ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫ

әл-Фараби атындағы Қазақ ұлттық университетінің

ХАБАРЛАРЫ

ИЗВЕСТИЯ

НАЦИОНАЛЬНОЙ АКАДЕМИИ
НАУК РЕСПУБЛИКИ КАЗАХСТАН
Казахский национальный
университет имени аль-Фараби

N E W S

OF THE ACADEMY OF SCIENCES
OF THE REPUBLIC OF
KAZAKHSTAN
al-Farabi Kazakh National University

SERIES
PHYSICO-MATEMATICAL

3 (343)

JULY – SEPTEMBER 2022

PUBLISHED SINCE JANUARY 1963

PUBLISHED 4 TIMES A YEAR

ALMATY, NAS RK

«ҚР ҰҒА Хабарлары. Физика-математикалық сериясы».
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

Меншіктеуші: «Қазақстан Республикасының Ұлттық ғылым академиясы» РҚБ (Алматы қ.). Қазақстан Республикасының Ақпарат және қоғамдық даму министрлігінің Ақпарат комитетінде 14.02.2018 ж. берілген **No 16906-Ж** мерзімдік басылым тіркеуіне қойылу туралы куәлік.
Тақырыптық бағыты: *ақпараттық коммуникациялық технологиялар сериясы.*
Қазіргі уақытта: *«ақпараттық технологиялар»бағыты бойынша ҚР БҒМ БҒСБК ұсынған журналдар тізіміне енді.*
Мерзімділігі: *жылына 4 рет.*
Тиражы: *300 дана.*
Редакцияның мекен-жайы: *050010, Алматы қ., Шевченко көш., 28, 219 бөл., тел.: 272-13-19*
http://www.physico-mathematical.kz/index.php/en/

Типографияның мекен-жайы: «Аруна» ЖК, Алматы қ., Мұратбаев көш., 75.

«Известия НАН РК. Серия физика-математическая».
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

Собственник: *Республиканское общественное объединение «Национальная академия наук Республики Казахстан» (г. Алматы).*
Свидетельство о постановке на учет периодического печатного издания в Комитете информации Министерства информации и общественного развития Республики Казахстан **No 16906-Ж** выданное 14.02.2018 г.
Тематическая направленность: *серия информационные коммуникационные технологии.*
В настоящее время: *вошел в список журналов, рекомендованных ККСОН МОН РК по направлению «информационные коммуникационные технологии».*
Периодичность: *4 раз в год.*
Тираж: *300 экземпляров.*
Адрес редакции: *050010, г. Алматы, ул. Шевченко, 28, оф. 219, тел.: 272-13-19*
http://www.physico-mathematical.kz/index.php/en/

Адрес типографии: ИП «Аруна», г. Алматы, ул. Муратбаева, 75.

News of the National Academy of Sciences of the Republic of Kazakhstan.
Physical-mathematical series.
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

Owner: RPA «National Academy of Sciences of the Republic of Kazakhstan» (Almaty). The certificate of registration of a periodical printed publication in the Committee of information of the Ministry of Information and Social Development of the Republic of Kazakhstan **No. 16906-Ж,** issued 14.02.2018
Thematic scope: *series information technology.*
Currently: *included in the list of journals recommended by the CCSES MES RK in the direction of «information and communication technologies».*
Periodicity: *4 times a year.*
Circulation: *300 copies.*
Editorial address: *28, Shevchenko str., of. 219, Almaty, 050010, tel. 272-13-19*
http://www.physico-mathematical.kz/index.php/en/

Address of printing house: ST «Aruna», 75, Muratbayev str, Almaty.

NEWS OF THE NATIONAL ACADEMY OF SCIENCES
OF THE REPUBLIC OF KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X
Volume 3, Number 343 (2022), 52-70
https://doi.org/10.32014/2022.2518-1726.139
УДК 372.851.02., 372.800.4.02

М.А. Болатбек[1], К.Б. Багитова[2*], Ш.Ж. Мусиралиева[2]

[1]Пассау университеті, Германия, Пассау;
[2]Әл-Фараби атындағы Қазақ ұлттық университеті, Қазақстан, Алматы.
E-mail: *kbbagitova@gmail.com*

КИБЕРҚАУІПСІЗДІК МӘСЕЛЕЛЕРІН ТАБИҒИ ТІЛДІ ӨҢДЕУ ӘДІСТЕРІ АРҚЫЛЫ ШЕШУ ТАҚЫРЫБЫНА ЖҮЙЕЛІК ШОЛУ

Аннотация. Қазіргі таңда ақпараттық-коммуникациялық Интернет желісі адамзат өмірінің ажырамас бөлігіне айналды. Адамдардың жұмысы, жеке өмірі және қаржысы Интернет, мобильді компьютерлер мен электрондық бұқаралық ақпарат құралдары әлеміне бет бұра бастады. Адамдар «Твиттер», «ВКонтакте», «Facebook» және т.с.с. әлеуметтік желілерді жаһандық байланыс орнату, пікір алмасу, білім алу және т.б. мақсаттарда пайдалануда. Өкінішке қарай, бұл кең таралған құбылыс бізді зиянды шабуылдарға, жеке өмірге қол сұғуға, алаяқтыққа және басқа да осындай қиыншылықтарды бұрынғыдан да көбейтеді.

Жеке пайдаланушылардың ғана емес, сонымен қатар ақпараттық ұйымдардың да бүкіл әлемдік кеңістікке белсенді қатысуы ұлттық қауіпсіздікті қамтамасыз ету бойынша ақпараттық-коммуникациялық технологиялар дамуының қазіргі тенденцияларына сәйкес келетін іс-шараларды ұйымдастыру қажеттілігін анықтайды. Сондықтан киберқауіпсіздік қауіпсіз және реттелген сандық әлемнің маңызды бөлігі болып табылады.

Google, Facebook және Twitter алпауыттары интернеттегі террористік мазмұнды жылдам анықтап, жою үшін жасанды интеллект (AI) технологиясын қолдануға уәде берді. IBM-де жоғарыда аталған әлеуметтік желілердегі барлық деректерді талдай алатын Watson әзірлемесі бар. Ресейде Платонның IT-авторы әлеуметтік желілерді

бақылау және қауіп-қатерлерді болжау жүйесін құруда. Германия үкіметі террористік актілерден кейін Интернеттегі террористермен күресу үшін ZITiS атты жаңа киберқауіпсіздік бөлімшесі құрылғанын жариялады. Мұндай жүйелер әзірге Қазақстанда жоқ. Осы себепті Интернеттегі веб-ресурстарға талдау ақпараттық қауіпсіздікті қамтамасыз етуші ұйымдар үшін аса өзекті болып табылады.

Түйін сөздер: киберқауіпсіздік, табиғи тілді өңдеу, әлеуметтік желі, интернет, қауіпсіздік, машиналық оқыту, терең оқыту, мәтінді жіктеу.

М.А. Болатбек[1], К.Б. Багитова[2*], Ш.Ж. Мусиралиева[2]

[1] Университет Пассау, Германия, Пассау;
[2]Казахский национальный университет им. аль-Фараби,
Казахстан, Алматы.
E-mail: *kbbagitova@gmail.com*

СИСТЕМАТИЧЕСКИЙ ОБЗОР ТЕМЫ РЕШЕНИЯ ЗАДАЧ КИБЕРБЕЗОПАСНОСТИ С ПОМОЩЬЮ МЕТОДОВ ОБРАБОТКИ ЕСТЕСТВЕННОГО ЯЗЫКА

Аннотация. В настоящее время информационно-коммуникационная сеть Интернет стала неотъемлемой частью жизни человечества. Работа, личная жизнь и финансы людей стали уходить в мир Интернета, мобильных компьютеров и электронных средств массовой информации. Люди используют социальные сети в «Твиттер», «ВКонтакте», «Facebook» и т.д. для глобального общения, обмена мнениями, получения знаний и т.д. К сожалению, это распространенное явление делает нас еще более уязвимыми для вредоносных атак, вторжений в личную жизнь, мошенничества и других подобных неприятностей.

