On Mobile Agents Resistance to Traffic Analysis

Kamil Kulesza¹  Zbigniew Kotulski²

Institute of Fundamental Technological Research
Polish Academy of Sciences, Warsaw, Poland

Konrad Kulesza
Rhodes University, Grahamstown, South Africa

Abstract

This paper will concern itself with a formulation of a traffic analysis problem for mobile agents. It is an interesting theoretical problem as well as a critical feature when using agents on a massive scale in decision making systems. The decision making systems are applied to demanding and complex environments such as stock markets. The mobile agents used are natural targets for attacks because they provide information for decision making. The resulting information can have a value measured in millions of dollars and information of such a high value attracts potential attacks. An efficient way to attack the user of decision making systems is to learn her strategy and respond in kind. In this respect even passive observation of agents can provide useful data, namely what information they are gathering. A common first defense is to provide anonymity for mobile agents. However, what happens when anonymity is gone? What information then becomes available and what steps will the user take? Yet, the problem has not been previously formulated for such a framework. We formulate it in terms of various factors used for traffic analysis. These factors originate from different side channels that provide information on the operating agents. At the end we state a paradox, which links an excessive use of countermeasures against traffic analysis with weakening system security.

Keywords: Mobile agents security, security protocols, traffic analysis, side channel attacks

¹ Email: kkulesza@ippt.gov.pl
² Email: zkotulsk@ippt.gov.pl
1 Introduction and Motivation

“Program a map to display frequency of data exchange, every thousand megabytes a single pixel on a very large screen. [...] Up your scale. Each pixel a million megabytes. At a hundred million megabytes per second, you begin to make out certain blocks in midtown Manhattan, outlines of hundred-year-old industrial parks ringing the old core of Atlanta.” – William Gibson in [1]

This paper is much different from crypito papers that we usually write. First of all, it does not contain any mathematical equations. Its main purpose is to present a fascinating topic, which still eludes a precise quantitative description. While it is no problem to present some mathematical model to make a paper look more scientific, it is almost impossible to build a good model for such a complex system.

1.1 Preliminaries

Mobile agents were at their peak in the late 90’s, which was closely correlated with the Internet boom. Now, when the preoccupation with the agents decreased, why is it worthy to bother with the mobile agents’ systems? The main reason for disappointment with the mobile agent technology is, that it failed to meet the expectations. It is true that individuals do not use mobile agents on a massive scale, for instance to shop for the lowest airlines fares [5]. One of the reasons was certainly the economy. But were mobile agents a case of failed technology or too high expectations? Or maybe mobile agents were put aside by new concepts in IT technology, e.g. grid computing? These are topics of ongoing dispute, which are far from being resolved.

We deal with mobile agents security because of two reasons:

• they are a fascinating and challenging theoretical concept (we also have a strong feeling that mobile agents will have their great comeback);
• there are applications for mobile agents, which are security critical.

Mobile agents benefit simultaneously from remote code execution, coupled with autonomy and adaptation to a changing environment. Mobile code offers a new paradigm in computing, which is well suited for increasingly interconnected environments. This paradigm opens new opportunities, but simultaneously creates new threats.

Mobile agents are excellent vehicles for modern decision making systems, see [2]. In the same paper another agents’ feature is discussed: close similarity to real life solutions and situations. In fact, we witness the situation that an increasing number of concepts, characteristic to real life, migrate into
cyberspace. This process is well visible in the field of security protocols [3]. Mobile agents based decision making systems may be perceived as a collection of protocols for distributed information acquisition and analysis. Reality and cyberspace enter into a feedback.

Intelligent mobile agents are the most refined form of decision systems that we have yet created. The agent systems can be considered as not only effective, but also user-friendly information technology tools, easy to accept by non-professional users [4].

In the modern world mobile agents are applied in the most demanding and complex environments, for instance stock markets. The resulting information can have a value measured in millions of dollars. When such high stakes are on the table, they always attract potential attackers. An efficient way to attack the user of such a system is to learn her strategy and respond with one’s own. The mobile agents are a natural target for the attack because they provide information for decision making.

While the mobile agent technology has been created for the users’ convenience and improvement in decision systems’ performance, it has introduced new risks into the process—the decision process can now be observed and influenced by the competitors.

This paper makes several contributions. First, we describe traffic analysis for mobile agents. Second, we show two realms; cyber space and reality, that business take place. The interplay between these two realms increases complexities of design and security analysis. Third, we outline side channels, many of which have their roots in system complexity. Finally, we state “over-protection paradox”, namely that an excessive use of countermeasures against traffic analysis weakens system security.

