УДК 004.9
https://doi.org/10.47533/2023.1606-146X.20

Ш. Ж. МУСИРАЛИЕВА, М. Ж. ШАЙЗАТ, А. К. БЕКЕТОВА, А. Б. МАНАСОВА*

Әл-Фараби атындағы Қазақ ұлттық университеті,
«Ақпараттық қауіпсіздік жүйесі» мамандығы, Алматы, Қазақстан
E-mail: manassova.akerke4493@gmail.com

БИТКОИН ЖЕЛІСІНДЕ КҮДІКТІ ТРАНЗАКЦИЯЛАРДЫ АНЫҚТАУ: ТАЛДАУ, БЕЛГІЛЕР ЖӘНЕ МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМДЕРІ

Бұл мақалада қазіргі кездегі ең танымал криптовалютаның бірі – биткоин туралы мәліметтер айтылады. Биткоин мен блокчейн технологияларының жұмыс қағидалары қарастырылады, оның ішінде Bitcoin-нің артықшылықтары мен кемшіліктері көрсетілген. Жұмыстың негізгі бағыты криптовалюталарды, атап айтқанда биткоинді заңсыз мақсаттарда пайдалануды талдауға бағытталған. Биткоин желісіндегі күдікті транзакцияларды анықтау үшін белгілер жинағын анықтауға ерекше назар аударылған. Зерттеу жұмысында көптеген ғылыми жарияланымдар мен мақалаларға шолу жасалып, талдау жүргізілді және осы зерттеулердің нәтижесінен бағалау кезінде Биткоин күдікті операцияларға арналған транзакциялардың негізгі атрибуттары болып табылатын белгілер анықталды. Күдікті биткоин транзакцияларды анықтайтын модельді оқыту үшін 100 000-ға жуық транзакциясы бар 15 белігден тұратын кіріс деректер жинағы (датесет) жасалды. Деректер жинағы қалыптасқаннан кейін кездейсоқ орман, логистикалық регрессия, k-ең жақын көршілер, шешім ағашы сияқты бірнеше машиналық оқыту алгоритмдеріне негізделген биткоин желісіндегі күдікті транзакцияларды анықтау үшін модельдер жасалды және оқытылды. Сонымен қатар, мақалада машиналық оқыту алгоритмдерінің салыстырмалы талдауының нәтижелері, сондай-ақ жоғары дәлдікті қамтамасыз ететін алгоритм таңдалған.

Түйін сөздер*: криптовалюта, блокчейн, биткоин, күдікті транзакциялар, белгілер.*

Әдебиетке шолу. Қазіргі таңда технологияның қарқынды дамуына байланысты әлемде криптовалюта деп аталатын цифрлық валюталардың жаңа формалары пайда болуда. Аталған криптовалюталардың ерекшеліктерінің бірі – олар банктер сияқты дәстүрлі қаржы институттарының қатысуынсыз жұмыс істей алады. Ерекше қарқынмен таралып жатқан криптовалюталардың бірі және ең танымалысы ол – биткоин. Биткоин 2009 жылы құрылған алғашқы сәтті енгізілген криптовалюта болып табылады [1]. Биткоин бағдарламалық хаттаманы қолданатын және қатысушыларға валютаны бір-бірімен тікелей айырбастауға мүмкіндік беретін орталықтандырылмаған жүйе негізінде жұмыс жасайды. Биткоинді құрудағы негізгі мақсат ашық цифрлық төлем жүйесін әзірлеу болды. Биткоиннің басты артықшылығы – ол аударымдарды жүзеге асыру үшін өзін-өзі қамтамасыз ететін алғашқы, сенімді және орталықтандырылмаған жүйені ұсынды [2].

Биткоиндер «Майнинг» деп аталатын процедураны қолдана отырып, сандарды кэштеуге байланысты күрделі тапсырмаларды орындау арқылы математикалық түрде жасалады [3]. Биткоиндерді өндірумен айналысатын адамдарды майнерлер деп атайды. Майнерлер транзакциялар журналын өңдейді және олардың заңдылығын растайды. Қазіргі кезде Bitcoin транзакциясын тексерудің орташа уақыты шамамен алты минутты құрайды, ал әлем бойынша жасалып жатқан барлық транзакциялардың жар-

* E-mail корреспондирующего автора: manassova.akerke4493@gmail.com

тысына жуығы дерлік осы аталған алты минут немесе одан да аз уақыт ішінде расталады [4].

Биткоин желісіндегі барлық транзакциялар блокчейн деп аталатын таратылған деректер кітабында жазылады. Блокчейн – бұл қауіпсіз деректер алмасу үшін жаңа тұжырымдамаларды енгізетін революциялық технология. Ол жалпы және таратылған желілерді пайдалана отырып, барлық аяқталған транзакцияларды қауіпсіз сақтауға мүмкіндік беретін блоктар тізбегінен тұрады [5].

Биткоин транзакциясын жүзеге асыру үшін келесі қадамдарды орындау қажет:

– Биткоиндерді сақтауға арналған әмиян жасаңыз.

– Алушының мекен-жайын анықтаңыз және аударым сомасын енгізіңіз.

