KONFIDENSIAL FAYLLARNING HARAKATINI ANIQLASH VA NAZORAT QILISH TIZIMI

Bozorov Obidjon Norqobilovich

Mirzo Ulug'bek nomidagi O'zbekiston Milliy Universiteti
E-mail: bozorov.obid@gmail.com

K E Y W O R D S

konfidensial fayllar, axborot xavfsizligi, real vaqtli monitoring, fayl harakati nazorati, kontekstual tahlil, DLP tizimi, xavf baholash modeli, TF-IDF algoritmi, Windows Servis, fayl kontenti tahlili, audit-log taqqoslash, ichki tahdidlarni aniqlash

A B S T R A C T

Bugungi raqamli transformatsiya jarayonida axborot xavfsizligini ta'minlash masalasi dolzarb muammolardan biri hisoblanadi. Ayniqsa, tashkilotlarda mavjud konfidensial ma'lumotlarning ichki foydalanuvchilar tomonidan tarqalib ketish xavfi ortib borayotgan bir paytda, ularning harakatini real vaqt rejimida nazorat qilish zarurati tobora kuchaymoqda. Ushbu maqolada Windows operatsion tizimi asosida ishlab chiqilgan servis yordamida konfidensial axborot fayllarining harakatini – jumladan, ularni yaratish, ko'chirish, o'chirish, ochish, qayta nomlash va boshqa faoliyatlarini kuzatish usuli taklif etiladi. Taklif etilgan yondashuv fayllarning kontentiga oid konfidensiallik mezonlarini matematik modellashtirish asosida aniqlaydi hamda fayllar harakati haqidagi loglarni tahlil qilish orqali samarali nazoratni ta'minlaydi. Maqolada fayl harakatini aniqlash algoritmi, matematik formulalar, real vaqtli monitoring mexanizmi va yaratilgan tizimning sinovdan o'tkazish natijalari keltirilgan. Tadqiqot natijalari taklif etilgan usul orqali konfidensial axborotni ishonchli nazorat qilish mumkinligini va mavjud tizimlarga nisbatan aniqlik va samaradorlikning oshishini ko'rsatadi.

Kirish

Raqamli texnologiyalar jadal rivojlanayotgan hozirgi davrda axborot xavfsizligini ta'minlash global miqyosdagi eng muhim masalalardan biriga aylanmoqda. Ayniqsa, tashkilot va muassasalarda mavjud bo'lgan konfidensial axborotlarning ichki foydalanuvchilar tomonidan tasodifiy yoki ataylab tarqalib ketish xavfi sezilarli darajada ortib bormoqda [1]. Bunday sharoitda, an'anaviy himoya vositalari (kirish huquqini cheklash, shifrlash, parollar, va h.k.) ko'pincha yetarli bo'lmay, axborot harakatlarini real vaqt rejimida kuzatib borish, ularni baholash va nazorat qilishga ehtiyoj kuchaymoqda.

Mazkur tadqiqotda Windows operatsion tizimi asosida ishlab chiqilgan maxsus servis yordamida konfidensial fayllarning harakatini — ya'ni ularni yaratish, o'chirish, ochish, qayta nomlash, tarmoq orqali yuborish, USB qurilmaga ko'chirish yoki bulutga yuklash kabi amallarni aniqlash va monitoring qilish usuli taklif etiladi. Faylga oid har qanday harakat quyidagi to'rtlik asosida ifodalanadi (1):

$$H_f(t,u,a,p) = (s(t),u(t),a(t),p(t)) \quad (1)$$

bu yerda:

– $s(t) \in \{0,1,2,3,4\}$ – fayl holati:

Bu model harakatlarni kontekstual ravishda baholash va tizimda yuz berayotgan barcha faoliyatni to'liq nazorat qilishga zamin yaratadi. Aynan shu asosda faylga nisbatan umumiy xavf darajasi quyidagi matematik ifoda bilan baholanadi (2):

$$R_f(t) = \sum_{i=0}^{4} w_i \cdot \delta_i(t) + \beta_1 \cdot \gamma_u(t) + \beta_2 \cdot \gamma_a(t) + \beta_3 \cdot \gamma_p(t) \quad (2)$$

bu yerda:

