Annales UMCS Informatica AI 7 (2007) 117-125

Annales UMCS
Informatica
Lublin-Polonia
Sectio AI

http://www.annales.umcs.lublin.pl/

Intelligent agents in support of internet security

Krystian Baniak*

Faculty of Electronics and Information Technology, Warsaw University of Technology, Nowowiejska 15/19, 00-665 Warszawa, Poland

Abstract

Internet today is the main medium rapidly delivering information that we need. The information is delivered via a vast number of services that make our life become faster and more conscious. The question is at what price? The solutions devised by humans are prone to errors, which furthermore entail a new type of security challenge – cyber crime. With this paper we try to show that autonomous intelligent agent based systems are able to support or even offload humans in bringing Internet Security to the next level. The related research is focused on analyzing network information flows, building knowledge base and applying reasoning techniques that would allow agents to track suspicious network activities, identify and profile users and finally indicate and prevent cyber crime. The main applications for this solution are forensics and law enforcement support.

1. Introduction

In real life agents are specialized individuals that act on somebody's behalf. Agents are typically highly specialized and undertake missions that require skills and knowledge that we often lack or possess insufficiently. Agent can be autonomous or cooperative. The latter is the type of an agent that works in distributed environments and performs actions that require teamwork and effective communication. In the field of security, agents create networks of individuals infiltrating given theatre of actions in order to gather information that is necessary to sustain protected entity's security.

Using this real life analogy, Artificial Intelligence researchers have long worked on applying proven agent scenarios to the computer science world. Today, intelligent agents are beneficial to distributed analysis, knowledge mining or cooperative reasoning. Computer Agents (software or hardware) are able to perform tedious and computationally demanding tasks like information

* Corresponding author: *e-mail address*: krystian.baniak@elka.pw.edu.pl

filtering, classification, and pattern/trend analysis – the tasks that would take humans lots of time otherwise.

We believe that due to its unique features, intelligent agent frameworks can leverage the Internet Security.

Internet, because of its open nature; anonymity of users (we know users only by their virtual identity); complexity of systems and finally imperfectness of human developed solutions creates great opportunity for perpetrators. Cyber crime is getting more serious and to maintain trust of Internet community, we have to augment its safety. Currently we have many security controls that inspect network communications like firewalls, intrusion protection and web filtering. The problem is that those controls focus on communication protocol sanity and cannot infer about the user motives that drive a given type of communication. The knowledge of this type may allow for profiling the user and building up his network of interests (social networks).

Intelligent agents can collect data, analyze it and derive patterns that resemble valid and suspicious classes of subscriber activities. This, however, is not as simple as analyzing network protocols. The task is much harder as machines will have to "sense" human intentions and decode groups of people sharing the same objectives (like pornography espionage nets). This, however, cannot be done without long term observation and distributed Internet wide agent cooperation across national boundaries.

We think that this type of system can help law enforcement in detecting and elimination of computer piracy, abuse, spam and espionage. Knowledge of human motives and behaviors in the network environment, as well as the security posture of observed parts of network, are beneficial in predicting potential future targets of offenders.

Similar attempts have been successfully used for predicting potential crime scenes by New York US [1] Police. Intelligent systems are used in many areas where they can offload humans in hard work of analyzing and discovering interconnections like in it is done for credit card fraud detection, money laundering. The Carnegie Mellon University research developed NetProbe application used for tracking online auction users that elevate reputation in the improper way (Network Detection via Propagation of Beliefs).

The next chapters will present the outline of our research on Intelligent Agents in the support of Internet Security.

2. Research goals

The main research goal is to implement Intelligent Agent framework proving the following thesis:

Autonomous intelligent agents, using data mining and artificial intelligence techniques, will be able to build a model of human behavior/activity in the Internet. This model will be used to detect, predict and prevent compound network/Internet scenarios that belong to category of cyber crime.

Our research is planned for the following phases:

- *Initiation Phase*,selecting knowledge representation solution, information gathering techniques and inference apparatus that will be used by agents.
- *Knowledge Base Building Phase*, accumulating the results of research and assessing the quality and usability of acquired knowledge. This phase is namely targeted on creating ontology of Internet as an information system.
- *Testing and Conclusions*,phase in which agent will be used in the simulated environment that is to validate its effectiveness.

