Real time distributed detection of network attacks

Komninos Theodoros[1], *Spirakis Paul*[2], *Tsaknakis Haralampos*[1]
komninos@cti.gr spirakis@cti.gr *tsaknak@cti.gr*

[1]Research Academic Computer Technology Institute,
N. Kazantzaki, University of Patras Campus, 26500, Rio, Patras, Greece
[2]University of Patras, Department of Computer Engineering, 26500, Rio, Patras, Greece

Summary

In this work, we provide a formal model of open systems under a global attack and of distributed intrusion detection processes. The types of attacks we consider share the characteristic that upon their initiation and while they are in progress, they produce sufficient network traffic (e.g. port scanning) so that local detectors can find sufficient evidence of the attack and report it. We call such attacks *bursty*. We also postulate *properties* of local detectors that allow the construction of a fast responding *global detector*. The global detector works in two levels and it is able to suitably combine local and, possibly, inconclusive information glimpses of a suspected ongoing attack in order to decide whether an attack is actually in progress or not, accompanying this decision by a confidence level value. Our overall scheme reduces the error probability *exponentially fast* to zero as a function of *the number* of (concurrent and almost simultaneously obtained in a distributed fashion) local reports with the only requirement that only a small fraction of them reflecting the true attack status (i.e. attack or no attack). We also provide a methodology for implementing consistent local detectors that were validated using experimental traffic data. The theoretical models for intrusion and intrusion detection described in our paper have been implemented in a distributed intrusion detection system that is currently operating in a real network.

Key words:
Distributed Intrusion Detection, Network Attacks, Alert Correlation, Hypothesis Testing

1. Introduction and motivation

Most of Internet attacks to networks *start* with a probe of a set of IP addresses, looking for vulnerable servers and services. A major difficulty in arguing about a good design of network intrusion detection systems (NIDS) is the incomplete, fuzzy, qualitative characterization of attack and detection concepts. Also, the attackers of today use a great variety of methods and procedures, which make at least incomplete (and easy to confuse) most AI-based systems that "learn" attack behaviours. On the other hand, extensive network audit sequences are only good for forensic analysis (e.g. to find the attack origin), but after the damage has happened. While we do not underestimate the value of systems based on "attack signatures", we feel that they should be complemented by on-line novel schemes that are able to deduce *a generic attack,* possibly, from generic local trails of "unusual" sequences of *protocol parameters* at early communication stages. In our work here we attempt to formalize and prove good properties of detection schemes based on a very common human paradigm: The path attackers or malicious programs that try to gain access or exploit a system or network, follow the - let's call it - 7Ps sequence, that is Probe, Penetrate, Poison, Persist, Propagate, Possess and finally Paralyze the system or network under attack. For every attack to succeed a proper first target must be found e.g. a vulnerable unpatched service, a system with weak configuration or password, application coding errors, etc. To find the first vulnerable system to start with, a probe is usually performed (the so called reconnaissance phase). This suggests a distributed, two-level, approach: Use local detectors at the system, just to notice suspect information that is quite *generic.* Then, use a second (more trusted) level, inside the system, to combine/enhance the evidence. In this paper we present and justify a model of open systems and distributed attack detection schemes. We postulate *properties* for good local detectors. We then show how to construct an efficient second-level scheme of decisions that reduces the probability of an incorrect decision to zero, *exponentially fast* with respect to the number of (simultaneous, from various places) local reports. We also provide a methodology for constructing consistent local detectors and indicate that this is possible based on experimental traffic data.

2. Related work

Due to its obvious practical importance, the issue of intrusion detection has received considerable attention among researchers and commercial vendors. A sample of relevant work and further references can be found in [1], [12], [15], [17], [23]. Most of this work focuses on the collection and analysis of traffic data with intrusion incidents, as well as on the definition of rules and methodologies for defending networks and systems from malicious users. As a result, a number of intrusion detection systems, that were proved useful and practical in a variety of situations, have been proposed and tested, such as those described in [18], [25], [26], [27]. Such systems monitor deviations of remote IP sources from certain predefined rules of conduct. When a remote source violates such a rule it is banned from the network. Rules are defined in a variety of ways involving several event attributes such as destination address, destination port, time windows, flags, etc. as well as certain ad-hoc thresholds believed to reflect malicious or abnormal activity in the sequence of events. The definition of appropriate rules is generally a complicated and delicate task. A major problem in many applications and case studies appears to be the high false alarm and missed detection rates ([5], [7], [23]).

A considerable amount of research in this area focuses on the analysis of traffic sequences using data mining techniques for building anomaly detection models ([5], [9], [14], [15], [20], [21]). The ongoing work on data mining for large traffic sequences seem promising to achieving deeper insight into the intrinsic characteristics of malicious events sequences. In a recent paper ([16]), a sequential hypothesis testing approach is adopted for monitoring each remote source IP address for suspected scanning activity based on the assumption that the rates of unsuccessful connections of benign and hostile remote sources are different. Results presented there indicate performance improvement, for appropriate parameter choices, over systems based on deterministic rules.