Активное участие не только отдельных пользователей, но и информационных организаций во всем мировом пространстве определяет необходимость организации мероприятий по обеспечению национальной безопасности, соответствующих современным тенденциям развития информационно-коммуникационных технологий. Поэтому кибербезопасность является важной частью безопасного и упорядоченного цифрового мира.

Гиганты Google, Facebook и Twitter пообещали использовать технологию искусственного интеллекта (ИИ) для быстрого обнаружения и

уничтожения террористического контента в Интернете. У IBM есть разработка Watson, которая может анализировать все данные из вышеупомянутых социальных сетей. В России IT-автор Платона создает систему мониторинга социальных сетей и прогнозирования рисков. Правительство Германии объявило о создании нового подразделения кибербезопасности под названием ZITiS для борьбы с террористами в Интернете после террористических актов. Таких систем в Казахстане пока нет. По этой причине анализ веб-ресурсов в Интернете является наиболее актуальным для организаций, обеспечивающих информационную безопасность.

Ключевые слова: кибербезопасность, обработка естественного языка, социальная сеть, Интернет, безопасность, машинное обучение, глубокое обучение, классификация текста.

M. Bolatbek[1], K. Bagitova[2*], Sh. Musiralieva[2]

[1]University of Passau, Germany, Passau;
[2]Al-FarabiKazakh National University, Kazakhstan, Almaty.
E-mail: *kbbagitova@gmail.com*

A SYSTEMATIC REVIEW ON CYBERSECURITY ISSUES USING NATURAL LANGUAGE PROCESSING TECHNIQUES

Abstract. Currently, the Internet information and communication network has become an integral part of human life. Work, personal life and finances of people began to go into the world of the Internet, mobile computers and electronic media. People use social networks Twitter, VKontakte, Facebook, etc. for global communication, exchange of opinions, gaining knowledge, etc. Unfortunately, this common phenomenon makes us even more vulnerable to malicious attacks, invasions of privacy, fraud and other similar troubles.

The active participation not only of individual users, but also of information organizations throughout the world space determines the need to organize measures to ensure national security that correspond to modern trends in the development of information and communication technologies. Therefore, cybersecurity is an important part of a secure and orderly digital world.

The giants Google, Facebook and Twitter have promised to use artificial intelligence (AI) technology to quickly detect and destroy terrorist content

on the Internet. IBM has a Watson development that can analyze all the data from the aforementioned social networks. In Russia, Plato's IT author creates a system for monitoring social networks and forecasting risks. The German government has announced the creation of a new cybersecurity unit called ZITiS to combat terrorists on the Internet after terrorist attacks. There are no such systems in Kazakhstan yet. For this reason, the analysis of web resources on the Internet is the most relevant for organizations providing information security.

Key words: cybersecurity, natural language processing, social network, Internet, security, machine learning, deep learning, text classification.

Кіріспе. Компьютерлік қауіпсіздік – есептеу құрылғыларын (компьютерлерді, смартфондарды және басқаларды), сондай-ақ компьютерлік желілерді (Интернетті қоса алғанда, жеке және жалпыға қолжетімді желілерді) қорғау үшін қолданылатын қауіпсіздік шаралары. Жүйелік әкімшілердің қызметі сандық жабдықтар, ақпараттық өріс және қызметтер кездейсоқ немесе рұқсатсыз кіруден, деректерді өзгертуден немесе жоюдан қорғалған және дамыған қауымдастықтағы компьютерлік жүйелерге тәуелділіктің артуына байланысты маңызды болып табылатын барлық процестер мен механизмдерді қамтиды.

Киберқауіпсіздік – деректердің құпиялылығын, тұтастығын және қол жетімділігін қамтамасыз ету үшін қауіпсіздік шараларын қолдану процесі. Жүйелік әкімші компьютерлердің, серверлердің жергілікті желісінің деректерін қоса алғанда, активтерді қорғауды қамтамасыз етеді. Сонымен қатар, ғимараттар мен ең бастысы қызметкерлер тікелей қорғауға алынады. Киберқауіпсіздікті қамтамасыз етудің мақсаты деректерді қорғау болып табылады. Деректердің қауіпсіздігін қамтамасыз ету мақсатында қарсы шаралар да қолданылуы мүмкін. Бұл шаралардың қатарына кіруді бақылау, қызметкерлерді оқыту, аудит және есеп беру, ықтимал тәуекелдерді бағалау, ену тестілері және авторизация талаптары кіреді (бірақ олармен шектелмейді).

Киберқауіпсіздік-бұл электрондық деректер мен ақпаратты қорғау. Бұл компьютерлер, ұялы телефондар, серверлер және желілер сияқты құрылғылардағы электрондық жүйелерді зиянды шабуылдардан қорғау.

Қазір кейбір елдерде киберқауіпсіздікті мектеп қабырғасынан бастап оқыту жоспарлануда. Мысалы, Ұлыбританияда оқушыларға Киберқауіпсіздік бойынша сабақтар ұсынылады, онда олар британдық компаниялар мен ұйымдардың хакерлердің желілік шабуылдарынан қауіпсіздігін қамтамасыз ету дағдыларын үйренеді. Оқушылармен

Киберқауіпсіздіктің нақты мәселелері және оларды шешу тәжірибесі қарастырылады. Бағдарлама 14-18 жас аралығындағы оқушыларға бағытталған.

Қиберқауіпсіздікке қатысты бірқатар деректерге тоқталатын болсақ:

1. Киберқауіпсіздіктің 95 пайызы адамның қателіктерінен болады. (Дүниежүзілік экономикалық форум).

2. Әлемдік ақпараттық қауіпсіздік нарығы 2028 жылы 366,1 миллиард долларға жетеді деп болжануда. (Fortune бизнес-аналитикасы).

3. 2020 жылы АҚШ кибер шабуылдардың 46 пайызын нысанаға алды, бұлкез-келген елге қарағанда ек іесе көп. (Microsoft Корпорациясы).

4. Бизнес жетекшілерінің 68 пайызы Киберқауіпсіздік бойынша тәуекелдер артып келеді деп санайды. (Акцент).

5. Орташаалғанда, Компания қалталарының тек беспайызы дұрыс қорғалған. (Варонис).

6. Деректердің ағып кетуі 22 жылы 2021 миллиард жазбаны анықтады. (Тәуекелге негізделген қауіпсіздік).

7. 2021 жылы бұзушылықтардың шамамен 70 пайызы қаржылық тұрғыдан негізделген, ал бес пайызданазы тыңшылыққа негізделген. (Verizon).

8. 2021 жылы бұзушылықтардың шамамен 40 пайызы фишингке, 11 пайызға жуығы зиянды бағдарламаларға және 22 пайызға жуығы бұзылуларға байланысты болды. (Verizon).

9. 2021 жылы 1862 деректердің ағуы тіркелді, бұл 2017 жылғы рекордтан 1506 ағып кетуден асып түсті. (CNET).

10. Зияндыэлектрондық пошта тіркемелерінің ең көп таралғантүрлері .docжәне -37 пайызы; келесі үлкені - .exe (19,5 пайыз). (Symantec).

