International Journal of Pure and Applied Mathematics

Volume 34 No. 1 2007, 37-48

ON A PROBABILISTIC MODEL OF INTRUSION DETECTION

Alexander Grusho[1], Elena Timonina[2],
Zeev (Vladimir) Volkovich[3] §, Zeev Barzily[4]

[1]Moscow State University
GSP-2, Leninskie Gory, Moscow, 119992, RUSSIA
e-mail: grusho@yandex.ru

[2]Russian State University for the Humanities
25 Kirovogradskaya, Moscow, 113534, RUSSIA
e-mail: eltimon@yandex.ru

[3,4]Software Engineering Department
ORT Braude College, P.O. Box 21982, Karmiel, ISRAEL
[3]e-mail: vlvolkov@ort.org.il
[4]e-mail: zbarzily@ort.org.il

[3]Department of Mathematics and Statistics
The University of Maryland
Baltimore County, USA

Abstract: In this paper we consider a probabilistic model of anomaly based intrusion detection systems. The model represents a network-like computer system by events' sequences in an appropriate functional space. This space is a kind of the Tychonoff product space. The model is described by means of a stochastic tree operated in a discrete time mode. Necessary and sufficient conditions for the existence of a strictly consistent sequence, of statistical tests, for unauthorized access detection are revealed. Modifications of these conditions, for the inferences by subsequences, are presented. The paper includes

Received: October 16, 2006

§Correspondence address: Software Engineering Department, ORT Braude College, P.O. Box 21982, Karmiel, ISRAEL

two appropriate examples that exhibit the results obtained.

AMS Subject Classification: 62F03, 62F05, 60G20, 62P99
Key Words: hypothesis testing, asymptotic properties of tests, generalized stochastic processes

1. Introduction

Intrusion detection has become a crucial defense tool in the today's information security systems due to the growing vulnerabilities of modern computing networks. Often procedures of the authentication and access control are not able to prevent hostile attacks.

Intrusion detection approaches can be divided into two groups: signature-based systems and anomaly based systems. Both approaches have disadvantages. For instants, signature-based mechanism can detect known attacks but cannot expose innovative intrusions. Conversely, anomaly based systems can find out intrusions but cannot identify any intrusion from the audit trails.

One of the main targets of intrusion is unauthorized access to valuable information. This action usually takes place in parallel to legal information processing. Although, it may be accompanied by alterations of the attributes (including statistical) distinguishing an attack from the normal system behavior.

Let us consider the problem of detecting illegal information outflow that can happen on the background of legal information activity. Intrusion can be identified by a sequence of generated events. An event can be understood as an element of the process which can be recognized and identified. Event's attributes consist of the contents, the time, the place and the result of the its realization. In particular, accesses to objects, the transmission of information and granting of access permission can be interpreted as events.

The events connected to manipulations with valuable resources normally take place during the standard system's activity. However outflow of these information is possible. In this situation the events accompanying treatment of the valuable resources probably are not presented in the traditional context or cause new no-conventional events. Valuable resources can be involved in numerous procedures in vary combinations. Hence, discovering of information leakage resting up a statistical procedure on the background of the legal activity is a complicated problem. Anomaly based systems are trying to solve such problems by means of investigating of events reflected in the audit data.

These systems are expected to detect abnormal behavior (anomalies) in a network. They perform within the hypothesis that attacks are distinct from "normal" (legitimate) activities and can, as a result be discovered, by systems created to identify such differences. Etalons of the normal behavior are constructed by profiles based on the historical data during periods of the normal activity. The system keeps event data and exploits a set of measures to uncover deviates of the monitored operation from the standard one. Parametric statistical measures fit the distributions of appropriate features to a certain prototype. Non-parametric ones build these distributions by real data observations.

In this paper we consider a probabilistic model of anomaly based intrusion detection systems. The model represents network-like computer system by events' sequences in appropriate functional space. This framework allows us to reveal conditions for the existence of suitable statistical rules for unauthorized access detection.

