Раздел 3. «IT-технологии, энергетика, автоматизация и вычислительная техника»

ҒТАМР 50.41.01
ӘӨЖ: 004.75
DOI: 10.4411/s030-034-427

А.Д. Тайсагатов

Қарағанды индустриялық университеті, Теміртау, Қазақстан
(E-mail: a.taisagatov@tttu.edu.kz)

Компьютерлік желілердегі ауытқуларды анықтаудың тиімді алгоритмдері

Бұл мақалада компьютерлік желілердегі ауытқуларды анықтау үшін қолданылатын тиімді алгоритмдер қарастырылады. Желілік трафиктегі қалыптан тыс өзгерістер жүйенің қауіпсіздігіне әсер етуі мүмкін, сондықтан ауытқуларды уақытылы анықтау маңызды. Зерттеу барысында машиналық оқыту әдістері, статистикалық талдау және сигнатуралық әдістер талданады. Алынған нәтижелер негізінде әртүрлі әдістердің тиімділігі бағаланып, олардың қолдану мүмкіндіктері қарастырылады.

Ключевые слова: Компьютерлік желілер, ауытқуларды анықтау, машиналық оқыту, статистикалық талдау, желілік қауіпсіздік.

Кіріспе

Қазіргі заманғы компьютерлік желілерде үлкен көлемде деректер алмасады, бұл өз кезегінде қауіпсіздік мәселелерін туындатады. Желілік ауытқулар көбінесе кибершабуылдардың, жабдықтың дұрыс істемеуінің немесе қалыпты емес трафиктің салдарынан орын алады. Сондықтан тиімді алгоритмдерді қолдану арқылы бұл ауытқуларды анықтау маңызды.
Желілік қауіпсіздікті қамтамасыз ету үшін заманауи әдістер мен технологиялар қолданылады. Олардың ішінде жасанды интеллектке негізделген жүйелер, машиналық оқыту алгоритмдері, статистикалық талдау және қолтаңбаға негізделген әдістер кеңінен таралған. Машиналық оқыту әдістері трафиктің қалыпты және қалыпты емес үлгілерін ажыратуға мүмкіндік береді, бұл шабуылдарды ерте кезеңде анықтауға және алдын алуға көмектеседі. Сонымен қатар, желілік аномалияларды анықтау үшін ережеге негізделген жүйелер, эвристикалық талдау және мінез-құлықтық модельдеу әдістері қолданылады.

Киберқауіпсіздік саласында қолданылатын негізгі тәсілдердің бірі – қолтаңбалық әдіс, онда белгілі бір зиянды әрекеттердің үлгілері сақталып, желілік трафик осы үлгілермен салыстырылады. Алайда, бұл әдістің негізгі кемшілігі – жаңа, белгісіз шабуыл түрлерін анықтай алмауы. Осыған байланысты, статистикалық талдау және аномалияны анықтау әдістері жоғары тиімділікті көрсетеді. Бұл тәсілдер қалыпты желілік әрекеттердің математикалық моделін құрып, кез келген ауытқуды нақты уақыт режимінде бақылауға мүмкіндік береді. Желілік аномалияларды анықтау үшін мәліметтерді алдын ала өңдеу, ерекшеліктерді таңдау, кластерлеу, жіктеу және аномалияларды талдау секілді кезеңдерден тұратын кешенді тәсілдер қолданылады. Әсіресе, үлкен көлемдегі деректермен жұмыс істеу кезінде тиімді алгоритмдерді пайдалану жүйенің өнімділігін арттырады және жалған оң нәтижелердің санын азайтады.

Қазіргі заманғы желілік инфрақұрылымдарда қауіпсіздікті қамтамасыз ету үшін гибридті әдістер де қолданылады, онда бірнеше тәсіл біріктіріліп, олардың артықшылықтары біріктіріледі. Бұл көпдеңгейлі қорғау жүйесін қалыптастырып, ықтимал шабуылдардың алдын алу мен оларды тиімді анықтауға мүмкіндік береді. Жалпы, желілік қауіпсіздікті қамтамасыз ету – үздіксіз жетілдіруді қажет ететін сала. Жаңа киберқауіптер мен шабуыл түрлеріне қарсы тұру үшін жасанды интеллект, машиналық оқыту және үлкен деректерді талдау әдістерін кеңінен пайдалану маңызды. Осы мақалада қарастырылған әдістер заманауи желілердегі қауіпсіздікті арттырудың негізгі құралдары ретінде қарастырылады.

