HAL
open science

Collection and analysis of attack data based on honeypots deployed on the Internet

Eric Alata, Marc Dacier, Yves Deswarte, Mohamed Kaâniche, Kostya Kortchinsky, Vincent Nicomette, Van Hau Pham, Fabien Pouget

To cite this version:

Eric Alata, Marc Dacier, Yves Deswarte, Mohamed Kaâniche, Kostya Kortchinsky, et al.. Collection and analysis of attack data based on honeypots deployed on the Internet. Workshop on Quality of protection (QoP 2005), Security Measurements and Metrics, Sep 2005, France. pp.79-92. hal-00140390

HAL Id: hal-00140390

https://hal.science/hal-00140390v1

Submitted on 6 Apr 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Collection and analysis of attack data based on honeypots deployed on the Internet

E. Alata[1], M. Dacier[2], Y. Deswarte[1], M. Kaâniche[1], K. Kortchinsky[3], V. Nicomette[1], V.H. Pham[2], and F. Pouget[2]

[1] LAAS-CNRS,
7 Avenue du Colonel Roche, 31077 Toulouse Cedex 4, France,
{ealata, deswarte, kaaniche, nicomett}@laas.fr
http://www.laas.fr
[2] Eurécom,
2229 Route des Crêtes, BP 193, 06904 Sophia Antipolis Cedex, France,
{dacier, pham, pouget}@eurecom.fr
http://www.eurecom.fr
[3] CERT-RENATER, c/o ENSAM, 151 Boulevard de lHôpital, 75013, Paris, France,
kostya.kortchinsky@renater.fr
http://www.renater.fr/Securite/CERT_Renater.htm

Abstract. The CADHo project (Collection and Analysis of Data from Honeypots) is an ongoing research action funded by the French ACI "Securité & Informatique" [1]. It aims at building an environment to better understand threats on the Internet and also at providing models to analyze the observed phenomena. Our approach consists in deploying and sharing with the scientific community a distributed platform based on honeypots that gathers data suitable to analyze the attack processes targeting machines connected to the Internet. This distributed platform, called *Leurré.com* and administrated by Institut Eurécom, offers each partner collaborating to this initiative access to all collected data in order to carry out statistical analyzes and modeling activities. So far, about thirty honeypots have been operational for several months in twenty countries of the five continents. This paper presents a brief overview of this distributed platform and examples of results derived from the data. It also outlines the approach investigated to model observed attack processes and to describe the intruders behaviors once they manage to get access to a target machine.

1 Introduction

Since the very first large distributed denial of service attacks launched in February 2000, an apparently increasing number of major security problems have been reported. In particular, a large number of worms have been observed during the last years. Surprisingly, the number of observed attacks does not seem to be influenced by the ever increasing deployment of efficient security protection tools, such as personal desktop firewalls. Is this apparent raise in the number of attacks

backed up by some undisputable data? If yes, what are the attack processes that lead to such phenomena?

As of today, we are unfortunately unable to answer these questions because of the lack of precise and unbiased data to assess the seriousness of the situation. A few qualitative indicators exist, such as, for instance, the yearly survey conducted by the Computer Security Institute (CSI) and the Federal Bureau of Investigations (FBI). However, these reports provide only high-level trends, based on statistical data obtained in various heterogeneous environments, without having a precise knowledge of the configuration of these environments. Also, the collected data is not rich enough to enable scientists to carry out rigorous analyses of the malicious behaviors at stake, and to model attack processes and their impact on the target systems security. Some companies, such as IBM, have access to a very large amount of security incident-related information collected from their customers, which, in theory, could be used to model and analyze the attack processes. In practice, however, all previous experience with such data has revealed that they are not suitable for that purpose. The main reasons lie in the complexity, diversity and dynamicity of the systems that are under scrutiny. Recently, various initiatives have been taken to monitor real world data related to malware and attacks propagation on the Internet. The Internet Telescopes, so-called blackholes/darknets and the DShield projects are among them. These projects provide valuable information for the identification and analysis of malicious activities on the Internet [2–4]. Nevertheless, such information is not sufficient to model attack processes and analyze their impact on the security of the targeted machines.