The paper is organized in the following form. In Section 2 we introduce and discuss the major model and proof the basic theorem according to the conditions for the existence of a strictly consistent sequence of statistical tests for anomaly behavior detection. Section 3 is devoted to a generalization of the presented construction to include inferences using periodic tracing of the system's performance. The last section contains two examples that exhibit the results obtained.

2. Model

Our model, of the anomaly based intrusion detection approach, deems the system as a stochastic tree. The starting even is placed in the root of the tree (root event). Nodes of the tree represent events which can occur after root event. The arches of the tree are directed from the root to the nodes. Each arch specifies that the destination node occurs only if the source node occurred. Two or more events can have the same contents. However, any event is unique characterized by its attributes like the time, the place in the hierarchy and the name. We consider a discrete time system and the events start and end only in this epochs. Several events can happen at once.

Let X_h denote the set of events at layer h. Namely:

$$X_h = \{x(h,1), ..., x(h,n_h)\},$$

then it is assumed that each subset $y_h \subset X_h$ can be observed. For instance, y_h can be the empty set, i.e. $y_h \in 2^{X_h} = Y_h$, where 2^{X_h} is the set of all subsets

of X_h. Thus, the system's behavior is described by a sequence of events' sets from different levels. We model the system by infinite sequences not including the root event:

$$(y_1,\ y_2, ...,\ y_h, ...)\ \ y_h \in Y_h,\ h = 1, 2, \dots .$$

On the space of these series

$$\prod_{h=1}^{\infty} Y_h = Y^{(1)},$$

we introduce the σ-algebra $\mathcal{A}$ generated by the cylindrical sets

$$(z_1, ..., z_n) \times \prod_{h \geq n+1} Y_h, : z_i \subseteq Y_i.$$

We denote

$$\prod_{h \geq n} Y_h = Y^{(n)}.$$

It is assumed that the standard system behavior is represented by a probability measure P_0 on the measurable space $(Y^{(1)}, \mathcal{A})$. Obviously, this measure is set by a consistent family of finite-dimensional distributions on the cylindrical sets:

$$P_{0,1,...,n}(z_1, ..., z_n) = P_{0,n}(z_1, ..., z_n), : z_i \subset Y_i.$$

An abnormal system's performance is characterized by a parametric set

$$\{P_{1,:\theta}, : \theta \in \Theta\},$$

of alternative distributions on $(Y^{(1)}, \mathcal{A})$, described by an appropriate family of consistent finite-dimensional distributions. Occurrence of related families can be caused by a variety of reasons such as information leakage, an attack's reflection or emerging of a new non recognized tasks. The problem of anomaly detection is to identify a deviation from the normal activity resting up numerous initial steps n of the monitored sequence. Note, that the distributions P_0 and $\{P_{1,:\theta}, : \theta \in \Theta\}$ are assumed to be known. In this context an agent constructs a test T_n for the hypothesis

$$H_0^{(n)} : P = P_{0,\,n} \tag{1}$$

against the general alternative

$$H_1^{(n)} : P \in \{P_{1,:\theta}, : \theta \in \Theta\}. \tag{2}$$

We denote α_n is the significance level of the test and S_n the critical area and $W_n(\theta)$, $\theta \in \Theta$ is the test power.

Definition 1. A test sequence $T_n,\ n = 1, 2, ...$ is called strictly consistent if:

— if the following hold:

$$\lim_{n\to\infty} \alpha_n = 0.$$

— for each $\theta \in \Theta$:

$$\lim_{n\to\infty} W_n(\theta) = 1. \tag{3}$$

— for each n

$$\overline{S}_n \times Y^{(n+1)} \supset \overline{S}_{n+1} \times Y^{(n+2)}, \tag{4}$$

where

$$\overline{S}_n = (\prod_{h=1}^{n} Y_h) \backslash S_n.$$

Note that each strictly consistent sequence is also a consistent one (see [2]). Now, we are going to state a theorem which provides sufficient and necessary conditions for the strictly consistency of a test sequence.

