WRAP: REAL VAQT REJIMIDA ISHLOVCHI, MASHINAVIY O‘QITISHNING BIR NECHTA ALGORITMLARIGA ASOSLANGAN ANOMAL HOLATLARNI ANIQLASH TIZIMI

Mirzayev O.N.
Yulchiyev I.U.
“Kiberxavfsizlik markazi” DM
Doi: https://doi.org/10.5281/zenodo.20136332

Abstrakt. Korporativ tizimlarda imtiyozli kirish boshqaruvi (Privileged Access Management — PAM) zamonaviy kiberxavfsizlikning markaziy elementlaridan biri hisoblanadi. Mavjud yechimlar (CyberArk, BeyondTrust, Apache Guacamole) imtiyozli sessiyalarni yozib olish va auditga e’tibor qaratadi, biroq sessiya davomida bajariladigan buyruqlarning xavfli xulq-atvorini real vaqtda aniqlash funksiyalari yetarli darajada rivojlanmagan. Insider tahdidlarning o‘rtacha aniqlanish vaqti 85 kunni tashkil qilishi (IBM, 2023) muammoning dolzarbligini ko‘rsatadi.

Ushbu maqolada biz WRAP (Web Remote Access Platform) — to‘liq web-asosli PAM tizimi va to‘rt xil oilaga mansub mashinaviy o‘qitish algoritmlarini birlashtirgan real vaqtdagi anomaliya aniqlash tizimini birlashtirilgan yagona platformani taqdim etamiz. Tizim Isolation Forest, One-Class SVM, Local Outlier Factor va Elliptic Envelope algoritmlarini parallel ravishda qo‘llaydi, ularning natijalarini sigmoid asosida normallashtiradi va uchta agregatsiya strategiyasi orqali birlashtiradi. Bundan tashqari, biz Linux PAM va Bash PROMPT_COMMAND mexanizmlariga asoslangan kam resurs talab qiluvchi agent arxitekturasini taklif etamiz — bu yechim foydalanuvchilarning mavjud SSH klientlari (PuTTY, OpenSSH, Bitvise) bilan moslikni saqlagan holda to‘liq audit va anomaliya tahlilini ta’minlaydi.

Sintetik dataset ustida (n = 500 normal sessiya + 50 anomal sessiya) o‘tkazilgan boshlang’ich tajribalar shuni ko‘rsatdiki, taklif qilingan ko‘p algoritmli yondashuv yagona Isolation Forest algoritmi bilan solishtirganda false positive ko‘rsatkichni 6% dan 2.3% gacha kamaytiradi. Bir sessiyani to‘liq tahlil qilish vaqti o‘rtacha 12 millisekundni tashkil qiladi — bu real-vaqt monitoring uchun yetarlicha tez. Tizim MIT litsenziyali kodda, Docker stackida to‘liq qayta hosil qilinadigan tarzda taqdim etilgan.

Kalit so‘zlar: imtiyozli kirish boshqaruvi, PAM, anomaliya aniqlash, ansambl o‘rganish, Isolation Forest, One-Class SVM, kiberxavfsizlik, SSH monitoring, ML xavfsizligi.

1. KIRISH

Zamonaviy korporativ infratuzilmalarda serverlar va xizmatlarga masofadan kirish — har kungi ish jarayonining ajralmas qismi. SSH (Secure Shell) va RDP (Remote Desktop Protocol) protokollari orqali har kuni minglab ulanishlar amalga oshiriladi. Statistikaga ko‘ra, har bir korporativ administrator kuniga o‘rtacha 50 dan ortiq sessiya yaratadi va ularning har birida 50–200 ta buyruq bajariladi (Verizon DBIR, 2023). Bu hajmdagi faollikni qo‘lda kuzatish va tahlil qilish — amaliy jihatdan imkonsiz.

Imtiyozli kirish boshqaruvi (Privileged Access Management — PAM) tizimlari aynan shu muammoni hal qilish uchun ishlab chiqilgan. Ammo bozorda mavjud yechimlarning aksariyati ikkita asosiy funksiyaga e’tibor qaratadi: kirishni ta’minlash va sessiyalarni yozib olish. Sessiya yopilganidan keyin yozuvni qaytarib ko‘rish mumkin, lekin real vaqtda — sessiya davomida — xavfli xatti-harakatlarni aniqlash imkoniyatlari deyarli mavjud emas.

