EVALUATIONS OF THE EFFECTIVENESS OF ANOMALY BASED INTRUSION DETECTION SYSTEMS BASED ON AN ADAPTIVE KNN ALGORITHM

Assosiate professor, PhD Evgeniya Nikolova, BFU
Assosiate professor, PhD Veselina Jecheva, BFU

Abstract: *The aim of the present paper is to present some evaluations of the effectiveness of IDS based on the k-Nearest Neighbor algorithm with Jaro and Jaro-Wincler distances, applied as metrics. The evaluation of the represented simulation results indicates the proposed methodology produces reliable and steady results.*

Keywords: *Intrusion Detection, Anomaly Based IDS, k-Nearest Neighbor algorithm.*

1. Introduction.

The main task of any detection system is recognizing an intrusion attempt. The decision whether an intrusion is present or not could be made in various ways. The typical procedure for creation an anomaly-based IDS contains the following steps:

- first, collect sets of normal data, clear of intrusions or attacks, and a set of test data, containing intrusions or attacks;
- second, train the intrusion detection system on the normal data, and then run it against the test data – the binary classification;
- third, measure the quality of the detection algorithm in terms of hit, miss, and false alarm rates.

The evaluation of intrusion detection systems (IDS) has begun to be an active topic over the last years [4], [8], [11], [9]. In [10] an IDS, based on the KNN algorithm, is presented. The proposed system applies Jaro (JD) and Jaro-Winkler distances (JWD) as measures of the closeness of the current activity to the normal one. In the present paper, our attention is drawn on the third step - the evaluation of the effectiveness of the algorithm, presented in [10]. Whatever the model is a detector's performance can be described by its receiver operating characteristic (ROC) curve. The ROC parameters are the probability of an alarm given an intrusion and the probability of an alarm given no intrusion. ROC curve is a plot of the detection probability versus false alarm rate. Another estimation of effectiveness is false discovery rate (*FDR),* which controls the expected proportion of incorrectly rejected null hypothesis: the traffic is normal. As a measure of the quality of binary classifications can be applied the Matthews Correlation Coefficient (*MCC*). Since the main advantage of anomaly based IDS is the potential to detect novel or unknown attacks, it is useful to estimate the ability of detecting old and new attacks.

The article is organized as follows: section 2 explains some measurement criteria for evaluations of the effectiveness of anomaly-based IDS, section 3 represents the results on our prototype, using the metric introduced in section 2. Finally, section 4 contains the main conclusions.

2. Measurement Criteria.

Common methods used for classification quality evaluation in the machine learning and information retrieval community are: accuracy, false discovery rate (*FDR*), Matthews Correlation Coefficient (*MCC*) and Receiver Operating Characteristic curve (*ROC*).

Measurement-based anomaly detection techniques have to contend with two types of errors. The true positive (sensitivity) is the probability that a statistical test will be positive for a true statistic. A false positive error occurs if a difference is declared when the traffic is normal. The true negative (specificity) is the probability that a statistical test will be negative for a negative statistic. On the other hand, a false negative error occurs if no difference is declared when the traffic is not normal.

Accuracy in reporting is a critical issue for intrusion detection systems. Accuracy is the degree of correctness of such system [13].

$$Accuracy = \frac{number\quad of\quad TP\quad +\quad number\quad of\quad TN}{numbers\quad of\quad TP + TN + FP + FN}$$

An accuracy of 100% means that the test identifies all anomalous and normal activity correctly. If an IDS raises an alarm for the legitimate activity of a user, then a false alarm is present. An intrusion detection system becomes more accurate as it detects more attacks and raises fewer false alarms.

False Discovery Rate (*FDR*) is the proportion of false positives among the declared differentially attacks ([2], [3]).

$$FDR = \frac{FP}{FP + TP}$$

Lippmann et.al. [11] proposed an approach, which takes into account the mis-classified sessions when calculating the false alarms rate.

The MCC is a correlation coefficient between the observed and predicted binary classifications, which returns a value between -1 and +1 ([1], [12]).

$$MCC = \frac{TP.TN - FP.FN}{\sqrt{TP + FP\quad TP + FN\quad TN + FP\quad TN + FN}}$$

A coefficient of +1 represents a perfect prediction, 0: an average random prediction and -1: the worst possible prediction.

The ROC curve is a method of graphically demonstrating the relationship between 1-TNR (true negative rate) and TPR (true positive rate) as x and y respectively ([5]). The best possible prediction method would yield a point in upper left corner (0,1) of the ROC space, representing 100% TPR (all TP are found) and 100% TNR (no FP are found). This point is called a perfect classification. The diagonal line (from the left bottom to the right corner) divides the ROC space in areas of good and bad classification. Points above this line indicate good classification results, while points below the line indicate wrong results.

