

ЖЕЛІЛІК ШАБУЫЛДАРДЫ АНЫҚТАУДА НЕЙРОНДЫҚ ЖЕЛІЛЕРДІ ҚОЛДАНУ

КЕРЕЕВ А.К.^{1*}, ТӘҢІРБЕРГЕН А.К.¹, АБДЫГАЛИЕВА Г.Б.²,
ИГИСИНОВ С.Ж.³, ТОЛЕУОВ Т.Ж.¹

***Кереев Адилжан Кутымович**¹ – PhD, информатика және ақпараттық технологиялар кафедрасының доценті, Қ.Жұбанов атындағы Ақтөбе өңірлік университеті, Ақтөбе қ., Қазақстан

E-mail: adiljan@mail.ru, <https://orcid.org/0000-0002-8283-5807>

Тәңірберген Айсұлу Көбейсіңқызы¹ – Магистр, Қ.Жұбанов атындағы Ақтөбе өңірлік университеті, Ақтөбе қ., Қазақстан

E-mail: aisulu21@mail.ru, <https://orcid.org/0009-0009-2936-4637>

Абдығалиева Гүлдария Болатбековна² – Магистр, информатика пәні мұғалімі, педагог-шебер, Фариза Оңғарсынова атындағы №41 мектеп-гимназиясы, Ақтөбе қ., Қазақстан

E-mail: guldariya78@mail.ru, <https://orcid.org/0009-0001-0552-8475>

Игисинов Сабит Жандарбекович³ – PhD, математика кафедрасының қауымдастырылған профессор м.а., М.Х. Дулати атындағы Тараз университеті, Тараз қ., Қазақстан.

E-mail: igisinovsabit@gmail.com, <https://orcid.org/0000-0002-4189-9142>

Толлеуов Тимур Жаксылыкович¹ – Магистр, информатика және АТ кафедрасының аға оқытушысы, Қ. Жұбанов атындағы Ақтөбе өңірлік университеті, Ақтөбе қ., Қазақстан

E-mail: timur_toleuov@mail.ru, <https://orcid.org/0000-0002-6050-1056>

Аңдатпа. Қазіргі уақытта кибершабуылдардың саны мен күрделілігінің артуы компьютерлік желілердің қауіпсіздігін қамтамасыз ету мәселесін ерекше өзекті етеді. Дәстүрлі сигнатуралық әдістер белгілі шабуылдарды анықтауда тиімді болғанымен, жаңа немесе өзгертілген шабуыл түрлерін анықтауда олардың мүмкіндіктері шектеулі. Осыған байланысты желілік трафикті интеллектуалды талдау әдістерін қолдану қажеттілігі туындайды. Зерттеудің мақсаты – машиналық оқыту әдістерін пайдалана отырып, желілік шабуылдарды анықтаудың тиімді моделін әзірлеу және оның өнімділігін бағалау. Эксперименттік база ретінде әртүрлі кибершабуыл түрлері мен қалыпты трафикті қамтитын CICIDS2018 деректер жиынтығы қолданылды. Зерттеу барысында деректерді алдын ала өңдеу, жетіспейтін және қате мәндерді жою, маңызды белгілерді іріктеу, белгілерді нормализациялау және көпклассты классификацияға арналған жасанды нейрондық желі моделін құру әдістері пайдаланылды. Ұсынылған модель екі жасырын қабаттан тұрады және ReLU активациясы, Dropout регуляризациясы және Adam оптимизаторын қолдану арқылы оқытылды. Эксперимент нәтижелері модельдің шамамен 96,5% дәлдікке, 0,93 macro-F1 және 0,96 weighted-F1 көрсеткіштеріне жеткенін көрсетті. Модель сирек кездесетін шабуылдарды қоса алғанда, әртүрлі шабуыл түрлерін тиімді анықтап, нақты уақыт режимінде желілік трафикті талдауға мүмкіндік береді. Зерттеу нәтижелері интеллектуалды шабуылдарды анықтау жүйелерін жетілдіруге және киберқауіпсіздік саласындағы заманауи технологияларды дамытуға мүмкіндік береді.

Түйін сөздер: киберқауіпсіздік, шабуылдарды анықтау жүйесі, машиналық оқыту, нейрондық желілер, желілік трафик, CICIDS2018

Кіріспе

Киберқауіптердің қарқынды өсуі жағдайында желілік шабуылдарды тиімді анықтау ақпараттық жүйелердің қауіпсіздігін қамтамасыз етуде аса маңызды міндетке айналууда. Бұл жұмыста әртүрлі желілік шабуылдар мен қалыпты трафик түрлерін қамтитын ауқымды деректер жиынтығы – CICIDS2018 датасеті негізінде шабуылдарды анықтау үшін жасанды нейрондық желілерді (ЖНЖ) қолдану қарастырылады. Зерттеудің мақсаты – желілік трафиктегі шабуылдарды анықтаудың дәлдігі мен жылдамдығын арттыру үшін ЖНЖ негізінде гибриді модельді әзірлеу және бағалау.

Тақырыптың өзектілігі өзгермелі жағдайлар мен жаңа шабуыл түрлеріне бейімделе алатын киберқауіптерден қорғаудың тиімді әдістерін әзірлеу қажеттілігімен түсіндіріледі. 80-нен астам түрлі шабуыл класын қамтитын CICIDS2018 датасеті әзірленген модельдерді тестілеу мен валидациялауға бірегей мүмкіндік береді. Осы жұмыста үш негізгі міндет қарастырылады: датасет деректерін талдау және алдын ала өңдеу, ЖНЖ-нің гибриді архитектурасын әзірлеу және ұсынылған модельдің өнімділігін тестілік деректер жиынтығында бағалау.

Зерттеу әдістемесі белгілерді шығару, нейрондық желі архитектурасын таңдау және гиперпараметрлерді оңтайландыру кезеңдерін қамтиды. Атап айтқанда, құрылымдалған және құрылымдалмаған деректерді тиімді өңдеуге мүмкіндік беретін көпқабатты персептрондар мен конволюциялық нейрондық желілерді біріктіретін гибриді модель қолданылады. Зерттеу нысаны – желілік трафикті терең талдау мен шабуылдарды анықтауға мүмкіндік беретін CICIDS2018 датасеті.

