«8D06301 – Ақпараттық қауіпсіздік жүйелері» білім беру бағдарламасы бойынша (PhD) философия докторы дәрежесін алу үшін ұсынылған Шайзат Медет Жанболатұлының «Bitcoin-транзакцияларын киберкриминалистикалық талдауға арналған әдістер мен машиналық оқыту модельдерін әзірлеу» тақырыбындағы диссертациялық жұмысына

АҢДАТПА

Зерттеу тақырыбының өзектілігі. Цифрлық активтер транзакцияларын талдау құқықтық реттеу, қаржылық мониторинг және цифрлық криминалистика үшін маңызды. Қазақстан Республикасы үшін бұл мәселе Қазақстан Республикасының «Қазақстан Республикасындағы цифрлық активтер туралы» Заңының қолданыста болуына байланысты ерекше өзекті, өйткені ол тиісті операцияларды бақылау және талдау құралдарына, оның ішінде қаржылық мониторинг пен құқық бұзушылықтарды тергеп-тексеру кезінде практикалық қажеттілік қалыптастырады.

Мәселенің халықаралық маңыздылығы виртуалды активтер мен тиісті қызметтерді жеткізушілерге қатысты FATF стандарттарын имплементациялау әлі де толық емес әрі біркелкі емес екендігімен расталады. Тәуекелдерді бағалау, қадағалау және Travel Rule талаптарын қолдану бойынша қиындықтар сақталуда, бұл транзакциялық мониторингтің талдамалық құралдарына деген қажеттілікті күшейтеді.

Зерттеудің өзектілігі криптовалюта ортасындағы заңсыз қызметтің ауқымымен де айқындалады. Chainalysis бағалауы бойынша, 2025 жылы заңсыз қызметпен байланысты мекенжайларға кемінде 154 млрд АҚШ долл. түскен. Мұндай операциялардың атрибуцияланған криптотранзакциялардың жалпы көлеміндегі үлесі салыстырмалы түрде аз болғанына қарамастан, олардың абсолюттік көлемі елеулі болып қалып отыр және тиісті ағындарды анықтау, талдау және құжаттау әдістерін дамытуды талап етеді.

Bitcoin желісін таңдау қатысушылардың псевдонимділігі сақталған жағдайда транзакциялық графты талдауға, қаражат қозғалысын қадағалауға және мінез-құлық заңдылықтарын зерттеуге мүмкіндік беретін желінің ашық архитектурасымен негізделген. Алайда блокчейн деректерінің қолжетімділігі оларды бірмәнді интерпретациялауға кепілдік бермейді. Қолданыстағы әдістердің нәтижелері таңбалаудың толықтығы мен атрибуциялау рәсімдеріне тәуелді, бұл оларды тексеру мен салыстыруды қиындатып, қате қорытындылар қаупін арттырады.

Осылайша, ғылыми проблема заңсыз қызмет белгілері бар Bitcoin-транзакцияларын анықтауға және талдау нәтижелерін транзакциялық, уақыттық және графтық контексті ескере отырып түсіндіруге мүмкіндік беретін әдістердің жеткілікті дәрежеде әзірленбеуінде.

Зерттеу тақырыбының ғылыми-зерттеу бағдарламаларымен байланысы: Диссертациялық жұмыс 2025–2027 жылдарға арналған «Жас ғалым» жас ғалымдарды гранттық қаржыландыру шеңберінде AP25793719

«Заңсыз криптовалюталық транзакцияларды талдау жүйесін әзірлеу» жобасы бойынша, «Ұлттық қауіпсіздік және қорғаныс, биологиялық қауіпсіздік» басым бағыты және «Ақпараттық қауіпсіздікті қамтамасыз ету» мамандандырылған бағыты аясында орындалды.

Зерттеудің мақсаты. Диссертациялық зерттеудің мақсаты Bitcoin-транзакцияларын киберкриминалистикалық талдауға арналған интерпретацияланатын белгілерді қалыптастыру әдістері мен машиналық оқыту модельдерін әзірлеу болып табылады.

