EFFECTIVENESS EVALUATION OF CLASSIFICATION TREE AND KULLBACK-LEIBLER DISTANCE-BASED ANOMALY INTRUSION DETECTION APPROACH

Veselina Jecheva, Burgas Free University, vessi@bfu.bg
Evgeniya Nikolova, Burgas Free University, enikolova@bfu.bg

Abstract. *The purpose of the paper is to present some evaluations of the effectiveness of IDS based on the classification tree and Kullback-Leibler distance-based approach to anomaly-based intrusion detection. The most used methods for this evaluation are: accuracy, Matthews correlation coefficient and ROC curve.*

Keywords: *Intrusion Detection, Anomaly Based IDS, Relative entropy, Receiver Operating Characteristic Curve, Accuracy, Matthews correlation coefficient, Error rate*

ОЦЕНКА НА ЕФЕКТИВНОСТТА НА СИСТЕМИ ЗА ОТКРИВАНЕ НА НАРУШЕНИЯ, ОСНОВАНИ НА КЛАСИФИКАЦИОННИ ДЪРВЕТА И РАЗСТОЯНИЕ НА КУЛБЕК-ЛЕЙБЛЕР

Веселина Жечева, Бургаски свободен университет, vessi@bfu.bg
Евгения Николова, Бургаски свободен университет, enikolova@bfu.bg

Резюме. *Настоящата статия представя някои оценки на ефективността на системи за откриване на нарушения, основани на класификационни дървета и разстояние на Kullback-Leibler. Най-използвани методи за такава оценка са: точност, корелационен коефициент на Матюс и ROC крива.*

Ключови думи: *Откриване на нарушения, системи за откриване на нарушения, основани на аномалии, ROC крива, корелационен коефициент на Матюс, процент на грешки, точност.*

I. Увод

Откриване на нарушения в информационната сигурност наричаме всички действия, свързани с наблюдение на събитията в дадена компютърна система и откриване на необичайни, неоторизирани и злонамерени действия от външни лица или легитимни потребители. Системата за откриване на нарушения (СОН) автоматизира този процес и следи данните в системата (мрежови или на определен хост), за да разграничи нарушенията и атаките от нормалните потребителски действия [1]. В зависимост от метода, използван за това разграничение, СОН се делят на две основни категории: базирани на злоупотреби и базирани на аномалии. Първият тип се основава на сигнатурен анализ на известни средства за атака в данните на системата. Тези системи притежават висока точност, но основният им недостатък е невъзможността да открият нови или варианти на съществуващи експлойти [12].

В настоящата статия се разглеждат СОН, базирани на аномалии. Тези системи извършват поведенчески анализ, като следят текущите данни в системата за необичайни действия, които могат да бъдат индикация за атака. Те се основават на идеята, че нарушенията предизвикват отклонение от нормалната работа на системата и

за целта използват предварително дефинирани профили на нормалната работа на потребителите в системата [11]. Ако текущите събития се отклоняват в значителна степен от тези профили, системата издава съобщение за атака.
Анализът на поведението на определени процеси в сървърите е често използван подход за поведенчески анализ в СОН [3]. Профилите на нормални действия се създават чрез събиране на данни за системни извиквания, които се изпълняват при нормална работа на системата. След това СОН следи за отклонение от тези профили в текущите данни за работата на системата.
Настоящата статия се основава на методология, описана в [10]. Предложеният подход включва изграждане на класификационни дървета, които служат за описание на нормалните действия в системата и са подредени на базата на преходните вероятности на съответните събития. След това системата следи текущите данни, състоящи се от номера на системни извиквания в системата и ги сравнява с вече записаните като използва разстоянието на Кулбек-Лейблер [6]. При наличие на значително отклонение от предварително съставените профили на нормални действия, системата издава сигнал за наличие на нарушение и маркира съответните действия като нарушение.

II. Симулационни експерименти

На базата на предложения модел са извършени редица симулационни експерименти. При експериментите са използвани данни, получени от проекта Computer Immune Systems Project, реализиран в Computer Science Department, University of New Mexico [13]. Разглежданите данни се отнасят до процеси (inetd, login, named, synthetic sendmail, ps), изпълнявани с администраторски права в определен период от време в Unix система, чието генериране е описано в [5]. Използваните данни представляват текстови файлове, които съдържат по 2 числа на всеки ред: идентификатор на изпълнявания процес (PID) и номера на системното извикване:

