The cybercrime process: an overview of scientific challenges and methods

Patrick LALLEMENT

Charles Delaunay Institute (ICD), CNRS Joint Research Unit « Sciences & Technologies for Risk Management » (UMR 6279), Université de Technologie de Troyes (UTT), 12 rue Marie Curie, CS 42060, 10004 Troyes Cedex

lallement@utt.fr

Abstract – The aim of this article is to describe the cybercrime process and to identify all issues that appear at the different steps, between the detection of incident to the final report that must be exploitable for a judge. It is to identify at all steps, issues and methods to address them.

Keywords: cybercrime, detection, forensic analysis

1. The cybercrime process

1.1 Definition of cybercrime

The cybercrime is defined in the penal law as a set of malicious acts that are committed against information systems or that make use of information and communication technologies. In the first subset we can class denial of services (DoS) attacks, theft or falsification of data. The second subset concerns fraud, child pornography, sexual harassment by the way of internet and all logistic support activities of organized criminalities. In France, the cyber criminality is not defined as a whole science, neither a field, there is no laboratory devoted to this transversal domain. However, the forensic process whose aim is to collect and process digital evidences raises different issues because systems are more and more complex and criminal strategies are continuously changing. The cybercrime field is generally viewed as an application domain for many communities concerned by information or data processing, decision-making aid, detection methods, sociology, networking, etc. That is why main issues are generally addressed in a fragmented way.

1.2 Cybercrime vs. security

The cybercrime process is initiated when the detection function identifies a situation or event as abnormal referring to the assumed security level and the security policy. This detection function can be executed in reactive mode as control function of a system or in a proactive mode by law enforcement actions such as internet flow or social networks supervision to look for suspicious contents. In the first case, the abnormal event (or state) is detected by processing control variables; in the second case, the nature of application contents carried though networks may alert about an illegal activities. The figure 1 shows out the cybercrime process regarding to the security process in the case of any information system. The detection function can be considered as common to both processes although the case of APT malwares (Advanced Persistent Threats), where investigations and system recovery take a long time and should be processed commonly and as to avoid alerting the intruding malware itself. The detection function must detect an abnormal situation and qualify it as malicious or not. If then, it generates an alerting event. Figure 2 present the different functions that will succeed to the detection and take place in the cybercrime process: investigation (collecting of clues, qualification of evidences), forensic analysis, argumentation and final reporting. The digital evidence has to be built from data collected on the (cyber) crime scene [1]. The digital evidence must show out a link between an attacker and a victim [2]. As consequence of their digital aspects, they may be heterogeneous, altered, uncertain and corrupted [3]. They have to be analyzed, interpreted and documented by forensic examiners such as they can be reliable and relevant to draw their conclusions for a court.