– Арнайы растау коды арқылы транзакцияны растаңыз.

– Майнерлерден транзакцияның расталуын күтіңіз.

– Транзакция расталғаннан кейін ол блокчейнге қосылады және қаражат алушының әмиянына аударылады.

Шынында, блокчейн – бұл таратылған мәліметтер базасы, мұнда әр қатысушыда тізбектелген блоктар ретінде сақталатын барлық деректердің көшірмелері болады. Әрбір блокта бүкіл тізбектің тұтастығын қамтамасыз ететін алдыңғы Блок туралы ақпарат бар. Әрбір блокта деректерді ауыстырудан қорғайтын алдыңғы блоктың хэші бар. Сондай-ақ, блокчейн деректердің қауіпсіздігін қамтамасыз ету және транзакциялардың түпнұсқалығын растау үшін криптографиялық алгоритмдерді пайдаланады. Сонымен қатар, блокчейн процестерді автоматтандыратын және мәмілелердің сенімділігі мен қауіпсіздігін қамтамасыз ететін смарт-келісімшарттар жасауға мүмкіндік береді. Осы мүмкіндіктердің арқасында блокчейн әртүрлі процестердің, соның ішінде қаржылық транзакциялардың, сатып алулардың, логистиканың, дауыс берудің және басқалардың қауіпсіздігін, ашықтығын және сенімділігін қамтамасыз етудің ең перспективалы технологияларының біріне айналды.

Делдалсыз жылдам және арзан транзакцияларды жүзеге асыру мүмкіндігі жеке пайдаланушыларды да, жосықсыз пайдаланушыларды да тартады. Бірақ, биткоин негізделген блокчейн технологиясы, жазбаларды өзгерту мүмкін еместігін және транзакциялардың ашықтықтығын қамтамасыз ететінін атап өткен жөн, өз кезегінде бұл кәдімгі ақша аударымдарының дәстүрлі әдістерімен салыстырғанда қылмыстық әрекеттерге онша жол бермейді.

Сонымен қатар, үкімет пен биржа сияқты реттеушілер бар, олар транзакцияларды бақылай алады және қылмыстық мақсатта жасалған күдікті операцияларды бұғаттай алады. Және транзакцияларды талдау және күдікті операцияларды анықтау арқылы қылмыстық әрекеттердің алдын алуға көмектесетін мамандандырылған компаниялар мен қызметтер бар.

Жалпы, блокчейн технологиясы және криптовалюталар, соның ішінде биткоин, жақсы және күмәнді мақсаттарда пайдаланылуы мүмкін. Бірақ технологияны дұрыс пайдалану көптеген артықшылықтар әкелуі мүмкін екенін есте ұстаған жөн, соның ішінде шығындарды азайту және әртүрлі салалардағы транзакциялардың тиімділігін арттыру.

Расымен де, криптовалюталарды, соның ішінде биткоинді пайдалану заңсыз операцияларға кейбір артықшылықтар береді және сонымен бірге үкіметтер мен

реттеушілердің мұндай операциялармен күресу мүмкіндіктерін қиындатады. Дегенмен, бұл мәселені шешудің бірнеше тәсілдері бар.

Біріншіден, кейбір елдер криптовалюта операцияларын бақылауға, соның ішінде пайдаланушыларды міндетті түрде сәйкестендіруге және транзакцияларды бақылауға бағытталған реттеуші шараларды әзірлеп, енгізді. Бұл заңсыз операцияларды жүзеге асыру мүмкіндіктерін шектеп, криптовалюта экономикасында ашықтықты қамтамасыз етуі мүмкін.

Екіншіден, заңсыз криптовалюта операцияларын анықтауға және бақылауға көмектесетін қызметтер бар. Мұндай қызметтерді үкіметтер мен реттеушілер күдікті транзакцияларды анықтау және заңсыз әрекеттердің жолын кесу үшін пайдалана алады.

Нәтижесінде, кейбір криптовалюта жобалары қазірдің өзінде заңсыз транзакциялармен күресу құралдарын әзірлеуде. Мысалы, кейбір жобалар блокчейнді қаржыландыру көздерін және қайырымдылық мақсаттарда аударылатын қаражаттарды бақылау үшін пайдаланады, бұл террористік ұйымдарды және басқа да заңсыз мақсаттарды қаржыландыру үшін криптовалюталарды пайдалану мүмкіндігін азайтады.

Осылайша, криптовалюталардың, соның ішінде биткоиннің заңсыз транзакциялар үшін кейбір артықшылықтары болса да, үкіметтер мен реттеушілер олармен күресудің бірқатар тәсілдерін қолдана алады.

Күдікті операцияны анықтау үшін белгілер жинағын анықтау. Бұл бөлімде криптовалюта саласындағы күдікті әрекеттердің белгілері, атап айтқанда биткоин желісінде немесе құқық қорғау органдарының әшкерелеуінен жалтару әрекеттері бар.