Раздел 3. «IT-технологии, энергетика, автоматизация и вычислительная техника»

Әдістер мен материалдар

Желідегі ауытқуларды анықтау үшін түрлі әдістер қолданылады, олардың әрқайсысы желілік трафиктің қалыпты және қалыптан тыс үлгілерін ажыратуда өзіндік ерекшеліктерге ие. Бұл әдістердің ішінде статистикалық талдау, машиналық оқыту, сигнатуралық әдістер және гибридті тәсілдер ерекше орын алады. Статистикалық талдау әдістері желілік трафиктің тарихи мәліметтеріне негізделеді. Бұл әдіс желідегі деректер ағынын талдап, олардың орташа мәні, дисперсиясы және басқа статистикалық сипаттамалары бойынша қалыпты үлгілерін анықтауға көмектеседі. Егер ағымдағы трафиктің параметрлері осы қалыпты үлгілерден айтарлықтай ауытқыса, онда ол ықтимал шабуыл немесе аномалия ретінде қарастырылады. Мысалы, желілік пакеттердің орташа ағынын төмендегі формула бойынша есептеуге болады:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

x_i - белгілі бір уақыттағы трафик көлемі
N - бақылау саны.

Аномалияны анықтау үшін статистикалық шекті анықтау маңызды. Ол стандартты ауытқу негізінде есептеледі:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

Егер ағымдағы трафиктің мәні $\mu \pm k\sigma$ диапазонынан шығып кетсе, онда ол күмәнді трафик ретінде қарастырылады. Мұнда k – пайдаланушының қалауы бойынша таңдалатын шек коэффициенті

Машиналық оқыту әдістері қадағаланатын (supervised) және қадағалаусыз (unsupervised) тәсілдерге бөлінеді. Қадағаланатын оқыту кезінде модельдер алдын ала белгіленген деректер жиыны бойынша оқытылып, қалыпты және аномальді трафикті ажырата алады. Ал қадағалаусыз оқытуда алгоритмдер деректер жиынының ішінен кластерлерді анықтап, олардың ерекшеліктеріне сүйене отырып аномалияларды табады. Мысалы, K-Means кластерлеу әдісі келесідей есептеледі:

$$J = \sum_{i=1}^{N}\sum_{j=1}^{k} \omega_{ij} \, ||x_i - c_j||^2$$

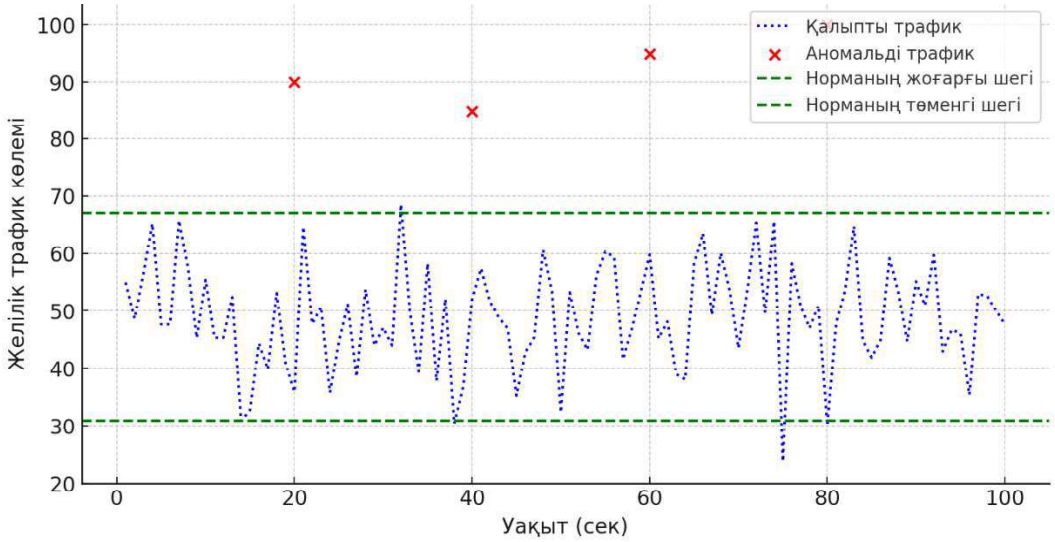
мұнда c_j – кластер орталығы, x_i – деректер нүктесі, ω_{ij} – деректердің белгілі бір кластерге тиесілігін анықтайтын айнымалы.