Another recent work ([10]) adopts a similar probabilistic approach addressing the issue of scalability in the network through state aggregations. In these approaches, a centralized processing scheme is adopted and simple fixed rate Bernoulli distributions are assumed to govern both the normal and the intrusive behaviours. Another work using probabilistic methods for network attack detection has been presented in [7]. There, a target tracking formulation was proposed based on a viewpoint of the attackers as dynamically evolving systems and making use of notions such as target state trajectories and sensor measurements in the intrusion detection setting.

3. Model and definitions

In our work, we model a computer network as an undirected graph $G = (V, E)$ in the natural way: vertices in $V(G)$ are computer nodes and edges in $E(G)$ are bidirectional communication links. If v is one of its vertices, the *eccentricity* of v, denoted by $e(v)$, is equal to the maximum distance from v to any vertex of G, i.e. $e(v) = max\{d(v,w) : w \in V(G)\}$ with $d(v,w)$ the distance between vertices v, w . The *radius* of G, denoted by $r(G)$, is the minimum eccentricity among all the vertices of G, i.e. $r(G) = \min\{e(v) : v \in V(G)\}$. The *diameter* of G, denoted by $d(G)$, is the maximum eccentricity among the vertices of G, i.e. $d(G) = max\{e(v) : v \in V(G)\}$. The *center* of G is the set of vertices of G with eccentricity equal to $r(G)$, i.e. $c(G) = \{v \in V(G) : e(v) = r(G)\}$.

The definition of eccentricity of the vertices of a graph as well as its center allow us to exploit the communication structure of a computer network in order to place, optimally local and global attack detectors: local detectors are placed at nodes of maximum eccentricity and a global detector (or more than one global detectors if necessary) are placed in one of the nodes belonging to the center of the network graph.

Definition 1. *An open system Σ is defined as a graph $G(V,E)$. Let V_F be the set of vertices in G with maximum eccentricity. The nodes $v \in V_F$ are called the doors of Σ.* We assume a global discrete time sequence $t = 0,1,2,\ldots$ perhaps unknown to Σ. The time sequence may be chosen to represent the arrival of events at the doors of the system.

Definition 2. *Given a vertex $v \in V_F$, an incoming event flow at v is a sequence of information items $X_0(v), X_1(v), \ldots, X_t(v)$, where $X_t(v)$ denotes the information item arriving at v from outside the system Σ at time t.* As events we consider here the incoming packets from sources outside of the system. In the sequel, we assume that each door v of Σ is equipped with a local mechanism (usually a piece of software) which is able to watch Δ consecutive incoming events at v and decide about a possible global attack to the open system with some confidence.

Definition 3. *A local detector $W(v)$ (for $v \in V_F$ of Σ) of window Δ is a decision process that, given a subsequence $S = \{X_{t_0}(v), X_{t_0+1}(v), \ldots, X_{t_0+\Delta-1}(v)\}$ of Δ incoming events at v, it outputs $z(W) = 1$ (i.e. an attack) or $z(W) = 0$ (a normal incoming flow), together with two parameters $p(W,S)$, $q(W,S)$ defined as follows*:

- $p(W,S)$ = probability that an "attack" decision is made, given that there is indeed an attack.
- $q(W,S)$ = probability that a "no attack" decision is made, given that indeed there is no attack.

Definition 4. *A local report of a local detector $W(v)$ at $v \in V_F$ is a triple (z,p,q) where $z \in \{0,1\}$ is an attack (or no attack) decision related to a subsequence S of incoming events, and p, q are the corresponding p(W,S) and q(W,S).*

Definition 5. *A local detector $W(v)$ is consistent if $\forall S$ the following hold:*
- $Prob\{z = 1/attack\} = p(W,S)$
- $Prob\{z = 0/no\ attack\} = q(W,S)$
- $p,q \geq 1/2$

Definition 6. *A local report (z,p,q) of a local detector is useless or erroneous if the true probabilities $Prob(z=1) = Prob(z=0) = 1/2$ (regardless of the reported p,q).*

A useless report provides no information whatsoever and is assumed to be unobservable in any single report (since if we know the useless reports we may simply ignore them) occurring only with some probability $1-f$ in the sequence of reports.

Assume now that x_T is a global unobserved variable of Σ so that $x_T = 1$ means that Σ is under global attack (at some time interval T) and $x_T = 0$ means that Σ is under no attack at that time interval.

Definition7. *A set $S_n(T)$ of n local reports $(z_1,p_1,q_1)...(z_n,p_n,q_n)$ related to sub-sequences of incoming events $S_1,S_2,...,S_n$ of the same time interval T is an independent set if*

$$Prob\{z_1,...,z_n/x_T\} = \prod_{i=1}^{n} Prob(z_i/x_T)$$

In our model, we assume that local detectors located at the doors of Σ produce an independent set S_n of n local reports (for a given global time interval T). We assume that these local reports are sent, in a distributed way, to a specific internal node J of Σ that belongs to the center of the network graph, equipped with a global decision making mechanism called here a juror.

