Work in progress

Detecting Early Worm Propagation Based on Entropy

Hanxun Zhou
Dept. of Information Science and Engineering
Northeastern University
Shenyang, China
zhouhx@neusoft.com

Yingyou Wen
Software Center
Northeastern University
Shenyang, China
wenyy@neusoft.com

Hong Zhao
Software Center
Northeastern University
Shenyang, China
zhaoh@neusoft.com

ABSTRACT

In this paper, we present a router-based system to identify worm attacks by computing entropy values of selected packet attributes. We first compute during a training phase a profile of entropy values of the selected packet attributes. Then Chebyshev's inequality is utilized after the training phase to calculate the normal bound of entropy value with a low probability of a false positive. The detector compares new data against the bound and generates an alert when the new input exceeds the normal bound. The detection accuracy and performance are analyzed using live traffic traces. The results indicate that this approach can be effective against current worm attacks.

Categories and Subject Descriptors

C.2.0 [**Computer-Communication Networks**]: General–Security and protection (e.g., firewalls)

General Terms

Keywords

network security; worm; worm detection; entropy; Chebyshev's inequality;

1. INTRODUCTION

Internet worms [1] have been a part of the world since the early days of the publicly available Internet. They have posed serious threats to the normal operation of the information infrastructure by destroying stored information, disrupting information transmission and improperly collecting and handling information. These malicious activities have caught the attention of the society through a series of high-profile incidents such as the CodeRed worm and Nimda worm in 2001, the Slammer worm.

In this paper, we present a router-based system to identify worm attacks by computing entropy values of selected packet attributes. The detection accuracy and performance are analyzed using live traffic traces. The results indicate that these methods can be effective against current worm attacks.

2. WORM DETECTION BASED ON

Conference name: Infoscale 2007, June 6-8, 2007, Suzhou, China
Copyright number (c): 978-1-59593-757-5

ENTROPY

2.1 Entropy

The concept of entropy in information theory describes how much information there is in a signal or event. An intuitive understanding of information entropy relates to the amount of uncertainty about an event associated with a probability distribution. In fact, Shannon's entropy is defined as a measure of the average information content associated with a random outcome. Shannon's definition[2] of information entropy makes this intuitive distinction mathematically precise. So we introduce entropy to worm detection. Shannon defines entropy in terms of a discrete random variable X, with possible states (or outcomes)x_1, x_2,.....x_n as:

$$H(X) = -\sum_{i=1}^{n} p(x_i)\log_2 p(x_i) \quad (1)$$

Where $p(x_i)$ is the probability of the ith outcome of X.

That is, the entropy of the event x is the sum, over all possible outcomes n of x, of the product of the probability of outcome i times the log of the inverse of the probability of outcome i (which is also called i's surprisal - the entropy of X is the expected value of its outcome's surprisal).

