New adaptive batch and sequential methods for rapid detection of network traffic changes with emphasis on detection of "Denial–of–Service" attacks.

Rudolf B. Blažek, Hongjoong Kim, Boris Rozovskii
University of Southern California, Center for Applied Mathematical Sciences
1042 W. 36th Place, DRB155
Los Angeles, CA 90089, U.S.A.
blazek@math.usc.edu, hjkim@math.usc.edu, rozovski@math.usc.edu

Abstract

Dans les réseaux informatiques, les attaques à grande échelle dans leur phase finale peuvent être facilement identifiées en observant les changements abrupts dans l'encombrement du réseau. Cependant, dans leur phase initiale, ces changements sont difficiles à détecter et ne se distinguent pas des fluctuations habituelles de ce dernier. Dans cet article, nous développons des méthodes efficaces de type séquentiel et de "groupe adaptatif" pour une détection plus rapide des attaques dites "attaques de refus de service". Ces méthodes se fondent sur l'analyse statistique des données venant de différents niveaux du protocole pour détecter d'infimes changements dans l'encombrement du reseau, qui sont typiques de ce genre d'attaques (i.e. détection du changement de point). Les deux méthodes séquentielles et de groupe utilisent un étalonnage par test statistique afin de parvenir à fixer un niveau de fausse–alerte. De plus, ces méthodes une fois combinées peuvent être utilisées pour détecter des attaques de "refus de service" dans un sens plus large. D'autre part, elles sont autodidactes, ce qui leur permet de s'adapter à une variété d'encombrements de réseau et de types d'utilisation. Des structures théoriques pour ces deux familles de tests, ainsi que des résultats issus de simulations sont presentés ci-dessous.

1. Introduction

In computer networks, large scale attacks in their final stages can readily be identified by observing very abrupt changes in the network traffic. But in the early stage of an attack, these changes are hard to detect and difficult to distinguish from usual traffic fluctuations. In this article, we develop efficient adaptive batch and sequential type methods for an early detection of attacks from the class of "Denial-of-Service Attacks". These methods employ statistical analysis of data from multiple layers of the network protocol for detection of very subtle traffic changes, which are typical for these kinds of attacks (i.e. change-point detection). Both the batch and sequential methods utilize thresholding of test statistics to achieve a fixed rate of false alarms. In addition, these methods can be used for detection of generalized denial-of-service attacks consisting from combinations thereof. Moreover, both methods are self-learning, which enables them to adapt to various network loads and usage patterns. Theoretical frameworks for both kinds of tests, as well as results of simulations, are presented.

2. Sequential and Batch Detection

The main idea is that the structure of an information system can be described by a stochastic model, and that a failure or an attack leads to an abrupt change of the structure. Among the main approaches to detecting such an event, the most important are *fixed–size batch detection*, and *sequential change–point detection*. In the latter setting the problem is formulated as detection of a change in the model as soon as possible after its occurrence. This approach is based on the change–point detection theory, which is a generalization and modification of Wald's sequential analysis (hypotheses testing).

Both approaches involve two performance indices to measure the performance of the method: the *rate of false alarms* and the *detection delay*. In contrast to standard statistical algorithms, *the sequential detection algorithms minimize the average detection delay for a given false alarm rate*. Therefore the sequential approach is preferable.

It is known that the CUSUM detection procedure introduced by Page (1954) and Shiryaev's detection procedure (Shiryaev, 1961) are both optimal when the observations are i.i.d. in pre-change and post-change modes. To be specific, both procedures minimize average delay to the change detection for a given false alarm rate (frequency of false alarms). Page's procedure may be considered as a Wald's test with reflection from zero barrier. Recent advances in the general change-point detection theory allow us to design similar detection procedures that are also optimal for general statistical models (no i.i.d. assumption is involved) in the sense of minimizing the average detection delay for a low false alarm rate. The sequential procedures in question are optimal, and at the same time have manageable computational complexity, since they are based on finite memory filters.

3. Measurable Characteristics of the Network traffic Flow

We observe information related to the headers, sizes and other characteristics of the received and transmitted packets. For example, in the transport layer we observe the number of TCP packets categorized by size or type (ACK, SYN, URG etc.), the numbers of UDP packets and their sizes, the source and destination port for each packet etc. In the application layer we plan to observe information about packets associated with a given application. Of interest is also the size of buffers related to received and sent SYN packets and similar information. Among other important characteristics of the network traffic flow are, for example, service delay and percentage of expired requests sent to a monitored network server.

In our current experiments, which are described below, we simultaneously observe the following network flow statistics:

$N_{\rm pt}^{k,i}$ ... total number packets of type pt with sizes in i-th bin received during the k-th time interval, and

$B^k_{\rm SYN}$ ... SYN packet induced buffer size observed at the end of the k-th time interval,

where the packet type pt is either ICMP, UDP, or TCP. The packets sizes are categorized into a number of size bins; the time intervals $T_1, \ldots, T_M$ are fixed size non-overlapping time intervals. Notice that all the statistics are chosen so that a large value suggests a possible attack.