Definition 8. *A juror J of Σ is a decision making procedure located at a node belonging to the center of Σ connected (through a path) to all the doors of Σ. When an independent set $S_n(T)$ of local reports arrives at J, then J outputs $y(S_n(T))=1$ (i.e. an "attack" decision) or $y(S_n(T))=0$ (i.e. a "no attack" decision) and two probabilities $p_J, q_J \in (0,1)$ so that:*

$$Prob\{y = 1/x_T = 1\} = p_J$$
$$Prob\{y = 0/x_T = 0\} = q_J$$

The quality of a juror mechanism can be measured by the number of independent local reports which is necessary for a safe global decision.

Definition 9. *Let be a juror mechanism and $S_n(T)$ an independent set of local reports (related to the same global time interval T) given to J. Assume that we want p_J, q_J to be at least $1-\frac{1}{T^a}$, for some fixed $a > 1$. Let $n(T)$ be the minimum number of independent consistent reports for this. We call $n(T)$ the efficiency of J with respect to T.*

Definition 10. *A juror J is fast responding when $n(T)$ is a logarithmic function of T i.e. when $n(T)=O(\log T)$.*

4. Our results

The main result of this paper is the construction of a simple to implement fast responding juror. We apply our model of distributed detection in order to implement a quick and dependable detection scheme of a single attack in open networks that are allowed to receive streams of Internet data at their doors. As an example, consider the independent set of local reports

$$S_6(T) = \begin{Bmatrix} \{1,0.6,0.6\},\{1,0.6,0.6\},\{1,0.5,0.5\}, \\ \{1,0.7,0.7\},\{1,0.7,0.55\},\{1,0.55,0.6\} \end{Bmatrix}$$

Our scheme is able to judiciously extract global attack information from even such a small set. Given the very large rates of local incoming events to actual Internet sub-networks, our distributed method is, in fact, a method very close to an ideal on-line "instant" global attack detection and it is very straightforward to implement.

5. Definition and analysis of a fast responding juror

Let $(z_i, p_i, q_i),\ i=1,2,...,n$, be the independent set S_n of consistent local reports that arrive at the juror node in a predetermined discrete time interval T. The juror will choose the situation y_J (1 or 0) according to the likelihood ratio of the two situations. From the independence of S_n we have that the probability distribution of the local reports $z_1, z_2, ...z_n$ conditioned on x_T is given by:

$$Prob(z_1,z_2,...z_n / x_T) = \prod_{i=1}^{n} Prob(z_i / x_T) =$$

$$= \left(\prod_{i=1}^{n} p_i^{z_i} (1-p_i)^{1-z_i} \right)^{x_T} \left(\prod_{i=1}^{n} q_i^{1-z_i} (1-q_i)^{z_i} \right)^{1-x_T}$$

We propose the following simple decision rule for the juror:

Juror Rule 1: *If*

$$Prob(z_1,z_2,...z_n/x_T=1) \geq Prob(z_1,z_2,...z_n/x_T=0)$$

then, choose $y_J=1$ *, else, choose* $y_J=0$.

By taking logarithms, the rule becomes:

If $\sum_{i=1}^{n} a_i z_i - b \geq 0$ *then choose* $y_J = 1$ *, else, choose* $y_J = 0$ *, with*

$$a_i = \log\left(\frac{p_i q_i}{(1-p_i)(1-q_i)} \right), i = 1,...n \text{, and}$$

$$b = \sum_{i=1}^{n} \log\left(\frac{q_i}{1-p_i} \right).$$

Thus, if u_i, v_i are independent binary random variables with $Prob(u_i = 1) = p_i$, $Prob(u_i = 0) = 1 - p_i$ and $Prob(v_i = 1) = 1 - q_i$, $Prob(v_i = 0) = q_i$, $i = 1,...,n$, we have for the juror :

$$p_J = Prob\left(\sum_{i=1}^{n} a_i u_i - b \geq 0 \right),$$

$$q_J = Prob\left(\sum_{i=1}^{n} a_i v_i - b < 0 \right)$$

Here, we will make the additional assumption that the mean values of the p_i, q_i for $i=1,...,n$ are bounded away from $\frac{1}{2}$, i.e. $\frac{1}{n}\sum_{i=1}^{n} p_i \geq \frac{1+\delta}{2}, \frac{1}{n}\sum_{i=1}^{n} q_i \geq \frac{1+\delta}{2}$ for some $\delta > 0$.

We now state and prove the main theorem:

Theorem 1. *The juror* J *defined by the above decision Rule 1 is fast responding.*

Proof. We will work with p_J first.