The CADHo project described in this paper is complementary to the above initiatives. It intends to address these issues by means of the following actions:

1. The project aims at deploying and sharing with the scientific community a distributed platform of honeypots [5] that gathers data suitable to analyze the attack processes targeting a large number of machines connected to the Internet.
2. The project aims at validating the usefulness of this platform by carrying out various analyses, based on the collected data, to characterize the observed attacks and model their impact on security. In particular, we will investigate how to use the modeling results to improve the design and validation of secure systems. Our objective consist in providing solid rationales for those who need to validate the fault assumptions they make when designing, for instance, intrusion tolerant systems.
3. Finally, the project aims at going beyond the study of the most frequent and automated attacks. Our objective consists here in investigating and modeling the behavior of malicious attackers once they have managed to compromise and get access to a new host. Indeed, we are not interested in monitoring all kinds of attackers. Instead, we want to monitor only those that are representative of large classes of attackers so that the knowledge derived from their observation is symptomatic of a large amount of real attacks. To fulfill this objective, we need to develop and deploy a sophisticated environment

that gives the attackers the "apparent" possibility of compromising a target system, under strict control and monitoring. This is a real challenge given the current the state of the art.

The CADHo project started in September 2004. The honeypot platform we have built has been deployed in thirty sites, from academia and industry, in twenty countries over the five continents. In the following sections, we describe our honeypot based data collection platform (called *Leurré.com*), and we present some examples of results obtained from the analysis of the data collected so far. In addition, the paper includes a preliminary discussion of the problems and the directions investigated for the modeling activities.

Section 2 presents the data collection environment *Leurré.com*. Section 3 provides a summary of the various analyses carried out on the collected data. Section 4 focuses on the modeling of attacks observed on the various honeypots deployed so far. Finally, Section 5 discusses the future evolution of the current platform toward the development of high-interaction honeypots that will enable us to model the behavior of attackers once they manage to control a target system and try to progress to defeat some security objectives.

2 The data collection environment *Leurré.com*

As mentioned in Section 1, one of the goals of the CADHo project is to share with the scientific community an open distributed platform to collect data from a large number of honeypots. This platform is deemed to evolve over the years, well beyond the end of the CADHo project. New partners are allowed to get access to the complete collected data set if and only if they agree to set up a honeypot on their premises, thus enriching the overall setup by their presence. Names of the partners are protected by a Non Disclosure Agreement that each participating entity must sign. We have developed all the required software to automate the various regular maintenance tasks (new installation, reconfiguration, log collection, backups, etc.) to ensure the long term existence of this set up.

A honeypot is a machine connected to a network but that no one is supposed to use. In theory, no connection to or from that machine should be observed. If a connection occurs, it must be, at best an accidental error or, more likely, an attempt to attack the machine. Recently, several approaches have been proposed to build environments where several honeypots are deployed. The generic term *honeynet* is used to represent them. The most visible honeynet project is the one carried out by the so called Honeynet Research Alliance [5, 6]. The Alliance is made of national entities. Some CADHo members are active members of the French one, the French Honeynet Project [7].

So far, most of the attention has been paid to implementation issues. Institut Eurécom has been working for more than a year on the definition of a low-interaction honeypot dedicated to the tasks explained here above. A first environment has been deployed, based on the VMware [8] technology. Based on

the acquired expertise during a one-year use of this environment and on the analyses carried out on the collected data, we are now convinced that, for the specific objectives of our project, the freely available software called *honeyd* [9] can be used instead of VMware. Indeed, it is known that the major drawback of *honeyd* is that an environment using that software can be remotely identified by a skilled attacker. This is less easy with VMware. Fortunately, data collected so far indicate that the risk of seeing attackers fingerprinting the environment under attack is negligible. This justifies the choice of a *honeyd* based solution.

Honeyd is a free software and it runs on various flavors of Linux and Windows. It does not consume a lot of resources and, therefore, old PCs can be used without any trouble. These are very interesting features since we are interested in building a large environment where many honeynets would run. The fact that we can add honeynets for almost no cost makes this solution very attractive. It is indeed unlikely that we could identify interested partners to join this platform on a voluntary basis otherwise.