– $\delta_i(t) = \begin{cases} 1, & agar\ holat\ iaktivbo'lsa \\ 0, & aks\ holda \end{cases}$;

– w_i – fayl holatlariga mos xavf og'irligi,

– $\gamma_u(t), \gamma_a(t), \gamma_p(t)$ – foydalanuvchi, dastur va vaqt konteksti uchun xavf darajalari,

– $\beta_1, \beta_2, \beta_3 \in [0,1]$ – ularning ta'sir

©O.N. Bozorov, F.R. Muhammadiyev

og'irliklari bo'lib, har bir kontekstual parametrning umumiy xavf bahosiga qo'shgan hissasini ifodalaydi.

Agar β_j qiymati 0 ga yaqin bo'lsa - ushbu parametr umumiy xavf bahosiga kam ta'sir qiladi; agar $\beta_j = 1$ bo'lsa - bu parametr xavf bahosiga maksimal darajada ta'sir ko'rsatadi. Shuningdek:

– $\gamma_u(t) = 1$, agar foydalanuvchi roli mehmon bo'lsa; 0 – agar admin bo'lsa;

– $\gamma_a(t) = 1$, agar dastur untrusted bo'lsa;

– $\gamma_p(t) = 1$, agar faoliyat ish vaqtidan tashqari sodir bo'layotgan bo'lsa.

Mazkur yondashuv real vaqt rejimida har bir fayl holatini nafaqat aniqlash, balki xavf darajasini aniqlash imkonini ham beradi.

Tadqiqotning maqsadi — konfidensial axborotlarning harakatini aniqlash va nazorat qilishning matematik asoslangan, real vaqtli, moslashuvchan va ishonchli usulini ishlab chiqish hamda uni dasturiy mahsulot shaklida amalga oshirishdan iborat.

Tadqiqotning vazifalari quyidagilardan iborat:

- konfidensial fayllarni semantik va matematik mezonlar asosida avtomatik aniqlash modelini ishlab chiqish;
- fayl harakatini to'liq aks ettiruvchi monitoring usullarini yaratish;
- fayl harakatlariga asoslangan xavf darajasini baholovchi modellarni ishlab chiqish;
- dasturiy servisni Windows muhitida testdan o'tkazish va amaliy natijalarni baholash.

Mavzuga oid adabiyotlar tahlili

Axborot xavfsizligini ta'minlash borasida ko'plab tadqiqotlar olib borilgan bo'lsa-da, konfidensial fayllarning real vaqt rejimida harakatini kontekstual tahlil asosida aniqlash va baholash muammosi hali to'liq yechim topmagan.

W. Stallings [2] o'zining "Information Security" kitobida axborotni himoya qilishning asosiy tamoyillarini (maxfiylik, yaxlitlik, mavjudlik) ta'riflab, xavfsizlik siyosatini shakllantirishga e'tibor qaratadi, biroq u real vaqtli harakatni tahlil qilish masalasini ochib bermaydi.

M. Bishop va D. Bailey [3] riskni baholash usullarini chuqur tahlil qilgan bo'lsa-da, ular ko'proq tashqi tahdidlar va reaktiv yondashuvlarga asoslangan. Aksincha, taklif etilgan model ichki foydalanuvchi harakatlarini proaktiv baholashga yo'naltirilgan.

OpenDLP va shunga o'xshash ochiq kodli tizimlar ko'pincha offline tahlil (diskni skanerlash) bilan cheklanadi va fayl harakatining konteksti (foydalanuvchi roli, dastur ishonchliligi, vaqt) hisobga olinmaydi. Shuningdek, matematik xavf modeli mavjud emas [8].

D. Jurafsky va J. Martin [6] tomonidan taklif etilgan TF-IDF yondashuvi matnni tasniflashda samarali bo'lsa-da, uni xavfsizlik tizimiga moslashtirishga doir tadqiqotlar kam. Ushbu maqolada aynan shu yondashuv konfidensial fayllarni aniqlash uchun moslashtirilgan.