Киберқауіпсіздік өте маңызды, себебі ол желідегі мәліметтерді оны ұрлағысы келетін және оны зиян келтіру үшін пайдаланғысы келетін кибершабуылдардан қорғауға қатысты процестерді қамтиды. Олардың қатарында құпия деректер, мемлекеттік және салалық ақпарат, жеке ақпарат, жеке басын анықтауға мүмкіндік беретін ақпарат, зияткерлік меншік және қауіпсіз медициналық ақпарат болуы мүмкін. Аталған деректерді қорғауға арналған алдыңғы қатарлы бағдарламалар мен киберқорғаныс механизмдерінің болуы өте маңызды. Қоғамның әр мүшесі ауруханалар мен басқа да денсаулық сақтау мекемелері, қаржылық қызметтер бағдарламалары және электр станциялары сияқты маңызды инфрақұрылымдардың қолданушысы болып табылады. Мысалы, қолданбаға кірген кезде немесе сандық денсаулық сақтау жүйелерінде құпия деректерді толтырған кезде аталған жүйелер,

желілер мен инфрақұрылымдар тиісті қорғанысқа ие болмаса, ол жердегі деректер теріс мақсатта пайдаланылуы мүмкін.

Жеке деңгейдегі кибершабуылдар адамның жеке басына қажетті ақпаратты ұрлауға және бопсалауға әкелуі мүмкін, бұл адамның өміріне айтарлықтай зиян келтіруі мүмкін.

Киберқауіпсіздіктің маңыздылығын айқындайтын себептерге тоқталып өтетін болсақ, ең алдымен жылдам кең жолақты байланыс, жетілдірілген гаджеттер және бұлтты есептеу сияқты технологияның қарқынды дамуы қосылған құрылғылардың көбеюіне әкелді. Бұл қараңғы Интернетте әр түрлі киберқылмыс түрлерін орындауға керемет мүмкіндік туғызды. Келесі себеп - технологияны пайдаланушылардың осалдығы. Қазіргі уақытта адамдардың көпшілігінің ақпараттық және коммуникациялық технологияларға көбірек сенім артатындығы киберқылмыскерлер үшін қылмыс жасау мүмкіндігінің тез өсіп келе жатқанын білдіреді. Бұлтты сақтаудың кеңеюі және әлеуметтік медианың өсуі сияқты факторлар көптеген адамдарды кибершабуылдарға осал етті. Бұл киберқауіпсіздікті бұрынғыдан да маңызды етеді. Сонымен қатар, банк деректемелері мен құпиясөздер сияқты құпия ақпаратты бұлтта сақтауға болады, бұл оның ұрлану қаупін арттырады. Сондай-ақ,әлеуметтік медианың өсуі де жеке деректердің алаяқтық жағдайда пайдаланылуының көбеюіне әкелді.

Соңғы зерттеулерге сәйкес, өткен жылы ұйым үшін киберқылмыстың орташа құны шамамен 13 миллион долларды құрады. Зерттеулер сонымен қатар қаржылық ақпарат, медициналық жазбалар, коммерциялық құпия, жеке мәліметтер және зияткерлік меншік сияқты ақпараттың күрт өсуін анықтады.

Компьютерлік вирустар өте тез таралуы мүмкін. Егер олар басқарылмаса, бұл бизнес үшін үлкен қиындықтар тудыруы мүмкін. Компьютерлік вирустар файлдар мен жүйелерді зақымдауы мүмкін. Сондықтан киберқауіпсіздікке байыпты қарау өте маңызды, өйткені бұл компьютерлік жүйені вирустардан құтқара алады.

Технологияның өсуі мен дамуы қара торды артта қалдырмады. Қараңғы веб-сайт – бұл тек мамандандырылған веб-шолғыштар арқылы қол жеткізуге болатын интернет-сайттардың құпия ынтымақтастығы. Ол негізінен интернет қызметін жасыру және пайдаланушылардың анонимділігі мен құпиялылығын сақтау үшін қолданылады.Қараңғы вебті заңды түрде қолдануға болады, бірақ ол көптеген заңсыз операциялардың орны ретінде де танымал. Есірткі мен адам саудасы, қаруды заңсыз тарату, бағдарламалық жасақтаманы тарату, заңсыз

аукциондар, қарақшылық және басқа да көптеген заңсыз әрекеттер қара торды қолданумен байланысты екені белгілі.

Киберқауіпсіздік өте маңызды, өйткені ол ұйымдарды ықтимал киберқауіптерден қорғайды. Технологияның дамуы көптеген адамдарды бұзу, деректерді ұрлау және бүлдіру, сондай-ақ өнеркәсіптік тыңшылық сияқты киберқылмыскерлердің әрекеттеріне осал етті. Киберқылмыс деңгейі өсуде, сондықтан киберқауіпсіздік болмаса, құпия ақпаратты, ақшаны немесе беделді жоғалту ықтималдығы артады.

Материалдар мен әдістер. Табиғи тілдегі мәтіндерді өңдеу (Natural Language Processing, NLP) – жасанды интеллект пен математикалық лингвистиканың жалпы бағыты. Ол табиғи тілдердегі мәтіндерді компьютерлік талдау және синтездеу мәселелерін зерттейді.

NLP-ды киберқауіпсіздік есептері үшін қолданудың артықшылықтарына келетін болсақ,

1. Терең оқыту алгоритмдері кез-келген ақпараттық корпустың құрылымдалмаған сипатын ескере отырып, Киберқауіпсіздік бойынша NLP деңгейінің жоғарылауының негізгі қозғаушы күші болып табылады.

2. NLP лексикалық талдауы үшін терең оқытуды қолдану екілік файлдар немесе бастапқы кодтар сияқты стандартты емес тілдердегі жағдайларды талдауға мүмкіндік береді.

3. NLP – бұл анықтау, тергеу және жауап беруді үйлестірудің өте пайдалы құралы, оның талдау және ақпарат алу мүмкіндіктерін ескере отырып тиімді пайдалануға болады.

Компьютерлер мен адам (табиғи) тілдерінің өзара әрекеттесуі тілді өңдеуге негізделгенін ескере отырып, тілдік деректердің үлкен көлемін өңдеу мен талдауда компьютерлердің біліктілігін арттыру маңызды бола түсуде. Осылайша, адам мен машинаның өзара әрекеттесуіндегі табиғи тілді өңдеудің рөлі, демек, киберқауіпсіздік әлемінде қарқын ала бастайды. Сонымен, NLP-ның киберқауіпсіздіктегі рөлі мен қолданылуы қандай болуы мүмкін?

(Spear-) фишинг алаяқтар алаяқтық құрбандарына арнайы электрондық пошталар, мәтіндік хабарламалар немесе телефон қоңырауларын жіберген кезде пайда болады. Негізгі ресурс ретінде әлеуметтік медианың пайда болуымен фишерлер одан да тиімді бола бастайды және NLP құралдары осы қосымшаларды қолдануға өте ыңғайлы. NLP ақпаратын жинау міндеттерін қолдана отырып, шабуылдаушылар әлеуметтік желілер мен басқа көздер арқылы жеке ақпаратты жинауды автоматтандырады, бұл шабуылды анықтауды қиындатады. Алайда, қорғаушылар осы шабуылдарға қарсы tdd сияқты құралдарды қолдана

алады. NLP-ді ұйымға қарсы фишингтік науқандарды анықтау үшін қолдануға болады.

Мүмкін болатын веб-шабуылдар мен қорғаныс механизмдерінің кең спектрі уақыт қатарлары немесе HTTP сұраулары / жауаптары болсын, терең оқыту мен NLP-ді қолдана алады. Шабуылдаушылар ақпарат жинау әдістерін немесе бәсекеге қабілетті машиналық оқытуды қолдана алады, ал қорғаушылар seq2seq atoencoders және басқа веб-қауіпсіздік модельдері сияқты NLP қосымшаларын қолдана отырып, өз позицияларын нығайта алады. Үнемі өсіп келе жатқан техникалық проблемаларға қарамастан, WAF және аномалияларды анықтау саласындағы нарық көшбасшылары өз өнімдеріне осындай мүмкіндіктерді енгізуді қарастыруда деп болжай аламыз.