Сигнатуралық әдістер белгілі бір зиянды әрекеттердің қолтаңбаларына негізделген. Бұл әдісте желілік шабуылдардың сипаттамалары алдын ала сақталып, трафик осы үлгілермен салыстырылады. Мысалы, белгілі бір IP мекенжайлардан немесе белгілі бір порттар арқылы келетін пакеттер қауіпті шабуылдардың белгілері ретінде қарастырылады. Мұндай жүйелер Snort немесе Suricata сияқты желілік қауіпсіздік құралдарында қолданылады.

Гибридті әдістер бірнеше тәсілді біріктіру арқылы тиімділікті арттырады. Мысалы, машиналық оқыту мен статистикалық талдау әдістерін бірге қолдану арқылы шынайы уақыттағы аномалияларды анықтауға мүмкіндік береді. Гибридті әдістер әсіресе үлкен көлемдегі деректерді өңдеуде және күрделі кибершабуылдарды анықтауда тиімді.

Төмендегі графикте қалыпты және аномальді желілік трафик көрсетілген. Қалыпты трафиктің диапазоны көк аймақпен белгіленген, ал күмәнді трафик қызыл нүктелермен берілген.

Раздел 3. «IT-технологии, энергетика, автоматизация и вычислительная техника»



Сурет 1. Желілік трафиктің қалыпты және аномальді үлгілері

Жоғарыдағы диаграммада қалыпты желілік трафик көк сызықпен көрсетілген, ал күмәнді немесе аномальді трафик қызыл нүктелермен белгіленген. Жасыл сызықтар қалыпты трафиктің шекарасын көрсетеді. Егер трафик осы шектерден шығып кетсе, онда ол аномалия ретінде қарастырылады.

Бұл әдістер желілік қауіпсіздікті қамтамасыз етуге және кибершабуылдардың алдын алуға көмектеседі. Статистикалық талдау және машиналық оқыту әдістерін біріктіру арқылы дәлірек және сенімді аномалияларды анықтау жүйесін құруға болады.

Нәтижелер мен пікірталас

Зерттеу барысында жоғарыда аталған әдістердің артықшылықтары мен кемшіліктері анықталды. Статистикалық әдістер нақты уақыттағы талдау үшін қолайлы, себебі олар желілік трафиктің қалыпты үлгілерін анықтап, шекті мәндерден ауытқуларды жылдам тіркей алады. Алайда, бұл әдістер жаңа шабуылдарды анықтауда шектеулі, себебі олар бұрыннан белгілі үлгілерге сүйенеді және шабуыл әдістері өзгерген жағдайда тиімділігі төмендейді. Сонымен қатар, статистикалық тәсілдер жалған оң нәтижелер беруі мүмкін, себебі олар қалыпты жүйелік өзгерістерді де аномалия ретінде қабылдайды.

Машиналық оқыту әдістері үлкен деректер жиынтығын өңдеуге және бұрын белгісіз болған жасырын шабуылдарды анықтауға мүмкіндік береді. Бұл әдістер, әсіресе, аномалияларды анықтау мен үлгілерді тану барысында жоғары дәлдік көрсетеді. Бірақ олардың тиімділігі оқу деректерінің сапасына тәуелді және дұрыс конфигурацияланбаған жағдайда жалған нәтижелер беруі мүмкін. Сонымен қатар, мұндай әдістер үлкен есептеу ресурстарын қажет етеді, әсіресе нақты уақыттағы мониторинг жүйелерінде.

Сигнатуралық әдістер дәстүрлі түрде кеңінен қолданылады, себебі олар белгілі зиянды әрекеттердің қолтаңбаларына негізделеді. Бұл әдістер танылған шабуылдарды жоғары дәлдікпен анықтай алады, сондай-ақ жалған оң нәтижелердің санын азайтады. Дегенмен, олардың басты кемшілігі – белгісіз шабуылдарға қарсы әлсіз болуы. Егер жаңа шабуылдар жүйеде тіркелмесе, оларды бұл әдіспен анықтау мүмкін емес. Осы себепті, сигнатуралық әдістер көбінесе толыққанды қорғаныс үшін басқа тәсілдермен бірге қолданылады.