Xulosa qilib aytganda, mavjud adabiyotlarda konfidensial fayllarni real vaqtli, kontekstual va matematik modellashtirilgan tarzda baholovchi yondashuv yetarli darajada yoritilmagan[7-10].

Nazariy va metodik asoslar

Tadqiqotda konfidensial axborotlarning harakatini monitoring qilish uchun taklif etilayotgan yondashuv axborot xavfsizligi, real vaqtli tizimlar va hisoblash texnikasining nazariy asoslariga tayangan. Tahlil natijalari ko'rsatmoqdaki, axborot sızıntilarining katta qismi ichki foydalanuvchilarning faoliyati natijasida sodir bo'ladi. Shu bois, real vaqt rejimida foydalanuvchining harakatini kontekstual jihatdan baholovchi matematik modellar zarur hisoblanadi.

Tadqiqotda asos sifatida quyidagi nazariy va metodik yondashuvlar qo'llanilgan:

Konfidensial fayllarni aniqlash algoritmi [1]da ko'rsatilgan. Bu model TF-IDF (Term Frequency–Inverse Document Frequency) usuliga asoslanib, har bir faylning konfidensiallik ko'rsatkichini hisoblab, belgilangan threshold qiymati asosida tasniflaydi.

1. Fayl harakatlarini aniqlash va loglash Windows Servis yordamida faylga oid quyidagi harakatlar asosida amalga oshirildi — yaratish, o'chirish, ko'chirib o'tkazish, ochish, qayta nomlash, USB qurilmalarga yoki bulutga uzatish — real vaqt rejimida aniqlanadi. Har bir holat logda quyidagi formatda yozildi: fayl ID-si, holat yuz bergan vaqt, foydalanuvchi roli, dastur ishonchliligi va fayl mazmuni.

2. Xavf darajasini baholovchi model harakat holatlari va kontekstual parametrlar asosida fayl uchun umumiy xavf darajasini [2] formulaga muvofiq aniqlanadi.

4. Baholash mezonlari formulalari

Model samaradorligini aniqlash uchun quyidagi statistik ko'rsatkichlar qo'llanildi:

– *Aniqlik (Accuracy):* $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$

(cc) BY

– *Soxta ijobiy holatlar ulushi (FPR):* $FPR = \frac{FP}{FP+TN}$

– *Soxta salbiy holatlar ulushi (FNR):* $FNR = \frac{FN}{FN+TP}$

– *F1 mezon:*

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

$$Precision = \frac{TP}{TP+FP},\ Recall = \frac{TP}{TP+FN}$$

– *Hodisaga o'rtacha reaktsiya vaqti (Response Time):*

$$T_{avg} = \frac{1}{n}\sum_{i=1}^{n}\left(t_{decision,i} - t_{event,i}\right)$$

Ushbu baholash mezonlari tizim samaradorligini matematik va statistik asosda baholash imkonini beradi.

Bu formulalar tizim samaradorligini matematik va statistik asosda baholash imkonini beradi.

5. Real vaqtda tizimning javob choralari xavf darajasi kritik chegaradan oshib ketsa, tizim quyidagi avtomatik choralarni ko'radi: ogohlantirish chiqaradi, harakatni bloklaydi yoki administratorga xabar yuboradi.

6. Eksperimental sinov va baholash natijalariga ko'ra tizim Windows 11 OT o'rnatilib sinovdan o'tkazildi. Jami 1321 ta test fayl (500 tasi konfidensial, 821 tasi oddiy) asosida sinovdan o'tkazildi. Foydalanuvchilarning uchta roli (admin, operator, mehmon) modellashtirildi va natijalar yuqoridagi ko'rsatkichlar asosida baholandi.

Mazkur metodik yondashuv orqali konfidensial axborotlar harakatini nazorat qilish tizimi algoritmik, dasturiy va matematik jihatdan asoslanadi.