The distributed platform Leurré.com itself is made of a potentially large number of identical honeynets deployed at the sites of our partners. All the honeynets are centrally managed to ensure that they have exactly the same configuration. This is very important if we want to keep the experiment under control. The data gathered by each honeynet are securely uploaded to a centralized database administrated by Institut Eurécom. This database contains, in a highly structured and efficient way, the complete content, including payload, of all packets sent to or from these honeynets. Furthermore, the collected data are enriched by additional information to facilitate their analysis, such as the IP geographical localization of packets source addresses, the OS of the attacking machine, the local time of the source, etc. In our context, each IP address interacting with the honeynets identifies an attacking machine. It is noteworthy that for attack processes that go through a chain of systems to attack a target, the IP address recorded in our database corresponds to the previous hop in the chain before reaching our honeynets, which does not necessarily correspond to the machine initiating the attack process.

Concretely, the distributed platform *Leurré.com* is constituted of three main components:

1. A set of computers connected to the Internet deployed at the partners sites, running *honeyd* with the same configuration. Each computer emulates three virtual machines running various operating systems (Linux RedHat, Windows 98, Windows NT) and services (ftp, web, etc.). All traffic received by or sent from each computer is saved in tcpdump files. A firewall ensures that connections cannot be initiated from the computer, only answers to external solicitations are allowed. Every day, a secured communication is established from a trusted machine during a short period of time to copy the tcpdump files archived on each computer. Integrity checks are also performed to ensure that the platform has not been compromised.

2. A centralized relational database where all the collected data are archived. All partners have the possibility to send queries to that database through a secure web interface.
3. A set of software programs that are used to collect, process and enrich the data collected from each platform. For instance, three different software are used for passive fingerprinting the OS of the attacking machines: *p0f* [10], *ettercap* [11], and *disco* [12]. *Maxmind* [13] is used to identify the geographic location of the attacks.

3 Data analysis: Summary of the main results

Several analyses have been carried out on the data collected from VMware and the *honeyd* based platforms. The results obtained from these analyses have been published in international conferences in the course of 2004-2005. An up to date list of publications on this topic can be found in [14]. In the following, we provide a short summary of the main conclusions and lessons learned from the data.

- The analyses reported in [15, 16] were based on the data collected from the initial VMware platform during a 10 month observation period. In particular, we have observed that the data exhibit a stable behavior from various perspectives, for instance with respect to the geographic location of the attacking machine, the target of the attack (virtual machine, port), etc. Such regularity suggests that there are some real values in using the data collected from honeypots to model attack processes and threats. Also, the data revealed the existence of two distinct sets of machines that targeted our honeypot platform. The first set of machines were only seeking to gather information on our environment, without really trying to perform an attack. Their activity mainly consisted in scanning our network in a systematic way considering a limited number of ports. The second set of machines (about 25%) were attacking only specific open ports of our honeypot. This suggests that they have already acquired such information from other machines belonging to first set (the scanning machines). It is noteworthy that such observation has been also confirmed by the more recent data collected from the *honeyd* distributed platform. However, we observe a higher proportion of *attacking* machines than *scanning* machines.
- A deep and thorough analysis of honeypots data is generally required to have a good understanding of malicious activity. In [17, 18], a new clustering algorithm is used to identify similar attack traces associated to attacking machines that are likely to use the same attack tool. The application of this algorithm to our honeypots data confirmed that it is very useful to highlight interesting phenomena that remain hidden if we analyze the data at a higher macroscopic level only (e.g., considering the number of attacks observed at the different ports without analyzing the root causes of the attacks).
- In the study reported in [19], we present a methodology to analyze the potential bias introduced by the use a low interaction *honeyd* platform compared

to the VMware based platform. We show that high interaction honeypots are useful to control the relevance of low interaction honeypot configurations, and that both interaction levels are required to build an efficient network of distributed honeypots.

- Finally, in [20], we present a comparative analysis of the attack processes observed on various platforms deployed at different geographic locations. In particular, we can highlight the three following observations:
  1. Some attack processes have been observed on all the platforms
  2. Some attack processes have been observed only on a subset of platforms
  3. Some attack processes have been observed only on a single platform

The results obtained suggest that the data observed from a single platform is not sufficient to characterize the malicious activity observed on the Internet. Based on the data collected so far, it seems that this is only possible for a minority of observed attacks. This result highlights the necessity to have a largely deployed distributed platform to observe the malicious activity carried out on the Internet in order to be able to derive meaningful and representative conclusions.

The results summarized above are based on the qualitative analysis of the collected data. Additional useful insights can be obtained by using mathematical modeling techniques, in particular with respect to the definition of appropriate models that can be used for prediction purposes. In the next sections, we discuss the objectives and the main problems related to this topic and we outline some examples of preliminary results to promote discussion.