4. Sequential Detection Methods

In our sequential detection method we use simultaneous thresholding of sequential statistics

$$S^{k,i}_{pt} = \max\{0, S^{k-1,i}_{pt} + N^{k,i}_{pt} - m^{k,i}_{pt}\},$$

where pt is one of the three packet types (ICMP, UDP, or TCP), and $m^{k,i}_{pt}$ represents a historical estimate of $E(N^{k,i}_{pt})$. If any statistic $S^{k,i}_{pt}$ exceeds a threshold $h^{k,i}_{pt}$, then an alarm message is sent to the decision making engine. In the case of the observed buffer sizes $B^k_{\rm SYN}$ we use a similar statistic $S^k_{\rm SYN} = \max\{0, S^{k-1}_{\rm SYN} + B^k_{\rm SYN} - m^k_{SYN}\}$.

As shown in the graph on the right in Figure 1, the information about the patterns of regular traffic flow is updated when a statistic $S^{k,i}_{pt}$ reaches and departs the 0 level. If the decision making engine reports that a previously issued alarm message was a false alarm, then the information about regular traffic patterns and thresholds will be updated accordingly (future development).

5. Batch–Sequential Approach

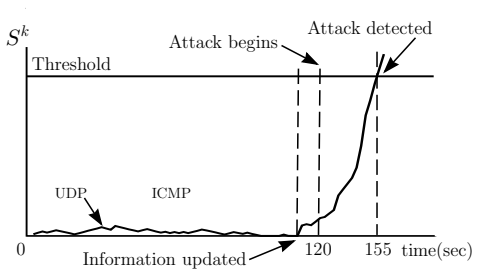
To utilize the advantages of the sequential methods also in the case of fixed–size batch detection, we use a batch–sequential modification inspired by the above sequential approach. For each packet type pt (i.e. ICMP, UDP, or TCP) we define a batch statistic

$$\chi^2_k = \sum_{i=1}^{M_{pt}} \frac{(N^{k,i}_{pt} - N^k_{pt} p^i_{pt})^2}{N^k_{pt} p^i_{pt}},$$

and for thresholding we use modified batch–sequential statistics

$$S_{pt,k} = \max\{0, S_{pt,k-1} + \chi^2_{pt,k} - \mu_{pt,k}\},$$

where N^k_{pt} is the total number of pt–packets received in the k-th interval, and p^i_{pt} are the historical probabilities for their i–th size bin. The main advantage of these statistics is that for each packet type we only observe one statistic which detects an anomalous departure from the recent distribution of packet sizes.



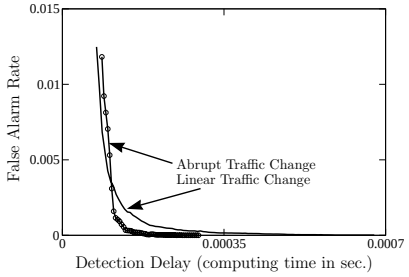


Figure 1: Left: An illustration of sequential change–point detection. Here one particular run of a simulated UDP DoS attack is shown. Right: A summary of the relationship between the detection delay and the false alarm rate for simulated SYN DoS attack. The detection delay is measured using the CPU time used by the detection algorithm.

6. DoS Attack Simulations

For simulations we have used a network with 100 nodes configured into a transit–stub topology. The network contained one transit domain, four transit nodes, and 12 stub domains with 96 nodes. Under regular conditions, the traffic consisted of approximately 5% ICMP packets, 15–20% UDP packets, and 75–80% TCP packets. The attacker's activity represented less than 1% of traffic. We have considered two scenarios for the simulated DoS attacks: with an abrupt traffic change and with a linear traffic change. Under the former scenario, the attacker's traffic abruptly changed to approximately 20% of all traffic, while under the former scenario the attacker's traffic rapidly increased to the level of 20% of all traffic in a linear fashion during a 60 second interval. Afterwards, the hostile traffic remained at the level of 20%. The graph on the right in Figure 1 depicts the observed relationship between the false alarm rate and the detection delay under both scenarios. The results of the simulations show that the algorithms work well in the sense that the simulated DoS attacks are detected in their early stages, well before the hostile traffic reaches it's full potential.

REFERENCES

Anděl, J. (1985), *Matematická Statistika. SNTL – Nakladatelství Technické Literatury, Praha, the Czech Republic* (In Czech).

Basseville, M. and Nikiforov, I.V. (1993), *Detection of Abrupt Changes: Theory and Applications. Prentice Hall, Englewood Cliffs.*

Brodsky, B.E. and Darkhovsky, B.S. (1993), *Nonparametric Methods in Change-Point Problems. Kluwer, Dordrecht.*

Shiryaev, A.N. (1963), *On optimum methods in quickest detection problems. Theory Probab. Appl.*, 8, 22–46.

Tartakovsky, A.G., *Extended asymptotic optimality of certain change-point detection procedures: non-i.i.d. case. Annals of Statistics* (submitted for publication).