Зиянды бағдарламалар мен кодтарды талдау туралы сөз болғанда, Endgame деректерді өңдеу мамандары зиянды кодты жақсы анықтау және түсіну үшін NLP-дің озық әдістерін қолданады. Олар зиянды бағдарламаларды талдауға арналған malicious Language Processing платформасын жасады. Оның мақсаты-NLP-ді жақсы кодта жасырылған зиянды кодты анықтауды автоматтандыру және жеделдету арқылы іске қосу. Лексикалық талдау тұжырымдамасын қолдана отырып, олар зиянды екілік файлдарды үлкен мәтін ретінде қарастырады. Осылайша, машиналар кодты оны орындамай-ақ "түсіне" алды. Сол сияқты, осалдықты экстраполяциялау арқылы осалдықты бағалау үшін NLP әдістерін кеңейтуге болады.

NLP көмегімен threat intelligence өнімдері бірнеше тілдегі сөздер мен техникалық деректердің мағынасын оқып, түсініп қана қоймайды, сонымен қатар заңдылықтарды анықтау үшін миллиардтаған мәліметтер нүктелерін қолданады.

NLP негізіндегі онтология IT-тәуекелдерді басқару мен кибер тұрақтылықты автоматтандыру мен біріктіруді қолдайды. Тәуекелдерді талдау және бағалау мәтіндік ақпаратты да қамтитынын ескере отырып, NLP пайдаланатын IT-тәуекелдерді басқаруға арналған өнімдер әртүрлі құрылымдар мен әдіснамалардың талаптарын IT-ортадан жинақталған деректермен салыстыра алады. Бұл өнімнің нормативтік және құқықтық талаптарға сәйкестігін қамтамасыз етуге мүмкіндік береді және құқық қорғау органдарымен қарым-қатынасты жеңілдетеді. Сонымен қатар, NLP контент-аналитикасының мүмкіндіктері нормативтік талаптардың өзгеруін тиімді бақылауға және талаптарға байланысты шығындарды бағалауды қолдауға мүмкіндік береді. NLP-ны қауіп-қатер модельдерін

жақсарту арқылы киберқауіпсіздік моделінің қаупін азайту үшін пайдалануға болады.

Нәтижелер. Электрондық коммерция және электрондық төлем жүйелері саласындағы соңғы жетістіктер несие карталарына қатысты қаржылық алаяқтық жағдайларының көбеюіне әкелді. Сондықтан несие картасындағы алаяқтықты анықтайтын құралдарды енгізу өте маңызды. Бұл мақалада белгілерді таңдау үшін генетикалық алгоритмді (GA) қолдана отырып, машиналық оқыту (ML) негізінде несие карталарын алаяқтықты анықтау механизмін ұсынған. Оңтайландырылған функцияларды таңдағаннан кейін, ұсынылған анықтау механизмі келесі машиналық оқыту жіктеуіштерін қолданады: шешім ағашы (DT), кездейсоқ орман (RF), логистикалық регрессия (LR), жасанды нейрондық желі (ANN) және аңғал Байес (NB). Тиімділікті тексеру үшін несие карталарының алаяқтықтарын анықтаудың ұсынылған механизмі еуропалық карта ұстаушыларынан алынған мәліметтер жиынтығын қолдана отырып бағаланады. Нәтиже ұсынылған тәсіл қолданыстағы жүйелерден асып түсетінін көрсетті (Ilebari etal., 2022:5).

Бұл мақалада (Alshehri et al., 2022:2) пайдаланушының мінез-құлқын талдаумен бірге машиналық оқытуды қолдана отырып, кибершабуылдарды анықтауға арналған жаңа платформа ұсынылған. Жақтаушы пайдаланушының мінез-құлқын оның әрекеттерін білдіретін оқиғалар тізбегі ретінде модельдейді. Содан кейін ұсынылған тізбектер жеке пайдаланушылардың ерекше мінез-құлқын анықтайтын белгілерді алу үшін нейрондық желінің қайталанатын моделіне орналастырылады. Осылайша, модель желідегі пайдаланушының мінез-құлқын қалыптастыру үшін тұрақты мінез-құлық жиілігін тани алады. Кейінгі процедура қайталанатын нейрондық желі белгісіз мінез-құлықты әдеттегі немесе тұрақты емес мінез-құлық ретінде жіктеу арқылы қалыптан тыс мінез-құлықты анықтайды. Ұсынылған құрылымның маңыздылығы кибершабуылдардың көбеюіне байланысты. Әдетте, ішкі шабуылдарды анықтау әлдеқайда қиын міндет, өйткені қауіпсіздік хаттамалары желідегі сенімді ресурстардан тұрады, соның ішінде пайдаланушылардан шабуылдарды тану қиынға соғады. Осылайша, пайдаланушының мінез-құлқын анықтауға болады және нәтижесінде қарапайым шаблондар әдеттегі желілік жұмыс процесін көрсететін терең заңдылықтарды тануға үйренеді. Эксперименттік нәтижелербұл тәсіл басқа тәсілдермен салыстырғанда жақсы нәтиже көрсеткенін және RNN-LSTM 1 көмегімен AUC 0,97 қол жеткізілгенін көрсетеді.

Заттар интернетінің (IoT) және киберфизикалық жүйелердің (CPS)

кеңеюін ескере отырып, киберқауіпсіздіктің ықтимал мәселелерін тиімді анықтау мен басқаруды дамыту ғана емес, сонымен қатар заттар интернетінің қауіпсіздігін қамтамасыз ету стандарттарына негізделген тиімді және бейімделгіш басқарумен байланысты мәселелерді шешу маңызды. Бұл зерттеу қолданыстағы стандарттарға кең және сыни зерттеулер жүргізеді және киберфизикалық желілерді кеңінен қолдануға қолдау көрсету үшін назар аудару керек бағыттарды анықтайды (Dong et al., 2022:7).

Фишинг – бұл сандық коммуникациядағы сенімді ұйым ретінде көрінетін құпия ақпаратты алуға бағытталатын алаяқтық әрекет. Бұл кибершабуылдың бір түрі, ол көбінесе сәтті болады, өйткені пайдаланушылар өздерінің осал тұстарын білмейді немесе тәуекелдерді түсіне алмайды. Бұл мақалада (Desolda et al., 2021:2) адам факторы мен фишинг саласындағы ең маңызды зерттеу жұмыстарының «бейнесін» салу мақсатында жүргізілген әдебиеттерге жүйелі шолу берілген. Әдебиеттерді жүйелі шолуда қарастырылған зерттеу мәселелеріне сәйкес алынған жарияланымдарды талдау фишингтік шабуылдардан қорғану үшін адам факторын қалай ескеру керектігін түсінуге көмектеседі.

Қауіп-қатер оқиғалары мен ымыраға келу индикаторларын бөлісу кибершабуылдарға қарсы тиімді қарсы шараларға қатысты тез және сыни шешім қабылдауға мүмкіндік береді. Алайда, қауіп-қатер туралы ақпарат алмасудың қолданыстағы шешімдері машинаны оқыту әдістерін (ML) қолдана отырып, қауіпті анықтау жүйелері (атап айтқанда, интрузияны анықтау жүйелері (IDS)) арасында ақпарат пен білімді оңай алмасуға мүмкіндік бермейді. Сонымен қатар, ML алгоритмдері үшін сенімді кірістерді жинаудың маңызды құрамдас бөлігі болып табылатын сарапшымен өзара әрекеттесу нашар қолдау табады. Осы мәселелердің барлығын шешу үшін ORISHA, қауіп-қатерді анықтау жүйелері мен ақпараттың басқа компоненттері арасында ынтымақтастық орнатуға мүмкіндік беретін ұйымдасқан ақпарат алмасу және ақпараттандыру платформасы ұсынылады. ORISHA-ға әр түрлі ұйымдарға тиесілі бірнеше қауіпті анықтау деңгейлерімен байланыс орнатуға мүмкіндік беретін зиянды бағдарламалар туралы ақпарат алмасу платформасының өзара байланысты даналары желісіне негізделген таратылған қауіп-қатерді талдау платформасы қолдау көрсетеді. Белгілі шабуылдарды анықтау сынағында жүргізілген эксперимент ұсынылған архитектураның дұрыстығын көрсетеді (Guarascio et al., 2022:6).

