the consensus

mechanism based on a dynamical model updating the fuzzy preference of the experts. Section four addresses the aggregation of attribute values using the Choquet integral. The last section is devoted to conclusions and perspectives on future work.

2 CONSENSUSUAL MODELING OF THE ATTACK TREE

The Fraud Interactive Detection Expert System (FIDES) has been introduced at first in Buoni et al. (2010) and then extended in Buoni et al. (2011).

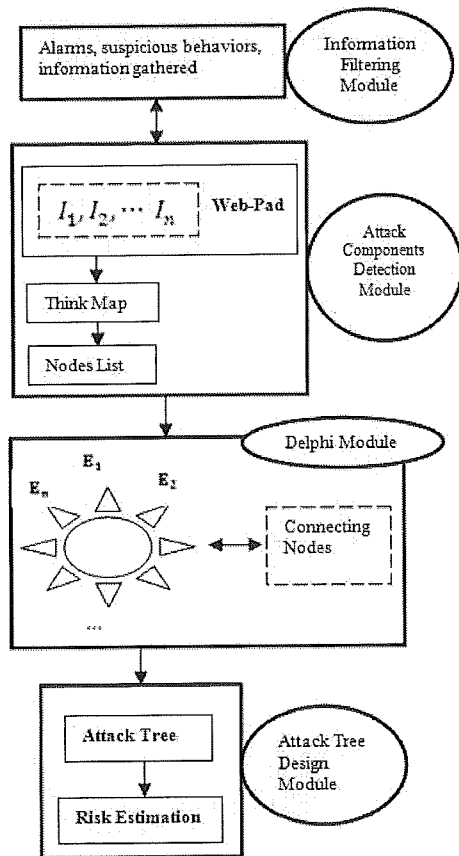


Figure 1: The architecture of FIDES

As shown in Fig.1, FIDES has four modules. In the Information Filtering Module, alarms, suspicious behaviours and information gathered by inspectors during their inspections or through whistle blowers, are evaluated.

All this information is then processed in the Attack Components Detection Module. Inspectors

using Web-Pad (Oxman 2004) can organize all this information in a think-map, i.e. a representation of a possible attack where three main elements are visible and connected: the action perpetrated the suspected person and suspicious behaviours, which might be connected with the other two elements.

Using the think map as a model, inspectors create nodes, which are elementary attacks, to be sent to the audit team experts.

The third module is founded on the Delphi method (Gordon 1994), it is an interactive and iterative method, typically based on questionnaires, where experts, supported by a moderator, try to refine their opinions after each round, in order to structure a description of the components of the fraud attack, based on an attack tree (Schneier, 1999).

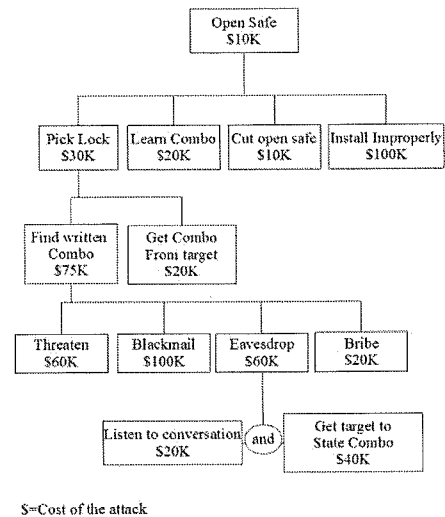


Figure 2: An example of attack tree (Schneier 1999)

The design of the attack tree starts from the set of possible nodes previously determined by the inspectors with the help of the think-maps. Accordingly, the moderator can choose which nodes are parents (V) (with descendant) and which ones are leaves (L) (without descendants, i.e. basic attack components) obtaining two sets $L = \{l_1, \dots, l_s\}$ and $V = \{v_1, \dots, v_t\}$.

Then, each expert is asked to elicit his/her own preference with respect to the strength of connections between the elements of L and V . The individual preferences are represented as fuzzy preference relations defined on the set $P = L \times V$.

The first goal to achieve is to find a consensual preference setting activating a Delphi session. To this aim, we introduce a dynamical consensus model based on the updating of the individual fuzzy preferences expressed by the group of experts on the set of pairs (i, v_j) . The modelling framework here used is that one described for first in Kacprzyk and Fedrizzi (1986) and then extended by Fedrizzi et al (1999) through the introduction of a consensual network dynamics that can be regarded as an unsupervised learning algorithm.