Гибридті әдістер әртүрлі тәсілдерді біріктіре отырып, ең жоғары тиімділікті қамтамасыз етеді. Олар статистикалық талдау, машиналық оқыту және сигнатуралық әдістердің артықшылықтарын біріктіріп, желідегі күмәнді белсенділікті неғұрлым дәл анықтауға мүмкіндік береді. Гибридті тәсілдер жүйенің қауіпсіздігін арттырып, жалған оң және жалған теріс анықтаулардың санын азайтады. Алайда, мұндай тәсілдерді енгізу күрделі инфрақұрылымды қажет етеді және жүйені үнемі жаңартып отыруды талап етеді.

Зерттеу нәтижелері көрсеткендей, желілік қауіпсіздікті қамтамасыз ету үшін бір ғана әдісті қолдану жеткіліксіз. Кибершабуылдардың үнемі өзгеріп отыруына байланысты әртүрлі әдістерді

Раздел 3. «IT-технологии, энергетика, автоматизация и вычислительная техника»

үйлестіріп пайдалану қажет. Гибридті тәсілдер желілік трафикті тиімді талдап, қалыптан тыс әрекеттерді жылдам анықтауға мүмкіндік береді. Болашақта жасанды интеллект пен үлкен деректерді өңдеу технологияларын кеңінен қолдану арқылы желілік аномалияларды анықтау жүйелерінің тиімділігін одан әрі арттыруға болады.

Қорытынды

Желілік қауіпсіздікті қамтамасыз ету үшін ауытқуларды анықтау әдістерін жетілдіру маңызды. Зерттеу көрсеткендей, машиналық оқыту және гибридті әдістер желідегі қауіптерді тиімді анықтауға мүмкіндік береді. Бұл әдістер үлкен деректерді өңдеуге, шабуылдарды ерте кезеңде анықтауға және жалған оң нәтижелердің санын азайтуға көмектеседі. Болашақта алгоритмдердің өнімділігін арттыру және нақты уақыттағы өңдеуді жақсарту үшін жаңа тәсілдерді қолдану қажет. Жасанды интеллект және терең нейрондық желілерді пайдалану арқылы күрделі шабуыл үлгілерін жылдам анықтауға және оларды алдын ала болжауға болады. Сонымен қатар, машиналық оқыту модельдерін үнемі жаңартып, шабуылдардың эволюциясына бейімдеу маңызды.

Үлкен деректерді өңдеу технологияларын жетілдіру арқылы желілік трафикті жылдам талдау мүмкіндігі артады. Бұлттық есептеулер мен үлестірілген жүйелерді қолдану арқылы жүктемені азайтып, деректерді өңдеу уақытын қысқартуға болады. Бақылаусыз оқыту әдістерін кеңінен енгізу арқылы белгісіз шабуылдарды анықтау тиімділігі артады, ал күшейтілген оқыту әдістерін қолдану шабуылдарға қарсы динамикалық жауап қайтару мүмкіндігін жақсартады.

Киберқауіпсіздікті қамтамасыз етуде блокчейн және криптографиялық әдістердің рөлі артып келеді. Блокчейн технологиясы арқылы желілік жазбалардың өзгертілмейтіндігін қамтамасыз етуге болады, ал гомоморфтық шифрлау деректерді қорғап, жүйенің қауіпсіздігін арттырады.

Желілік аномалияларды анықтау жүйелерінің тиімділігін арттыру үшін көпфакторлы талдау әдістерін енгізу маңызды. Желідегі әртүрлі параметрлерді, соның ішінде құрылғылардың мінез-құлық үлгілерін, қолданушы әрекеттерін және трафик статистикасын бір мезгілде талдау арқылы шабуылдарды дәлірек анықтауға болады.

Желілік қауіпсіздікті күшейту үшін автономды өзін-өзі оқытатын жүйелерді дамыту да маңызды бағыт болып табылады. Мұндай жүйелер шабуылдарды анықтап қана қоймай, олардан қорғану стратегияларын да автоматты түрде бейімдей алады. Бұл жүйелерге жасанды интеллект және машиналық оқыту технологияларын біріктіру арқылы желілік инфрақұрылымның қауіпсіздігін жоғарылатуға болады. Жалпы, киберқауіптердің үнемі өзгеріп отыруына байланысты желілік аномалияларды анықтау әдістерін жетілдіру тоқтаусыз процесс болып табылады. Жаңа технологияларды енгізу, тиімді алгоритмдерді дамыту және жүйелерді үздіксіз жаңарту арқылы желілік қауіпсіздікті жоғары деңгейде ұстап тұруға болады.

