JOURNAL OF INTERNATIONAL SCIENTIFIC RESEARCH
Volume 3, Issue 5, October 2025 Online ISSN: 3030-3508
https://spaceknowladge.com

SHAXSGA DOIR MOLIYAVIY MA'LUMOTLARNI RUXSATSIZ CHIQIB KETISHINI ANIQLASH VA BARTARAF ETISH

Bozorov Adham Norqobil o'g'li
Toshkent iqtisodiyot va pedagogika instituti o'qituvchisi
E-mail: adhambozorov1996@gmail.com
Turgunova Ruxshona Xurshed qizi
Toshkent iqtisodiyot va pedagogika instituti talabasi
E-mail: rxurshedovna06@gmail.com

MAQOLA MALUMOTI

MAQOLA TARIXI:

Received:20.11.2025
Revised: 21.11.2025
Accepted:22.11.2025

KALIT SO'ZLAR:

shaxsga doir moliyaviy ma'lumotlar, data leakage, DLP, formal model, logistik regressiya, NER, Luhn algoritmi, konfidensiallik darajasi

ANNOTATSIYA:

Maqolada shaxsga doir moliyaviy ma'lumotlarni (bank kartasi rekvizitlari, hisob raqamlari, kredit shartnomalari va boshqalar) ruxsatsiz chiqib ketishini aniqlash va oldini olishning matematik modeli yoritiladi. Hujjatlar fazosi, konfidensiallik darajasi, moliyaviy atributlar vektori va ruxsatsiz chiqib ketish hodisasi formal ko'rinishda ta'riflanadi. Deterministik (pattern va shablonlar, Luhn algoritmi) hamda ehtimollik yondashuvlari (ko'p sinfli logistik regressiya, NER – Named Entity Recognition) asosida aniqlash usullari taklif etiladi. Shuningdek, DLP tizimining samaradorligini TP, FP, FN, TN ko'rsatkichlari orqali baholash uchun asosiy metrikalar (Accuracy, Precision, Recall, F1-score) keltiriladi. Taklif etilgan matematik model bank va moliyaviy tashkilotlar uchun shaxsga doir ma'lumotlar xavfsizligini oshirishga xizmat qiladi.

Kirish. Raqamli iqtisodiyot sharoitida onlayn-banking, mobil ilovalar, fintech platformalar va raqamli to'lov xizmatlarining jadal rivojlanishi shaxsga doir moliyaviy ma'lumotlar hajmini keskin oshirmoqda. Bank kartalari, hisob raqamlari, kredit shartnomalari, mijozlar reyestrlari, to'lovlar jurnallari, mobil ilova orqali tranzaksiyalar tarixi kabi axborotlar ko'plab tashkilotlar serverlarida saqlanadi va qayta ishlanadi. Ushbu

ma’lumotlarning ruxsatsiz uchinchi shaxslarga chiqib ketishi (data leakage) fuqarolarga moliyaviy zarar, kredit tarixiga putur yetishi, firibgarlik operatsiyalarining ko‘payishi bilan birga, bank va moliya institutlari uchun ham sezilarli reputatsion va huquqiy xavflarni yuzaga keltiradi.

Shaxsga doir moliyaviy ma’lumotlar xalqaro standartlarda (PCI DSS, GDPR, ISO/IEC 27001 va boshqalar) eng yuqori himoya talab qiluvchi “sensitive payment data” sifatida ko‘riladi. Shu bois, bunday ma’lumotlar bilan bog‘liq axborot oqimlarini nazorat qilish, ruxsatsiz chiqib ketish holatlarini erta aniqlash va bartaraf etish uchun qat’iy formal model va matematik apparatga asoslangan DLP (Data Loss Prevention) yechimlari zarur.

Mazkur ishda shaxsga doir moliyaviy ma’lumotlarning formal modeli, ruxsatsiz chiqib ketish hodisasining matematik ifodalanishi, deterministik va ehtimollik usullari yordamida aniqlash mexanizmlari hamda tizim samaradorligini baholash mezonlari ko‘rib chiqiladi.