Соңғы жылдары қол жеткізілген жылдам технологиялық прогрестің арқасында көптеген адамдар өздерінің өмір салтын дәстүрлі бизнес тәсілдерінен электронды ресурстарға ауыстыруда. Аталған үдеріс берілген мақалада "қаскүнемдер" деп аталатын киберқылмыскерлердің назарын аударды (және әлі де жалғасуда), олар Интернет құрылымын фишинг сияқты киберқылмыскерлерді пайдаланушыларды жеке мәліметтерді, соның ішінде жеке ақпаратты, банктік және несие карталарын ашуға алдау үшін қолданады. Сенімді ұйымдардың заңды веб-сайттарының көшірмелері арқылы деректер, идентификаторлар, парольдер және маңызды ақпараттар ұрланады. Қазіргі таңда орын алған COVID-19 пандемиясы бұрын-соңды болмаған жағдай. Нәтижесінде, көптеген адамдар осы қауіпті жағдай туралы сенімді ақпарат жинауға тырысып, кибершабуылдарға осал болу үстінде. Өкінішке орай, осы жағдайды пайдаланып, пандемияға байланысты нақты шабуылдардың саны күрт өсті. Осы себепті корпорациялар мен киберқауіпсіздік зерттеушілері осы өсіп келе жатқан мәселені шешу үшін үнемі тиімді және инновациялық шешімдерді әзірлеуі керек. Қара тізімдерді, визуалды эффектілерді, эвристиканы және басқа да қорғаныс шешімдерін қолдану сияқты фишингпен күресудің бірнеше тәсілдері қазірдің өзінде қолданылып жатқанына қарамастан, олар фишинг шабуылдарының алдын алуға тиімді бола алмайды. Берілген мақалада авторлар COVID-19-мен байланысты домендік атауларды зиянды немесе заңды деп жіктеу үшін шектеулі функцияларды қолданатын машиналық оқыту модельдерін ұсынады. Алынған алғашқы нәтижелер домендік атаулардан алынған лексикалық белгілердің аз жиынтығы модельдерге жоғары балл алуға мүмкіндік беретіндігін, сонымен қатар, функция ретінде қосалқы домендер деңгейінің саны болжамдарға үлкен әсер етуі мүмкін екендігін көрсетеді (Mvula et al., 2022:3).

Киберқауіптердің үдемелі нашарлауы жағдайында олар туралы ақпаратты (CTI) ашық бастапқы қауіп-қатер туралы ақпаратты жариялау платформаларынан (OSTIPs) жинау ақпараттық қауіпсіздік қызметкерлеріне қоғамдық пікірді нақты анықтауға, төтенше жағдайларды жеңуге және тіпті қазіргі заманғы тұрақты қауіптерге қарсы тұруға көмектеседі. Алайда, жиіұсынылатын ақпарат көлемінің тез өсуіне байланысты CTI қолмен жинау тиімсіз болып шықты. OSTIPs-те жарияланған мақалалар құрылымданбаған, бұл CTI жазбаларын тек табиғи тілдерді өңдеу әдістерімен (NLP) автоматты түрде жинаудың шұғыл қажеттілігіне әкеледі. Осы шектеулерді жою үшін, берілген мақалада NLP әдісін, машиналық оқыту әдісін және киберқауіпсіздік

қауіптері туралы білімді біріктіретін көп типті кеңестерге (GCO) негізделген CTI жазбаларын құрудың автоматты тәсілі ұсынылған. Эксперимент нәтижелері GCO ұсынған мақалаларды жіктеудің және киберқауіпсіздік туралы мәліметтерді (CSIs) 93%-дан асатын дәлдікпенанықтайтынын көрсетеді, нәтижесінде Neo4j негізіндегі CTI деректер базасында жасалған жазбалар зиянды қауіптер тобын анықтауға көмектеседі (Sun et al., 2021:18).

Киберқауіпсіздік бойынша сарапшылар өз жұмысында NVD сияқты мәліметтер базасында сақталған білімге сүйенеді, бірақ бұл қауіптер мен осалдықтар туралы жалғыз ақпарат көзі емес. Бұл ақпараттың көп бөлігі әлеуметтік медиа арналары арқылы келеді. Бұл мақалада авторлар қауіпсіздік мамандары мен қарапайым пайдаланушылар онтологиялық көзқараста әртүрлі білім көздерін біріктіру арқылы семантикалық желі технологияларынан пайда көре алады деп мәлімдейді. Олар осалдықтардың онтологиясына негізделген, бірақ NLP құралдарымен толықтырылған, әлеуметтік желілерде киберқауіпсіздікке қатысты ақпаратты анықтауға және әртүрлі деректер көздеріне сұраныстарды іске қосуға арналған жүйені ұсынады. Биомедициналық салада дәлелденген киберқауіпсіздікті қамтамасыз ету үшін семантикалық желі технологиясының трансформациялық күші бағаланып, талқыланады (Aranovich et al., 2021:8).

Осалдықтардың ауырлығын тез және дәл бағалау және кибершабуылдарға қарсы шаралардың басымдықтарын анықтау үшін осалдықтар мен шабуылдар туралы ақпарат жинау қажет. Жалпы осалдықтар – оқиғаларды тізімдейтін сөздік, жалпы шабуыл үлгілерін тізімдеу және жіктеу-шабуыл үлгілерінің сөздігі. Шабуыл үлгілерін жалпы осалдықтарға тікелей сәйкестендіру және жіктеу қиын, өйткені олар әрқашан тікелей байланысты бола бермейді. Бұл жұмыста сөздіктер арасындағы ортақ байланыстарды тікелей іздеу тәсілін ұсынады. Содан кейін ұқсастық шаралары мен танымал алгоритмдердің жиынтығы болып табылатын бірнеше шаблондар, мысалы, терминдердің жиілігі – құжаттың кері жиілігі, әмбебап сөйлем кодтаушысы және BERT сөйлемдері ұсынылған тәсілді қолдана отырып эксперименталды түрде бағаланады. Жүргізілген эксперименттер term frequency–inverse document frequency алгоритмі ең жақсы жалпы өнімділікті қамтамасыз ететіндігін растайды (Kanakogi et al., 2022:6).

Адамның қарым-қатынасы олардың өзара әрекеттесуінің негізіндегі эмоцияларға байланысты. Әлеуметтік медианы қолданудың өсуі мәтіндік деректердің өзара байланысын интернеттегі өзара әрекеттесу арқылы

анықтауға болады. Мұндай өзара әрекеттесулер мәтіндік хабарламалар, электрондық пошталар және әлеуметтік медиа хабарламалары түрінде қол жетімді мәтіндік деректердің көптігіне әкеледі. Адамның қарым-қатынасын анықтау және талдау киберқауіпсіздіктен бастап қоғамдық денсаулыққа дейінгі көптеген қосымшалар үшін пайдалы. Бұл мақалада авторлар RIEA (эмоцияны талдау арқылы қарым-қатынасты анықтау) деп аталатын әдісті ұсынады, олардың арасындағы әңгімені талдау арқылы бірнеше зияткерлік агенттер арасындағы қатынасты анықтауға болады. Берілген жұмыстың мақсаты – эмоцияларды шығару және оларды қатынастар жиынтығына көрсету және уақыт өте келе қатынастардың қалай өзгеретінін талдау үшін когнитивті психология және табиғи тілді өңдеу (NLP) ұғымдарын біріктіру. Авторлар көптеген әңгімелерді талдау үшін психологиялық модельдерді қолданады және эмоциялар мен қатынастардың сәйкестігін анықтау үшін машиналық оқыту әдістерін қолданады. Олар жіктеу үшін ең жақсы масштабтау әдісін қолдана отырып, төрт түрлі ассоциативті сыныпты және төрт тіркеме стилін қолданады. Алынған нәтижелер RIEA тұлғааралық қатынастарды 85% дәлдікпен дұрыс анықтай алатындығын көрсетеді. Бағалау RIEA сөйлесулерден тұлғааралық қатынастарды дәл анықтай алатындығын және күрделі қатынастарды анықтау үшін кеңейтілуі мүмкін екенін көрсетеді. Бұл зерттеу сонымен қатар эмоционалды мінез-құлықтағы өзгерістердің уақыт өте келе қарым-қатынастың дамуына әсерін көрсетеді (Qamar et el., 2021:7).