The purpose of this paper is to present a new source of risk. It arises from the agent systems’ potential vulnerabilities and leads to a perceived or acknowledged business risk that may need to be contained. As far as we know, the problem was not investigated from such an angle. To be able to formulate it, first we need to provide some information on mobile agents security. This will allow us to make security assumptions needed for a more precise description of the problem.

1.2 Mobile Agent Security

Mobile agents security falls into a set of problems with mobile security, which were nicely outlined by Roger Needham in [6]. In the paper he presents the development of security methods from a historical perspective. At first security was designed for immobile environments, next mobile technologies (e.g. agents) appeared and a security gap was created. Although great efforts
have been made to close the gap, the major problem is in the paradigm. The historical foundation was good as long as “nothing move[d]”, see [6]. Hence, it is not surprising that currently mobile agents security is an active field of research, with many challenges still ahead, see [7,8]. The most recent survey on mobile agents security can be found in [22].

The problems with security protocols involving mobile agents, prompted Volker Roth to write about programming Satan’s agents [9], in the fashion similar to Ross Anderson’s Satan’s computer [10]. Roth’s article shows the seriousness of the situation. However, since mobile agents security is still a young discipline, it is assumed that in time, many of the problems will be resolved. Therefore, this paper will make two assumptions: (1) protocols used are secure and (2) the agents outside trusted hosts do not leak any other information than possibly about their presence.

In the agent system we have the following parties: the owner for the agents, the mobile agents, the hosts (locations visited by the agents), the adversary. In the security model that we consider:

- Agents can move freely between hosts;
- In public agents travel in encrypted form, the same applies to the data acquired by the agent;
- Hosts are secure locations, which means that an adversary cannot compromise a host’s security.

The paper discusses how to perform successful traffic analysis in such an environment. Moreover, any effort to increase a level of protection against traffic analysis can result in opening windows of opportunity for some side channel attack. Before outlining these concepts in Section 3, first a few remarks about traffic analysis itself.

## 2 Traffic Analysis

“Thus, what is of supreme importance in war is to attack the enemy’s strategy. Next best is to disrupt his alliances by diplomacy. The next best is to attack his army.” – Sun Zi in [11]

### 2.1 Remarks on Traffic Analysis in Security

In the networking world there are many schemes using different techniques in order to enhance resistance to traffic analysis. Yet, a majority of them share one common assumption about the network: a topology consisting of point-to-point links. This approach works nicely for the cable networks. However, when
it comes to mobile security it fails miserably, due to the lack of point-to-point links. Although the problem of resistance to traffic analysis has been around for some time (e.g. [12,13]), only recently Matt Blaze et al. have managed to formulate it for the wireless environment, see [14].

A majority of the schemes’ first and often only line of defense is anonymity (e.g. [7,15]). Although there is a whole continuum for degrees of anonymity (see [16]), to simplify the model it is assumed that for a mobile agent it is a binary value and that it can be lost only once (see [17]).

However, what is the situation when the anonymity is gone? In such a situation, what remains to be protected is information accessed, collected and analyzed by decision making systems. The way to handle traffic analysis may be derived from drawing conclusions from real world cases.

2.2 Traffic Analysis in the Real World

This section serves the purpose of introducing traffic analysis in a wider context, and to draw attention to a more general, strategic agenda. Also, it can be nicely translated into a very precise language of game theory.

Passive observation seems to be as old as espionage itself, which claims to be the second oldest profession. The case of traffic analysis for intelligence agents was described in detail by Peter Wright in “Spycatcher: The Candid Autobiography of a Senior Intelligence Officer” [18]. In the book he describes how Russian agents were performing successful traffic analysis on British counterintelligence services in London during the Cold War. There are also accounts of the technical side of the story, traffic analysis depended heavily on monitoring radio transmissions between counterintelligence officers.

The accounts given in the book concerned a multi-layer traffic analysis. The first-level traffic analysis provided information on what data was being collected by the opponent. The second-level analysis employed strategy. When data was gathered over a long period of time, it permitted the Soviets to draw conclusions on what information had been acquired by counterintelligence. This, together with the knowledge of their own operations, developed an accurate picture about British secret services’ level of knowledge and strategy. It also allowed for estimating what the other party does not know, to find a so-called knowledge complement. Such a reasoning seems at first to be very complicated, but it serves the ultimate goal “to attack the enemy’s strategy”.

