Securing Critical Infrastructures, Grenoble, October 2004

A COMPARISON OF CONVENTIONAL AND ONLINE FRAUD

B. Thomas (*), J. Clergue (*), A. Schaad (*), M. Dacier (+)
(*) SAP Research, Sophia Antipolis, France, (+) Institut Eurécom, Sophia Antipolis, France

1 Introduction

Fraud is a growing problem experienced by most organisations [1] as well as affecting the general public. Impersonation of an individual, using identity information stolen from them, is the fastest growing crime in the UK [2]. There are also other, very different, kinds of fraud. Some forms of fraud that are often omitted from consideration are research misconduct and plagiarisms, but these are also frauds, despite the fact they normally do not lead to prosecution [3].

Over the past decade the use of global communications and the Internet as a means of conducting business has increased in popularity. A study of fraud is needed in light of this new context. The contribution of this paper is to highlight that fraud is likely to be just as prevalent, if not more so, in this new online environment as in the conventional scenario. Furthermore, additional research effort is needed to treat fraud in this environment. It presents significant new challenges, which means that the techniques used to treat conventional cases of fraud cannot easily be applied to it.

In order to study the general phenomena of fraud, it is necessary to consider what the subject consists of. This is explored in Section 2. In Section 2.1 a comparison is made of fraud definitions already in use, and a one-fits-all definition is proposed. Following this, the potential for a fraud taxonomy is discussed, by presenting requirements for taxonomies in Section 2.2 which are then applied to existing taxonomies of fraud in Section 2.3. Section 2.4 proposes a first step towards a general fraud taxonomy. The decision tree presented is over-simplified, but is used to identify general classes of conventional fraud sharing similar properties. Section 2.5 gives examples of these classes. To determine how fraud could be treated in the online context, Section 3 reviews the existing research in fraud prevention and detection. This highlights that although a lot of work has been done, in most cases only one industry sector at a time was considered, without studying the entire domain of fraud as a single phenomenon. The paper then considers how fraud in the online world compares to the conventional cases. Section 4.1 gives examples of online fraud for each of the fraud classes identified by the tree. Section 4.2 shows the benefits of using electronic systems rather than humans to handle fraud. Following this, Section 4.3 presents challenges of using existing fraud techniques in the online scenario. Finally, Section 5 gives a summary and the conclusion of the paper.

2 What is Fraud?

As mentioned above, confusion exists over which acts are covered by the term ‘fraud’. Before a detailed study can begin, the subject must be clearly defined.

2.1) Definitions of Fraud

There is no single definition of fraud that is used universally. Many authors define fraud only within the domain in which they are working: telecoms fraud [4, 5, 6]; academic and research fraud [3, 7]; phishing frauds [8, 9], computer fraud [10].

Telecoms fraud has been categorised into two classes [4, 5]. These are “*swindling*” or “*subscription fraud*” – obtaining and using an account without the intent to pay; and “*impersonation*” or “*superimposed fraud*” – taking over a legitimate user account, so fraudulent activity becomes superimposed over normal user transactions. These two definitions have also been combined to yield one that covers other cases of fraud as well: “*A deliberate act of obtaining access to services and resources by false pretenses with no intention of paying*.” [6]

Academic misconduct is often not considered in fraud studies. Researchers may be pressured to publish regularly and produce results, in order to secure funding for future work, thus they sometimes misrepresent or totally fabricate research [3]. There are many forms of research misconduct: plagiarism; fabrication; falsification; ghost authoring; undisclosed conflicts of interest; with-holding research; misrepresentation about originality [3].

‘Phishing’ definitions are often similar [8, 9], referring to impersonation of legitimate company communications using faked websites, emails and phone calls. The impersonation is used to steal identity information (account numbers, passwords, usernames, etc), which enable fraud for financial gain to be committed with the victim’s identity, for example, to access their online bank account [8, 9].

So-called ‘computer frauds’ have similar properties to other frauds, but the act itself is carried out either by “*input, alteration, deletion or suppression of computer data*” or by *“interference with the functioning of a computer system*” [11].

Some general definitions of fraud state that it leads to an "*economic benefit*" [10, 11], but this is not the case for some of the examples above. In place of a general definition, one source gives definitions of what it considers to be the main classes of fraud [12]. These classes are covered by the dictionary-style definitions used [10, 13, 14, 15, 16, 17], which vary in length and complexity, but all contain the notion that there was an intentional act of deception, used to gain a benefit (for some party). The most succinct definition is adopted here: "*A deception deliberately practiced to secure unfair or unlawful gain*" [13, 14]. In order to apply this definition, some key terms require clarification.

Here, deception refers to:

- The act of deceiving: to be false; to fail to fulfil; to cheat; to ensnare; to cause to accept as true what is false or invalid [15].

Whilst this paper defines unfair or unlawful gain as:

- Any benefit of the deception which the guilty party should not normally have: money or promises of money; unethical gifts or preferential treatment; increased reputation; academic award; services; material goods; etc… In contrast, information which could enable them to gain a benefit is insufficient to demonstrate the party intended to gain.

Under these semantics, it is the belief of this paper that the adopted definition covers the smaller classes of fraud discussed above, along with each of the alternative dictionary definitions. The adopted definition also accounts for "*occupational fraud*" (sometimes called insider fraud) which is defined to be: "*The use of one's occupation for personal enrichment through the deliberate misuse or misapplication of the employing organization's resources or assets*" [1]. Since this definition of occupational fraud does not mention deception, one might ask whether deceit covers misuse of company resources. This paper concludes that it does, since deception, as defined above, includes failing to fulfil, and an employee misusing resources fails to fulfil the expectations their employer has of them.

2.2) Properties of Taxonomies

Since fraud covers such a broad scope, a taxonomy or classification system may be useful to generalise similar properties of existing frauds. Such a system may suggest where cases could be treated similarly, as in the animal kingdom, where hierarchies of species and genera show animals sharing similar genealogy or physiological characteristics. In order to be both useful and correct, taxonomies need certain properties.

Several complimentary studies of requirements for taxonomies have been made [10, 18, 19]. This paper adopts the work of Krsul [18], as it is clear, concise, non-specific to computer security, covers much of what is said by the other authors and gives examples of taxonomies which fail to satisfy the required properties.