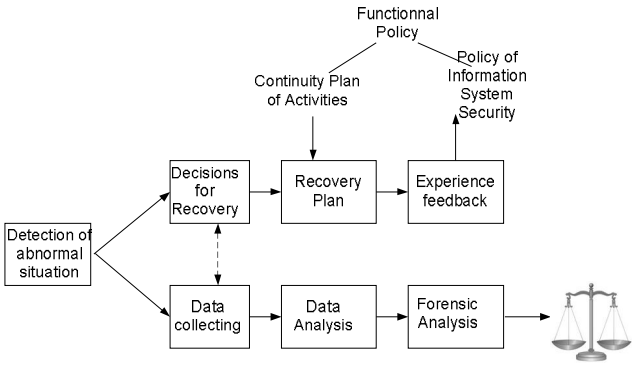

Fig. 1: Cybercrime process vs. security process

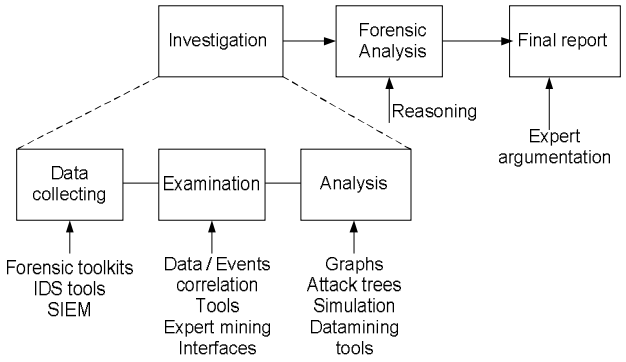

Fig. 2: Cybercrime process

2. Detection

Detection comes within the competence of forensic experts, but generally, it cannot be performed humanly, due to the complexity of the system under control, the volume of information to process, the velocity of some attacks, their critical aspects and the predefined scenario that they can use to run. An automatic detection system is then relevant to cope with many of these challenges to react with a more or less expert way, a more or less proactive scheme; they can only generate burglar alarms (less) or be able to characterize and counter the malicious event (more).

2.1 Intrusion detection

It is made with Intrusion Detection Systems (IDS) [4]. The IDS objective is to detect an abnormal state and to qualify it as intrusive state or not, and if so, to trigger an alert. IDS can report alert in IDMEF format [5] based on XML syntax, which can be useful to organize a cooperative surveillance for distributed systems, and perform alerts correlations. This is of greater importance as countermeasure against some attacks which runs according to a pre-defined scenario. In this case, an IDS-level detection can be completed with a real-time layer to predict what is happening and prevent some severe attacks (distributed DoS, rootkits, worms) that can spread using a slow and/or sophisticated propagation scheme (botnets, rootkits). This layer developed for IPS (Intrusion Prevention Systems) requires reasoning methods at a global level as for IPS. They use generally Bayesian approaches and variables from the network. Bayesian networks offer a powerful way for modeling, representing and reasoning with complex information and have been proposed to process alert correlations systems [6]. For large systems, this reasoning layer can represent a fast mining challenge, using complex time-stamped events [7].

Detection can also be performed in a signal mode, using statistical approaches that present the advantage to avoid a priori knowledge (comparing with Bayesian approaches) [8]. The variables used are: the traffic rate, abnormal packets, CPU utilization, etc. A likelihood function can be built and the challenge is to minimize the false positive and the false negative ratio.

Concerning intrusion detection, the remaining issue lies in the description of complex situations; a language has been proposed by the ANR PLACID project (2007-2011) with the Intrusion Detection Description Logic (IDDl) [9] to describe intrusions, it is IDMEF-compatible, but it is limited to handle information about alerts, topology and vulnerabilities.

2.2 Fraud detection

The detection challenge can be assimilated as a classification problem between legitimate and fraudulent transactions. The methods used can be supervised or not. In supervised mode, models request learning to distinguish between legitimate and fraudulent transactions. Because fraudulent ones are less frequent (< 1%) than the others, they are worse learned and therefore, the classification quality is decreased. Artificial neural networks have been largely proposed in the 90s, Support Vector Machine (SVM) [10] [11] but their efficiency is largely depending on the type of transaction considered. In [12] authors have compared the different classification methods among various applications. More recent works have suggested a fusion approach with different methods, to filter the current transactions with a level of suspicion, to use the Dempster-Shafer theory to quantify an overall belief for a transaction, to use history and a Bayesian learner to classify suspicious transactions [13].

In non-supervised mode the learner doesn't use any a priori class. It must be designed to the specific context: insurances, payment, telecoms… Methods proposed are based on graphs, decision trees, neural networks, fuzzy rules.

Some works suggest a combination of supervised and non-supervised approaches. In addition to the unit fraud detection problem, a correlation between them may be necessary to identify organized group frauds. The more recent techniques aim to integrate business rules and social networks data.

2.3 Suspicious content detection

2.3.1 Steganalysis

It makes reference to data dissimulated behind a legal flow (voice, video). Detection of hidden data remains a difficult issue because it exploits opportunities given by the coding techniques. Detection methods in signal mode have been proposed [14] [15].