Bu zamonaviy tahdidlar landshafti uchun jiddiy muammodir. IBM Cost of a Data Breach Report 2023 hisobotiga ko‘ra, insider tahdidlarining o‘rtacha aniqlanish vaqti 85 kunni tashkil qiladi, bu pulli zararni bir qancha million dollargacha yetkazishi mumkin. Ko‘p hollarda hujum sessiya davomida “normal” ko‘rinadi — masalan, autentifikatsiyadan o‘tgan foydalanuvchi “sudo” komandasi bilan tizim fayllarini o‘qiydi va tashqi serverga uzatadi. Hech bir individual qadam darhol shubha tug‘dirmaydi, biroq ularning umumiy ketma-ketligi noodatiy xatti-harakatni ko‘rsatadi.

Adabiyotda mashinaviy o‘qitish (ML) asosidagi anomaliyani aniqlash usullari keng o‘rganilgan. Chandola va boshqalar (2009) o‘z survey maqolasida anomaliya aniqlashning umumiy taksonomiyasini berishadi. Ammo SSH sessiyalari kontekstida — ya’ni buyruqlar oqimi va xulq-atvor xususiyatlari asosida real vaqtda anomaliya aniqlash — yetarli darajada o‘rganilmagan. Bundan tashqari, mavjud ilmiy ishlar va bozordagi kommersiya yechimlari o‘rtasida sezilarli farq mavjud: ilmiy ishlar nazariy modellarni taklif etadi, bozordagi yechimlar esa amaliy lekin yagona algoritm asosida ishlaydi.

1.1. Muammoning bayoni

Hozirgi PAM tizimlari va mashinaviy o‘qitish asosidagi anomaliyani aniqlash usullari o‘rtasida quyidagi bo‘shliqlar mavjud:

• Mavjud PAM yechimlarining hech birida sessiya ichidagi xavfli xulq-atvorni real vaqtda ML asosida aniqlash funksiyasi yo‘q.

• Bozordagi yechimlar foydalanuvchilarni o‘zining proksi yoki jump host tizimi orqali ulanishga majbur qiladi, bu esa mavjud ish jarayonlarining buzilishiga va “shadow IT” holatlarining yuzaga kelishiga olib keladi.

• Ilmiy ML ishlari ko‘pincha bitta algoritm bilan cheklanadi va uyg‘unlashgan yondashuvni sinab ko‘rmaydi.

• “Reproducible research” nuqtai nazaridan, kompleks PAM + ML tizimlarining ochiq kodi va to‘liq “stack”i mavjud emas.

1.2. Maqolaning hissalari

Ushbu maqolada quyidagi asosiy hissalar taqdim etiladi:

• To‘liq web-asosli PAM platformasi (WRAP) — SSH, RDP, SFTP va VNC protokollarini brauzer orqali yagona interfeysda taqdim etadi.

• To‘rt xil algoritmik oilaga (tree, boundary, density, statistical) mansub algoritmlarni birlashtirgan SSH-mahsus ansambl arxitekturasi.

• Linux PAM va Bash PROMPT_COMMAND ga asoslangan tashqi SSH klientlarini kuzatish uchun kam resurs talab qiluvchi agent arxitekturasi.

• Sigmoid funksiyasiga asoslangan “skor” normallashtirish, kelishuv koeffitsienti (agreement ratio) tushunchasi va uchta agregatsiya strategiyasi (average, max, majority).

• To‘liq qayta tiklanadigan (reproduktiv) tadqiqotlar to‘plami: Docker Compose orqali ishga tushiriladigan, MIT litsenziyali ochiq kod, sintetik ma’lumotlar to‘plami va qayta o‘qitish (retraining) API bilan.

1.3. Maqolaning tuzilishi

Maqolaning qolgan qismi quyidagicha tashkil etilgan. 2-bo‘limda mavjud yondashuvlar va adabiyot tahlili keltirilgan. 3-bo‘limda WRAP tizimining umumiy arxitekturasi tavsiflanadi. 4-bo‘lim ko‘p algoritmli ansambl yondashuvining tafsilotlarini taqdim etadi. 5-bo‘lim PAM agent arxitekturasini bayon qiladi. 6-bo‘lim eksperimental natijalarni va tahlilni o‘z ichiga oladi. 7-bo‘lim ish bo‘yicha umumiy muhokama va cheklovlarni keltiradi. 8-bo‘lim xulosa va kelajak ishlari haqida.