Consider the random variable $U = \sum_{i=1}^{n} a_i u_i - b$. The probability of error is given by:

$$1-p_J = Prob(U<0) = Prob(e^{-tU}>1), for\, t>0$$

From the above relationship and by making use of Markov's inequality $(Prob(X>1) \leq E(X), X \geq 0)$ and the independence assumption, we obtain (after some algebra):

$$1-p_J \leq E(e^{-tU}) = e^{\sum_{i=1}^{n} h(t,p_i,q_i)}$$

where:

$$h(t,p_i,q_i) = log\left(q_i^t (1-p_i)^{1-t} + (1-q_i)^t p_i^{1-t} \right)$$

Note that the function $h(\,)$ defined above is convex with respect to t and also that $h(0,p_i,q_i) = h(1,p_i,q_i) = 0$ and $\frac{\partial h}{\partial t} < 0$ for $t = 0$, while $\frac{\partial h}{\partial t} > 0$ for $t = 1$. Therefore, $h()$ has a global minimum for some $t \in (0,1)$ which is always negative and can be computed by setting $\frac{\partial h}{\partial t} = 0$.

Furthermore, $h()$ is jointly concave with respect to p_i, q_i for $p_i, q_i \geq 1/2$ and $t \in (0,1)$. This can be verified by taking the Hessian $\nabla^2_{p_i,q_i} h$ and showing that it is negative definite for the given bounds. Also, $h()$ is monotonically decreasing with respect to p_i, q_i, which can be verified by taking $\frac{\partial h}{\partial p_i}$, $\frac{\partial h}{\partial q_i}$ and showing them negative for the given bounds. So, we have:

$$\sum_{i=1}^{n} h(t,p_i,q_i) \le n\, h(t,\frac{1}{n}\sum_{i=1}^{n} p_i,\frac{1}{n}\sum_{i=1}^{n} q_i) \le$$

$$\le n\, h(t,\frac{1+\delta}{2},\frac{1+\delta}{2})$$

Minimizing the above function with respect to t , we finally get the following upper bound for the probability of error:

$$1-p_J \le \left(\sqrt{1-\delta^2}\right)^n$$

Since we want $p_J \ge 1-\frac{1}{T^a}$ (i.e. $1-p_J \le \frac{1}{T^a}$), it is sufficient to have the smallest n such that $\left(\sqrt{1-\delta^2}\right)^n < \frac{1}{T^a}$, i.e. $n = a\frac{\log T}{\log \rho}+1$ for $\rho = \frac{1}{\sqrt{1-\delta^2}} > 1$, i.e. $n=O(\log T)$ as needed.

For the other probability of error, $1-q_J$, we work in an entirely analogous way to show that $1-q_J \le \left(\sqrt{1-\delta^2}\right)^n$ again.

Fault tolerance

A basic underlying assumption of the above result was that the juror can distinguish between consistent and non-consistent (or useless) reports and take into account the ones that are consistent. However, the latter assumption is not realistic in many cases as it cannot capture situations with random transmission errors in the network between local agents and jurors, unsuspected faults of the local decision mechanisms, or, more importantly, cases where intelligent adversaries are trying to mislead the detection process by creating erroneous artificial traffic. In such cases, the application of the exact decision rule based on the reported p_i and q_i may lead to uncontrollably large probabilities of error.

The presence of unknown non consistent reports randomly spread among the set of reports received by a juror necessitates a modification of the decision rule to take into account this uncertainty. Any such modification will result in more fault tolerant detectors at the expense of some degradation in detection performance. The issue is whether one can find fast responding jurors (having as efficiency a logarithmic function of time) that are fault tolerant at the same time. Another issue is the computation of the quantitative trade-off between tolerance and performance, i.e. how much performance should be sacrificed in order to achieve a certain level of fault tolerance.

We show that it is possible to have fast responding jurors that are also fault tolerant up to a given threshold in the rate of non consistent reports. The performance degradation (as a function of the threshold rate) for such fault tolerant jurors affects only the constant term of the logarithmic function of time.

Let us assume that from all the reports received by a juror a fraction f of them are consistent and the rest $1-f$ are erroneous (or useless) spread uniformly in the stream of reports. If we don't know which reports are consistent, the probabilities p_i,q_i of each report should be replaced by

$$p_i^{'}(f) = fp_i + (1-f)\frac{1}{2}, \quad q_i^{'}(f) = fq_i + (1-f)\frac{1}{2}$$

in order to reflect the true reliability status of the sequence of reports received by a juror.

Definition 11. *A juror is called f_0 - tolerant if it is fast responding for any rate f of consistent reports such that $f \ge f_0$.*

Assuming that the mean values of the reported p_i,q_i for $i=1,\ldots,n$ are bounded away from $\frac{1}{2}$, i.e. $\frac{1}{n}\sum_{i=1}^{n} p_i \ge \frac{1+\delta}{2}$, $\frac{1}{n}\sum_{i=1}^{n} q_i \ge \frac{1+\delta}{2}$ for some $\delta>0$ (as before) we can express the following result:

Theorem 2. *The juror defined by Rule 1 applied with $p_i=p_i^{'}(f_0)$ and $q_i=q_i^{'}(f_0)$ for each received report is f_0 - tolerant.*

Proof. Indeed, if we follow similar steps as in the proof of Theorem 1 replacing p_i,q_i by $p_i^{'}(f_0)$, $q_i^{'}(f_0)$ and observe the properties of the functions involved, we will finally get the following bounds for the probabilities of error:

$$1-p_J \le \left(\sqrt{1-(\delta f_0)^2}\right)^n, \quad 1-q_J \le \left(\sqrt{1-(\delta f_0)^2}\right)^n$$

The efficiency of the above juror $n(T)$ (smallest number of reports in a time interval T such that $1-q_J \le \frac{1}{T^a}$) is obtained as:

$$n(T) = \frac{2a\log T}{\log\left(\frac{1}{1-\left(\delta f_0\right)^2}\right)} = O\left(\log T\right)$$

So, the juror is fast responding (as its efficiency is a logarithmic function of time) as long as the rate f of consistent reports satisfies $f \ge f_0$. The performance degradation of the rule (in comparison to the situation where all reports used by the juror are consistent) is manifested by the term δf_0 which affects only the constant factor of the above relationship.