The key area of application for researched solution is law enforcement and cyber crime prevention. It is intended to equip the agent framework with capability of active scan targeting for evidence and information by using vulnerabilities. However, this approach will be used only towards suspects and offending network nodes. The Internet subscribers have their rights – privacy and anonymity – that we cannot neglect. Profiling, as performed by agents, has to be at an appropriate abstraction level. It cannot disclose the sensitive, personal information of the Internet user. However, the profile has to be enough good to classify and track the subscriber. Only the case of evident violation can lead to thorough inspection and evidence gathering. This process will be further referred to as revocable anonymity case.

3. Characteristics of agent operation model

Agents are the computer systems characterized by the following attributes [2]:

- Autonomy: they pursue its target without supervision.
- Social ability: they form structures and operate cooperatively; communicate among themselves and interact with humans.
- Reactivity: perception of surrounding environment. In our case it means response to the information flow event in the Internet.
- Pro-activeness, goal driven response to agent's environment.

Intelligent Agent System consists of the following important elements:

- Architecture: organization schema and roles and responsibilities of agents.
- Protocol: communication protocol – agent's language.
- Goals: *modus operandi* of agent system.
- Knowledge representation: description method of acquired knowledge that enables inference and data manipulation.

– Inference model: core intelligence element; method of data manipulation that enables agent to operate independently and accomplish its goals.

3.1. Architecture and goals

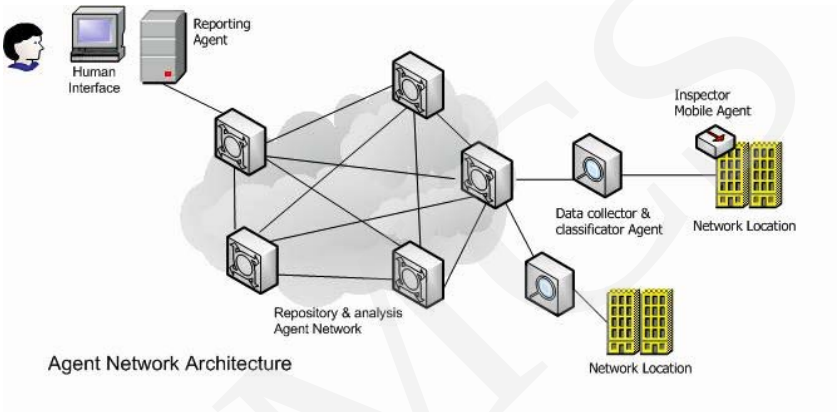


Fig. 1. Agent system architecture

Agent framework, proposed in this paper, uses several types of agents. This resembles society in which different types of individuals perform dedicated, specialized tasks – roles:

– *Stationary agents*; planted on trusted platforms, they are divided into the following categories:
  – *Workers*: Collector Agents that specialize in data collection, pre-classification and executing feedback action ordered by specialists. Workers are able to assess vulnerability level of the observed object. They may use genetic algorithms to excel in accuracy of their observations[1] [3].
  – *Specialists*: agents that constitute the core of all framework, responsible for storing accumulated knowledge in appropriate representation, analysis and feedback actions
  – *Reporters:* agents that make an interface for humans, they are able to explore distributed database of specialists and prepare reports and notifications for the human operator.
– *Mobile agents*: special agents infiltrating potential targets (suspicious network entities). Those agents work in an intrusive way as they use the penetration testing techniques and exploit weaknesses of targets to harvest us much evidence as possible. They only engage their mission in the case

[1]Similar approach has been proposed for vulnerability assessment by the research ATIRP symposium [3].

the target is considered suspect. They bring much concern with privacy and ethics as they act like spies.

Currently Collector Agent has the following architecture. It uses database that allows for classification of network traffic. The collector agent observes network flows and uses appropriate parsers to decode the required information. Identified network entities and their characteristics are stored in the profile database. Profiles are created and updated constantly while new information appears.