2. ADABIYOTLAR TAHLILI

2.1. Imtiyozli kirish boshqaruvi tizimlari

PAM tizimlari korporativ xavfsizlik bo‘yicha NIST SP 800-53 va CIS Controls v8 tomonidan asosiy nazorat sifatida tan olingan. Gartner 2023 Magic Quadrant for PAM hisobotida lider sifatida CyberArk, BeyondTrust va Delinea tilga olingan. Ochiq manba dunyosida Apache Guacamole yagona to‘liq PAM darajasidagi yechim hisoblanadi, biroq u faqat web-proxy funksiyasini bajaradi.

Ilmiy manbaalarda imtiyozli sessiyalarni kuzatish va audit qilish bo‘yicha o‘tkazilgan tadqiqotlar asosan ikki yo‘nalishda jamlangan: (1) keylogging va shell history asosida foydalanuvchi xulq-atvori tahlili (Cui va boshqalar, 2014); (2) auditd va syslog log fayllaridan keyingi tahlilning ML asosida o‘tkazilishi (Khan va boshqalar, 2019). Ushbu ishlarning ikkalasi ham real vaqtda emas, qaytarib tahlil (post-hoc analysis) yondashuviga asoslangan.

2.2. Anomaliya aniqlash algoritmlari

Anomaliya aniqlash algoritmlarining klassik taksonomiyasi Chandola va boshqalar (2009) tomonidan taqdim etilgan. Ular algoritmlarni statistik, masofa asosidagi, zichlik asosidagi, klaster asosidagi va spektral usullariga ajratadi.

Isolation Forest, Liu va boshqalar (2008) tomonidan taklif etilgan, tree-based oilaga mansub bo‘lib, hozirda eng keng qo‘llaniladigan anomaliya aniqlash algoritmlaridan biri hisoblanadi. Uning asosiy g’oyasi — anomaliyalarni tasodifiy daraxtlar orqali tezroq izolyatsiya qilish. Algoritm O(n log n) vaqt murakkabligi bilan ishlaydi va katta o‘lchovli ma’lumotlarga moslashgan.

One-Class SVM, Schölkopf va boshqalar (2001) tomonidan rivojlantirilgan, boundary-based oilaga mansub. RBF kerneli orqali normal ma’lumotlar atrofida nochiziqli giperplane chegara o‘rganadi. Kichik datasetlarda yaxshi natija beradi, ammo gamma va nu parametrlariga sezgir.

Local Outlier Factor (LOF), Breunig va boshqalar (2000) tomonidan kiritilgan, density-based oilaga mansub. Har bir nuqtaning lokal zichligini qo‘shnilari bilan solishtiradi va mahalliy zichlik anomaliyalarini topadi. Bu — globally normal ko‘rinadigan, lekin lokal kontekstda noyob bo‘lgan nuqtalarni aniqlash uchun foydali.

Elliptic Envelope, Rousseeuw va Driessen (1999) tomonidan taklif etilgan Minimum Covariance Determinant (MCD) estimatori asosida ishlaydi. Ma’lumotlarning Gaussian taqsimotda ekanligini taxmin qiladi va Mahalanobis masofa orqali statistik chegaralarni aniqlaydi.

2.3. Ansambl yondashuvlari

Aggarwal va Sathe (2017) o‘zlarining "Outlier Ensembles" monografiyasida ansambl yondashuvining nazariy asoslarini bayon qilgan. Ularning ko‘rsatishicha, har xil algoritmlar har xil tip anomaliyalarni aniqlaydi va ularni birlashtirish umumiy aniqlikni oshiradi. Goldstein va Uchida (2016) PLOS ONE jurnalida e’lon qilgan keng qamrovli taqqoslash tadqiqotida ko‘p o‘lchovli ma’lumotlar uchun turli anomaliya aniqlash algoritmlarini empirik baholaganlar va ensemble yondashuvining ustunligini tasdiqlaganlar.

Biroq SSH sessiyalari kontekstida ansambl yondashuvini qo‘llash bo‘yicha aniq tadqiqotlar deyarli yo‘q. Mavjud ishlarning aksariyati yagona algoritm (ko‘pincha Isolation Forest) ga asoslangan va turli algoritmlarning komplementar tabiati e’tibordan chetda qolgan.