Осылайша, бұл зерттеу машиналық оқытудың заманауи әдістерін қолдана отырып, шабуылдарды анықтау жүйелерінің тиімділігін арттыру мәселесін шешуге бағытталған. Жұмыс нәтижелері киберқауіпсіздік технологияларын одан әрі дамытуға және желілік трафикті талдаудың жаңа тәсілдерін енгізуге ықпал етуі мүмкін. Бұл салада ЖНЖ-нің гибриді моделін қолдану ақпараттық қауіпсіздік жүйелерін жетілдіру үшін жаңа мүмкіндіктер ашады.

Соңғы жылдары сигнатуралық әдістер мен жасанды нейрондық желілер алгоритмдерін біріктіретін гибриді шабуылдарды анықтау жүйелеріне (IDS) деген қызығушылық тұрақты түрде артып келеді. Сигнатуралық талдау белгілі қауіптерді дәл анықтауы және жалған іске қосылу деңгейінің төмендігі арқасында әлі де жоғары маңызға ие [1], алайда жаңа немесе модификацияланған шабуылдар пайда болған жағдайда оның тиімділігі айтарлықтай төмендейді [2].

Нейрондық желілерге негізделген тәсілдер осы шектеулердің орнын толтыруға бағытталған. Белгілерді іріктеу әдістеріне және рекурренттік нейрондық желілерді қолдануға негізделген зерттеулер, соның ішінде CICIDS2018 деректер жиынтығын пайдалану арқылы, шабуылдарды жіктеу сапасының жақсарғанын көрсетеді [3]. Терең сенім желілеріне (DBN) негізделген архитектураларды қолдану жоғары өлшемді деректерді тиімді өңдеуге мүмкіндік беретінін дәлелдейді [4].

Жүйенің тұрақтылығы мен дәлдігін едәуір арттыруға бірнеше машиналық оқыту алгоритмдерін дауыс беру немесе мета-интеграция механизмдері арқылы біріктіретін гибриді ансамбльдік IDS жүйелері мүмкіндік береді [5; 6]. Сонымен қатар федеративті оқыту және дифференциалдық құпиялық әдістері бағытындағы зерттеулер де қарқынды дамуда, бұл деректер құпиялылығын сақтай отырып, таратылған ортада IDS жүйелерін оқытуға мүмкіндік береді [7].

Нейрондық желілерге негізделген IDS жүйелерінің adversarial-әсерлерге тұрақтылығы мәселесіне де ерекше назар аударылады, себебі қарсы әрекет етуші шабуылдар жүйенің сенімділігін едәуір төмендетуі мүмкін [8]. Гибриді архитектураларда сигнатуралық компонент бастапқы сүзгілеу қызметін атқарып, нейрондық модульге түсетін жүктемені азайтады және жүйенің жалпы жалпылау қабілетін арттырады [2; 7].

Үлкен деректер жиынтықтарымен, соның ішінде CICIDS2018 деректерімен жұмыс істеуде белгілерді оңтайландыру және іріктеу маңызды рөл атқарады. Зерттеулер рекурсивті белгілерді алып тастау әдістерін қолданатын гибриді машиналық оқыту архитектураларының тиімділігін көрсетеді, бұл жоғары дәлдікті сақтай отырып есептеу шығындарын азайтуға мүмкіндік береді. Үлкен деректер платформаларын және HDFS сияқты таратылған сақтау жүйелерін пайдалану IDS жүйелерінің ауқымдылығын қамтамасыз етіп, оларды бұлттық және шеткі есептеу орталарымен біріктіруге мүмкіндік береді [9].

Қосымша зерттеулер тобы автомобиль желілері мен IoT жүйелерін қоса алғанда, нақты қауіп сценарийлерін қарастырады, мұнда шабуылдарды анықтау үшін жаңа буындағы мультимодальды және генеративті архитектуралар қолданылады [10]. Машиналық оқыту алгоритмдерін салыстырмалы талдау нәтижелері әмбебап шешімнің жоқ екенін көрсетіп, IDS жүйелерін нақты датасет сипаттамалары мен шабуыл сценарийлеріне бейімдеудің маңыздылығын айқындайды [11; 12].

Қазіргі зерттеу нәтижелері нейрондық желілер мен сигнатуралық талдауды біріктіретін гибриді IDS жүйелері жоғары дәлдік, тұрақтылық және бейімделгіштік қамтамасыз ететінін дәлелдейді. CICIDS2018 датасеті негізінде жүргізілген зерттеулер мұндай жүйелердің белгілі және белгісіз шабуылдарды тиімді анықтай алатынын көрсетеді, бұл гибриді тәсілдің жаңа буындағы киберқауіпсіздік жүйелерін құрудағы болашағын айқындайды.

Зерттеу материалдары мен әдістері

Бұл зерттеуде желілік шабуылдарды анықтау жүйесінің тиімділігін әзірлеу және бағалау үшін нақты желілік трафик сценарийлері мен әртүрлі кибершабуыл түрлерін қамтитын CICIDS2018 деректер жиынтығы пайдаланылды.

CICIDS2018 деректер жиынтығы желілік шабуылдарды анықтау жүйелерін зерттеуде кеңінен қолданылатын эталондық деректердің бірі болып табылады. Ол Canadian Institute for Cybersecurity (CIC) институты мен Communications Security Establishment (CSE) ұйымының бірлесуімен әзірленген. Деректер жиынтығын құрудың негізгі мақсаты – басқарылатын эксперименттік ортада шынайы желілік трафик пен қазіргі заманғы кибершабуыл сценарийлерін модельдеу.

Деректер бұлттық инфрақұрылымда орналастырылған корпоративтік желіні имитациялау жағдайында алынған. Эксперименттік желі ұйымның бірнеше бөлімшелерін, серверлерді және клиенттік компьютерлерді қамтыды, бұл пайдаланушылар мен желілік сервистердің типтік әрекеттерін қайта құруға мүмкіндік берді.

Желілік трафик бастапқыда PCAP форматында жазылып, кейін CICFlowMeter құралының көмегімен өңделді. Бұл құрал желілік ағындардың статистикалық сипаттамаларын шығарады. Әрбір ағын үшін пакеттердің параметрлерін, ағын сипаттамаларын, пакеттер арасындағы уақыт интервалдарын, тақырып параметрлерін және TCP-флагтар есептегіштерін сипаттайтын шамамен 80 белгі қалыптастырылады. Әрбір желілік ағынға трафик түрін (қалыпты немесе шабуыл) көрсететін класс белгісі беріледі, бұл деректер жиынтығын машиналық оқыту алгоритмдерін оқыту және тестілеу үшін қолайлы етеді.

