УДК 004.056.53 : 004.492.3

Комар М.П.
Тернопільський національний економічний університет

Побудова ієрархічного класифікатора комп'ютерних атак на базі багатоканальних нейромережевих детекторів

Запропоновано підхід до побудови сукупного класифікатора для ієрархічної класифікації атак на інформаційні телекомунікаційні мережі на основі багатоканальних нейромережевих детекторів з використанням методу головних компонент. Це дало можливість зменшити розмірність аналізованої інформації при незначній втраті інформативності за рахунок стиснення вхідної інформації для отримання найбільш інформативних ознак, а також класифікувати типи і класи атак за рахунок об'єднання навчених на певний тип атаки нейромережевих детекторів. Запропонований підхід дозволив усунути конфлікти в роботі нейромережевих детекторів.

Ключові слова: інформаційні телекомунікаційні мережі, комп'ютерні атаки, сукупний класифікатор, ієрархічна класифікація атак, нейронна мережа, багатоканальний нейромережевий детектор, метод головних компонент.

Komar Myroslav
Ternopil National Economic University

CONSTRUCTION OF THE HIERARCHICAL CLASSIFIER FOR COMPUTER ATTACKS ON THE BASIS OF MULTICHANNEL NEURAL NETWORK DETECTOR

The approach of collective classifier construction for hierarchical attacks' classification for the information telecommunication network based on multichannel neural network detector by using the method of principal components is presented. This made it possible to reduce the dimension of the analyzed data with negligible loss of information because of the input data compression for obtaining the most informative features and also to classify types and classes of attacks by combining neural network detectors trained for a certain type of attack. Proposed approach allowed to eliminate conflicts in the work of neural network detectors.

Keywords: information telecommunication networks, cyber attacks, collective classifier, hierarchical classification of attacks, neural network, multichannel neural network detector, method of principal components.

Вступ

Останнім часом відбулося злиття комп'ютерних мереж, інформаційних і телекомунікаційних технологій, що дозволило утворити сучасні інформаційні телекомунікаційні мережі (ІТМ), які є складною розподіленою системою, що характеризується наявністю множини взаємодіючих ресурсів, системних та прикладних інформаційних і телекомунікаційних процесів. У таких умовах важливою науково-технічною задачею є забезпечення цілісності, достовірності та конфіденційності інформації. Інформаційні телекомунікаційні мережі піддаються різного роду загрозам, і користувач не може бути впевнений у захищеності важливої інформації, оскільки кіберзлочинці продовжують удосконалювати і розробляти методи і засоби організації мережевих атак.

Одним з гучних злочинів у сфері інформаційної безпеки, на думку експертів видання Forbes [1] стала атака Anonymous на платіжні системи MasterCard, Visa і Paypal, коли ті в кінці 2010 року відмовилися приймати платежі для сайту WikiLeaks. Втрати від атаки склали $5,5 млн. Ще одним гучним злочином була атака на Citibank в червні 2011 року. Тоді хакери викрали $2,7 млн. з рахунків 3400 клієнтів банку.

Таким чином, пошук та реалізація підходів забезпечення цілісності, достовірності та конфіденційності інформації є важливою науково-технічної задачею.

Постановка задачі

Залежно від техніки, що використовується при здійсненні несанкціонованих дій на комп'ютерну систему, виділяють чотири основні класи мережевих атак (denial of service, user-to-root, remote-to-local, probe), кожен з яких складається з декількох типів [2]. DOS (denial of service, відмова в обслуговуванні) атаки. Це мережеві атаки, направлені на виникнення ситуації, коли в системі, що атакується, відбувається відмова в обслуговуванні. Дані атаки характеризуються генерацією великого об'єму трафіку, що приводить до перевантаження і блокування сервера. Виділяють шість типів DOS-атак: back, land, neptune, pod, smurf, teardrop. U2R (user-to-root) атаки передбачають отримання зареєстрованими користувачами привілеїв локального суперкористувача (адміністратора). Виділяють чотири типи U2R-

атак: buffer_overflow, loadmodule, perl, rootkit. R2L (remote-to-local) атаки характеризуються отриманням доступу незареєстрованого користувача до комп'ютера з боку віддаленої машини. Виділяють вісім типів R2L-атак: ftp_write, guess_passwd, imap, multihop, phf, spy, warezclient, warezmaster. Probe-атаки грунтуються на процесі сканування мережевих портів віддаленої машини з метою отримання конфіденційної інформації. Виділяють чотири типи Probe-атак: ipsweep, nmap, portsweep, satan [2].