For a not increasing family of cylindrical sets, $\mathcal{I} = \{I_n\},\ n = 1, 2, ...,\ I_n \subset Y^{(1)}$, we denote

$$A_0(\mathcal{I}) = \bigcap_{n=1}^{\infty} I_n. \tag{5}$$

Theorem 1. *A strictly consistent test sequence for testing (1) against (2) exists if and only if there is a non-increasing sequence of cylindrical sets, $\mathcal{I} = \{I_n\},\ n = 1, 2, ...,\ I_n$, such that:*

1. $P_0(Y^{(1)} \backslash A_0(\mathcal{I})) = 0$;

2. for each $\theta \in \Theta$, a set $A_1(\theta) \subset Y^{(1)}$ can be picked with the properties:

(a) $A_0(\mathcal{I}) \cap A_1(\theta) = \emptyset$;

(b) $P_{1,:\theta}(Y^{(1)} \backslash A_1(\theta)) = 0$.

Proof. Sufficiency. Let us represent

$$I_n = D_n \times Y^{(n+1)}, D_n \subset \prod_{h=1}^{n} Y_h,$$

and let us take $S_n = \overline{D}_n$ as the critical areas of the constructed tests. Here

$$\overline{D}_n = (\prod_{h=1}^{n} Y_h) \backslash D_n.$$

The sequence $I = \{I_n\}$ is non increasing thus

$$D_n \times Y^{(n+1)} \supset D_{n+1} \times Y^{(n+2)},$$

i.e. the sequence S_n satisfies (4).

Let us prove (3). It is easy to see that

$$\overline{D_n \times Y^{(n+1)}} = \overline{D}_n \times Y^{(n+1)},\ n = 1, 2, ...$$

and

$$\overline{D}_n \times Y^{(n+1)} = \overline{D_n \times Y^{(n+1)}} = Y^{(1)} \backslash (D_n \times Y^{(n+1)}) \subset Y^{(1)} \backslash A_0(\mathcal{I}).$$

In addition,

$$Y^{(1)} \backslash A_0(I) = \bigcup_{n=1}^{\infty} \overline{I}_n = \bigcup_{n=1}^{\infty} (S_n \times Y^{(n+1)}),$$

and

$$P_0(S_n \times Y^{(n+1)}) \leq P_0(Y^{(1)} \backslash A_0(I)).$$

On the other hand, according to the definition of the significance level we get

$$0 = P_0(S_n \times Y^{(n+1)}) = P_{0,n}(S_n) \leq \alpha_n.$$

It is obvious that

$$\lim_{n\to\infty} W_n(\theta) = \lim_{n\to\infty} P_{1,\theta,n}(S_n) = \lim_{n\to\infty} P_{1,\theta}(S_n \times Y^{(n+1)})$$

$$= P_{1,\theta}(\bigcup_{n=1}^{\infty} (S_n \times Y^{(n+1)})) = P_{1,\theta}(Y^{(1)} \backslash A_0(\mathcal{I})).$$

By the theorem's conditions a set $A_1(\theta)$ satisfying

$$A_0(\mathcal{I}) \cap A_1(\theta) = \emptyset$$

exists for each $\theta \in \Theta$, i.e.

$$A_1(\theta) \subseteq Y^{(1)} \backslash A_0(\mathcal{I}).$$

It gives us

$$1 = P_{1,\theta}(A_1(\theta)) \leq P_{1,\theta}(Y^{(1)} \backslash A_0(\mathcal{I})).$$

Sufficiency is proved.