2.3.2 Peer to peer networks (P2P)

Content analysis in network to detect malicious activities will concern the internet in general but more precisely social networks and P2P exchanges. In this case, the detection problem becomes rather an identification problem (of paedophile activity for ex.). Data to examine are of a huge volume, they are also dynamic and in the case of P2P networks, there is no central authority. A random and not computer-aided flow inspection is not possible because a large amount of data is necessary to build an evidence of illegal activities. Moreover and at the difference to the previously mentioned detection methods, no history is there available because illegal behaviours always try to be undetectable by using encrypting tools or specific key-words. Neither learning nor statistical methods are efficient here. Approaches proposed are rather inference based on expert (law enforcement)-defined rules to detect and process queries [16]. IP addresses are relied to UDP flows to identify the users. Nevertheless, the computer-aided and automated tools to state that a given user is for ex. a paedophile stands a legal problem, the expert should always have the last word and automated tools should be viewed as processing resource to cope with the large amount of data.

2.3.3 Social networks

The detection issue is there doubled: it is to detect communities on micro-blogging platforms and then to detect specific breaches in violation of citizen protection (fraud, illegal content dissemination, attack to underage, etc.). To help law enforcement people, a processing chain must then associate, in detection and investigation modes, the content analysis of publications and conversations and also the analysis of relations between actors, while it could capture knowledge about the structure, the behaviour and the practices of criminals. Social networks pose complex issues concerning contents and network analysis and also a visualization challenge. It is due to the large amount of data to process (for ex. 465 Millions of tweeter accounts, 175 Millions of tweets per day), the velocity (< 1 minute) and the variety of data (structured and not-structured).

Annotation of texts from social networks is difficult, due to the flow processing and to the downgraded linguistic nature of messages and conversations, which are also multi-languages and multi-domains. Approaches used are rather symbolic, statistic, but the most promising seems to be the mix of them.

The networks analysis uses graph-based representations. There is no consensus to describe and quantify the dynamic of graphs, and to describe how the information does propagate along them. Several works have studied how to retrieve comprehensive information from the structure of static cyber-communities from complex networks [17]. The identification issue for dynamic communities is now addressed by two ways: 1) a dynamic graph can be viewed as a succession of static graphs, each of them representing a state of the dynamic graph at a given moment. In each static graph (i.e. at each time) it is the possible to determine communities with more or less independencies. It is then necessary to retrieve correspondences between communities along the time to restore the temporal evolution. More complex rules have to be defined to identify fusion, scission, appearance and extinction of communities [18] [19] [20]. 2) Specific algorithms have to be designed to detect communities in dynamic graphs [21] [22].

The social networks analysis methods are borrowed from the graph theory and are completed with many works about data and text-mining to process data extracted from social networks messages and relations, indicators and aggregates computed from social graphs and the dynamics of exchanges. Techniques developed recently from the "pervasive computing" domain give interesting perspectives [23]. In this frame, social networks users are viewed as "sensors" that give information about its environment. New sensors can then enhance the already existing sensors. The more recent works suggest to use data-fusion techniques, Complex Event Processing (CEP) engines, time-sequences association and analysis, spatiotemporal patterns to detect events (alarm reporting, weak signals).

2.4 Synthesis

In all cases, the detection issues have to cope with a large amount of heterogeneous data. In some cases, there are also serious time constraints. Most of methods proposed are similar to those for decision-making. Indeed, the reasoning associated to detection consist in doing classification between normal, abnormal and suspicious cases, and then to decide if the suspicious case is normal or not, using supplementary data such as history, learning techniques and quantitative methods developed in the artificial intelligence field (fuzzy rules, neural networks). Most of them are used to detect intrusion or frauds. Methods that are designed with a generic approach are rare [24], probably because specific information (contexts, experience, behaviours) is necessary to reduce false positive and false negative ratios.

3. Investigations

The response to incident process can be split in several steps: data-gathering, examination, analysis, reporting [25] [26]. Data are of different volatility as defined in [27], from very volatile (network traffic, RAM) to persistent (logs files), they may be heterogeneous in terms of sources (network, system), format, uncertain (incomplete, unclear), not structured (rough data), encrypted. They may also have been falsified. In many cases, they represent a large amount of data to process, i.e. exceed the human processing in a limited time. The main challenge for first responders and analyzers is to assure evidence

conditioning and to keep track of all operations they have done.

