TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI HUZURIDAGI ILMIY DARAJALAR BERUVCHI DSc.13/30.12.2019.T.07.02 RAQAMLI ILMIY KENGASH

TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI

AXMEDOVA NOZIMA FARXOD QIZI

KATTA HAJMLI MA’LUMOTLARNING HAYOTIY SIKLIDA AXBOROTNI HIMOYALASHNING MODEL VA ALGORITMI

05.01.05 – Axborotni himoyalash usullari va tizimlari. Axborot xavfsizligi

TEXNIKA FANLARI BO‘YICHA FALSAFA DOKTORI (PhD) DISSERTATSIYASI AVTOREFERATI

Toshkent – 2025

UO‘K 004.056.52

Texnika fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi avtoreferati mundarijasi

Оглавление автореферата диссертации доктора философии (PhD) по техническим наукам

Contents of dissertation abstract of doctor of philosophy (PhD) on technical science

TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI HUZURIDAGI ILMIY DARAJALAR BERUVCHI DSc.13/30.12.2019.T.07.02 RAQAMLI ILMIY KENGASH

TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI

AXMEDOVA NOZIMA FARXOD QIZI

KATTA HAJMLI MA'LUMOTLARNING HAYOTIY SIKLIDA AXBOROTNI HIMOYALASHNING MODEL VA ALGORITMI

05.01.05 – Axborotni himoyalash usullari va tizimlari. Axborot xavfsizligi

TEXNIKA FANLARI BO'YICHA FALSAFA DOKTORI (PhD) DISSERTATSIYASI AVTOREFERATI

Toshkent – 2025

Texnika fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasining mavzusi O‘zbekiston Respublikasi Oliy ta’lim, fan va innovatsiyalar vazirligi huzuridagi Oliy attestasiya komissiyasida B2025.1.PhD/T5293 raqam bilan ro‘yxatga olingan.

Dissertatsiya Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universitetida bajarilgan.

Dissertatsiya avtoreferati uch tilda (o‘zbek, rus, ingliz (rezyume)) Ilmiy kengash veb-sahifasida (www.tuit.uz) va «ZiyoNet» axborot-ta’lim portalida (www.ziyonet.uz) joylashtirilgan.

Ilmiy rahbar:	**Tashev Komil Axmatovich** texnika fanlar nomzodi, dotsent
Rasmiy opponentlar:	**Botirov Fayzulla Baxtiyorovich** texnika fanlar doktori, dotsent
	Ibroximov Azizbek Ravshanbek o‘g‘li texnika fanlari bo‘yicha falsafa doktori
Yetakchi tashkilot:	**“UNICON.UZ” MCHJ**

Dissertatsiya himoyasi Toshkent axborot texnologiyalari universiteti huzuridagi DSc.13/30.12.2019.T.07.02 raqamli Ilmiy kengashning 2025 yil “25”.oktabr soat 10:00 dagi majlisida bo‘lib o‘tadi. (Manzil: 100084, Toshkent shahri, Amir Temur ko‘chasi, 108-uy. Tel.: (+99871) 238-64-15; e-mail: info@tuit.uz).

Dissertatsiya bilan Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti Axborot-resurs markazida tanishish mumkin (369 – raqam bilan ro‘yxatga olingan).
(Manzil: 100084, Toshkent, Amir Temur ko‘chasi, 108-uy. Tel.: (+99871) 238-64-15).

Dissertatsiya avtoreferati 2025 yil “15” oktabrda tarqatildi.
(2025 yil “15” oktabrdagi 10 - raqamli reestr bayonnomasi).

B.Sh. Maxkamov
Ilmiy darajalar beruvchi ilmiy kengash raisi, iqtisodiyot fanlari doktori, professor

M.S. Saitkamolov
Ilmiy darajalar beruvchi ilmiy kengash ilmiy kotibi, iqtisodiyot fanlari doktori, dotsent

D.Ya. Irgasheva
Ilmiy darajalar beruvchi ilmiy kengash qoshidagi ilmiy seminar raisi, texnika fanlari doktori, professor

KIRISH (falsafa doktori (PhD) dissertatsiyasi annotatsiyasi)

Dissertatsiya mavzusining dolzarbligi va zarurati. Dunyo miqyosida katta hajmli ma’lumotlar iqtisodiy va ijtimoiy sohalarning deyarli barcha tarmoqlariga chuqur kirib borib, raqamli muhitda internet foydalanuvchilari uchun keng ko‘lamli imkoniyatlar yaratishga zamin hozirlamoqda. Ushbu texnologiyadan foydalanayotganlar sonining ortib borishi bilan birga unga nisbatan tahdidlar va hujumlar soni ham sezilarli darajada oshmoqda. “Terranova Security” kompaniyasi tomonidan berilgan ma’lumotlarga ko‘ra, 2023-yilning ikkinchi choragidan 2024-yilning ikkinchi choragigacha bo‘lgan davrda katta hajmli ma’lumotlar tizimlaridagi axborot xavfsizligining buzilishi asosan axborot tizimidan ruxsatsiz foydalanishlar natijasida yuzaga kelgan. Bunday holatlar ma’lumotlarning tizimdan noqonuniy chiqib ketishiga olib kelgan bo‘lib, ulardan 65 foizdan ortig‘i anonimlashtirilmagan shaklda bo‘lganligi aniqlangan[1]. Shuningdek, tadqiqot natijalari kelgusida katta hajmli ma’lumotlar tizimlarida axborot xavfsizligini oshirishga qaratilgan milliy strategiyalar va xalqaro hamkorlik tashabbuslarini yanada rivojlantirishda muhim ahamiyat kasb etadi.

Dunyo miqyosida katta hajmli ma’lumotlar hayotiy siklida axborotni himoyalashning tahdid modellari hamda kiberxavfsizlik tahdidlaridan himoyalanish mexanizmlarini ishlab chiqish bo‘yicha ilmiy tadqiqotlar olib borilmoqda. Jahonda katta hajmli ma’lumotlar doirasida axborot xavfsizligini buzilishiga olib keluvchi tizimga ruxsatsiz kirishlarning nazoratini kuchaytirish va himoya darajasini oshirish maqsadida foydalanishlarni boshqarish hamda ma’lumotlarni anonimlashtirishga asoslangan yangi model, usul va algoritmlarni ishlab chiqish dolzarb vazifa sifatida ilmiy-tadqiqot ishlari olib borilmoqda. Bu borada so‘nggi yillarda katta hajmli ma’lumotlar bilan bog‘liq axborot xavfsizligi masalalari muhim ilmiy yo‘nalishlardan biriga aylangan. Katta hajmli ma’lumotlar tizimlarining murakkab va ochiq muhitda faoliyat yuritishi, ularning zaif jihatlarini aniqlash hamda samarali himoya mexanizmlarini ishlab chiqishga alohida e’tibor berilmoqda.

Respublikamizda katta hajmli ma’lumotlarda foydalanishni boshqarish, atributga asoslangan foydalanishni boshqarish, axborot xavfsizligi risklarini boshqarish, anonimlikni ta’minlashga doir modellarning takomillashtirilishi bo‘yicha bir qator samarali ishlar amalga oshirilgan. Raqamli texnologiyalarning jadal rivojlanishi bilan bog‘liq holda katta hajmli ma’lumotlarda axborot xavfsizligini ta’minlash ustuvor yo‘nalishlardan biri hisoblanib, “Raqamli O‘zbekiston–2030” kiberxavfsizlik salohiyatini oshirish, fuqarolarning shaxsiy ma’lumotlarini himoya qilish, bank va to‘lov tizimlarida zamonaviy biometrik va firibgarlikka qarshi texnologiyalarni joriy etish hamda katta hajmli ma’lumotlarni anonimlashtirish talablari asosida yig‘ish va saqlash”[2] bo‘yicha aniq vazifalar belgilangan. Belgilangan vazifalarning samarali bajarilishini ta’minlashda katta hajmli ma’lumotlar hayotiy sikli doirasida axborotni himoyalash, ularning modeli va algoritmini ishlab chiqish, anonimlashtirish jarayonida ma’lumot yo‘qotilishini minimallashtirishga qaratilgan ishlarni amalga oshirish muhim ahamiyat kasb etadi.

[1] Terranova Security kompaniyasi ma’lumotlari

[2] O‘zbekiston Respublikasi Prezidentining 2020-yil 05-oktabrdagi ““Raqamli O‘zbekiston — 2030” strategiyasini tasdiqlash va uni samarali amalga oshirish chora-tadbirlari to‘g‘risida”gi PF-6079-son Farmoni

Shu bilan birga, O‘zbekiston Respublikasi Prezidentining 2022-yil 28-yanvardagi “2022–2026-yillarga mo‘ljallangan Yangi O‘zbekistonning Taraqqiyot strategiyasi to‘g‘risisda”gi PF-60-son Farmoni”, 2024-yil 21-iyundagi “Raqamli kriminalistika sohasida ilmiy-tadqiqot faoliyatini tashkil etish chora-tadbirlari to‘g‘risida”gi PQ-229-son Qarori hamda O‘zbekiston Respublikasi Davlat xavfsizlik xizmati raisining 2023-yil 4-sentabrdagi “Kiberxavfsizlik va muhim axborot infratuzilmasi obyektlarining kiberxavfsizligini ta’minlash darajasini baholash tartibi to‘g‘risida”gi 91-sonli buyrug‘i bilan tasdiqlangan nizomda va boshqa normativ-huquqiy hujjatlarda belgilangan vazifalarni amalga oshirishga mazkur dissertatsiya tadqiqoti ma’lum darajada xizmat qiladi.

Tadqiqotning respublika fan va texnologiyalari rivojlanishining ustuvor yo‘nalishlariga mosligi. Mazkur tadqiqot respublika fan va texnologiyalar rivojlanishining IV. “Axborotlashtirish va axborot-kommunikatsiya texnologiyalarini rivojlantirish” ustuvor yo‘nalishi doirasida bajarilgan.

Muammoning o‘rganilganlik darajasi. Katta hajmli ma’lumotlar xavfsizligini ta’minlashda foydalanishni boshqarish va ma’lumotlarni anonimlashtirishga qaratilgan yondashuvlar asosida ishlab chiqilgan modellar, usullar va algoritmlarni qo‘llash bo‘yicha Pietro Colombo, Jim Longstaff, Sofia Zebboudj, Joanne Noble, Rabah Brahami va boshqa olimlar tomonidan ilmiy izlanishlar olib borilmoqda. Katta hajmli ma’lumotlarda foydalanishni boshqarish uchun rollar, attributlar va matritsalar kabi usullarini hamda ma’lumotlarni anonimlashtirishda K-anonimlashtirish usullarini qo‘llagan holda Victor C. M. Leung, Domingo-Ferrer, U. Narayanan, S. Varshney, L. T. Yang, C. Liu kabi xorijiy olimlar tomonidan ilmiy izlanishlar olib borilgan bo‘lib, hozirda ular boshchiligidagi ilmiy maktablarning tadqiqotchilari va izlanuvchilari tomonidan davom ettirilmoqda. MDH davlatlari olimlaridan A.A. Klimov, A.V. Karpov, Y.A. Zaitsev kabilar ma’lumotlarni sirqib chiqishidan himoyalash bo‘yicha ilmiy izlanishlar olib borganlar.

Katta hajmli ma’lumotlarda axborot xavfsizligini ta’minlashga qaratilgan dasturiy mahsulotlarni ishlab chiqish va tajribadan o‘tkazish bo‘yicha Terranova Security, Eclipse, Cisco, Arkansas public safety solutions (APSS), One Identity tashkilotlarining yetakchi mutaxassislari tomonidan foydalanuvchilarning foydalanish vakolatlarini belgilash va nazorat qilish hamda ma’lumotlarni anonimlashtirishni ta’minlovchi dasturiy mahsulotlarni ishlab chiqishga ixtisoslashtirilgan ilmiy-amaliy tadqiqot ishlari olib borilmoqda.

O‘zbekistonda S.K. Ganiyev, M.M. Karimov, D.Y. Irgasheva, K.F. Kerimovlar boshchiligidagi ilmiy jamoalar, tomonidan kompyuter tarmoqlarida ruxsatsiz foydalanishlarni boshqarish, vakolatlarni taqsimlash va foydalanuvchilarning rollarini belgilash, ma’lumotlar bazasiga qaratilgan tahdid va hujum modellari, usullari va algoritmlarini takomillashtirish va yangi himoya usullarini ishlab chiqish bo‘yicha ilmiy izlanishlar olib borilgan.

Shu bilan birga katta hajmli ma’lumotlarning hayotiy siklida tahdid modeli, foydalanishlarni boshqarish, risklarni boshqarish va ma’lumotlarni anonimlashtirishda ularning yo‘qolishini kamaytirishga qaratilgan modellar, usullar va algoritmlar yetarlicha tadqiq etilmagan.

Dissertatsiya tadqiqotining dissertatsiya bajarilgan oliy ta’lim yoki ilmiy tadqiqot muassasasining ilmiy-tadqiqot ishlari rejalari bilan bog‘liqligi. Dissertatsiya tadqiqoti Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universitetining ilmiy-tadqiqot ishlari rejasining 101128871 - DEBSEUz-ERASMUS-EDU-2023-CBHE “Development of the targeted Educational program for Bachelors in Solar Energy in Uzbekistan” mavzusidagi hamda AL-662204323 “Gazlama va gilamlardagi fraktal tuzilishli milliy naqshlarning modellari, algoritmlari va dasturiy majmualarini ishlab chiqish” mavzusidagi loyihalar doirasida bajarilgan.

Tadqiqotning maqsadi katta hajmli ma’lumotlarning hayotiy siklida axborotning sirqib chiqishini va ruxsatsiz foydalanilishni oldini olishga imkon beruvchi himoyalashning model va algoritmlarini ishlab chiqishdan iborat.

Tadqiqotning vazifalari:

katta hajmli ma’lumotlar tizimida foydalanishni boshqarish modelini takomillashtirish;

katta hajmli ma’lumotlar tizimida foydalanishni boshqarish jarayonining arxitekturasini ishlab chiqish;

katta hajmli ma’lumotlar tizimida axborot xavfsizligi risklarini boshqarish modelini takomillashtirish;

katta hajmli ma’lumotlar tizimida axborotni konfidensialligini ta’minlashga imkon beruvchi anonimlashtirish algoritmini ishlab chiqish.

Tadqiqotning obyekti sifatida katta hajmli ma’lumotlar tizimi olingan.

Tadqiqotning predmetini katta hajmli ma’lumotlarning hayotiy siklida foydalanishni boshqarish, axborotni himoyalash modellari va algoritmlari tashkil etadi.

Tadqiqotning usullari. Tadqiqot jarayonida diskret matematika, tasodifiy jarayonlarni modellashtirish, ehtimollar nazariyasi, to‘plamlar nazariyasi, obyektga yo‘naltirilgan dasturlash, mashinali o‘qitish usullaridan foydalanilgan.

Tadqiqotning ilmiy yangiligi quyidagilardan iborat:

katta hajmli ma’lumotlar tizimida subyekt va obyekt atributlari to‘plamiga xavfsizlik atributlarini hamda bajarilishi talab etiladigan beshta shartlarni kiritish asosida foydalanuvchilarga vakolat taqdim etish aniqligini oshirishga imkon beruvchi boshqaruv modeli takomillashtirilgan;

katta hajmli ma’lumotlar tizimlarida himoya tizimini xavfsizlik talablari asosida boshqa tizimlar bilan integratsiya qilish imkonini beruvchi atributga asoslangan shifrlashning ochiq kalitlarni tarqatish siyosatini qo‘llovchi foydalanishni boshqarishning arxitekturasi ishlab chiqilgan;

katta hajmli ma’lumotlar tizimida yangi turdagi risklarni aniqlash, baholash va boshqarish jarayonida klassik noravshan matematik baholash usulini qo‘llash orqali xavfsizlikni ta’minlashning dinamik siklini hosil qilish imkonini beruvchi risklarni boshqarish modeli takomillashtirilgan;

boshlang‘ich massa markazini tanlash hamda o‘rtacha markaz usuliga ikki o‘rtachali klasterlash hamda “greedy” elementlarini kiritish asosida ma’lumotlarning foydalanuvchanligini saqlash imkonini beruvchi K-anonim-lashtirish algoritmi takomillashtirilgan.

Tadqiqotning amaliy natijalari quyidagilardan iborat:

katta hajmli ma’lumotlar tizimida foydalanishni boshqarishning xavfsizlik modellari, jumladan, atributga asoslangan takomillashtirilgan M-ABAC modeli asosida foydalanishni boshqarishning dasturiy vositasi ishlab chiqilgan;

katta hajmli ma’lumotlar tizimida, ikki o‘rtachali klasterlash hamda “greedy” algoritmlari asosida, anonimlashtirish algoritmining dasturiy vositasi ishlab chiqilgan.