Necessity. Let test sequence $\{T_n\}$, $n = 1; 2; ..;$ be a strictly consistent tests' sequence. We introduce

$$D_n = \overline{S_n}, \quad n = 1, 2, ...$$

and consider

$$I_n = D_n \times Y^{(n+1)}.$$

It is easy to see that:

$$I_n \supset I_{n+1}, \quad n = 1, 2, ...$$

Along with the theorem conditions we see

$$\lim_{n\to\infty} P_{0,n}(S_n) \leq \lim_{n\to\infty} \alpha_n = 0.$$

At the same time

$$\lim_{n\to\infty} P_{0,n}(S_n) = P_0(\bigcup_{n=1}^{\infty}(S_n \times Y^{(n+1)})) = P_0(Y^{(1)} \backslash A_0(\mathcal{I})),$$

where $I = \{I_n\}$. For each $\theta \in \Theta$ we obtain, from the theorem's requirements,

$$\lim_{n\to\infty} W_n(\theta) = P_{1,\theta}(\bigcup_{n=1}^{\infty}(S_n \times Y^{(n+1)})) = 1.$$

Therefore, for every $\theta \in \Theta$, we can put

$$A_1(\theta) = \bigcup_{n=1}^{\infty}(S_n \times Y^{(n+1)}) = Y^{(1)} \backslash A_0(\mathcal{I}).$$

The theorem is proved. □

3. Inferences via Subsequences

Typically, a network cannot be monitored all the time by an intrusion detection agent. Therefore one of the appropriate strategies appears to be the selective periodic tracing of the system's performance. In this section we modify by taking into account the possibility of the selective tracing.

Let us introduce a binary sequence $\varepsilon = \{\varepsilon_i\}$ such that $\varepsilon_i = 0$ in the case where an agent cannot inspect at time i, and $\varepsilon_i = 1$ otherwise. We denote by $\varepsilon^{(1)}$ the subsequence of all ones in ε and consider

$$Y_{\varepsilon}^{(1)} = \prod_{i_k \in \varepsilon^{(1)}}^{\infty} Y_{i_k}.$$

The mapping

$$F_\varepsilon : Y^{(1)} \longrightarrow Y_\varepsilon^{(1)}$$

converts the set $Y^{(1)}$ to the thinned, according to ε, set $Y_\varepsilon^{(1)}$. Each set Y_i can be endowed with the discrete topology. Subsequently, the spaces $Y^{(1)}$ and $Y_\varepsilon^{(1)}$ can be considered as topological spaces with the product Tychonoff topology (see, for example [3]). This topology is produced by the cylindrical subsets. Moreover, the compactness of the sets Y_i implies the compactness of the sets $Y^{(1)}$ and $Y_\varepsilon^{(1)}$. It easy to see, that the mapping F_ε is continuous for each ε, since the preimage of each cylindrical subset in $Y_\varepsilon^{(1)}$ is also a cylindrical subset of $Y^{(1)}$.

Let $\mathcal{A}_\varepsilon$ be the minimal σ -algebra generated by the cylindrical subsets of $Y_\varepsilon^{(1)}$.

Lemma 1. *For each binary sequence the function F_ε is $(\mathcal{A}_\varepsilon, \mathcal{A})-$measurable.*

Proof. A proof of the lemma is similar to the proof of Lemma 4 in (see [1]). For any cylindrical set I_n the prototype $F_\varepsilon^{-1}(I_n) = I'$ is also a cylindrical set. Here, Cartesian products of the spaces Y_i are fed between any two subsequent elements of the initial part, corresponding to the zero elements of the sequence.

To conclude the proof we take advantage of a method offered in (see [4], Chapter 2). Let D be the system of all measurable sets from A_ε whose prototypes lay in A. The class D is closed under countably infinite intersections, countably infinite unions and under the complement, i.e. D is a $\sigma-$algebra which includes all cylindrical sets. It follows that D is a subclass of $\sigma-$algebra A_ε, generated by the cylindrical sets. It means that the set D includes the minimal $\sigma-$algebra generated by the cylindrical sets of the space $Y_\varepsilon^{(1)}$. Hence $D = A_\varepsilon$ and the lemma is proved.

