Finding Reliable Threshold for Intrusion Detection*

W. Szpankowski
Department of Computer Science
Purdue University
W. Lafayette, IN 47907
http://www.cs.purdue.edu/people/spa

April 25, 2001

*joint work with P. Flajolet, Y. Guivarc'h, and B. Vallee, to appear in *ICALP 2001*, Crete, Greece.

Pattern Matching Approach to Intrusion Detection

Recently, pattern matching found many applications the intrusion detection.

The main idea is to search in **audit file** (called further a *text*) for certain **patterns** (known also as **signatures**) representing suspicious activities that might be indicative of an intrusion by an outsider, or misuse of the system by an insider. The key to this approach is to recognize that these patterns are **subsequences**, not substrings (i.e., consecutive symbols).

Example:

The pattern `date` occurs as a subsequence in the text `hidden pattern`, in fact *four* times, but *not even once* as a string.

Motivation

There are several well-recognized difficulties in doing misuse intrusion detection using pattern matching. Two important problems are:

(1) The first one is of an **algorithmic nature** and boils down to find an efficient and fast pattern matching algorithm that recognizes all (or some) occurrences of subsequences.

(2) The second problem is **more fundamental** and it reduces to the following questions:

> *How many occurrences of a signature constitute a real attack?*
>
> *When a subsequence can be viewed a signature?*

In other words, how to set a **threshold** so that we can detect only real intrusions and avoid false alarms? It is clear that **random** errors occurs (actually quite often) and setting the threshold too low will lead to unrealistic number of false alarms. On the other hand, setting the threshold too high may result in missing some attacks, which is even more dangerous.

Hidden Pattern Occurrences

We fix an alphabet $\mathcal{A} := \{a_1, a_2, \ldots, a_r\}$. The text is $T_n = t_1 t_2 \cdots t_n$. A particular matching problem is specified by a pair:

$$(\mathcal{W}, \mathcal{D})$$

where the *pattern* $\mathcal{W} = w_1 \cdots w_m$ is a word of length m; and the *constraint* $\mathcal{D} = (d_1, \ldots, d_{m-1})$ is an element of $(\mathbf{N}^+ \cup \{\infty\})^{m-1}$.

An m-tuple $I = (i_1, i_2, \ldots, i_m)$ $(1 \leq i_1 < i_2 < \cdots < i_m)$ satisfies the constraint $\mathcal{D}$ if

$$i_{j+1} - i_j \leq d_j,$$

in which case I is called a *position*.

Let $\mathcal{P}_n(\mathcal{D})$ be the set of all positions subject to the separation constraint $\mathcal{D}$, satisfying furthermore $i_m \leq n$.

An *occurrence* of pattern $\mathcal{W}$ in the text T_n of length n subject to the constraint $\mathcal{D}$ is a position $I = (i_1, i_2, \ldots, i_m)$ of $\mathcal{P}_n(\mathcal{D})$ for which $t_{i_1} = w_1$, $t_{i_2} = w_2, \ldots, t_{i_m} = w_m$.

Let $\Omega_n(\mathcal{D})$ be the number of $\mathcal{W}$ occurrences in T. Mathematically speaking, we must investigate

$$\Omega_n(\mathcal{D}) = \sum_{I \in \mathcal{P}_n(\mathcal{D})} X_I$$

where $X_I := [\![\mathcal{W} \text{ occurs at position } I \text{ in } T_n]\!]$ with $[\![B]\!] = 1$ if the property B holds, and $[\![B]\!] = 0$ otherwise (Iverson's notation).

Mathematical Formulation of the Problem

Our interest lies in determining reliably the threshold α_{th} such that if the number $\Omega_n(\mathcal{W})$ of subsequence occurrences in T exceeds α_{th}, then with high probability, say $1-\beta$ (e.g., one can set $\beta = 10^{-5}$), there is a **real** intrusion. We assume that the text T is generated by a random source (e.g., memoryless or Markov source) while the pattern $\mathcal{W}$ is *given*.