Ғылыми жарияланымдар мен мақалаларға шолу мен талдау, әдебиеттерге шолу және ашық дереккөздерді зерттеу барысында бағалау кезінде негізгі атрибуттар болып табылатын келесі белгілер анықталды сонымен қатар күдікті немесе күмәнді транзакциялар үшін Bitcoin транзакциясын бағалау кезінде негізгі белгілері болып табылады:

• Көптеген биткоиндерді ондаған және жүздеген транзакцияларға бөлудің көптеген операциялары көбінесе тарихы жоқ әмияндар арқылы орындалады, яғни жақында ашылған немесе қазірдің өзінде әрекетсіз болып саналатын биткоин әмияндары, басқаша айтқанда, оларды бір күндік әмияндар немесе транзиттік әмияндар деп атауға болады. Егер бәрі болмаса да, осы типтегі биткоин әмияндарының көпшілігі тек 2 рет қолданылады. Бірінші рет биткоиндерді (ақшаны) алу үшін, екінші рет алынған биткоиндерді (ақшаны) басқа әмиянға жіберу үшін. Мұндай операциядан кейін биткоин әмияндары пайдалануды тоқтатады. Бір пайдаланушыда көптеген биткоин әмияндары болуы мүмкін;

• Криптовалютаны аудару немесе айырбастау операциялары, қағаз ақшасы жағдайындағыдай, операцияларды міндетті тіркеу немесе хабарламалар беру үшін белгіленген шекті мәндерден аспайтын шағын сомаларға немесе ірі сомаларға бөлу арқылы жүзеге асырылады. Толығырақ айтатын болсақ, биткоиндердің үлкен санын көптеген аз мөлшердегі биткоиндерге бөлу және биткоиндерді кейінгі алушыларға аудару.

• Бөлінген биткоиндер транзакцияның қысқа тізбегінен кейін шамамен 5-6 транзакция биткоиндер бір әмиянға қайта жиналады;

• Биткоиндерді әртүрлі әмияндарға жіберу өрт кезіндегі тәртіппен және өте қысқа мерзімде, мысалы, 30 минутпен 1 сағат арасында жүзеге асырылады;

• Комиссияны қоскanда алынған сомма және жіберілген сомма бірдей сомма болып табылады. Өйткені, мұндай транзакцияларды жүргізетін адамдар бұл әмияндарды енді пайдаланбайтынын біледі, сондықтан олар барлық ақшаны (биткоин) аударады;

• Биткоинді криптобиржаға енгізу, содан кейін кез-келген нәрсені сатып алу немесе басқа әмиянға аудару сияқты қосымша операцияларсыз биржадан жылдам шығарып алу.

Жоғарыда аталған белгілердің барлығы ғылыми жарияланымдар мен мақалалардан алынған оннан астам практикалық мысалдарды талдау барысында анықталды. Бірден атап өткім келеді, белгілердің біреуінің болуы транзакцияны заңсыз немесе қылмыстық деп санауға негіз болмайды. Көбінесе ықтимал қылмыстық әрекетке күдік қарапайым пайдаланушының мінез-құлқы тұрғысынан қисынды негіздемесі жоқ транзакцияларда бірнеше белгілердің болуын тудырады. Белгілердің болуы одан әрі мониторинг пен егжей-тегжейлі талдау жүргізуге түрткі болуы керек.

Датасеттің қалыптасуы және сипаттамасы. Биткоин үшін күдікті транзакциялардың негізгі белгілерін анықтағаннан кейінгі келесі кезең, деректермен жұмыс: деректерді іздеу және жинау, жиналған деректерді талдау және өңдеу, модельді оқыту үшін датасет кіріс белгілерін қалыптастыру.

Зерттеу жұмысының шегінде биткоинді ұрлауға, биткоинді жылыстатуға байланысты белгілі жағдайларды зерттеу арқылы 5 мыңға жуық заңсыз транзакциялардан тұратын датасет құрылды. Жоғарыда көрсетілген мысалдардың бірі бұл 2018 жылдың қыркүйегінде орын алған Bitcoin ұрлығы, хакерлер Zaif биржасының әмиянына рұқсатсыз кіріп, 5966 Bitcoin ұрлаған [7]. Сонымен қатар, деректер жинағындағы транзакциялар санын ұлғайту үшін таңбаланған транзакциялары бар әлемдегі ең үлкен деректер жинағы «Elliptic DataSet» биткоиндік транзакциялар саласындағы танымал деректер жинағы зерттелді. Деректер жиынтығы жалпы сомасы 6 миллиард долларды құрайтын 200 000 транзакцияны қамтиды. Бұл деректер жинағы қауымдастықтарға жиналған деректерді криптовалюталардағы қаржылық қылмысты анықтау үшін пайдалануға мүмкіндік беру мақсатында жасалған [10].