Since we are mainly concerned with applications from the field of economy, it is time to get rid of espionage stories and provide a business-related example. An excellent account of traffic analysis operation is given in “Wall Street”, the movie directed by Oliver Stone (see [19]). Let us provide an outline of the plot, since it is essential for further discussion. The story is based on the
famous Ivan Boesky case in 80’s, when a stock market tycoon was nailed by SEC (Securities and Exchange Commission) with charges of insider trading. In the picture stock market tycoon Gordon Gekko (Michael Douglas) sends his young apprentice Bud Fox (Charlie Sheen) to observe Sir Larry Wildman (Terrence Stamp). The later is a powerful British investor planning some deal in the US. Gekko wants to learn about the deal. Following Sir Larry Wildman for all the day Bud Fox finds what investment banks Brit was talking to. He also found where the investor flew his jet to after the talks. The apprentice was unable to get any detail of conversations, since these were carried out in secure locations. When Bud Fox came to report his master, he was apologizing for the poor results. But for Gordon Gekko it was enough information. Knowing the names of investment banks and people involved he was sure that Sir Larry Wildman was interested in some heavy industry enterprise. All these, combined with the Wildman’s plane destination (some location in Pennsylvania) revealed his adversary the company itself—Anacott Steel. Such a reasoning was possible, since Gordon Gekko knew well Sir Larry Wildman. He combined his knowledge with all the information obtained by the traffic analysis. As a result he derived and implemented strategy that allowed him to make millions of dollars on the stocks market. It was not only matter of insider trading or greenmail operation; it was demonstration of the highest skills—a successful attack on the enemy’s strategy.

2.3 Traffic analysis, global market and mobile agents

In the old days traffic analysis in the reality was different from that in data security. There were some occasional interactions, mainly in the SIGNIT (signal intelligence) field. A good instance is a technical side of the story described in “Spycatcher ...” [18]. In case of modern financial markets it is very difficult to separate real world activities from cyberspace ones. In the days before markets were interconnected the things were much simpler (e.g. “Wall Street”). The core business activities and financial market, although closely related, were functionally separated.

Nowadays, all exchanges form one global market which never sleeps—there always is some major exchange that trading is going on. The volume of “virtual money” moving around a planet many times exceeds the global gross product. Moreover, global markets are very volatile, so big money is made and lost in a matter of seconds. The future of individual companies, industries and whole nations depends on what happens on the markets. In this respect the cyberworld has a big impact on reality. With the development of commodities markets and rapid growth in derivatives trading, everything can be subject to some market valuation and consequently traded. This especially
concerns the information, which on one side fuels the markets, while on the other can be also a traded commodity (at least in some specific forms, like intellectual property). So far, global markets are the most complex environment created by humans. Nobody fully understands all the interactions that take place and there is nobody to control them. One of the best tools to handle such a complexity is to use intelligent mobile agents for information gathering and possibly trading. Out of these species, the most efficient ones are the agents with evolution driven intelligence, see [2]. Their main goal is to collect information for their owner; their very survival depends on their ability to do so. In such a framework the agent may have a personal interest in making things happen. He will do anything to survive: he will evolve, find the shortest path through the network, lie to you or fight his way with other agents. Such a philosophy closely mirrors the real world.

3 Mobile Agents Traffic Analysis

In the previous chapter we described the state of the art in traffic analysis. First, we discussed it in the context of cyberspace (data security). Next, we recalled cases from the real world. As was stated earlier, we witness situations where reality and cyberspace enter into a feedback. Joining both realms and applying them to mobile agents will allow us to present the main contribution of this paper. It is a good moment to recall our assumption on mobile agents security: (1) protocols used are secure and (2) the agents outside trusted hosts do not leak any other information than possibly about their presence.

Now we are ready to formulate the traffic analysis problem for mobile agents. We focus on the specific situation when agents are used on a massive scale, for instance to acquire information for decision making systems. In the proposed framework, where large numbers of free roaming mobile agents are employed in the complex network (possibly the whole Internet), traffic analysis of the agents resembles a case for the wireless environment. Hence, we advocate using an approach proposed by Matt Blaze et al. in [14].

3.1 The System and Threat Models

The owner of agents has a simple goal, to collect data without leaking information about itself. The owner, during the process, can encounter two adversary types: listening adversary and active (e.g. Byzantine) adversary. The adversary goals are more complex:

- Collecting the data on the opponent’s (owner) level of information;
- Collecting information on opponent’s knowledge complement (Section 2.2);
Collecting information of the opponent’s patterns of behavior and the reactions to certain stimulus;
Collecting information on the opponent’s strategy (ultimate goal).