Әлеуметтік медиа қылмыстарды жасау және анықтау үшін қолданылады. Автоматтандырылған әдістерді қолдана отырып, қылмысты ашуды да кеңейтуге болады. Қылмыскерлердің көптеген адамдарды қамту қабілеті бұл аймақты жиі зерттеу тақырыбына айналдырды, сондықтан әлеуметтік платформаларда жасалған нақты қылмыстарды қарастыратын бірнеше сауалнамалар жүргізілді. Осы уақытқа дейін әлеуметтік желілердегі қылмыстардың барлық түрлерін, олардың ұқсастықтарын, сондай-ақ олардың ашылуын қарастыратын шолу мақаласы болған жоқ. Қылмыстар мен оларды ашу әдістерінің ұқсастығын көрсету домендер арасында әдістер мен деректерді жіберуге мүмкіндік береді. Осылайша, зерттеудің берілген мақсаты – әлеуметтік желілерде жасалған қылмыстарды құжаттау және олардың ұқсастықтарын қылмыс таксономиясы арқылы көрсету. Сонымен қатар, бұл сауалнама жалпыға қол жетімді деректер жиынтығын құжаттайды (Drury et al., 2022:4).

Фишинг жыл сайын миллиардтаған доллар шығындарға алып

келеді және интернет-экономикаға үлкен қауіп төндіреді. Фишингтік шабуылдар қазіргі уақытта көбінесе электрондық пошта арқылы жүзеге асырылады. Фишингтік электрондық поштаны анықтаудың қазіргі зерттеу тенденциясын жақсы түсіну үшін бірнеше зерттеу жұмыстары жүргізілді. Алайда, бұл мәселені әр түрлі тұрғыдан бағалау маңызды. Тек бірнеше баламаларды зерттей отырып, жіктеу және оқыту мақсаттары үшін NLP әдістерін қолдануға жарық берген бір жұмысты қоспағанда, бірде-бір сауалнама фишингті анықтау үшін табиғи тілдерді өңдеу әдістерін (NLP) қолдануды ешқашан жан-жақты зерттемеген. Бұл олқылықтың орнын толтыру үшін берілген зерттеудің мақсаты фишингтік электрондық пошталарды анықтау үшін NLP қолдану бойынша зерттеулерді жүйелі түрде шолу және қорытындылау болып табылады. Алдын-ала анықталған критерийлер негізінде 2006 жылдан 2022 жылға дейін жарияланған 100 ғылыми мақала іріктеліп, талданды. Авторлар NLP көмегімен фишинг хаттарын анықтау бойынша негізгі зерттеу салаларын, фишингті анықтау үшін электрондық пошталарда қолданылатын машиналарды оқыту алгоритмдерін, фишинг хаттарындағы мәтіндік функцияларды, фишинг хаттарында қолданылған мәліметтер жиынтығы мен ресурстарды және бағалау критерийлерін зерттеген. Көптеген жіктеу алгоритмдерінің ішінде фишинг хаттарын анықтау үшін тірек векторлық машиналар (SVM) кеңінен қолданылады. NLP-тің ең көп қолданылатын әдістері-TF-IDF және сөздерді ендіру. Сонымен қатар, фишингтік электрондық поштаны анықтау әдістерін салыстырмалы талдау үшін ең көп қолданылатын мәліметтер жиынтығы-Nazario phishing corpus. Ұсынылған жұмыстарды талдау NLP әдістерін қолдана отырып, араб тіліндегі фишингтік электрондық пошталарда көп жұмыс жүргізілмегенін көрсетті (Salloum et al., 2022:22).

SMS арқылы фишингтік алаяқтық смартфондардың кеңінен қолданылуына және мобильді интернет технологиясының қол жетімділігіне байланысты жиі кездеседі. Құрылымданбаған қысқа мәтіндерді талдау арқылы фишингтік SMS-хабарламаларды анықтау жасанды интеллектке негізделген киберқауіпсіздік саласындағы күрделі міндет болып табылады. Табиғи тілді өңдеумен біріктірілген машиналық оқытуға негізделген әдістер фишинг пен заңды SMS хабарламаларының арасындағы айырмашылықты анықтауда үлкен әлеуетке ие. Бұл мақалада авторлар анықтамалық мәліметтер базасында бірнеше заманауи машиналарды оқыту алгоритмдерімен тәжірибе жасады. Сонымен қатар, фишингті анықтаудың автоматты стратегиясын құру

үшін NLP негізінде белгілерді алу және белгілерді таңдау кезеңдері кіреді. Белгілерді алу және таңдаудан кейін қолданылған кезде тірек векторларының машиналық жіктеуіші заңды SMS үшін F1 99,08% және дәлдігі 98,39% құрады. Тексерілген әдістердің тиімділігі бақылау жиынтығындағы танымал бағалау көрсеткіштерінің көмегімен бағаланған (Ulfath et al., 2021:8).

Кибернетикалық осалдықтарды анықтаудың дәстүрлі әдістерінің кемшіліктері жаңа қауіптерді анықтауға, оларды ортақ осалдықтардың (CVE) жазбаларына тіркеуге және оларды жалпы осалдықтарды бағалау жүйесі (CVSS) арқылы бағалауға қажетті уақытпен байланысты. Бұл проблемаларды әлеуметтік медиа мен ашық бастапқы деректерге негізделген осалдықтарды ерте анықтау жүйелері арқылы жеңілдетуге болады.Бұл жұмыста киберқауіпсіздік туралы жаңалықтардағы кибернетикалық осалдықтар Open Source Intelligence (OSINT) көмегімен ерте кибернетикалық қауіптерді автоматты түрде анықтау жүйесінің бөлігі ретінде анықтауға бағытталған модель ұсынылады. Машиналарды оқытудың үш моделі киберқауіпсіздік туралы мақалаларды тиісті (яғни қауіпсіздікке жаңа қауіп төндіретін) немесе маңызды емес деп жіктеудің берік негізін құру үшін 1000 таңбаланған жаңа мақалалар жиынтығында оқытылды: тірек векторлық машинасы, аңғал Байес классификаторы және BERT моделі. BERT моделі сынақ жиынтығында 88,45% орташа дәлдікпен ең жақсы өнімділікті көрсетті. Алынған тәжірибелер табиғи тілдерді өңдеу модельдері (NLP) киберқауіпсіздік туралы жаңалықтар мақалаларынан тиісті ақпаратты алу үшін осалдықты ерте анықтау жүйелері үшін қолайлы таңдау болып табылады деген қорытындыға әкеледі (Iorga et al., 2020:5).