Виявлення і класифікація мережевих атак на комп'ютерну систему відбувається на основі аналізу інформації, що передається по каналах передачі даних ІТМ. Виділяють 41 параметр мережного з'єднання, які, у свою чергу, об'єднані у три групи: вбудовані параметри, параметри контенту, параметри трафіку [2].

Для вирішення задачі виявлення і класифікація атак на ІТМ на основі аналізу мережевого трафіку розроблено досить багато підходів, зокрема з використанням штучних нейронних мереж. Аналіз відомих публікацій [3–15], а також проведені експериментальні дослідження [16–17] показують здатність запропонованих підходів виявляти мережеві атаки.

Проте, частина атак залишаються такими, що практично не детектуються. Багато в чому це виникає унаслідок того, що база з'єднань KDD-99 [2], яка використовується для тестів містить недостатню кількість записів про ці атаки і, відповідно, дані, на яких можна навчити нейронну мережу. Для того, щоб поліпшити якість навчання нейронних мереж на «рідкісних» типах атак, пропонується використовувати метод головних компонент [18, 19] для скорочення розміру інформації при навчанні та аналізі мережевого трафіку.

Основна частина

Метод головних компонент (principal component analysis, PCA) [18, 19] є одним з основних способів зменшення розмірності даних при мінімальній втраті інформації. Для визначення числа головних компонент запропоновано використати критерій відносної інформативності

$$J = \frac{\lambda_1 + \lambda_2 + ... + \lambda_p}{\lambda_1 + \lambda_2 + ... + \lambda_n}, \quad (1)$$

де λ_i – кількість інформації в *i*-й компоненті.

Аналізуючи за допомогою формули (1) розподіл кількості інформації, що міститься в кожній подальшій компоненті *n*, визначається число головних компонент *p*, які доцільно використовувати для подальшого аналізу без істотної втрати відносної інформативності *J*.

Сукупним нейромережевим детектором є такий детектор, який складається з множини нейромережевих детекторів, кожен з яких навчений на певному типі атак, що дозволяє детектувати клас і тип атаки. Оскільки виділяють 22 різновиди мережевих атак , які об'єднані в чотири класи (*DoS*, *U2R*, *R2L* і *Probe*) [2], то для побудови сукупного класифікатора використовуються 22 нейромережевих детектори, які навчені на виявлення кожного з 22 різновидів мережевих атак.

Узагальнену схему функціонування сукупного нейромережевого детектора для виявлення мережевих атак з використанням методу головних компонент можна представити таким чином (рис. 1).

При такому підході, кількість n вхідних нейронів нейронної мережі, що використовується в якості детектора, дорівнює 12. Вхідною інформацією є 12 перших головних компонент, які подаються на 22 нейромережеві детектори, де і відбувається її визначення до класу мережевої атаки або до класу нормального з'єднання.

На рис. 2 представлений приклад схеми побудови сукупного класифікатора для ієрархічної класифікації атак, який базується на стисненні інформації з використанням методу головних компонент і на такому об'єднанні нейромережевих детекторів, щоб нейтралізувати конфлікти між ними. Для нейтралізації конфліктів між детекторами використовується евклідова відстань між вхідним образом і ваговими векторами нейронів-переможців кожного з детекторів. Детектор, який має мінімальну евклідову відстань, є переможцем в конкурентній боротьбі і визначає клас і тип атаки.