3.1 Data collecting and forensic analysis of terminals

One must distinguish tools that can only collect data and those which can also process and analyze them at a first level. There are toolkits from markets that enable to collect digital evidences from the computer (RAM, DISK) while respecting advices for it [28]. The use of market-standardized tools provides generally more guarantees about their reliability and the integrity of data collected (comparing to ad hoc tools developed by experts themselves for ex.). As available tools one can mention the Digital Forensic Framework (DFF) [29], X-ways forensic [30] for live (RAM, registers) and post-mortem (connections, data, metadata, files launched by processes) analysis, Internet Evidence Finder for internet-related data gathering, XRY and UFED Cellebrite for smartphones and GPS terminals [31] [32].

3.2 System forensic analysis

After intrusion attacks, data have to be collected from network equipment (logs files, traffic) and from the system files. Networks data are generated by tools that have been developed for another usage than security: packets sniffing, traffic analyzing, connectivity testing [33]. In [34] authors have also pointed out the difference between the objectives of auditing tools designers and objectives of forensic analysts. Existing tools have been developed to analyze back tracks (IP addresses, mail counts, web resources) that can be used to give relevant information about the attackers' localization. Other tools enable to analyze files, emails and collect information about systems and running applications in order to prevent spamming [35]. Security Information and Event Management tools are combined tools and platforms designed to collect, analyze, correlate security events in order to produced synthetic reporting [36]. They are event-oriented tools and they use threats databases. They need to be enhanced with data and knowledge from intrusions tests, attack trees, with knowledge about specific architectures, system configurations and security policies.

Big data architectures associated with virtualization and emulation techniques, data-mining tools such as those based on large graphs constitute a set of processing aids for large amounts of data. Recent tools such as Picviz Inspector [37] are able to process large logs files; they can be viewed as pre-analysis tools, able to reduce the initial entropy of possible ways of analyzing.

3.3 Attack trees reconstruction

One of the data-gathering interests is to be able to replay the events running by simulation. Events correlation tools can be used to synthetize and reduce the large amounts of IDS-raised alerts and to realize high-level analysis tasks such as foiling attacks plans and scenario, impact analyzing [38]. Some of these tools try to correlate multi-sources indices and events with the aim to go back in attacks and security incidents action-plans [39]. Graphic modeling tools used for the attack trees reconstruction are generally derived from those already used in reliability studies: failure trees, vulnerability trees, attack graphs but they are limited to their static aspect. The dynamic feature of cyber-attacks requires other approaches such as attack modeling with Petri Nets (ex. PENET tool [40]), goal-inducing attack chains [41], which consider events sequences rather than individual events, dynamic Bayesian networks [42] where temporal properties of attacks are considered, adapted attacks trees for systems with dynamic aspects [43] and at last, the Boolean Driven Markov Processes (PDMP) formalism, designed by EDF (Electricité de France) for reliability analysis, which has been proposed for attack trees analysis [44]; for this purpose, the dependency notion (represented by a directed arc in the graph) has been adapted to an attack sequence relation.

3.4 Synthesis and scientific challenges

The most important issues are: 1) to define a standardized representation language for encoding the events; 2) the intelligent sampling among the large amount of pieces of information (physical pieces, files), sampling that could be expert-driven with his own criteria or semi-automated with specific algorithms (optimization, decision making); 3) information modeling and visualizing in a synthetic way, presenting the analysis outputs in an intuitive mode to help the experts in their reasoning.

4. Forensic analysis

4.1 General challenge

Let's consider the general case of a distributed system submitted to an intrusion attack. Most of previously mentioned tools enable to collect indices and tracks, to store and preserve them for processing such as correlations or scenario reconstitution. At this step, the expert has to use his own reasoning to form his own opinion. Only a few works have proposed reasoning tools to help experts, to propose and evaluate hypothesis not only from technical data but also from knowledge about context, behaviours, etc. The scientific challenge is then to build a reasoning scheme that can be able to produce an exploitable report for adjudication, using heterogeneous data that may be uncertain and at different semantic levels. Figure 3 shows out this process.

The formalism and the tools that are used in causal analysis are generally the same as for diagnosis. But the aim of diagnosis is to identify a faulty component of a system for repairing or replacement. As difference, the

forensic process needs to build hypothesis and to verify their plausibility. In [45], the authors propose an approach based on the expert knowledge and that uses fuzzy logic for network forensic. In [46], authors use Bayesian networks to verify hypothesis and constraints in forensic analysis. In [47], authors propose also a Bayesian approach but it is limited to specific attacks.