Taxonomic characteristics "*… are the properties or characteristics of the objects to be classified*." The required properties for each taxonomic characteristic used are:

"*Objectivity: The features must be identified from the object known and not from the subject knowing. The attribute being measured should be clearly observable."*

"Determinism: There must be a clear procedure that can be followed to extract the feature."

"Repeatability: Several people independently extracting the same feature for the object must agree on the value observed."

"Specificity: The value for the feature must be unique and unambiguous."

Finally, two requirements exist for the usefulness of the taxonomy. It must have explanatory value: through separation or ordering specimens of the domain, generalizations can be made. It must also have predictive value: one can predict the existence of unseen specimens by extrapolating from the known specimens. Specifically, this last property is particularly interesting for fraud, as new schemes regularly appear which were never considered before.

2.3) Taxonomies of Fraud

Having considered what one desires from a taxonomy, existing attempts at fraud taxonomies can be evaluated. Two sources considered a broad enough field of frauds to be candidates for a general fraud taxonomy [10, 12]. One approaches fraud from the criminal investigation viewpoint [12]. Some classes presented there come from a description of the victim: Financial Institution Fraud; Gaming Fraud; Communications Fraud; Utility Fraud; Insurance Fraud; Government Fraud. On the other hand, the class Business Frauds describes who the fraudster is. The remaining classes, Confidence Fraud and Investment Fraud, describe perpetration methods: breaching a relationship of trust; receiving investments under false pretences or stock market manipulation. Such a mixture of classification criteria makes new specimens harder to classify. In addition, the system provides no clear way to classify cases of academic fraud, research misconduct, and misuses of resources. Finally, the method to classify a specimen is not made explicit, so the system lacks determinism.

Another paper creates a taxonomy for computer fraud [10]. It classifies cases using two criteria: perpetration platform (whether the fraudsters acted without authorization or exceeding an authorization they had); and perpetration method (whether the attack involves data manipulation or program manipulation). The taxonomy diagrams given by that paper are presented in Tables 1 and 2. These can be read as classification trees, where the root node is omitted, and moving left-to-right indicates movement towards the leaves. This system satisfies many of the required properties of taxonomies (Section 2.2), but is too specific to electronic methods. For example, in many real-world fraud cases it is problematic to classify even a single step as "*without authorisation (WOA)*" or merely "*exceeding authorisation*". Often one can argue either way. An example would be sending a letter that promises money when some advance fees are sent by the victim: sending a letter is allowed, but the sender may or may not be authorised to promise the monies specified, and they are certainly not authorised to defraud. Another difficulty is how to classify a fraud when the steps mix perpetration platforms, as in the example just given. Due to these difficulties the authors of this paper do not believe a good starting point for a general fraud taxonomy was found.

WOA	Masquerade	Imperso-nation	Password attacks	Guess
				Crack
				Harvest
			Password trafficking	
		Spoofing attacks		
	Vulnerability Exploitation	Software		
		Personnel		
		Communications		
		Physical		
Exceeding Authorization				

Table 1: *Taxonomy of Computer Fraud – Perpetration Platform [10]*

Data	Insert	Improper data	
		Data improperly	
	Improper obtaining or use		
	Integrity attacks		
	Availability attacks		
Program	Run attacks	Without authorization	
		In excess of authorization	
		Improper Parameters	
		Transit attacks	Interruption
			Interception
			Modification
			Fabrication
	Integrity attacks		
	Availability attacks		

Table 2: *Taxonomy of Computer Fraud – Perpetration Method [10]*

2.4) Classifying types of fraud

The goal of this paper is just to identify frauds sharing similar properties. It does not therefore create its own fraud taxonomy, as this would be a major amount of work in itself. Instead, a first step towards the goal is taken by creating a decision tree to classify fraud examples (Figure 1). This has the form the authors expect a taxonomy satisfying the requirements (Section 2.2) to take, but lacks some formal definitions. The tree would need further elaboration of the criteria encapsulated in the decision boxes to become a fully deterministic taxonomy. However, the tree does identify major classes of conventional fraud, by considering criteria which could be useful for fraud prevention. The tree shows that some frauds which do not belong to a single industry sector have similar properties, indicating that frauds previously thought of as different, could be treated in similar ways.

The decision tree (Figure 1) classifies the relationship between victim and fraudster with respect to organisational boundaries. For prevention purposes this is useful, since the form of this relationship indicates how much control the victim can have over the actions of the fraudsters.

Identification of who is a fraudster (F) and who is a victim (V) can sometimes be difficult: several possible role assignments may exist. This paper considers a fraudster to be a party who deceives and gains some benefit, conversely a victim is any party who is deceived; benefit and deception being defined as above. For the sake of simplicity, the tree does not consider whether a fraud is committed by a single party or multiple parties who collaborate. Instead, all parties involved in committing the act are considered together as the fraudster.

The fraudster-victim relationship is hard to define in frauds where multiple victim parties exist. This paper handles these frauds by splitting them to consider only one victim's viewpoint at a time: what they lose, what relation the fraudster has to them, etc. The decision tree can then be applied for each victim independently.

Some of the terms used by internal nodes of the tree require further explanation. An organisation can be: government; business or company; group of related companies; group of companies working in collaboration; public-sector organisation (e.g. health service; benefits agency); charitable organisation. If the victim could be considered to be several of these, for example, a company within a larger group, the smallest organisation covering the greatest number of victims and fraudsters in the fraud is considered to be the victim organisation.