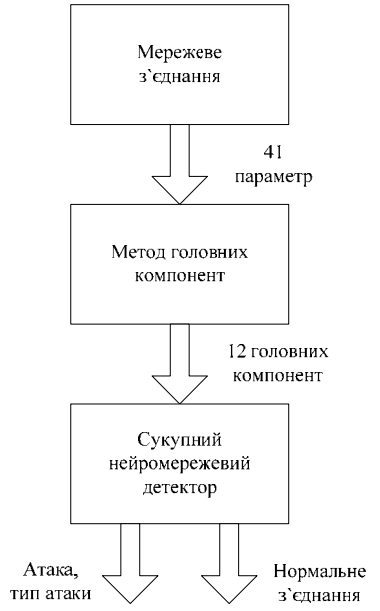

Рис. 1. Взаємодія PCA і сукупного нейромережевого детектора

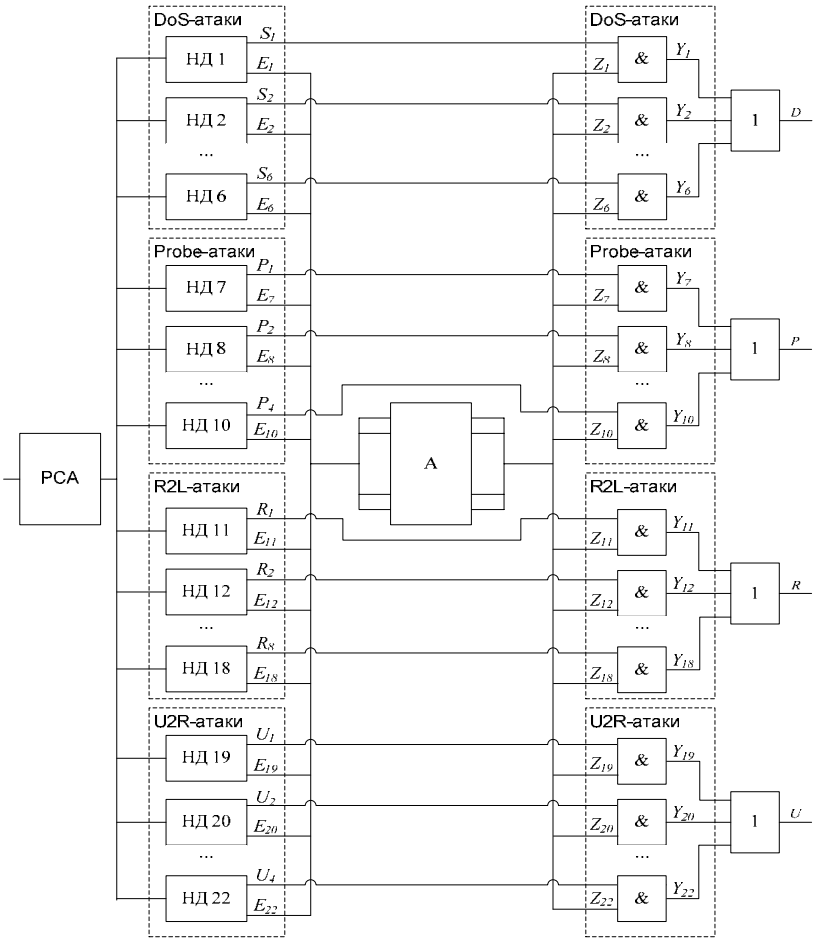

Рис. 2. Схема сукупного класифікатора для ієрархічної класифікації атак на ІТМ

Розглянемо функціонування сукупного класифікатора. Стиснений набір вхідних даних розмірністю 12 поступає на нейромережеві детектори, кожен з яких навчений на відповідний тип атак. В результаті, якщо детектор виявляє атаку, то вихідне значення його першого виходу встановлюється в одиничне значення. Для усунення конфліктів в роботі такого класифікатора, коли декілька детекторів встановлюються в одиничний стан, на другий вихід кожного детектора передається мінімальна евклідова відстань між вхідним образом і ваговими векторами відповідного детектора:

$$E_j = \min_j D_j = \min_i \sqrt{(x_1 - w_{1j})^2 + (x_2 - w_{2j})^2 + ... + (x_{12} - w_{12j})^2} \ . \qquad (2)$$

Інформація про мінімальну евклідову відстань поступає з кожного детектора на арбітр, який визначає детектор з номером k, що має мінімальну евклідову метрику:

$$E_k = \min E_j,\ j = \overline{1,22} \ . \qquad (3)$$

В результаті k-й вихід арбітра встановлюється в одиничний стан, а решта виходів – в нульовий стан:

$$Z_i = \begin{cases} 1, & якщо\ i = k \\ 0, & інакше \end{cases} \ . \qquad (4)$$