Limiting case. What happens if f_0 becomes small tending to zero? Given that $\log\left(\frac{1}{1-\left(\delta f_0\right)^2}\right)\approx\left(\delta f_0\right)^2$ for small f_0, the above relationship can be simplified as follows: $f_0=\frac{\sqrt{2a}}{\delta}\sqrt{\frac{\log T}{n(T)}}=O\left(\sqrt{\frac{\log T}{n(T)}}\right)$

Therefore, it is possible to have fast responding $f_0-tolerant$ jurors even in cases where f_0 tends to 0 as long as its rate of decrease to 0 is slower than $O\left(\sqrt{\frac{\log T}{n(T)}}\right)$. The form of the decision rule for such a juror as $f_0\to 0$ is obtained as a limiting case of an $f_0-tolerant$ juror as follows:

$$\sum_{i=1}^{n}\left(z_i-\frac{1}{2}\right) \mid \begin{array}{l}\geq 0\ decide\ "attack" \\ <0\ decide\ "no\ attack"\end{array}$$

This result indicates that in cases where an adversary (a hostile user or nature) has the ability to "inject" erroneous reports in the stream of reports arriving at the juror at very high rates, the best way to go is to employ the above $f_0-tolerant$ decision rule which treats all reports equally regardless of the reported p_i,q_i and the associated with them weighting factors. This is equivalent to majority voting. Such a juror will still arrive at a correct decision with high probability as long as the adversary's actions cannot corrupt the stream of reports arriving at the juror faster than $1-O\left(\sqrt{\frac{\log T}{n(T)}}\right)$.

6. Consistent local detectors

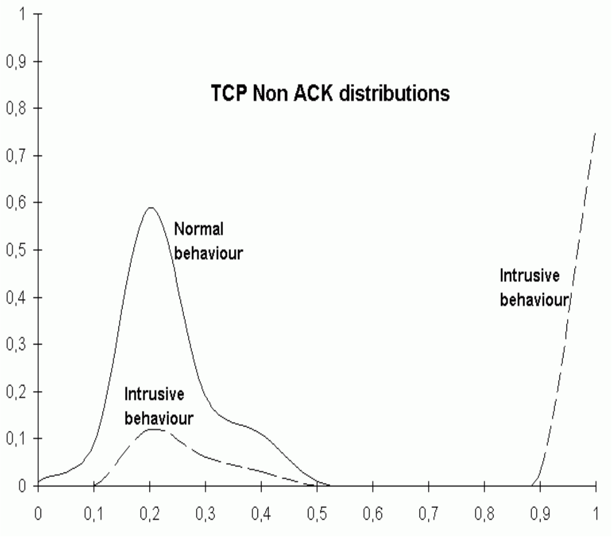
We briefly describe here the methodology that we are using for the construction of consistent local detectors. Given a sequence of locally observed events $X1, X2, X3, \ldots$ we divide it into equal segments of N events each. We call each segment an ``epoch". Given a property G (which is a set of rules defined among certain attributes of the events and which is believed to offer information for distinguishing locally attacks from non-attacks), we can measure the empirical distribution of the proportion of events satisfying these rules. Let N_G^r denote the number of epochs for which the proportion of events satisfying property G is equal to r. Then, if TN is the total number of epochs we define the following empirical distribution:

$$\mathrm{Prob}(r)=\frac{N_G^r}{TN}, r\in[0,1].$$

Such a distribution can be determined for two different situations: (a) when an attack is present and (b) when no attack is present.

Determining which properties G are appropriate to consider for constructing such distributions (and subsequently for building local detectors), was based on studying the behavior of some common types of attacks against hosts in a computer network, as well as on extensive experimentation with simulated traffic data. For example, host scanning is an important element of several attacks, including most malicious programs epidemics. Such attacks can be considered as bursty and, thus, are within our model. Several worms (Code Red-II, Nimda, etc) propagate via scanning other hosts for vulnerabilities. In scanning for TCP ports, the attacker sends several special packets to various destinations. A host can then watch the specified protocol negotiations during the *start* of a communication and can observe the special flagged packets probing a particular port and figure out (for example) if a remote address is trying to open communication to a closed port.