Adabiyotlar tahlili. Raqamli iqtisodiyot konsepsiyasi 1990-yillarning oxirida Don Tapscott tomonidan kiritilgan bo‘lib, u iqtisodiy faoliyatning tobora ko‘proq qismi internet texnologiyalari, raqamli ma’lumotlar va kommunikatsiya tarmoqlariga asoslangan yangi iqtisodiy tuzilmani ifodalaydi. OECD ta’rifiga ko‘ra, raqamli iqtisodiyot — bu raqamli texnologiyalar, ma’lumotlar va ulanishlarni ishlab chiqarish, tarqatish va iste’mol qilishga asoslangan iqtisodiyotdir . Shu nuqtai nazardan, shaxsga doir moliyaviy ma’lumotlar xavfsizligini ta’minlash, ruxsatsiz chiqib ketishlarni aniqlash va oldini olish masalalari ham raqamli iqtisodiyotning muhim tarkibiy qismi sifatida qaraladi. Chunki moliyaviy ma’lumotlar oqimi tez va keng miqyosda elektron kanallar orqali amalga oshadi, va ularning himoyasi nafaqat bank mijozlari, balki butun moliyaviy tizim uchun zarurdir.

Ilmiy adabiyotlarda Data Loss Prevention (DLP) tizimlari va shaxsga doir moliyaviy ma’lumotlarni aniqlash mexanizmlari ko‘plab tadqiqotlarda yoritilgan. Masalan, Shabtai va boshqalar DLP tizimlari orqali ruxsatsiz chiqib ketishlarni aniqlash va oldini olishning turli usullarini tahlil qilgan bo‘lib, deterministik yondashuvlar (pattern va shablonlar, Luhn algoritmi) hamda ehtimollik asosidagi yondashuvlar (logistik regressiya, NER) samaradorligini ta’kidlaydi. Ular DLP tizimining samaradorligini baholashda TP, FP, FN va TN ko‘rsatkichlari hamda Accuracy, Precision, Recall, F1-score kabi metrikalardan foydalanishni tavsiya qiladilar. Shuningdek, Campanile va boshqalar NER (Named Entity Recognition) metodini qo‘llash orqali hujjat matnidan sezgir moliyaviy entitetlarni ajratib olish usullarini ko‘rib chiqqanlar. Bu yondashuv deterministik patternlarga qaraganda kontekstni chuqurroq hisobga olish imkonini beradi va DLP tizimlarining aniqlik va

to‘liqlik darajasini oshiradi. Shu bilan birga, klassifikatsiya va ehtimollik modellarini birlashtirish orqali ruxsatsiz chiqib ketish ehtimolini aniqlash va xavfni oldindan baholash imkoniyati yuzaga keladi.

Xulosa qilib aytganda, ilmiy adabiyotlar tahlili shuni ko‘rsatadiki, shaxsga doir moliyaviy ma’lumotlarni himoya qilish masalas i nafaqat texnologik, balki iqtisodiy va huquqiy kontekstda ham dolzarbdir. Deterministik va ML asosidagi metodlarni birlashtirgan formal matematik model DLP tizimlarining samaradorligini oshirishga, moliyaviy xavfsizlikni ta’minlashga va raqamli iqtisodiyot sharoitida bank va moliyaviy tashkilotlar uchun ishonchli himoya mexanizmlarini yaratishga xizmat qiladi.

Shaxsga doir moliyaviy ma’lumotlarning formal modeli. Hujjatlar fazosi va konfidensiallik darajasi

Tizimdagi barcha hujjatlar va fayllar to‘plamini

$$\mathcal{D} = \{d_1, d_2, \dots, d_N\}$$

bilan belgilaymiz. Har bir hujjat:

$$d = (x, m)$$

ko‘rinishda tasvirlanadi, bu yerda:

- x– hujjatning matnli yoki strukturalangan mazmuni (matn, jadval, XML, JSON va hokazo),
- m– metadata (yaratilgan sana, muallif, fayl turi, kanal, o‘lcham va boshqalar).