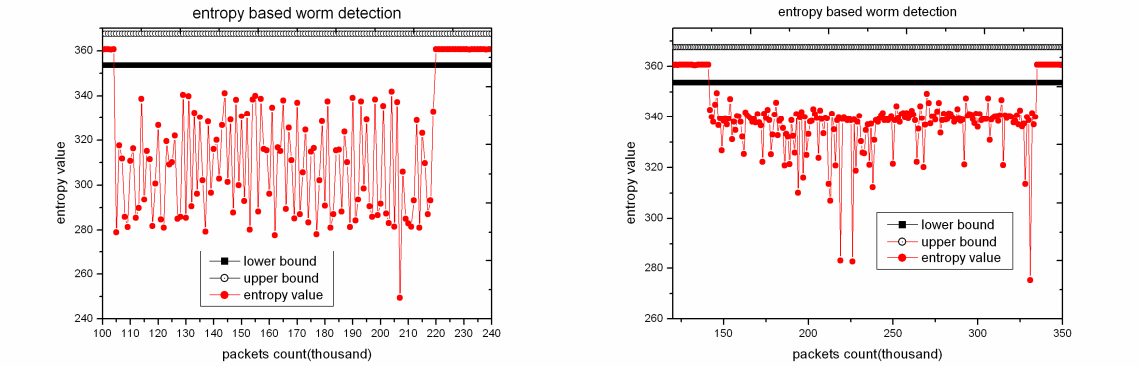
2.2 Chebyshev's Inequality

For networks with various infection-like (but normal) behavior, the simple detection heuristic discussed above may not work. We use Chebyshev's inequality to determine whether the simple detection heuristic can be used. For a given random variable, Chebyshev's inequality can provide a bound on the probability that the value lies outside a certain distance from the variable's mean. During the training phase, we approximate the mean m and the variance s^2 of the real distribution by computing the sample mean $\overline{m}$ and the sample variance s^2 for the entropy values e_1 , e_2 …… e_n .The intuition of applying Chebyshev's inequality is that we can set a threshold d so that we get a bound on the probability that a value (denoted as x) deviates from the mean μ more than the threshold d. However, in most of cases, the upper bound or lower bound is enough to detect the outbreak of worms. In particular, a Chebyshev's inequality can be represented as follows:

$$p(|x - m| > d) < \frac{s^2}{d^2} \quad (2)$$

Where p is probability.

Work in progress



(a) CodeRed II worm detection (b) Slammer worm detection

Figure 1. worm detection

2.3 Learning

The entropy values learned with Chebyshev's inequality is very easy to implement as an incremental version with only some information stored. An incremental version of this method is particularly useful from the point of performance Moreover, the algorithm will improve in accuracy as time moves forward and more data is learned.

3. EXPERIMENT EVALUATION

To evaluate the efficacy of our detection algorithm, we conducted the simulation experiments using two traces collected from Neusoft (trace-1) and Northeastern University(trace-2), respectively.

We split each trace file into two chronological parts for training and detecting worms respectively, in particular, the first 80% of the trace was used as training data and the rest 20% was used for detecting worms. This simulates the process of learning from the historical data and applying the learned model to current and future data. The learning results are shown in Table Ⅰ. We set the detection threshold (d) to be 7 and 10 with a false positive rate below 1.045522% and 2.080527%,respectively. The proposed algorithm was then applied on the tested traces without adding attacks. Testing results showed that no false alarms were reported in this experiment.

TABLE I. LEARNING RESULTS

name	m	s^2
tarce-1	360.120564	0.512306
tarce-2	**42.486988**	**2.080527**

In the worm detection experiments, a mixture of both normal and worm traffic is used in the proposed algorithm. Because we did not have traces that contain actual worm traffic, we inserted the same sets of worm traffic into the cleaned test set. Two worm datasets, which include CodeRed II and Slammer worm respectively, were utilized. Both of the worm samples were collected from real traffic as they appeared in the wild. Because our algorithm only considers packets of the worm, the worm sets were inserted at random places in the test data.

For this round of experiments, each dataset was also split into two distinct chronologically-ordered portions, one for training and the other for testing, following the 80%-20% rule. The worm traffic is combined with the 20% of trace files to make the final test traffic for our evaluation. Figure 1 shows the output (entropy values) of an entropy detector examining the destination IP address packet attribute under CodeRed II and Slammer worm attack. Before the attack begins, the entropy values of destination IP address entropy still fall within the bound calculated from chebyshev's inequality. During the attack, the entropy values fall below the lower bound abruptly all the time. And the lower bound stays stable all the time. Furthermore, there were no false-positives in such experiment

4. CONCLUSION

In this paper, we present a router-based system to identify worm attacks by computing entropy distributions of selected packet attributes. The detection accuracy and performance are analyzed using live traffic traces. The results indicate that entropy based detection of massive network events is a powerful, efficient and fast approach to detect massive network events like worm outbreaks and to provide an analysis of the characteristics of worm traffic. As such it is very suitable for use in an early warning system. Our future work will include combining other network trace statistics to detect worm attacks and stealthy worms.

5. ACKNOWLEDGMENTS

This work is supported by the national natural science foundation of China under Grant Nos. 60602061 and the national high-tech research and development plan of China under Grant Nos. 2006AA01Z413.

6. REFERENCES

[1] Stuart Staniford, Vern Paxson, and Nicholas Weaver.How to 0wn the internet in your spare time. In Proceedings of the 11th USENIX Security Symposium,August 2002.

[2] C.E. Shannon, and W. Weaver, The Mathematical Theory of Communication, University of Illinois Press, 1963.