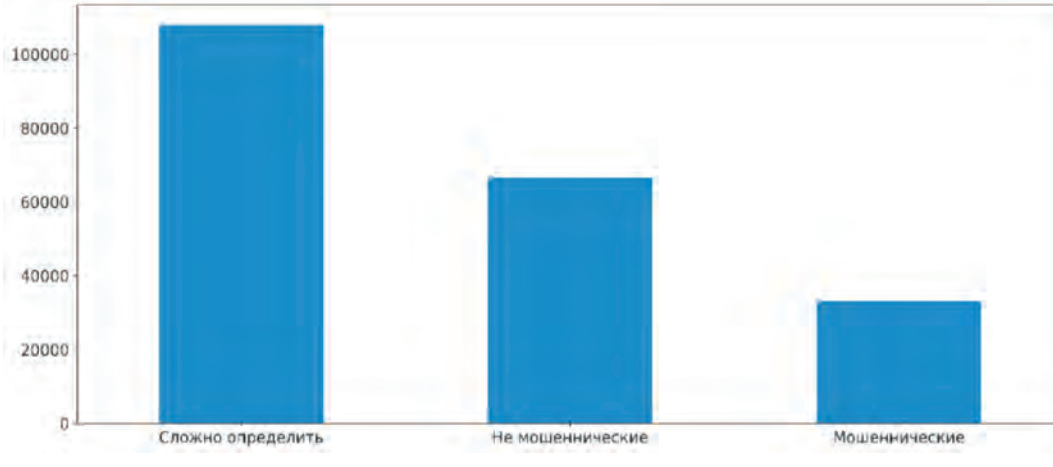
Эллиптикалық деректер жиынтығы Bitcoin транзакцияларын «заңсыз емес» (алаяқтық, зиянды бағдарламалар, террористік ұйымдар, төлем бағдарламалары, Понци схемалары және т.б.) санатындағы объектілермен, заңды санатқа жататын нақты объектілермен (биржалар, әмиян жеткізушілері, кеншілер, заңды қызметтер және т. б.) салыстырады, сондай-ақ белгісіз санаты да бар. Олардың пайыздық мөлшерлемесі: 2% заңсыз, 21% заңды және 77% белгісіз деп белгіленді.Elliptic деректер жинағы теңгерілмсіз (4 545 «заңсыз» және 42 019 «заңды» транзакциялар) деген маңызды бақылау жасалды. Сондықтан, келесі тапсырма мәмілелердің жоғарыда аталған элементтерінің мәнін табу ғана емес, сонымен қатар белгісіз операциялар үшін мәмілелерді заңды немесе заңсыз деп анықтау әрекеті болды.

Биткоин транзакциясының кіріс деректерін 1-кестеден көруге болады: 'transaction', ‘weight’, 'block_hash', 'block_height', 'block_time', 'confirmations', 'count_in', 'count_out', 'fiat_rate', 'hash', 'pool_time', 'size', 'total_in', 'total_out', . 'risk_score'.

Кесте 1 – Кіріс деректер

Атрибут атауы	Деректер түрі	Сипаттама
TRANSACTION	STRING	Белгілі бір транзакцияны анықтау үшін қолданылатын бірегей идентификатор
BLOCK_HASH	STRING	Хэш блогы
BLOCK_HEIGHT	INTEGER	Блоктың биіктігі
CONFIRMATIONS_	INTEGER	Желідегі транзакцияны растау
COUNT_IN	INTEGER	Алынған транзакциялардың жалпы саны
COUNT_OUT	INTEGER	Жіберілген транзакциялардың жалпы саны
FIAT RATE	FLOAT	Транзакциялар кезінде АҚШ долларындағы ақша
INPUTS	STRING	BTC жіберген мекенжайлар
OUTPUTS	STRING	BTC алған мекенжайлар
POOL_TIME	TIMESTAMP	Транзакцияны растау уақыты
TOTAL_IN	NUMERIC	Жіберілген BCT жалпы саны
TOTAL_OUT	NUMERIC	Алынған BTC жалпы саны
SIZE	INTEGER	Бұл транзакцияның жалпы мөлшері
WEIGHT	INTEGER	Бұл транзакцияның салмағы
RISK_SCORE	BOOLEAN	Күдікті транзакцияларды бағалау

Модельді құру және оқыту. Датасетті дайындағаннан кейін келесі кезең модельді оқыту болды. Модельді оқыту үшін кездейсоқ орман, шешім ағаштары, логистикалық регрессия сияқты бірнеше машиналық оқыту алгоритмдері қолданылды. Оқыту және сынақ үлгілеріне бөлу 80% - дан 20% - ға дейінгі арақатынасқа негізделген. Нәтижесінде оқыту үшін 1 суретте көрсетілгендей 79 618 кездейсоқ элементтер пайдаланылды, ал сынақ үшін 19905 элемент қолданылды.



Сурет 1 – Әрбір санаттың саны

Модельді оқытудың ерекшеліктерін анықтау үшін мәндер арасындағы байланысты немесе корреляция анықталды.