Tadqiqot natijalarining ishonchliligi tanlab olingan tahdid modeliga ko‘ra o‘tkazilgan tahlillar, model va algoritmlarni amalga oshirishdan olingan natijalar, belgilangan sharoitda qo‘lga kiritilgan hisoblashlar natijalari bilan izohlanadi.

Tadqiqot natijalarining ilmiy va amaliy ahamiyati. Tadqiqot natijalarining ilmiy ahamiyati katta hajmli ma’lumotlarning hayotiy siklida tahdid modeli, atributlari va ularga qarashi chora mexanizmlarini tahlil qilish, atributlarga asoslangan foydalanishlarni boshqarish modelini hamda axborot xavfsizligi risklarini boshqarish modelini takomillashtirish, himoya tizimlarini bir-biriga integratsiya qilish imkonini beruvchi foydalanishlarni boshqarish arxitekturasini va ma’lumotlar maxfiyligini ta’minlovchi takomillashtirilgan K-anonimlashtirish algoritmini ishlab chiqish orqali axborotni himoyalash tizimlarining samaradorligini oshirishga xizmat qilishi bilan izohlanadi.

Tadqiqot natijalarining amaliy ahamiyati katta hajmli ma’lumotlar tizimida foydalanishlarini boshqarish modeli, axborot xavfsizligi risklarini boshqarish modeli va ma’lumotlarni xavfsizligini ta’minlash maqsadida takomillashtirilgan K-anonimlashtirish algoritmi asosida ishlab chiqilgan dasturiy vosita yordamida katta hajmli ma’lumotlarning hayotiy siklida axborotni himoyalash tizimlarini samaradorligini oshirishga imkon berishi bilan izohlanadi.

Tadqiqot natijalarining joriy qilinishi. Katta hajmli ma’lumotlar tizimida takomillashtirilgan modellar va taklif etilgan algoritmlar asosida ishlab chiqilgan dasturiy vositalar bo‘yicha olingan ilmiy natijalar asosida:

katta hajmli ma’lumotlar tizimida subyekt va obyekt atributlari to‘plamiga xavfsizlik atributlarini hamda bajarilishi talab etiladigan beshta shartlarni kiritish asosida atributlarga asoslangan foydalanishni boshqarishning takomillashtirilgan modeli bo‘yicha ishlab chiqilgan “BDSS dasturi” dasturiy vositasi “Kiberxavfsizlik markazi” davlat unitar korxonasining amaliy faoliyatiga joriy qilingan (Raqamli texnologiyalar vazirligining 2024-yil 6-noyabrdagi 24-8/7511-son ma’lumotnomasi). Ilmiy tadqiqot natijasi foydalanuvchilarga vakolat berish aniqligini 0,2% oshirish, ruxsat etilmagan vakolatlarni taqdim etish xatoligini 0,8 foizga kamaytirish imkonini bergan;

katta hajmli ma’lumotlar tizimlarida himoya tizimini xavfsizlik talablariga asosan boshqa tizimlar bilan integratsiya qilish uchun atributga asoslangan shifrlashning ochiq kalitlarni tarqatish siyosatini qo‘llovchi ishlab chiqilgan foydalanishni boshqarishning arxitekturasi asosida ishlab chiqilgan dasturiy vosita “Kiberxavfsizlik markazi” davlat unitar korxonasining amaliy faoliyatiga joriy qilingan (Raqamli texnologiyalar vazirligining 2024-yil 6-noyabrdagi 24-8/7511-son ma’lumotnomasi). Dissertatsiya ishida taklif etilgan modellar va algoritmlar asosida ishlab chiqilgan dasturiy vositadan foydalanish katta hajmli

ma’lumotlarda foydalanishni boshqarish jarayonida foydalanuvchiga taqdim etilgan har qanday noto‘g‘ri vakolatlar sababli yuzaga keluvchi axborot xavfsizligining tashkil etuvchilari hisoblangan maxfiylikka, butunlikka va foydalanuvchanlikka qaratilgan tahdidlardan himoyalash imkoniyati paydo bo‘lgan;

katta hajmli ma’lumotlar tizimida yangi turdagi risklarni aniqlash, baholash va boshqarish jarayonida klassik noravshan matematik baholash usulini qo‘llash orqali takomillashtirilgan risklarni boshqarishning ko‘p bosqichli modeli asosida ishlab chiqilgan “BDSS dasturi” dasturiy vositasi “Radioaloqa, radioeshittirish va televideniye markazi” davlat unitar korxonasining amaliy faoliyatiga joriy qilingan (Raqamli texnologiyalar vazirligining 2024-yil 6-noyabrdagi 24-8/7511-son ma’lumotnomasi). Ilmiy tadqiqot natijasida dasturiy vosita subyektlarga foydalanish vakolatini belgilash uchun mavjudlariga nisbatan 3 sekund kamroq vaqtni talab etgan;

katta hajmli ma’lumotlar tizimlarida axborotni sirqib chiqishini oldini olish va ma’lumotlarni foydalanuvchanligini saqlash maqsadida ikki o‘rtachali klasterlash hamda “greedy” elementlari asosida takomillashtirilgan K-anonimlashtirish algoritmining “Big Data security analyzer dasturi” dasturiy vositasi “UZINFOCOM Davlat axborot tizimlarini yaratish va qo‘llab-quvvatlash bo‘yicha yagona integrator” MCHJning faoliyatiga joriy qilingan (Raqamli texnologiyalar vazirligining 2024-yil 6-noyabrdagi 24-8/7511-son ma’lumotnomasi). Ilmiy tadqiqot natijasida dasturiy vosita anonim jadvallarni shakllantirish uchun mavjudlariga nisbatan 0,03 sekund kam vaqt talab etgan.

Tadqiqot natijalarining aprobatsiyasi. Mazkur tadqiqot natijalari 3 ta xalqaro va 7 ta respublika ilmiy-amaliy anjumanlarida muhokamadan o‘tkazilgan.

Tadqiqot natijalarining e’lon qilinganligi. Dissertatsiyaning mavzusi bo‘yicha jami 19 ta ilmiy ish chop etilgan, jumladan, O‘zbekiston Respublikasi Oliy attestatsiya komissiyasining dissertatsiyalarning asosiy ilmiy natijalarini chop etish tavsiya etilgan ilmiy nashrlarida 6 ta maqola nashr etilgan hamda EHM uchun yaratilgan 3 ta dasturiy vositalarni qaydlash guvohnomalari olingan.

Dissertatsiyaning tuzilishi va hajmi. Dissertatsiya tarkibi kirish, to‘rtta bob, xulosa, foydalanilgan adabiyotlar ro‘yxati va ilovalardan iborat. Dissertatsiya hajmi 112 betni tashkil etadi.

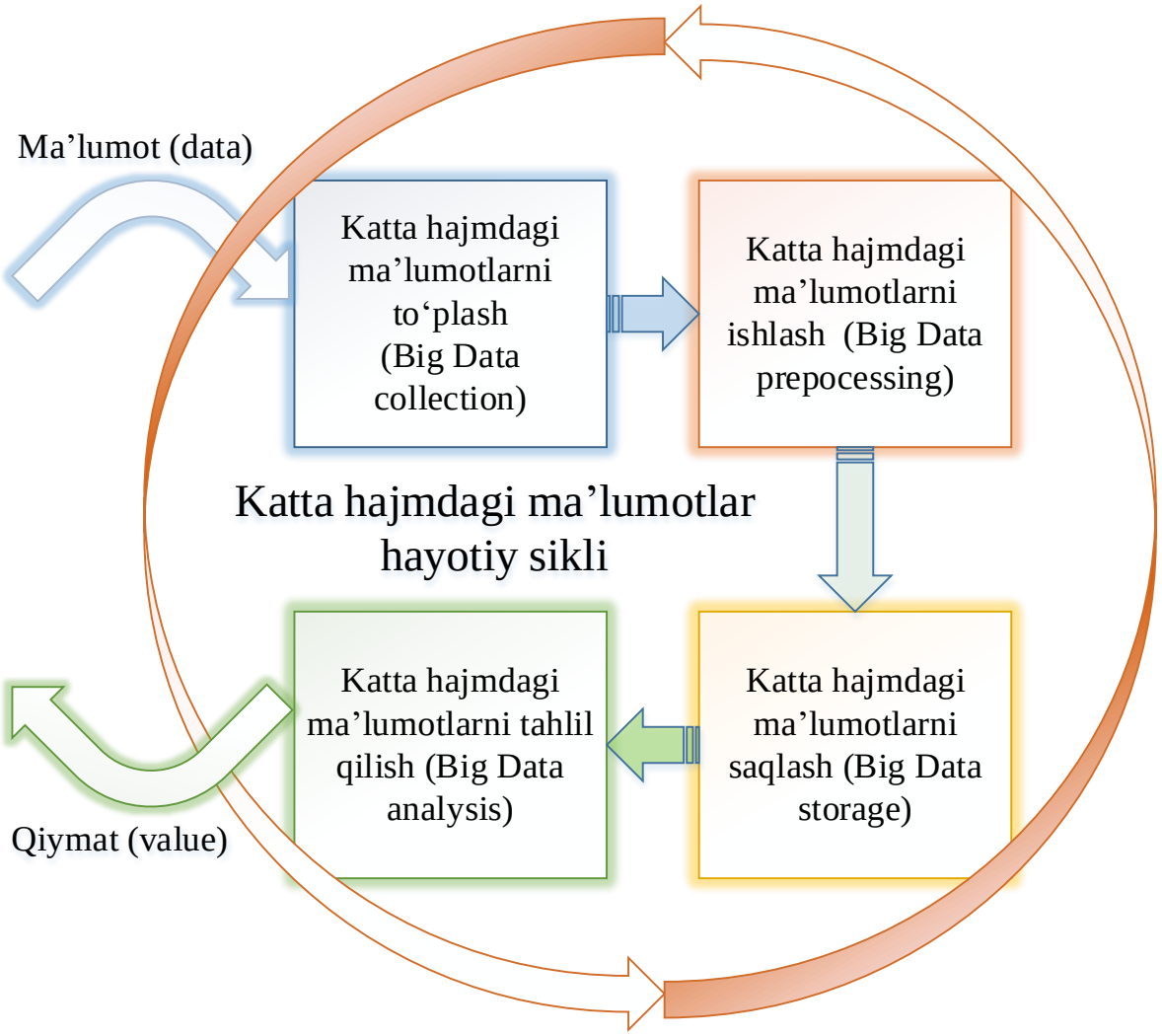
DISSERTATSIYA ISHINING ASOSIY MAZMUNI

Kirishda dissertatsiya mavzusining dolzarbligi va zarurati asoslangan, maqsad va vazifalar shakllantirilgan, tadqiqot obyekti va predmeti aniqlangan, tadqiqotning O‘zbekiston Respublikasi ilm-fan va texnologiyalarni rivojlantirishning ustuvor yo‘nalishlariga mosligi aniqlashtirilgan, tadqiqotning ilmiy yangiligi va amaliy natijalari bayon etilgan, olingan natijalarining ishonchliligi asoslangan, olingan natijalarning nazariy va amaliy ahamiyati ochib berilgan, amaliyotga joriy etilgan tadqiqot natijalari, nashr etilgan ishlar va dissertatsiya ishining tuzilmasi bo‘yicha ma’lumotlar keltirilgan.

Dissertatsiyaning **“Katta hajmli ma’lumotlarning hayotiy siklidagi xavfsizlik muammolari”** deb nomlangan birinchi bobida katta hajmli ma’lumotlarning texnologik xususiyatlari, ularning hayotiy sikli bosqichlari va har

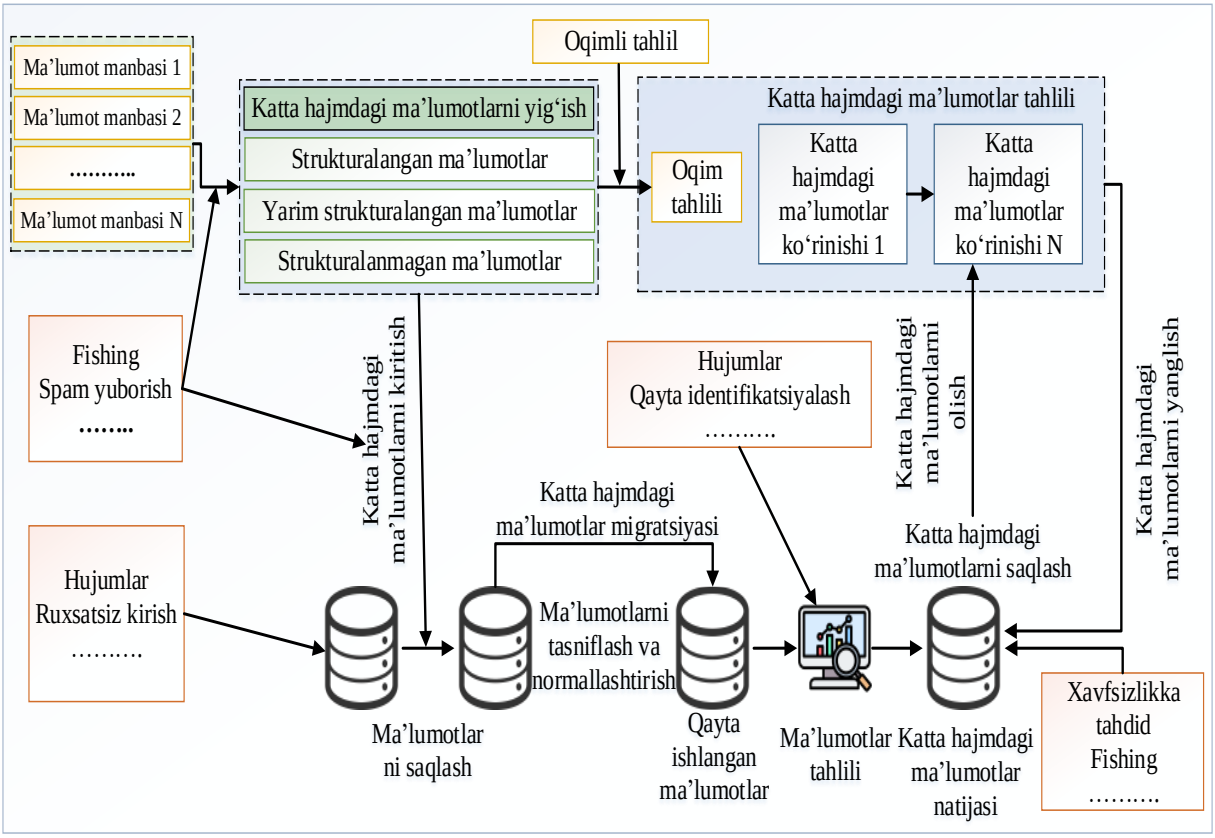
bir bosqichda yuzaga keluvchi tahdidlar tahlil qilingan. Tahdidlar, zaifliklar va aktivlarning o‘zaro bog‘liqligi asosida axborot xavfsizligi tahdid modeli ishlab chiqilgan. Ma’lumotlar tahdidlarga nisbatan qanday darajada himoyalanganini aniqlashda (maxfiylik, yaxlitlik, foydalanuvchanlik) baholovchi mezonlari keltirilgan.

*Birinchi paragraf*da katta hajmli ma’lumotlar texnologiyalarining 5V modeli asosida tavsifi, KHMda ma’lumotlarni to‘plash mexanizmi, KHM hayotiy siklning bosqichlari hamda axborot xavfsizligini ta’minlash nuqtai nazaridan hayotiy sikli bo‘yicha (1-rasmda) strukturaviy tahlil keltirilgan.



1-rasm. Axborot xavfsizligini ta’minlash nuqtai nazardan katta hajmli ma’lumotlar hayotiy sikli

*Ikkinchi paragraf*i katta hajmli ma’lumotlarning hayotiy siklidagi tahdid modeli (2-rasm), shuning asosida har bir bosqichda yuzaga keladigan kiberxavfsizlik tahdidlari hamda KHMni boshqarish tizimi uchun tahdid modeliga bag‘ishlangan.



2-rasm. Katta hajmli ma’lumotlarning hayotiy siklidagi tahdid modeli

*Uchinchi paragraf*da oltita boshqichdan - tahdid ishtirokchilari, hujum usullari, hujum nishonlari, hujum oqibatlari, kuzatiladigan anomaliyalar va qarshi choralardan iborat bo‘lgan KHMdagi kiberxavfsizlik tahdidi atributlari va ulardan himoyalanish mexanizmi tavsiflangan.