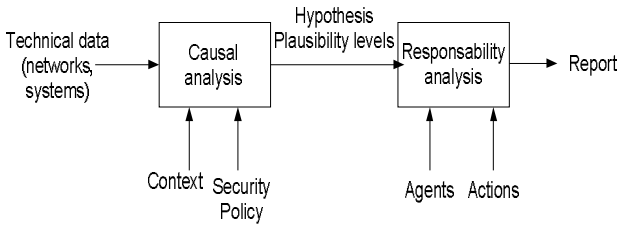

Fig. 3: Forensic Analysis

4.2 Causal analysis

Causes to effects relations are often represented with causal graphs [48]. Causal Bayesian networks (CBN) provide a suitable and interesting modeling and representation power [49] [50]. The difference with classical BN is that in the case of BN, A → B means that the probability of event A, i.e. p(A), has an influence on p(B), which doesn't mean a causal relation. With CBN, it means that A is a cause of B. A probabilistic definition of causality has been defined in [51] but it doesn't integrate any time representation. For cyber-attack, we have previously underlined that the occurrence time of events is an important attribute to use. For that reason other definitions of causality are preferable such as in [52]. In [53], authors make a distinction between endogenous (with random values) and exogenous variables (with fixed values) to establish structural equations of causality. Probabilities are not sufficient to deal with uncertainty. The possibility theory [54] or more particularly possibilistic networks [55] have been proposed to represent and handle uncertainty distribution related to incomplete variables.

4.3 Responsibility and argumentation logic

Responsibility has not to be confused with causality. If A is a cause and B is an abnormal event, A is the result of an agent action who is the direct cause of the abnormal event or who could have prevented it. Previous works in the artificial intelligence field [56] have proposed logical formalisms to reason about responsibility that emerge from agents behaviours, so that it could be able to answer to questions as: Who is the direct cause of A? What are the most plausible causes of A? Has B a direct effect on A? At what degree? An indirect effect? At what degree is an agent responsible of A?

As there is often a need of explanation for cause and responsibility attributions, an argumentation system is required. In the abstract argumentation approach [57], the argumentation is built using graphs. Nodes represent arguments (which are elementary objects) and direct arcs represent attack relations. In the argumentation logic approach, the representation is based on logical relations between arguments which have been built from pieces of information [58].

5. General synthesis

5.1 Classes of problems and tools

The data processing approaches required for an efficient detection and investigation have to take into account the characteristics of malicious actions. There are three of them: the willpower of concealment (i.e. to avoid detection), the operating scenario and the individual or collective behaviours that are characteristic of cybercrime classes. For usual forensic challenges, these attributes can be affected as in Table 1.

TAB. 1: Attackers characteristics

	Concealment	Scenario	Behaviour
Weak signals	X		
Steganography	X		
APT	X	X	
Botnets, rootkits		X	
Fraud		X	
Social networks	X		X
P2P networks	X		X

The concealment problem implies to decrease detection levels in such a way to be able to detect weak signals but the induced risk is to increase false positive and false negative ratios. Attacks that use a predefine scenario require the use of a more important amount of data from control of systems and networks, from attacks history, from intrusion tests, and to process real-time correlations. Behaviours aspects need to use data from contexts, sociological studies and expert knowledge.

5.2 Technological limits

For protection, detection and investigation, the greatest challenge is to develop process chains that can collect and analyze time-limited, sizeable, heterogeneous data about systems, networks and applications. The Table 2 displays how these characteristics will concern cybercrime cases.

The potential information of the available data is not exploited due to technological limits. Data-mining tools (especially classification algorithms) have scalability constraints. The only perspective lies in the big data

technology that gives an interesting opportunity to store large volume of data with intelligent query, absorb sporadic input flows without bottleneck effects, and propose adapted analysis and visualization tools, so-called Big Analytics (BA) and Visual Analytics (VA) respectively. To be BA-compatible, algorithms have to be scalable, because they behave linearly vs. data size or because they can be parallelized (some non-supervised clustering processes for ex.) or because they can be adapted to a massive parallelization (scoring methods or neural network-based algorithms).