4 Modeling based on the collected data

Honeypots are generally attacked by different attackers from several geographic locations all over the world. In other words, the time when the attacks are launched, their source and consequences are not known in advance. Also, the vulnerabilities exploited by the attackers and the attack scenarios might differ significantly. On the other hand, the attack results might depend on the state of the target system when the attack is initiated. All these factors are uncertainty sources that have to be taken into account in the analysis and modeling tasks carried out on the data collected from the honeypots. Statistical and probabilistic analysis techniques are well suited to take into account such uncertainties in order to: i) characterize the attackers behavior and the attack scenarios, and ii) build stochastic models and evaluate quantitative measures reflecting targeted system capacity to resist to attacks.

The data collected from the honeypots can be processed in various ways to characterize the attack processes and perform predictive analyses. For example, we can build stochastic models characterizing the frequency and the distribution of attacks taking into account the geographic location of the attackers, the IP addresses of the attacking machines, the vulnerabilities exploited, the severity of the consequences of the attacks on the target system and data, etc. In particular, modeling activities can be used to fulfill the following objectives:

1. Identify the probability distributions that best characterize the attack occurrence and attack propagation processes.
2. Analyze whether the data collected from different platforms exhibit similar or different malicious attack activities.
3. Model the time relationships that may exist between attacks coming from different sources (or to different destinations).
4. Predict the occurrence of new waves of attacks on a given platform based on the history of attacks observed on this platform as well as on the other platforms.

The approach adopted in the CADHo project to fulfill these objectives consists in exploring the application of statistical analysis and probabilistic modeling techniques that are traditionally used to model and evaluate the dependability of software and hardware based systems using data collected in operation, and extending their use to the data collected from the honeypots.

For the sake of illustration, we present in the following simple preliminary models based on the data collected from our honeypots. The examples address: i) the time-evolution modeling of the number of attacks observed on different honeypot platforms deployed so far, and ii) the analysis of potential correlations for the attack processes observed on the different platforms taking into account the geographic location of the attacking machines and the relative contribution of each platform to the global attack activity. We remind that in our context, an attacking machine is identified by an IP address interacting with our honeypots, which does not necessarily correspond to the machine initiating the attack process (see Section 2).

The data collection period considered for the examples corresponds to 46 weeks. We take into account the attacks observed on 14 honeypot platforms among those deployed so far. The honeypots selected correspond to those that have been active for almost the whole considered period. The total number of attacks observed on these honeypots is 816476. These attacks are not uniformly distributed among the platforms. In particular, the data collected from three platforms represent more than fifty percent of the total attack activity.

Let us denote by:

- $Y(t)$ the function describing the evolution of the number of attacks per unit of time observed on all the honeypots during the observation period,
- $X_j(t)$ the function describing the evolution of the number of attacks per unit of time observed on all the honeypots during the observation period for which the IP address is located in country j.

In a first stage, we have plotted, for various time periods, $Y(t)$ and the curves $X_j(t)$ corresponding to different countries j. Visual inspection showed surprising similarities between $Y(t)$ and some $X_j(t)$. To confirm such empirical observations, we have then decided to rigorously analyze the phenomena using mathematical linear regression models.

Considering a linear regression model, we have investigated if $Y(t)$ can be estimated from the combination of the attacks described by $X_j(t)$, taking into

account a limited number of countries j. Let us denote by $Y^*(t)$ the estimated model.

Formally, $Y^*(t)$ is defined as follows:

$$Y^*(t) = \sum \alpha_j X_j(t) + \beta \qquad j = 1, 2, ..k \tag{1}$$

Constants α_j and β correspond to the parameters of the linear model that provide the best fit with the observed data, and k is the number of countries considered in the regression.

The quality of fit of the model is measured by the statistics R^2 defined by:

$$R^2 = \frac{\sum (Y^*(i) - Y_{av})^2}{\sum (Y(i) - Y_{av})^2} \tag{2}$$

$Y(i)$ and $Y^*(i)$ correspond to the observed and estimated number of attacks for unit of time i, respectively. Y_{av} is the average number of attacks per unit of time, taking into account the whole observation period.