Dissertatsiyaning **“Katta hajmli ma’lumotlar tizimida foydalanishni boshqarish modeli va arxitekturasi”** deb nomlangan ikkinchi bobida atributga asoslangan foydalanishni boshqarishning takomillashtirilgan M-ABAC modeli ishlab chiqilgan. Model foydalanuvchi atributlari bilan bir qatorda xavfsizlik kontekstini ham hisobga oladi. Foydalanishni boshqarish qarorlari foydalanuvchining atributlaridan tashqari, aniqlangan xavf darajasiga asoslangan holda belgilanadi. Shuningdek, tizim arxitekturasi, komponentlararo axborot oqimi va qaror qabul qilish moduli tavsiflangan.

*Birinchi paragraf*da katta hajmli ma’lumotlar tizimida subyekt va obyekt atributlari to‘plamiga xavfsizlik atributlarini kiritish hamda beshta shartning bajarilishi orqali takomillashtirilgan foydalanishni boshqarishning ABAC modeli (M-ABAC) taklif etilgan. Ushbu shartlar quyida keltirilgan:

Birinchi shart. KHMda obyekt atributi foydalanishni boshqarish jarayonida ishtirok etadigan obyekt xususiyatlari haqidagi ma’lumotlarni tavsiflash uchun ishlatilishi kerak (Bu xavfsizlik modelining asosiy konsepsiyasidir. KHMdagi obyekt atributlari atribut nomi va atribut qiymati bo‘yicha aniqlanadi. Atributlarga qiymat berish jarayoni atributlarni belgilash deb ataladi).

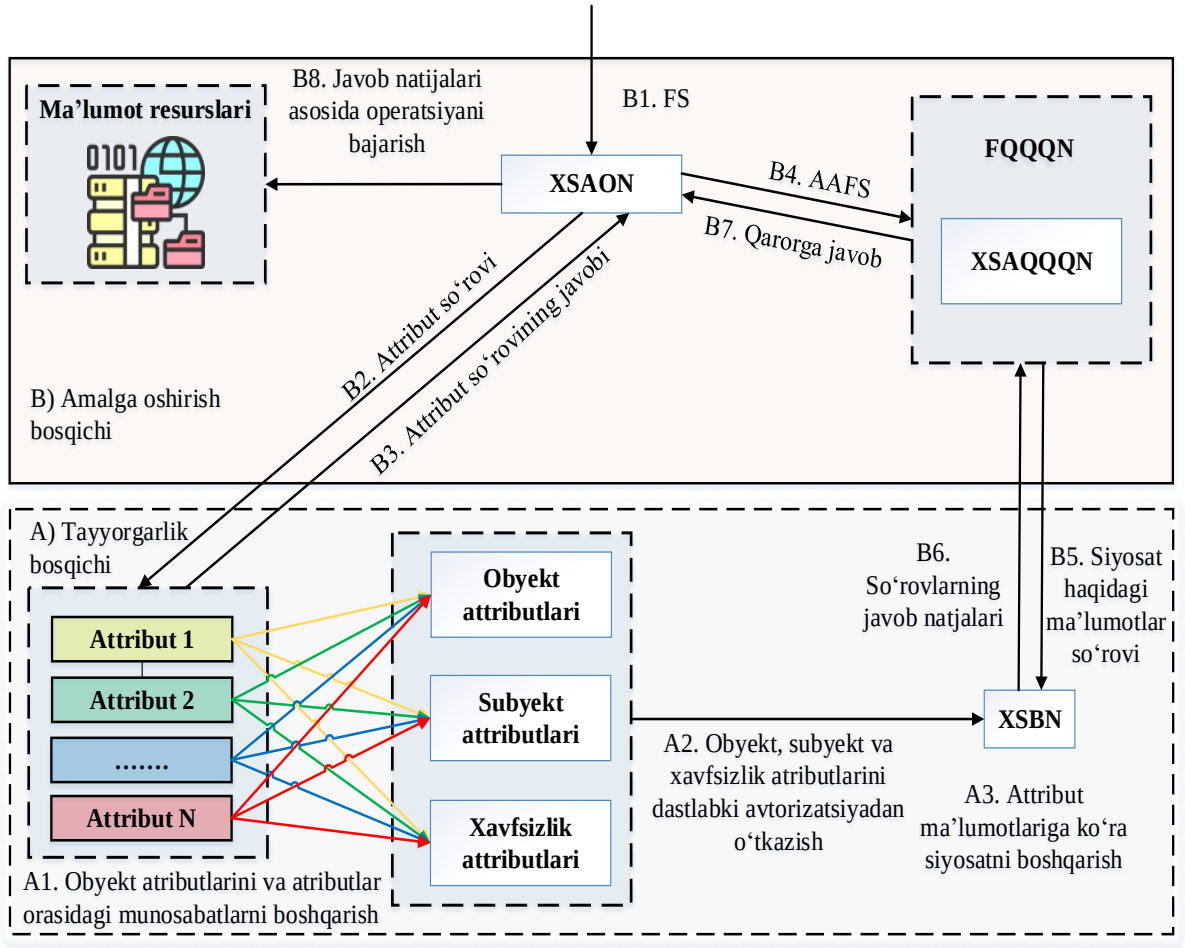
Ikkinchi shart. KHMda umumiy hisoblangan atributlar tanlangan obyektda foydalanishni boshqarish uchun o‘ziga xos cheklash ma’lumotlarini tavsiflash uchun ishlatilishi zarur. KHMda umumiy hisoblangan atributlardan xavfsizlikni

ta’minlashda foydalanish ABAC modelining ishonch zanjirini yaratishga asos bo‘la olmaydi. Tanlangan atributlar xavfsizlik parametrini ifodalashi kerak.

Uchinchi shart. KHMda ABAC modelining ishonch zanjirini yaratishda tanlangan atribut subyekt foydalanadigan resurslarda mavjud bo‘lishi hamda kamida ikki guruhga tegishli bo‘lishi kerak (Misol uchun atribut subyektga hamda u foydalanadigan resursiga tegishli bo‘lishi).

To‘rtinchi shart. KHMda atributlar to‘plami ma’lum bir obyektni tavsiflashi shart. KHMda xavfsizlikni ta’minlashda tanlangan atributlar to‘plami $X_A = (x_{a1}, x_{a2}, \dots x_{an})$ statik bo‘lishi va vaqt o‘tishi bilan o‘zgarmasligi kerak. Bu yerda, X_A-xavfsizlik atributlari to‘plami.

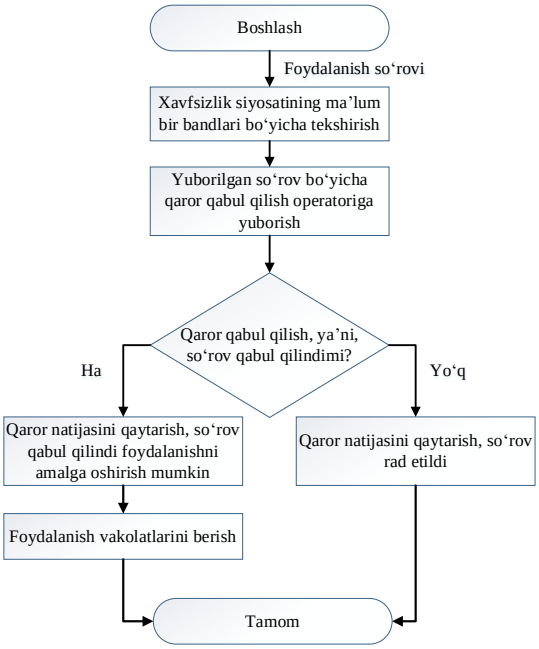
Beshinchi shart. KHMda atributga asoslangan foydalanishni boshqarish uchun xavfsizlik atributlar to‘plami xavfsizlik atributlaridan tashqari kamida bitta subyekt atributini va bitta obyekt atributini o‘z ichiga olgan bo‘lishi shart.



Xavfsizlik siyosatga asoslangan qaror qabul qilish nuqtasi - XSAQQQN
Foydalanish so‘rovi - FS
Xavfsizlik siyosatni amalga oshirish nuqtasi - XSAON
Foydalanishning qaror qabul qilish nuqtasi - FQQQN
Xavfsizlik siyosatni boshqarish nuqtasi - XSBN
Attributga asoslangan foydalanish so‘rovi - AAFS

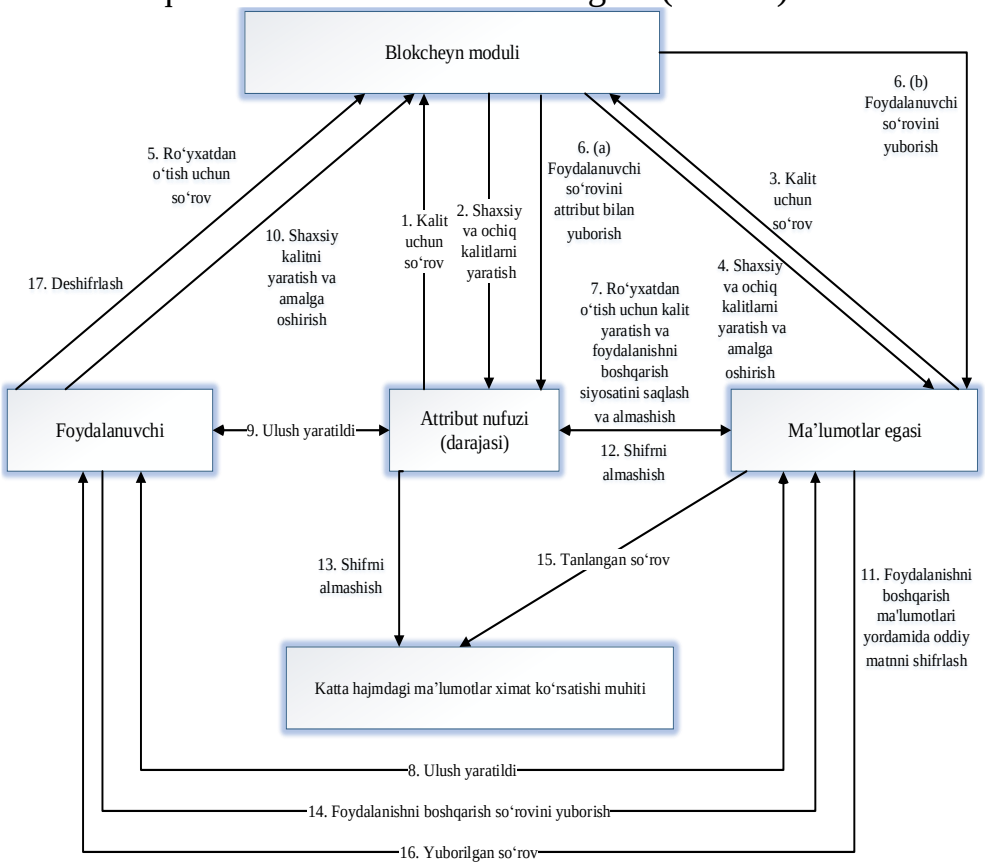
3-rasm. KHMda foydalanishlarni boshqarish konseptual modeli tuzilmasi

Taklif etilgan M-ABAC konseptual modeli tuzilmasi 3-rasmda, foydalanishni boshqarishda qaror qabul qilish va amalga oshirish algoritmining blok sxemasi esa 4-rasmda keltirilgan.



4-rasm. Foydalanishni boshqarishda qaror qabul qilish va amalga oshirish algoritmining blok sxemasi

*Ikkinchi paragraf*da KHMda foydalanishni boshqarish modeli taklif etilganida va integratsiya bo‘ladigan tizimlar yaratilganida yoki yangi tizimlar bilan birgalikda ishlash zaruriyati tug‘ilganda model komponentlari o‘zgarmasligi uchun foydalanishni boshqarish arxitekturasi taklif etilgan. (5-rasm)

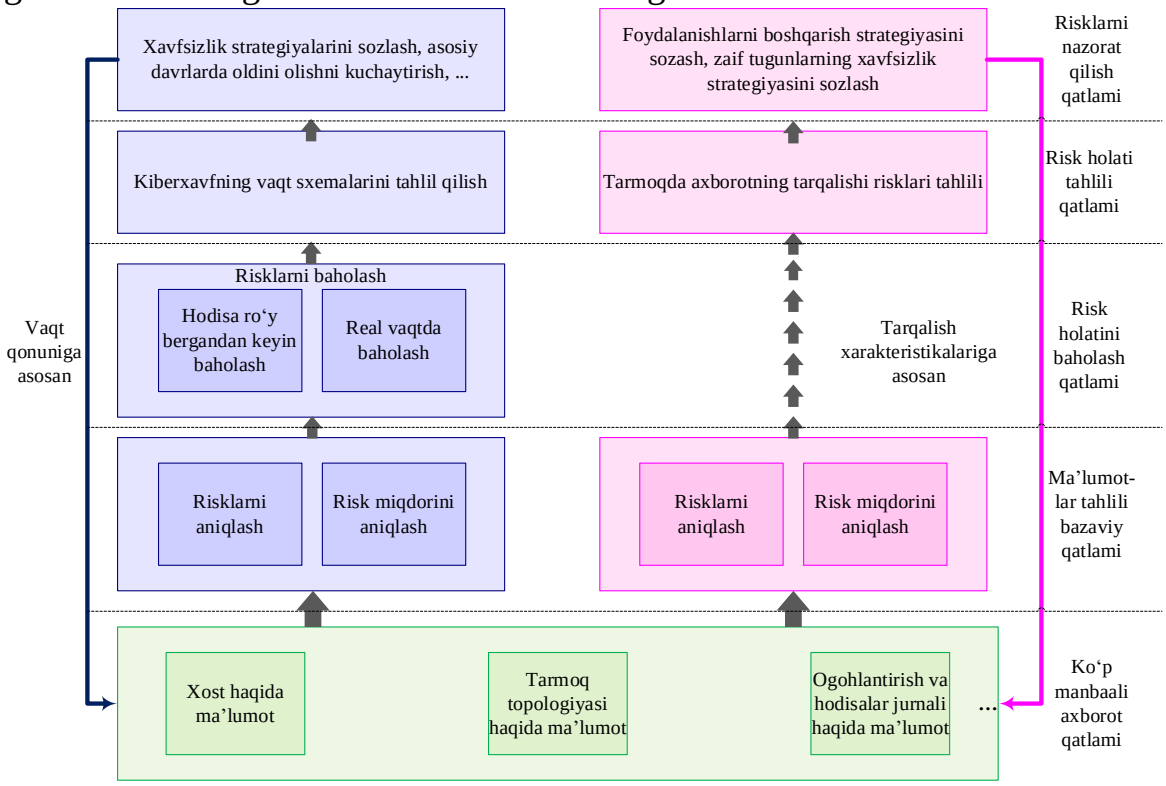


5-rasm. Kalit siyosatli atributga asoslangan shifrlashni madadlovchi foydalanishni boshqarish arxitekturasi

Uchinchi paragraf KHM uchun riskka moslashtirilgan foydalanishni boshqarishning 4 ta modulli mexanizmiga bag‘ishlangan. Unga ko‘ra har bir ruxsat so‘rovi uchun riskni baholash kiritiladi va shu orqali foydalanishni boshqarishda riskga asoslangan yondashuv qo‘llaniladi.

Dissertatsiyaning **“Katta hajmli ma’lumotlar tizimida risklarni boshqarish, anonimlikni ta’minlash model va algoritmi”** deb nomlangan uchinchi bobida ko‘p sathli axborot xavfsizligi risklarini baholash modeli ishlab chiqilgan. Modelda aktivlarning maxfiyligi, yaxlitligi va foydalanuvchanligi, zaifliklari hamda tahdidlar asosida risk darajasi aniqlanadi va tegishli choralar ko‘riladi. Shuningdek, risk darajalari aniqlangandan so‘ng, foydalanuvchilar haqidagi shaxsiy ma’lumotlarni himoyalash maqsadida 2 o‘rtachali klasterlash va “greedy” algoritmlariga asoslangan yangi anonimlashtirish algoritmi taklif etilgan.

*Birinchi paragraf*da axborot xavfsizligi risklarini boshqarishni beshta qatlam hamda ikkita o‘lchov - risk rivojlanishining vaqtinchalik xususiyatlari va risk tarqalishining fazoviy xususiyatlari asosida amalga oshiruvchi katta hajmli ma’lumotlarda risklarni boshqarishning ko‘p bosqichli konseptual modeli taklif etilgan hamda uning tuzilishi 6-rasmda keltirilgan.



