Defending Financial Infrastructures Through Early Warning Systems: The Intelligence Cloud Approach*

[Extended Abstract]

Giorgia Lodi, Leonardo Querzoni, Roberto Baldoni
Dip. di Informatica e Sistemistica
Sapienza University of Rome
last_name@dis.uniroma1.it

Mirco Marchetti, Michele Colajanni
Dipartimento di Ingegneria dell'Informazione
Università di Modena e Reggio Emilia
name.last_name@unimore.it

Vita Bortnikov, Gregory Chockler, Eliezer Dekel, Gennady Laventman, Alexey Roytman
IBM Haifa
last_name@il.ibm.com

ABSTRACT

Recent evidence of successful Internet-based attacks and frauds involving financial institutions highlights the inadequacy of the existing protection mechanisms, in which each instutition implements its own isolated monitoring and reaction strategy. Analyzing on-line activity and detecting attacks on a large scale is an open issue due to the huge amounts of events that should be collected and processed. In this paper, we propose a large-scale distributed event processing system, called *intelligence cloud*, allowing the financial entities to participate in a widely distributed monitoring and detection effort through the exchange and processing of information locally available at each participating site. We expect this approach to be able to handle large amounts of events arriving at high rates from multiple domains of the financial scenario. We describe a framework based on the intelligence cloud where each participant can receive early alerts enabling them to deploy proactive countermeasures and mitigation strategies.

Categories and Subject Descriptors

D.2.8 [**Software Engineering**]: Software Architectures; D.4.4 [**Operating Systems**]: Communications Management—*Network communication*; K.4.2 [**Computers and Society**]: Social Issues—*Abuse and crime involving computers*; K.6.5 [**Management of Computing and Information Systems**]: Security and Protection

1. INTRODUCTION

The trend towards the "webification" of critical financial services, such as home banking, online trading, remote payments, improves 24h service availability and user-friendliness. On the other hand, it exposes such services and the supporting IT infrastructure to massive, coordinated Internet-based attacks and frauds, that are not being effectively countered by any single organization.

The urgent need for more effective monitoring and security solutions is widely recognized in other critical infrastructures. Several technologies and best practices enable thorough analysis of the events related to a specific domain (e.g. the network traffic within an ISP [6]). However, current monitoring approaches are inadequate to deal with coordinated and distributed attacks on a large scale. Even well-protected and highly secure financial institution networks are vulnerable to Distributed Denial of Service attacks [10, 12, 7, 11, 9], and to complex and coordinated frauds involving multiple actors spread over different countries [8]. In these cases, the monitoring and detection systems whose scope is limited to each inidividual organization are unable to detect attacks and provide early alerts. To be effective, the monitoring activities have to involve multiple participants possibly distribued over disparate organizational, administrative, and geographical domains. Those partipants will collectively generate massive amounts of event data whose processing can no longer be effectively accomplished by the existing centralized solutions.

As an alternative, we introduce a novel distributed event aggregation and correlation system to monitor widely distributed infrastructures, with the aim of providing early detection of attacks, frauds and threats. Our approach leverages early work on IBM System S[5], which is a distributed system supporting parallel execution of the event processing flows on a cluster of machines. In this paper, we propose extending this approach into a massively distributed event processing system, dubbed *intelligence cloud*, to provide the necessary infrastructure for early detection of distributed and coordinated attacks through the analysis and correlation of events coming from multiple, distributed and possibly heterogeneous sources. We discuss challenges associated with realizing this vision, which include dealing with the system scale, and the ability to deal with high incoming event rates in a timely fashion.

The roadmap of the paper is as follows. In Section 2 we present some of the vulnerabilities reported in the financial sector so far. In Section 3 we introduce the notion of intelligence cloud and provide basic insights on it. In Section 4 we apply intelligence clouds to Financial ecosystem. This work is encapsulated inside a large European funded project, namely CoMiFin [2], whose aim is to design a monitoring middleware system for protection of financial critical infrastructures.