Деректер жиынтығына Brute Force, Botnet, DDoS, DoS, Infiltration, Heartbleed, сондай-ақ веб-шабуылдар (SQL Injection және XSS) сияқты заманауи желілік шабуыл түрлері енгізілген. Шабуыл сценарийлерінің әртүрлілігі мен алынған белгілер санының көптігі CICIDS2018 деректер жиынтығын интеллектуалды шабуылдарды анықтау жүйелерін әзірлеу және салыстырмалы талдау жүргізу үшін кеңінен қолдануға мүмкіндік береді.

Кесте 1. CICIDS2018 деректер жиынтығындағы шабуыл түрлері

№	Шабуыл түрі	Қысқаша сипаттамасы
1	Benign	Қалыпты желілік трафик
2	DoS	Қызмет көрсетуден бас тарту шабуылы
3	DDoS	Таратылған қызмет көрсетуден бас тарту шабуылы
4	Brute Force	Құпиясөздерді толық перебор әдісі арқылы табу
5	Botnet	Құрылғыларды ботнет желісіне жұқтыру
6	Infiltration	Зиянкестің ішкі желіге енуі
7	Web Attack – SQL Injection	Зиянды SQL-код енгізу
8	Web Attack – XSS	Кросс-сайттық скрипт орындау
9	Heartbleed	OpenSSL протоколындағы осалдық

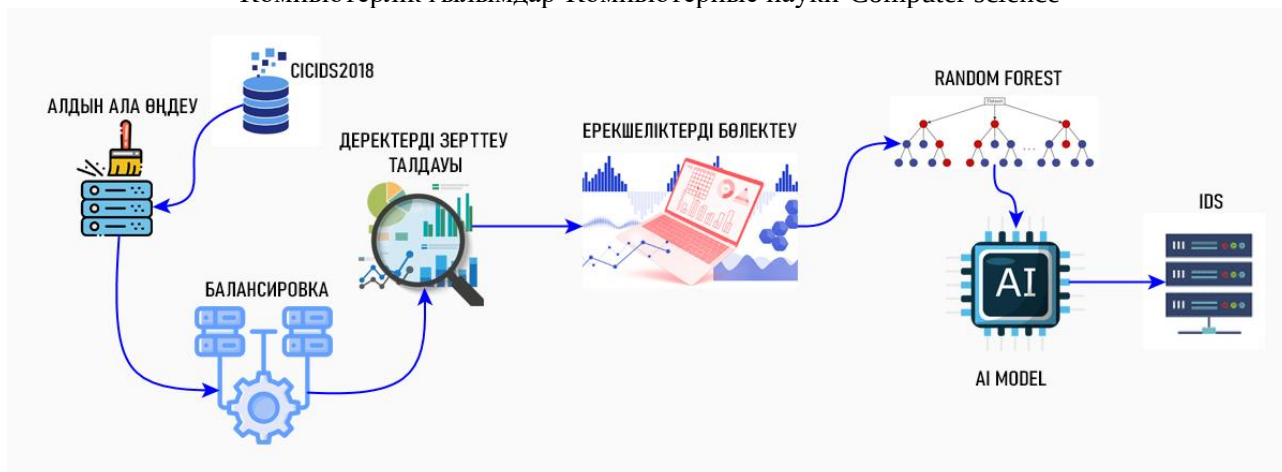
Деректер жиынтығын машиналық оқыту алгоритмдерінде қолданбас бұрын алдын ала өңдеу кезеңі жүргізілді. Бірінші кезеңде жетіспейтін немесе қате мәндері бар жазбалар жойылды. Кейін жіктеу үшін маңызды ақпарат бермейтін белгілер алынып тасталды.

Келесі қадамда сандық белгілерді бір масштабқа келтіру және машиналық оқыту алгоритмдерінің тұрақтылығын арттыру мақсатында нормализация жүргізілді. Сонымен қатар трафик түрін көрсететін Label категориялық белгісі сандық форматқа түрлендірілді.

Алдын ала өңдеу кезеңі аяқталғаннан кейін деректер оқыту және тестілеу жиынтықтарына бөлінді. Оқыту жиынтығы шабуылдарды анықтау моделін құру үшін пайдаланылды, ал тест жиынтығы әзірленген әдістің тиімділігін бағалау үшін қолданылды.

Нәтижелер және оларды талқылау

Дайындалған деректер негізінде желілік сессияларды көпкласстық классификациялау үшін жасанды нейрондық желі (ANN) моделі әзірленіп, оқытылды (1-сурет).



Сурет 1. CICIDS2018 деректер жинағын пайдаланып, басып кіруді анықтау жүйесі (IDS) үшін жасанды интеллект моделін даярлау және оқыту

Өртүрлі архитектуралармен жүргізілген эксперименттерден кейін келесі модель конфигурациясы тандалды:

- Кіріс қабаты: өлшемі 40 (қолданылатын белгілер санына сәйкес).
- Бірінші жасырын қабат: толық байланысқан (Dense) қабат, 64 нейрон, ReLU активациясымен. Одан кейін Dropout қабаты - 0.5 ықтималдықпен, яғни оқыту кезінде нейрондардың жартысы өшіріледі (регуляризация үшін).
- Екінші жасырын қабат: Dense, 32 нейрон, ReLU, одан кейін тағы Dropout (0.5).
- Шығыс қабаты: Dense, 9 нейрон, бұл 9 класқа (8 шабуыл + норм) сәйкес келеді, softmax активациясымен.

Осылайша, желі орташа өлшемдегі 2 жасырын қабаттан тұрады. Тереңірек желілер (3 және одан көп қабаттар) біздің деректерде айтарлықтай жақсарту бермеді, бірақ оқыту уақытын едәуір арттырды және артық үйренуге (overfitting) бейім болды. Екі қабат модель үшін жеткілікті болып шықты, бұл тапсырманың салыстырмалы түрде «сызықтық» табиғатын көрсетеді (көптеген кластар қарапайым шешуші беттермен ажыратылуы мүмкін). Dropout 0.5 – біршама агрессивті мән, бірақ іс жүзінде жалпылауды айтарлықтай жақсартты, сондықтан жоғары күйде қалдырылды.