Hujjat konfidensiallik darajasi quyidagi sinflardan biriga tegishli bo‘ladi:

$$\mathcal{Y} = \{\text{"ochiq"}, \text{"konfidensial"}, \text{"maxfiy"}\}.$$

Amaliy jihatdan bu sinflar sonli ko‘rinishga o‘tkaziladi:

$$\text{"ochiq"} \to 0, \text{"konfidensial"} \to 1, \text{"maxfiy"} \to 2;$$

va natijada hujjat uchun konfidensiallik darajasi

$$C(d) \in \{0,1,2\}$$

ko‘rinishda aniqlanadi. Bu belgilash keyinchalik klassifikatsiya va siyosatlarni avtomatlashtirishda qulaylik yaratadi.

Shaxsga doir moliyaviy atributlar vektori

Shaxsga doir moliyaviy ma’lumotlar uchun maxsus atributlar to‘plamini quyidagicha belgilaymiz:

$$\mathcal{A} = \{\text{KARTA\_RAQAM, CVV, PIN, HISOB\_RAQAM,}$$
$$\text{TRX\_ID, INN, KREDIT\_SUMMA, ... }\}.$$

Har bir hujjat uchun ushbu atributlarning mavjudligini ko‘rsatuvchi indikator vektor quramiz:

$$z(d) = (z_1, z_2, \dots, z_p), z_i \in \{0,1\}.$$

Bu yerda $z_i = 1$, agar d hujjatda mos moliyaviy atribut aniqlansa, aks holda $z_i = 0$. Mazkur vektor hujjatning “moliyaviy sezgirlik darajasi”ni formal tasvirlashga imkon beradi.

Ruxsatsiz chiqib ketish hodisasining matematik modeli

DLP tizimi nuqtai nazaridan qiziqarli obyekt – bu ma’lum hujjatni ma’lum kanal orqali tashqi tomonga uzatish hodisasidir. Uni quyidagicha modellaymiz:

$$e = (u, d, c, t),$$

bu yerda:

- $u \in \mathcal{U}$– foydalanuvchi yoki xodim,
- $d \in \mathcal{D}$– uzatilayotgan hujjat,
- $c \in \mathcal{C}$– aloqa kanali (USB, e-mail, web, printer, messenjer va boshqalar),
- t– hodisa sodir bo‘lgan vaqt tamg‘asi.

Ruxsatsiz chiqib ketish holatini ikkilik tasodifiy o‘zgaruvchi sifatida belgilaymiz:

$$L = \begin{cases} 1, & \text{agar hodisa real ruxsatsiz chiqib ketish bo‘lsa,} \\ 0, & \text{aks holda.} \end{cases}$$

Bizni qiziqtiradigan kattalik – hodisa uchun ruxsatsiz chiqib ketish ehtimoli:

$$P(L = 1 \mid e) = P(L = 1 \mid u, d, c, t).$$

DLP tizimi maqsadi – ushbu ehtimol yuqori bo‘lgan hodisalarni aniqlash va ularga texnik hamda tashkiliy choralarni qo‘llashdir (bloklash, ogohlantirish, loglash va h.k.), lekin ushbu ishda alohida risk balli formulasi kiritilmaydi, balki aniqlash mexanizmlarining o‘zi formalashtiriladi.

Deterministik va ML asosida chiqib ketishlarni aniqlash

Pattern va shablonlar bo‘yicha aniqlash

Kartalar raqami, hisob raqami, kredit shartnomasi raqami kabi ma’lumotlar odatda aniq raqamli/shablonli tuzilishga ega. Har bir moliyaviy atribut uchun pattern-funksiya aniqlaymiz:

f_i(d) = \begin{cases} 1, & \text{agar } d \text{ hujjatda } i\text}-pattern aniqlansa,\\ 0, & \text{aks holda}. \end{cases}

Patternlar muntazam ifodalar (regex) yoki raqamlar uzunligi va formatiga asoslanadi. Masalan, bank kartasi raqami 16 xonali bo‘ladi, turli BIN diapazonlari ma’lum brendlarga mos keladi.