TAB. 2 : Data characteristics

	Volume	Dynamicity	Heterogeneity
Weak signals	X		X
Steganography	X		
APT	X		X
Botnets, rootkits	X	X	X
Fraud	X	X	
Social networks	X	X	X
P2P networks	X	X	X

The visualization methods for multi-dimensional data are generally based on projection operations with the constraint of an efficient interactivity. The forensics needs to correlate variables with time attributes, which requires new approaches based on graphs and graphs matrix [59].

6. Conclusion

This paper has surveyed the most significant challenges concerning the forensic process as they are presented to the scientific community, especially concerning detection methods and forensic analysis. These challenges are permanently changing with behaviours and action modes. As proposed methods run according to reactive principles they do always lean on a strong survey about cybercrime features. This inventory of methods reveals that the digital forensic process is not addressed as a whole. Supplementary efficiency could probably be gained by designing global responses that involve all required competences.

References

[1] E. Casey, *Digital evidence and computer crime: forensic science, Computers and the Internet*, Academic Press, 2000

[2] S.J. Wang, C.H. Yang, *Gathering digital evidence in response to information security incidents*, IEEE Int. Conf. on Intelligence and Security Informatics, Lectures Notes in Computer Science (LNCS), Atlanta, Georgia, USA May 2005

[3] E. Casey, *Error, Uncertainty and Loss in Digital Evidence*, Int. J. of Digital Evidence, Vol. 1, n°2, 2002

[4] A. Patcha and J-M. Park, *An overview of anomaly detection techniques: existing solutions and latest technological trends*, Computer Networks, 51, pp. 3448-3470, 2007

[5] IETF, *The Intrusion Detection Message Exchange Format (IDMEF)*, Request for Comments RFC 4765, Internet Engineering Task Force (IETF), 2007

[6] K. Tabia and P. Leray, *Bayesian Network-Based Approaches for severe Attack Prediction and Handling IDSs' reliability*, Proc. of IPMU'10, pp. 632-642, 2010

[7] H. Tong, Y. Sakurai, T. Eliassi-Rad and Ch. Faloutsos, *Fast-Mining of Complex Time-Stamped Events*, Int. Association of Computer Investigative Specialist, Forensic procedures, www.iacis.com, accessed in oct. 2012

[8] L. Fillâtre and I. Nikiforov, *Asymptotically Uniformly Minimax Detection and Isolation in Network Monitoring*, IEEE Trans. On Signal Processing, July 2012, Vol. 60, n°7, pp. 3357-3371

[9] S. Yahi, S. Benharfat and T. Kenaza, *Conflicts Handling in Cooperative Intrusion Detection: A descriptive Logic Approach*, 22nd IEEE Int. conf. on tools with Artificial Intelligence, ICTAI 2010, Arras, pp. 360-362, 2010

[10] J. Kim, A. Ong and R. Overill, *Design of an Artificial Immune System as a Novel Anomaly Detector for Combating Financial Fraud in Retail Sector*, Proc. of the Congress on Evolutionary Computation, pp. 405-412, 2003

[11] R.C. Chen, M.L. Chiu, Y.L. Huang and L.T. Chen, *Detecting Credit Card Fraud by using Questionnaire-responded transaction model based on support vector machine*, Proc. Of the 5th int. Conf. on Intelligent Data Engineering and Automated Learning, Vol. 3177, Oct. 2004, pp. 800-806

[12] M.F.A. Gadi, X. Wang and A.P. Do Lago, *Credit Card Fraud Detection with Artificial Immune System*, P.J. Bentley, D. Lee, and S. Jung (Eds.): ICARIS 2008, LNCS 5132, pp. 119-131, 2008

[13] S. Panigrahi, A. Kundu, S. Sural and A.K. Majumdar, *Card Fraud Detection: A Fusion Approach using Dempster-Shafer Theory and Bayesian Learning*, Information Fusion, 2009

[14] R. Cogranne, C. Zitzmann, L. Fillâtre, I. Nikiforo, F. Retraint and Ph. Cornu, *Reliable Detection on Hidden Information Based on Non-Linear Local Model*, IEEE Workshop On statistical Signal Processing, 4p, 28-30 June, Nice, 2011.