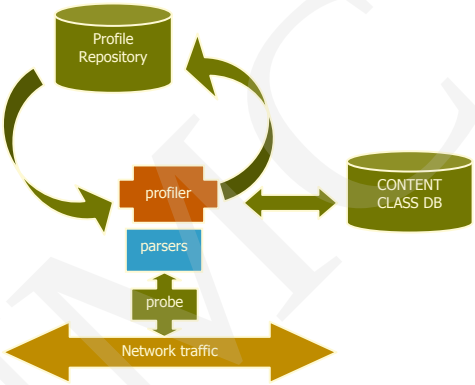


Fig. 2. Collector Agent architecture

Hierarchical agent frameworks have also been proposed by the Purdue University (AFFID) or Iowa University (JAM) researchers.

3.2. Protocol

Message exchange is crucial for every agent framework. As long as the Internet is open access network, agents have to communicate in a secure way. The other problem is time synchronization that is extremely important when agent correlate events occurr in the distributed environment.

To satisfy the protocol requirements public key cryptography, and time synchronization mechanisms have to be adopted. Similar solutions can be found in the area of agent based intrusion detection systems [3].

3.3. Knowledge representation

"Knowledge representation is a multidisciplinary subject that applies theories and techniques from the other three fields: Logic, Ontology and Computation Support, according to John Sowa [4]. Knowledge represents agent's perspective of the analyzed environment. Network related behavior, that can be observed in the Internet can be easily described in the form of scenarios or frames. Agents have or create frames that further enable them to infer on the new events. This

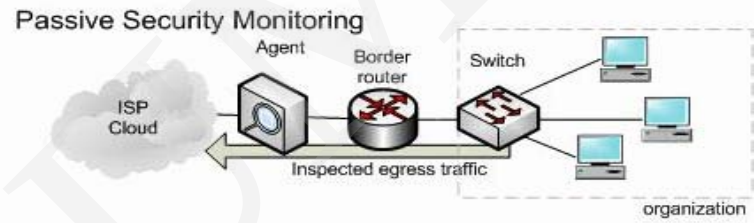
approach is similar to human operation (at least from the perspective of creationism theory).

The other significant technique is social nets analysis. Agents may use this approach for the analysis of interconnections between subscribers and to tackle organized crime.

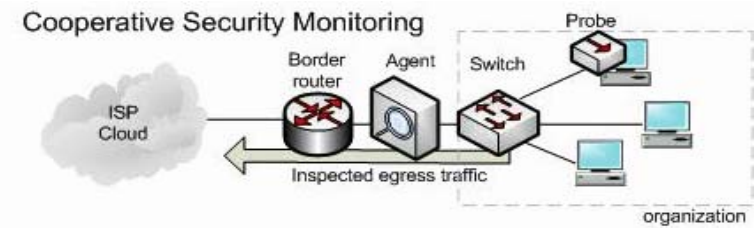
4. Security challenges

Safety of security subsystem is a very important goal. Security perspective is multidimensional and covers the following attributes:

- Confidentiality and Integrity of exchanged messages and accumulated knowledge
- Accuracy of accumulated information; resistance to modification.
- Privacy protection: agents cannot collect personal information without appropriate authorization.
  - Agents can collect data passively using ISP[2] interface.



  - Agent can collect data cooperatively using the monitored organization interface



- Survivability of agent knowledge that consists of information and evolved reasoning methodologies which implies redundancy and distributed storage.

Those attributes are further translated into research challenges that include:

- Secure communication (agent language).
- Trust among agents, making decisions under the condition of uncertainty.
- Stealth of operation: do not interfere with normal traffic.
- Security of execution: execute on the trusted platforms.

[2]ISP – Internet Service Provider

– Preservation of state and information: availability and redundancy.
– Legality of operation: privacy concern: revocable anonymity of subscriber.

Revocable anonymity – our research goal is to preserve the Internet user privacy. However, in the case of confirmed security offence, the observed Internet entity can be a subject of thorough analysis for evidence gathering. This state requires appropriate law regulations to be applicable. The outcome of this state is the condition where information about entity identifies it – its anonymity is revoked. This state is triggered by the collector agent after it identifies offensive behavior.