Another proof of the lemma consists of the following: F_ε is a continuous function, in the Tychonoff product topology, because the preimage of any open set is also open. Then this function is measurable in the Tychonoff product spaces endowed with the Borel $\sigma-$algebras. These $\sigma-$algebras coincide with A and A_ε since the appropriate topologies have a countable topological basis.

Based on this lemma we can introduce probability distributions on the measurable space $(Y_\varepsilon^{(1)}, A_\varepsilon)$:

$$P_0'(A) = P_0(F_\varepsilon^{-1}(A));\ P_{1,\theta}' = P_{1,\theta}(F_\varepsilon^{-1}(A)) : \theta \in \Theta.$$

Studied problem is reduced like (1) and (2) to the testing of the hypothesis

$$H_0^{(n)} : P = P'_{0,\,n} \tag{6}$$

against the general alternative

$$H_1^{(n)} : P \in \{P'_{1,:n,\theta},: \theta \in \Theta\}. \qquad \square \tag{7}$$

Theorem 2. *If there is a strictly consistent sequence* $T_n,\ n = 1, 2, ..$ *for testing (1) against (2) and there exist sets* $A_0(\mathcal{I})$, $A_1(\theta),: \theta \in \Theta$ *satisfying the conditions 1 and 2 of Theorem 1 such that*

$$F_\varepsilon(A_0(\mathcal{I})) \cap F_\varepsilon(A_1(\theta)) = \emptyset$$

for every $\theta \in \Theta$. *Then a strictly consistent sequence* $T_n,\ n = 1, 2, ..$*of tests for testing (6) against (7) exists.*

Proof. Let us take a set $A_0(I)$ where $I = \{I_n\},\ n = 1, 2, ...,$ is a non-increasing sequence of cylindrical sets. It is easy to see that

$$\mathcal{I}' = \{I'_n\} = \{F_\varepsilon\,(I_n)\}$$

is a non-increasing sequence of cylindrical sets in $Y_\varepsilon^{(1)}$. We introduce

$$B_0(\mathcal{I}') = \bigcap_{n=1}^{\infty} I'_n,$$

and show that

$$F_\varepsilon(A_0(\mathcal{I})) = B_0(\mathcal{I}').$$

If $\omega \in A_0(I)$ then due to (5) $\omega \in I_n$ and $F_\varepsilon(\omega) \in I'_n = F_\varepsilon(I_n)$. Therefore

$$F_\varepsilon(\omega) \in \bigcap_{n=1}^{\infty} I'_n = B_0(\mathcal{I}'),$$

and $F_\varepsilon(A_0(I)) \subset B_0(I')$.

Conversely, if $\omega' \in B_0(I')$, then $\omega' \in F_\varepsilon(I_n), n = 1, 2, ...$ It implies

$$F_\varepsilon^{-1}(\omega') \cap I_n \neq \emptyset, \qquad n = 1, 2, ... \tag{8}$$

The set $F_\varepsilon^{-1}(\omega') \cap A_0(I)$ is a closed set in the topology of $Y^{(1)}$, due to the fact that $F_\varepsilon^{-1}(\omega')$ is a closed one too. It provides (see, for example [3]) that $F_\varepsilon^{-1}(\omega') \cap A_0(I)$ is a compact set which can be represented by the form

$$F_\varepsilon^{-1}(\omega') \cap A_0(I) = \bigcap_{n=1}^{\infty} (F_\varepsilon^{-1}(\omega') \cap I_n).$$

If this set is empty, then due to its compactness there is a finite subfamily of I having the empty intersection (see [3]). Because of the monotonicity of the sequence I a number N exists such that

$$F_\varepsilon^{-1}(\omega') \cap I_N = \emptyset.$$

But it contradicts (8). Since

$$F_\varepsilon^{-1}(\omega') \cap A_0(\mathcal{I}) \neq \emptyset.$$

Thus, we can see, that $\omega \in A_0(I)$ exists, being a preimage of ω', according to F_ε and $B_0(I) \subset F_\varepsilon(A_0(I))$. The proved equality and the theorem conditions show us that the sufficient conditions of Theorem 1 are held. Application of this theorem complites the proof. □