Бұл жұмыста киберқауіпсіздік оқиғалары туралы ақпаратты алатын және киберқауіпсіздік туралы мәліметтерді графикке интеграциялаудың түпкі мақсаты бар семантикалық модельді толтыратын жүйені ұсынылады. Ол 2017-2019 жылдардағы ағылшын тіліндегі 1000 жаңалықтан тұратын жаңа корпуста оқытылды, олар оқиғаларға негізделген егжей-тегжейлі аннотациялармен белгіленген және кибершабуылдарды да, осалдықтармен байланысты оқиғаларды да қамтиды. Ұсынылатын модель оқиғаның бес түрін, олардың семантикалық рөлдерін және оқиғаға қатысты дәлелдердің 20 түрін (мысалы, файл, құрылғы, бағдарламалық жасақтама, ақша) анықтайды. CASIE терең нейрондық желілерге назар аудара отырып, әртүрлі тәсілдерді қолданады және бай лингвистикалық мүмкіндіктер мен сөздерді ендіруді қамтуы мүмкін. Авторлар оқиғаларды анықтаудың әр компонентімен

тәжірибе жүргізген және нәтижелер әрбір ішкі жүйенің жақсы жұмыс істейтінін көрсетеді (Satyapanich et al., 2020:2).

Хакерлік форумдар және басқа да әлеуметтік платформалар киберқауіпсіздік қауіптері туралы маңызды ақпаратты қамтуы мүмкін. Бірақ осы көздерден тиісті қауіп-қатер туралы ақпаратты алу үшін қолмен талдауды қолдану көп уақытты қажет ететін және қателікке бейім процесс болып табылады, ол айтарлықтай ресурстарды бөлуді қажет етеді. Бұл мақалада авторлар хакерлік форумдардағы қауіп-қатер туралы ақпаратты тез табу үшін машиналық оқыту әдістерінің әлеуетін зерттейді. Нақты хакерлік форумнан алынған мәтіндік деректерді қолдана отырып, авторлар мәтінді конвульсиялық нейрондық желі әдістерімен жіктеудің тиімділігін машиналық оқытудың дәстүрлі тәсілдерімен салыстырады.Олар машиналық оқытудың дәстүрлі әдістері, мысалы, тірек векторлар машинасы, конвульсиялық нейрондық желілердің алгоритмдерімен салыстырылатын өнімділіктің жоғары деңгейін қамтамасыз ете алатындығын анықтады (Deliu et al., 2017:8).

Талқылау. Қазіргі уақытта Интернеттегі киберқауіпсіздік туралы мәліметтер тез өсуде, бірақ олардың көпшілігі қауіпсіздікті талдау үшін уақытында түсіну қиын және автоматтандырылған қауіпсіздік жүйелерін тікелей пайдалануға жарамсыз құрылымданбаған мәтіндік мәліметтер болып табылады. Киберқауіпсіздік туралы ақпаратты құрылымданбаған мәтіндік көздерден құрылымдық көріністерге нақты уақыт режимінде автоматты түрде ауыстыру киберқауіпсіздік талдаушыларына кибернетикалық жағдайды жақсы түсінуге көмектеседі. Аталған нысандарды тану (NER) құрылымданбаған деректерді құрылымдық деректерге айналдыра алады. Жақында Transformers (BERT) ұсынған Bidirectional Encoder Representations деп аталатын тілдік бейнелеу моделі NLP-нің әртүрлі тапсырмаларында айтарлықтай жақсартуларға қол жеткізді. Бұл мақалада авторлар Bert және оның бүкіл әлемдегі BERT жетілдірілген нұсқасын (Bert wwm) NER киберқауіпсіздік есебіне қолданады. Олар BERT моделін BiLSTM-CRF архитектурасымен біріктіреді және эксперимент көрсеткендей, ұсынылатын әдіс F1 дәлдігі, Recallбағалауы бойынша қазіргі қолданыстағы модельге қарағанда жоғары өнімділікті қамтамасыз ететінін көрсетті (Zhou et al., 2021: 2).

Трафикті анықтау соңғы жылдары интрузияны анықтау жүйелерінде (IDS) маңызды рөл атқаруда. Бұл мақалада табиғи тілді өңдеу арқылы (NLP) пакеттік деңгейдегі трафикті анықтаудың жаңа тәсілі ұсынылған, ол ендірудің үлгісі ретінде қарапайым кон-

трастты ұсыныс қосымшаларын (SimCSE) қолданады. Жаңа тәсіл шикі пакеттік мәліметтерге негізделген трафиктің ерекшеліктерін зерттеуге мүмкіндік береді. Ұсынылатын әдісті бағалау үшін екі белгілі мәліметтер жиынтығымен тәжірибелер жүргізілді. Зиянды әрекетті анықтау үшін ұсынылатынмодель USTC-TFC2016 деректер жиынтығында 99,99% дәлдікке қол жеткізді, ал виртуалды жеке желі (VPN) қызметін анықтау үшін ұсынылатын модель ISCXVPN2016 деректер жиынтығында 99,98% дәлдікке қол жеткізді. Сонымен қатар, алынған модель нөлдік күндік шабуылдарды анықтау үшін тиімді екендігі анықталды, бұл модельдің бұрын байқалмаған шабуылдарды анықтау қабілетін көрсетеді. Тәжірибелер көрсеткендей, ұсынылатын тәсіл желілік трафикті тиімді анықтай алады және көптеген басқа заманауи әдістерден асып түседі (Bar et al., 2022:5).

Біреуді өз білімімен бөлісуге сендіру әдісі әлеуметтік инженерия ретінде белгілі. Әлеуметтік инженерлер адамдардың құнды ақпарат алмасудың салдарын білмеуіне, сондай-ақ олардың жүйелері мен ақпараттық технологиялар инфрақұрылымын қауіпсіздік шабуылдарынан қорғау туралы білімінің болмауына сүйенеді. Бұл шабуылдарды ұйым қызметкерлері үшінші тарап агенттігі арқылы жүзеге асыра алады. Олар қаржылық пайда немесе кек алу үшін ұйым ережелерін бұзады. Зиянкес зардап шеккендердің құпия ақпаратын жинау үшін әртүрлі тактикаларды қолданады, бұл – әлеуметтік инженерияға шабуыл жасау әдісі. Құпия ақпаратты заңсыз алу процесі қылмыстық әрекет болып табылады. Берілген зерттеу табиғи тіл процесі (NLP) арқылы белгісіз көзден немесе URL мекен-жайынан алынған хабарламаның спам немесе заңды екенін анықтау үшін құрылым ұсынады. COVID-19 кезінде көптеген адамдар интернетті күнделікті іс-әрекеттері үшін ол жердегі қауіпсіздік қатерлерін білместен қолдана бастады. Бұл шабуылдаушылардың осы құрбандарды нысанаға алуына және шабуылдарын тиімді түрде орындауына түрткі болды. SEA-бұл киберқауіпсіздікке қарсы шабуылдың бір түрі, ол бақылау жүйелерін бұзу үшін адамдардың табиғи қызығушылығын пайдаланады және сәттіліктің жоғары пайызына ие. Берілген зерттеудің мақсаты – COVID-19 пандемиясының әлеуметтік инженерлік шабуылдардың кеңеюіне қалай жол ашқаны, сондай-ақ осы шабуылдарды анықтау мен жеңілдетудің әртүрлі әдістері туралы егжей-тегжейлі зерттеу (Shalke et al., 2022:4).

Қорытынды. Бұл мақалада қазіргі таңда кейбір қауіпсіздік қатерлері мен кибершабуылдарға байланысты өте маңызды болып табылатын

киберқауіпсіздік мәселесі көтеріледі. Кибер қауіпсіздікке анықтама беріліп, қазіргі таңда жиі кездесетін шабуыл түрлері келтіріледі. Сонымен қатар, табиғи тілді өңдеу әдістері арқылы киберқауіпсіздіктің кей мәселелерін шешуге болатыдығы көрсетіліп, ағымдағы жұмыстарға кең әдеби шолу жасалған. Ғылыми жұмыс № AP06851248 «Мәтіндегі экстремистік бағытты анықтау үшін веб-ресурстардағы семантикалық талдау модельдерін, алгоритмдерін құрастыру және кибер-криминалистика құрал-жабдықтарын әзірлеу» атты жобаның аясында орындалды.