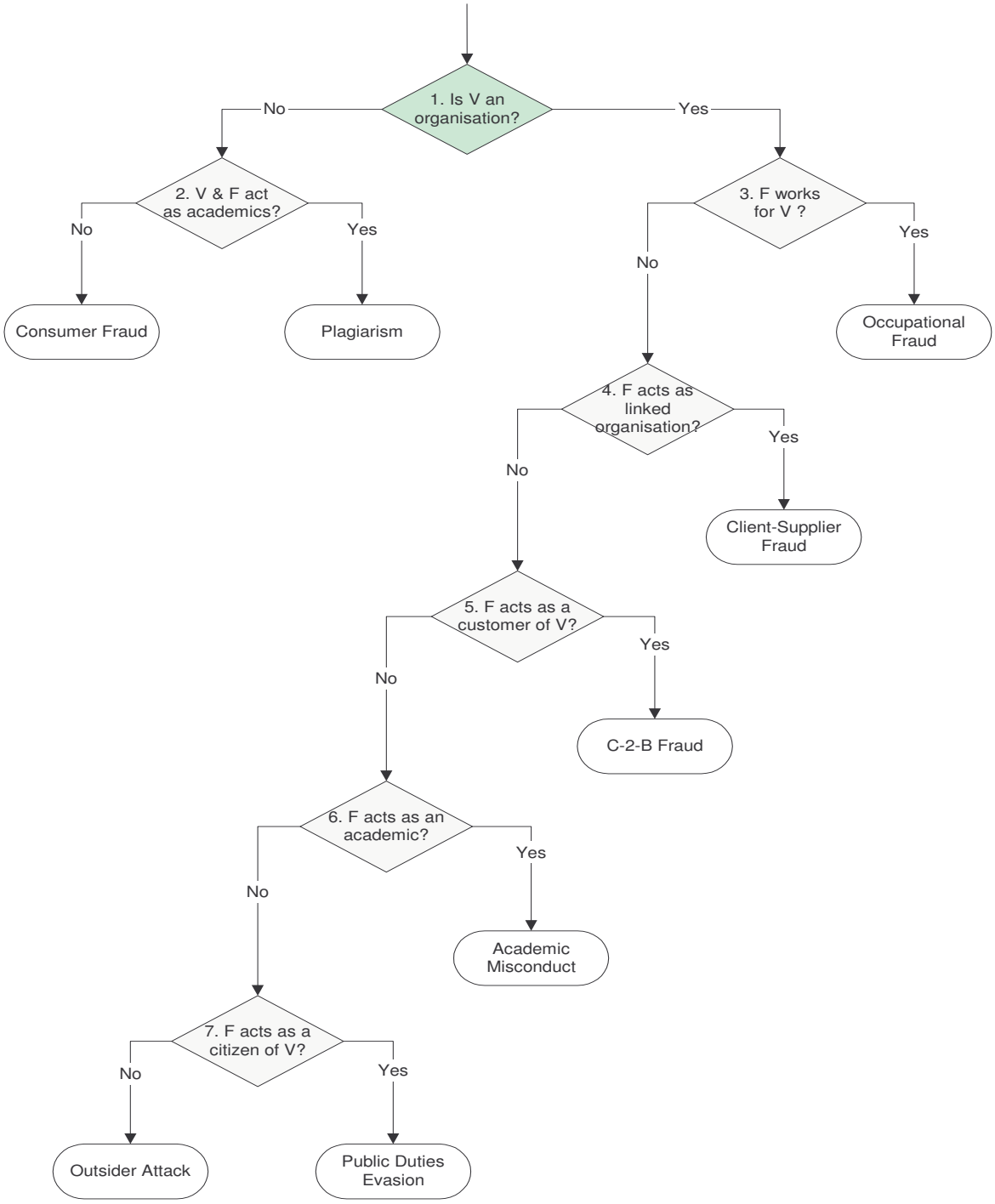

***Figure 1*:** Decision tree to classify fraudster-victim relationship

When someone 'acts as an academic' during a fraud, they either: publish their work; study at an educational institution; lecture at an educational institution.

In the decision nodes, a 'linked organisation' refers to a client or supplier of some other organisation, i.e. there is a contract between them. Someone 'acts as a customer of' a business, if they behave like an individual member of the public purchasing something or using a service as a normal customer would. Academic organisations are those that: publish; organise research conferences; award academic qualifications. Someone 'acts as a citizen of' a government, if they interact with a publicly-funded entity in the collection of duties (taxes and excise duties) or the distribution of welfare benefits (healthcare, unemployment benefits and so on).

The leaves give names to the fraud classes. The terms not discussed previously, or immediately obvious from the tree description, are as follows. 'Consumer fraud' involves a member of the public being deceived. 'Plagiarism' is an academic using someone else's work without crediting them. 'Outsider attacks' refer to frauds suffered by an organisation, perpetrated by a party with whom they have no explicit relationship. They may be attacked by

someone working on behalf of a competitor who wants to damage reputation or steal intellectual property. It could also be an individual who wants to damage reputation, or who gains access to company resources to commit fraud as though they were internal to the organisation.

2.5) Fraud examples

In order to further illustrate the meanings of the tree leaves, the following gives an example of a conventional fraud for each of the classes presented.

Consumer fraud: Art forgery. A New York gallery "*bought authentic paintings, then hired artists to carefully copy them, down to markings on the backs and frames of the canvases*", they then sold the copies as genuine, keeping the originals to sell on at a later date [20].

Plagiarism: A medical researcher copied text from a research application he had reviewed as part of a confidential application review process he was involved in. He used the text without citation to apply for a grant for his own work [21].

Occupational fraud: Skimming. A supermarket branch was installed with six checkout tills. The store manager set up a seventh till, and collected all the proceeds from the till for himself without showing the sales records [22].

Client-supplier fraud: A chip distributor receives an order from a newly set-up computer retail company which has established a line of credit at a known bank. The company, which is phoney, sells the chips on the black market for a fraction of the retail price, closes down and disappears, leaving the distributor out of pocket [23].

C-2-B fraud: Claims fraud. A survey by Direct Line Insurance in the UK showed that 1 in 20 participants had "*falsely claimed that a personal technology item, such as a mobile phone, camera or walkman, was lost or stolen in order to upgrade it with a newer version*" [24].

Public duties evasion: A Nanny, who earned £600 per week, claimed that she was unemployed for a period of 3 years. As a result, she avoided tax, and received Income Support and Housing Benefit (to pay for the rent of her home) in addition to her wage. The loss caused to the UK government was £26,000 [25].

Academic misconduct: In August 1996, London doctors claimed they had re-implanted an ectopic pregnancy, and a baby had been born as a result. The first author was a world famous expert on Ultrasonography in Obstetrics and assistant editor to the journal that published the article. The second author was editor of the same journal and president of the Royal College of Obstetricians and Gynaecologists. Following accusations of fraud, an investigation showed: the patient did not exist; the second author claimed he had never read the article, despite putting his name to it (he said this was common practice); other papers by the first author also contained falsified data [3].