2-суретте барлық транзакциялар үшін барлық мүмкіндіктер арасындағы корреляция көрсетілген. Кейбір белгілердің қатты тәуелділігі, сондай-ақ block_height, block_time, fiat_rate, pool_time, confirmations сияқты теріс тәуелділігі бар екенін көруге болады.

	risk_score	block_height	block_time	confirmations	count_in	count_out	fiat_rate	pool_time	size	total_in	total_out	weight
risk_score	1.00	0.04	0.05	-0.05	0.03	0.09	0.08	0.05	0.07	0.02	0.02	0.05
block_height	0.04	1.00	1.00	-1.00	0.15	0.08	0.79	1.00	0.15	-0.02	-0.02	0.09
block_time	0.05	1.00	1.00	-1.00	0.15	0.08	0.79	1.00	0.15	-0.02	-0.02	0.09
confirmations	-0.05	-1.00	-1.00	1.00	-0.15	-0.08	-0.79	-1.00	-0.15	0.02	0.02	-0.09
count_in	0.03	0.15	0.15	-0.15	1.00	0.08	0.18	0.15	0.85	0.00	0.00	0.83
count_out	0.09	0.08	0.08	-0.08	0.08	1.00	0.09	0.08	0.55	0.03	0.03	0.55
fiat_rate	0.08	0.79	0.79	-0.79	0.18	0.09	1.00	0.79	0.18	-0.02	-0.02	0.11
pool_time	0.05	1.00	1.00	-1.00	0.15	0.08	0.79	1.00	0.15	-0.02	-0.02	0.09
size	0.07	0.15	0.15	-0.15	0.85	0.55	0.18	0.15	1.00	0.02	0.02	0.99
total_in	0.02	-0.02	-0.02	0.02	0.00	0.03	-0.02	-0.02	0.02	1.00	1.00	0.02
total_out	0.02	-0.02	-0.02	0.02	0.00	0.03	-0.02	-0.02	0.02	1.00	1.00	0.02
weight	0.05	0.09	0.09	-0.09	0.83	0.55	0.11	0.09	0.99	0.02	0.02	1.00

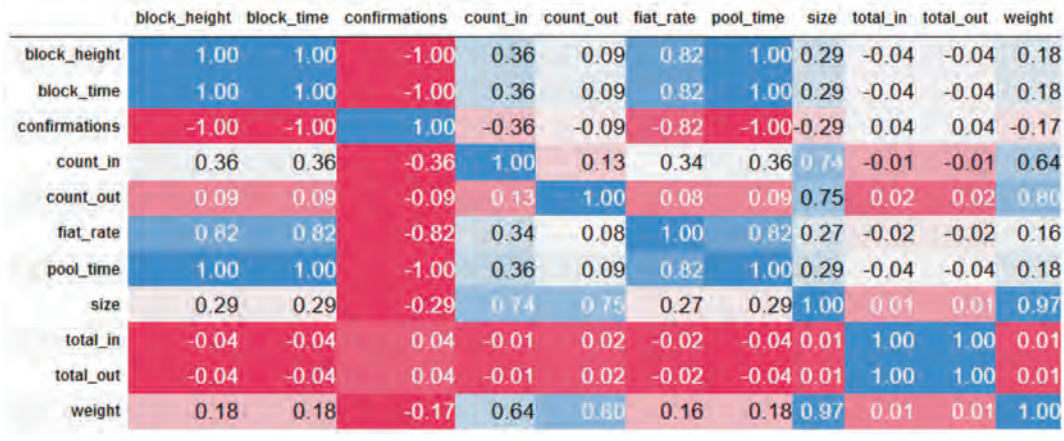
Сурет 2 – белгілер арасындағы корреляция

3-суретте таза транзакциялар үшін барлық белгілер арасындағы корреляция көрсетілген.

	block_height	block_time	confirmations	count_in	count_out	fiat_rate	pool_time	size	total_in	total_out	weight
block_height	1.00	1.00	-1.00	-0.04	-0.09	0.83	1.00	-0.04	-0.05	-0.05	-0.04
block_time	1.00	1.00	-1.00	-0.04	-0.09	0.83	1.00	-0.04	-0.05	-0.05	-0.04
confirmations	-1.00	-1.00	1.00	0.05	0.09	-0.83	-1.00	0.05	0.04	0.04	0.05
count_in	-0.04	-0.04	0.05	1.00	0.03	-0.03	-0.04	0.97	0.01	0.01	0.97
count_out	-0.09	-0.09	0.09	0.03	1.00	-0.05	-0.09	0.05	0.02	0.02	0.05
fiat_rate	0.83	0.83	-0.83	-0.03	-0.05	1.00	0.83	-0.04	-0.06	-0.06	-0.04
pool_time	1.00	1.00	-1.00	-0.04	-0.09	0.83	1.00	-0.04	-0.05	-0.05	-0.04
size	-0.04	-0.04	0.05	0.97	0.05	-0.04	-0.04	1.00	0.01	0.01	1.00
total_in	-0.05	-0.05	0.04	0.01	0.02	-0.06	-0.05	0.01	1.00	1.00	0.01
total_out	-0.05	-0.05	0.04	0.01	0.02	-0.06	-0.05	0.01	1.00	1.00	0.01
weight	-0.04	-0.04	0.05	0.97	0.05	-0.04	-0.04	1.00	0.01	0.01	1.00