3. Results of the Effectiveness

The conducted experiments indicate that the proposed methodology produces results with high level of accuracy, since all obtained values are between 81,45% and 98,44% for all examined processes (see [10]). Since our data set contains sequences of system calls, our approach takes into account the rate of mis-classified system calls during the work of the privileged processes for the *FDR* calculation.

Table 1. The False Discovery Rate values depending on *K* when *L=7*

Processes	Distance	*K*=10	*K*=20	*K*=30
synthetic sendmail	*JD*	3,55%	8,71%	5,94%
	JWD	3,08%	7,15%	6,27%
login	*JD*	1,52%	0,00%	1,07%
	JWD	0,48%	0,07%	0,15%
named	*JD*	6,14%	2,21%	2,06%
	JWD	6,15%	2,21%	2,53%
xlock	*JD*	2,62%	8,10%	5,94%
	JWD	2,45%	7,61%	8,11%

Table 2. *FDR values, when L=10 and 13*

		L=10			*L*=13		
Process	Distance	*K*=10	*K*=20	*K*=30	*K*=10	*K*=20	*K*=30
login	*JD*	1,30%	0,00%	0,25%	2,95%	0,59%	0,14%
	JWD	1,34%	0,33%	0,11%	2,99%	0,44%	0,29%

The results of FDR are shown in Table 1 depending on the different values of K and the applied distance. The best FDR results were obtained for the process login and to a certain extent for the process named. The relatively high FDR values for the processes synthetic sendmail and xlock when K=20. Table 2 contains the values of FDR for the process login, where L=10 and 13. As we can see, the proposed method distinguishes the intrusion behavior from the normal one with very good false rate, since all obtained values belong to the interval (0%, 3%).

The comparison between the *ROC* curves for the process login at *K*=10, 20 and 30 and *L*=7 is represented in Figure 1. It could be seen that detection rate in all cases is high, at *K*=10 we obtain lower false positive rate, compared to the cases, where *K*=20 and 30.

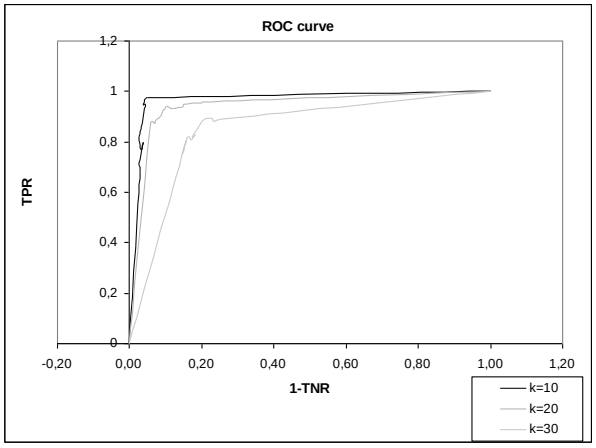

Figure 1. *ROC* curves for the process login at *K*=10, 20 and 30 and *L*=7

The *MCC* was calculated for each sequence in the training set, and the values were presented in Table 3. Since the high MCC means better classification, the obtained values of *MCC*, which are greater than 51,80%, indicate that the proposed methodology achieves significant correlation between normal data profiles and the observed data. The best results were obtained for the process xlock where *K*=30. The obtained *MCC* values support the assumption, proposed by Forrest et.al [6, 7], that the system call sequences are reliable discriminator between normal and abnormal system activity.

Table 3. The *MCC* values depending on *K* when *L=7*

Processes	Distance	*K*=10	*K*=20	*K*=30
synthetic sendmail	*JD*	63,86%	82,09%	78,43%
	JWD	64,44%	77,58%	77,87%
login	*JD*	51,80%	61,77%	65,00%
	JWD	52,22%	64,98%	68,49%
named	*JD*	54,74%	81,11%	63,66%
	JWD	54,99%	65,03%	59,53%
xlock	*JD*	69,51%	85,69%	94,05%
	JWD	72,44%	87,14%	92,06%

Table 4 presents the *MCC* values for the process login, which measure the predictive performance of the proposed methodology where *L*=10 and 13. The obtained values are between 42,88% and 74%, which means balanced results, as a coefficient of +1 represents a perfect prediction.

Table 4. *MCC values, when L=10 and 13*

		L=10			*L*=13		
Process	Distance	*K*=10	*K*=20	*K*=30	*K*=10	*K*=20	*K*=30
login	*JD*	46,48%	47,97%	74,37%	46,13%	51,54%	60,73%
	JWD	42,88%	47,67%	68,09%	45,23%	52,50%	50,50%

Tables 5, 6 and 7 show the performance of the algorithm while detecting old and new attacks. An old attack is defined as an attack identical to an attack that was present in the set S_2, while a new attack is an attack that was not present in that set. Any sequence, which is not found in this set, is called a failing. Any individual failing could indicate anomalous behavior, or it could be a sequence, which was not included in the training data. The rare sequences that have very low frequency are also taken into account when we test the traces of the current activity. As indicated in the tables, the results demonstrate the ability of the created IDS to detect effectively and accurately the old attack sequences. It correctly classified more than 91,67% of the sequences from both sets S_1 and S_2. An advantage of the proposed method is the ability to detect previously known attacks, as well as novel intrusions.