6-rasm. Katta hajmli ma’lumotlarda axborot xavfsizligi risklarini boshqarishning ko‘p bosqichli konseptual modelining tuzilishi

*Ikkinchi paragraf*da maxfiylikni saqlashga qaratilgan modellarning samaradorligini aniqlash uchun D2D texnologiyasida (a, k) – anonimlashtirish modeliga asoslangan algoritm qo‘llanilib, amaldagi K- anonimlashtirish algoritimi asosida yangi algoritm ishlab chiqish zarurati aniqlangan.

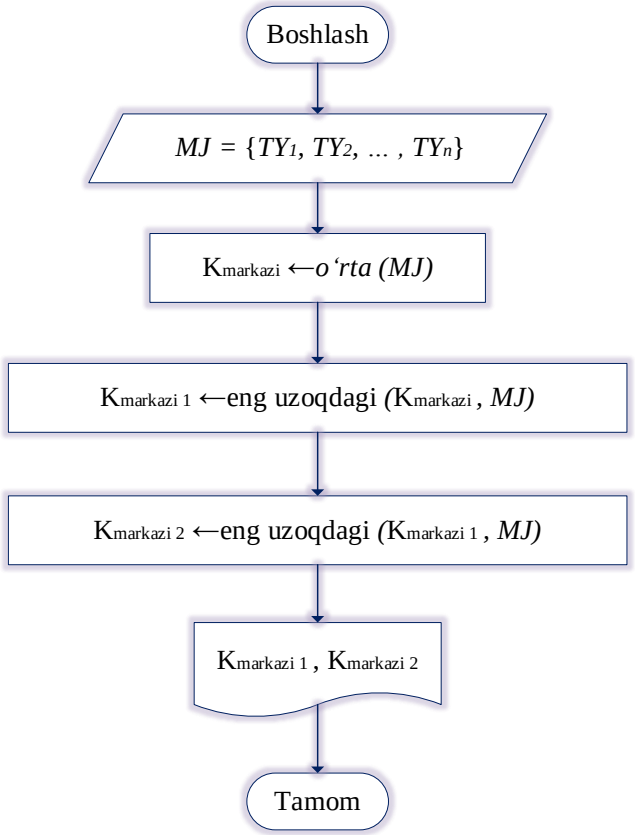
*Uchinchi paragraf*da foydalanuvchilar maxfiyligining ma’lum darajada ta’minlab, nisbatan maxfiylik va foydalanuvchanlik balansini saqlash imkonini beruvchi, ikki o‘rtachali klasterlash hamda “greedy” algoritmlari asosida

anonimlashtirish algoritmi ishlab chiqilgan. Ushbu algoritm 2 qismdan iborat bo‘lib, ularning blok-sxemalari mos ravishda 7-rasm va 8-rasmlarda keltirilgan.

Ushbu algoritmning 1-qismida boshlang‘ich massa markazi aniqlanadi. Klasterlash effektini yaxshilash hamda barqarorligini ta’minlashga 2 o‘rtachali (means) algoritm uchun boshlang‘ich massa markazini tanlashda o‘rtacha markazni hisoblash talab etiladi. Tasniflash uchun har bir klasterlashda ikkita boshlang‘ich massa markazini tanlash kerak. Boshlang‘ich massa markazini tanlash uchun o‘rtacha massa markazi usulidan foydalanish mumkin. Bunda xavfsizlik atributlari uchun ma’lumotlar jadvalidagi t sonli atributning o‘rtacha qiymati quyidagicha aniqlanadi:

$$o'rtacha\ (N_t) = \frac{\sum_{i=1}^{n} U_i}{n} \quad , \qquad (1)$$

bu yerda, U_i - i-chi to‘plam ostidagi t-sonli atributning qiymati, n - jadvaldagi to‘plamlar soni.



7-rasm. Boshlang‘ich massa markazini aniqlash algoritmining blok-sxemasi

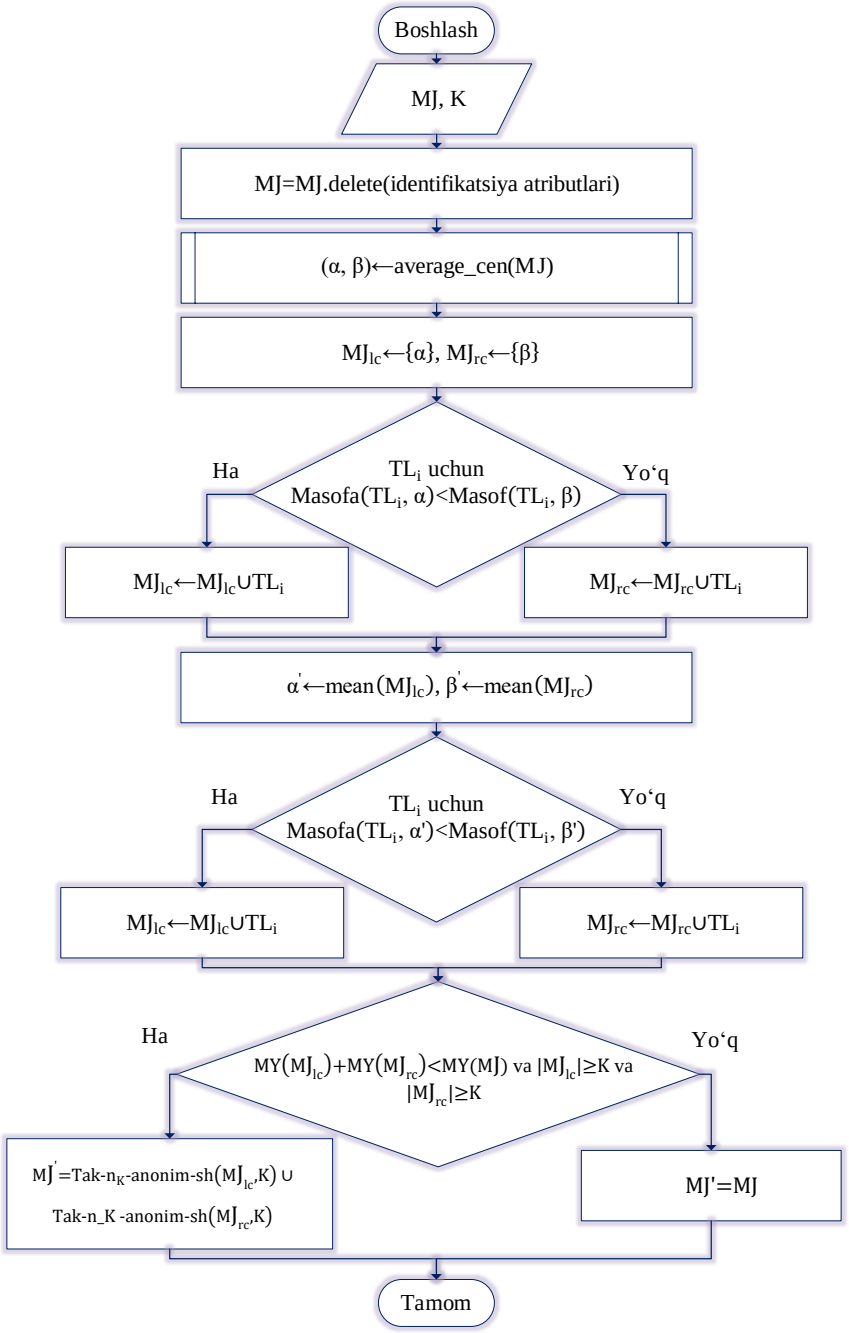
Tasniflashdan so‘ng, har bir sinf uchun klasterlash markazini qayta tanlash zarur. Yangilangan massa markazi esa (2) ifoda yordamida hisoblanadi.

$$K'_{markazi} = \{o'rtacha(N_1), o'rtacha(N_2) \dots o'rtacha(N_m), o'rtacha(C_1), o'rtacha(C_2), .., o'rtacha(C_n)\} \qquad (2)$$

Katta hajmli ma’lumotlarda maxfiylikni himoya qilish sharti bilan ma’lumotlar yo‘qotilishini minimallashtiradigan ekvivalentlik sinfini yaratiladi. K −anonimlik algoritmining ta’rifiga asoslanib, ma’lumotlarning maxfiyligini kafolatlash uchun qayta ishlangan ma’lumotlar jadvali quyidagi ifoda orqali hisoblanadi.

$$P\left(\frac{t_i}{MJ'}\right) \leq \frac{1}{K} \qquad (3)$$

Bu yerda t_i foydalanuvchi ma’lumotlari jadvalidagi i-yozuvni, MJ' esa klasterlash va umumlashtirish jarayonidan keyingi ma’lumotlar jadvalini bildiradi. Ma’lumotlar yo‘qotish miqdori imkon qadar kichik bo‘lishini ta’minlash uchun algoritm quyidagi tengliklini qanoatlantirishi kerak.



8-rasm. Takomillashtirilgan K-anonimlashtirish algoritmining blok-sxemasi

$$K = \arg\min MY(MJ') \qquad (4)$$

Bu yerda $MY(MJ')$ - umumlashtirish jarayonidan keyingi MJ' ma’lumotlar jadvalidagi ma’lumotlarning yo‘qolishi darajasini ko‘rsatadi. Shunga asoslanib, asl ma’lumotlarni takomillashtirilgan 2 o‘rtachali (means) va greedy algoritmlari asosida tasniflanadi.

Dissertatsiyaning **“Takomillashtirilgan modellar, algoritm samaradorligini baholash va amaliyotga tatbiq etish natijalari”** deb nomlangan to‘rtinchi bobida ishlab chiqilgan model va algoritmlar asosida yaratilgan dasturiy vositaning amaliy sinov natijalari keltirilgan.

Sinovlar real tashkilotlarda o‘tkazilib, foydalanuvchi noto‘g‘ri vakolat olishi holati 0,8 foizgacha kamaytirilgani, anonimlashtirish natijasida ma’lumotlarni

1-jadval

Turli xildagi foydalanishni boshqarish modellarini taklif etilgan M-ABAC modeli bilan taqqoslash

T/r	Mezonlar	DAC	MAC	RBAC	ABAC	R-BAC	M-ABAC
1	*2*	*3*	*4*	*5*	*6*	*7*	*8*
1.	Eng kam imtiyozni qo‘llash asosida foydalanishni boshqarish imkonining mavjudligi	-	-	+	+	+	+
2.	Dinamik vazifalarni ajratish imkonining mavjudligi	-	-	+	+	-	+
3.	Berilgan vakolatlarni qayd etib borish imkonining mavjudligi	+	+	+	+	+	+
4.	Beriladigan vakolatlarni sintaktik va semantik madadlash modulining mavjudligi	-	-	-	-	-	+
5.	Umumiy boshqaruv siyosatiga amal qilish imkonining mavjudligi	-	-	-	-	+	+
6.	Konfiguratsiyaning mosla-shuvchanlik imkonining mavjudligi	-	-	+	-	-	+
7.	Operativ (tezkor) holatda vaziyatdan xabardorlik qilish imkonining mavjudligi	-	-	-	-	+	-
8.	Autentifikatsiya funksiyasi bilan birlashtirish imkonining mavjudligi	-	-	-	-	-	-
9.	Operatsion tizimga muvofiqlik darajasi (Operatsion tizimga integratsiya qilish imkonining mavjudligi)	+	-	+	-	-	+
10.	Foydalanishni boshqarish funksiyalarini sinovdan o‘tkazish va tekshirish imkonining mavjudligi	-	-	-	-	+	+
11.	Passiv va aktiv ma’lumotlar oqimlarini madadlash imkonining mavjudligi	-	-	-	-	-	-
12.	Katta hajmli vertikal va gorizontal doirani madadlash imkonining mavjudligi	-	-	-	-	-	-
13.	Vakolatlarni dinamiklik imkonining mavjudligi	+	-	-	+	+	+
14.	Qatlamlar bo‘ylab foydalanuvchining hisob ma’lumotlarini o‘tkazish imkonining mavjudligi	-	-	-	-	+	+
15.	Masshtablilik imkonining mavjudligi	-	-	+	-	-	+
16.	Atributlarni boshqarishda moslashuvchanlik imkonining mavjudligi	-	-	-	+	+	+

himoyalash samaradorligi oshgani ko‘rsatilgan. Modelning mavjud ochiq platformalar bilan solishtirma jadvali, samaradorlik ko‘rsatkichlari grafik va jadval ko‘rinishida keltirilgan. Taklif etilayotgan modelning KHMni ishlash jarayonida foydalanishni boshqarish uchun qo‘llaniladigan modellar bilan talab etilgan mezonlar asosida, taqqoslash natijalari 1-jadvalda keltirilgan.KHMda foydalanishni boshqarish modellari asosida ishlab chiqilgan dasturiy vositalarning real vaqt rejimida subyektning (foydalanuvchining) foydalanish vakolatini belgilashga ketgan vaqt sarfi natijalari 2-jadvalda keltirilgan.

2-jadval

Subyekt so‘roviga javob qaytarish vaqti bo‘yicha testlash natijalari (soat.minut.sekund=soat.min.sek)

T/r	Foydalanishni boshqarish vositalari	Foydala-nishni boshqarish modeli	So‘rov yuborish vaqti (soat.min.sek.)	So‘rovga javob olish vaqti (soat.min.sek.)	Foydala-nish vakolatini belgilashga ketgan vaqt sarfi
1.	BDSS	M-ABAC	15.30.24	15.31.14	50 sekund
2.	Access Control Manager	DAC	15.36.13	15.37.12	59 sekund
3.	ZKBio Access IVS	MAC	15.43.19	15.44.15	56 sekund
4.	ZKAccess 3.5	RBAC	16.05.12	16.06.26	74 sekund
5.	HikCentral Access Control	ABAC	16.15.47	16.16.39	52 sekund
6.	BEWARD Access Control	R-ABAC	16.31.53	16.32.58	65 sekund

Ishlab chiqilgan dasturiy vosita tomonidan tashkilot xodimlariga berilgan vakolatlarni aniqligini tekshirish natijalari 3-jadvalda keltirilgan.

3-jadval

Foydalanuvchilarning tizimdan foydalanishda berilgan vakolatlar aniqligi bo‘yicha olingan natijalar

T/r	Foydalanishni boshqarish vositalari	Tizimdan foydalanishga ruxsat olgan foydalanuvchilar	Berilgan vakolat aniqligi (foizda va foydalanuvchi sonida)	Berilgan vakolat xatoligi (foizda)
1.	BDSS	114	95,6 (109 ta)	4,4
2.	Access Control Manager	110	94,5 (104 ta)	5,5
3.	ZKBio Access IVS	127	93,7 (119 ta)	6,3
4.	ZKAccess 3.5	131	93,9 (123 ta)	6,1
5.	HikCentral Access Control	95	94,7 (90 ta)	5,3
6.	BEWARD Access Control	109	95,4 (104 ta)	4,6

Foydalanuvchilarning tizimdan foydalanishlari uchun ruxsat etilmagan vakolatlar aniqligi bo‘yicha olingan natijalar 4-jadvalda keltirilgan.

4-jadval

Foydalanuvchilarning tizimdan foydalanishlari uchun ruxsat etilmagan vakolatlar aniqligi bo‘yicha olingan natijalar

T/r	Foydalanishni boshqarish vositalari	Tizimdan foydalanishga ruxsat etilmagan foydalanuvchilar	Ruxsat etilmagan vakolat aniqligi (foizda va foydalanuvchi sonida)	Ruxsat etilmagan vakolat xatoligi (foizda)
1.	BDSS	36	91,7 (33 ta)	8,3
2.	Access Control Manager	40	87,5 (35 ta)	12,5
3.	ZKBio Access IVS	23	87 (20 ta)	13
4.	ZKAccess 3.5	19	89,5 (17 ta)	10,5
5.	HikCentral Access Control	55	90,9 (50 ta)	9,1
6.	BEWARD Access Control	41	90,2 (37 ta)	9,8

Belgilangan ko‘rsatkichlarning birinchi, ya’ni dastlabki ma’lumotlarni anonimlashtirish natijasida shakllantirilgan jadvallardagi xatoliklar bo‘yicha testlash natijalari 5-jadvalda keltirilgan.