2. FINANCIAL VULNERABILITIES

The need for early detection of illicit activities in the financial context is highlighted by the effectiveness of coordinated and distributed attack strategies. Cross-domain interactions, spanning different organization boundaries are in place in financial contexts. These interactions involve heterogeneous infrastructure systems such as telecommunication supply, banking, and credit card companies, that possibly generate heterogeneous data and contribute to the birth of a *global financial ecosystem*. This trend towards globally integrated enterprises show that a crucial requirement to be met is to design a monitoring system of the financial ecosystem that allows financial institutions to raise their situation-awareness. In particular, in this section we focus on two examples of recent distributed, coordinated and cross-domain attacks against financial institutions: a payment card fraud and an Internet-based attack.

The first example that we consider is a fraud carried out in 2008 using 100 compromised payment cards [8]. A network of coordinated attackers have been able to use these cards to retrieve cash from 130 different ATMs in 49 countries worldwide, totaling 9 million of US dollars. The high degree of coordination in this attack is testified by the astonishing fact that it took only half an hour to be executed. By targeting several geographically distributed ATMs belonging to different financial institutions, attackers have been able to evade all the local monitoring techniques used for detecting anomalies in payment card usage patterns. The fraud has been detected only later, after aggregating all the information gathered locally by each financial institution involved in the payment card scam. A distributed and cross-domain event processing and monitoring system would have detected the anomalous usage pattern of the counterfeit credit card much earlier, thus stopping the attack and mitigating its damage.

A huge amount of financial transactions generates traffic that is also carried over publicly accessible communication mediums such as Internet. In this scenario, financial infrastructures are inherently exposed to a variety of cyber attacks and frauds. A typical coordinated and distributed attack is represented by a Distributed Denial of Service (DDoS), able to render web-based financial services unreachable from legitimate users. There are several cases reported in the press and many other “less publicized”. In this section we specifically refer to a DDoS observed in 2007 in North Europe that we are considering as a testbed case in CoMiFin [2]. The DDoS attack[1] targeted a credit card company and two DNS servers through syn flood and smurf techniques, and lasted several days. Internet availability of the targeted infrastructure has been restored only after several trial-and-error activities carried out manually by network administrators of the attacked systems and of their Internet Service Providers (ISPs).

A coordinated and cross-domain event processing technique would be extremely helpful by providing early detection of threats and automatic and effective communication among all the parties involved in the attack. For example, in case of low variability in the source addresses and protocols of the network packets used to carry out the DDoS attacks, a global monitoring system can cluster them in a limited number of classes characterized by specific features. This classification can be disseminated by the monitoring system to all the interested parties (both within the same financial institution and across different domains), thus helping in defining appropriate filtering rules that can reduce the volume of DDoS attacks. Again, all the information related to the source addresses from which the DDoS attack appears to be generated can be automatically disseminated by the global monitoring system to the ISP of the attacked financial institution, as well as to other ISPs involved in the ecosystem. Currently, ISP involvement, known to be the most effective mean to counter DDoS as it can filter the offending traffic when it is close to its source, is carried out through human intervention that sets up new filtering rules. This process is error-prone, introduces unnecessary delays in the application of DDoS mitigation procedures, and is slow to react to possible changes in the attack pattern. A global monitoring system can speed up such a process. In addition, a global monitoring system can also help financial institutions and ISPs that are not directly involved in a DDoS attack: although these parties cannot do anything to mitigate the attack when it is in progress, the acquired knowledge of past attacks through the use of a global monitoring system can be required for early identification of known attack sources (e.g., amplificators, bots) that participate in successive illicit activities.

3. INTELLIGENCE CLOUD

Effective protection against those attacks requires global cooperation among the participating entities to share the information about emerging threats, and collectively offer the computational powers to discover and contain those threats. Intelligence cloud proposes an architecture to realize this vision. In a nutshell, it can be seen as consisting of two main parts: an infrastructure for fast and secure dissemination of the primitive event information produced by the local monitoring software at each individual participant (e.g., [13, 17, 21]); and a globally distributed event processing system. In the remainder of this section, we mostly focus on the distributed event processing part of intelligence cloud. The monitoring part is discussed in Section 4.