Table 5. The algorithm accuracy when K=10 and L=7

Processes	Distance	Old attacks	Failing sequences	Rare sequences
synthetic sendmail	*JD*	98,85%	80%	90%
	JWD	98,85%	80%	90%
login	*JD*	97,86%	87%	96%
	JWD	97,87%	88%	95%
named	*JD*	98,68%	71%	90%
	JWD	98,65%	71%	90%
xlock	*JD*	96,90%	72%	91%
	JWD	96,90%	72%	91%

Table 6. The algorithm accuracy when K=20 and L=7

Processes	Distance	Old attacks	Failing sequences	Rare sequences
synthetic sendmail	*JD*	98,78%	80%	91%
	JWD	98,78%	80%	91%
login	*JD*	97,17%	83%	93%
	JWD	97,87%	82%	94%
named	*JD*	96,87%	80%	90%
	JWD	96,57%	80%	90%
xlock	*JD*	96,57%	83%	91%
	JWD	96,72%	82%	92%

Table 7. The algorithm accuracy when K=30 and L=7

Processes	Distance	Old attacks	Failing sequences	Rare sequences
synthetic sendmail	*JD*	98,86%	81%	91%
	JWD	98,86%	81%	91%
login	*JD*	97,14%	84%	93%
	JWD	97,87%	84%	92%
named	*JD*	91,67%	81%	91%
	JWD	91,67%	81%	91%
xlock	*JD*	95,06%	83%	92%
	JWD	95,08%	84%	92%

Conclusion

Supervised network intrusion detection has been an area of active research for many years. The present paper proposes an anomaly-based approach, which applies some data mining techniques for the normal data description and the kNN algorithm during the detection phase. The experimental results with a host-based dataset demonstrate that the proposed method is robust and effective while detecting the violations of the computer security.

References:

1. Baldi P., Brunak S., Chauvin Y., Andersen CAF, Nielsen H., Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics* 16, 2000, pp. 412–424.
2. Benjamini,Y. and Hochberg,Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B, 57, 289–300.
3. Benjamini,Y. and Hochberg,Y. (2000) On the adaptive control of the false discovery rate in multiple testing with independent statistics. J. Edu. Behav. Stat., 25, 60–83.
4. Durst R., T. Champion, B.Witten, E. Miller, and L. Spagnuolo, 'Testing and evaluating computer intrusion detection systems', *Communications of the ACM*, **42**(7), 53–61, (1999).
5. Ferri C., N. Lachinche, S. A. Macskassy, A. Rakotomamonjy, eds. (2005). Second Workshop on ROC Analysis in ML.
6. Forrest S., S.A. Hofmeyr, A. Somayaji, T.A. Longtaff, A sense of self for Unix processes, In Proceedings of the 1996 IEEE Symposium on Security and Privacy, IEEE Computer Society Press, Los Alamitors, CA, pp.120-128.
7. Forrest S., S.A. Hofmeyr, A. Somayaji, Intrusion detection using sequences of system calls, *Journal of Computer Security*, Vol. 6, 1998, pp. 151-180.
8. Gaffney J.E., J.W. Ulvila, *Evaluation of Intrusion Detectors: A Decision Theory Approach*, The IEEE Symposium on Security and Privacy, 2001.
9. McHugh J., 'Testing intrusion detection systems: A critique of the 1998 and 1999 darpa intrusion detection system evaluations as performed by liconln laboratory', *ACM Transactions on Information and System Security*, **3**(4), 262–2944, (2000).
10. Jecheva V., E. Nikolova, An adaptive KNN algorithm for anomaly intrusion detection, МК Взаимодействието теория-практика: ключови проблеми и решения, том 3, 2011, 227-231.
11. Lippmann R., D. Fried, I. Graf, J. Haines, K. Kendall, D. McClung, D. Webber, S. Webster, D. Wyschograd, R. Cunninghan, M. Zissan, "Evaluating Intrusion Detection Systems: the 1998 DARPA off-line Intrusion Detection Evaluation", Proceedings of the DARPA Information Survivability Conference and Exposition, IEEE Computer Society Press, Los Alamitos, CA, 12-26, 2000.
12. Matthews, B.W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim*. Biophys. Acta, Vol. 405, pp. 442-451.
13. Taylor J. R., An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements. University Science Books, 1999, pp.128-129.