На виходах логічних елементів «І» визначається тип атаки:

$$Y_i = F_i Z_i \ , \qquad (5)$$

$$\text{де } F_i = \begin{cases} S_i, якщо\ i = \overline{1,6} \\ P_i, якщо\ i = \overline{7,10} \\ R_i, якщо\ i = \overline{11,18} \\ U_i, якщо\ i = \overline{19,22} \end{cases}$$

Виходи логічних елементів «АБО» визначають клас атаки:

$$D = \underset{i=1}{\overset{6}{\forall}} Y_i, \quad P = \underset{i=7}{\overset{10}{\forall}} Y_i, \quad R = \underset{i=11}{\overset{18}{\forall}} Y_i, \quad U = \underset{i=19}{\overset{22}{\forall}} Y_i, \qquad (6)$$

де *D – DoS*-атака; *P– Probe*-атака; *R– R2L*-атака; *U– U2R*-атака.

На рисунку 3 представлений розподіл в тривимірному просторі перших головних компонент різних мережевих атак і нормальних з'єднань, що візуально демонструє взаємозв'язок між даними, які розглядаються.

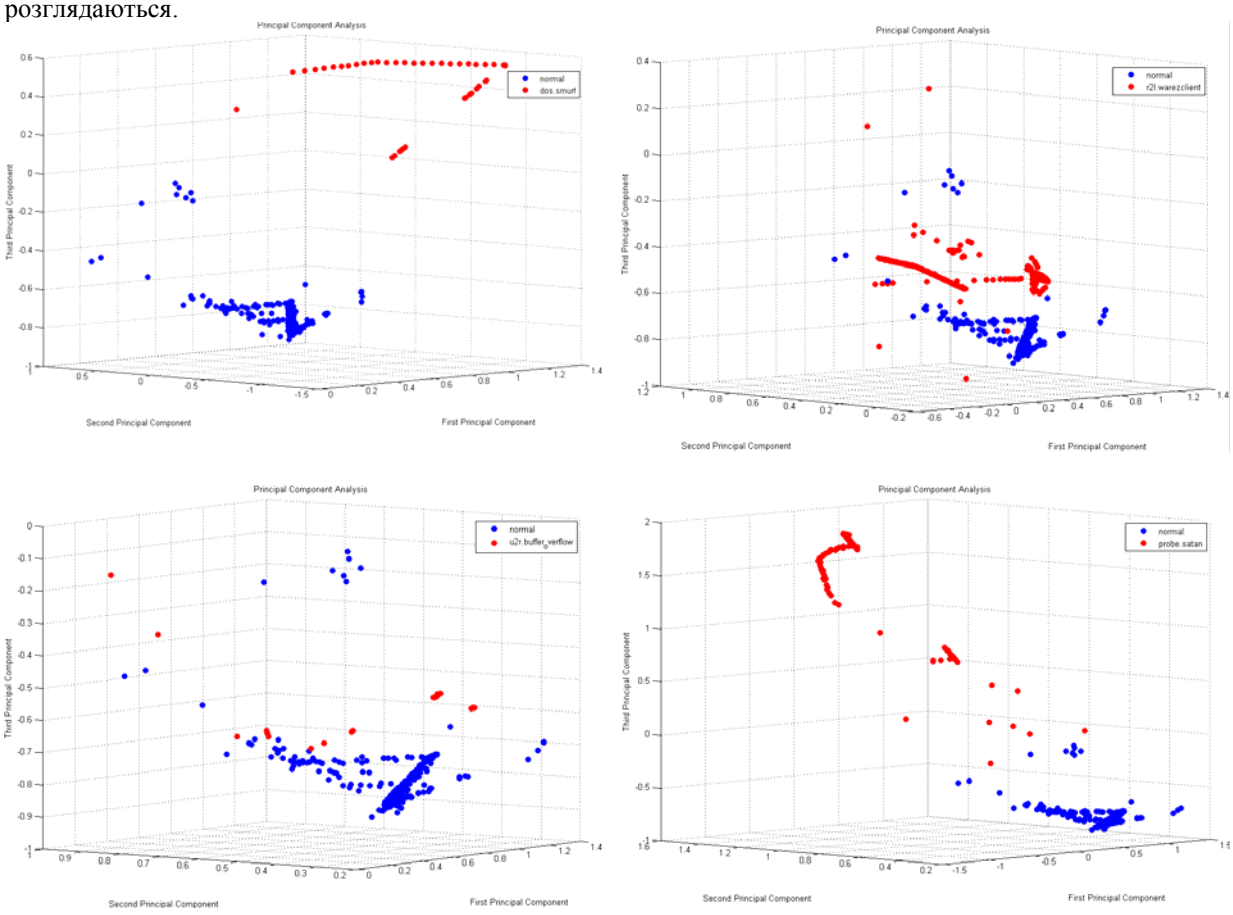

Рис. 3. Розподіл мережевих атак і нормальних з'єднань

Як видно з рис. 3, даний розподіл має нелінійний характер. Застосування методу головних компонент для зменшення розмірності даних, що описують мережевий трафік, показало, що, одна головна компонента містить 52,40% інформації, дві головних компоненти містять вже 71,67% інформації, три – 88,37% і т. д. Перших 12 головних компонент містять більше 99% інформації про мережевий трафік. У останніх 30 компонентах міститься менше 1% інформації, і, з міркування доцільності, їх можна виключити з аналізу [20].

Результати експериментальних досліджень показали, що при використанні методу головних компонент для попередньої обробки інформації про мережеві з'єднання, достовірність виявлення атак на ІТМ підвищується. Структура нейронної мережі у випадку використання PCA – 12-10-2 (12 вхідних нейронів, 10 нейронів прихованого шару і 2 вихідних нейрона). Отримані результати представлені в таблиці 1.