We have experimented by constructing local distributions over a small window of observation Δ using simulated traffic data for the two situations (attack and non-attack) and for various sizes of the epoch size N. The property considered in these experiments was based on attribute values flag =``ACK" or ``ACK/PUSH" versus values of flag =``non-ACK". The results obtained indicate a significant difference in the shapes of the distributions between the two situations. A snapshot of such distributions that were obtained experimentally is shown in the following figure:



The next step is to use standard results in statistics ([22]) to build local detectors. Let us denote by H_1 and H_0 the hypotheses ``attack" and ``not attack" respectively and let $f_1(r), f_0(r)$ denote the two distributions respectively.

Given an observation sequence $r_1, r_2, \dots, r_n$ and a false alarm constraint $1-q$, the local decision rule takes the following form of a likelihood ratio test:

Decide H_1 *if* $L(r_1, r_2, \dots, r_n) = \prod_{i=1}^{n} \frac{f_1(r_i)}{f_0(r_i)} \geq d$,

else, decide H_0.

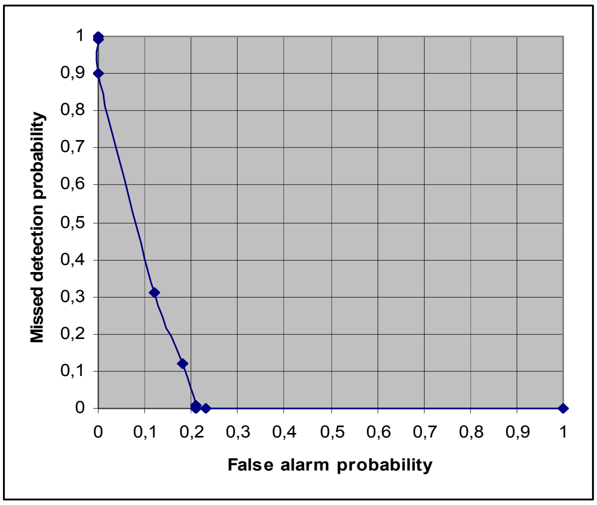
The threshold d above and the detection power p are computed as functions of the specified false alarm probability $1-q$ according to the following equations:

$$\sum_{L(r_1,r_2,\dots,r_n)\geq d} \prod_{i=1}^{n} f_0(r_i) = 1-q$$

$$p = \sum_{L(r_1,r_2,\dots,r_n)\geq d} \prod_{i=1}^{n} f_1(r_i).$$

The application of the above procedure leads to a decision, either H_1 or H_0, accompanied by a pair of probabilities p and q expressing the confidence level in that decision.

The performance of a local detector can be depicted by plotting the probability of missed detection ($1-p$) versus the probability of false alarm ($1-q$) (a form of the ROC curve of a detector). As an example, in the figure below we show the ROC curve obtained from the experimental distributions of the previous figure:



The above figure indicates that it is possible to build local detectors that are not only consistent (as is our major concern here) but also of high information content, since, we can have, for example, reports with missed detection probability 0.2 and false alarm probability 0.15, i.e. with $p=0.8$ and $q=0.85$, which are very desirable. However, it should be pointed out that during the early stages of the transition from one situation ("non-attack") to the other ("attack") the differences are not so large. In order to take that into consideration we have considered classes of distributions for each situation (*attack* and *non-attack*) and performed the local hypothesis testing based on the least favorable pair of distributions within these classes, i.e. the pair that leads to the worst possible detection performance of the local detector for a given false alarm rate. This procedure leads to smaller detection power. Nevertheless, it was necessary for obtaining robust local detectors. Furthermore, since the complexity of the computations in performing the hypothesis testing depend mainly on the forms of the two distributions, we considered approximations by using exponential families of distributions. Such approximations seemed to adequately fit the experimental data while facilitating the computations involved.

Our experiments were conducted using as a basis a network intrusion detection system operating at CTI for the protection of the organization's network hosts from malicious scanning. This system, called "Helena" ([18]), was initially implemented as a distributed system based on monitoring scanning attempts of *unused* TCP ports (i.e. TCP ports with no corresponding service running on the host) reported by agent modules installed in several hosts of the network. Incidents reported by agent modules to a central node, called ``juror", are combined by checking the overall scanning behaviour banning from the network those IP source addresses whose behavior violate a predefined threshold. This system is a special case of the logic of the system proposed in this paper. Currently we are extending the logic of the system "Helena" to incorporate the model proposed here.

7. The overall performance of our distributed detector

Since our juror J is fast responding (as established in Section 5) it is crucial that:

- The $\Theta(\log T)$ local reports to arrive to J fast.
- The local reports are independent.

To satisfy the first requirement, we can place the juror at a node belonging to the center of $G(V,E)$. This can be done in advance after taking into account the topology of the network. To satisfy the second requirement, we can safely assume that different local detectors are independent since they observe different local incoming information flows, conditioned only on the event of a global attack or not. This was assumed in our definition of *conditional* independence.

8. Conclusions

We proposed here a model and a very fast way of distributed detection of attacks on an open network. We proved that even small and not reassuring (in any strong way) distributed glimpses of information can be effectively and concurrently combined in a way that reduces *exponentially* to zero the probability of error in our decision, when the number of reports grows.