R^2 represents the proportion of total variation about the average explained by the regression. Indeed, R is the correlation factor between the estimated model and the observed values. The closer the R^2 value is to 1, the better the estimated model fits the collected data.

We have applied this model considering linear regressions involving one, two or more countries. Surprisingly, the results reveal that a good fit can be obtained by considering the attacks from one country only. For example, the models providing the best fit taking into account the total number of attacks from all the platforms are obtained by considering the attacks issued from UK, USA, Russia or Germany only. The corresponding R^2 values are of the same order of magnitude (0.944 for UK, 0.939 for USA, 0.930 for Russia and 0.920 for Germany), denoting a very good fit of the estimated models to the collected data. This result is confirmed by several statistical tests that provided significant *p-values* indicating that the data appear to be consistent with the linear regression model. For example, the estimated model obtained when considering the attacks from Russia only is defined by equation (3):

$$Y^*(t) = 44.568 X_1(t) + 1555.67 \tag{3}$$

$X_1(t)$ represents the evolution of the number of attacks from Russia. Figure 1 plots the evolution of the observed and estimated number of attacks per unit during the data collection period considered in this example. The unit of time corresponds to 4 days. It is noteworthy that, similar conclusions are obtained if we consider another granularity for the unit of time, for example one day, or one week.

These results are even more surprising that the attacks from Russia and UK represent only a small proportion of the total number of attacks (1.9% and 3.7% respectively). Concerning the USA, although the proportion is higher (about 18%), it is not significant enough to explain the linear model.

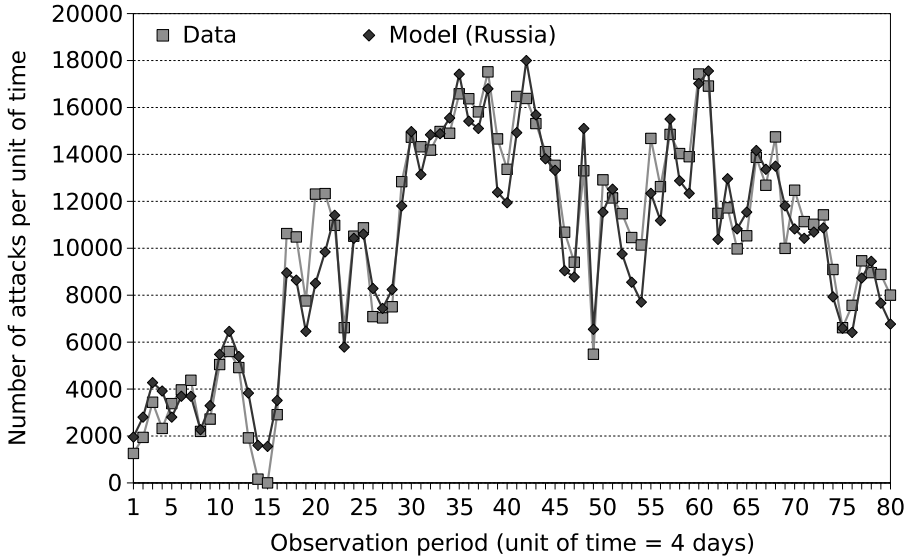

Fig. 1. Evolution of the number of attacks per unit of time observed on all the platforms and estimated model considering attacks from Russia only

The fact that the linear regression models considering attacks originating only either from UK or USA, or Russia provide a good fit to the collected data, is related to the fact that the corresponding curves present similar trends. This is illustrated on Figure 4 which represents the evolution of the Laplace factor considering the data collected from the honeypot platforms with all source countries included (Figure 2a), and the data corresponding to attacks from UK, USA or Russia only (Figures 2b, 2c and 2d). It clearly shows that there exists a striking similarity between all the curves.

As detailed in [21], the Laplace factor $u(i)$ computed at unit of time i is based on all the data observed before i. This explains the smooth evolution of the Laplace curve compared for example to Figure 1. The global and the local trends exhibited by the data are identified respectively by analyzing the sign and the variation (increase or decrease) of the Laplace factor curve. From a practical point of view, the curves presented on Figure 4 can be analyzed as follows:

- values oscillating between -2 and 2 indicate a stable behavior (i.e., there is no significant trend toward an increase or a decrease of the number of attacks per unit of time)
- positive values > 2 (respectively, negative values < -2) suggest a global trend towards an increase (respectively a decrease) of the intensity of attacks.
- decreasing or increasing values of the Laplace factor over a subinterval indicate a local decrease or increase of the intensity of attacks, respectively, for that subinterval.