Таблиця 1

Порівняльний аналіз результатів виявлення атак

Атака	без PCA, %	PCA, %	Відхилення, %
DoS_Back	99,5	99,5	0,0
DoS_Land	90,5	100,0	+9,5
DoS_Neptune	100,0	100,0	0,0
DoS_Pod	98,1	98,1	0,0
DoS_Smurf	100,0	100,0	0,0
DoS_Teardrop	100,0	100,0	0,0
Probe_Ipsweep	7,1	65,2	+58,1
Probe_Nmap	54,5	100,0	+45,5
Probe_Portsweep	99,6	99,9	+0,3
Probe_Satan	99,3	99,3	0,0

Таким чином, для успішного аналізу мережевого трафіку досить використовувати 12 перших головних компонент, в яких міститься більше 99% інформації про мережеве з'єднання, а не 41 параметр. Це дозволяє істотно прискорити як процес навчання нейромережевого детектора, так і процес аналізу мережевого трафіку. Для цього, до виділених з мережевого трафіку даних необхідно застосувати спочатку метод головних компонент, а потім подати отримані дані на вхід нейронної мережі.

Отже, застосування методу головних компонент дозволило зменшити розмірність аналізованої інформації в 3,4 рази при втраті інформативності 0,8%.

Висновки

Запропоновано підхід до побудови сукупного класифікатора для ієрархічної класифікації атак на ІТМ на основі багатоканальних нейромережевих детекторів, що дало можливість зменшити розмірність аналізованої інформації в 3,4 рази при втраті інформативності 0,8% за рахунок стиснення вхідної інформації на основі методу головних компонент для отримання найбільш інформативних ознак, а також класифікувати типи і класи атак за рахунок об'єднання навчених на певний тип атаки нейромережевих детекторів. Запропонований підхід дозволив усунути конфлікти в роботі нейромережевих детекторів.