In the above probabilistic setting the number of subsequence occurrences Ω_n is a random variable. How to find the threshold α_{th}? Observe that if we set $\alpha_{\text{th}} = \mathbf{E}[\Omega_n]$, where $\mathbf{E}[\Omega_n]$ is the average value of Ω_n, then **random** number of the signature detection will lead to an alarm. We definitely do **not** want this to happen. We rather find $\delta > 0$ such that

$$\Pr\{|\Omega_n - \mathbf{E}[\Omega_n]| > \delta \mathbf{E}[\Omega_n]\} < \beta$$

where β is very small (say $\beta = 10^{-5}$). The above says that the probability of deviating from $(1 \pm \delta)\mathbf{E}[\Omega_n]$ is smaller than β. Now we set the threshold

$$\alpha_{\text{th}} = (1+\delta)\mathbf{E}[\Omega_n].$$

This assure us that with very low probability (namely, β) we start a false alarm or undetect an attack.

Theoretical Results

Theorem 1. *Consider a general constraint $\mathcal{D}$ and the number of occurrences $\Omega_n \equiv \Omega_n(\mathcal{D})$. The mean and variance of Ω_n satisfy*

$$\begin{aligned} \mathbf{E}[\Omega_n] &= \frac{\pi(\mathcal{W})}{b!}\Big(\prod_{j\in\mathcal{F}} d_j\Big)\, n^b \left(1+O(\frac{1}{n})\right), \\ \mathbf{Var}[\Omega_n] &= \sigma^2(\mathcal{W}) n^{2b-1}\left(1+O(\frac{1}{n})\right), \end{aligned}$$

where $\mathcal{F}$ is the set of j such that $d_j < \infty$, and the "variance coefficient" $\sigma^2(\mathcal{W})$ involves the autocorrelation $\kappa(\mathcal{W})$

$$\begin{aligned} \sigma^2(\mathcal{W}) &= \frac{\pi^2(\mathcal{W})}{(2b-1)!}\,\kappa^2(\mathcal{W}) \\ \kappa^2(\mathcal{W}) &:= \sum_{(I,J)\in\mathcal{B}_2^{[1]}} \left(\frac{1}{\pi(\mathcal{W}_{I\cap J})}-1\right). \end{aligned}$$

The set $\mathcal{B}_2^{[1]}$ is the collection of all pairs of occurrences (I,J) that intersect exactly once.

Limiting Distribution

Theorem 2. *The random variable Ω_n asymptotically follows the Central Limit Law:*

$$\lim_{n\to\infty} \Pr\left\{\frac{\Omega_n - \mathbf{E}[\Omega_n]}{\sqrt{\mathbf{Var}[\Omega_n]}} \leq x\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt \qquad (1)$$

In words, Ω_n is approximately distributed as a normal distribution with mean $\mathbf{E}[\Omega_n] = O(n^b)$ and $\mathbf{Var}[\Omega_n] = O(n^{2b-1})$.

Large Deviation for the Constrained Case

To find the threshold α_{th} we need a result in large deviations. Let now $d_j = d < \infty$ for all j.

Theorem 3. *In the constrained case, large deviations from the mean have exponentially small probability:*

$$\Pr\{|\Omega_n - \mathbf{E}[\Omega_n]| > \delta \mathbf{E}[\Omega_n]\} \leq 2 \exp\left(-\frac{\pi^2(w)}{m^2 d^2}\frac{n\delta^2}{2}\right),$$

Thus to compute the threshold

$$\alpha_{\text{th}} = (1+\delta)\mathbf{E}[\Omega_n]$$

we need δ that can be found from

$$2 \exp\left(-\frac{\pi^2(w)}{m^2 d^2}\frac{n\delta^2}{2}\right) = \beta$$

where β is the (given) *confidence level* (e.g., $\beta = 10^{-5}$). For this we need to know $\pi(\mathcal{W})$ which must be computed experimentally from the audit file (text). We shall use several novel estimation technique (e.g., context-tree weighting algorithm) to estimate $\pi(\mathcal{W})$.