PID	1393	1393	...	1393
Системни извиквания	112	19	...	105

Всеки използван набор данни съдържа данни за нормална работа на системата, както и данни, съдържащи нарушения в сигурността. На базата на данните за нормална работа на системата са генерирани класификационните дървета, които представляват база с профили на нормална работа на системата. Данните, съдържащи нарушения в сигурността, се разглеждат като данни за текуща работа на системата и чрез описаната методология в тях се търсят отклонения от нормалната работа на системата. Симулационните експерименти са извършени като всеки набор от данните за нормална работа, се обхожда чрез плъзгащ прозорец с дължина L=7, като при това се построяват класификационните дървета с височина L. След това се сканират данните за текуща работа на системата и се изчислява разстоянието на Кулбек-Лейблер между двата набора данни, което се използва като индикатор за степента на близост между тях.

III. Ефективност на предложената методология

1. Резултати от работата на системата.

Целта на СОН е сканиране на данните за работата на наблюдавана система и

реализирането на двоична класификация, т.е. да се определи дали дадена последователност от наблюдения принадлежи към една от двете групи – множество от нормални дейности на системата или множество от отклонения от нормалната дейност (нарушения). За всяко възможно наблюдение тестът, чрез който се реализира тази класификация, може да допуска два типа грешки – грешка от първи род (false positive - FP) и грешка от втори род (false negative - FN). FP се допуска, когато едно събитие се отчита като нарушение, но всъщност това е нормална дейност, докато FN е грешката, която се допуска, когато настъпва нарушение, но то не е класифицирано като такова. При двоичната класификация са възможни следните четири резултата, представени в Таблица 1:

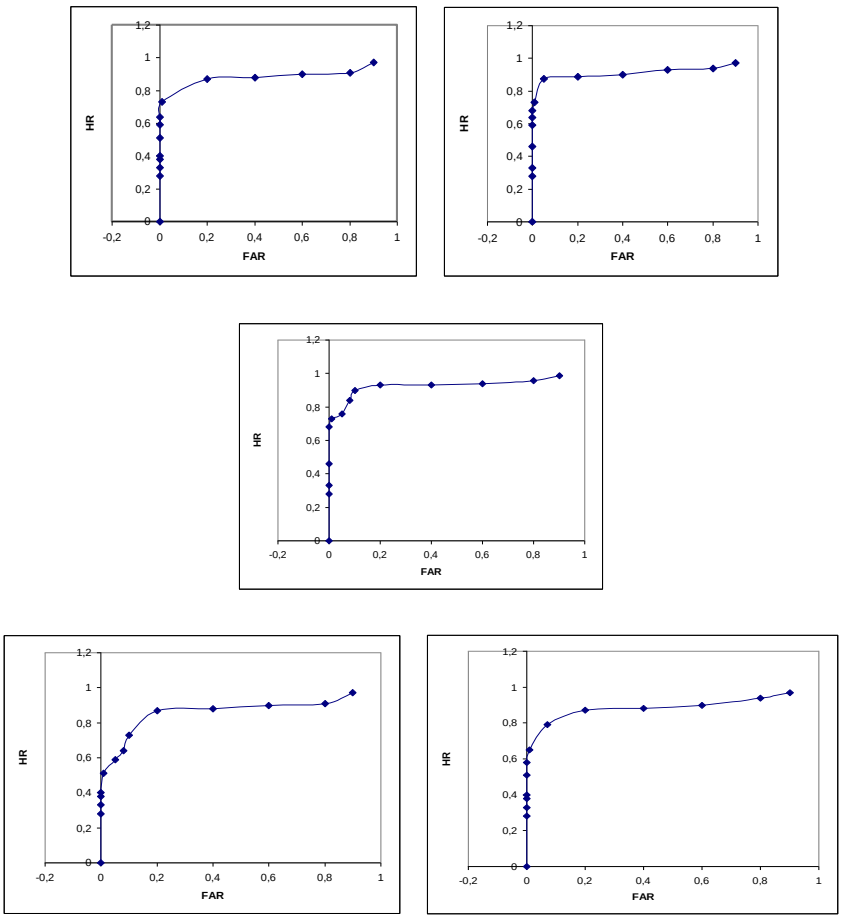
Резултат от системата / Наличие на нарушение	Да	Не
Да	*TP*	*FN*
Не	*FP*	*TN*

Табл. 1

В таблицата TP (true positive) и TN (true negative) са коректно класифицираните нормални действия и нарушения съответно. Оценката на ефективността и настройката на СОН се нуждае от баланс между тези четири стойности. За да се направи оценка на ефективността на предложената методология, използваща разстоянието на Кулбек-Лейблер, са приложени някои статистически методи. Като показатели за измерване на точността на класификацията са използвани ROC крива, корелационен коефициент на Матюс, процент на грешки и точност.