It can be noticed that all the curves in Figure 4 present similar trends. In particular, significant trend changes occur almost around the same units of time (e.g., 15, 45, 54).

We have applied similar analyses by respectively considering each honeypot platform in order to investigate if similar conclusions can be derived by comparing their attack activities per source country to their global attack activities. The results are summarized in Table 3. The second column identifies the source country that provides the best fit. The corresponding R^2 value is given in the third column. Finally, the last three columns give the R^2 values obtained when considering UK, USA, or Russia in the regression mode.

It can be noticed that the quality of the regressions measured when considering attacks from Russia only is generally low for all platforms (R^2 less than 0.80). This indicates that the property observed at the global level is not visible when looking at the local activities observed on each platform. However, for the majority of the platforms, the best regression models often involve one of the three following countries: USA, Germany or UK, which also provide the best regressions when analyzing the global attack activity considering all the platforms together. Two exceptions are found with P6 and P8 for which the observed attack activities exhibit different characteristics with respect to the origin of the attacks (Taiwan, China), compared to the other platforms.

The trends discussed above have been also observed when considering a different granularity for the unit of time (e.g., 1 day or 1 week) as well as different data observation period.

To summarize, two main observations can be derived from the results presented above:

1. Some trends exhibited at the global level considering the attack processes on all the platforms together are not observed when analyzing each platform individually (this is the case for example of attacks from Russia). On the other hand, we have observed the other situation where the trends observed globally are also visible locally on the majority of the platforms (this is the case for example of attacks from USA, UK and Germany)
2. The attack processes observed locally on each platform are very often highly correlated with the attack processes originating from a particular country. The country providing the best regressions locally, does not necessarily yield good regressions when considering other platforms or at the global level. These trends seem to result from specific factors that govern the attack processes observed from each platform.

A thorough analysis of the collected data is currently carried out in order to find a sound justification of the observed trends, taking into account the different attributes characterizing the attacks. Moreover, a particular emphasis is put on the elaboration of stochastic models that can be used from a predictive point of view to forecast the attack activities to be observed on a given platform based on past observations on the same platform and on the other platforms.

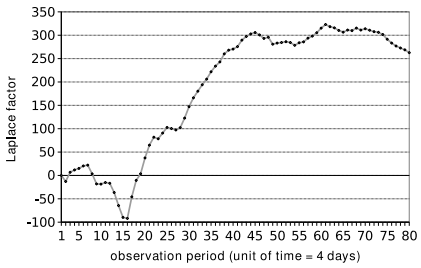

Fig. 2a. All countries

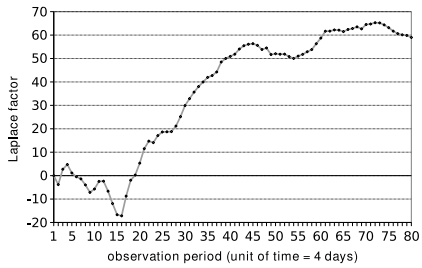

Fig. 2b. UK

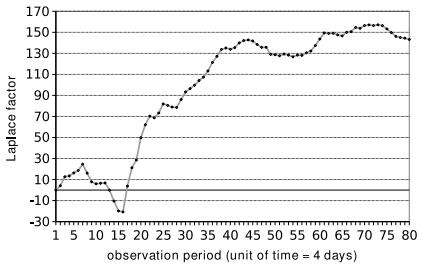

Fig. 2c. USA

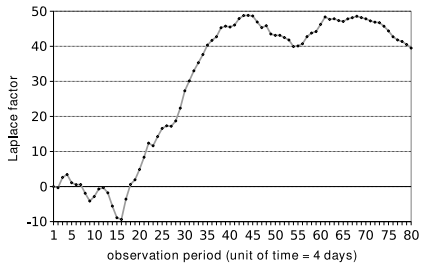

Fig. 2d. Russia

Fig. 2. Laplace Factor evolution considering attacks from all platforms, all countries included (2a), or only attacks from UK (2b), USA (2c) or Russia (2d)