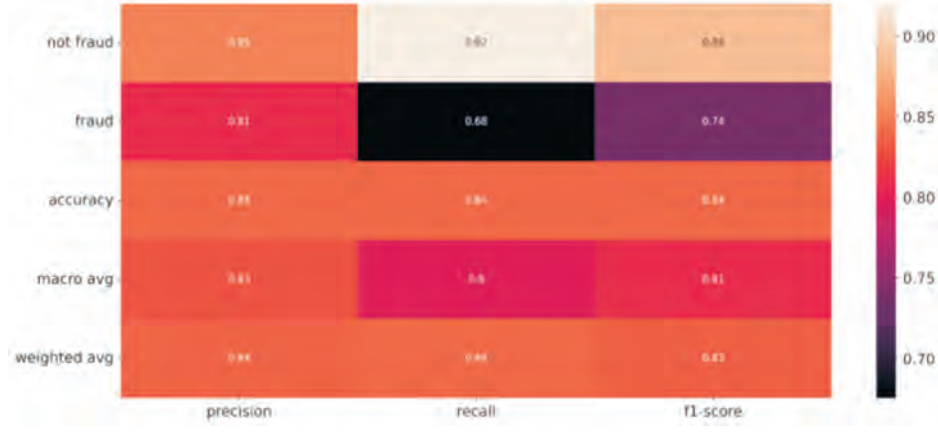
Сурет 3 – алаяқтық емес транзакциялар үшін белгілер арасындағы корреляция

4-суретте алаяқтық транзакциялар үшін барлық белгілер арасындағы корреляция көрсетілген.



Сурет 4 – алаяқтық транзакциялар үшін белгілер арасындағы корреляция

5-ші суретте кездейсоқ орман алгоритмі таңдалды. Кездейсоқ орман өте жақсы нәтиже көрсетеді. Дәлдік шамамен 84 пайызды құрайды.



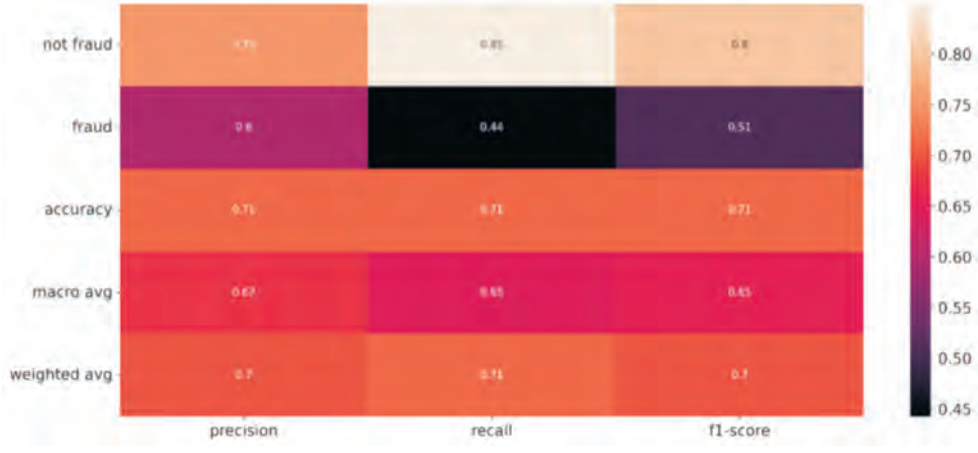
Сурет 5 – Кездейсоқ орман алгоритмінің жіктелуі туралы есеп

Реттелген кездейсоқ орман алгоритмнің дәлдігін арттыру үшін кесте 2-гі параметрлер таңдалады және қолданылды:

Кесте 2 – Кездейсоқ орман алгоритмінің параметрлері

'n_estimators':,	1757
'min_samples_split':	5
'min_samples_leaf':	2
'max_features':,	auto
'max_depth':	200
'criterion':	entropy
'bootstrap':	True

Ал 6 суреттен логистикалық регрессия алгоритмі функционалды алгоритмдер арасындағы ең нашар нәтижені көрсетеді, 66 пайызы транзакцияны анықтауға ғана жарамды.

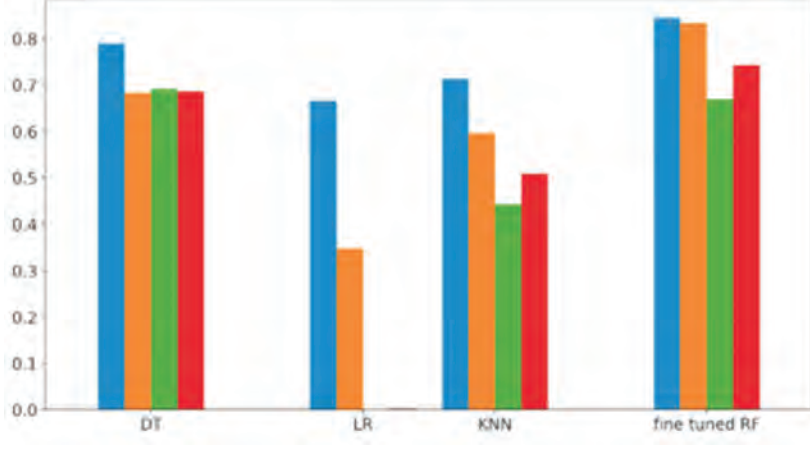


Сурет 6 – Логистикалық регрессия алгоритмінің жіктелуі туралы есеп

Кесте 3 – Нәтижелер және негізгі көрсеткіштерді салыстыру.