2. Оценки на резултатите от работата на системата.

Понятието попадение (hit) се използва, ако тестът правилно регистрира дадено събитие като нарушение. Може да се изчисли степен на попадение (hit rate - HR) като отношение на броя на попаденията в сесиите с нарушения и общия брой на сесии с нарушения в тестваните данни, а относителна стойност на фалшиви сигнали (false-alarm rate - FAR) като отношение на броя на фалшивите сигнали към общия брой на истински нормални данни. ROC кривата (Receiver Operating Characteristic Curve) [4] представя нагледно способността СОН да разграничи принадлежността на наблюденията към едно от двете множества – нормални дейности или нарушения. Тя представя графично отношението на степента на попадение и относителната стойност на фалшиви сигнали при различни прагови стойности. Колкото графиката е по-близо до горния ляв ъгъл, с попадение 100% и 0% относителна стойност на фалшиви сигнали, толкова по-добра е разпознавателната способност на СОН. Следователно, ROC кривата показва цялостно ефективността на класификационните способности на даден тест.



Фиг. 1. *ROC крива* за процесите named, inetd, ps, login и synthetic sendmail

За СОН с перфектна класификационна способност графиката на ROC крива преминава през горния ляв ъгъл. На Фигура 1 са представени графиките на ROC криви за процесите named, inetd, ps, login и synthetic sendmail, когато предложената методология прилага разстоянието на Кулбек-Лайблер. Тази СОН постига 89% HR при 20% FAR за процеса named, 90 HR% при 36% FAR за процеса inetd, 93% HR при 20% FAR за процеса ps, 87% HR при 20% FAR за процеса login и 87% HR при 20% FAR за процеса synthetic sendmail. Тъй като площта под ROC кривите за процесите named и ps е от 0,9 до 1, този метод дава отлични резултати. Площта под ROC кривите за процесите inetd, login и synthetic sendmail е между 0,8 и 0,9, което означава, че този метод дава добри резултати при класификация.

3. Точност на системата.

Точност (Accuracy) е степента на съответствие на брой открити аномалии от метода при дадена непозната последователност с реалния брой аномалии в данните [2].

$$Точност = \frac{TP + TN}{Брой \quad на \quad входящите \quad редици}$$

Колкото по-висока е стойността на точността, толкова тестът идентифицира нарушения и нормална активност с по-висока прецизност. Представеният метод за откриване на нарушения може да установи в определена непозната последователност нормално или аномално поведение с точност, чиито стойности за процесите inetd, login, named, synthetic sendmail и ps са представени в Таблица 2. Таблицата ясно показва, че нашият

подход показва голяма ефективност по отношение на точността, след като всички стойности са над 0,8447.

Процес	*Точност (Accuracy)*
inetd	0,9631
login	0,9546
named	0,8681
synthetic sendmail	0,8447
ps	0,9780

Табл. 2. *Точност* за процесите inetd, login, named, synthetic sendmail и ps

Като мярка за качеството на двоична класификация може да се използва *корелационният коефициент на Матюс* (*Matthews correlation coefficient - MCC*) [9]:

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

MCC=1 представлява 100% точна прогноза, а *MCC*=-1 представлява възможно най-лошата прогноза.

Процес	*MCC*
inetd	0,8542
login	0,7810
named	0,6265
synthetic sendmail	0,8732
ps	0,6720

Табл. 3. *MCC* за процесите inetd, login, named, synthetic sendmail и ps

Получените стойности на *MCC,* представени в Таблица 3, за всички процеси, принадлежат на интервал (0,6265; 0,8732), което означава балансирани резултати, тъй като най-добрата стойност на коефициентът е 1. Тези резултати показват, че е на лице значима корелация между профилите, описващи поведението на системата и изследваните последователности на системна активност, когато се прилага методологията, основана на разстоянието на Кулбек-Лайблер.

Процент на грешките (*Error rate*) се изчислява със следната формула:

$$Процент\ \ на\ \ грешките = \frac{Общ\ \ броя\ \ на\ \ наблюденията}{Общ\ \ брой\ \ нарушения}$$

където *Общ брой нарушения* е броят на откритите нарушения от предложения метод. Стойностите на тази величина за изследваните процеси са представени в таблица 4. Те дават информация за честотата на срещане на нарушения в експерименталните данни от предложения метод.