[15] R. Cogranne, C. Zitzmann, F. Retraint, L. Fillâtre, Ph. Cornu and I. Nikiforov, *A Cover Image Model for Reliable Steganalysis*, 15p, Information Hidding, 18-20 May, Pragua, 2011

[16] M. Latapy, C. Magnien and R. Fournier, *Quantifying Paedophile Activity in Large P2P System*, Information Processing and Management, 2012

[17] V.D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, *Fast unfolding of community hierarchies in large networks*, Journal of Statistical Mechanics, 2008

[18] S. Asur, S. Parthasarathy and D. Ucar, *An event-based framework for characterizing the evolutionary behavior of interaction graphs*, Proc. of the 13th ACM Trans. on the Int. Conf. on Knowledge Discovery and Data Mining (KDD), pp. 913-921, 2007.

[19] F. Gilbert, P. Simonetto, F. Zaidi, F. Jourdan and R. Bourqui, *Communities and hierarchical structures in dynamic social networks : Analysis and visualization*, Social Network Analysis and Mining, Vol. 1, pp. 83-95, Springer Wien, 2011

[20] M. Oliveira and J. Gama, *Understanding Clusters' Evolution*, Proc. of Ubiquitous Data Mining (UDM), Workshop in conjunction with the 19th European Conference on Artificial Intelligence - ECAI 2010 in Lisbon, Portugal, August 16-20, pp. 1-6, Lisbon, 2010

[21] T. Aynaud and J.-L. Guillaume, *Static community detection algorithms for evolving networks*, WiOpt Workshop on Dynamic Networks, pp. 508-514, 2010.

[22] T. Aynaud and J.-L. Guillaume, *Multi-Step Community Detection and Hierarchical Time Segmentation in Evolving Networks*, Proc. of the 5th SNA-KDD Workshop Social Network Mining and Analysis, August 21, San Diego, 2011.

[23] A. Rosi, M. Mamei, F. Zambonelli, S. Dobson, G. Stevenson and J. Ye, *Social sensors and pervasive services : Approaches and perspectives*, IEEE Int. Conf. on the Pervasive Computing and Communications (PERCOM) Workshops, Seattle (WA), pp. 525-530, 21-25 March, 2011

[24] K. Yamanishi, J. Takeuchi, G. Williams and P. Milne, *On-Line Unsupervised Outlier Detection Using Finite Mixtures with Discounting Learning Algorithms*, Data Mining and Knowledge Discovery, Vol. 8, pp. 275-300, 2004

[25] National Institute of Standards and Technology (NIST), *Guide to Integrating Forensic Techniques into Incident Response*, Special publication 800-86, 2006

[26] International Association of Computer Investigative Specialist, *Forensic Procedure*, www.iacis.com/, accessed in Oct. 2012

[27] IETF, *Guidelines for Evidence Collection and Archiving*, Request for Comments RFC 3227, Internet Engineering Task Force (IETF), 2002

[28] National Institute of Standards and Technology (NIST), *Disk Imaging Tool Specification*, V3.1.6, Oct. 2001

[29] Digital Forensic Framework (DFF), http://www.digital-forensic.org/, accessed in Oct. 2012

[30] X-Ways Forensics, http://www.x-ways.net/forensics/, accessed in Oct. 2012

[31] Micro Systemation XRY, http://www.msab.com/, accessed in Oct. 2012

[32] Cellebrite, http://www.cellebrite.com/, accessed in Oct. 2012

[33] N. Meghanathan, S. Reddy Allam and L.A. Moore, *Tools and Techniques for Network forensics*, Int. J. of Networks Security & Its Applications (IJNSA), Vol. 1, n°1, 2009

[34] S. Peisert, S. Karin, M. Bishop and K. Marzullo, *Principles-driven forensic Analysis*, Proc. of the New Security Paradigms Workshop, NSPW'05, pp. 85-93, Lake Arrowhead, CA, Sept. 2005