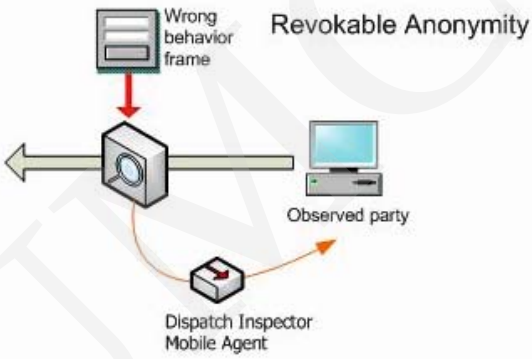


Fig. 3. Revocable anonymity concept

5. Considered techniques

In this chapter we outline the techniques selected as feasible and beneficial in the research process.

Sources of information inspected by Collector Agents:

– web traffic, IP address and domain names of destinations as well as HTTP header values i.e.: host, referrer, url,
– messaging traffic
  – email messages
  – Internet communicator traffic
  – IP telephone voice calls
– DNS queries
– vulnerability scans that estimate threat level for a given network node,
– Public databases of urls and domain names:
  – IP black lists
  – URL classification databases

Knowledge representation – this technique is essential for storing the acquired knowledge about entities and traffic flows. It is a formal way of representing the analyzed environment that facilitates decision making and

inference about the environment. Information regarding Internet entities are best described by profiles. In the case of network activities and relationships we think that the following techniques can be beneficial:

- hierarchical frames (concept originally introduced by Marvin Minsky, MIT)
- social networks – graph of entity interconnections

Inference techniques that are targeted for research:

- Reasoning based on matching known scenarios to the inspected situation.
- Social nets analysis for detecting complex nets of interconnection and relation among Internet users.

6. Ethics and law considerations

Internet Security analysis cannot affect the Internet user's privacy. The privacy and anonymity are the rights of subscribers mandated by legislative bodies in many countries:

- US Electronic Communication Privacy Act ECPA
- EU OECD Guidelines (Organization for Economic Cooperation and Development)

Currently, there is no unified method of tackling cyber crime in the Internet. The problem consists in discrepancies of applicable law per country and a lack of funds and specialists in the third world countries. International efforts toward cyber crime include

- US Mutual Legal Assistance Treaties (MLAT)
- Interpol (EU border control)
- UN Agreements

Those treaties and international agreements help pursuing crime committed against the victims of a given party on the territory of another country.

The conclusion is that for successful operation, Agent Platform requires global coverage. It also has to be supported by law; otherwise its results cannot be used as evidence in the case of offence. Generally, unlawfully gathered evidence is hearsay.

7. Related work

At the Drexel University, a group of researchers implemented (2003) *Secure Mobile Agents on Ad Hoc Wireless Networks* – the system providing security for mobile agents in the WiFi networks. The solution uses meta-reasoning and machine learning techniques. It is capable of reconfiguring the network to maintain system integrity and security and block rogue wireless clients. The agents communicate securely with the aid of public key cryptography.

8. Conclusion

Intelligent Agents have proven its usability in many fields of science like knowledge mining, security policy enforcement, intrusion detection. Using agents in the approach toward creating Internet subscriber activity profiles and improving overall security posture of this open nature communication system is very promising. Fist of all we think we will be able to create common ground for representation of Internet information flow (ontology). Secondly, having this done, we can detect complex security incidents involving multiple cross-boundary threats.

References

[1] Chari S.N., Cheng P., *BlueBoX*: *A Policy-Driven, Host-Based Intrusion Detection System*, ACM Transactions on Information and System Security, 6(2) (2003) (May).

[2] Wooldridge M., Jennings N.R., *Intelligent Agents Theory*. Knowledge Engineering Review, October (1994).

[3] Conner M., Patel C., Little M., *Genetic Algorithm/Artificial Life Evolution of Security Vulnerability Agents*. Army Research Laboratory Federal Laboratory 3rd Annual Symposium on Advanced Telecommunications & Information Distribution Research Program (ATIRP), February (1999).

[4] Sowa J.F., *Knowledge Representation: Logical, Philosophical and Computational Foundations*. Brooks Cole Publishing, (2000).