Outsider attack: Intellectual property theft. Time Group, a UK PC company was judged to have "*deliberately and disgracefully*" copied and marketed a design for i-desks (from a separate smaller company) – slim PCs built into desks – having believed that "*[the designer of the i-desk] would [not] have the ability seriously to contest any infringement*" [26].

3 Existing Fraud Prevention and Detection Methods

For the conventional fraud world, much work has researched methods for prevention and detection. To judge if these are applicable to the new online scenario, a full understanding of their contributions is necessary. Fraud recovery is not considered within the scope of this paper, as preventing frauds before they happen and recognising when they occur are the first steps in limiting the damage caused.

3.1) Fraud Prevention

In the initial implementation of bank ATMs (Automatic Teller Machines), the C-2-B and occupational fraud that occurred was often not the result of technical attacks, such as cryptanalysis, but were due to bugs in implementation or incorrect use of the system [27]. For example, ATM keys were split in half, with a different manager to enter each half. Unfortunately, the halves were sometimes given to an engineer to type in, defeating the intended dual control [27, 28]. Bank accounting and general book-keeping systems further protect against fraud with duplicated record-keeping and duty separations: dual control and functional control. One problem is that even when dual control is enforced on system users, interfaces between systems can be single points of failure, which allow breaking the dual control requirement [28]. Smaller-value occupational frauds in banks were mostly committed by opportunists who discovered workarounds in security controls or lax auditing controls (e.g. not auditing change of address) [28]. In all of the high-value occupational fraud cases "*Either the victim's top managers were negligent… or they were the perpetrators*" [28].

Service payment systems which have no return communication channel from the end-user system (e.g. an energy meter) back to the supplier system, presented new challenges for fraud prevention, since effective monitoring of the end-user becomes impossible. An example case is the South African electricity prepayment system [29]. The C-2-B frauds they encountered were prevented in later versions of the system by making meters

more resistant to physical attack, controlling token refunds more strictly, or increasing protection against token replay. Client-supplier frauds committed by token vendors could have been prevented by: restricting vendor access rights to only what they needed; better designed and implemented balancing systems; storing better audit evidence in locations not vulnerable to attack or theft [29]. In addition to the fraud prevention properties already stated above, security requirements engineering of an online-gaming system also found that for client-supplier and C-2-B fraud prevention one should: use book-balancing controls; authenticate both ends of client-server transactions; ensure system faults cannot lead to any kind of benefit for an individual who could cause them [30].

Occupational fraud often occurs in company payroll systems (this is termed payroll fraud), but measures exist to prevent it: ensure proper segregation of payroll preparation, disbursement, and distribution functions; use direct bank transfers instead of cheques; require positive ID if delivering cheques; don't allow some combinations of duplicate data (such as bank details and addresses) in payroll systems; cross-check payroll with other employee records before payment [31]. Other occupational frauds can be prevented by ensuring proper separation of duties between employees: do not leave a single company accountant responsible for all functions regarding supplier payment [32]; do not give the programmer of a bank's computer system access to all live customer accounts [33]. Another effective way of preventing occupational fraud is by deterring employees from committing fraud [1]. Examples of deterrence methods are: creating a work environment in which employees feel both able and obliged to report frauds committed by colleagues; showing that fraudsters are appropriately punished.

3.2) Fraud Detection

Historically, occupational frauds were detected after the event by auditors discovering suspicious records and then investigating, or by complaints from customers which subsequently lead to an investigation. There are common warning signs of occupational fraud for auditors or competent internal accountants: unexpected trends in the account ledgers; suspect supplier addresses (e.g. post office boxes); suspect supplier names; supplier address matching an employee's address; late reporting of deposits; books failing to balance; complaints by customers or suppliers [22, 32]. Regression analysis is the traditional method used to detect trends in figures differing from historic behaviour [34]. Payroll frauds have other 'red flags': payslips without deductions; payroll costs exceed budgeted amount [31]. In general, close scrutiny and balancing of account ledgers is the most effective way to manually detect occupational fraud [22, 31, 32, 33]. However, most occupational frauds are uncovered by tips from other employees [1]. Auditors have legal requirements to detect and report fraud whilst undertaking a financial statement audit, and there is a fraud detection process they should follow which includes: staff interviews; fraud risk analysis; careful unbiased scrutiny of the actions of managers [35].

Many existing systems automatically detect fraud by analysing transaction logs. Most are targeted at C-2-B telecoms fraud, but a wide range of techniques are in use [4, 5, 6, 36, 37, 38, 39]. Thresholding has been used: looking for unexpectedly large values, and applying the same thresholds to all accounts. This was error-prone and inflexible [36]. Many systems track accounts using profiles: statistical summaries of account history. In these systems, comparison of a user's profile with a profile known to be fraudulent and/or the user's historic profiles is used to identify if the user commits a fraud, after each completed transaction [36, 37, 38]. Statistical methods can be used to classify profiles as fraudulent or legitimate [36], as can supervised neural networks, where network training occurs on fraudulent profiles (with fraud cases manually identified) and new customer profiles [37]. Feed-forward neural networks have been compared to statistical methods appropriate for the task: probability density function estimation via Gaussian methods; a Bayesian network model [38]. The advantage of the two statistical methods is that less expert knowledge is required, none for the Gaussian method, and analysis can be done after each transaction, rather than requiring a fixed number of transactions to analyse [38]. Bayesian networks have also been compared to statistical discriminant analysis, regression trees, and classification trees (the last two are data mining techniques) [39]. It is claimed that Bayesian methods could handle larger quantities of data, and are not based on inaccurate statistical assumptions. On the layer above detection systems, expert systems have been used to decide what action to take on an alert: the aim is to minimise wrongful accusation of honest customers [39].

The above systems handle 'Impersonation' telecoms frauds (see Section 2.1), but the 'Swindling' type (another C-2-B case) were also studied [4]. A formal notation was created for these frauds using predicates, along with a means of quantifying an entity's 'intention to deceive' from their actions. A decision algorithm combines this deceiving intention measurement with statistical and transition analysis methods, as seen above, to detect swindlers [4]. Rule-based detection methods for telecoms fraud, which use 'red flags' of fraudulent activity (e.g. international calling increases tenfold) to set off alerts when matched to an account, have been developed which use automatic discovery of the flags [5]. For this task, extensions to existing data mining classification techniques were needed to enable discovery of fraud patterns from historic

account data stored on two levels: the fixed user information data and the constantly varying transaction data [5]. In previous classification techniques the fixed data dominated the varying data, leading to fraud patterns based solely on user data, e.g. a fraudster's name would be a red flag [5].