The point of departure is a set of individual fuzzy preference relations defined on $P = \{p_1, \dots, p_M\}$ for each expert in the set $E = \{e_1, \dots, e_N\}$. The fuzzy preference relation of expert e_i , R_i , is given by its membership function $\mu_i: P \times P \rightarrow [0,1]$ such that

- $\mu_i(p_k, p_l) = 1$, if p_k is definitely preferred over p_l ,
 - $\in (0.5, 1)$, if p_k is preferred over p_l ,
 - $= 0.5$, if there is indifference between p_k and p_l
 - $\in (0, 0.5)$, if p_l is preferred over p_k ,
 - $= 0$, if p_l is definitely preferred over p_k ,
- where $i = 1, \dots, N$ and $k, l = 1 \dots M$.

Each individual fuzzy preference relation R_i can be represented by a matrix $[r_{kl}^i]$, $r_{kl}^i = \mu_i(a_k, a_l)$, which is commonly assumed to be reciprocal, that is $r_{kl}^i + r_{lk}^i = 1$. Clearly, this implies $r_{kk}^i = 0.5$ for all $i = 1, \dots, N$ and $k = 1, \dots, M$.

In the soft consensus model each expert is represented by a pair of connected nodes, a primary node and a secondary node. The N primary nodes form a fully connected sub network and each of them encodes the preference of a single expert. The N secondary nodes, on the other hand, encode the individual preferences originally declared by the experts and each of them is connected only with the associated primary node.

Moreover, for the sake of simplicity, let us assume that the alternatives available are only two, that is $M=2$, which means that each (reciprocal) individual fuzzy preference relation R_i , has only one degree of freedom, denoted by $x_i = r_{12}^i$. Accordingly, the preference originally declared by expert e_i will be denoted s_i .

The iterative process of preference transformation corresponds to the gradient dynamics of a cost function W , depending on both the present and the original network configurations. The value of W combines a measure V of the overall disagreement in the present network configuration and a measure U of the overall change from the original network configuration.

The diffusive interaction between primary nodes i and j is mediated by the interaction coefficient $v_{ij} \in (0,1)$, whereas the inertial interaction between primary node i and the associated secondary node is mediated by the interaction coefficient $u_i \in (0,1)$,

$$v_{ij} = f'((x_i - x_j)^2) \text{ and } u_i = f'((x_i - s_j)^2) \quad (1)$$

The values of the interaction coefficients are given by the derivative of a scaling function f (see Figure 3).

The diffusive component of the network dynamics results from the consensual interaction between each node x_i and the remaining $N - 1$ nodes $x_{j \neq i}$ in the network. The aggregated effect of these $N - 1$ interactions can be represented as a single consensual interaction between node x_i and a virtual node \bar{x}_i containing a particular weighted average of the remaining preference values.

The interaction coefficient $v_i \in (0,1)$ of this aggregated consensual interaction controls the extent to which expert e_i is influenced by the remaining experts in the group.

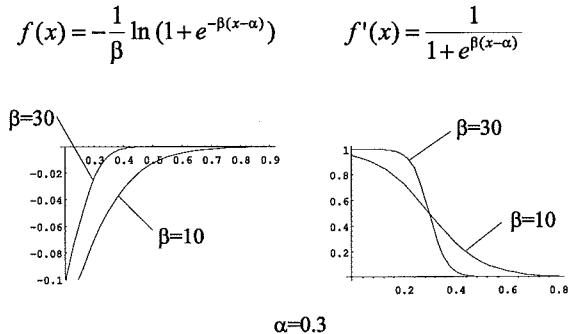


Figure 3. Scaling function f and sigmoid function f'

In our soft consensus model the value v_i , as well as the weighting coefficients $v_i \in (0,1)$ in the definition of \bar{x}_i as given below, depend non-linearly on the standard Euclidean distance between the opinions x_i and x_j ,

$$v_i = \sum_{j \neq i} v_{ij} / (n-1)$$

$$\bar{x}_i = \frac{\sum_{j \neq i} v_{ij} x_j}{\sum_{j \neq i} v_{ij}}$$

The individual disagreement cost $V(i)$ is given by $V(i) = \sum_{j \neq i} V(i, j) / (n-1)$ where $V(i, j) = f((x_i - x_j)^2)$ and the individual opinion changing cost is

$$U(i) = f((x_i - s_j)^2). \quad (2)$$

Summing over the various experts we obtain the collective disagreement cost V and inertial cost U , $V = \frac{1}{4} \sum_i V(i)$ and $U = \frac{1}{2} \sum_i U(i)$, where $1/4$ and $1/2$ are conventional multiplicative factors.