Platform	Country providing the best model	R^2 Best model	R^2 UK	R^2 USA	R^2 Russia
P1	Germany	0.895	0.873	0.858	0.687
P2	USA	0.733	0.464	0.733	0.260
P3	Germany	0.722	0.197	0.373	0.161
P4	Germany	0.874	0.869	0.872	0.608
P5	UK	0.861	0.861	0.699	0.656
P6	Taiwan	0.796	0.249	0.425	0.212
P7	Germany	0.754	0.630	0.624	0.631
P8	China	0.746	0.303	0.664	0.097
P9	Germany	0.738	0.574	0.412	0.389
P10	Germany	0.708	0.510	0.546	0.087
P11	USA	0.912	0.787	0.912	0.774
P12	SPAIN	0.791	0.620	0.727	0.720
P13	USA	0.870	0.176	0.870	0.111
P14	USA	0.874	0.659	0.874	0.517

Fig. 3. Estimated models for each platform: correlation factors obtained for the countries providing the best fit and for UK, USA and Russia

5 High-interaction honeypots

The honeypots that we have already deployed in the context of this project belong to the family of so-called "low interaction honeypots". This means that their design is such that attackers have not the possibility, at any point in time, to actually get access to the machine they are attacking. This property is enforced by the fact that there is no real machine. Instead, targets are implemented by means of virtual machines without any real operating system or server to compromise. Thus, hackers can only scan ports and send requests to fake servers without ever succeeding in taking control over them.

In the CADHo project, we are also interested in running experiments with "high interaction" honeypots where attackers can really compromise the targets. Collecting data from such honeypots would enable us to study the behaviors of attackers once they have managed to get access to a target. Obviously enough, we do not want to let them use these machines for launching attacks against third party machines. Instead, we will devise a simulated environment within which they could evolve. An important feature of the environment we are planning to build is that it will "select" the attackers that we will, or will not, let compromise our machines. Indeed, we are not interested in monitoring all kinds of attackers. On the contrary, we want to monitor only those that are representative of large classes of attackers so that the knowledge derived from their observation is symptomatic of a large amount of real attacks. Such high interaction honeypots will be deployed within a limited number of highly controlled environments.

The experiments and the data that will be collected based on the high-interaction honeypots will enable us to address two distinct objectives. First, we are interested in better understanding the attack scenarios, in particular those carried out by skilled intruders. This acquired knowledge will be useful to build concrete responses and to develop tools to counter this form of attack, which is known to be very costly, but which has received little attention up to now. Second, we want to propose concrete and efficient techniques to assess the impact of such ongoing attacks on the security of the targeted system. Along this line, we propose to use observations from this setup to validate a theoretical model initially developed in our previous work on quantitative analysis of operational security in the 90s [22, 23]. The original method is a probabilistic one that differs from classical qualitative approaches (red book, ITSEC, common criteria, etc.). The core of the method lies in a so called privilege graph which highlights the various possibilities offered to an intruder to increase his privileges thanks to identified vulnerabilities or features of the system he has access to. We have shown how to use this model to derive probabilistic estimations of the ability of a system to resist attacks. These estimations are expressed as a *mean effort to security failure* (METF, similar to the MTTF measure for reliability), assessing the effort necessary for an attacker to realize a violation of a given security policy. The effort is considered as a multi-dimensional variable, taking into account the attacker competence and knowledge, the time needed to prepare and perform the attack, the efficiency of the protection mechanisms (e.g., the difficulty to guess

a given password), etc. An automatic tool has been developed to compute these measures, and has been used for a campaign of more than one year on a relatively complex system (a network of several hundred workstations in an academic environment). The results have been analyzed in detail [23], giving convincing arguments on the interest of the method, and the significance of the quantitative measures. The limitations of that approach reside in the absence of real world validation of the assumptions made about the behaviors of the intruders. Common sense has dictated our design but a more rigorous approach requires running some experiments to validate our claims. Thanks to high-interaction honeypots, this is something that now becomes feasible and something that we aim to do within the CADHo project.

6 Conclusion

The distributed data collection platform *Leurré.com* based on honeypots has been operational for many months. The data collected so far and our preliminary analyses have revealed that very interesting observations and conclusions can be derived from this data with respect to the attack activities observed on the Internet. Our objective is to deploy a large number of honeypots all around the world, in various places, in order to get comprehensive data that will allow to derive meaningful results reflecting the main phenomena that characterize the malicious activities on the Internet. It is our wish to share with the scientific community the data contained in our database. We invite all teams interested in using our data for analytical purposes to join us. All partners who accept to deploy one honeypot in their premises are allowed to have access to the database.