Research has also considered fraud detection in IP-based (Internet Protocol-based) telecoms services, where the records used in the detection methods above are not available, and several layers of actors provide service: content; payment; network access; etc. A model of fraud intended to be applied to all IP-based services including telecoms, has been developed and applied to online-gaming fraud [6]. The model is too specific to the IP-based client-supplier, C-2-B, and consumer frauds to be useful in the general fraud case, and some of its categorisations could be hard to judge objectively. The same work also considered whether Intrusion Detection Systems (IDS) can be generalised and adapted to create combined Fraud and Intrusion Detection systems, as Fraud Detection Systems (FDS) are a specialisation of IDS. Some components were found to be general enough to be shared between the applications; some would need separate designs for each purpose [6]. It has further been suggested that lessons learned in IDS research can be applied to give more effective FDS, and vice-versa [17].

In tax evasion, a form of public duties evasion, data-mining classification was applied to pick the best combination of organisations to audit, to simultaneously minimise costs and maximise returns [40]. Healthcare insurance claims fraud, a client-supplier fraud where fraudulent claims come from healthcare providers, was for a long time not submitted to large-scale data analysis to detect fraud [41]. Several commercial systems have used a combination of behavioural heuristics (similar to red flags) and basic statistical and knowledge-discovery techniques to detect fraud in this domain [41].

In the financial trading world, managers have historically had problems evaluating financial risks. For example, this allowed the occupational fraud which lead to the Barings Bank collapse. A totally different, agent-based system has been developed for this domain, which monitors traders' activities and warns when risks increase faster than expected [42]. Different autonomous agents populate the system and collaborate to: monitor trades; calculate risks; gather data from internal and external systems; report warning signs to managers and traders. These agents must share a picture of the organisational framework they are working in, and this can be specified using UML (the Unified Modelling Language) [42]. Such a system has the potential to be rapidly adapted for another problem (for example intrusion detection) by changing the UML specification and the data available to the agents.

3.3) Discussion

The above methods almost exclusively consider fraud in one industry sector at a time. It is the belief of this paper that such solutions are not suited to the online business world where: techniques and components are often shared between systems used in diverse situations; most organisations conduct business via the web; component-based developments (reusable components fitted together to create domain-specific solutions) are becoming increasingly popular. Instead, generic approaches, treating fraud across all sectors, would likely reduce the amount of research needed, and increase effectiveness of the methods developed.

4 Fraud in the Online World

As already stated, systems of critical infrastructure are currently moving to a new environment where: many systems are connected by the Internet; increasing numbers of transactions are conducted electronically across organisational boundaries; similar computing components are used across many diverse industries. Consideration should thus be given to how fraud in this online environment compares to cases in the conventional world. One obvious difference is that acts in the online world feel less consequential and more anonymous, for example an employee may think their organisation loses nothing if they misuse the Internet connection and that in any case, the Internet is used so much that it will not be monitored. Thus, they may feel there are greater opportunities to commit fraud undetected.

4.1) Online fraud examples

To show online frauds corresponding to all the previously identified classes of conventional fraud are possible, examples of Internet-based frauds are given below. As before in Section 2.5, an example for each leaf of the decision tree is provided.

Consumer fraud: Buyer-beware fraud. Prior to the release of Microsoft's 'X-Box' a small website offered the product at an attractive price, mostly advertising on Internet bulletin boards. The merchant accepted payments and promised delivery soon after the release date. When the release came and went customers revisited the site to find it was gone, and no trace could be found of the individual who had received payments for goods he had no intention of delivering [43].

Plagiarism: A researcher from Boston Medical Center downloaded a DNA sequence from an Internet database and claimed it to be a sequencing of another gene he was working on [44].

Occupational fraud: A UK survey found that 70% of people have stolen important information from work, e.g. sales contact data, normally to take on to their next job. Most of those who stole

information e-mailed it to a personal account from the workplace [45].

Client-supplier fraud: Imagine an online Lotto service which has tickets sold to customers through a network of retailers, with the main Lotto company being responsible for administering the draw. A retailer store may be able to withhold revenue if the gambling software does not allow the Lotto to be able to verify all transactions. The retailer would report customer wins as normal, but not report some of the customer losses that would have created revenue for the Lotto [6, 30].

C-2-B fraud: Early online shopping cart programs contained vulnerabilities which could allow fraudsters to alter the prices they were paying for products, so they could get products substantially cheaper than they were supposed to [46].

Public duties evasion: The tax collection agency in the UK, the Inland Revenue (http://www.inlandrevenue.gov.uk/) has recently made it possible to fill out many tax claims online. Apart from the conventional understatement of liabilities, the example of online C-2-B fraud above suggests electronic manipulations of the form might also facilitate fraud.

Academic misconduct: Allegations have been made that German researchers downloaded an unrelated picture from the Internet, and then claimed in a publication that it was a photo showing fused tumour cells from their work in a cancer vaccine [47].

Outsider attack: Research published by Nature showed GM (Genetically Modified) pollen could contaminate crops much further away than was previously expected. On the day of publication, a thread on a web forum was started, claiming that the research was biased as the authors were anti-GM activists. Posts grew from here and as a result an Internet petition was started and submitted to Nature who retracted the article in the face of public pressure, saying it should not have been published. A later investigation showed the research was accurate and the people who started the forum thread were working for a GM seed company. Nature was deceived into helping the GM seed company [48].

The examples above show that the major classes of conventional fraud are found in the online world: either examples are known or clear opportunities exist which criminals have not yet exploited.

4.2) Benefits Brought by Electronic Systems

When treating online fraud, there are benefits of working in an environment of electronic systems. Much existing computer security research has considered controlling access to electronic resources (access control). Access policies can be formally defined, enabling the construction of formal proofs that the protocols have the intended properties. This gives confidence that correct implementations of the policies (verified as correct) will always enforce the required access control [28]. In contrast to humans, computer programs are deterministic, social pressures do not apply, so computers will always apply the necessary checks, unlike a train conductor, for example, who may not check all tickets because he is too tired or does not want to annoy people he believes he has already checked. Additionally, computers are faster at checking and not prone to boredom when doing repetitive tasks. This means computer checks can occur more reliably than human checks. Checking of a computer system can often be done without user interaction, as in most office logging tools, so checking need not inconvenience users or hinder them in carrying out their normal business activities.