Luhn algoritmi yordamida karta raqamini tekshirishda 16 xonali raqam $(d_1, d_2, \dots, d_{16})$sifatida olinib, maxsus transformatsiya asosida yig‘indi Shisoblanadi. Agar

$$S \bmod 10 = 0$$

sharti bajarilsa, mazkur raqam karta raqamiga mos kelishi ehtimoli yuqori deb olinadi.

Hujjat bo‘yicha barcha patternlar natijalarini

$$p(d) = (f_1(d), f_2(d), \dots, f_m(d))$$

vektori ko‘rinishida jamlash mumkin. Ushbu deterministik vektor keyinchalik ML modelga xususiyat sifatida kiritiladi.

ML asosida hujjat klassifikatsiyasi

Hujjat matnidan xususiyatlar vektori olish uchun xaritalash funksiyasini aniqlaymiz:

$$\varphi: \mathcal{D} \to \mathbb{R}^n, \mathbf{x} = \varphi(d).$$

Bu yerda **x**tarkibiga TF-IDF, n-gram, patternlar soni, fayl tipi kabi parametrlar kiritilishi mumkin.

Konfidensiallik sinfini aniqlash uchun ko‘p sinfli logistik regressiya (softmax) modeli qo‘llanadi:

$$P(Y = k \mid \boldsymbol{x}) = \frac{exp\,(\boldsymbol{\theta}_k^\top \boldsymbol{x})}{\sum_{j=0}^{2} exp\,(\boldsymbol{\theta}_j^\top \mathbf{x})}, k = 0,1,2.$$

Natijada hujjatning prognoz qilingan sinfi:

$$\hat{y}(d) = \arg\max_k P(Y = k \mid \mathbf{x})$$

ko‘rinishda aniqlanadi. Model parametrlari kross-entropiya yo‘qotish funksiyasini minimallashtirish orqali o‘qitiladi.

NER (Named Entity Recognition) yordamida entitetlarni ajratish

Matnni tokenlar ketma-ketligi sifatida qaraymiz:

$$x = (w_1, w_2, \dots, w_T).$$

Har bir token uchun teg (entitet turi):

$$z_t \in \{\text{B-KARTA, I-KARTA, B-HISOB, B-MOL\_SUM, O, ... }\}.$$

NER modelining vazifasi – berilgan matn uchun eng ehtimolli teglar ketma-ketligini topish, ya’ni $P(\mathbf{z} \mid x; \theta)$ni maksimum qilishdir. BERT yoki LSTM-CRF kabi modellar yordamida shaxsga doir moliyaviy entitetlar (karta raqami, hisob raqami, kredit summa va boshqalar) aniqlanadi va ular asosida $\mathbf{z}(d)$indikator vektori to‘ldiriladi. Bu deterministik patternlarga qaraganda kontekstni chuqurroq hisobga olish imkonini beradi.

Siyosatni optimallashtirish va samaradorlikni baholash

Kutilayotgan zarar va siyosat

Shaxsga doir moliyaviy ma’lumotlar ruxsatsiz chiqib ketishi bilan bog‘liq har bir hodisa uchun kutiladigan zarar:

$$\text{Cost}(e) = \alpha \cdot \text{Moliyaviy zarar} + \beta \cdot \text{Obro‘ga zarar} + \gamma \cdot \text{Huquqiy jarimalar}$$

ko‘rinishida baholanishi mumkin. DLP siyosati πtanlanganida, kutilayotgan umumiy zarar:

$$\mathbb{E}[\text{Damage} \mid \pi] = \sum_{e \in \mathcal{E}} P(L = 1 \mid e, \pi) \cdot \text{Cost}(e)$$

ko‘rinishda yoziladi. Bu ifoda DLP siyosatini tanlashda xavfni miqdoriy baholash imkonini beradi. Amaliyotda xavfsizlik bilan birga foydalanish qulayligi ham muhim bo‘lganligi sababli, siyosatga qo‘shimcha “usability” xarajati ham hisobga olinadi; lekin mazkur ish doirasida risk balli va qaror qoidalarining alohida matematik formulalari keltirilmaydi.