Модельдің параметрлер саны: Бірінші қабатта: $40 \times 64 + 64 = 2624$ салмақ, екінші қабатта: $64 \times 32 + 32 = 2080$, шығыс қабатында: $32 \times 9 + 9 = 297$. Барлығы шамамен 5000 параметр – бұл нейрожелі үшін өте аз, және бұл жоғары жылдамдық пен деректерге төмен талап тұрғысынан тиімді (артық үйрену қаупі де төмен). Бұл шағын көлем – саналы таңдау: егер желі жүздеген мың параметрден тұрса, ол миллион оқыту мысалында артық үйреніп кетуі мүмкін еді, сонымен қатар жұмыс кезінде көп есептеу ресурстарын талап етер еді. Біздің модель ықшам және жылдам орындалады: веб-интерфейсте бір сессияны классификациялау кідірісі <1 мс.

Оқыту гиперпараметрлері:

– *Оптимизатор:* Adam (Adaptive Moment Estimation) – нейрожелілер үшін ең жақсы әмбебап оптимизаторлардың бірі. Әдепкі параметрлер қолданылды: бастапқы $learning\ rate = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$.

– *Шығын функциясы:* Categorical Crossentropy (көпклассты меткалар үшін кросс-энтропия), бұған дейін сипатталғандай..

– *Батч өлшемі:* 256. Эксперименттер көрсеткендей, 256 кезінде тұрақты және жылдам жақындау байқалды. Төменірек (64) мәнде оқыту көбірек шуылдады, жоғарырақ (512) кезінде – сәл баяу жақындады. 256 – оңтайлы болып шықты, сонымен қатар GPU-ны жақсы жүктеді.

– *Эпохалар саны:* валидациялық таңдамадағы қателік мониторингіне негізделе отырып Early Stopping қолданылды: егер валидациялық шығын 5 эпох қатарынан азаймаса, оқыту тоқтатылды. Модель ең жақсы күйіне шамамен 25 эпоха ішінде жетті. ~20-эпохадан кейін

валидациялық дәлдік ~ 0.967 деңгейінде тұрақтанды және ары қарай тербеліп тұрды, ал 26-шы эпохада нашарлай бастады – Early Stopping іске қосылып, ең аз шығын кезіндегі салмақтар қайтарылды.

– *Класс салмақтары*: бұрын айтылғандай, сирек кластарға шығын функциясында үлкен салмақ берілді. Нақтырақ: «Heartbleed» және «Infiltration» – салмақ ~ 50 (олар өте аз), «Bot» – 5, «WebAttack» – 2, қалғандары – шамамен 1. Бұл Keras ішіндегі `class_weight` параметрінде ескерілді. Бұл болмаса, Heartbleed және Infiltration бойынша recall валидацияда нөлге тең еді. Салмақтарды енгізгеннен кейін модель оларды аз да болса тани бастады.

Оқыту процесі: Оқыту Nvidia GPU-де (GTX 1080) жүргізілді – бір эпоха шамамен 700k мысалмен ~ 5 секундқа созылды (желі шағын болғандықтан – жеткілікті жылдам). 25 эпоха ~ 2 минутты алды, оған валидация да кіреді. CPU-де (i7) бұл шамамен 1 минут/эпоха (жалпы 25 минут), бұл да қолайлы.

Эпохалар бойынша шығындар мен дәлдік:

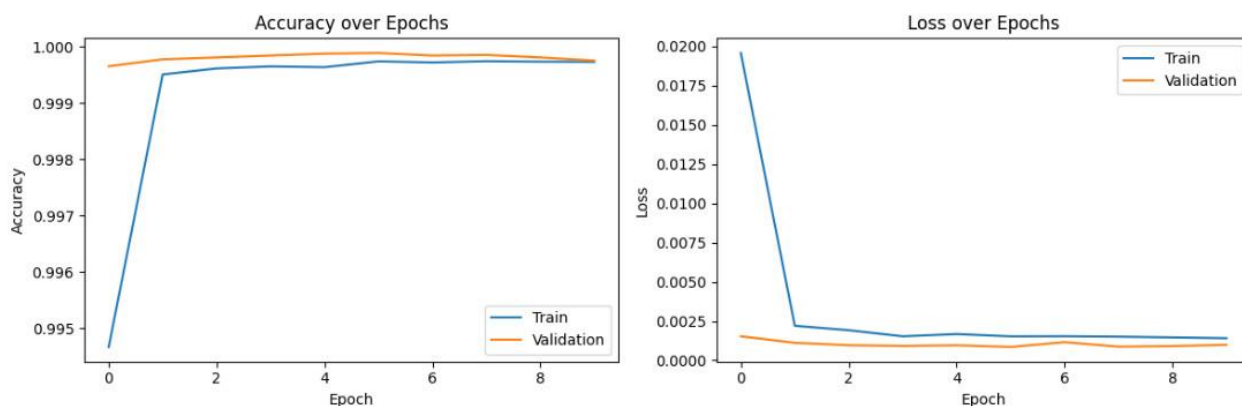
–1-эпоха: train loss ~ 0.55 , val loss ~ 0.43 , val accuracy ~ 0.88 (желі бірдеңе үйрене бастады)

–10-эпоха: train loss ~ 0.17 , val loss ~ 0.19 , val accuracy ~ 0.945

–20-эпоха: train loss ~ 0.13 , val loss ~ 0.13 , val accuracy ~ 0.967

–25-эпоха: train loss ~ 0.11 , val loss ~ 0.135 , val accuracy ~ 0.964 (кішкене артық үйрену белгілері байқалады – val loss минимумы ~ 0.13 -тен кейін аздап өсе бастады)

Ерте тоқтату (Early Stopping) модель салмақтарын 20–22 эпоха аралығында бекітті - бұл кезде валидациялық шығын минималды, ал дәлдік максималды болды (2-сурет).