№	Name	Accuracy	Precision	Recall	F1-score
1	Decision Tree	0.789249	0.681771	0.691323	0.686514
2	Logistic Regression	0.665845	0.347826	0.001205	0.002402
4	KNN	0.713618	0.595658	0.442302	0.507651
5	Fine-tuned Random Forest	0.844966	0.833677	0.669027	0.742332

7-ші суретте және 3-ші кестеде көрініп тұрғандай, барлық алгоритмдердің нәтижесін көруге болады.



Сурет 7 – Алгоритмдерді салыстыру

Жұмыс барысында 4 машиналық оқу алгоритмдерді қолданып, 4 модель жасалды. Ең жақсы нәтижені реттелген кездейсоқ орман алгоритмі көрсетті.

Қорытынды. Жүргізілген жұмыс барысында күдікті немесе күмәнді деп танылған транзакциялар үшін Bitcoin транзакциясын бағалауда негізгі белгілері анықталды және талданды.

Датасеттің теңгерімсіздігін болдырмау үшін биткоинді ұрлауға, биткоинді жылыстатуға байланысты белгілі жағдайларды зерттеу арқылы 5 мыңға жуық заңсыз транзакцияларды қамтитын датасет құрылды. Сондай-ақ «Elliptic DataSet» деректер жинағы зерттелді. 2 деректер жиынын біріктіру арқылы бір негізгі деректер жинағы қалыптасты. Сонымен қатар, Crystal және Blockchain платформаларына жүгіну арқылы 200 мың биткоин транзакциялары үшін «Python» бағдарламалау тілінде деректерді автоматты түрде жинауға арналған персер жазылды. Әр түрлі 4 машиналық оқыту алгоритмдерін қолдана отырып, күдікті транзакцияларды анықтау үшін 4 модель жасалды және оқытылды. Модельдердің әрқайсысы 15 параметрі бар 100 мыңға жуық деректерде оқытылды. Нәтижесінде кездейсоқ орман негізінде құрылған ең жоғары дәлдікке ие модель таңдалды. Бұл жұмыс келесі жоба аясында орындалды және бұл зерттеуді Қазақстан Республикасы Ғылым және жоғары білім министрлігінің Ғылым комитеті (ЖТН AP19676342) қаржыландырды.

ӘДЕБИЕТ

1 Nakamoto S. Bitcoin: A Peer-to-Peer Electronic Cash System. http://bitcoin.org/bitcoin.pdf (датао бращения 05.09.2022).

2 Что такое биткоин и как он работает? https://forklog.com/cryptorium/chto-takoe-bitkoin (дата обращения: 09.09.2022).

3 Что такое криптовалюта и как она применяется? https://www.kaspersky.ru/resource-center/definitions/what-is-cryptocurrency (дата обращения 05.09.2022).

4 Что такое Биткоин транзакции: как проверить транзакцию биткоин? https://phemex.com/ru/academy/what-is-bitcoin-transaction (дата обращения: 27.09.2022).

5 Салах, К.; Рехман, МХУ; Низамуддин, Н .; Аль-Фукаха, А. Блокчейн для ИИ: обзор и открытые исследовательские задачи. IEEE Access 2019, 7 , 10127–10149.

6 Ajello N. Fitting a Square Peg in a Round Hole: Bitcoin, Money Laundering, and the Fifth Amendment Privilege Against Self-Incrimination // Brooklyn Law Review. 2015. Volume 80. Issue 2.

7 Взломали биржу Zaif. https://incrypted.com/exchange-zaif-was-hacked (дата обращения: 23.09.2022).

8 Как отслеживаются транзакции в сети биткоина? https://forklog.com/cryptorium/kak-otslezhivayutsya-tranzaktsii-v-seti-bitkoina (дата обращения: 07.09.2022).

9 Crypto Firms Can't Outrun the Travel Rule. https://www.coindesk.com/layer2/2022/01/12/crypto-firms-cant-outrun-the-travel-rule/ (дата обращения: 06.10.2022).

10 Elliptic Data Set. Bitcoin Transaction Graph. https://www.kaggle.com/ellipticco/elliptic-data-set (датаобращения: 06.10.2022).

REFERENCE

1 Nakamoto S. Bitcoin: A Peer-to-Peer Electronic Cash System. http://bitcoin.org/bitcoin.pdf (Access date 05.09.2022).

2 Chto takoe bitkoini kak on rabotaet? [What is Bitcoin and how does it work?] https://forklog.com/cryptorium/chto-takoe-bitkoin (Access date: 09.09.2022).

3 Chto takoe kriptovalyuta i kak ona primenyaetsya? [What is cryptocurrency and how is it applied?] // https://www.kaspersky.ru/resource-center/definitions/what-is-cryptocurrency. (Access date 05.09.2022).