The full cost function W is then

$$W = (1 - \lambda)V + \lambda U \text{ with } 0 \leq \lambda \leq 1. \quad (3)$$

The consensual network dynamics, which can be regarded as an unsupervised learning algorithm, acts on the individual preference x_i through the iterative process

$$x_i \rightarrow x'_i = x_i - \varepsilon \frac{\partial W}{\partial x_i}. \quad (4)$$

We can analyse the effect of the two dynamical components V and U separately. The dissensus cost V induces a non-linear process of diffusion based on the gradient term

$$\frac{\partial V}{\partial x_i} = v_i(x_i - \bar{x}_i). \quad (5)$$

As a result, the iterative step of the non-linear diffusion mechanism corresponds to a convex combination (with sufficiently small ε) between the opinion value x_i and the weighted average \bar{x}_i of the remaining preference values x_j ,

$$x'_i = (1 - \varepsilon v_i)x_i + \varepsilon v_i \bar{x}_i. \quad (6)$$

The inertial cost, on the other hand, leads to a non-linear mechanism which opposes changes from the original opinions x_i , by means of the gradient term

$$\frac{\partial U}{\partial x_i} = u_i(x_i - s_i). \quad (7)$$

The full dynamics associated with the cost function $W = (V + U)/2$ acts iteratively on each decision maker i through convex combinations of the opinion value x_i , the average opinion value \bar{x}_i , and the original opinion value s_i .

$$x'_i = (1 - \varepsilon(v_i + u_i))x_i + \varepsilon v_i \bar{x}_i + \varepsilon u_i s_i.$$

Accordingly, the expert e_i is in dynamical equilibrium, in the sense that $x' = x_i$, if the following stability equation holds,

$$x_i = (v_i \bar{x}_i + u_i s_i) / (v_i + u_i), \quad (8)$$

that is, if the present preference value x_i coincides with an appropriate weighted average of the original preference s_i and the average preference value \bar{x}_i .

3 CHOQUET-BASED VALUATION

In many applications of attack trees, information about attributes is commonly associated to the leaves and one of the main problem to be solved becomes how to promulgate the information up the tree until it reaches the root node. Unfortunately, most of aggregation operators introduced in the literature, e. g. OWA operators (Yager 2008), don't take care of the possible interactions between the nodes. One way to overcome this drawback is to introduce the Choquet integral (Choquet, 1953; Grabisch et al., 2010) whose distinguished feature is to be able to take into account the interaction between nodes, ranging from redundancy (negative interaction) to synergy (positive interaction).

Moreover, the estimation of the attributes' values is usually based on data type depending on subjective judgements, most commonly represented by natural language expressions. Following Zadeh (1978, 1979), here we assume to translate these expressions into the mathematical formalism of possibility measures and to represent the numeric imprecision of attributes' values using unimodal LR fuzzy numbers, as fuzzy subsets of the set of real numbers (Dubois and Prade, 1987).

Definition 1. An unimodal LR fuzzy number A is defined by

$$A(x) = \begin{cases} L\left(\frac{a-x}{a_1}\right) & \text{for } a - a_1 \leq x \leq a, \\ R\left(\frac{x-a}{a_2}\right) & \text{for } a \leq x \leq a + a_2, \\ 0 & \text{else,} \end{cases} \quad (9)$$

where $a \in \mathbb{R}$ is the peak of A , $\alpha > 0$ and $\beta > 0$ are the left and the right spread, respectively, and $L, R: [0, 1] \rightarrow [0, 1]$ are two strictly decreasing continuous shape function such that $L(0) = R(0) = 1$ and $L(1) = R(1) = 0$.

Extending the Choquet integral to a fuzzy domain several forms of information can be handle at the same time, i.e. crisp data, interval values,

fuzzy numbers and linguistic variables (Yang, 2005).

At first, the Choquet integral is defined for a measurable interval-valued function (Aumann, 1965), and then it's extended to fuzzy integrand using the alpha-cuts (Grabisch, 1995).

From now on, we introduce the following notations:

- I the set of interval numbers (rectangular fuzzy numbers)
- $N = \{1, 2, \dots, n\}$ a set of elements
- $F: N \rightarrow I$ an interval-valued function
- $F_L(i)$ and $F_R(i)$ respectively the left end point and the right end point of the interval $F(x)$
- \mathcal{F} the set of all unimodal LR-type fuzzy numbers
- $[{}^L A^\alpha, {}^R A^\alpha]$, the alpha-cut of fuzzy number A
- $\Phi: N \rightarrow F$ a unimodal LR fuzzy-valued function
- \mathcal{F} -tree, an attack tree whose leaves' values are unimodal LR fuzzy numbers

The following definitions are due to Yang (2005):

Definition 2. $F(i)$ is measurable if both $F_L(i)$ and $F_R(i)$ are measurable functions.