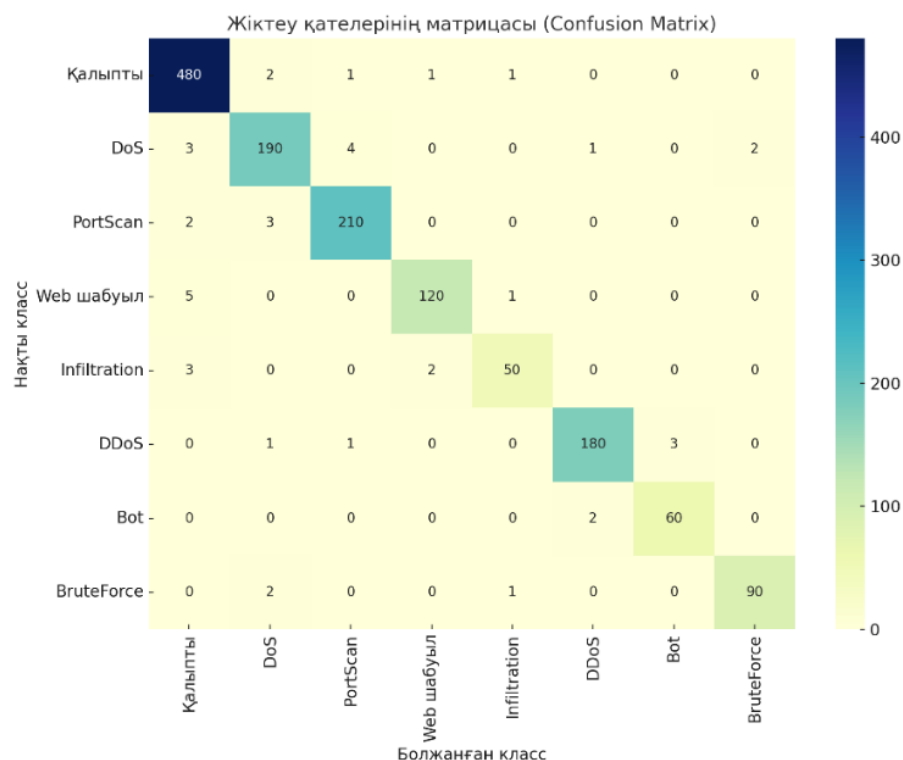


Сурет 2. Модельдің сапасын бағалау

Жалпылама метрикалар: F1-score-ның кластары бойынша орташа мәні (macro-F1) шамамен 0.93 құрайды, бұл барлық категориялар бойынша precision мен recall-дың жақсы теңгерілгенін көрсетеді. Ағындар санына қарай өлшенген орташа мән (weighted-F1) одан да жоғары - шамамен 0.96, себебі ең көп таралған кластар (қалыпты трафик, PortScan, DoS) өте дәл анықталады. Осылайша, тестілік деректерде модель шамамен 96.5% дәлдікке жетеді, ал шабуылдарды анықтау толықтығы (жалпы) да 95-97% аралығында. Салыстыру үшін, сигнатуралық жүйе Suricata сол деректер жиынында (ашық ережелерді қолданып) шамамен 90% жиынтық дәлдік көрсетті - кейбір күрделі шабуылдар мүлде танылмады, ал біздің модель олардың елеулі бөлігін анықтай алды.

Маңыздысын атап өткен жөн - ANN моделі барлық негізгі шабуыл түрлерін, соның ішінде сирек кездесетіндерді де сәтті анықтайды: мысалы, ол «Infiltration» (жасырын ену) және Heartbleed эксплойтын да ұстай алды, ал бұл шабуылдар көбінесе классикалық IDS үшін қиындық тудырады. Нейрожелінің жалған позитивті срабатываниялары да аз: барлық кластар үшін $\text{precision} \geq 0.88$, бұл срабатываниялардың басым бөлігі – нақты шабуылдар екенін білдіреді. Нәтижелер график түрінде қателік матрицасында 3-суретте көрсетілген, онда салмақтың негізгі бөлігі диагональ бойынша орналасқан, ал негізгі шатасулар шабуылдардың іргелес кластары арасында жүреді (мысалы, кейбір Brute Force сессиялары қателесіп веб-

шабуыл ретінде классификациялануы мүмкін және керісінше - бұл қисынды, өйткені екеуі де кіру әрекеттерінің сериясы арқылы көрінеді).



Сурет 3. Қателік матрицасы

Көрсетілген архитектурасы бар нейрожелі (2 жасырын қабат, ReLU, Dropout) желілік трафик ағындарын шабуыл түрлері бойынша классификациялауда жоғары сапа көрсетті. Регуляризация әдістерін (Dropout, L2) қолдану оқыту және тест деректері бойынша дәлдіктің бірдей жоғары деңгейіне жетуге мүмкіндік берді, яғни артық үйренуден сақтап қалды. Оқыту процесі жылдам өтті (бірнеше минут шамасында) - бұл модельдің ықшамдығымен және Adam оптимизаторын қолданумен байланысты. Қорытынды модель Keras .h5 форматында, белгілерді нормализациялау шкалаларымен бірге, келесіде жүйеге енгізу үшін дискке сақталды.

ANN моделін оқытқаннан кейін мұқият сынақтан өткізілді: біріншіден, оқытуда қолданылмаған кейінге қалдырылған деректер жиынында, екіншіден, нақты шабуылдарды еліктейтін жасанды жасалған сценарийлерде. Бұл сынақтар модельдің жұмысқа жарамдылығына көз жеткізу және оны сигнатуралық IDS-пен салыстырғанда тиімділігін бағалау мақсатында жүргізілді.

Жоғарыда келтірілгендей, датасеттің ресми тестілік ішкі жиынында модель шамамен 96.5% дәлдікке (Assurance) жетеді. Ешбір шабуыл класы назардан тыс қалмады: тіпті ең сирек түрлері де анықталған шабуылдардың ішінде бар (Recall > 0.75). Бұл модельдің шабуыл заңдылықтарын жалпылап үйренгеніне, тек нақты үлгілерді «жаттап алмағанына» сенімділік береді. Мысалы, модель PortScan-ның жаңа сессияларын 99% дәлдікпен анықтады, дегенмен IP-адресер мен сканерлеу реттілігі бұрынғылардан өзгеше болуы мүмкін еді – модель сканерлеу трафигінің сипатын таныды (бірнеше қысқа қосылымдардың әртүрлі порттарға қатар орын алуы). Сондай-ақ, «Infiltration» мысалы айтарлықтай маңызды: шамамен 36 ағынның ішінен модель 29-ын дұрыс классификациялады, 7-сін өткізіп алды – ал Suricata негізгі ережелермен ешбірін анықтамады, өйткені «Infiltration» - бұл файл арқылы жүзеге асқан күрделі әлеуметтік инженерлік шабуыл, және оның қарапайым желілік сигнатурасы жоқ.

Модельді пайдалануға жақын жағдайда бағалау үшін, тестілік полигон іске қосылды - бұл оқшауланған виртуалды орта, онда желілік трафик жасанды түрде генерацияланды, оның

ішінде қалыпты сессиялар да, бақыланатын шабуылдар да болды (Kali Linux құралдары мен скрипттер арқылы). Suricata және біздің ANN моделіміз бұл трафикте параллель жұмыс істеді. Мақсат – олардың қорытындыларын салыстыру және модельдің деректер жиыннан тыс та жоғары тиімділігін сақтауын тексеру. Атап айтқанда, тесттер жүргізіліп, 3 кестедегідей нәтижелер алынды.