2.4. PAM va shell monitoring

Linux Pluggable Authentication Modules (PAM), Morgan va Kukuk (2010) tomonidan rasmiy hujjatlashtirilgan, autentifikatsiya, sessiya va parol siyosatini standartlashtiruvchi infratuzilma. PAM modullari /etc/pam.d/ konfiguratsiyasi orqali boshqariladi va session hookalari (session_open, session_close) sessiya hayotiy davrining muhim nuqtalarini taqdim etadi. Bu mexanizm an’anaviy ravishda audit va logging maqsadlarida foydalaniladi.

PROMPT_COMMAND mexanizmi Bash Reference Manual da hujjatlashtirilgan va shell foydalanuvchisi har bir buyruqdan keyin bajariladigan funksiyani aniqlashga imkon beradi. Bu mexanizm history-asosli command logging va audit uchun keng qo‘llaniladi. Ammo bu ikki mexanizmni (PAM va PROMPT_COMMAND) yagona monitoring tizimida birlashtirish va markaziy ML platformasiga ulash bo‘yicha tadqiqotlar ochiq adabiyotda yetarlicha aks ettirilmagan.

3. WRAP TIZIMINING UMUMIY ARXITEKTURASI

WRAP modulli, mikroservis asosidagi arxitekturada qurilgan. Tizim quyidagi asosiy komponentlarni o‘z ichiga oladi:

3.1. Komponentlar

Komponent	Texnologiya	Vazifa
Frontend	Angular 17, xterm.js	Foydalanuvchi interfeysi, brauzer ichida terminal/RDP
Backend	NestJS 10, Prisma 5	REST API, WebSocket, autentifikatsiya, RBAC, audit
ML servis	Python 3.11, FastAPI, scikit-learn 1.5	Ko‘p algoritmli anomaliya aniqlash
Database	PostgreSQL 15	Asosiy ma’lumotlar, audit log, anomaly alerts
Cache	Redis 7	Session token, rate limiting
Vault	HashiCorp Vault	Credential shifrlangan saqlash (KV v2)
Object storage	MinIO (S3-compatible)	Sessiya yozuvlari (recording)
RDP proxy	Apache guacd 1.5.5	RDP, VNC protokol konvertatsiyasi
PAM agent	Bash, curl, PAM	Tashqi SSH klientlarini kuzatish

3.2. Ma’lumot oqimi

Foydalanuvchi WRAP tizimiga ikki xil yo‘l bilan ulanishi mumkin:

• Web interfeys orqali (asosiy yo‘l): foydalanuvchi brauzerda Angular SPA’ga kiradi, JWT bilan autentifikatsiyadan o‘tadi, target hostga ulanish so‘rovini yuboradi. Backend Vault’dan credentialni o‘qiydi, target hostga SSH (ssh2 kutubxonasi orqali) ulanish ochadi va bu ulanishni xterm.js’ga WebSocket orqali proxylaydi.

• Tashqi SSH klient orqali (yangi yondashuv): foydalanuvchi PuTTY yoki OpenSSH orqali bevosita target hostga ulanadi. Target hostda o‘rnatilgan WRAP agent PAM va PROMPT_COMMAND mexanizmlari orqali har bir buyruqni WRAP backend ga uzatadi.

Har ikkala yo‘lda sessiya yopilganda backend AnomalyService ML mikroservisiga REST POST so‘rovini yuboradi. ML servis 4 algoritm yordamida natijalarni hisoblaydi va ularni databazaga AnomalyAlert va AlgorithmScore yozuvlari ko‘rinishida saqlaydi.

3.3. Xavfsizlik mexanizmlari

WRAP quyidagi xavfsizlik mexanizmlariga ega:

• Autentifikatsiya: lokal (parol + bcrypt 12 round), LDAP/Active Directory, OIDC, TOTP 2FA.

• Avtorizatsiya: Role-Based Access Control (RBAC) — SuperAdmin, Admin, Operator, Auditor, User rollari; har bir endpoint @RequirePermissions dekoratori bilan himoyalangan.

• Credential storage: HashiCorp Vault KV v2, AES-256 shifrlangan, faqat sessiya boshlanish vaqtida o‘qiladi va plain-text hech qachon databazaga yoki frontendga uzatilmaydi.

• Audit: barcha mutation va sezgir o‘qish amallari audit_logs jadvaliga yoziladi (actor, action, resource, result, ip, userAgent, metadata).