Зерттеу міндеттері. Қойылған мақсатқа қол жеткізу үшін диссертациялық жұмыста мынадай міндеттер шешіледі:

1. Машиналық оқыту мен графтық талдауға негізделген Bitcoin-транзакцияларын киберкриминалистикалық талдаудың қолданыстағы тәсілдерін талдау.
2. Bitcoin-транзакцияларының қолданыстағы деректер жиындарын және транзакциялық графтарды құру тәсілдерін зерттеу, олардың негізгі шектеулерін анықтау.
3. Elliptic және Elliptic++ анонимдендірілген деректер жиындары үшін генетикалық алгоритмге негізделген белгілерді іріктеу әдісін әзірлеу.
4. Заңсыз Bitcoin-транзакцияларын анықтауға арналған кестелік және графтық модельдерді салыстырмалы бағалау.
5. Bitcoin ашық блокчейнінің деректері негізінде семантикасы ашық белгілері бар Bitcoin-транзакцияларының мамандандырылған деректер жиынын қалыптастыру.
6. Жергілікті ортадан ауытқу белгілерін, қаражатты кейінгі жұмсаудың уақыттық белгілерін және алдыңғы транзакциялық тарих белгілерін қамтитын криминалистикалық тұрғыдан интерпретацияланатын белгілер жүйесін әзірлеу.
7. Қалыптастырылған деректер жиынында машиналық оқыту модельдерін әзірлеу және бағалау, әзірленген белгілердің үлесін анықтау және жіктеу нәтижелерін интерпретациялау.
8. Транзакцияларды жіктеуді, мекенжайлық, мінез-құлықтық және графтық талдауды біріктіретін Bitcoin-транзакцияларын киберкриминалистикалық талдаудың бағдарламалық жүйесін әзірлеу.

Зерттеу объектісі: Bitcoin-транзакцияларын киберкриминалистикалық талдау процесі.

Зерттеу пәні: Bitcoin-транзакцияларын киберкриминалистикалық талдауға арналған белгілерді қалыптастыру және іріктеу әдістері, сондай-ақ машиналық оқыту модельдері.

Зерттеу әдістері. Жұмыста жүйелік және салыстырмалы талдау, генетикалық алгоритм, деректерді талдау, машиналық оқыту және модельдерді интерпретациялау әдістері қолданылды.

Зерттеудің ғылыми жаңалығы

Практикалық кейстерді талдау кезінде анықталған қаражат қозғалысының үлгілерін формалдайтын Bitcoin-транзакцияларының

криминалистикалық тұрғыдан интерпретацияланатын 61 дескрипторынан тұратын жүйе ұсынылды. Белгілер графтық ортадан жергілікті ауытқуды, транзакция шығыстарын кейінгі жұмсау динамикасын және ағымдағы сипаттамалардың алдыңғы транзакциялық тарих бойынша күтілетін мәндерден ауытқуын сипаттайды. Оларды 26 базалық белгіге қосу тексерілген барлық модельдердің сапасын арттырды: PR-AUC орташа өсімі 3.97 п. т., ал F1-score – 4.26 п. т. құрады. Абляциялық талдау белгілердің барлық үш тобының үлесін растады.

Анонимдендірілген Elliptic сипаттамаларының орнына семантикасы ашылған 87 белгісі бар қалыптастырылған деректер жиынын пайдаланатын CatBoost негізіндегі заңсыз Bitcoin-транзакцияларын анықтау моделі әзірленді. Elliptic графтық құрылымы, уақыттық кесінділері мен таңбалануы сақталды, транзакциялардың ашық идентификаторлары қалпына келтірілді, ал белгілерді есептеу тәртібі формалды түрде анықталды. Модель PR-AUC 90.27 %, F1-score 83.90 % және Recall 81.11 % орташа мәндеріне жетті.

Генетикалық іздеуге және ішкі жиындарды заңсыз класс бойынша F1-score көмегімен Random Forest моделімен бағалауға негізделген анонимдендірілген белгілерді іріктеудің қаптамалық әдісі ұсынылды. Іріктеу тек оқыту уақыттық кесінділерінде орындалады. Elliptic жиынында әдіс белгілер кеңістігін 166-дан 53 белгіге дейін қысқартты және Precision 99.37 %, Recall 73.13 % және F1-score 84.26 % көрсеткіштерін қамтамасыз етіп, жарияланған RF+GCN көрсеткіштерінен тиісінше 2.3, 5.6 және 4.7 п. т. асып түсті. Elliptic++ жиынында 183 белгінің 85-і іріктелді, ал F1-score толық жиын нәтижесінен 1.57 п. т. жоғары болды.