At this point it is good to recall our assumptions for the security model:
Agents can move freely between hosts;
In public agents travel in encrypted form, the same applies to the data acquired by the agent;
Hosts are secure locations, which means that an adversary cannot compromises a host’s security.

These assumptions mirror a real life situation, where agents can operate from the diplomatic post using diplomatic status as additional protection (e.g. [18]). We have to make same assumptions as made by intelligence services, that agents are under constant surveillance.

The mobile agents generate traffic in two ways, they can exchange data with their owner and proliferate themselves through the network. Such a situation creates great opportunities for traffic analysis, since not only each of the ways can be separately analyzed, but also their interactions can be investigated. This would result in a multi-layer traffic analysis (see Section 2.2). To illustrate this point consider a situation when only movements of the agents are observed. Data can be collected by:
Following the agents through the network;
Tracing an agent’s route backwards;
Observing some key nodes (e.g. database hosts).

As the result the volumes of data are collected from which activity patterns can be extracted. In addition it is always good to have some extra information, which can speed up the analysis. The case from “Wall Street” is a good instance.

The proposed remedy follows from an analogy with the real world that the best agents do not communicate with the masters. They act autonomously, because information exchange is the most vulnerable element of any intelligence operation (see [18]). Making use of this practice, we suggest that mobile agents exchange information only in secure locations, ideally at the owner’s own host.

3.2 Further Threats and Countermeasures

In the cyberworld an agent consists of bits, which can be freely copied. Although protection techniques are available (see [22]), they have serious limi-
tations. So, we consider a situation that agents can be captured. This can be done simply by copying or replicating the agent, the process that neither agent’s owner nor the agent himself be aware about. This is a major difference with respect to the real world, where usually a master knows that an agent was captured. In cyberworld it is left to the adversary to decide whether she would disclose the fact of an agent’s capture (e.g. by making traceable use of acquired information).

Once an adversary has the agent (or technically speaking the copy of the agent), she can interrogate him. The main goal is to learn all information that the agent possesses. However, sometimes it is more feasible to manipulate an agent for one’s own advantage. In such a situation an agent can be used as the medium to obtain more information from the agent’s master.

One protection can be the “need to know principle” used by all intelligence services. It can be nicely illustrated by the old saying: “The less information you have, the shorter is your interrogation time”. For mobile agents it can be implemented within SPECNAZ framework as proposed in [2].

But this still might be insufficient against more sophisticated attacks. For instance, consider an adversary making use of a fake owner’s host, with full corresponding environment. The game can be carried out much further, since an agent may be fed with the data and released in order to mislead his owner. Also the information that he provides might be correct, but aiming to provoke some action. Because the problem is fuzzy, technical countermeasures are difficult to design. One of the methods would be to use state appraisal functions (see [22]), which make sure that agent’s data are not tampered with.

The general defense method should copy real life and use agent’s intelligence. Intelligent mobile agents can be more difficult to confuse, but they are also more difficult to control. If you are intelligent, you can lie in a convincing way. Survival driven agents evaluate their situation in the context of their own best interest, see [2]. This makes them efficient, especially in the global market context, but also increases certain risks (double agents, turning agents and so on).

Knowing that agents are under constant surveillance, an owner may develop countermeasures. For instance, various strategies can be employed to increase the volume of traffic, with an artificial increase in random and non-meaningful traffic (so-called “white noise”).

In the other example, agents drop information only at secure locations and are not in regular contact with the owner. Usually, mobile agents generate traffic in two ways, they exchange data with their owner and proliferate themselves through the network. This results in multi-layer traffic analysis (Section 2.2). In the real world the best agents do not communicate with the
owner. They act autonomously, because information exchange is the most vulnerable element of any intelligence operation (e.g. [18]). Instead agents should exchange information only in secure locations, ideally at the owner’s own host.

Unfortunately, there is always a price to pay for the countermeasures (see [20]). It comes either in system security or performance and accuracy. For instance, a traffic volume big enough to prevent traffic analysis might be unfeasible in terms of other constraints (e.g. cost, bandwidth available). However, a really serious threat follows the observation, that overdoing countermeasures might enable side channel attacks.

### 3.3 Side Channel Attacks

A side channel attack (a.k.a covert channels interface) uses some secondary data about the object investigated to deduce its main properties. An excellent example is a whole class of attacks on the smartcards based on the power analysis (e.g. [21]). In this case cryptographic functions performed by the smartcard are not attacked directly (e.g. by breaking algorithms). The power consumption of the device is measured and on this basis the statistical information about “the patterns” in the smartcard operations are obtained. This attack has recently proven to be quite successful, see [21].