Кесте 2. Алынған нәтижелер

№	Шабуыл сценарийі	Генерация құралы	Suricata нәтижесі	ANN моделінің нәтижесі	Негізгі бақылаулар
1	Порттарды сканерлеу (Port Scan)	Nmap (slow-scan, random-scan)	Қарқынды сканерлеуді анықтайды, алайда slow-scan кейде төмен қарқындылыққа байланысты танылмайды	PortScan ретінде дұрыс классификациялайды, Recall ≈ 0.98	Модель шабуыл жылдамдығы төмен болса да порттарды жүйелі сканерлеуді анықтайды
2	DDoS шабуылы	hping3 (ICMP flood)	Эвристикалық белгілер жинақталғаннан кейін ескерту береді	DoS/DDoS шабуылын дерлік бірден анықтайды (1–2 сек ішінде)	ANN ағын статистикасын талдау арқылы жылдамырақ әрекет етеді
3	Brute Force (SSH)	Hydra	«Possible SSH Brute Force» ескертуін береді	BruteForce ретінде жоғары сенімділікпен классификациялайды, Recall ≈ 1.00	Шабуылды анықтау нәтижелері толық сәйкес келеді
4	Web Attack (SQL Injection)	SQL injection payload	UNION SELECT сигнатурасы бойынша анықталады	WebAttack ретінде классификациялайды	Екі әдіс те шабуылды сәтті анықтады
5	Web Attack (XSS)	XSS payload	Анықталмады (сигнатура жоқ)	WebAttack ретінде классификациялайды	ANN сигнатуралардан тыс жаңа шабуылды анықтайды

Қорытынды

Сынақ нәтижелері модельдің дәлдік көрсеткіштері деректер жиынында қандай болса, нақты трафикке де солай таралатынын растады. ANN моделі Suricata-ның ағындық деректерін дұрыс өңдейді: есептелетін барлық белгілер оқыту кезіндегі масштабпен сәйкес келеді, ал болжамдар тұрақты түрде күтілетін нәтижелермен үйлеседі (оқыту жиынынан тыс жерде жүйелі ығысу немесе сапаның нашарлауы байқалмады).

Сынақ барысында ANN-нің инференс режиміндегі жұмыс жылдамдығы өлшенді. GPU бар серверде бір классификация ~ 0.5 миллисекунд, ал CPU-де - шамамен 5 мс алды. Бұл - желілік сессиялардың ұзақтығымен салыстырғанда өте аз, сондықтан модель нақты уақыт режимінде жұмыс істей алады, тіпті 1 Гбит/с және одан жоғары жылдамдықтарда, және секундына мыңдаған оқиғаны өңдеуге үлгереді. Suricata параллель жұмыс істей отырып, біршама жүктемеге ұшырайды (өйткені оған әрбір пакетті талдау керек), ал модель ANN тек әр сессия соңында агрегатталған деректерді алады, сондықтан ANN-нің жүктемесі минималды (CPU < 5%).

Кейінге қалдырылған және нақты деректердегі тексеру нәтижелері бойынша, әзірленген ANN моделі талап етілетін дәлдік пен тұрақтылыққа жеткені туралы қорытынды жасауға болады. Ол кейбір шабуылдарды анықтауда дәстүрлі сигнатуралық талдаудан тиімдірек екенін көрсетті (әсіресе жаңа немесе типтік емес шабуылдарда) және бұл ретте жалған дабылдар деңгейі ұқсас болып қалды. Suricata-мен бірге жұмыс істегенде (жүйеде солай жүзеге асырылған), модель қатерлерді толығырақ қамтиды: не бір-бірін іске қосу растайды, не толықтырады – егер бір жүйе шабуылды танымаған болса. Бұл - архитектура мен оқыту әдістерін дұрыс таңдағанымызды растайды.

Әдебиеттер тізімі

1. Kumar B. P., Giri M., Devish R., Dhanush A. V. Mitigating unknown DDoS attacks with RPL: a computationally efficient IDS framework // 2025 4th International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS). Ernakulam, India, – IEEE, 2025. – P. 555-560. URL: <https://ieeexplore.ieee.org/document/11135873>
2. Liu X., Cao Y. Research on network intrusion detection techniques based on feature selection model and recurrent neural network // Proceedings of the 8th International Conference on Cyber Security and Information Engineering (ICCSIE 2023), Putrajaya, Malaysia, 22–24 September 2023. — New York: Association for Computing Machinery, 2023. — P. 45–51. — DOI: [10.1145/3617184.3617781](https://doi.org/10.1145/3617184.3617781).
3. Priya S. et al. Performance analysis comparison on various cyber-attack datasets by relating a deep belief network model on an intrusion detection system (IDS) // Information Technology in Industry. – 2021. – Vol. 9. – №. 3. – P. 608-613. URL: <https://it-in-industry.org/index.php/itii/article/view/600>
4. Mhawi D. N., Aldallal A., Hassan S. Advanced feature-selection-based hybrid ensemble learning algorithms for network intrusion detection systems // Symmetry. – 2022. – Vol. 14. – №. 7. – Art. 1461. DOI: [10.3390/sym14071461](https://doi.org/10.3390/sym14071461).
5. Emanet S., Baydogmus G. K., Demir O. An ensemble learning based IDS using Voting rule: VEL-IDS // PeerJ Computer Science. – 2023. – Vol. 9. – P. e1553. DOI: [10.7717/peerj-cs.1553](https://doi.org/10.7717/peerj-cs.1553).
6. Markovic T., Leon M., Buffoni D., Punnekkat S. Random forest with differential privacy in federated learning framework for network attack detection and classification // Applied Intelligence. – 2024. – Vol. 54. – №. 17. – P. 8132-8153. DOI: [10.1007/s10489-024-05589-6](https://doi.org/10.1007/s10489-024-05589-6).
7. Mukhef H. A., Ali H. H., Ibrahim Z. S. Adversarial robustness of network IDS (structured data) // Journal of Al-Qadisiyah for Computer Science and Mathematics. – 2025. – Vol. 17. – №. 4. – P. 209–224. DOI: [10.29304/jqcs.2025.17.42554](https://doi.org/10.29304/jqcs.2025.17.42554)
8. Kumar A., Kumar V., Singh Bhadauria A. P. Optimizing intrusion detection in edge computing network: a hybrid ML approach with recursive feature elimination // International Journal of Intelligent Engineering and Systems. – 2025. – Vol. 18. – №. 1. – P. 36–46. DOI: [10.22266/ijies2025.0229.04](https://doi.org/10.22266/ijies2025.0229.04).
9. Jadhav A. D., Pellakuri V., Ahire P. G., Chaugule A. A., Tiwari H. U. Distributed two phase intrusion detection system using machine learning techniques and underlying big data storage and processing architecture-HDFS // Inteligencia Artificial. – 2025. – Vol. 28. – №. 76. – P. 124-148. DOI: [10.4114/intartif.vol28iss76pp124-148](https://doi.org/10.4114/intartif.vol28iss76pp124-148).
10. Zhang S., Wang J., An D. Diffusion multimodal distillation collaboration: a generative equilibrium framework for efficient vehicle networking intrusion detection // IEEE Internet of Things Journal. – 2026. – Vol. 13, №. 9. – P. 19491–19505. DOI: [10.1109/JIOT.2026.3664119](https://doi.org/10.1109/JIOT.2026.3664119).
11. Budilaksono S. et al. Comparison of data mining algorithm: PSO-KNN, PSO-RF, and PSO-DT to measure attack detection accuracy levels on intrusion detection system // Journal of Physics: Conference Series. – IOP Publishing, 2020. – Vol. 1471. – №. 1. – Art. 012019. DOI: [10.1088/1742-6596/1471/1/012019](https://doi.org/10.1088/1742-6596/1471/1/012019).
12. Salih A. A., Abdulazeez A. M. Evaluation of classification algorithms for intrusion detection system: A review // Journal of Soft Computing and Data Mining. – 2021. – Vol. 2. – №. 1. – P. 31-40. DOI: [10.30880/jsdm.2021.02.01.004](https://doi.org/10.30880/jsdm.2021.02.01.004).
13. Tahri R. et al. A comparative study of machine learning algorithms on the UNSW-NB15 dataset // ITM Web of Conferences. – EDP Sciences, 2022. – Vol. 48. – Art. 03002. DOI: [10.1051/itmconf/20224803002](https://doi.org/10.1051/itmconf/20224803002).