The primary objective of the intelligence cloud is to leverage the computational and storage resources available at each participant attached to the cloud to mine the event streams

[1]The following text has been agreed to be disclosed in a public form within the CoMiFin project.

delivered by the underlying event propagation substrate for potentially dangerous patterns of activity and other anomalies. For financial institutions, electronic transactions are measured in millions of some currency. These transactions take fractions of seconds to execute. The sooner the cloud recognizes the threat the smaller the loss. Sending all the information to a centralized location and then processing it for many hours (as it is the case with many of the existing distributed event processing systems) is therefore, not acceptable. The online processing mechanism has to be distributed, and the underlying logic should be capable of reaching accurate conclusions even in the presence of incomplete and partially corrupt event inputs.

Event processing in the intelligence cloud leverages the concept of dynamic event processing network of [5]. Each node is capable of performing some basic analytics. The analytics is performed on the stream of events that go through it. The basic analytics is designed to perform its job in minimal time so that it does not affect the speed of the event stream. The results of the processing are then directed to the next node in the event processing network. The next node might perform a different analytic operation on the events coming from the originating node. In this way, we get a parallelization of the event processing that enables the intelligence cloud to handle massive amounts of events [5]. The dynamic assignment of analytics to nodes and direction of the incoming event streams to the processing locations is handled by a scheduler component (see for example [15, 18]). The scheduler takes the specification of the event processing logic, and assigns processing tasks to individual machines within the intelligence cloud cluster. The task placement is based on the desired optimization objectives (such as latency, throughput and high availability), and constrained by the current CPU load, memory consumption at the partipating machines as well as the data locality and affinity.

To facilitate parallelization, the processing logic is specified as a flow, which is a directed acyclic graph (DAG) of elementary processing operators. The flows are specified in a high-level language, such as Spade [5], and translated into an intermediate representation that can be understood by the scheduler. One new avenue we intend to explore is the programming frameworks for specifying flows, which will be powerful enough to describe sophisticated processing algorithms yet concise and simple to use.

One promising candidate for such programming framework is Jaql [3], which is a simplified query language for processing semi-structured date based on the Java Script Object Notation (JSON) data model [4] developed in IBM. In the current implementation, Jaql is compiled into a flow of map-reduce tasks, which can then be executed on a Hadoop cluster [1]. Extending Jaql to support event processing flows as well as enhancing the map-reduce paradigm to support input data furnished as a continuous event stream are among the new directions we are exploring in the CoMiFin project detailed in the next section.

4. INTELLIGENCE CLOUD FOR FINANCIAL ECOSYSTEM

In the considered testbed CoMiFin scenario [2], a financial ecosystem has end-points at participating institutions

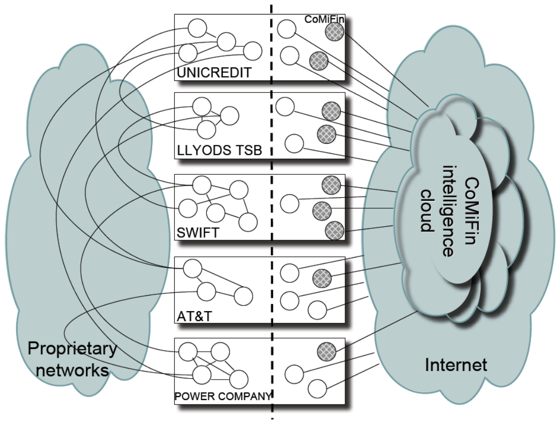

Figure 1: The intelligence cloud in a financial scenario

(e.g., financial institutions, electric utilities, communication providers, and others). In addition, there will be internal CoMiFin nodes that will together form the intelligence cloud as shown in Figure 1. Events will flow into the cloud from the various end-points. These raw event data will be processed within the CoMiFin infrastructure and turned into intelligence for event detections. The produced cumulative intelligence will then be disseminated to the interested participants. The main premise is that nowadays the banks and financial institutions cannot be considered isolated entities. A large fraction of the everyday financial transaction traffic involves multiple participants often spread all over the world. As a result, even a single compromised location becomes a threat to the business integrity of many unrelated financial bodies potentially residing in distant geographies and administrative domains.