Зерттеудің теориялық маңыздылығы жергілікті ортадан ауытқу белгілерін, кейінгі жұмсаудың уақыттық белгілерін және алдыңғы транзакциялық тарих белгілерін формалдау арқылы Bitcoin-транзакцияларын интерпретацияланатын киберкриминалистикалық талдаудың теориялық-әдістемелік негіздерін дамытуында. Ұсынылған жүйе бұл белгілерді транзакциялық және құрылымдық-графтық сипаттамалармен біріктіреді, бұл транзакция параметрлерін, оның графтағы орнын, жергілікті ортадан ауытқуын, кейінгі жұмсау динамикасын және алдыңғы транзакциялық тарих бойынша есептелген күтілетін мәндерден ауытқуын ескеруге мүмкіндік береді. Әзірленген белгілер тікелей блокчейн деректері бойынша есептеледі және мазмұндық интерпретацияға ие.

Зерттеудің практикалық маңыздылығы әзірленген бағдарламалық-талдамалық құралдарды Bitcoin-транзакциялары туралы деректерді дайындау, оларды талдау және жіктеу, модель нәтижелерін интерпретациялау және алынған қорытындыларды сараптамалық тексеруге ұсыну үшін қолдану мүмкіндігінде. Әзірленген құралдарды қаржылық мониторинг бөлімшелері, құқық қорғау органдары, криптовалюталық биржалар және кірістерді жылыстатуға қарсы іс-қимыл жасайтын әрі криптовалюталарды пайдалану арқылы жасалған қылмыстарды тергейтін ұйымдар қолдана алады.

Қорғауға ұсынылатын негізгі ережелер:

1. Криминалистикалық гипотезаларды формалдайтын Bitcoin-транзакцияларының криминалистикалық тұрғыдан интерпретацияланатын белгілер жүйесі заңсыз транзакцияларды анықтау сапасын арттырады: оны базалық белгілер кеңістігіне қосу тексерілген барлық модельдер үшін PR-AUC және F1-score өсімін қамтамасыз етті, бұл ретте ең жоғары өсім PR-AUC бойынша 5.19 п. т. және F1-score бойынша 7.46 п. т. құрады.
2. CatBoost алгоритміне негізделген заңсыз Bitcoin-транзакцияларын анықтау моделі семантикасы ашық транзакциялық, құрылымдық-графтық және криминалистикалық тұрғыдан интерпретацияланатын белгілері бар қалыптастырылған деректер жиынында оқытылуымен және Elliptic анонимдендірілген белгілерін пайдаланбауымен ерекшеленеді, бұл оны нақты Bitcoin-транзакцияларын бастапқы талдау үшін қолдануға мүмкіндік береді; қол жеткізілген PR-AUC, F1-score және Recall мәндері тиісінше 90.27 %, 83.90 % және 81.11 % құрады.
3. Анонимдендірілген белгілері бар блокчейн деректер жиындарына арналған, белгілердің ішкі жиындарын генетикалық іздеуге және оларды заңсыз класс бойынша F1-score көмегімен Random Forest моделі арқылы қаптамалық бағалауға негізделген белгілерді іріктеу әдісі Elliptic деректер жиынына қолданылған кезде бастапқы 166 белгінің орнына 53 белгіден тұратын ішкі жиынды қалыптастыруға мүмкіндік берді, бұл ретте іріктелген ішкі жиындағы Random Forest моделінің Precision, Recall және F1-score мәндері Elliptic бастапқы бенчмаркінің жарияланған RF+GCN моделі көрсеткіштерінен тиісінше 2.3, 5.6 және 4.7 п. т. жоғары болды.

Диссертациялық жұмыс нәтижелерінің апробациясы. Зерттеу нәтижелері әл-Фараби атындағы ҚазҰУ-дың ақпараттық технологиялар және жасанды интеллект факультеті мен киберқауіпсіздік және криптология кафедрасының ғылыми семинарларында талқыланды, сондай-ақ «Фараби әлемі» конференциясында «. Биткоин желісіндегі ауытқуларды анықтау әдісі» баяндамасында ұсынылды (Алматы, 2022).