Information about the authors:

Bolatbek M.A. – University of Passau, researcher, Passau, Germany. Senior Lecturer, Department of Information Systems, Al-Farabi Kazakh National University, PhD, Almaty, Kazakhstan, E-mail: *bolatbek.milana@gmail.com*, *https://orcid.org/0000-0002-2153-180X*;

Bagitova K.B. – Doctoral student of the Department of Information Systems of the Al-Farabi Kazakh National University, Almaty, Kazakhstan. E-mail: *kbbagitova@gmail.com*, *https://orcid.org/0000-0003-1587-1995;*

Mussiraliyeva Sh.Zh. – Candidate of Physical and Mathematical Sciences, Head of the Department of Information Systems of the Al-Farabi Kazakh National University, Almaty, Kazakhstan. E-mail: *mussiraliyevash@gmail.com*, *https://orcid.org/0000-0001-5794-3649*.

REFERENCES:

Alshehri A., Khan N., Alowayr A. and Yahya M. Alghamdi, “Cyberattack detection framework using machine learning and user behavior analytics,” Computer Systems Science and Engineering, vol. 44, no.2, pp. 1679–1689, 2023.

Aranovich R., Wu M., Yu D., Katsy K., Ahmadnia B., Bishop M., Filkov V. and Sagae K. 2021. Beyond NVD: Cybersecurity meets the Semantic Web. In New Security Paradigms Workshop (NSPW ‘21). Association for Computing Machinery, New York, NY, USA, 59–69. https://doi.org/10.1145/3498891.3501259.

Bar R. and HajajC., “SimCSE for Encrypted Traffic Detection and Zero-Day Attack Detection,” in IEEE Access, vol. 10, pp. 56952-56960, 2022, doi: 10.1109/ACCESS.2022.3177272.

Deliu I., Leichter C., Franke K.. (2017). Extracting cyber threat intelligence from hacker forums: Support vector machines versus convolutional neural networks. 3648-3656. 10.1109/BigData.2017.8258359.

Desolda G.S., Ferro L., Marrella A., Catarci T. and Francesca Costabile M.. 2021. Human Factors in Phishing Attacks: A Systematic Literature Review. ACM Comput. Surv. 54, 8, Article 173 (November 2022), 35 pages. https://doi.org/10.1145/3469886.

Dong S., Cao J., Flynn D. et al. Cybersecurity in smart local energy systems: requirements, challenges, and standards. Energy Inform 5, 9 (2022). https://doi.org/10.1186/s42162-022-00195-7.

Drury B., Drury S.M., Arafatur Rahman Md, Ullah I., A social network of crime: A review of the use of social networks for crime and the detection of crime, Online Social Networks and Media,Volume 30,2022,100211,ISSN 2468-6964, https://doi.org/10.1016/j.osnem.2022.100211.

Guarascio M., Cassavia N., Sergio Pisani F., Manco G., Boosting Cyber-Threat Intelligence via Collaborative Intrusion Detection, Future Generation Computer Systems, Volume 135, 2022, Pages 30-43, ISSN 0167-739X,https://doi.org/10.1016/j.future.2022.04.028.

Ileberi E., Sun Y. & Wang Z. A machine learning based credit card fraud detection using the GA algorithm for feature selection. J Big Data 9, 24 (2022). https://doi.org/10.1186/s40537-022-00573-8.

Iorga D., Dragos-Georgian C., Octavian G., Cristian S., Mihai D., Razvan R. (2020). Early Detection of Vulnerabilities from News Websites using Machine Learning Models. 1-6. 10.1109/RoEduNet51892.2020.9324852.

Kanakogi K., Hironori W., Yoshiaki F., Shinpei O., Takao O., Takehisa K., Hideyuki K., Atsuo H. and Nobukazu Y. 2022. "Comparative Evaluation of NLP-Based Approaches for Linking CAPEC Attack Patterns from CVE Vulnerability Information" Applied Sciences 12, no. 7: 3400. https://doi.org/10.3390/app12073400.

Mvula P.K., Branco P., Jourdan G., Viktor H.L., COVID-19 malicious domain names classification, Expert Systems with Applications, Volume 204, 2022, 117553, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2022.117553.

Qamar S., Mujtaba H., Majeed H. et al. Relationship Identification Between Conversational Agents Using Emotion Analysis. Cogn Comput 13, 673–687 (2021). https://doi.org/10.1007/s12559-020-09806-5.

Salloum S., Gaber T., Vadera S. and Shaalan K., "A Systematic Literature Review on Phishing Email Detection Using Natural Language Processing Techniques," in IEEE Access, vol. 10, pp. 65703-65727, 2022, doi: 10.1109/ACCESS.2022.3183083.

Satyapanich, Taneeya, Francis Ferraro and Tim Finin. "CASIE: Extracting Cybersecurity Event Information from Text." AAAI (2020).

Shalke C.J. and Achary R., "Social Engineering Attack and Scam Detection using Advanced Natural Langugae Processing Algorithm," 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), 2022, pp. 1749-1754, doi: 10.1109/ICOEI53556.2022.9776697.

Sun, Tianfang, Pin Yang, Mengming Li, and Shan Liao. 2021. "An Automatic Generation Approach of the Cyber Threat Intelligence Records Based on Multi-Source Information Fusion" Future Internet 13, no. 2: 40. https://doi.org/10.3390/fi13020040.

Ulfath Rubaiath E & Sarker, Iqbal & Chowdhury, Mohammad & Hammoudeh, Mohammad. (2021). Detecting Smishing-Attacks Using Feature Extraction and Classification Techniques. 10.1007/978-981-16-6636-0_51.

Zhou S., Liu J., Zhong X. and Zhao W., "Named Entity Recognition Using BERT with Whole World Masking in Cybersecurity Domain," 2021 IEEE 6th International Conference on Big Data Analytics (ICBDA), 2021, pp. 316-320, doi: 10.1109/ICBDA51983.2021.9403180.

МАЗМҰНЫ

СОДЕРЖАНИЕ

CONTENTS

Publication Ethics and Publication Malpracticein
the journals of the National Academy of Sciences of the Republic of Kazakhstan

For information on Ethics in publishing and Ethicalguidelines for journal publication see http://www.elsevier.com/publishingethics and http://www.elsevier.com/journal-authors/ethics.

Submission of an article to the National Academy of Sciences of the Republic of Kazakhstan implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see http://www.elsevier.com/postingpolicy), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any otherlanguage, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The National Academy of Sciences of the Republic of Kazakhstan follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct (http://publicationethics.org/files/u2/New_Code.pdf). To verify originality, your article may be checked by the Cross Check originality detection service http://www.elsevier.com/editors/plagdetect.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the National Academy of Sciences of the Republic of Kazakhstan.

The Editorial Board of the National Academy of Sciences of the Republic of Kazakhstan will monitor and safeguard publishing ethics.

Правила оформления статьи для публикации в журнале смотреть на сайтах:
www:nauka-nanrk.kz
http://physics-mathematics.kz/index.php/en/archive
ISSN2518-1726 (Online),
ISSN 1991-346X (Print)

Директор отдела издания научных журналов НАН РК *А. Ботанқызы*
Заместитель директор отдела издания научных журналов НАН РК *Р. Жәлиқызы*
Редакторы: *М.С. Ахметова, Д.С. Аленов*
Верстка на компьютере *Г.Д.Жадыранова*

Подписано в печать 15.09.2022.
Формат 60х881/8. Бумага офсетная. Печать – ризограф.
17,5 п.л. Тираж 300. Заказ 3.

Национальная академия наук РК
050010, Алматы, ул. Шевченко, 28, т. 272-13-19