Література

1. Рейтинг наиболее громких кибератак последнего времени [Електронний ресурс]. – Режим доступу: http://internetua.com/sostavlen-reiting-naibolee-gromkih-kiberatak-poslednego-vremeni – Назва з екрану.
2. KDD Cup 1999 Data [Електронний ресурс]. – Режим доступу: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html – Назва з екрану.
3. Cannady J. Artificial neural networks for misuse detection / J. Cannady // Proceedings of the 21st national information systems security conference. – Arlington (USA), 1998. – P. 368-381.
4. Mukkamalaa S. Intrusion detection using an ensemble of intelligent paradigms / Mukkamalaa, A.H. Sung, A. Abraham // Journal of Network and Computer Applications. – 2005. – Vol. 28(2). – P.167-182.
5. Lorenzo-Fonseca I. Intrusion detection method using neural networks based on the reduction of characteristics / I. Lorenzo-Fonseca, F. Maciá-Pérez, F. Mora-Gimeno [et al.] // LNCS. – 2009. – Vol. 5517. – P. 1296-1303.
6. Grediaga A. Application of neural networks in network control and information security / A. Grediaga, F. Ibarra, F. García [et al.] // LNCS. – 2006. – Vol. 3973. – P. 208-213.
7. Zhang C. Comparison of BPL and RBF Network in intrusion detection system / C. Zhang, J. Jiang, M. Kamel // LNCS (LNAI). – 2004 – Vol. 2639. – P.460-470.
8. Cannady J. Applying CMAC-based online learning to intrusion detection / J. Cannady // Proceedings of the International Joint Conference on Neural Networks (IJCNN). – 2000. – Vol. 5. – P. 405-410.
9. Debar H. A neural network component for an intrusion detection system / H. Debar, M. Becker, D. Siboni // IEEE Computer Society Symposium on Research in Security and Privacy. – 1992. – P.240-250.
10. Cheng E. Network-based anomaly detection using an Elman network / E. Cheng, H. Jin, Z. Han, J. Sun // Networking and Mobile Computing, Lecture Notes in Computer Science, Springer, Berlin/Heidelberg. – 2006. – Vol. 3619. – P. 471-480.
11. Höglund A.J. A computer host-based user anomaly detection system using the self-organizing map / A.J. Höglund, K. Hätönen, A.S. Sorvari // Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks (IJCNN'00). – 2000. – Vol. 5. – P. 411-416.
12. Ramadas M. Detecting anomalous network traffic with self-organizing maps / M. Ramadas, S. Ostermann, B. Tjaden // LNCS. – 2003. – Vol. 2820. – P.36-54.
13. Sarasamma S.T. Hierarchical Kohonen net for anomaly detection in network security / S.T. Sarasamma, Q.A. Zhu, J. Huff // IEEE Transaction on Systems, Man and Cybernetics. – Part B 35 (2). – 2005. – P. 302-312.
14. Jirapummin C. Hybrid neural networks for intrusion detection system / C. Jirapummin, N. Wattanapongsakorn, P. Kanthamanon // Proceedings of the 2002 International Technical Conference On Circuits/Systems, Computers and Communications, Thailand. – 2002. – P. 928-931.
15. Horeis T. Intrusion detection with neural networks – Combination of self-organizing maps and radial basis function networks for human expert integration / T. Horeis // Tech. report, University of Passau, 2003 [Електронний ресурс]. – Режим доступу: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.106.191&rep=rep1&type=pdf – Назва з екрану.
16. Комар М.П. Методы искусственных нейронных сетей для обнаружения сетевых вторжений / М.П. Комар // Збірник тез сьомої Міжнародної науково-технічної конференції «Інтернет – Освіта – Наука» (ІОН-2010). – Вінниця (Україна), 2010. – С. 410–413.
17. Комар М.П. Система анализа сетевого трафика для обнаружения компьютерных атак / М.П. Комар // Вестник Брестского государственного технического университета: (Серия: физика,

математика и информатика). – 2010. – №5. – С. 14–16.

18. Jolliffe I. Principal component analysis / I.T. Jolliffe. – Springer, 2010. – 516 p.

19. Shilpa Lakhina. Feature Reduction using Principal Component Analysis for Effective Anomaly–Based Intrusion Detection on NSL-KDD / Shilpa Lakhina, Sini Joseph, Bhupendra Verma // International Journal of Engineering Science and Technology. – 2010. – Vol.2, № 6. – P. 1790–1799.

20. Komar M. Intelligent system for detection of networking intrusion / M. Komar, V. Golovko, A. Sachenko, S. Bezobrazov // Proceedings of the 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS-2011). – Prague (Czech Republic), 2011. – Vol.1. – P. 374-377.

References

1. Rating of the largest recent cyber attacks [Electronic resource]. – Mode of access: http://internetua.com/sostavlen-reiting-naibolee-gromkih-kiberatak-poslednego-vremeni – Title from the screen (In Russian).

2. KDD Cup 1999 Data [Electronic resource]. – Mode of access: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html – Title from the screen.

3. Cannady J. Artificial neural networks for misuse detection / J. Cannady // Proceedings of the 21st national information systems security conference. – Arlington (USA), 1998. – P. 368-381.