Процес	*Процент на грешките*

	(Error rate)
inetd	114,25
login	93,98
named	49,77
synthetic sendmail	18,21
ps	67,30

Таблица 4. Процент на грешки за процесите inetd, login, named, synthetic sendmail и ps

IV. **Заключение**

Системи за откриване на нарушения се въвеждат за първи път в началото на 90-те. До този момент научните изследвания в тази област се фокусират върху възможността за откриване на нови атаки и свеждане до минимум на фалшивите аларми. Оценката на ефективността на СОН не е активна тема до 1998, когато MIT Lincoln Laboratories дава една такава оценка [7, 8].

За да се направи такава оценка, най-често се използва *ROC крива.* Основният проблем на този метод е, че при него не се отчита „цената” на погрешно класифицирани наблюдения (*misclassification*). Обикновено „цената” на *FN* е много по-висока от „цената” на *FP*, което предполага намирането на други подходи за оценка на ефективността на СОН. От друга страна прекалено големият брой FP понижава доверието в системата, което значи, че този вид грешки също трябва да бъдат сведени до минимум. Други оценки на ефективността са *Точност* (*Accuracy*), *Процент на грешките* (*Error rate*), а като мярка за качеството на двоична класификация най-често се използва *MCC*. Използвайки ги, получените резултати при оценяване на предложената методология, основана на разстоянието на Кулбек-Лайблер, ни водят до извода, че при работата на тази СОН се получават надеждни и стабилни резултати.

Тъй като основното предимство на СОН, базирани на аномалии, е потенциалът им за откриване на нови или непознати атаки, то би било полезно да се оцени възможността им за откриване както на известни, така и на нови атаки.

References

[1] Abraham, A., J. Thomas, *Distributed Intrusion Detection Systems: A Computational Intelligence Approach*, Applications of Information Systems to Homeland Security and Defense, Idea Group Inc. Publishers, USA, Chapter 5, pp. 105-135, 2005.

[2] Baldi P., Brunak S., Chauvin Y., Andersen C.A., Nielsen H., *Assessing the accuracy of prediction algorithms for classification: An overview,* Bioinformatics, Vol. 16, pp. 412–424, 2000.

[3] Ghosh A.K., A. Schwartzbard, M. Schatz, *Learning Program Behavior Profiles for Intrusion Detection*, Proceedings of the 1st Workshop on Intrusion Detection and Network Monitoring, pp. 51–62, 1999.

[4] Ferri C., N. Lachinche, S. A. Macskassy, A. Rakotomamonjy, eds., *Second Workshop on ROC Analysis in ML*, 2005.

[5] Forrest S., S.A. Hofmeyr, A. Somayaji, T.A. Longtaff, *A Sense of Self for Unix Processes*, Proceedings of the 1996 IEEE Symposium on Security and Privacy, IEEE Computer Society Press, Los Alamitors, CA, pp.120-128, 1996.

[6] Han, Te Sun & Kobayashi, Kingo, *Mathematics of Information and Coding*, American Mathematical Society. pp. 19–20, 2002.

[7] Lippmann R., D. Fried, I. Graf, J. Haines, K. Kendall, D. McClung, D. Weber, S. Webster, D. Wyschogrod, R. Cunningham, M. Zissman, *Evaluating intrusion detection systems: The*

1998 DARPA off-line intrusion detection evaluation, In Proceedings of the DARPA Information Survivability Conference and Exposition, Los Alamitos, California, USA, 2000. IEEE Computer Society Press.
[8] R. Lippmann, J. W. Haines, D. J. Fried, J. Korba, K. Das, *The 1999 darpa off-line intrusion detection evaluation*, Computer Networks, 34(4):579–595, 2000
[9] Matthews B.W., *Comparison of the predicted and observed secondary structure of T4 phage lysozyme*, Biochim. Biophys. Acta 1975, 405, 442-451.
[10] Nikolova E., V. Jecheva, *Classification Tree and Kullback-Leibler Distance-based Anomaly Intrusion Detection Approach*, Knowledge - traditions, innovations, perspectives, Burgas Free University, 2013
[11] Qiao Y., X.W. Xin, Y. Bin, S. Ge, *Anomaly intrusion detection method based on HMM*, IEEE Electronic Letters Online No: 20020467, 2002.
[12] Tran T.P., T. Jan, A. J. Simmonds, *A Multi-Expert Classification Framework for Network Misuse Detection*, Proceeding of Artificial Intelligence and Soft Computing, ISBN 0-88986-610-4, 2006.
[13] University of New Mexico’s Computer Immune Systems Project, http://www.cs.unm.edu/~immsec/systemcalls.htm.