5.2. Klassifikatorni baholash ko‘rsatkichlari

Shaxsga doir moliyaviy ma’lumotlarni aniqlash masalasini “maxfiy ma’lumot bor/yo‘q” ko‘rinishiga keltirib, quyidagi to‘rt ko‘rsatkich aniqlanadi:

- **TP (True Positive)** – real maxfiy hujjat to‘g‘ri aniqlangan holatlar,
- **FP (False Positive)** – aslida maxfiy bo‘lmagan hujjat noto‘g‘ri ravishda maxfiy deb belgilangan holatlar,
- **FN (False Negative)** – maxfiy hujjat aniqlanmay qolgan holatlar,
- **TN (True Negative)** – maxfiy bo‘lmagan hujjat to‘g‘ri oddiy deb baholangan holatlar.

Asosiy metrikalar:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN},$$

$$\text{Precision} = \frac{TP}{TP + FP},$$

$$\text{Recall} = \frac{TP}{TP + FN},$$

$$F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}.$$

DLP tizimlari uchun ayniqsa Recall muhim, chunki FN – aniqlanmay qolgan ruxsatsiz chiqib ketishlar sonini bildiradi. Shu bilan birga FP ko‘rsatkichining ortib ketmasligi tizimning amaliy qulayligi uchun zarur, shuning uchun Precision ham nazorat qilinadi. Bu

metrikalar turli modellarning (pattern-only, ML-only, ML+NER) samaradorligini solishtirish va eng maqbul yechimni tanlashga yordam beradi.

Xulosa. Maqolada shaxsga doir moliyaviy ma'lumotlarni chiqib ketishini aniqlash va bartaraf etish masalasi uchun formal matematik model taklif etildi. Hujjatlar fazosi $\mathcal{D}$, konfidensiallik darajasi $\mathrm{C}(\mathrm{d})$, moliyaviy atributlar vektori $\mathbf{z}(\mathrm{d})$hamda ruxsatsiz chiqib ketish hodisasi $\mathrm{e} = (\mathrm{u}, \mathrm{d}, \mathrm{c}, \mathrm{t})$formal ko'rinishda ifodalandi. Deterministik usullar (pattern va Luhn algoritmi) yordamida moliyaviy atributlarni aniqlash, ko'p sinfli logistik regressiya asosida hujjatlarni "ochiq/konfidensial/maxfiy" sinflarga ajratish va NER modeli yordamida matndan sezgir moliyaviy entitetlarni ajratib olish mexanizmlari bayon qilindi.

Shuningdek, DLP siyosatini tanlashda kutilayotgan zarar tushunchasi va klassifikatorni baholash uchun Accuracy, Precision, Recall, F1 kabi metrikalar keltirildi. Taklif etilgan matematik asoslar shaxsga doir moliyaviy ma'lumotlar bilan ishlovchi bank va moliyaviy tashkilotlarda DLP tizimlarini loyihalash, baholash va takomillashtirishda nazariy poydevor bo'lib xizmat qiladi.

Foydalanilgan adabiyotlar

1. A. Shabtai, Y. Elovici, L. Rokach, A Survey of Data Leakage Detection and Prevention Solutions, Springer (2012).

2. PCI Security Standards Council, Payment Card Industry Data Security Standard (PCI DSS), Version 4.0, PCI SSC (2022).

3. ISO/IEC 27001:2022, Information Security, Cybersecurity and Privacy Protection — Information Security Management Systems — Requirements, ISO (2022).

4. L. Campanile, M.S. de Biase, S. Marrone, F. Marulli, M. Raimondo, L. Verde, Sensitive Information Detection Adopting Named Entity Recognition: A Proposed Methodology, Lect. Notes Comput. Sci. 13378 (2022) 377–388.

5. L. Hernández Aros, L.X. Bustamante Molano, F. Gutierrez-Portela, J.J. Moreno Hernández, M.S. Rodríguez Barrero, Financial fraud detection through the application of machine learning techniques: a literature review, Humanit. Soc. Sci. Commun. 11 (2024) Article 36-06.