5-jadval

Dastlabki ma’lumotlarni anonimlashtirish natijasida shakllantirilgan jadvallardagi xatoliklar bo‘yicha testlash natijalari

T/r	Algoritm nomi	Aniqligi (%)	Xatolik (%)
1.	Takomillashtirilgan $K-$anonimlashtirish algoritmi	96,56	3,44
2.	EM algoritmi	95,84	4,16
3.	Greedy algoritmi	95,26	4,74
4.	$k-$anonimlashtirish algoritmi	96,51	3,49
5.	l-Diversity algoritmi	96,54	3,47
6.	$(k,e)-$anonimlashtirish algoritmi	95,41	4,59
7.	$t-$Closeness algoritmi	94,28	5,72
8.	$(X,Y)-$Privacy algoritmi	95,14	4,86
9.	MultiR $k-$anonimlashtirish algoritmi	93,78	6,22
10.	Distributional Privacy algoritmi	94,85	5,15

Belgilangan ko‘rsatkichlarning ikkinchisi, ya’ni, anonim jadvallarni shakllantirishga ketgan vaqt sarfi bo‘yicha testlash natijalari 6-jadvalda keltirilgan.

6-jadval

Anonim jadvallarni shakllantirishga ketgan vaqt sarfi bo‘yicha testlash natijalari

T/r	Algoritm nomi	Vaqt sarfi (sekunda)
1.	Takomillashtirilgan $K-$anonimlashtirish algoritmi	0,05 s.
2.	EM algoritmi	0,15 s.
3.	Greedy algoritmi	0,08 s.
4.	$k-$anonimlashtirish algoritmi	0,58 s.
5.	l-Diversity algoritmi	0,18 s.
6.	$(k, e)-$anonimlashtirish algoritmi	0,07 s.
7.	$t-$Closeness algoritmi	0,54 s.
8.	$(X, Y)-$Privacy algoritmi	0,35 s.
9.	MultiR $k-$anonimlashtirish algoritmi	0,47 s.
10.	Distributional Privacy algoritmi	0,51 s.

Belgilangan ko‘rsatkichlarning uchinchisi, ya’ni qayta identifikatsiyalash talab etiladigan to‘plamlar soni bo‘yicha testlash natijalari 7-jadvalda keltirilgan.

7-jadval

Qayta identifikatsiyalash talab etiladigan to‘plamlar soni bo‘yicha testlash natijalari

T/r	Algoritm nomi	Dastlabki to‘plamlar soni	Qayta identifikatsiyalash talab etiladigan to‘plamlar soni
1.	Takomillashtirilgan $K-$anonimlashtirish algoritmi	10	2
2.	EM algoritmi	10	3
3.	Greedy algoritmi	10	3
4.	$k-$anonimlashtirish algoritmi	10	4
5.	l-Diversity algoritmi	10	5
6.	$(k, e)-$anonimlashtirish algoritmi	10	3
7.	$t-$Closeness algoritmi	10	5
8.	$(X, Y)-$Privacy algoritmi	10	4
9.	MultiR $k-$anonimlashtirish algoritmi	10	4
10.	Distributional Privacy algoritmi	10	3

Testlash natijalariga ko‘ra takomillashtirilgan K- anonimlashtirish algoritmi dastlabki ma’lumotlarni anonimlashtirishda 96,56 foiz aniqlik bilan anonimlashtirishga hamda boshqa dasturiy vositalarga nisbatan 0,02 foizga yaxshiroq natija qayd etish imkonini bergan. Bundan tashqari dastlabki ma’lumotlarni anonimlashtirish natijasida shakllantirilgan jadvallardagi xatoliklar

bo‘yicha 3,44 foiz ko‘rsatkich qayd etilib, amaldagi dasturiy vositalarga nisbatan 0,03 foizga xatolikni kamaytirish imkonini bergan.

XULOSA

“Katta hajmli ma’lumotlarning hayotiy siklida axborotni himoyalashning model va algoritmi” mavzusidagi dissertatsiya ishi bo‘yicha olib borilgan tadqiqotlar natijasida quyidagi xulosalar taqdim etildi:

1. Katta hajmli ma’lumotlar muhitini himoya qilish uchun uning hayotiy siklidagi tahdidlar va hujumlar tahlil qilindi. Xavfsizlik zanjiri modeli sifatida katta hajmli ma’lumotlar hayotiy siklidagi tahdid modeli taklif etildi. katta hajmli ma’lumotlar hayotiy sikli har bir bosqichida axborot xavfsizligini ta’minlash talablarini bajarish maqsadida katta hajmli ma’lumotlarni boshqarish tizimidagi tahdidlar tasniflandi. Ma’lumotlar fragmentining hayotiy sikli, ma’lumotlar bazasini boshqarish tizimlariga nisbatan bo‘ladigan tahdid manbalari, katta hajmli ma’lumotlarni qayta ishlash va saqlash tizimlariga nisbatan bo‘ladigan tahdid manbalarini inobatga olgan xavfsizlik tahdid modeli taklif etildi.

2. Foydalanishni boshqarishning ABAC modeli subyekt va obyekt atributlari to‘plamiga xavfsizlik atributlarini kiritish orqali takomillashtirildi. Natijada taklif etilgan model asosida ishlab chiqilgan “BDSS dasturi” dasturiy vositasi foydalanuvchilarga vakolat berish aniqligini 0,2 foiz oshirish, ruxsat etilmagan vakolatlarni taqdim etish xatoligini 0,8 foizga kamaytirish imkonini bergan. Subyektlarga foydalanish vakolatini belgilash uchun esa mavjud dasturiy vositalarga nisbatan 3 sekund kamroq vaqt talab etilgan.

3. Kalit siyosatli atributga asoslangan shifrlashni madadlovchi foydalanishni boshqarishning arxitekturasi ishlab chiqildi. Natijada ushbu arxitektura yangi blokni blokcheynga qo‘shishdan oldin tekshirish orqali faqat tasdiqlangan ma’lumotlar manbalari o‘rtasida aloqa o‘rnatilishini ta’minlagan.

4. Axborot xavfsizligi risklarini boshqarish modeli, risk darajalarini aniqlashda klassik noravshan matematik baholash usulini qo‘llash orqali takomillashtirildi. Natijada katta hajmli ma’lumotlar tizimida noaniq risklarni baholash imkoni yaratilgan.

5. Ikki o‘rtachali klasterlash hamda “greedy” algoritmlari asosida anonimlashtirish algoritmi ishlab chiqildi. Ilmiy tadqiqot natijasida ishlab chiqilgan algoritmning “Big Data security analyzer dasturi” dasturiy moduli dastlabki ma’lumotlarni 96,56 foiz aniqlik bilan anonimlashtirish imkonini bergan va anonim jadvallarni shakllantirish uchun mavjudlariga nisbatan 0,03 sekund kam vaqt talab etgan.

6. Dissertatsiya ishida taklif etilgan model va algoritm asosida ishlab chiqilgan dastur modullar, undan foydalanish uchun texnik talablar, dasturiy mahsulot tajribaviy sinovdan o‘tkazildi, tashkilotlardan olingan natijalar keltirildi. Ushbu natijalarga ko‘ra ishlab chiqilgan dasturiy vosita tashkilotning axborot tizimidagi barcha foydalanuvchilar vakolatlarini dinamik o‘zgartirish, tashkilotning axborot tizimiga ulangan barcha axborot tizimlariga integratsiya bo‘lish imkonini berishi aniqlandi.

НАУЧНЫЙ СОВЕТ DSc. 13/30.12.2019.T.07.02 ПО ПРИСУЖДЕНИЮ УЧЕНЫХ СТЕПЕНЕЙ ПРИ ТАШКЕНТСКОМ УНИВЕРСИТЕТЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

ТАШКЕНТСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

АХМЕДОВА НОЗИМА ФАРХОД КИЗИ

МОДЕЛЬ И АЛГОРИТМ ЗАЩИТЫ ИНФОРМАЦИИ В ЖИЗНЕННОМ ЦИКЛЕ БОЛЬШИХ ДАННЫХ

05.01.05 – Методы и системы защиты информации. Информационная безопасность

АВТОРЕФЕРАТ ДИССЕРТАЦИИ
ДОКТОРА ФИЛОСОФИИ (PhD) ПО ТЕХНИЧЕСКИМ НАУКАМ

Ташкент – 2025

Тема диссертации доктора философии (PhD) по техническим наукам зарегистрирована в Высшей аттестационной комиссии при Министерстве Высшего образования, науки и инноваций Республики Узбекистан за B2025.1.PhD/T5293.

Диссертация выполнена в Ташкентском университете информационных технологий имени Мухаммада ал-Хоразмий.

Автореферат диссертации на трех языках (узбекский, русский, английский (резюме)) размещен на веб-странице Научного совета (www.tuit.uz) и на Информационно-образовательном портале «ZiyoNet» (www.ziyonet.uz).

Научный руководитель:	**Ташев Комил Ахматович** Кандидат технических наук, доцент
Официальные оппоненты:	**Ботиров Файзулла Бахтиёрович** доктор технических наук, доцент
	Иброхимов Азизбек Равшанбек угли доктор философии по техническим наукам
Ведущая организация:	**ООО «UNICON.UZ»**

Защита диссертации состоится «25» октября 2025 года в 10:00 на заседании Научного совета DSc. 13/30.12.2019.T.07.02 при Ташкентском университете информационных технологий. (Адрес: 100084, г. Ташкент, ул. Амира Темура, 108. Тел.: (99871) 238-64-15; e-mail: info@tuit.uz).

С диссертацией можно ознакомиться в Информационно-ресурсном центре Ташкентского университета информационных технологий (регистрационный номер № 369). (Адрес: 100084, г. Ташкент, ул. Амира Темура, 108. Тел.: (99871) 238-64-15).

Автореферат диссертации разослан «15» октября 2025 года.
(протокол рассылки № 10 от «15» октября 2025 года.)

Б.Ш. Махкамов
Председатель научного совета по присуждению ученых степеней, д.э.н., профессор

М.С. Саиткамолов
Ученый секретарь научного совета по присуждению ученых степеней, д.э.н., доцент

Д.Я. Иргашева
Председатель научного семинара при научном совете по присуждению ученых степеней, д.т.н., профессор

ВВЕДЕНИЕ (аннотация диссертации доктора философии (PhD))

Актуальность и востребованность темы диссертации. В мировом масштабе большие данные глубоко проникают практически во все отрасли экономики и социальной сферы, создавая основу для широких возможностей в цифровой среде для пользователей интернета. Однако вместе с ростом числа пользователей этих технологий значительно увеличивается количество угроз и атак. Согласно данным компании Terranova Security, в период с второго квартала 2023 года по второй квартал 2024 года нарушения информационной безопасности в системах больших данных в основном происходили вследствие несанкционированного использования информационных систем. Такие инциденты приводили к незаконному утечке данных из систем, причём более 65% этих данных оказались в неанонимизированном виде[1]. Результаты подобных исследований имеют важное значение для дальнейшего развития национальных стратегий и международных инициатив по повышению информационной безопасности в системах больших данных.

В мировом масштабе ведутся научные исследования, направленные на разработку моделей угроз для защиты информации в жизненном цикле больших данных, а также механизмов противодействия киберугрозам. В рамках больших данных в качестве актуальной задачи в мире проводятся научные исследования по разработке новых моделей, методов и алгоритмов на основе управления доступом и анонимизации данных с целью повышения уровня защиты и повышения уровня контроля за несанкционированным доступом к системе, приводящим к нарушению информационной безопасности. В последние годы вопросы информационной безопасности, связанные с большими данными, стали одним из важнейших научных направлений. В связи со сложной и открытой средой систем больших данных особое внимание уделяется выявлению их уязвимостей и разработке эффективных механизмов защиты.

В республике проведен ряд эффективных работ по управлению доступом к большим данным, управлению доступом на основе атрибутов, управлению рисками информационной безопасности, совершенствованию моделей обеспечения анонимности. В связи с бурным развитием цифровых технологий обеспечение информационной безопасности в больших данных рассматривается как одно из приоритетных направлений, и в стратегии «Цифровой Узбекистан-2030» поставлены конкретные задачи по повышению потенциала кибербезопасности, защите персональных данных граждан, внедрению современных биометрических и антифрод-технологий в банковские и платежные системы, сбору и хранению больших данных с учетом требований анонимизации[2]. Для обеспечения эффективной реализации указанных задач важны защита информации в рамках жизненного цикла больших данных, разработка их модели и алгоритма, а также реализация

[1]Данные компании Terranova Security

[2] Указ Президента Республики Узбекистан, от 05.10.2020 г. № УП-6079 «Об утверждении Стратегии «Цифровой Узбекистан-2030» и мерах по ее эффективной реализации»

работ, направленных на минимизацию потерь информации в процессе анонимизации.

Наряду с этим, данное диссертационное исследование в определенной степени служит реализации задач, установленных в нормативно-правовых актах, включая Указ Президента Республики Узбекистан №PF-60 от 28 января 2022 года «О стратегии развития Нового Узбекистана на 2022–2026 годы», Постановление Кабинета Министров Республики Узбекистан №PQ-229 от 21 июня 2024 года «О мерах по организации научно-исследовательской деятельности в области цифровой криминалистики», Приказ Председателя Службы государственной безопасности Республики Узбекистан №91 от 4 сентября 2023 года «Об утверждении Положения об оценке уровня обеспечения кибербезопасности и безопасности объектов критической информационной инфраструктуры Республики Узбекистан», а также других нормативно-правовых документов, регулирующих соответствующую деятельность.

Соответствие исследования приоритетным направлениям развития науки и технологий республики. Данное исследование выполнено в соответствие с приоритетным направлением развития науки и технологий Республики IV. «Информатизация и развитие информационно-коммуникационных технологий».

Степень изученности проблемы. Научные исследования по применению моделей, методов и алгоритмов, разработанных на основе подходов, направленных на разграничение доступа и анонимизацию данных, для обеспечения безопасности больших данных ведут Пьетро Коломбо, Джим Лонгстафф, София Зеббудж, Джоанн Нобл, Рабах Брахами и другие ученые. С использованием таких методов, как роли, атрибуты и матрицы для разграничения доступа к большим объемам данных и методов К-анонимизации для анонимизации данных, научные исследования проводились зарубежными учеными, такими как Виктор К. М. Леунг, Доминго-Феррер, У. Нараянан, С. Варшней, Л. Т. Ян, Ч. Лю, и в настоящее время продолжаются исследователями и учениками научных школ под их руководством. Ученые из стран СНГ, такие как А. А. Климов, А. В. Карпов, Ю. А. Зайцев, проводили научные исследования по защите данных от утечек.

Ведущие специалисты Terranova Security, Eclipse, Cisco, Arkansas Public Safety Solutions (APSS), One Identity проводят специализированные научно-практические исследования по разработке и тестированию программных продуктов, направленных на обеспечение информационной безопасности больших объемов данных, специализируясь на разработке программных продуктов, обеспечивающих определение и контроль прав доступа пользователей и анонимизацию данных.

В Узбекистане научные коллективы под руководством С.К. Ганиева, М.М. Каримова, Д.Ю. Иргашевой, К.Ф. Керимова проводили научные исследования по управлению несанкционированным доступом в компьютерных сетях, распределению прав и определению ролей пользователей, совершенствованию моделей угроз и атак, методов и

алгоритмов, направленных на базы данных, а также разработке новых методов защиты.

В то же время, модели, методы и алгоритмы, направленные на минимизацию потерь данных в процессе управления угрозами, управления доступом, управления рисками и анонимизации данных на протяжении всего жизненного цикла больших данных, недостаточно изучены.

Связанность диссертационного исследования с научно-исследовательскими планами высшего учебного заведения, в котором была выполнена диссертация. Диссертационное исследование выполнено в рамках плана научно-исследовательских работ Ташкентского университета информационных технологий имени Мухаммада аль-Хоразмий по проектам 101128871 — DEBSEUz — ERASMUS-EDU-2023-CBHE «Разработка целевой образовательной программы по солнечной энергетике для бакалавриата в Узбекистане»; AL-662204323 «Разработка моделей, алгоритмов и программных комплексов для фрактальных структур национальных узоров на тканях и коврах».

Целью исследования разработка моделей и алгоритмов защиты, предотвращающих утечку информации и несанкционированное использование данных в жизненном цикле больших данных.

Задачи исследования:

усовершенствование модели управления доступом в системе больших данных;

разработка архитектуры процесса управления доступом в системе больших данных;

усовершенствование модели управления рисками информационной безопасности в системе больших данных;

разработка алгоритма анонимизации, обеспечивающего конфиденциальность информации в системе больших данных.

Объектом исследования В качестве объекта исследования выбрана система больших данных.

Предметом исследования являются модели и алгоритмы управления доступом и защиты информации в жизненном цикле больших данных.