Definition 3. The Choquet integral of $F(i)$ with respect to a fuzzy measure μ is defined as

$$\int F d\mu = \{ \int G d\mu \mid G(i) \in F(i) \quad \forall i \in N, \text{ and } G(i) \text{ (measurable)} \}.$$

Definition 4. $\Phi(i)$ is measurable if its alpha-cuts $\Phi^\alpha(i)$ are measurable interval-valued functions for every $\alpha \in (0, 1]$.

Definition 5. Given a measurable fuzzy-valued function $\Phi(i)$ on N and a fuzzy measure μ on 2^N , the Choquet integral of $\Phi(i)$ with respect to μ is defined as

$$\int \Phi d\mu = \bigcup_{0 \leq \alpha \leq 1} \alpha \int \Phi^\alpha d\mu. \quad (10)$$

Accordingly, the calculation of the Choquet integral with a fuzzy-valued function depends on the calculation of the Choquet integral with interval-valued functions, and the following proposition can be proved (Grabisch, 1995).

Proposition 1. Given the measurable interval-valued function Φ^α and the fuzzy measure μ on 2^N , the Choquet integral of Φ^α with respect to μ is

$$\int \Phi^\alpha d\mu = [\int \Phi_L^\alpha d\mu, \int \Phi_R^\alpha d\mu]. \quad (11)$$

Therefore (5.2) becomes

$$\int \Phi d\mu \bigcup_{0 \leq \alpha \leq 1} \alpha = [\int \Phi_L^\alpha d\mu, \int \Phi_R^\alpha d\mu]. \quad (12)$$

Consider now an \mathcal{F} -tree whose leaves' values are unimodal LR fuzzy numbers.

To prove that the root value is still an unimodal LR fuzzy number, we introduce the following

Proposition 2. The Choquet integral of unimodal LR fuzzy numbers is still an unimodal LR fuzzy number.

Proof. A generic unimodal LR fuzzy number A is characterized by an alpha-cut $[{}^L A^\alpha, {}^R A^\alpha]$, where L^α and R^α are strictly monotonic continuous functions (with respect to α).

Consider now a set of unimodal LR fuzzy numbers $\{A_1, \dots, A_k\}$. If we aggregate these fuzzy numbers through Choquet integral with respect to a fuzzy measure μ , we obtain a fuzzy number A characterized by the alpha-cut $[{}^L A^\alpha, {}^R A^\alpha]$, where,

$$\begin{aligned} {}^L A^\alpha &= C_\mu [{}^L A_1^\alpha, \dots, {}^L A_k^\alpha], \\ {}^R A^\alpha &= C_\mu [{}^R A_1^\alpha, \dots, {}^R A_k^\alpha]. \end{aligned}$$

In fact, from the strict monotonicity of the Choquet integral, and given that the lower bound of each alpha-cut is less than the relative upper bound, we have $[{}^L A^\alpha < {}^R A^\alpha]$.

Moreover, if we consider $0 \leq \alpha_1 \leq \alpha_2 \leq 1$ since ${}^L A_1^{\alpha_1} < {}^L A_1^{\alpha_2}$ and ${}^R A_1^{\alpha_1} > {}^R A_1^{\alpha_2} \quad \forall i = 1, \dots, k$, from the strict monotonicity of the Choquet integral we have

$${}^L A^{\alpha_1} < {}^L A^{\alpha_2} \quad {}^R A^{\alpha_1} > {}^R A^{\alpha_2}.$$

Then L^α and R^α are strictly monotonic functions (with respect to α). Moreover, since Choquet integral is a continuous aggregation function, all L_i^α and R_i^α are continuous functions $\forall i = 1, \dots, k$, and the composition of continuous functions is

continuous, then it follows that L^α and R^α are continuous functions (with respect to α).

Then, as an immediate consequence of Prop. 2, starting from the leaves and carrying on a bottom up Choquet aggregation, the obtained tree root's value is again an unimodal (continuous) LR fuzzy number.