4. Mukkamalaa S. Intrusion detection using an ensemble of intelligent paradigms / Mukkamalaa, A.H. Sung, A. Abraham // Journal of Network and Computer Applications. – 2005. – Vol. 28(2). – P.167-182.

5. Lorenzo-Fonseca I. Intrusion detection method using neural networks based on the reduction of characteristics / I. Lorenzo-Fonseca, F. Maciá-Pérez, F. Mora-Gimeno [et al.] // LNCS. – 2009. – Vol. 5517. – P. 1296-1303.

6. Grediaga A. Application of neural networks in network control and information security / A. Grediaga, F. Ibarra, F. García [et al.] // LNCS. – 2006. – Vol. 3973. – P. 208-213.

7. Zhang C. Comparison of BPL and RBF Network in intrusion detection system / C. Zhang, J. Jiang, M. Kamel // LNCS (LNAI). – 2004 – Vol. 2639. – P.460-470.

8. Cannady J. Applying CMAC-based online learning to intrusion detection / J. Cannady // Proceedings of the International Joint Conference on Neural Networks (IJCNN). – 2000. – Vol. 5. – P. 405-410.

9. Debar H. A neural network component for an intrusion detection system / H. Debar, M. Becker, D. Siboni // IEEE Computer Society Symposium on Research in Security and Privacy. – 1992. – P.240-250.

10. Cheng E. Network-based anomaly detection using an Elman network / E. Cheng, H. Jin, Z. Han, J. Sun // Networking and Mobile Computing, Lecture Notes in Computer Science, Springer, Berlin/Heidelberg. – 2006. – Vol. 3619. – P. 471-480.

11. Höglund A.J. A computer host-based user anomaly detection system using the self-organizing map / A.J. Höglund, K. Hätönen, A.S. Sorvari // Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks (IJCNN'00). – 2000. – Vol. 5. – P. 411-416.

12. Ramadas M. Detecting anomalous network traffic with self-organizing maps / M. Ramadas, S. Ostermann, B. Tjaden // LNCS. – 2003. – Vol. 2820. – P.36-54.

13. Sarasamma S.T. Hierarchical Kohonen net for anomaly detection in network security / S.T. Sarasamma, Q.A. Zhu, J. Huff // IEEE Transaction on Systems, Man and Cybernetics. – Part B 35 (2). – 2005. – P. 302-312.

14. Jirapummin C. Hybrid neural networks for intrusion detection system / C. Jirapummin, N. Wattanapongsakorn, P. Kanthamanon // Proceedings of the 2002 International Technical Conference On Circuits/Systems, Computers and Communications, Thailand. – 2002. – P. 928-931.

15. Horeis T. Intrusion detection with neural networks – Combination of self-organizing maps and radial basis function networks for human expert integration / T. Horeis // Tech. report, University of Passau, 2003 [Electronic resource]. – Mode of access: http://citeseerx.ist.psu.edu/viewdoc/ download?doi=10.1.1.106.191&rep=rep1&type=pdf – Title from the screen.

16. Komar M. Methods of artificial neural networks for network intrusion detection / M. Komar // Proceedings of the Seventh International Scientific Conference "Internet - Education - Science - 2010", - Vinnitsa (Ukraine). – 2010. – P. 410-413 (In Ukrainian).

17. Komar M. System for analyzing network traffic to detect computer attacks / M. Komar // Herald Brest State Technical University. Physics, mathematics, computer science. – 2010. – №5. – P. 14-16 (In Russian).

18. Jolliffe I. Principal component analysis / I.T. Jolliffe. – Springer, 2010. – 516 p.

19. Shilpa Lakhina. Feature Reduction using Principal Component Analysis for Effective Anomaly–Based Intrusion Detection on NSL-KDD / Shilpa Lakhina, Sini Joseph, Bhupendra Verma // International Journal of Engineering Science and Technology. – 2010. – Vol. 2, № 6. – P. 1790–1799.

20. Komar M. Intelligent system for detection of networking intrusion / M. Komar, V. Golovko, A. Sachenko, S. Bezobrazov // Proceedings of the 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS-2011). – Prague (Czech Republic), 2011. – Vol.1. – P. 374-377.