Методы исследования. В исследовании использовались методы дискретной математики, моделирования случайных процессов, теории вероятностей, теории множеств, объектно-ориентированного программирования и машинного обучения.

Научная новизна исследования заключается в следующем:

в системе больших данных усовершенствована модель управления доступом, которая позволяет повысить точность предоставления полномочий пользователям за счёт включения в совокупность атрибутов субъектов и объектов также атрибутов безопасности и пяти обязательных условий, подлежащих выполнению;

в системах больших данных разработана архитектура управления доступом, поддерживающая политику распределения открытых ключей шифрования, основанного на атрибутах, что обеспечивает возможность

интеграции системы защиты с другими системами в соответствии с требованиями безопасности;

в системе больших данных усовершенствована модель управления рисками, которая позволяет формировать динамический цикл обеспечения безопасности за счёт применения классического нечеткого математического метода оценки при выявлении, оценке и управлении новыми типами рисков;

усовершенствован алгоритм K-анонимизации, позволяющий сохранить доступность данных путём введения элементов двухсреднего кластерирования и «greedy» к методу выбора начального центра масс и методу среднего центра.

Практические результаты исследования заключаются в следующем:

разработано программное средство управления доступом на основе усовершенствованной атрибутной модели M-ABAC в системе больших данных;

создано программное средство для анонимизации в системе больших данных, основанное на алгоритме кластеризации «2-means» и алгоритме «greedy».

Достоверность результатов исследования. Достоверность результатов исследования подтверждается проведенным анализом в соответствии с выбранной моделью угроз, результатами, полученными при реализации модели и алгоритмов, а также результатами расчетов, полученными в заданных условиях.

Научная и практическая значимость результатов исследования.

Научная значимость результатов исследования заключается в том, что они способствуют повышению эффективности систем защиты информации за счет анализа модели угроз, атрибутов и механизмов реагирования в жизненном цикле больших данных, усовершенствования модели разграничения доступа на основе атрибутов и модели управления рисками информационной безопасности, разработки архитектуры разграничения доступа, позволяющей интегрировать системы защиты между собой, а также разработки усовершенствованного алгоритма K-анонимизации, обеспечивающего конфиденциальность данных.

Практическая значимость результатов исследования объясняется тем, что программное средство разработанное на основе модели управления доступом, модели управления рисками информационной безопасности и усовершенствованного алгоритма K-анонимизации для обеспечения безопасности информации в системе больших данных позволяет повысить эффективность систем защиты информации в жизненном цикле больших данных.

Внедрение результатов исследования. На основе полученных научных результатов по программным средствам, разработанным на основе усовершенствованных моделей и предложенных алгоритмов в системе больших данных:

программное средство «BDSS dasturi», разработанное на основе усовершенствованной модели управления доступом на основе атрибутов за

счет включения атрибутов безопасности в набор атрибутов субъекта и объекта, а также пяти условий, которые должны быть соблюдены в системах больших данных, внедрено в практическую деятельность ГУП «Центр кибербезопасности» (Справка Министерства цифровых технологий от 06 ноября 2024 года № 24-8/7511). Результат научных исследований позволил повысить точность предоставления прав доступа пользователям на 0,2% и снизить погрешность предоставления прав доступа неавторизованным пользователям на 0,8%;

программное средство основанное на разработанной архитектуре управления доступом для интеграции системы защиты с другими системами в соответствии с требованиями безопасности в системах больших данных, поддерживающая шифрование на основе атрибутов с политикой распределения открытых ключей, внедрено в практическую деятельность ГУП «Центр кибербезопасности» (Справка Министерства цифровых технологий от 06 ноября 2024 года № 24-8/7511). В диссертации подчеркивается возможность использования программного средства, разработанного на основе предложенных моделей и алгоритмов, для защиты от угроз конфиденциальности, целостности и доступности, являющихся основами информационной безопасности, которые возникают вследствии ненадлежащего предоставления полномочий пользователю при управлении доступом к большим данным, а также даются предложения и рекомендации.

программное средство «BDSS dasturi», разработанное на основе многоэтапной модели управления рисками информационной безопасности, усовершенствованной путем применения классического метода нечеткой математической оценки в процессе выявления, оценки и управления новыми видами рисков в системах больших данных, внедрено в практическую деятельность ГУП «Центр радиосвязи, радиовещания и телевидения» (Справка Министерства цифровых технологий от 06 ноября 2024 года № 24-8/7511). В результате научных исследований программное средство заняло на 3 секунды меньше времени на назначение прав доступа субъектам;

программное средство «Big Data security analyzer dasturi», усовершенствованного, в целях предотвращения утечек информации в системах больших данных и сохранения доступности данных, алгоритма k-анонимизации на основе алгоритмов кластеризации «2-means» и «greedy» в системах больших данных внедрено в практическую деятельность ООО «UZINFOCOM — единый интегратор по созданию и поддержке государственных информационных систем» (Справка Министерства цифровых технологий от 06 ноября 2024 года № 24-8/7511). В результате научных исследований программное средство заняло на 0,03 секунды меньше времени на формирование анонимных таблиц, чем существующие.

Апробация результатов исследования. Результаты исследований обсуждались на 3 международных и 7 республиканских научно-практических конференциях.

Публикация результатов исследования. По теме диссертации опубликовано 19 научных работ, в том числе 6 статей в научных изданиях,

рекомендованных ВАК Республики Узбекистан для публикации основных научных результатов диссертаций, получено 3 свидетельства о регистрации программных средств, разработанных для ЭВМ.

Структура и объем диссертации. Диссертация состоит из введения, четырех глав, заключения, списка литературы и приложений. Объем диссертации составляет 112 страниц.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИСЕРТАЦИИ

Во введении устанавливается актуальность и необходимость темы диссертации, формулируются цели и задачи, определяются объект и предмет исследования, выясняется соответствие исследования приоритетным направлениям развития науки и техники Республики Узбекистан, описывается научная новизна и практические результаты исследования, обосновывается достоверность полученных результатов, раскрывается теоретическая и практическая значимость полученных результатов, приводятся сведения о внедренных в практику результатах исследования, опубликованных работах и структуре диссертации.

В первой главе диссертации под названием «Проблемы безопасности в жизненном цикле больших данных» анализируются технологические характеристики больших данных, этапы их жизненного цикла и угрозы, возникающие на каждом этапе. Разработана модель угроз информационной безопасности, основанная на взаимозависимости угроз, уязвимостей и активов. Представлены критерии оценки для определения степени защищенности данных от угроз (конфиденциальность, целостность, удобство использования). В первом параграфе представлено описание технологий больших данных на основе модели 5V, механизма сбора данных в ИТ-системе, этапов жизненного цикла ИТ, а также структурный анализ жизненного цикла с точки зрения обеспечения информационной безопасности (рисунок 1).

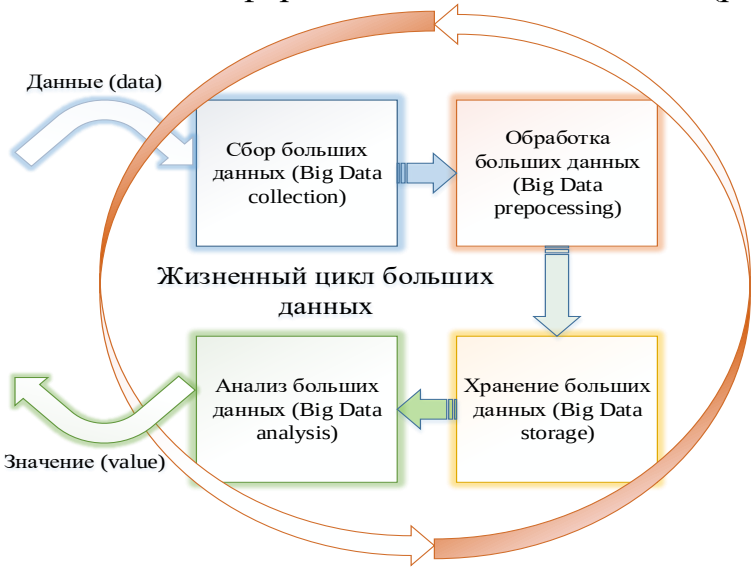


Рисунок 1. Жизненный цикл больших данных с точки зрения обеспечения информационной безопасности

Второй параграф посвящен модели угроз жизненного цикла больших данных (рисунок 2), угрозе кибербезопасности, возникающей на каждом этапе, модели и угрозы для систем управления большими данными.

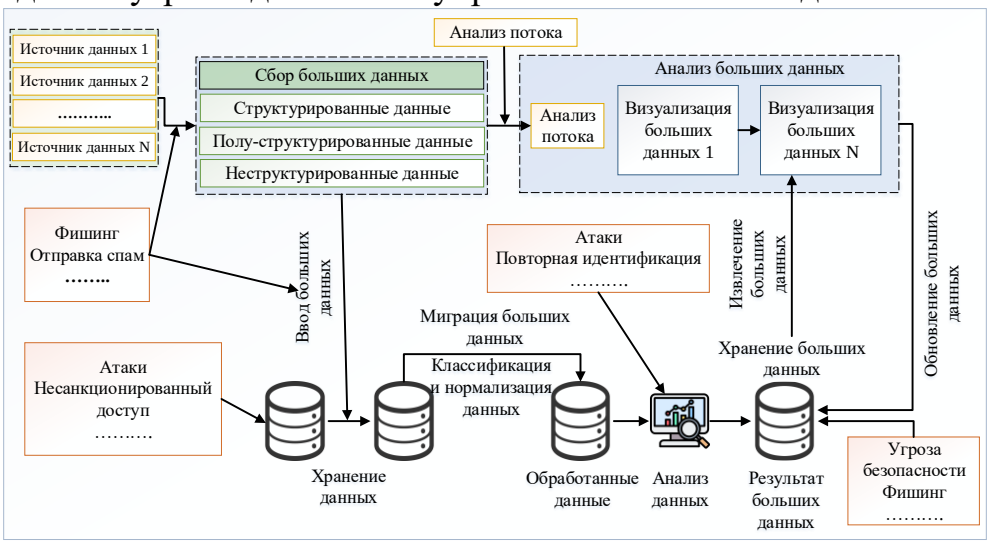


Рисунок 2. Модель угроз в жизненном цикле больших данных

В третьем параграфе описываются атрибуты угроз кибербезопасности в больших данных и механизмах защиты от никс, состоящих из шести этапов — субъекты угроз, методы атаки, цели атаки, последующие атаки, наблюдаемые аномалии и контрмеры.

Во второй главе диссертации под названием «Модель и архитектура контроля доступа в системах больших данных» разрабатывается усовершенствованная модель контроля доступа на основе атрибутов M-ABAC. Модель учитывает контекст безопасности, а также атрибуты пользователя. Решения по контролю доступа определяются на основе выявленного уровня риска в дополнение к атрибутам пользователя. Также описываются архитектура системы, межкомпонентный информационный поток и модуль принятия решений.

В первом параграфе предлагается усовершенствованная модель контроля доступа ABAC (M-ABAC) в системе больших данных путем включения атрибутов безопасности в набор атрибутов субъекта и объекта и выполнения пяти условий. Эти условия перечислены ниже:

Первое условие. В AAC атрибут объекта должен использоваться для описания информации о свойствах объекта, участвующего в процессе контроля доступа (Это основная концепция модели безопасности. Атрибуты объекта в AAC определяются именем атрибута и значением атрибута. Процесс присвоения значений атрибутам называется назначением атрибута).

Второе условие. В KHM атрибуты, которые обычно вычисляются, должны использоваться для описания конкретной информации об ограничениях для управления доступом к выбранному объекту. Использование атрибутов, которые обычно вычисляются в KHM для обеспечения безопасности, не может быть основой для создания цепочки доверия в модели ABAC. Выбранные атрибуты должны представлять параметр безопасности.

Третье условие. При создании цепочки доверия в КНМ выбранный атрибут должен присутствовать в ресурсах, используемых субъектом, и принадлежать как минимум к двум группам (например, атрибут должен принадлежать субъекту и ресурсу, который он использует).

Четвертое условие. В КНМ набор атрибутов должен описывать конкретный объект. Набор атрибутов, выбранных в КНМ для обеспечения безопасности

$X_A=(x_{a1},x_{a2},\ldots x_{an})$, должен быть статическим и не меняться со временем. Здесь X_A — это набор атрибутов безопасности.

Пятое условие. Для управления доступом на основе атрибутов в КНМ набор атрибутов безопасности должен включать как минимум один атрибут субъекта и один атрибут объекта в дополнение к атрибутам безопасности.

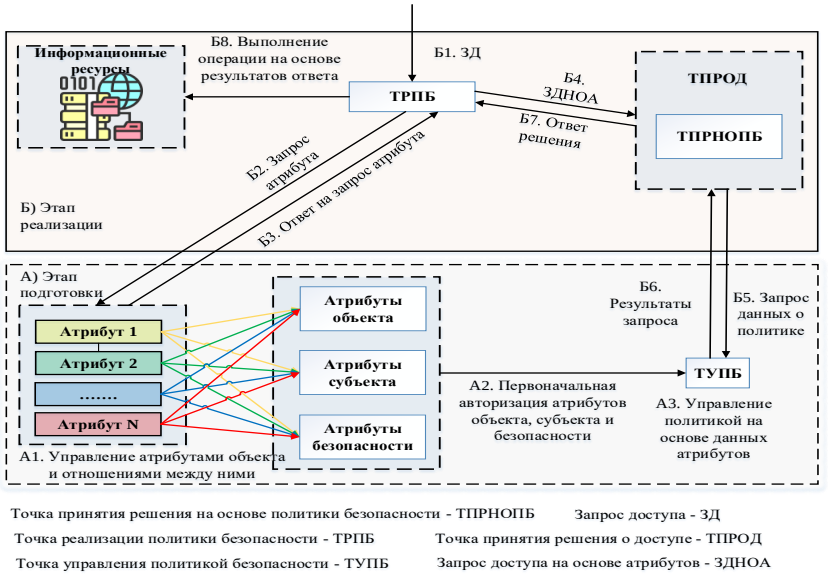


Рисунок 3. Структура концептуальной модели управления доступом в больших данных

Структура предлагаемой концептуальной модели M-ABAC представлена на рисунке 3, а структурная схема алгоритма принятия и реализации решений по контролю доступа — на рисунке 4.

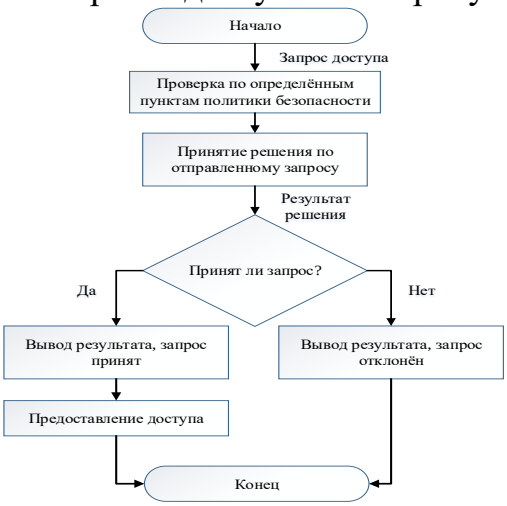


Рисунок 4. Блок-схема алгоритма принятия решений и его реализации при управлении доступом

Во втором параграфе предлагается архитектура управления доступом, в которой компоненты модели не меняются, когда применяется модель управления доступом, когда создаются системы для интеграции или когда возникает необходимость совместной работы с новыми системами в системе больших данных. (Рисунок 5).