Диссертациялық жұмыс бойынша жарияланымдар: Диссертациялық зерттеу нәтижелері 5 ғылыми жұмыста жарияланды, оның ішінде 2 мақала – Scopus деректер базасында индекстелетін журналдарда, 2 мақала – Қазақстан Республикасы Ғылым және жоғары білім министрлігінің Ғылым және жоғары білім саласындағы сапаны қамтамасыз ету комитеті ұсынған басылымдарда және 1 жұмыс – халықаралық ғылыми конференция материалдарында:

Scopus деректер базасында индекстелетін журналдарда жарияланған ғылыми мақалалар:

1. Shaizat M., Mussiraliyeva S., Tereikovskyi I. Forensically Interpretable Graph Descriptors for Improved Illicit Bitcoin Transaction Detection // International Journal of Computer Network and Information Security. – 2026. – Vol. 18, No. 3. – P. 63–84. Scopus деректері бойынша журналдың процентилі – 72.

2. Shaizat M., Mussiraliyeva S. Enhanced Identification of Illicit Bitcoin Transactions Through Genetic Algorithm-Based Feature Selection // Eastern-European Journal of Enterprise Technologies. – 2025. – Vol. 4, No. 9 (136). – P. 34–42. Scopus деректері бойынша журналдың процентилі – 57.

Ғылым және жоғары білім саласындағы сапаны қамтамасыз ету комитеті ұсынған басылымдарда жарияланған ғылыми мақалалар:

3. Мусиралиева Ш. Ж., Шайзат М. Ж., Бекетова А. К., Абайұлы Е., Манасова А. Б. О методе идентификации подозрительных транзакций в Биткоин сети / Identification of Suspicious Transactions in the Bitcoin Network // Известия Национальной академии наук Республики Казахстан. Серия физико-математическая. – 2023. – № 1 (345). – С. 154–169.
4. Мусиралиева Ш. Ж., Шайзат М. Ж., Бекетова А. К., Манасова А. Б. Биткоин желісінде күдікті транзакцияларды анықтау: талдау, белгілер және машиналық оқыту алгоритмдері // Вестник Национальной инженерной академии Республики Казахстан. – 2023. – № 3 (89).

Халықаралық конференциялар материалдарында жарияланған ғылыми жұмыстар:

5. Елтай Ж., Шайзат М., Бекетова А. Биткоин желісіндегі ауытқуларды анықтау әдісі // Материалы Международной научной конференции студентов и молодых ученых «Фараби әлемі». – Алматы, Казахстан, 2022. – С. 42.

Диссертацияның көлемі мен құрылымы

Диссертация кіріспеден, төрт бөлімнен, қорытындыдан және пайдаланылған дереккөздер тізімінен тұрады. Бірінші бөлімде Bitcoin-транзакцияларын киберкриминалистикалық талдаудың негіздері, қолданыстағы әдістер мен деректер жиындары қарастырылған. Екінші бөлімде генетикалық алгоритмге негізделген анонимдендірілген белгілерді іріктеу әдісі әзірленіп, кестелік және графтық модельдер салыстырылған. Үшінші бөлімде семантикасы ашық белгілері бар деректер жиыны қалыптастырылып, интерпретацияланатын дескрипторлар әзірленген және машиналық оқыту модельдері бағаланған. Төртінші бөлімде Bitcoin-транзакцияларын киберкриминалистикалық талдаудың гибридті архитектурасы әзірленіп, оның веб-қосымша түріндегі іске асырылуы сипатталған. Жұмыстың жалпы көлемі 41 сурет пен 20 кестені қоса алғанда, 114 бетті құрайды. Пайдаланылған дереккөздер тізімі 104 атауды қамтиды.

Автордың жеке үлесі

Автордың жеке үлесі ғылыми әдебиеттерді талдаудан, белгілерді қалыптастыру және іріктеу әдістерін әзірлеуден, деректер жиынын дайындаудан, машиналық оқыту модельдерін құрудан, ұсынылған тәсілді бағдарламалық іске асырудан және алынған нәтижелерді интерпретациялаудан тұрады. Сондай-ақ автор диссертация тақырыбы бойынша ғылыми жарияланымдар дайындады.