Experiments

We took the full text of Hamlet and search for the pattern ("*The law is Gaussian*"):

$$\mathcal{W} = \mathsf{thelawisgaussian}$$

and its mirror image $\tilde{w}$, corresponding to $m = 16$.

Theory predicts $\Omega_n = 1.330\,10^{48}$.

Experiments give $E = \ 1.365\,10^{48}$.

a deviation of less than 4% from what is expected.

More detailed analysis is presented in the next two tables.

Tables

Value of d	Expected (E)	Occurred (Ω)	Ratio Ω/E
12	2.76777E+01	0	0.00
13	9.19522E+01	0	0.00
14	2.79468E+02	693	2.47
15	7.86648E+02	1,526	5.46
18	1.21199E+04	31,385	2.58
20	5.88656E+04	124,499	2.11
25	1.67306E+06	2,527,148	1.51
30	2.57769E+07	40,001,940	1.55
40	1.92891E+09	2,757,171,648	1.42
50	5.48229E+10	76,146,232,395	1.38
∞	1.33098E+48	1.36554E+48	1.025

What we observe is more frequent occurences than what we predict—although the discrepancy is not large. The phenomenon is perceptible for low values of d. The reason is clearly that the pattern is chosen from English and there is some advantage to having it contains frequent groups like '**the**'. A similar phenomenon is observed for **brigitte** who has an advantage (*cf.*, bring, bright, brim, etc.).

Tables

To check this interpretation we look at the same pattern but in reverse, namely, $w =$ naissuagsiwaleht. Then, we find:

Value of d	Expected (E)	Occurred (Ω)	Ratio Ω/E
14	2.79468E+02	371	1.32
15	7.86648E+02	2,379	3.02
18	1.21199E+04	14,123	1.16
20	5.88656E+04	41,066	0.69
25	1.67306E+06	1,277,584	.76
30	2.57769E+07	25,631,589	.99
40	1.92891E+09	2,144,491,367	1.11
50	5.48229E+10	48,386,404,680	0.88
∞	1.33098E+48	1.38807E+48	1.042

The pattern has fewer letters that form "meaningful" groups and the deviation from what is expected no longer displays a systematic bias.

More Experiments

To find **signatures** we must decide how to choose the gaps d. If we choose d too large, then almost **all** words occur as hidden patterns. If we choose d too small, then we may miss some patterns. We want to choose d such that the occurrence of $\mathcal{W}$ is statistically significant.

In the previous experiment we set d constant. Our analysis predicts that the pattern might start occurring near $d = 10$, while its presence is unlikely for smaller values, $d < 10$. In fact, w starts occurring at $d = 14$ while $\tilde{w}$ starts at $d = 13$—a deviation of some 30–40% from what the model predicts.

		$w =$ thelawisgaussian		$\tilde{w} =$ naissuagsiwaleht	
d	Expected (E)	Occurred (Ω)	Ω/E	Occurred (Ω)	Ω/E
13	9.195E+01	0	0.00	18	0.19
14	2.794E+02	693	2.47	371	1.32
15	7.866E+02	1,526	5.46	2,379	3.02
18	1.211E+04	31,385	2.58	14,123	1.16
20	5.886E+04	124,499	2.11	41,066	0.69
30	2.577E+07	40,001,940	1.55	25,631,589	0.99
50	5.482E+10	76,146,232,395	1.38	48,386,404,680	0.88
∞	1.330E+48	1.36554E+48	1.03	1.38807E+48	1.04

What's Next?

More theory:

1. Precise large deviations to get a more precise threshold.

2. ADAPTIVITY. Since the attacker may change his/her signature, we must quickly react to it. To analyze it, we consider a **set** of hidden words instead of a single hidden word.

More Experiments:

1. How to estimate $\pi(\mathcal{W})$?

2. Implementation for a real audit file.

3. Selection of signatures.