• Recording: SSH sessiyalari asciicast formatida, RDP sessiyalari Apache Guacamole formatida MinIO‘ga shifrlangan holda saqlanadi.

• Network: Nginx reverse proxy, rate limiting (login uchun 10 req/min), CORS, HTTPS (production).

4. KO‘P ALGORITMLI ENSEMBLE YONDASHUVI

4.1. Xususiyatlar fazosi

Har bir SSH sessiyasini 12 o‘lchovli xususiyatlar vektori bilan ifodalaymiz. Xususiyatlar to‘rt toifaga bo‘lingan:

Vaqt	Statistika	Xavf indikatorlari	Boshqa harakatlar
hour_of_day	total_commands	sudo_count	file_read_count
day_of_week	unique_commands	dangerous_cmd_count	compression_count
session_duration_min	commands_per_min	network_cmd_count	script_exec_count

Bu xususiyatlar to‘plami quyidagi tahdid kategoriyalarini qamrab oladi:

• Data exfiltration (compression_count + network_cmd_count): kimdir fayllarni siqib, tashqi serverga uzatishga urinishi.

• Privilege escalation (sudo_count + dangerous_cmd_count): root huquqlariga o‘tib, tizim fayllarini o‘zgartirishga urinish.

• Reconnaissance (file_read_count + network_cmd_count): tizim haqida ma’lumot to‘plash.

• Off-hours access (hour_of_day + day_of_week): ish vaqtidan tashqari kirish.

4.2. Algoritmlar va ularning oilalari

To‘rt xil algoritmik oilaga mansub algoritmlar tanlandi (ensemble diversity ta’minlash uchun):

Algoritm	Oila	Asosiy g’oya	Manbai
Isolation Forest	Tree	Tasodifiy bo‘linish bilan izolyatsiya	Liu et al. (2008)
One-Class	Boundar	RBF kerneli	Schölkopf

SVM	y	orqali chegara o‘rganish	et al. (2001)
Local Outlier Factor	Density	Lokal zichlikni qo‘shnilar bilan solishtirish	Breunig et al. (2000)
Elliptic Envelope	Statistical	MCD asosida Mahalanobis masofa	Rousseeuw & Driessen (1999)

4.3. Skor normallashtirish

Har bir algoritm o‘zining xom skorini turli diapazonda qaytaradi. Ularni umumiy [0, 1] diapazoniga keltirish uchun har bir algoritm uchun moslashtirilgan sigmoid funksiyalar qo‘llaniladi. Umumiy formula:

$\sigma(s) = 1 / (1 + \exp(k \cdot (s - s_0)))$

bu yerda k — sezgirlik koeffitsienti, s_0 — algoritmga xos chegara. Empirik testlar asosida quyidagi qiymatlar tanlandi:

Isolation Forest: $k = 5$, $s_0 = 0$

One-Class SVM: $k = 1$, $s_0 = 0$

LOF: $k = 3$, $s_0 = -1.5$

Elliptic Envelope: $k = 0.5$, $s_0 = 0$

4.4. Agregatsiya strategiyalari

To‘rttala algoritmning natijalarini birlashtirish uchun uchta strategiya taqdim etiladi:

• Average (o‘rtacha): aggregated_score = mean(s_1, s_2, s_3, s_4). Eng barqaror natijani beradi va false positive ni kamaytiradi.

• Maximum: aggregated_score = max(s_1, s_2, s_3, s_4). Konservativ yondashuv — har qanday algoritm anomaliya deb topgan sessiya shubhali deb belgilanadi. Yuqori sezuvchanlik talab qilinadigan hollarda foydali.

• Majority voting: agar 4 ta algoritmdan kamida 3 tasi anomaliya deb topsa, sessiya anomaliya hisoblanadi. Robustlikni oshiradi.

Standart sozlama sifatida average tanlandi. Foydalanuvchi (xavfsizlik analitiki) so‘rov vaqtida agregatsiya usulini API parametri orqali tanlashi mumkin.