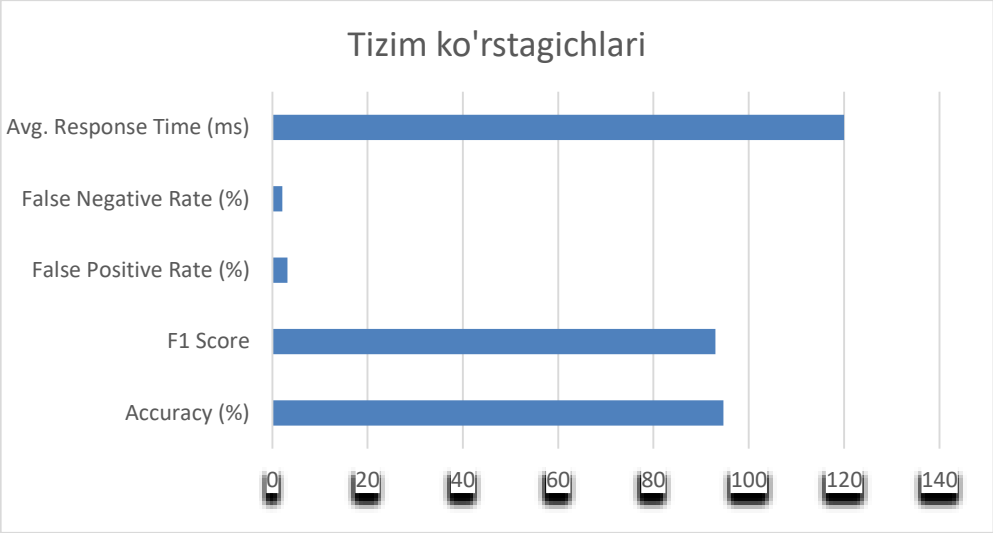
Natijalar va tahlil

Yartilgan tizim test sinovlari davomida tizimning aniqligi, reaktsiya vaqti va noto'g'ri klassifikatsiya holatlari statistik jihatdan tahlil qilindi. Quyidagi asosiy natijalarga erishildi:

- Aniqlik (Accuracy): 94.7%;
- F1 mezoni: 0.93;
- Soxta ijobiy (False Positive) holatlar ulushi: 3.2%;
- Soxta salbiy (False Negative) holatlar ulushi: 2.1%;
- O'rtacha reaktsiya vaqti: 120 millisekund.

Tahlil natijalariga ko'ra, tizim konfidensial fayllarni chiqib ketishini 95% aniqlik bilan farqlay olgan. Ayniqsa, mehmon foydalanuvchilar tomonidan amalga oshirilgan harakatlar yuqori xavfli deb baholangan holatlarning 97% ini muvaffaqiyatli aniqlagan. Real vaqtli ishlash samaradorligi o'rtacha 100–130 ms oralig'ida bo'lib, bu tizimni real amaliyotda qo'llash uchun yetarli ekanligini ko'rsatdi.

Shuningdek, tizim loglarida tahlil qilingan ma'lumotlar asosida foydalanuvchilarning rolga qarab faoliyatini o'rgandi. Bu holat anomal foydalanish holatlarini aniqlashda foydali bo'ladi.



1-rasm. Tizimning asosiy baholash ko'rsatkichlari

Diagramma tizim samaradorligi yuqoriligini, xususan, aniqlik va F1 mezonining yuqori, soxta klassifikatsiya darajasi esa past bo'lganligini yaqqol ko'rsatadi. Ushbu grafik vizual vosita sifatida foydalanuvchilar va tizim adminlari uchun qaror qabul qilishda foydali bo'lishi mumkin.

Ushbu natijalar taklif etilgan model va yondashuvni amaliy tizimlarda qo'llash mumkinligini va u mavjud xavfsizlik tizimlarini

(cc) BY

to‘ldiruvchi muhim modul sifatida xizmat qilishi mumkinligini ko‘rsatadi. va yondashuvni amaliy tizimlarda qo‘llash mumkinligini va u mavjud xavfsizlik tizimlarini to‘ldiruvchi muhim modul sifatida xizmat qilishi mumkinligini ko‘rsatadi.

Tadqiqot doirasida ishlab chiqilgan tizim ochiq kodli va keng tarqalgan vosita — **OpenDLP** bilan taqqoslandi. Quyidagi 1-jadvalda ikki tizimning asosiy funksional ko‘rsatkichlari yoritilgan:

1-jadval.