4 Chto takoe Bitkoin tranzakcii: Kak proverit' tranzakciyu bitkoin? [What are Bitcoin Transactions?: How to verify a Bitcoin transaction?] https://phemex.com/ru/academy/what-is-bitcoin-transaction (Access date: 27.09.2022).

5 Salah, K.; Rekhman, MHU; Nizamuddin, N.; Al'-Fukaha, A. Blokchejndlya II: obzor i otkrytye issledovatel'skie zadachi. [Blockchain for AI: Overview and open research tasks.] IEEE Access 2019, 7, 10127–10149.

6 Ajello N. Fitting a Square Peg in a Round Hole: Bitcoin, Money Laundering, and the Fifth Amendment Privilege Against Self-Incrimination // Brooklyn Law Review. 2015. Volume 80. Issue 2.

7 Vzlomali birzhu Zaif. [The Zaif exchange was hacked.] https://incrypted.com/exchange-zaif-was-hacked (Access date: 23.09.2022).

8 Kak otslezhivayutsya tranzakcii v seti bitkoina? [How are transactions tracked on the Bitcoin network?] https://forklog.com/cryptorium/kak-otslezhivayutsya-tranzaktsii-v-seti-bitkoina (Access date: 07.09.2022).

9 Crypto Firms Can't Outrun the Travel Rule. https://www.coindesk.com/layer2/2022/01/12/crypto-firms-cant-outrun-the-travel-rule/ (Access date: 06.10.2022).

10 Elliptic Data Set. Bitcoin Transaction Graph. https://www.kaggle.com/ellipticco/elliptic-data-set (Access date: 06.10.2022).

Ш. Ж. МУСИРАЛИЕВА, М. Ж. ШАЙЗАТ, А. К. БЕКЕТОВА, А. Б. МАНАСОВА

Казахский национальный университет имени аль-Фараби, специальность «Системы информационной безопасности»,Алматы, Казахстан
E-mail: manassova.akerke4493@gmail.com

ИДЕНТИФИКАЦИЯ ПОДОЗРИТЕЛЬНЫХ ОПЕРАЦИЙ В СЕТИ БИТКОИН: АНАЛИЗ, ПРИЗНАКИ И АЛГОРИТМЫ МАШИННОГО ОБУЧЕНИЯ

В данной статье дан анализ одному из самых популярных на сегодняшний день криптовалют – биткоину. Рассматриваются принципы и характер работы биткоина и блокчейн технологий, где перечисляются преимущества и недостатки биткоина. Внимание в работе акцентируется на анализе использования криптовалют, а именно биткоина в незаконных и преступных целях. Особое внимание уделяется идентификации подборки признаков для определения подозрительной деятельности в сети биткоина. В ходе исследовательской работы были проведены обзор и анализ множества научных публикаций и статьи, а из этих исследований были выявлены признаки, которые являются ключевыми атрибутами при оценке биткоин -транзакции на предмет подозрительных операций. На основе этих признаков был сформирован входной датасет из 15 атрибутов и около 100 000 транзакциях для дальнейшего использования в создании модели, которая будет оценивать биткоин транзакции на предмет подозрительности. С использованием этого набора данных были созданы и обучены модели для идентификации подозрительных операций в биткоин-сети на основе нескольких алгоритмов машинного обучения, такие как случайный лес, логистическая регрессия, Метод k-ближайших соседей, дерево решений. Также в статье приведе-

ны сравнения результатов алгоритмов машинного обучения, а также выбран лучший алгоритм, который показал лучшую точность.

Ключевые слова: *блокчейн, криптовалюта, биткоин, подозрительные транзакции, характеристики блокчейна.*

SH. MUSSIRALIYEVA[1], M. SHAIZAT[1], A. BEKETOVA[1], A. MANASSOVA[1]

[1]Al-Farabi Kazakh National University, specialty «Information Security Systems», Kazakhstan, 050040, Almaty, 71 Al-Farabi Avenue.
E-mail: manassova.akerke4493@gmail.com

IDENTIFICATION OF SUSPICIOUS TRANSACTIONS IN THE BITCOIN NETWORK: ANALYSIS, FEATURES, AND MACHINE LEARNING ALGORITHMS

This article talks about one of the most popular cryptocurrencies today - bitcoin. The principles of bitcoin and blockchain technologies are considered, where the proc and cons of bitcoin are listed. The focus of the research work is on the analysis of the use of cryptocurrencies, namely bitcoin for illegal and criminal purposes. Particular attention is paid to identifying a set of signs to identify suspicious activity in the bitcoin network. During research work reviewed and analyzed many scientific publications and articles, and from these studies, signs were identified that are key attributes when evaluating a Bitcoin transaction for suspicious or questionable transactions. Based on these features, an input dataset of 15 attributes with about 100,000 transactions was generated for further use in building a model that will identify suspicious bitcoin transactions. Using this dataset, models were created and trained to detect suspicious transactions in the bitcoin network based on several machine learning algorithms, such as random forest, logistic regression, k-nearest neighbors, decision tree. The article also provides comparisons of the results of machine learning algorithms, and in addition, selects the best algorithm that showed the best accuracy.

Key words*: Cryptocurrency, blockchain, bitcoin, suspicious transactions, features.*