[35] T. Eggendorfer, *Methods to identify spammers*, Proc. of the 1st Int. Conf. on forensic applications and techniques in telecommunications, Information and Multimedia, e-Forensics'08, 7p, Adelaide, Australia, Jan. 21-23, 2008

[36] AlienVault, http://www.alienvault.com/, accessed in Oct. 2012

[37] Picviz Labs, http://www.picviz.com/, accessed in Oct. 2012

[38] C.Kruegel, F. Valeur and G. Vigna, *Intrusion detection and correlation : Challenges and solutions*, Springer, 2004

[39] T. Samuel and P.M. Chen, *Backtracking intrusions*, Proc. of the 9th ACM Symposium on Operating Systems principles, pp. 223-236, NY, USA, 2003

[40] CyberPower, PENET tool, *powercyber.ece.iastate.edu/penetintro.html, accessed in Oct. 2012*

[41] C.Phillips and L. Painton Swiler, *A graph-based system for network-vulnerability analysis*, Proc. of New Security Paradigms Workshop, pp. 71-79, 1998

[42] M. Frigault, L. Wang, A. Singhal and S. Jajodia, *Measuring Network Security using dynamic bayesian network*, Proc. of the ACM workshop on quality protection, QoP'08, Alexandria, USA, pp. 23-30, ACM, NY, 2008

[43] P.A. Khand, *System level security modeling using attack trees*, Proc. of 2nd Int. Computer Control and Communication Conf. (IC4), pp. 1-6, Feb. 17-18, Karachi, 2009

[44] L. Piètre-Cambacédès and M. Bouissou, *The promising potential of the BDMP formalism for security modeling*, supplemental volume of the proc. of the 39th annual IEEE.IFIP Int. Conf. on Dependable Systems and Networks (DSN 2009), Estoril, Portugal, June 2009

[45] M.Y.K. Kwan, K.P. Chow, F.Y.W. Law, P.K.Y. Lai, *Computer Forensics using Bayesian Network: A case study*, HKU CS Technical Report, TR-2007-12, Hong-Kong University, 2007

[46] N. Liao, S. Tian, T. Wang, *Network Forensics based based on fuzzy logic and expert system*, J. Computer Communications archive, Vol 32, Issue 17, 2009

[47] T. Duval, B. Jouga and L. Roger, XMeta, *A Bayesian Approach for computer forensics*, ACSAC, Tucson USA, 2004

[48] J. Pearl, *Causality: models, reasoning and inference*, Cambridge University Press, NY, USA, 2000

[49] A. Darwish, *Modeling and Reasoning with Bayesian Networks*, Cambridge University Press, NY, 2009

[50] F.V. Jensen and T.D. Nielsen, *Bayesian Networks and Decision Graphs*, Springer 2007

[51] I.J. Good, *A causal Calculus I*, British J. of the Philosophy of Science, 11, pp. 305-318, 1961

[52] P. Suppes, *A probabilistic Theory of Causality*, Amsterdam, 1970

[53] J. Halphern and J. Pearl, *Causes and explanations: A structural model approach, part 1: Causes*, British J. of the Philosophy of Science 56, pp. 843-887, 2005

[54] D. Dubois, H.T. Nguyen and H. Prade, *Possibility Theory, probability and fuzzy sets*, in Fundamentals of Fuzzy Sets, D. Dubois and H. Prade Eds, Kluwer Academics Publishers, 2000

[55] S. Benferhat and S. Smaoui, *Possibilistic causal networks for handling interventions: a new propagation algorithm*, Proc. of the AAAI'07 pp. 373-378, 2007

[56] L. Cholvy, F. Cuppens and C. Saurel, *Towards a logical formalization of responsibility*, Proc. of the 6th Int. Conf. on AI and Law, pp. 233-242, ACM Press, 1997.

[57] P. Besnard and A. Hunter, *Elements of Argumentation*, MIT Press, 2008

[58] P.M. Dung, *On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-Person Games*, AI vol. 77, n°2, pp. 321-358, 1995

[59] J. Bertin, *Semiology of Graphics: Diagram, Networks, Maps*, University of Wisconsin, Press, Madison, (W.J. Berg transl.), 1983