In our financial setting, regular transaction activities that financial institutions perform are isolated from the activities carried out by the global monitoring system. Specifically, financial transactions are usually handled through established protocols (e.g, SWIFT) for inter-financial institutions activities or through proprietary secure internal or external networks that financial institutions maintain (e.g., between bank branches as shown in Figure 1). In contrast, in order to exploit the functionalities offered by the global monitoring system, each participating actor can subscribe to the monitoring infrastructure by signing a *basic agreement*. With this agreement, every actor can clearly define which resources it is willing to publicly provide to the intelligence cloud participants, and the conditions under which the resource sharing will occur. In addition, the basic agreement can provide the participants with a set of basic services including billing, a list of partners that form the intelligence cloud, a list of specific available services provided by the monitoring infrastructure. A subscriber to the monitoring system will interface to the financial end-point and provide controlled information produced by its own internal monitoring software. In return, the financial institution will get business intelligence produced by the intelligence cloud. The information will be tagged with a confidence tag and will be acted upon at the discretion of the financial institution. To make an example, we might think that a yellow tag will indicate that

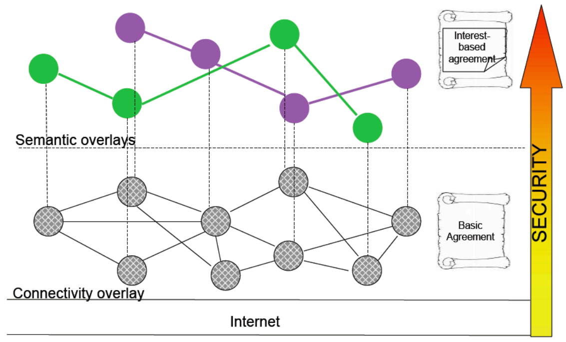

Figure 2: The monitoring system: communication infrastructure

a DDoS attack (or any other cyber attack) was recognized but still needs more validation. A red tag will indicate that a DDoS attack (or any other anomaly) was recognized with high confidence. Financial institutions and any other critical infrastructure should filter the received warnings to the level they are comfortable with.

From a communication infrastructure point of view, signing the basic agreement implies the construction of a connectivity overlay network (actually forming a graph with some connectivity degree k [14, 20]) that connects all the resources of the participants to the intelligence cloud agree to share (the bottom level of Figure 2).

Internet and a k-connected overlay allow us to maximize the availability of the overall system leveraging on the Internet business continuity (IP reconfigurability) and the k connectivity of the overlay (possibly resilient to k failures).

Once joining the intelligence cloud infrastructure, more secure agreements can be signed by a subset of participants. We call them *interest-based agreements* that are used in order to allow participants to subscribe to so-called *semantic exclusive rooms*. These rooms are virtual spaces where participants can share interest-based events and information at the highest level of security (e.g., in Figure 2 the arrow named security). Interest-based events and information include faults notification, service interruptions, DDoS and any other cyber attacks. The interest-based agreements describe the quality of the services, the data formats, and any security aspects related to the sharing of interest-based events and information. From a communication infrastructure point of view, semantic rooms are structured as *semantic overlays* that are built on top of the connectivity one. The top level of Figure 2 illustrates semantic overlays that embody interest groups.

5. CONCLUDING REMARKS

Realizing monitoring system at an Internet scale to detect possible attacks and frauds to the financial infrastructure requires massive data processing capabilities in a timely manner. We propose the vision of the intelligence cloud to handle with this complexity. The intelligence cloud is formed by a flexible distributed event dissemination substrate integrated with a dynamic event processing system that parallelizes event computation and correlation. The paper pointed out main challenges associated with the realization of an internet-scale intelligent cloud which are as follows:

- sustaining high event loads to produce meaningful response within well-defined (and reasonably short) time boundaries.
- high-level programming frameworks for specifying processing logic, which should be powerful enough to describe complex processing rules while being simple and extensible.
- effective processing heuristics capable of accurate detection of complex events even in the presence of incomplete and partially corrupt event inputs.