Extensions of this work that we are currently examining involve:

- Multiple attack strategies.
- Further investigation about what kinds of events can make consistent local detectors.
- Techniques of combining randomization (e.g. sampling, noise) with the local reports in order to increase independence and to break the *local appearance* of *systematic confusion patterns.*

Acknowledgments

The authors would like to express their cordial thanks to Mr. E. Valeodis, Mr. Ch. Yannakopoulos and Prof. Y. Stamatiou for their help during the preparation of this paper.

References

[1] Allen J., Christie A., et.al.: State of Practice of Intrusion Detection Technologies. Tech. report, Carnegie Mellon University http://www.cert.org/archive/pdf/99tr028.pdf

[2] Axelsson S.: Intrusion Detection Systems: A Survey and Taxonomy. Dept. of Computer Eng., Chalmers Univ. of Technology.

[3] Bace R.: Intrusion Detection. Macmillan Tech. Publishing (2000).

[4] Barbara D., Wu N., Jajodia S.: Detecting Novel Network Intrusions Using Bayes Estimators. Proc. of the first SIAM International Conference on Data Mining (SDM'01) (2001).

[5] Bloedorn E., Talbot L.,et. al.: Data Mining applied to Intrusion Detection: MITRE Experiences. IEEE International Conference on data Mining (2001).

[6] Boyd S. & Vandenberghe L.: Convex Optimization. Cambridge University Press (2004).

[7] Burroughs D.J., Wilson L.F., Cybenko G.: Analysis of Distributed Intrusion Detection Systems Using Bayesian Methods. IPCCC (2002).

[8] Chernoff H.: A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. Annals of Mathematical Statistics, 23 (1952) 493--507.

[9] Clifton C., Gengo G.: Developing Custom Intrusion Detection Filters Using Data Mining. Military Communications International Symposium (MILCOM2000) (2000).

[10] Compella R.R., et. al..: On Scalable Attack Detection in the Network. IMC'04 (Oct 2004) Italy.

[11] Cover T.M., Thomas J.A.: Elements of Information Theory. Wiley (1991).

[12] Debar H., Dacier M., Wespi A.: A Revised Taxonomy for Intrusion Detection Systems. Annales des Telecommunications **55** (7-8): 361-378.

[13] Johansen K., Lee S., CS424 Network Security: Bayesian Network Intrusion Detection (BNIDS) (2003): http://www.cs.jhu.edu/fabian/courses/CS600.424/course_papers/samples/Bayesian.pdf

[14] Julisch K.: Mining Alarm Clusters to Improve Alarm Handling Efficiency. Proc. 17th Annual Computer Security Applications Conference (2001).

[15] Julisch K.: Data Mining for Intrusion Detection. In ``Applications of Data Mining in Computer Security'', Kluwer Academic Publisher, Boston (2002).

[16] Jung J., Paxson V., et.al.: Fast Portscan Detection Using Sequential Hypothesis Testing. Proc. of the IEEE Symposium on Security and Privacy (May 2004).

[17] Komninos T., Spirakis P.: Dare the Intruders, Ellinika Grammata and CTI Press (2003).

[18] Komninos T, Spirakis P., Stamatiou et.al..: A Software Tool for Distributed Intrusion Detection in Computer Networks (Helena) (Best Poster presentation in PODC 2004).

[19] Kreidl O.P., Frazier T. M.: Feedback Control Applied to Survivability: A host-Based Autonomic Defence System. IEEE Trans. On Reliability, Vol. 53, No. 1 (March 2004).

[20] Lee W., Xiang D.: Information-Theoretic Measures for Anomaly Detection. Proc. of the 2001 IEEE Symposium on Security and Privacy (2001).

[21] Lee W., Stolfo S.J., Mok K.W.: Mining Audit Data to Build Intrusion Detection Models. Proc. IEEE Symposium on Security and Privacy (May 1999).

[22] Lehmann E.L.: ``Testing Statistical Hypotheses'', 2nd edition, Springer Texts in Statistics.

[23] Lippmann R.P., Fried D.J., et.al..: Evaluating Intrusion Detection Systems: The 1998 DARPA off-line Intrusion Detection Evaluation. Proc. of the 2000 DARPA Information Survivability Conference and Exposition.

[24] Mukherjee Biswanath, Heberlein Todd L., and Levitt Karl N.: Network intrusion detection. IEEE Network, **8** (3):26--41 (May/June 1994).

[25] Paxson V.: Bro: a system for detecting network intruders in real time. Computer Networks, 31 (23--24): 2435--2463, Amsterdam, Netherlands (1999).

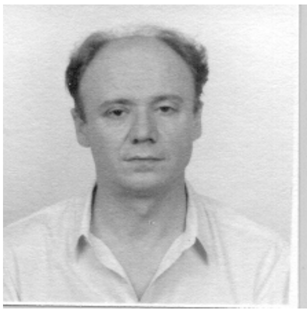
[26] Porras P.A., Neumann P.G.: EMERALD: Event monitoring enabling responses to anomalous live disturbances. National Information Systems Security Conference, Baltimore, MD (October 1997).