Әдебиеттер тізімі

1 Chandola, V., Banerjee, A., and Kumar, V. (2009) Anomaly detection: A survey. ACM Computing Surveys (CSUR), 41, 15:158.

2 Ahmed M, Mahmood A N, Hu J. A Survey of Network Anomaly Detection Techniques[J]. Journal of Network and Computer Applications, 2015, 60:19-31.

3 Yu W, Aggarwal CC, Ma S, Wang H. On anomalous hotspot discovery in graph streams. In: Proceedings of the 13th IEEE International Conference on Data Mining (ICDM), Dallas, TX, 2013.

4 Ide, T. and Kashima, H., Eigenspace-Based Anomaly Detection in Computer Sys- tems, ACM SIGKDD 2004, pp.440-449.

5 Eslami M, Zheng G, Eramian H, et al. Anomaly detection on bipartite graphs for cyber situational awareness and threat detection[C]// 2017 IEEE International Conference on Big Data (Big Data). IEEE, 2017.

6 ISCXIDS2012)[OL].https://www.unb.ca/cic/datasets/ids.html Canadian Institute for Cybersecurity.

7 Denning, D. E. (1987). An Intrusion-Detection Model. IEEE Transactions on Software Engineering.

8 Lakhina, A., Crovella, M., & Diot, C. (2004). Diagnosing Network-Wide Traffic Anomalies. ACM SIGCOMM.

9 Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly Detection: A Survey. ACM Computing

Surveys.

10 Sommer, R., & Paxson, V. (2010). Outside the Closed World: On Using Machine Learning for Network Intrusion Detection. IEEE Symposium on Security and Privacy.

А.Д. Тайсагатов

Эффективные алгоритмы обнаружения аномалий в компьютерных сетях

В данной статье рассматриваются эффективные алгоритмы, используемые для обнаружения аномалий в компьютерных сетях. Аномальные изменения в сетевом трафике могут влиять на безопасность системы, поэтому своевременное обнаружение аномалий важно. В исследовании анализируются методы машинного обучения, статистического анализа и техники, основанные на сигнатурах. На основе полученных результатов оценивается эффективность различных методов и рассматриваются возможности их применения.

Ключевые слова: Компьютерные сети, обнаружение аномалий, машинное обучение, статистический анализ, безопасность сети.

A.D. Taisagatov

Effective Algorithms for Detecting Anomalies in Computer Networks

This article examines the effective algorithms used for detecting anomalies in computer networks. Abnormal changes in network traffic can impact the security of the system, so timely detection of anomalies is important. The study analyzes machine learning methods, statistical analysis, and signature-based techniques. Based on the obtained results, the effectiveness of different methods is evaluated, and their application possibilities are considered.

Key words: Computer networks, anomaly detection, machine learning, statistical analysis, network security.

References

1 Chandola, V., Banerjee, A., and Kumar, V. (2009) Anomaly detection: A survey. ACM Computing Surveys (CSUR), 41, 15:158.

2 Ahmed M, Mahmood A N, Hu J. A Survey of Network Anomaly Detection Techniques[J]. Journal of Network and Computer Applications, 2015, 60:19-31.

3 Yu W, Aggarwal CC, Ma S, Wang H. On anomalous hotspot discovery in graph streams. In: Proceedings of the 13th IEEE International Conference on Data Mining (ICDM), Dallas, TX, 2013.

4 Ide, T. and Kashima, H., Eigenspace-Based Anomaly Detection in Computer Sys- tems, ACM SIGKDD 2004, pp.440-449.

5 Eslami M, Zheng G, Eramian H, et al. Anomaly detection on bipartite graphs for cyber situational awareness and threat detection[C]// 2017 IEEE International Conference on Big Data (Big Data). IEEE, 2017.

11 ISCXIDS2012)[OL].https://www.unb.ca/cic/datasets/ids.html Canadian Institute for Cybersecurity.

12 Denning, D. E. (1987). An Intrusion-Detection Model. IEEE Transactions on Software Engineering.

13 Lakhina, A., Crovella, M., & Diot, C. (2004). Diagnosing Network-Wide Traffic Anomalies. ACM SIGCOMM.

14 Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly Detection: A Survey. ACM Computing Surveys.

15 Sommer, R., & Paxson, V. (2010). Outside the Closed World: On Using Machine Learning for Network Intrusion Detection. IEEE Symposium on Security and Privacy.