In addition to this, the general achievements in security engineering have helped to aid the prevention and detection of fraud [28]. Cryptographic primitives exist to aid enforcement of key security properties. Strong authentication techniques increase confidence that the entity on the other side of a communications link is the intended party. Encryption and decryption systems make it possible for information to be obscured, meaning it can be kept confidential from parties who should not have access; it can then be sent across public communications channels knowing that anyone intercepting the message will have difficulty reading it. Message authentication checks can ensure that a message has not been altered in transit. Furthermore, certain protocols give guarantees that a party can not deny having done something later (non-repudiation), e.g. they can not deny having accepted an agreement, or having received and read a message. More details of these properties are discussed in [28].

4.3) Challenges for Treatment of Online Fraud

Having seen that a similar scope exists for fraud in the online world, one should ask whether conventional fraud treatment methods can be adapted and reapplied. Unfortunately, fraud treatment in the online world faces many new challenges, which highlight that new solutions are needed in this domain. As an eye-opener, 10 examples of the difficulties are given below, but this is by no means an exhaustive list.

1 Anonymity: Anonymity is a property often required in web transactions, either due to public pressure or privacy legislation. This means detecting fraud can not rely on detecting identity.

2 Ubiquity: A traditional ‘red flag’ for credit card frauds used to be the same card being used in geographically separated locations at the same time.

For example, transactions in Tokyo and New York within a short space of time would normally lead to the card being suspended. However, with online retailers, it becomes possible to buy from companies in these two locations at almost the same time, thus the old detection paradigm no longer applies. An additional problem of ubiquity is that a party can appear to be several different people at the same time, for example they can login to IRC, chat rooms, web forums, and so on with multiple different usernames. This allows a single person to appear to be a whole crowd of people, so they can apply social pressure to create the right environment for a fraud by themselves (e.g. reporting a 'great investment opportunity' that many people seem to be willing to engage in).

3 Speed: With the advent of web-based services, decisions happen much faster due to the speed of computer processing and the speed at which web business can move. It is often the case in Internet business that decisions need to be taken quickly: opportunities are short-lived. In addition people do not appear to retain the same psychological link between action and consequence. For example, some people will always click 'OK' on a popup window to get rid of it, they may not realise they have just downloaded and installed a malicious program onto their computer, or agreed to pay someone money. In some cases fraud is now just a click away.

Aside from this, electronic communication is fast, making it possible for a fraudster to contact many potential victims in an instant. Even if the more wary targets notice a scam they will not have time to communicate a warning to the others before it is too late. Another speed-related problem is that a weakness allowing a system to be compromised can often be exploited quickly. The concept of a 'flash worm', which could compromise a million machines in a second, is described in a paper published at WORM 2004 [50]. Thus, a fraudster can quickly compromise a large number of web servers which are trusted by the public (their usual news sites, entertainment providers, web mail clients, etc) to set up an environment to lure victims: altering or inserting news stories; adding commands to HTML pages to automatically download and install some malicious code when read.

4 Parallelism: Electronic systems offer parallelism. This is either true parallelism, through using many systems at once, or concurrency (logical parallelism) by multi-tasking many threads of execution on the same processor. A fraudster can thus run the same fraud against many victims in parallel, giving a greater chance they will succeed in some cases. He can also attack one victim in parallel with a large number of attacking systems, as has happened in the Distributed Denial of Service attacks (DDoS) which have been used to make the content of certain web hosts inaccessible. Also, each of his attacking systems can target further attacks at different victims, as is the case in worm propagation. A single fraudster can now have the attacking power he would previously only have had with many collaborators.

5 CPU Speed and Parallelism: With the CPU speed available today, and the ability to run tasks in parallel, a fraudster can run many frauds with small inconsequential gains (e.g. a few cents) which will mount up very quickly without a long period of exposure. He can disappear before any criminal investigation begins.

6 The Environment is Your Enemy: In the online world, it is hard to know what is trustworthy. For example, when an advert appears in a popup window on a website the user trusts, they may believe it has been reviewed by the site and approved as 'safe', but this is often not the case. Similarly, they may be tempted to open an e-mail sent by a known party, but many of the first e-mail viruses spread by sending themselves to everyone in an infected party's address book. There are also many points in the communications infrastructure that can be attacked: DNS attacks mean typing a trusted web address can redirect the browser to a totally different server; routing of communications crosses many unknown systems which can be attacked to redirect or intercept traffic; communications components can be replaced for non-trustworthy ones without noticeable effect; WiFi access points can be impersonated.

The stealth with which information can be collected has also increased. For example, communications can be listened to without either communicator realising. This can lead directly to fraud, for example if someone gives out their credit card number in an unencrypted channel. Alternatively, information capture can lead indirectly to a benefit for a fraudster. For example, if he witnesses an IP address from within a company accessing a pornographic website, he can communicate to the originator and extort money from them threatening that, "all e-mail addresses in your domain will be mailed with the site address and your IP so they will find you and punish you".

7 Fraud Cost vs. Benefit has Changed: Unfortunately, with the advent of the extremely cheap communications offered by the Internet, which means, for example, that it costs the same to send an e-mail to 1 person as to 100, the fraudster now has to invest less (often just an Internet connection and time) to commit a fraud yielding the same or greater benefit as before. This can be witnessed through the prolific nature of spam (unsolicited emails).

8 Available Victims: The Internet connects huge numbers of people located all over the world. This means everyone is a potential victim of fraud, making it hard to warn them effectively. In addition, frauds can cross jurisdictional and geographical boundaries, so different laws and regulations apply to

different parties, and investigation requires international cooperation. Finally, people who use websites may do so only once, never to come back. This means there is less scope for profiling customers to say whether a request comes from a legitimate customer or not.