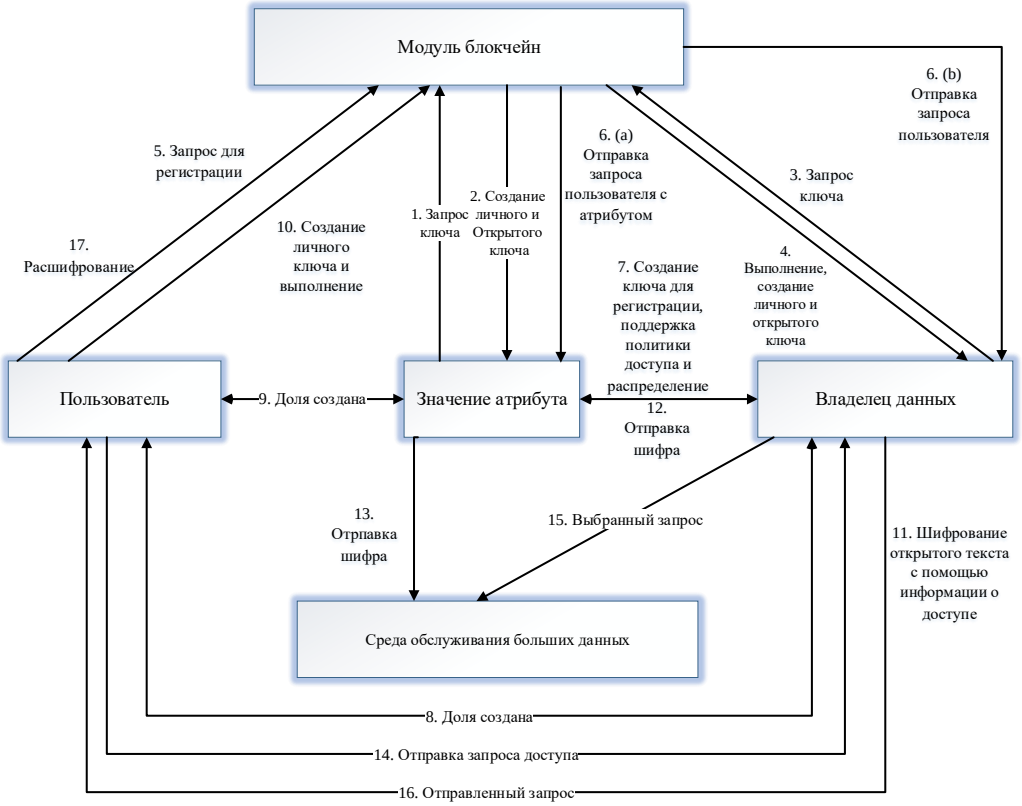


Рисунок 5. Архитектура управления доступом, поддерживающая шифрование на основе атрибутов с политикой ключа

Третий параграф посвящен 4-модульному механизму риск-адаптированного контроля доступа для KHM. Согласно ему, для каждого запроса на разрешение вводится оценка риска, тем самым применяя риск-ориентированный подход к контролю доступа.

В третьей главе диссертации под названием «Модель и алгоритм управления рисками, анонимность в системах больших данных» разработана многоуровневая модель оценки рисков информационной безопасности. Модель определяет уровень риска на основе конфиденциальности, целостности и удобства использования активов, уязвимостей и угроз, и принимаются соответствующие меры. Также после определения уровней риска предлагается новый алгоритм анонимизации на основе 2-средней кластеризации и «жадных» алгоритмов для защиты персональной информации о пользователях.

В первом параграфе предлагается многоуровневая концептуальная модель управления рисками в больших данных, которая реализует управление рисками информационной безопасности на основе пяти слоев и двух измерений — временных характеристик развития риска и пространственных характеристик распределения риска, а ее структура представлена на рисунке 6.

Рисунок 6. Структура многоэтапной концептуальной модели управления рисками информационной безопасности в больших данных

Во втором параграфе алгоритм на основе модели (a,k)-анонимизации в технологии D2D используется для определения эффективности моделей сохранения конфиденциальности, а также выявляется необходимость разработки нового алгоритма на основе текущего алгоритма K-анонимизации.

В третьем параграфе разрабатывается алгоритм анонимизации на основе кластеризации с двумя средними и «жадных» алгоритмов, который обеспечивает определенный уровень конфиденциальности пользователя и позволяет поддерживать баланс относительной конфиденциальности и удобства использования. Данный алгоритм состоит из 2 частей, структурные схемы которых представлены на рисунках 7 и 8 соответственно.

В первой части данного алгоритма определяется начальный центр масс. Для улучшения эффекта кластеризации и обеспечения устойчивости необходимо вычислять средний центр при выборе начального центра масс для алгоритма с двумя средними. Для классификации необходимо выбрать два начальных центра масс в каждой кластеризации. Для выбора начального центра масс можно использовать метод среднего центра масс. В этом случае среднее значение номера атрибута t в таблице данных по атрибутам безопасности определяется следующим образом:

$$\text{среднее } (N_t) = \frac{\sum_{i=1}^{n} U_i}{n} \quad , \qquad (1)$$

где U_i — значение t-го атрибута в i-м наборе, а n — количество наборов в таблице.

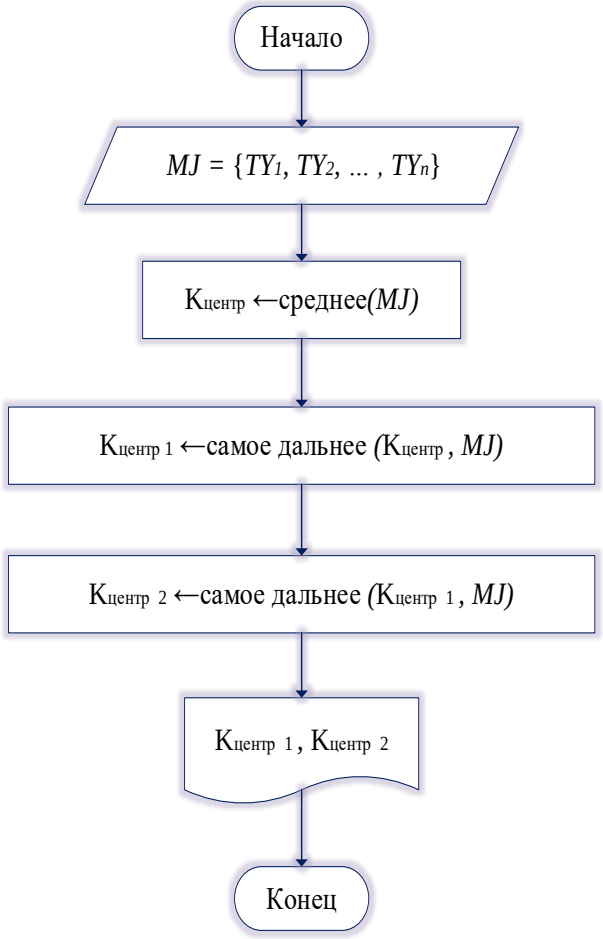


Рисунок 7. Блок-схема алгоритма определения начального центра масс

После классификации необходимо заново выбрать центр кластеризации для каждого класса. Обновленный центр масс рассчитывается с использованием выражения (2).

$$K'_{\text{центр}} = \{\text{среднее}(N_1), \text{среднее}(N_2) \dots \text{среднее}(N_m), \text{среднее}(C_1), \text{среднее}(C_2), .., \text{среднее}(C_n)\} \quad (2)$$

Создается класс эквивалентности, который минимизирует потерю данных, сохраняя конфиденциальность в больших данных. На основе определения алгоритма K-анонимности обработанная таблица данных для гарантии конфиденциальности данных рассчитывается по следующему выражению.

$$P\left(\frac{t_i}{MJ'}\right) \leq \frac{1}{K} \quad (3)$$

Здесь t_i обозначает i-ю запись в таблице пользовательских данных, а MJ^' обозначает таблицу данных после процесса кластеризации и обобщения. Чтобы гарантировать, что объем потери данных будет как можно меньше, алгоритм должен удовлетворять следующему уравнению.

$$K = \arg\min MY(MJ') \quad (4)$$

Здесь MY(MJ^') — уровень потери данных в таблице данных MJ^' после процесса обобщения. Исходя из этого, исходные данные классифицируются на основе улучшенных алгоритмов 2-means и жадного алгоритма.

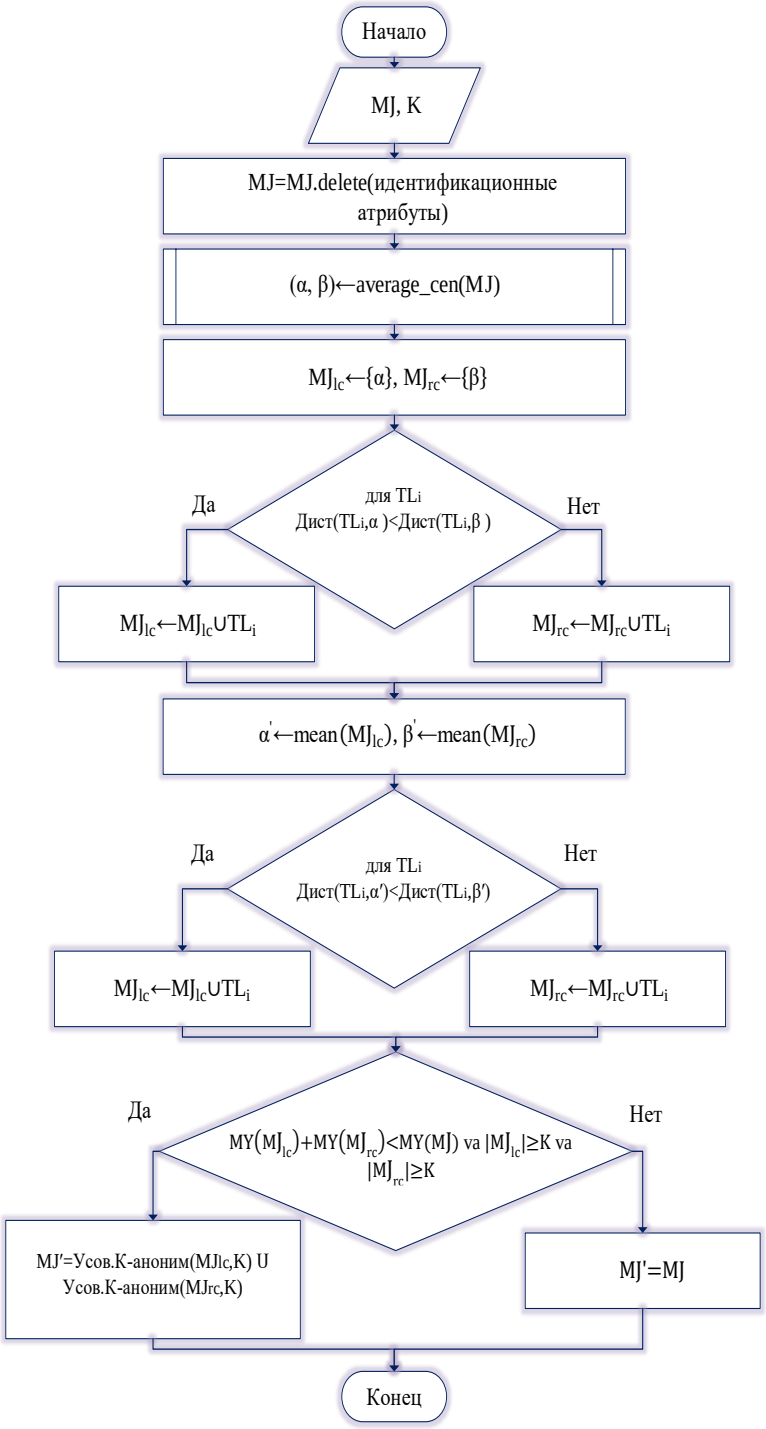


Рисунок 8. Блок-схема усовершенствованного алгоритма K-анонимизации

В четвертой главе диссертации под названием «Усовершенствованные модели, оценка эффективности алгоритмов и результаты практического применения» представлены результаты практического тестирования программного средства, созданного на основе разработанной модели и алгоритмов. Тестирование проводилось в реальных организациях и показало, что частота ошибок авторизации пользователей снизилась до 0,8 процента, а эффективность защиты данных возросла в результате анонимизации. Сравнительная таблица модели с существующими открытыми платформами, показатели эффективности представлены в виде графиков и таблиц.

Таблица - 1

Сравнение различных моделей управления доступом с предлагаемой моделью M-ABAC

№	Критерии	DAC	MAC	RBAC	ABAC	R-BAC	M-ABAC
1.	Возможность управления доступом на основе принципа наименьших привилегий	-	-	+	+	+	+
2.	Возможность динамического разделения обязанностей	-	-	+	+	-	+
3.	Возможность учета предоставленных прав доступа	+	+	+	+	+	+
4.	Наличие модуля синтаксической и семантической поддержки предоставленных прав доступа	-	-	-	-	-	+
5.	Соответствие общей политике управления	-	-	-	-	+	+
6.	Гибкость настройки	-	-	+	-	-	+
7.	Возможность оперативного (быстрого) оповещения о ситуации	-	-	-	-	+	-
8.	Интеграция с функцией аутентификации	-	-	-	-	-	-
9.	Уровень совместимости с операционной системой (возможность интеграции в операционную систему)	+	-	+	-	-	+
10.	Возможность тестирования и проверки функций контроля доступа	-	-	-	-	+	+
11.	Поддержка пассивных и активных потоков данных	-	-	-	-	-	-
12.	* Поддержка масштабной вертикальной и горизонтальной области действия	-	-	-	-	-	-
13.	Динамичность авторизации	+	-	-	+	+	+
14.	Возможность переноса учетных записей пользователей между слоями	-	-	-	-	+	+
15.	Масштабируемость	-	-	+	-	-	+
16.	Гибкость в управлении атрибутами	-	-	-	+	+	+

Результаты затрат времени программных средств, разработанных на основе моделей управления доступом на определение прав доступа субъекта (пользователя) в режиме реального времени в системе больших данных, представлены в таблице 2.

Таблица - 2

Результаты теста на время ответа на запрос субъекта
(час.минута.секунда=час.мин.сек)

№	Средство управления доступом	Модель управления доступом	Время отправки запроса (час.мин. сек.)	Время ответа на запрос (час.мин. сек.)	Время, затраченное на назначение прав доступа
1.	BDSS	M-ABAC	15.30.24	15.31.14	50 секунд
2.	Access Control Manager	DAC	15.36.13	15.37.12	59 секунд
3.	ZKBio Access IVS	MAC	15.43.19	15.44.15	56 секунд
4.	ZKAccess 3.5	RBAC	16.05.12	16.06.26	74 секунд
5.	HikCentral Access Control	ABAC	16.15.47	16.16.39	52 секунд
6.	BEWARD Access Control	R-ABAC	16.31.53	16.32.58	65 секунд

Полученные результаты по точности назачения прав доступа, предоставленных разработанным программным средством пользователям при использовании системы представлены в таблице 3.

Таблица - 3

Полученные результаты по точности назачения прав доступа, предоставленных пользователям при использовании системы

№	Средство управления доступом	Количество пользователей, получивщих доступ на ресурсы системы	Точность назначения прав доступа (в процентах и количестве пользователей)	Погрешность назначения прав доступа (в процентах)
1.	BDSS	114	95,6 (109)	4,4
2.	Access Control Manager	110	94,5 (104)	5,5
3.	ZKBio Access IVS	127	93,7 (119)	6,3
4.	ZKAccess 3.5	131	93,9 (123)	6,1
5.	HikCentral Access Control	95	94,7 (90)	5,3
6.	BEWARD Access Control	109	95,4 (104)	4,6

Полученные результаты по точности отказа в доступе пользователям на использование системы представлены в таблице 4.

Таблица - 4

Полученые результаты по точности отказа в доступе пользователям на использование системы

№	Средство управления доступом	Количество пользователей, отказанным в доступе на использование системы	Точность отказа в доступе (в процентах и количестве пользователей)	Погрешность отказа в доступе (в процентах)
1.	BDSS	36	91,7 (33)	8,3
2.	Access Control Manager	40	87,5 (35)	12,5
3.	ZKBio Access IVS	23	87 (20)	13
4.	ZKAccess 3.5	19	89,5 (17)	10,5
5.	HikCentral Access Control	55	90,9 (50)	9,1
6.	BEWARD Access Control	41	90,2 (37)	9,8

Результаты тестирования по первому из заданных показателей, то есть по ошибкам в таблицах, сформированных в результате анонимизации исходных данных, приведены в Таблице 5.

Таблица - 5

Результаты тестирования на погрешность в таблицах сформированных в результате анонимизации исходных данных

№	Название алгоритма	Точность (%)	Погрешность (%)
1.	Усовершенствованный алгоритм $K-$анонимизации	96,56	3,44
2.	Алгоритм EM	95,84	4,16
3.	Алгоритм Greedy	95,26	4,74
4.	Алгоритм $k-$ анонимизации	96,51	3,49
5.	Алгоритм l-Diversity	96,54	3,47
6.	Алгоритм $(k,e)-$ анонимизации	95,41	4,59
7.	Алгоритм $t-$Closeness	94,28	5,72
8.	Алгоритм $(X,Y)-$Privacy	95,14	4,86
9.	Алгоритм Multi R $k-$ анонимизации	93,78	6,22
10.	Алгоритм Distributional Privacy	94,85	5,15

Результаты тестирования по второму показателю — времени, затраченному на формирование анонимных таблиц — приведены в Таблице 6.