References

1. Kumar B. P., Giri M., Devish R., Dhanush A. V. Mitigating unknown DDoS attacks with RPL: a computationally efficient IDS framework // 2025 4th International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS). Ernakulam, India, – IEEE, 2025. – P. 555-560. URL: <https://ieeexplore.ieee.org/document/11135873>

2. Liu X., Cao Y. Research on network intrusion detection techniques based on feature selection model and recurrent neural network // Proceedings of the 8th International Conference on Cyber Security and Information Engineering (ICCSIE 2023), Putrajaya, Malaysia, 22–24 September 2023. — New York: Association for Computing Machinery, 2023. — P. 45–51. — DOI: [10.1145/3617184.3617781](https://doi.org/10.1145/3617184.3617781).
3. Priya S. et al. Performance analysis comparison on various cyber-attack datasets by relating a deep belief network model on an intrusion detection system (IDS) // Information Technology in Industry. — 2021. — Vol. 9. — №. 3. — P. 608-613. URL: <https://it-in-industry.org/index.php/itii/article/view/600>
4. Mhawi D. N., Aldallal A., Hassan S. Advanced feature-selection-based hybrid ensemble learning algorithms for network intrusion detection systems // Symmetry. — 2022. — Vol. 14. — №. 7. — Art. 1461. DOI: [10.3390/sym14071461](https://doi.org/10.3390/sym14071461).
5. Emanet S., Baydogmus G. K., Demir O. An ensemble learning based IDS using Voting rule: VEL-IDS // PeerJ Computer Science. — 2023. — Vol. 9. — P. e1553. DOI: [10.7717/peerj-cs.1553](https://doi.org/10.7717/peerj-cs.1553).
6. Markovic T., Leon M., Buffoni D., Punnekkat S. Random forest with differential privacy in federated learning framework for network attack detection and classification // Applied Intelligence. — 2024. — Vol. 54. — №. 17. — P. 8132-8153. DOI: [10.1007/s10489-024-05589-6](https://doi.org/10.1007/s10489-024-05589-6).
7. Mukhef H. A., Ali H. H., Ibrahim Z. S. Adversarial robustness of network IDS (structured data) // Journal of Al-Qadisiyah for Computer Science and Mathematics. — 2025. — Vol. 17. — №. 4. — P. 209–224. DOI: [10.29304/jqcs.2025.17.42554](https://doi.org/10.29304/jqcs.2025.17.42554)
8. Kumar A., Kumar V., Singh Bhadauria A. P. Optimizing intrusion detection in edge computing network: a hybrid ML approach with recursive feature elimination // International Journal of Intelligent Engineering and Systems. — 2025. — Vol. 18. — №. 1. — P. 36–46. DOI: [10.22266/ijies2025.0229.04](https://doi.org/10.22266/ijies2025.0229.04).
9. Jadhav A. D., Pellakuri V., Ahire P. G., Chaugule A. A., Tiwari H. U. Distributed two phase intrusion detection system using machine learning techniques and underlying big data storage and processing architecture-HDFS // Inteligencia Artificial. — 2025. — Vol. 28. — №. 76. — P. 124-148. DOI: [10.4114/intartif.vol28iss76pp124-148](https://doi.org/10.4114/intartif.vol28iss76pp124-148).
10. Zhang S., Wang J., An D. Diffusion multimodal distillation collaboration: a generative equilibrium framework for efficient vehicle networking intrusion detection // IEEE Internet of Things Journal. — 2026. — Vol. 13, №. 9. — P. 19491–19505. DOI: [10.1109/JIOT.2026.3664119](https://doi.org/10.1109/JIOT.2026.3664119).
11. Budilaksono S. et al. Comparison of data mining algorithm: PSO-KNN, PSO-RF, and PSO-DT to measure attack detection accuracy levels on intrusion detection system // Journal of Physics: Conference Series. — IOP Publishing, 2020. — Vol. 1471. — №. 1. — Art. 012019. DOI: [10.1088/1742-6596/1471/1/012019](https://doi.org/10.1088/1742-6596/1471/1/012019).
12. Salih A. A., Abdulazeez A. M. Evaluation of classification algorithms for intrusion detection system: A review // Journal of Soft Computing and Data Mining. — 2021. — Vol. 2. — №. 1. — P. 31-40. DOI: [10.30880/jscdm.2021.02.01.004](https://doi.org/10.30880/jscdm.2021.02.01.004).
13. Tahri R. et al. A comparative study of machine learning algorithms on the UNSW-NB15 dataset // ITM Web of Conferences. — EDP Sciences, 2022. — Vol. 48. — Art. 03002. DOI: [10.1051/itmconf/20224803002](https://doi.org/10.1051/itmconf/20224803002).

ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ОБНАРУЖЕНИЯ СЕТЕВЫХ АТАК

КЕРЕЕВ А.К.^{1*}, ТӘҢІРБЕРГЕН А.К.¹, АБДЫГАЛИЕВА Г.Б.²,
ИГИСИНОВ С.Ж.³ ТОЛЕУОВ Т.Ж.¹

*Кереев Адилжан Кутымович¹ — PhD, доцент кафедры информатики и информационных технологий, Актюбинский региональный университет имени К. Жубанова, г. Ақтөбе, Казахстан

E-mail: adiljan@mail.ru, <https://orcid.org/0000-0002-8283-5807>

Тәңірберген Айсұлу Көбейсінкызы¹ — Магистр, преподаватель кафедры информатики и ИТ, Актюбинский региональный университет имени К. Жубанова, г. Ақтөбе, Казахстан

E-mail: aisulu21@mail.ru, <https://orcid.org/0009-0009-2936-4637>

Абдығалиева Гульдариya Болатбековна² – Магистр, учитель информатики, педагог-мастер, Школа-гимназия №41 имени Фаризы Онгарсыновой, г. Ақтөбе, Казахстан

E-mail: guldariya78@mail.ru, <https://orcid.org/0009-0001-0552-8475>

Игисинов Сабит Жандарбекович³ – PhD, и.о. ассоциированного профессора кафедры математики, Таразский университет имени М.Х. Дулати, г. Тараз, Казахстан.

E-mail: igisinovsabit@gmail.com, <https://orcid.org/0000-0002-4189-9142>

Толуев Тимур Жаксылыкович¹ – Магистр, старший преподаватель кафедры информатики и ИТ, Актюбинский региональный университет имени К. Жубанова, г. Ақтөбе, Казахстан

E-mail: timur_toleuov@mail.ru, <https://orcid.org/0000-0002-6050-1056>

Аннотация. В настоящее время рост количества и сложности кибератак делает обеспечение безопасности компьютерных сетей одной из наиболее актуальных задач. Традиционные сигнатурные методы эффективны при выявлении известных атак, однако их возможности ограничены при обнаружении новых или модифицированных типов атак. В связи с этим возникает необходимость применения интеллектуальных методов анализа сетевого трафика. Целью данного исследования является разработка и оценка эффективной модели обнаружения сетевых атак с использованием методов машинного обучения. В качестве экспериментальной базы использован набор данных CICIDS2018, содержащий различные типы кибератак и нормальный сетевой трафик. В ходе исследования были применены методы предварительной обработки данных, нормализации признаков и построения модели искусственной нейронной сети для многоклассовой классификации. Предложенная модель включает два скрытых слоя и обучается с использованием регуляризации Dropout и оптимизатора Adam. Экспериментальные результаты показали, что модель достигает точности около 96,5% и эффективно обнаруживает различные типы атак. Полученные результаты подтверждают перспективность применения методов машинного обучения для совершенствования систем обнаружения вторжений и развития современных технологий кибербезопасности.

Ключевые слова: кибербезопасность, система обнаружения вторжений, машинное обучение, нейронные сети, сетевой трафик, CICIDS2018.

APPLICATION OF NEURAL NETWORKS FOR NETWORK ATTACK DETECTION

KEREEV A.K.^{1*}, TANGIRBERGEN A.K.¹, ABDYGALIEVA G.B.²,
IGISINOV S.ZH.³, TOLEUOV T.ZH.¹

***Kereev Adilzhan Kutymovich**¹ – PhD, associate professor, department of informatics and information technologies, K. Zhubanov Aktobe regional university, Aktobe, Kazakhstan.

E-mail: adiljan@mail.ru, <https://orcid.org/0000-0002-8283-5807>

Tangirbergen Aisulu Kobeyinsky¹ – Master's degree, lecturer, department of informatics and information technologies, K. Zhubanov Aktobe regional university, Aktobe, Kazakhstan.

E-mail: aisulu21@mail.ru, <https://orcid.org/0009-0009-2936-4637>

Abdygalieva Guldariya Bolatbekovna² – Master's degree, informatics teacher, teacher-master, №. 41 School-gymnasium named after Fariza Ongarsynova, Aktobe, Kazakhstan.

E-mail: guldariya78@mail.ru, <https://orcid.org/0009-0001-0552-8475>

Iginov Sabit Zhandarbekovich³ – PhD, department of mathematics, acting associate professor, M.Kh. Dulaty Taraz university, Taraz, Kazakhstan.

E-mail: igisinovsabit@gmail.com, <https://orcid.org/0000-0002-4189-9142>

Toleuov Timur Zhaksilykovich¹ – Master's degree, senior lecturer, department of informatics and information technologies, K. Zhubanov Aktobe regional university, Aktobe, Kazakhstan.

E-mail: timur_toleuov@mail.ru, <https://orcid.org/0000-0002-6050-1056>

Abstract. Currently, the increasing number and complexity of cyberattacks make ensuring the security of computer networks one of the most critical challenges in modern information systems. Traditional signature-based methods are effective in detecting known attacks; however, their capabilities are limited when identifying new or modified attack types. Therefore, the use of intelligent methods for network traffic analysis becomes necessary. The aim of this study is to develop and evaluate an effective model for network attack detection using machine learning techniques. The CICIDS2018 dataset, which includes various types of cyberattacks and normal network traffic, was used as the experimental basis. The research involved data preprocessing, feature normalization, and the construction of an artificial neural network model for multiclass classification. The proposed model consists of two hidden layers and is trained using Dropout regularization and the Adam optimizer. Experimental results show that the model achieves an accuracy of approximately 96.5% and effectively detects different types of attacks. The findings demonstrate the potential of machine learning methods for improving intrusion detection systems and advancing modern cybersecurity technologies.

Key words: cybersecurity, intrusion detection system, machine learning, neural networks, network traffic, CICIDS2018.