The algorithm proceeds as described below. First of all, the alpha-cuts of each unimodal LR fuzzy number in the leaves will be considered, using a suitable grid. The procedure receives the extremes of the alpha-cut, and computes the aggregated value for both the lower and the upper bounds. Increasing the values of alpha in between $[0,1]$, the two computed values form and interval included in the previous ones (for lower value of alpha). Thus the obtained intervals form the alpha-cuts of the fuzzy root, i.e. the required solution.

4 CONCLUSIONS

In this paper, at first we developed a consensual network dynamics aiming at supporting the negotiation process of a group of experts involved in the description of a fraudulent attack through a tree structure.

Secondly, assuming that the leaves of the attack tree are equipped with attribute values represented by LR fuzzy numbers, we propose an algorithm for aggregating these values using the Choquet integral, whose distinguished feature is to be able to take into account the interaction between nodes.

Future work will be devoted to the introduction of case-based reasoning techniques combined with multi criteria models to improve the joint evaluation of risk and uncertainty of the attacks useful for estimating the prevention costs.

REFERENCES

- Aumann R. J. (1965). Integrals of set-valued functions, *Journal of Mathematical Analysis with Applications*, 12, 1-12, 1965.
- Buoni, A., Fedrizzi, M., Mezei, J. (2010). A Delphi-based approach to fraud detection using attack trees and fuzzy numbers. In *Proceeding of the International Association for Scientific Knowledge*. Oviedo, November 8-10. E-Alt & InterTic.
- Buoni, A., Fedrizzi, M., Mezei, J. (2011). Combining attack trees and fuzzy numbers in a multi-agent approach to fraud detection, *International Journal of Electronic Business* (forthcoming).
- Choquet G. (1953). Theory of capacities, *Annales de l'Institut Fourier*, 5, 131-295.
- Chou, C.L-Y, Du, T., Lai, S. V, (2007), Continuous auditing with a multi-agent system. *Decision Support Systems*, 42 (4) 2274-2292.
- Dubois, D. and Prade, H., (1987). Fuzzy numbers: An overview, *Analysis of Fuzzy Information - Vol. I: Mathematics and Logic*, J. Bezdek, ed., CRC Press, Boca Raton, 3-39.
- Kacprzyk, J. and Fedrizzi, M. (1988). A "soft" measure of consensus in the setting of partial (fuzzy) preferences. *European Journal of Operational Research*, 34, 316-325.
- Fedrizzi, M., Fedrizzi, M., and Marques Pereira, R.A. (1999). Soft consensus and network dynamics in group decision making. *International Journal of Intelligent Systems*, 14, 63-77.
- Gordon, T. J. (1994). The Delphi method in futures research methodology. AC/UNU Millenium Project, Washington, AC/UNU.
- Grabisch M., Nguyen H. T., Walker E. A. (1995). *Fundamentals of Uncertainty Calculi, with Applications to Fuzzy Inference*. Kluwer, Boston, MA.
- Grabisch, M., Labreuche (2010). A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid, *Annals of Operations Research*, 175, 247-286.
- Hand, D. J. (2007). Statistical techniques for fraud detection and evaluation. Available at: <http://langtech.jrc.ec.europa.eu/mmdss2007/htdocs/Presentations/Docs/MMDSS_Hand_PUBLIC.pdf>
- KPMG Fraud survey (2009). Available at: <<http://www.kpmginstitutes.com/aci/insights/2009/pdf/kpmg-fraud-survey-2009.pdf>>.
- Oxman, R. (2004). Think-maps: teaching design thinking in design education. *Design Studies*. Vol. 25, Number 1.
- Schneier, B. (1999). Attack trees. Available at: <<http://www.schneier.com/paper-attacktrees-ddjft.html>>.
- Yager, R. R. (2006). OWA trees and their role in security modelling using attack trees. *Information Sciences*, 176, 2933-2959.
- Yang R., Wang Z., Heng P. A., and Leung K. S. (2005). Fuzzy numbers and fuzzification of the Choquet integral, *Fuzzy Sets and Systems*, 153, 95-113.
- Wang, D.G., Li, T., Liu, S.J.L, Liang, G., Zhao, K. (2008). An immune multi-agent system for network intrusion. *Proceedings of the third International Symposium on Intelligence Computation and Applications*, (LNCS 5370, Springer-Verlag Berlin Heidelberg), 436-445.
- Zadeh, L. (1978). Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems* 1, 3-28.
- Zadeh, L. (1979). A theory of approximate reasoning. In Hayes, J., Michie, D., and Mikulich, L., editors, *Machine Intelligence 9*, Halsted Press, New York, 149-194.
- Zhang, L.S., Zhou, N., Wu, J.X. (2008). The fuzzy integrated evaluation of embedded system security. *International Conference on Embedded Software and Systems*, 157-162.