Таблица - 6

Результаты тестирования на затраченное время на генерацию анонимных таблиц

№	Название алгоритма	Затраченное время (в секундах)
1.	Усовершенствованный алгоритм $K-$анонимизации	0,05 с.
2.	Алгоритм EM	0,15 с.
3.	Алгоритм Greedy	0,08 с.
4.	Алгоритм $k-$ анонимизации	0,58 с.
5.	Алгоритм l-Diversity	0,18 с.
6.	Алгоритм $(k,e)-$ анонимизации	0,07 с.
7.	Алгоритм $t-$Closeness	0,54 с.
8.	Алгоритм $(X,Y)-$Privacy	0,35 с.
9.	Алгоритм Multi R $k-$ анонимизации	0,47 с.
10.	Алгоритм Distributional Privacy	0,51 с.

Результаты тестирования по третьему показателю, а именно по количеству наборов данных, требующих повторной идентификации, приведены в Таблице 7.

Таблица - 7

Результаты тестирования по количеству наборов, требующих повторной идентификации

№	Название алгоритма	Изначальное количество наборов	Количество наборов, требующих повторной идентификации
1.	Усовершенствованный алгоритм K-анонимизации	10	2
2.	Алгоритм EM	10	3
3.	Алгоритм Greedy	10	3
4.	Алгоритм k- анонимизации	10	4
5.	Алгоритм l-Diversity	10	5
6.	Алгоритм (k,e)- анонимизации	10	3
7.	Алгоритм t-Closeness	10	5
8.	Алгоритм (X,Y)-Privacy	10	4
9.	Алгоритм Multi R k-анонимизации	10	4
10.	Алгоритм Distributional Privacy	10	3

Согласно результатам тестирования, усовершенствованный алгоритм K-анонимизации обеспечивает анонимизацию исходных данных с точностью 96,56%, что на 0,02% превышает показатели других программных решений. Кроме того, по показателю ошибок в таблицах, сформированных в результате анонимизации, зафиксировано значение 3,44%, что на 0,03% ниже уровня ошибок по сравнению с существующими средствами.

ЗАКЛЮЧЕНИЕ

В результате проведенных исследований по диссертационной работе на тему «Модель и алгоритм защиты информации в жизненном цикле больших данных» сделаны следующие выводы:

1. Для обеспечения безопасности в среде больших данных был проведён анализ угроз и атак на каждом этапе жизненного цикла данных. В качестве модели цепочки безопасности была предложена модель угроз. Для выполнения требований информационной безопасности на каждом этапе жизненного цикла были классифицированы угрозы в системе управления большими данными. Предложена модель угроз информационной безопасности, учитывающая жизненный цикл фрагментов данных, источники угроз, направленных на системы управления базами данных, а также источники угроз в отношении систем обработки и хранения больших данных.

2. Усовершенствована модель контроля доступа ABAC за счет введения в набор атрибутов субъекта и объекта атрибутов безопасности. В результате разработанный на основе предложенной модели программное средство «BDSS dasturi» позволил повысить точность выдачи разрешений пользователям на 0,2 % и снизить погрешность выдачи несанкционированных разрешений на 0,8 %. На назначение прав доступа субъектам потребовалось на 3 сек. меньше времени, чем при использовании существующих программных инструментов.

3. Разработана архитектура контроля доступа, поддерживающая шифрование на основе политики ключевого атрибута. В результате данная архитектура обеспечила установление связи только между одобренными источниками данных путем проверки нового блока перед его добавлением в блокчейн.

4. Усовершенствована модель управления рисками информационной безопасности за счет использования классического метода нечеткой математической оценки для определения уровней рисков. В результате появилась возможность оценивать неопределенные риски в системе данных большого объема.

5. Разработан алгоритм анонимизации на основе алгоритмов «2-means» и «greedy». Разработанный в результате научного исследования программный модуль «Big Data security analyzer dasturi» позволил анонимизировать исходные данные с точностью 96,56 % и затратил на формирование анонимных таблиц на 0,03 сек. меньше времени, чем существующие.

6. На основе предложенной в диссертации модели и алгоритма были разработаны программные модули, определены технические требования для их использования, а также проведены экспериментальные испытания программного продукта. Представлены результаты, полученные от организаций. Согласно этим результатам, разработанное программное средство позволяет динамически изменять полномочия всех пользователей информационной системы организации и обеспечивает возможность интеграции с другими информационными системами, подключёнными к информационной инфраструктуре организации.

SCIENTIFIC COUNCIL AWARDING SCIENTIFIC DEGREES DSc.13/30.12.2019.T.07.02 AT TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES

TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES

AKHMEDOVA NOZIMA FARKHOD QIZI

METHOD AND ALGORITHM OF PROTECTION OF INFORMATION IN THE BIG DATA LIFECYCLE

05.01.05 –Methods and systems of information protection. Information security.

DISSERTATION ABSTRACT FOR THE DOCTOR OF PHILOSOPHY DEGREE (PhD) OF TECHNICAL SCIENCES

Tashkent - 2025

The theme of dissertation of doctor of philosophy (PhD) on technical sciences was registered at the Supreme Attestation Commission at the Ministry of Higher Education, Science and Innovations of the Republic of Uzbekistan under number B2025.1.PhD/T5293.

The dissertation has been prepared at the Tashkent University of information technologies named after Muhammad al-Khwarizmi.

The abstract of the dissertation is posted in three languages (Uzbek, Russian, English (resume)) on the Scientific Council website www.tuit.uz and on the website of «ZiyoNet» Informaton and Educational portal www.ziyonet.uz.

Scientific adviser:	**Tashev Komil Akhmatovich** Candidate of Technical Sciences, Associate Professor
Official opponents:	**Botirov Fayzulla Bakhtiyorovich** Doctor of Technical Sciences, Associate Professor
	Ibrokhimov Azizbek Ravshanbek ogli Doctor of Philosophy of Technical Sciences
Leading organization:	**LLC "UNICON"**

The defence will take place on «25» of october 2025 at 10:00 a.m. at the meeting of Scientific Council DSc.13/30.12.2019.T.07.02 at Tashkent University of information technologies. (Address: 100084, Tashkent city, Amir Temur Street, 108. Tel.: (99871) 238-64-15; e-mail: info@tuit.uz).

The dissertation could be reviewed in the Information Resource Centre of Tashkent university of information technologies named after Muhammad al-Khwarizmi. (Registration number № 369). (Address: 100084, Tashkent city, Amir Temur str., 108. Tel.: (99871) 238-64-15).

The abstract of dissertation is distributed on «15» october 2025.
(Protocol at the register № 10 on «15» october 2025)

B.Sh. Makhkamov
Chairman of the Scientific Council awarding scientific degrees, doctor of economical sciences, professor

M.S. Saitkamolov
Scientific secretary of Scientific Council awarding scientific degrees, doctor of economical sciences, associate professor

D.Ya. Irgasheva
Chairman of the academic Seminar under the Scientific Council awarding scientific degrees, doctor of technical sciences, professor

INTRODUCTION (abstract of PhD dissertation)

The aim of the research is to develop protection models and algorithms that prevent information leakage and unauthorized use of data in the lifecycle of big data.

The object of the research is the big data system.

The scientific novelty of the research work:

the attribute-based access control model has been improved by including security attributes in the set of subject and object attributes and five conditions that must be met in big data systems;

an access control architecture has been developed in big data systems that supports the key-policy attribute-based encryption, which enables the integration of the protection system with other systems in accordance with security requirements;

a risk management model has been improved, which allows the formation of a dynamic security cycle through the use of a classical fuzzy mathematical assessment method in the identification, assessment, and management of new types of risks in the big data system;

the model of information security risk management was improved by applying the classical method of fuzzy mathematical evaluation in the process of identification, evaluation and management of new types of risks in big data systems;

the K-anonymization algorithm has been improved, allowing to preserve data availability by applying elements of two-means clustering and “greedy” to the method of selecting the initial center of mass and the average center method.

Implementation of research results. On the basis of obtained scientific results on software tools developed on the basis of improved models and proposed algorithms in the big data system:

software tool “BDSS dasturi”, developed on the basis of an improved attribute-based access control model by including security attributes in the set of subject and object attributes, as well as five conditions that must be met in big data systems, implemented in the practical activities of the State Unitary Enterprise “Cyber Security Center” (Reference of the Ministry of Digital Technologies dated November 06, 2024 y. № 24-8/7511). The result of the scientific research made it possible to increase the accuracy of granting access rights to users by 0.2% and reduce the error of granting access rights to unauthorized users by 0.8%;

software tool based on the developed access control architecture for integration of the protection system with other systems in accordance with the security requirements in big data systems, supporting key-policy attribute-based encryption, implemented in the practical activity of State Unitary Enterprise “Cyber Security Center” (Reference of the Ministry of Digital Technologies from November 06, 2024 y. № 24-8/7511). The thesis emphasizes the possibility of using a software tool developed on the basis of the proposed models and algorithms to protect against threats to confidentiality, integrity and availability, which are the foundations of information security, which arise due to improper granting of privileges to the user when managing access to big data, and also gives suggestions and recommendations.

software tool “BDSS dasturi”, developed on the basis of a multi-stage model

of information security risk management, improved by applying the classical method of fuzzy mathematical evaluation in the process of identification, evaluation and management of new types of risks in big data systems, implemented in the practical activities of the State Unitary Enterprise “Center of Radio Communication, Broadcasting and Television” (Reference of the Ministry of Digital Technologies dated November 06, 2024 y. № 24-8/7511). As a result of scientific research, the software tool took 3 seconds less time to assign access rights to subjects;

software tool “Big Data security analyzer dasturi” of k-anonymization algorithm improved in order to prevent information leaks in big data systems and preserve data availability on the basis of “2-means” and ‘greedy’ clustering algorithms in big data systems was implemented in practical activity of “UZINFOCOM - single integrator for creation and support of state information systems” LLC (Reference of the Ministry of Digital Technologies dated November 06, 2024 y. No. 24-8/7511). As a result of scientific research, the software tool took 0.03 seconds less time to form anonymous tables than the existing ones.

The structure of the dissertation. The dissertation consists of an introduction, four chapters, conclusion, references and appendices. The volume of the dissertation is 112 pages.

E’LON QILINGAN ISHLAR RO‘YXATI
СПИСОК ОПУБЛИКОВАННЫХ РАБОТ
LIST OF PUBLISHED WORKS

I bo‘lim (I частъ; I part)

1. Akhmedova N., Tashev K, Kalauov S. Big Data Technologies in Transport Planning // International Conference on “Information Science and Communications Technologies (ICISCT 2021)”. Tashkent – 2021. -5 p. (3) Scopus, (ОАК Раёсатининг қарори 30.10.2021 й. №308/6).

2. Akhmedova N. The Role of Big Data in Intelligent Transport Systems // Scientific journal “Research and education” Volume 2, Issue 4. Tashkent-2023. ISSN: 2181-3191. –P. 101-108. (14) ResearchBib.

3. Akhmedova N. A study of security problems in big data and their solutions // J.:“Chemical technology, control and management” 2020, No4(94) International scientific and technical journal, ISSN 1815-4840, E-ISSN 2181-1105 (05.00.00; №12).

4. Axmedova N.F., Tashev K.A. Transport tizimlarining raqamli transformatsiyasida Big Data texnologiyasi // “Муҳаммад ал-Хоразмий авлодлари” илмий-амалий ва ахборот-таҳлилий журнали 2(24)/2023, Тошкент-2023й. – B. 212-218 (05.00.00; №10).

5. Ахмедова Н., Ташев К., Рустамов Д. Классификация угроз кибербезопасности в больших данных и механизм защиты от атак // “ВЕСТНИК ТУИТ” Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universitetining ilmiy-texnika va axborot-tahliliy jurnali 2 (70)/2024. Toshkent-2024. –C. 88-99 (05.00.00; №31).

6. Axmedova N.F., Rustamov D.T., Jafarov M.M. Kiberxavfsizlik tahdidi atributlari va ulardan himoyalanish mexanizmi // Ахбороткоммуникациялар: тармоқлар-технологиялар-ечимлар. АК:ТТЕ № 3 (71)/2024. –B. 45-52 (05.00.00; №2).

7. Axmedova N.F. Katta hajmli ma’lumotlarni anonimlashtirish: D2D tarmoqlarda axborot xavfsizligi yondashuvi // “Муҳаммад ал-Хоразмий авлодлари” илмий-амалий ва ахборот-таҳлилий журнали. № 2(32), 2025. Тошкент -2025 й. -B. 110-113 (05.00.00; №10).

II bo‘lim (II частъ; II part)

8. Akhmedova N. Mechanism for digital transformation of intelligent transport systems // 5th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks “ICICV 2023”, India, Tirunelveli-2023, -P. 406-416.

9. Akhmedova N.F., Tashev K.A. Methods for Securing Big Data // International Conference on “Trends in Sustainable Computing and Machine Intelligence” (ICTSM 2023). October 5-6. Bangkok - 2023. –P. 397-413.

10. Akhmedova N., Mirzaev D. Attribute Based Access Control Method in Big Data Technologies // The eight International Conference on Future Networks & Distribution Systems. ICFNDS ’24, Marakech-2024. -P. 459-464.

11. Ахмедова Н., Рустамова С. Дунё мамлакатларида йўл ҳаракати хавфсизлигини таъминлашда ахборот коммуникация технологияларининг қўлланилиши // “Йўл ҳаракати хавфсизлигини таъминлашнинг долзарб муаммолари ва бу соҳада ҳуқуқбузарликларнинг профилактикаси” мавзусидаги халқаро илмий-амалий конференция материаллари тўплами, Тошкент – 2021 й., -Б. 262-267.

12. Akhmedova N., Khamzayev J. Problems of intelligent transport system and solutions with big data // International conference “Recent advances in intelligent information and communication technologies “ISPC-2022””, Tashkent-2022, -P. 34-42.

13. Akhmedova N.F., Tashev K.A. Theory of the development process of a software module that analyzes and ensures the security of big data in intelligent transport systems // International scientific and technical conference “digital technologies: problems and solutions of practical implementation in the spheres” April 27-28. Tashkent-2023. –P. 825-831.

14. Akhmedova N. Unveiling the Power and Challenges of Big Data Platforms // “Axborot texnologiyalari va kommunikatsiyalari sohasida axborot xavfsizligi va kiberxavfsizlik muammolari” Respublika ilmiy-amaliy anjumani ma’ruzalar to‘plami. Toshкent-2023. –P. 87-93.

15. Akhmedova N. Safeguarding the big data frontier: exploring existing solutions for security challenges // “Kiberxavfsizlikni ta’minlash va axborot texnologiyalari sohasidagi jinoyatlarga qarshi kurashishni takomillashtirish istiqbollari” Respublika ilmiy-amaliy konferensiya materiallari to‘plami. Toshkent – 2023. –P. 110-116.

16. Ахмедова Н. Технологии и проблемы безопасности Big Data // Сборник докладов республиканской научно-практической конференции “Проблемы применения современных информационных, коммуникационных технологий и IT-образования”. Самарканд-2025. –С. 328-330.

17. Axmedova N.F. “BDSS” дастури. O‘zbekiston Respublikasi Adliya Vazirligi. EHM uchun yaratilgan dasturning rasmiy ro‘yxatdan o‘tkazilganligi to‘g‘risidagi guvohnoma. №DGU 24182, 28.03.2023.

18. Axmedova N.F., Tashev K.A. “Big Data security analyzer” дастури. O‘zbekiston Respublikasi Adliya Vazirligi. EHM uchun yaratilgan dasturning rasmiy ro‘yxatdan o‘tkazilganligi to‘g‘risidagi guvohnoma. №DGU 24183, 28.03.2023.

19. Axmedova N.F., Tashev K.A. “Secure Big Data in ITS” дастури. O‘zbekiston Respublikasi Adliya Vazirligi. EHM uchun yaratilgan dasturning rasmiy ro‘yxatdan o‘tkazilganligi to‘g‘risidagi guvohnoma. №DGU 24184, 28.03.2023.

Avtoreferat «Muhammad al-Xorazmiy avlodlari» ilmiy jurnali tahririyatida o‘zbek, rus va ingliz tillaridagi matnlarining mosligi tekshirildi.

Bosmaxona litsenziyasi:

Bichimi: 84x60 $^{1}/_{16}$. «Times New Roman» garniturasi.
Raqamli bosma usulda bosildi.
Shartli bosma tabog‘i: 2,75. Adadi 100 dona. Buyurtma № 31/24.

Guvohnoma № 851684.
«Tipograff» MCHJ bosmaxonasida chop etilgan.
Bosmaxona manzili: 100011, Toshkent sh., Beruniy ko‘chasi, 83-uy.