[27] Roesch M.: Snort: Lightweight intrusion detection for networks. Proc. 13th Conference on Systems Administration (LISA-99), pp. 229--238, Berkeley, CA (Nov. 7--12 1999) USENIX Association.

Mr. Theodore Komninos received his MSc in Computer and Networks Security (2000) from Department of Computer Engineering and Informatics, School of Engineering, University of Patras, Greece, his MBA (1993) from the Hellenic Management Association. He owns a diploma in Computer Engineering and Informatics (1989) from the University of Patras and a diploma in Civil Engineering (1986) from the same University. In 1989 he joined RACTI and he is now member of the BoD and Director of Educational Technology Sector, Director of Systems & Networks Support Sector and Director of Networking and Information Systems Security Sector. He is lead auditor and Special Advisor for Information, Systems and Network Security for the Greek Ministry of Education, a member of Network Specialists of the Greek Research Network (GRNet-member of GEANT) and has extensive experience of CSF programs. Mr. Komninos is supervising Postgraduate Diploma Thesis and Master Thesis in the area of Information, Systems & Network Security, Distributed Networking Intrusion Detection Systems and Information Warfare. His research interests include Information, Systems & Network Security, Distributed Networking Intrusion Detection Systems, Information Warfare, Design of innovative environments for Intrusion Detection Systems, and New Technologies in Education. He is also co-author of the book “Strengthening Security of Systems and Networks. Dare the intruders”, Greek Letters-CTI Press, 2003 (in Greek) and author of the 4th Chapter titled “Choosing the right computer equipment for secondary schools” in the book “Time is the Judge” (pony translation of Greek title).

Prof. Paul Spirakis (google: Paul Spirakis) born in 1955, obtained his PhD from Harvard University, USA, in 1982. He served as a postdoctoral researcher at Harvard University and as an assistant professor at New York University, (the Courant Institute**).** He was appointed as a Full Professor in the Department of Computer Science and Engineering of Patras University (Greece) in 1990.

Paul Spirakis was honored several times with international prizes and grants (e.g. NSF), also the top prize of the Greek Mathematics Society. He was acknowledged between the top 50 scientists worldwide in Computer Science with respect to "The best Nurturers in Computer Science Research", published by B. Kumar and Y.N. Srikant, ACM Data Mining, 05. He was appointed as a Distinguished Visiting Scientist of Max Planck Informatik. Paul Spirakis is the Director of the Research Academic Computer Technology Institute (RA.CTI). His research interests include Algorithms and Complexity and interaction of Complexity and Game Theory. He has extensively published in most of the important Computer Science Journals and most of the significant refereed conferences. He has edited various conference proceedings and is currently an Editor of Several Prestigious Journals. He has published two books through Cambridge University Press, and eight books in Greek. Paul Spirakis was the Greek National Representative in the Information Society Research Programme (IST) from January 1999 till June 2002. He was elected unanimously as one of the two vice-Presidents of the Council of the European Association for Theoretical Computer Science (EATCS). He has been a member of ISTAG (Information Society Technologies Advisory Group) a prestigious body of about 40 individuals advising EU for research policy, form January 2003 to January 2005. He consults for the Greek State, the European Union and several major Greek Computing Industries.



Dr. Tsaknakis Haralampos received his Ph.D. and M.Sc. degrees in Electrical Engineering from the University of Connecticut, U.S.A., in 1986 and 1983 respectively, and a Diploma in Electrical and Mechanical Engineering from the Aristotelian University of Thessaloniki, Greece, in 1979. From 1997 until today he has been with the Research Academic Computer Technology Institute (R.A.CTI), Greece, where he is currently the Project Manager of the Technical Consultancy Program between CTI and the Greek Ministry of Education, managing the implementation of a large variety of IT projects. At R.A.CTI, he is also involved in the development of models and methods for the security of computer systems and networks. From 1995 to 1997, he had been with Intrasoft S.A. (now Intrasoft International) as a senior consultant for business development of the company’s IT products and services in the European market and as Project Manager of large IT projects. From 1986 to 1994 he was with Alphatech, Inc., U.S.A., (now BAE Systems – Advanced Information Technologies) as a senior research engineer designing innovative special purpose algorithms for large scale Command Control Communication Surveillance and Intelligence Systems. From 1981 to 1986 he was a research and teaching assistant at the University of Connecticut, U.S.A. working towards his Ph.D. at the Department of Electrical Engineering & Computer Science. From 1980 to 1981 he was a research scientist at the Department of Electrical Engineering of the Aristotelian University of Thessaloniki, Greece, where he conducted research in the area of Electromagnetics. Dr. Tsaknakis has published several papers in international journals and conferences in subjects including Estimation & Filtering, Data Fusion, Combinatorial Optimization, Control Theory, Probability, Stochastic Processes & Statistics. He received awards from the Greek National Fellowship Foundation (1973-74, 1976-77), from the University of Connecticut, U.S.A. (1981-82) and from Alphatech, Inc., U.S.A., (1991). His main current research interests are in mathematical modeling, optimization, computational algorithms and distributed detection systems.