4. Examples

4.1. Example 1

Let us suppose that, within the conditions of Theorem 1, the support of the measures P_0 is finite. It means that the information's transition from the root can be made by a finite sequence of events. Therefore, the support of the measure P_0 is a closed set and, A_0 is an intersection of a not growing sequence of cylindrical sets. The essential requirement is that the support of any alternative, $A_1(\theta)$, is disjoint with the support of A_0. From the security point of view, it suggests that all admissible ways from the set A_0 are checked, and it is made sure that these prompts are not connected to the information leakage actions. In this case, according to Theorem 1, there is a strictly consistent sequence for testing the hypothesis (1) against the alternative (2). The information leakage is, obviously, represented by the alternative hypothesis.

Note, that this conclusion does not depend on a probability model. In any case, such a sequence can be created. This results from the fact that, the acceptance region, of the hypothesis $H_0^{(n)}$, is constructed, for every n, as a collection of all initial intervals, of the length n, from A_0. The hypothesis $H_0^{(n)}$ is rejected, if there is a chain of events which is not included in this acceptance

set. For a periodic sequence, the information leakage occurrence has to be checked, in each round, to prove its absence in the whole sequence.

An existence of a strictly consistent test sequence does not guarantee the nonappearance of the information leakage even when A_0 is finite. It is caused by the fact, that an occurrence of a sequence, not contained in A_0, can be shown only for a large n. On the other hand, an existence of a strictly consistent test sequence provides some evidence to the protection's efficiency.

Let us present a modification of this example by employing Theorem 2. We assume that an IDS agent can trace events, according to some random sequence ϵ, containing an infinite set of ones. Let

$$F_\varepsilon(A_0) \cap F_\varepsilon(A_1(\theta)) = \emptyset, \tag{9}$$

for every $\theta \in \Theta$. For a Bernoulli distributed series ε, the empty intersection occurs with probability 1 and, $F_\varepsilon(A_0)$ is finite being a limit of a cylindrical sets' sequence. Here, Theorem 2 yields the existence of a strictly consistent test sequence.

4.2. Example 2

Let A_0 be a cylindrical set. Namely a number n exists such, that

$$A_0 = I_n = D_n \times Y^{(n+1)}.$$

Obviously, A_0 is a limit set of a not growing sequence of cylindrical sets. A strictly consistent test sequence can be found, if

$$A_1(\theta) \cap A_0 = \emptyset,$$

for every $\theta \in \Theta$. It means that an initial part, of length n, of any admissible alternative cannot be a subset of D_n. In this case, a strictly consistent test sequence can be defined by $T_n = T_1$, $n = 2, 3, ...$ It accepts the null hypothesis if an observed events' sequence belongs to D_n. Otherwise, the hypothesis is rejected. Here again, the significance level is zero and the power of the test is one.

Such a situation arises, when the manipulations influence completely dissipates during n steps. Therefore the deviation from the normal behavior can be detected only within an initial interval having a length n.

Let a sequence F_μ be chosen such that (9) holds. Recall, $F_\mu(I_n) = I'_n$ is a cylindrical set. Theorem 2 supplies a strictly consistent test sequence.

Acknowledgments

This work was supported by the Russian Foundation for Basic Research, Grant No. 04-01-00089.

References

[1] A. Grusho, A. Kniazev, E. Timonina, Detection of illegal information flow, In: *Proceedings of Third International Workshop on Mathematical Methods, Models, and Architectures for Computer Network Security, MMM-ACNS, St. Petersburg: Springer – LNCS 3685* (2005), 235-244.

[2] E.L. Lehmann, *Testing Statistical Hypotheses*, Springer Texts in Statistics, Springer, Second Edition (1997).

[3] J. Munkres, *Topology*, Prentice Hall, Second Edition (1999).

[4] A.N. Shiryaev, *Probability*, Springer, Second Edition (1995).