9 The 'Slashdot Effect': When a site is reviewed by a popular site such as Slashdot (http://slashdot.org/), it can receive a flood of legitimate requests as a result. "*Often so many people follow up a mention by clicking the associated link that a page is impossible to reach.*" [49]. These kinds of phenomena may look like malicious attacks or frauds, but a victim must understand the cause of such an event, be it malicious or legitimate, in order to deal with it appropriately. Whilst research in fields like marketing, criminology and sociology aid understanding of these events in the conventional world, there has so far been less research into how individuals behave online.

10 Schizophrenia: On the Internet, computers act as representatives for humans. Users must trust their PCs, which carry out actions on their behalf which they themselves would not be capable of. An analogy is the trust one must give to an interpreter when visiting a foreign bank. However, one does not know if a machine is compromised or what it is doing in the background, and can not understand or even view the signals the PC emits. Therefore, humans enter a dangerous situation when they give trust to their machines.

5 Summary and Conclusion

This paper attempts to provide a comparison of conventional fraud and online fraud. In particular, in sections 2, 3, and 4 this paper has covered:

- Definitions of fraud.
- The need for fraud taxonomies and existing attempts.
- The identification of classes of conventional frauds with similar properties.
- Existing research in fraud prevention and detection.
- The potential for online fraud.
- Benefits of using electronic systems to treat fraud.
- Challenges for treating fraud in the online world.

This paper concludes that there is a vast scope for fraud in the conventional world, and a similar scope for fraud in the online world. In addition, the fraud treatment methods found have not yet considered the entire fraud problem, but instead focussed on a single industry sector at a time. These methods have also not given much consideration to the online domain. Certain properties of the online domain mean the existing fraud treatment methods cannot be re-used there. Thus, additional research is needed to secure critical infrastructures against fraud. Systems of critical infrastructure which work in the online world have no reliable protection against fraud at this point in time. Although security techniques exist which could be used as a starting point to protect against fraud, time must be invested before systems which span different kinds of businesses, or which act in the online world, can be secured.

References

[1] Association of Certified Fraud Examiners, *2004 Report to the Nation on Occupational Fraud and Abuse*, http://www.cfenet.com/pdfs/2004RttN.pdf visited July 9th 2004.

[2] Leyden, J., "Fraudsters Prey on Apathetic Brits", *The Register*, March 18th 2004 http://www.theregister.co.uk/2004/03/18/fraudsters_prey_on_apathetic_brits/ visited July 21st, 2004.

[3] Smith, R. (Editor of the British Medical Journal), *Fraud in Medical Research*, talk given September 2001. http://bmj.bmjjournals.com/talks/fraud.ppt visited July 6th 2004.

[4] Bhargava, B., Y. Lu and Y. Zhong, "Fraud Formalization and Detection", *Proc. of Data Warehousing and Knowledge Discovery Conference (DaWaK)*, September 2003.

[5] Rosset, S., U. Murad, E. Neumann, Y. Idan, and G. Pinkas, "Discovery of Fraud Rules for Telecommunications – Challenges and Solutions", *Proc. of ACM KDD-99 (Knowledge Discovery In Databases)*, San Diego, USA, 1999, pp 409-413.

[6] Kvarnstrom, H., E. Lundin and E. Johnsson, "Combining Fraud and Intrusion Detection – Meeting New Requirements", *Proc. of 5th Nordic Workshop on Secure IT Systems (NordSec2000)*, October 2000.

[7] Decker, C. and C. Burgess, *A Taxonomy of Plagiarism and Academic Fraud*, July 6th 2003. http://whorfin.ucr.edu/~curt/plagiarism/tax.html visited July 12th 2004.

[8] The Anti-Phishing Working Group, http://www.antiphishing.org/ visited July 7th 2004.

[9] Financial Services Technology Consortium, *Project Proposal: FSTC Counter-Phishing Initiative*, http://www.antiphishing.org/FSTC_Phishing_Prospectus_Final.pdf visited July 6th 2004.

[10] Vasiu, L. and I. Vasiu, "Dissecting Computer Fraud: From Definitional Issues to a Taxonomy", *Proc. of the 37th Hawaii International Conference on System Sciences – 2004 (IEEE)*, Hawaii, 2004.

[11] Council of Europe, *Final Draft Convention on Cyber-crime (2001)*, http://conventions.coe.int/Treaty/en/Treaties/Html/185.htm visited July 7th 2004.

[12] National White Collar Crime Center (NW3C), Federal Bureau of Investigation, *IFCC report 2002*, December 31st 2002. http://www.nw3c.org/downloads/data_trends_report2002.pdf visited July 6th, 2004.

[13] Commerce-Database.com, *Fraud Definition*, http://www.commerce-database.com/legal-terms/fraud.htm visited July 12th 2004.

[14] Legal Definitions.com, *Fraud*, http://www.legal-definitions.com/fraud.htm visited July 12th 2004.

[15] Merriam-Webster, *Merriam-Webster Online Dictionary*, http://www.m-w.com/cgi-bin/dictionary?book=Dictionary&va=deception visited July 12th 2004.

[16] Cambridge University Press, *Cambridge Advanced Learner's Dictionary*, http://dictionary.cambridge.org/define.asp?key=20050&dict=CALD visited July 12th 2004.

[17] Lundin, E., *Aspects of employing fraud and intrusion detection systems*, Thesis for Licentiate of Engineering, Department of Computer Engineering, Chalmers University of Technology, Göteburg, Sweden, 2002.

[18] Krsul, I.V., *Software Vulnerability Analysis*, DPhil Thesis, Purdue University, May 1998, https://www.cerias.purdue.edu/tools_and_resources/bibtex_archive/archive/98-09.pdf visited July 12th 2004.

[19] Lough, D.L., *A Taxonomy of Computer Attacks with Applications to Wireless Networks*, DPhil Thesis in Computer Engineering, Virginia Polytechnic Institute and State University, April 2001, http://scholar.lib.vt.edu/theses/available/etd-04252001-234145/unrestricted/lough.dissertation.pdf visited July 12th 2004.

[20] *The Associated Press*, "Two Charged in Fake-Art Scheme", June 15th 2004, http://www.msnbc.msn.com/id/5219352/ visited August 4th 2004.

[21] Dept of Health and Human Services, "Findings of Scientific Misconduct", *The Federal Register*, 66:237, December 10th 2001, pp 63709, http://ori.dhhs.gov/html/misconduct/xiong.asp visited August 11th 2004.