Each side channel makes use of a different measure of some patterns resulting from the main activity of the systems under attack. Let us provide a few possible side channels, for the agents’ mode of operation described:

- Time spent at the host by the agent;
- Power or resources used by the agent;
- Changes in the visible agent’s characteristics (e.g. the size of the traveling agent);
- Host communication with the agent’s owner, for instance billing for the information used;
- Way that agent was hosted (priorities, status, security level, etc.). For instance in the case described in Section 2.2 the only intelligence services communications were encrypted, while all other (e.g. police, fire services) were in plain text, which allowed to separate them from the data stream.

It is interesting to note that almost all the attacks result from the countermeasures described in the previous section. The side channels are used to make the countermeasures transparent, like they were never in place. It can be shown that many of the above outlined attack opportunities originate from the countermeasures against traffic analysis. This leads to the paradox that
a side channel attack may result from the excessive use of countermeasures. For instance, when smartcards with build-in countermeasures against power analysis attacks were tested for electro-magnetic emission (e.g. [21]), it was found that they leak a lot of information. In our case consider owners requiring agents to drop information at secure locations. It forces mobile agents to carry all collected information with them. As the agent acquires data, his size will change. Although all information is encrypted, it will provide data that the host database was used. In practice it is very difficult to anticipate in advance all possible side channels. This would require complete knowledge of all the system parameters in every state of operation. Such a requirement can be easily brought to a much more philosophical question: will our understanding of Nature ever be complete?

In summary; in order to protect oneself against traffic analysis one needs to avoid any “patterns”. The type of sideline pattern can be very difficult to predict in advance. Hence, the owner has to “submerge” the activity (e.g. information requested) into the ocean of statistically non-distinguishable activities, for instance see [15]. Still, there is no guarantee that some unexpected attack resulting from a newly found sideline would not appear. The more countermeasures against traffic analysis are used, the greater the chance for more side channels. The lesson learned is that an increasing supply of privacy may benefit the attacker and hence be counterproductive.

4 Conclusions

We discussed mobile agents resistance to traffic analysis. We focused on mobile code used on a massive scale in decision making systems, with a special emphasis on trading on financial markets. In this context we discussed the interplay between two overlapping realms of cyberspace and real world. As the result, we could present the implication of these interactions on mobile agents and their owner. It became clear that such a perspective, allows to describe previously unseen complications in the system design and analysis. For instance, many side channels might appear and they are very difficult to control. This led us to a paradox of overprotection, when introducing more security into the system makes it more vulnerable for side channel attacks. As it was stated at the very beginning of this paper building a good mathematical model for the described problems is a challenging task. It seems that such a model should make use of methods from various branches of mathematics, for instance:

- Mathematical methods of AI, e.g. using fuzzy sets/logic to build traffic analysis expert systems;
• Abstract algebra, e.g. modeling agents’s and environment properties as operations within an algebraic structure;
• Game theory, e.g. see approach proposed by S. Bistarelli et al. in [3], also consult his paper in the same ENTCS volume;
• Graph theory methods (see [23]), e.g.: movement graphs/networks, graph knowledge presentation, graph grammars.

We think that an “overprotection paradox” can be generalized to all systems with a complexity exceeding a certain threshold. Claiming that adding more protection mechanisms to the system actually weakens it, seems to be very counterintuitive. The research community has so far largely considered building security as the graduate process of introducing more and more secure components that protect against different threats. However, as we demonstrated above such a perception can be false. The key issue is the interaction between various parts of a complex system, with many hidden couplings, which were not foreseen by the designers. When the system operations span different realms such interactions become extremely complicated and intractable. In such a situation, instead of adding more complexity, we would rather advocate applying the KISS principle. While there is more then one way to read this acronym, we recommend Keep It Simple and Safe.

We hope that in this short paper, we were able to sketch the problem in the way, which makes it an interesting topic for further investigations.

Acknowledgement
The paper is dedicated to the memory of Konrad Kulesza (1979-2003)—researcher, philosopher and warrior. The ideas presented originate from Konrad and Kamil discussions during Kamil’s visit to Rhodes University, Grahamstown, South Africa.

The authors want to thank Lynn Moses for editing the manuscript. We also thank Maurice ter Beek and Maciek Stanczyk for their help with typesetting.

References