OpenDLP va taklif etilgan tizimning funksional taqqoslanishi.

Ko‘rsatkichlar	OpenDLP	Taklif etilgan tizim
Monitoring rejimi	Offline (diskni skanerlash)	Real vaqtli
Kontekstual tahlil (vaqtli, rolli, dastur ishonchliligi)	Mavjud emas	Mavjud
Real vaqtli bloklash	Mavjud emas	Mavjud
Fayl kontentini baholash	Mavjud (regex, keywords)	Mavjud (TF-IDF asosida)
Matematik xavf modeli	Mavjud emas	Mavjud
Moslashtirish darajasi	Cheklangan	Moslashuvchan
Ochiq kodli	Ha	Yo‘q (yopiq prototip)

Ushbu taqqoslashdan ko‘rinib turibdiki, OpenDLP tizimi ko‘proq ma’lumotlarni yig‘ish va offline tahlilga qaratilgan bo‘lib, real vaqtli kontekstual nazorat va xavfni matematik modellashtirish imkoniyatlariga ega emas. Taklif etilgan yondashuv esa ushbu kamchiliklarni to‘ldirib, fayllar ustidan to’liq nazorat qilish mexanizmini taqdim etadi.

Xulosa va tavsiyalar

Ushbu tadqiqot doirasida konfidensial fayllarning harakatini real vaqt rejimida kuzatib boruvchi va baholovchi tizim ishlab chiqildi. Taklif etilgan model harakat holati, foydalanuvchi roli, dastur ishonchliligi va vaqt konteksti kabi omillarni inobatga olgan holda xavf darajasini aniqlashga asoslangan. Matematik modellashtirish va eksperimental sinovlar tizim samaradorligini yuqori darajada ekanini ko‘rsatdi.

Ushbu tadqiqot ishi natijalari asosida quyidagi tavsiyalar qilish mumkin:

1. Tashkilotlarda ushbu tizimni joriy etish orqali ichki axborot oqimlarini nazorat qilish va axborot sızıntilarining oldini olish mumkin.
2. Foydalanuvchi faoliyatining xavf profilini tahlil qilish orqali proaktiv choralar ishlab chiqilishi lozim.
3. Mashinali o‘rganish algoritmlari bilan integratsiya qilish orqali xavf baholashning qayta o’qitiladigan tizimini ishlab chiqilishi mumkin.

Foydalanilgan adabiyotlar (References)

1. Juraev G., Bozorov O. Using TF-IDF in text classification //American Institute of Physics Conference Series. – 2023. – T. 2789. – №. 1. – C. 050017.
2. Stallings W. Effective cybersecurity: a guide to using best practices and standards. – Addison-Wesley Professional, 2018.
3. Neumann, P. G., Bishop, M., Peisert, S., & Schaefer, M. (2010, May). Reflections on the 30th Anniversary of the IEEE Symposium on Security and Privacy. In 2010 IEEE Symposium on Security and Privacy (pp. 3-13). IEEE.
4. ISO/IEC 27001:2022. Information security, cybersecurity and privacy protection — Information security management systems — Requirements.
5. Axelsson S. The base-rate fallacy and the difficulty of intrusion detection // ACM Transactions on Information and System Security. 2000. Vol. 3, No. 3. P. 186–205.
6. Jurafsky D., Martin J.H. Speech and Language Processing. 3rd ed. Stanford University, 2022.
7. Han J., Kamber M., Pei J. Data Mining: Concepts and Techniques. Morgan Kaufmann, 2012.
8. Microsoft Docs. FileSystemWatcher Class – .NET [Electronic resource]. URL: https://learn.microsoft.com/en-us/dotnet/api/system.io.filesystemwatcher (accessed: 05.01.2025).
9. Anderson R. Security Engineering: A Guide to Building Dependable Distributed Systems. Wiley, 2020.
10. Graesser, A.C., McNamara, D.S., Louwerse, M.M. et al. Coh-Metrix: Analysis of text on cohesion and language. Behavior Research Methods, Instruments, & Computers 36,

(cc) BY

193–202 (2004).
https://doi.org/10.3758/BF0319556

(cc) BY