[22] Wells, J.T., *Enemies Within*, http://www.cfenet.com/resources/Articles/ViewArticle.asp?ArticleID=13 visited July 1st 2004.

[23] Wells, J.T., "The World's Dumbest Fraudsters", *Journal of Accountancy*, May 2003. http://www.cfenet.com/resources/Articles/ViewArticle.asp?ArticleID=38 visited August 3rd 2004.

[24] Leyden, J., *Brits Cheat Insurance to Get Gadget Upgrades*, The Register, April 15th 2004, http://www.theregister.co.uk/2004/04/15/insurance_fraud_rife/ visited August 3rd 2004.

[25] BBC News Online, *Benefit Fraud Nanny Facing Prison*, April 2nd 2004. http://news.bbc.co.uk/1/hi/england/nottinghamshire/3595211.stm visited August 3rd 2004.

[26] Smith, T., *Court Slaps Time Group for Ripping Off PC-in-a-Desk Design*, The Register, December 23rd 2003, http://www.theregister.co.uk/2003/12/24/court_slaps_time_group/ visited August 3rd, 2004.

[27] Anderson, R.J., "Why Cryptosystems Fail", *Proc. of ACM 1st Conference in Computer and Communications Security*, Virginia, November 1993. Available from http://www.cl.cam.ac.uk/users/rja14/ visited July 12th 2004.

[28] Anderson, R.J., *Security Engineering*, Wiley, 2001, ISBN 0471389226.

[29] Anderson, R.J. and S.J. Bezuidenhoudt, *On the Reliability of Electronic Payment Systems*, 1996. http://www.cl.cam.ac.uk/ftp/users/rja14/meters.ps.gz visited June 30th 2004.

[30] Anderson, R.J., *How to Cheat at the Lottery*, 1999, http://www.cl.cam.ac.uk/~rja14/lottery/lottery.html visited June 30th 2004.

[31] Wells, J.T., *Keep Ghosts Off the Payroll*, http://www.cfenet.com/resources/Articles/ViewArticle.asp?ArticleID=21 visited July 1st 2004.

[32] Wells, J.T., *Protect Small Businesses*, http://www.cfenet.com/resources/Articles/ViewArticle.asp?ArticleID=27 visited July 1st 2004.

[33] Wells, J.T., *Lapping It Up*, http://www.cfenet.com/resources/Articles/ViewArticle.asp?ArticleID=22 visited July 1st 2004.

[34] Mercer, L.C.J., "Fraud Detection via Regression Analysis", *Computers & Security*, 9:4, pp 331-388, 1990.

[35] Pany, K. and R. Whittingham, *Fraud in a Financial Statement Audit: What Every Auditing Student Should Know About SAS No. 99*, The American Institute of Certified Public Accountants (www.aicpa.org).

[36] Cahill, M., D. Lambert, J. Pinheiro and D. Sun, "Detecting Fraud In The Real World", *Handbook of Massive Datasets, J. Abello et al. (Editors)*, Kluwer Academic Publishers, 2002, pp 911-929. http://cm.bell-labs.com/cm/ms/departments/sia/doc/hmds.pdf visited August 3rd 2004.

[37] Moreau, Y. and J. Vandewalle "Detecting Fraud in Mobile Communications using Supervised Neural Networks", *Proc. of SNN '97: Europe's Best Neural Networks Practice*, April 1997.

[38] Taniguchi, M., M. Haft, J. Hollman, and V. Tresp, "Fraud detection in communications networks using neural and probabilistic methods", *Proc. of 1998 IEEE Int. Conf. in Acoustics, Speech and Signal Processing (ICASSP'98)*, Vol. 2, pp 1241-1244.

[39] Ezawa, K.J. and T. Schuermann, "Fraud/Uncollectible Debt Detection Using a Bayesian Network Based Learning System", *Proc. of Eleventh Conf. on Uncertainty in Artificial Intelligence*, Montreal, August 1995. pp 157-166.

[40] Bonchi, F., F. Giannotti, G. Mainetto and D. Pedreschi, "A Classification-Based Methodology for Planning Audit Strategies in Fraud Detection", *Proc. of ACM KDD-99 (Knowledge Discovery In Databases)*, San Diego, California, 1999.

[41] Major, J.A. and D.R. Riedlinger, "EFD: A Hybrid Knowledge/Statistical-Based System for the Detection of Fraud", *The Journal of Risk and Insurance*, 2002, 69:3, pp 309-324.

[42] Wang, H., J. Mylopoulos and S. Liao, "Intelligent Agents and Financial Risk Management Systems", *Communications of the ACM*, 45:3, March 2002, pp 83-88.

[43] NACHA – The Electronic Payments Association, *Internet Payments Fraud: A Primer for Merchants and Financial Institutions*, The Internet Council, February 3rd 2003.

[44] Rhoades, L.J., "Dept of Health and Human Sciences, Findings of Misconduct", *Federal Register*, August 6th, 2003 68:151, pp 46642, http://ori.dhhs.gov/html/misconduct/karunakaran.asp visited August 4th 2004.

[45] BBC News Online, *Firms Become Digital Detectives*, April 16th 2004. http://news.bbc.co.uk/1/hi/technology/3626205.stm visited August 3rd 2004.

[46] ISS E-Security Alert, *Form Tampering Vulnerabilities in Several Web-Based Shopping Cart Applications*, February 1st 2000, http://xforce.iss.net/xforce/alerts/id/advise42 visited August 3rd 2004.

[47] Office of Research Integrity (U.S.), *German University Investigating Allegations Against Cancer Researchers*, September 2001, 9:4, http://ori.dhhs.gov/html/resources/germanuniv.asp visited August 4th 2004.

[48] Kaminsky, D., "Méthodologie d'analyse d'attaques informationnelles à base de cas réels", *Proc. of Symposium sur la Sécurité des Technologies de l'Information et des Communications (SSTIC 2004)*, Rennes, June 2004.

[49] BBC News Online, *The Dangers of Having a Good Idea*, May 5th 2003, http://news.bbc.co.uk/1/hi/technology/2995343.stm visited August 4th 2004.

[50] Staniford, S., D. Moore, V. Paxson and N. Weaver, "The Top Speed of a Flash Worm", *Proc. of WORM04*, San Diego, USA, October 2004.