Intelligence cloud is being currently developed in the context of the CoMiFin project where we are designing new strategies to make this massive processing system even trusted and robust.

6. ACKNOWLEDGMENTS

We would like to thank all the CoMiFin partners that provided us with invaluable comments and real data concerning cyber attacks. This research has been partly funded by the EU project CoMiFin (Communication Middleware for Monitoring Financial Critical Infrastructures (IST-225407)).

7. REFERENCES

[1] http://hadoop.apache.org/
[2] http://www.comifin.eu/
[3] http://www.jaql.org/
[4] http://www.json.org/
[5] System S, http://domino.research.ibm.com/comm/research_projects.nsf/pages/esps.index.html
[6] AT&T "Protect your business by preventing Internet attacks", September 2004, http://www.corp.att.com/emea/docs/pb/internet_protect.pdf
[7] ChronoPay Suffers DDoS Attack, http://www.kommersant.com/p876309/r_500/electronic_payment_processing
[8] FBI investigates 9 Million ATM scam, http://www.myfoxny.com/dpp/news/090202_FBI_Investigates_9_Million_ATM_Scam
[9] Liberty Reserve is down under DDoS attack, http://www.ecommerce-journal.com/news/libertyreserve_what_is_going_on
[10] National Australia Bank hit by DDoS attack, http://www.zdnet.com.au/news/security/soa/National-Australia-Bank-hit-by-DDoS-attack/0,130061744,339271790,00.htm
[11] Netcraft, Payment Gateway StormPay Battling Sustained DDoS Attack, http://news.netcraft.com/, 10th February, 2006
[12] Update: Credit card firm hit by DDoS attack, http://www.computerworld.com/securitytopics/security/story/0,10801,96099,00.html
[13] R. Baldoni, R. Beraldi, V. Quema, L. Querzoni, and S. Tucci-Piergiovanni, "TERA: topic-based event routing for peer-to-peer architectures", *In Proc. of the 2007 ACM international conference on Distributed event-based systems*, 2007
[14] R. Baldoni, S. Bonomi, L. Querzoni, and S. Tucci-Piergiovanni, "Investigating the Existence and the Regularity of Logarithmic Harary Graphs", *In Proc. of the IEEE International Symposium on Reliable Distributed Systems*, 2008 (extended version to appear in Theoretical Computer Science).
[15] N. Bansal, R. Bhagwan, N. Jain, Y. Park, D. S. Turaga, C. Venkaramani,"Towards Optimal Operator Placement in Partial-Fault Tolerant Applications", *IEEE Infocom* 2008, April, Phoenix, AZ
[16] D. Bickson, Y. Tock, O. Shental, D. Dolev, "Polynomial Linear Programming with Gaussian Belief Propagation", *In Proc. 46th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, USA, September 2008.
[17] G. Chockler, R. Melamed, Y. Tock, R. Vitenberg "SpiderCast: An Interest-Aware Unstructured Overlay for Topic-Based Publish/Subscribe", *LADIS* 2008.
[18] F. Fu, D. S. Turaga, O. Verscheure, M. Van der Schaar, and L. Amini, "Configuring networked classifiers in distributed and resource constrained stream processing systems", *In Proc. of ICASSP* 2007.
[19] Girdzijauskas, G. Chockler, Melamed, Y. Tock. "Gravity: An Interest-Aware Publish/Subscribe System Based on Structured Overlays". *In Proc. of DEBS'08* (fast abstract), Rome, July 2008.
[20] R. Melamed and I. Keidar, "Araneola: A Scalable Reliable Multicast System for Dynamic Environments". *Journal of Parallel and Distributed Computing (JPDC)* 68(12), December 2008.
[21] Y. Vigfusson, H. Abu-Libdeh, M. Balakrishnan, K. Birman, Y. Tock, "Dr. Multicast: Rx for Datacenter Communication Scalability", *In Proc. of HOTNETS '08* 2008.