As summarized in the paper, the data collected can be analyzed from several perspectives, using qualitative as well as quantitative analysis and modeling techniques. Regarding modeling activities, there are several open issues that need to be addressed in future research in order to be able to build stochastic models that can be used to quantify security or to analyze from a predictive point of view the level of threat and the types of attack processes carried out on the Internet. We believe that the data collected from our honeypots, in particular, high interaction honeypots, will be very useful to identify realistic assumptions and build models that reflect the observed activities. The preliminary models discussed in this paper and the experiments that we are planning to carry out with high interaction honeypots constitute a starting point toward this objective.

References

1. ACI "Sécurité et Informatique", http://acisi.loria.fr/
2. M. Bailey, E. Cooke, F. Jahanian, J. Nazario, and D. Watson, "The Internet Motion Sensor: A Distributed Blackhole Monitoring System", Proc. 12th Annual Network and Distributed System Security Symposium (NDSS), San Diego, CA, Feb. 2005.
3. Home Page of the CAIDA Project, http://www.caida.org/, last visited 06/2005
4. DShield Distributed Detection System homepage, http://www.honeynet.org/, last visited 06/2005

5. L. Spitzner, Honeypots: Tracking Hackers, Addison-Wesley, ISBN from-321-10895-7, 2002
6. Home Page of the Honeynet Project, http://www.honeynet.org/, last visited 06/2005
7. Home page of the French Honeynet Project, http://honeynet.rstack.org/
8. VMWARE, Home page, http://www.vmware.com/
9. Honeyd Home page, http://www.citi.umich.edu/u/provos/honeyd/
10. p0f passive fingerprinting tool homepage, http://lcamtuf.coredump.cx/p0f-beta.tgz
11. The ettercap tool home page, http://ettercap.sourceforge.net/
12. The Disco tool home page, http://www.altmode.com/disco/
13. MaxMind GeoIP Country Database Commercial Product: http://www.maxmind.com/app/products/
14. F. Pouget, Publications web page, http://www.eurecom.fr/ pouget/papers.htm
15. M. Dacier, F. Pouget, H. Debar, "Honeypots: Practical Means to Validate Malicious Fault Assumptions on the Internet", Proc. 10th IEEE International Symposium Pacific Rim Dependable Computing (PRDC10), Tahiti, March 2004, pages 383-388.
16. M. Dacier, F. Pouget, H. Debar, "Attack Processes found on the Internet", Proc. OTAN Symposium on Adaptive Defense in Unclassified Networks, Toulouse, France, April 2004.
17. F. Pouget, M. Dacier, "Honeypot-based Forensics", Proc. AusCERT Asia Pacific Information Technology Security Conference (AusCERT2004), Brisbane (Australia), May 2004.
18. F. Pouget, M. Dacier, V. H. Pham, "Towards a Better Understanding of Internet Threats to Enhance Survivability", Proc. International Infrastructure Survivability Workshop (IISW04), Lisbon (Portugal), December 2004.
19. F. Pouget, T. Holz, "A Pointillist Approach for Comparing Honeypots", Proc. Conference on Detection of Intrusions and Malware & Vulnerability Assessment (DIMVA 2005), Vienna (Austria), July 2005.
20. F. Pouget, M. Dacier, V. H. Pham, "Leurré.com: On the Advantages of Deploying a Large Scale Distributed Honeypot Platform", Proc. E-Crime and Computer Evidence Conference (ECCE 2005), Monaco, Mars 2005.
21. K. Kanoun, M. Kaâniche, J-C. Laprie, "Qualitative and Quantitative Reliability Assessment", IEEE Software, Vol. 14, n2, pages 74-86, 1997.
22. M. Dacier, Y. Deswarte, M. Kaâniche, "Models and tools for quantitative assessment of operational security", Proc. 12th International Information Security Conference (IFIP SEC'96), Samos (Greece), May 1996, pages 177-186
23. R. Ortalo, Y. Deswarte, M. Kaâniche, "Experimenting with Quantitative Evaluation Tools for Monitoring Operational Security", IEEE Transactions on Software Engineering, Vol.25, N5, pages 633-650, September/October 1999.