Turku Centre for Computer Science

TUCS Dissertations

118. **Alexey Dudkov**, Chip and Signature Interleaving in DS CDMA Systems
119. **Janne Savela**, Role of Selected Spectral Attributes in the Perception of Synthetic Vowels
120. **Kristian Nybom**, Low-Density Parity-Check Codes for Wireless Datacast Networks
121. **Johanna Tuominen**, Formal Power Analysis of Systems-on-Chip
122. **Teijo Lehtonen**, On Fault Tolerance Methods for Networks-on-Chip
123. **Eeva Suvitie**, On Inner Products Involving Holomorphic Cusp Forms and Maass Forms
124. **Linda Mannila**, Teaching Mathematics and Programming – New Approaches with Empirical Evaluation
125. **Hanna Suominen**, Machine Learning and Clinical Text: Supporting Health Information Flow
126. **Tuomo Saarni**, Segmental Durations of Speech
127. **Johannes Eriksson**, Tool-Supported Invariant-Based Programming
128. **Tero Jokela**, Design and Analysis of Forward Error Control Coding and Signaling for Guaranteeing QoS in Wireless Broadcast Systems
129. **Ville Lukkarila**, On Undecidable Dynamical Properties of Reversible One-Dimensional Cellular Automata
130. **Qaisar Ahmad Malik**, Combining Model-Based Testing and Stepwise Formal Development
131. **Mikko-Jussi Laakso**, Promoting Programming Learning: Engagement, Automatic Assessment with Immediate Feedback in Visualizations
132. **Riikka Vuokko**, A Practice Perspective on Organizational Implementation of Information Technology
133. **Jeanette Heidenberg**, Towards Increased Productivity and Quality in Software Development Using Agile, Lean and Collaborative Approaches
134. **Yong Liu**, Solving the Puzzle of Mobile Learning Adoption
135. **Stina Ojala**, Towards an Integrative Information Society: Studies on Individuality in Speech and Sign
136. **Matteo Brunelli**, Some Advances in Mathematical Models for Preference Relations
137. **Ville Junnila**, On Identifying and Locating-Dominating Codes
138. **Andrzej Mizera**, Methods for Construction and Analysis of Computational Models in Systems Biology. Applications to the Modelling of the Heat Shock Response and the Self-Assembly of Intermediate Filaments.
139. **Csaba Ráduly-Baka**, Algorithmic Solutions for Combinatorial Problems in Resource Management of Manufacturing Environments
140. **Jari Kyngäs**, Solving Challenging Real-World Scheduling Problems
141. **Arho Suominen**, Notes on Emerging Technologies
142. **József Mezei**, A Quantitative View on Fuzzy Numbers
143. **Marta Olszewska**, On the Impact of Rigorous Approaches on the Quality of Development
144. **Antti Airola**, Kernel-Based Ranking: Methods for Learning and Performance Estimation
145. **Aleksi Saarela**, Word Equations and Related Topics: Independence, Decidability and Characterizations
146. **Lasse Bergroth**, Kahden merkkijonon pisimmän yhteisen alijonon ongelma ja sen ratkaiseminen
147. **Thomas Canhao Xu**, Hardware/Software Co-Design for Multicore Architectures
148. **Tuomas Mäkilä**, Software Development Process Modeling – Developers Perspective to Contemporary Modeling Techniques
149. **Shahrokh Nikou**, Opening the Black-Box of IT Artifacts: Looking into Mobile Service Characteristics and Individual Perception
150. **Alessandro Buoni**, Fraud Detection in the Banking Sector: A Multi-Agent Approach

TURKU CENTRE *for* COMPUTER SCIENCE

Joukahaisenkatu 3-5 B, 20520 Turku, Finland | www.tucs.fi



University of Turku

Faculty of Mathematics and Natural Sciences

- Department of Information Technology
- Department of Mathematics and Statistics

Turku School of Economics

- Institute of Information Systems Science



Åbo Akademi University

Division for Natural Sciences and Technology

- Department of Information Technologies

ISBN 978-952-12-2801-8
ISSN 1239-1883

Alessandro Buoni

Alessandro Buoni

Alessandro Buoni

Fraud Detection in the Banking Sector

Fraud Detection in the Banking Sector

Fraud Detection in the Banking Sector: A Multi-Agent Approach