4.5. Kelishuv koeffitsienti (Agreement Ratio)

Yangi tushuncha sifatida har bir baho uchun agreement_ratio hisoblanadi:

agreement_ratio = (yakuniy qaror bilan rozi bo‘lgan algoritmlar soni) / (algoritmlar umumiy soni)

Masalan, 4 ta algoritmdan 4 tasi "anomaliya" deb belgilasa, agreement_ratio = 1.0. Agar 3/4 bo‘lsa — 0.75. Bu ko‘rsatkich xavfsizlik analitikiga ikkinchi darajadagi ishonch ko‘rsatkichi sifatida xizmat qiladi: agreement = 1.0 va riskLevel = CRITICAL bo‘lgan alertlar darhol e’tibor talab qiladi, agreement = 0.5 esa qo‘shimcha tekshiruv talab qiladi.

5. PAM-INTEGRATSIYALASHGAN AGENT ARXITEKTURASI

5.1. Motivatsiya

Web-asosli PAM yechimlarining asosiy cheklovi — ular foydalanuvchini maxsus kanal (brauzer yoki klient) orqali ulanishga majburlay. Bu mavjud ish jarayonlarini buzadi va shadow IT xavfini oshiradi. Biz buni hal qilish uchun target server tomonida ishlovchi kam resurs talab qiluvchi agent arxitekturasini taklif etamiz.

5.2. Komponentlar

Agent uchta asosiy skriptdan iborat:

• session-open.sh — PAM open_session hook tomonidan chaqiriladi. POST /api/agent/session/open so‘rovini yuboradi va qaytayotgan sessionId’ni vaqtinchalik faylga yozadi.

• wrap-monitor.sh — /etc/profile.d/ ga joylashtiriladi. PROMPT_COMMAND orqali har buyruqdan keyin send-command.sh ni nosinxron chaqiradi.

• session-close.sh — PAM close_session hook tomonidan chaqiriladi. ML tahlilni ishga tushiradi.

5.3. PAM integratsiyasi

/etc/pam.d/sshd fayliga quyidagi qatorlar qo‘shiladi:

```
session optional pam_exec.so log=/var/log/wrap-pam.log \
    /opt/wrap-agent/session-open.sh
session optional pam_exec.so log=/var/log/wrap-pam.log \
    /opt/wrap-agent/session-close.sh
```

optional kalit so‘zi muhim ahamiyatga ega: agar WRAP platforma vaqtinchalik javob bermasa, foydalanuvchi SSH ga kira olmaslikning oldi olinadi (fail-open approach).

5.4. PROMPT_COMMAND mexanizmi

Bash shell foydalanuvchiga PROMPT_COMMAND o‘zgaruvchisi orqali har bir buyruqdan keyin bajariladigan funksiya belgilash imkonini beradi. Agent buni quyidagicha amalga oshiradi:

```
_wrap_send_command() {
  local exit_code=$?
  local last_cmd=$(history 1 | sed ‘s/^[[:space:]]*[0-9]*[[:space:]]*//’)
  [[ -z "$last_cmd" ]] && return $exit_code
  [[ "$last_cmd" == "$_WRAP_LAST_CMD" ]] && return $exit_code
  _WRAP_LAST_CMD="$last_cmd"
  ${WRAP_AGENT_DIR}/send-command.sh  "$_WRAP_SESSION_ID"  "$last_cmd"
&
  return $exit_code
}
export PROMPT_COMMAND="_wrap_send_command; ${PROMPT_COMMAND}"
```

Nosinxron rejim (& operatori) muhim — foydalanuvchi shell’ning ish samaradorligiga ta’sir qilmaslik uchun.

5.5. Sessiya manbasini farqlash

Backend tomonida har bir sessiya source maydoni bilan belgilanadi (SessionSource enum):

```
enum SessionSource { WRAP, EXTERNAL }
```

Bu xavfsizlik analitikiga sessiyaning qaysi kanal orqali kelganini ko‘rsatadi. WRAP — ichki web terminal orqali, EXTERNAL — tashqi SSH klient (PuTTY, OpenSSH va h.k.) orqali.

6. EKSPERIMENTAL NATIJALAR

6.1. Tajriba muhiti

Barcha tajribalar quyidagi muhitda o‘tkazildi:

- Hardware: Intel Xeon E5-2670 v3 (12 yadro, 2.3 GHz), 32 GB RAM, SSD storage.
- OS: Ubuntu 22.04 LTS, Linux kernel 5.15.
- Container runtime: Docker 24.0, Docker Compose v2.
- ML stack: Python 3.11, scikit-learn 1.5.0, NumPy 1.26, FastAPI 0.110.

6.2. Dataset

Boshlang’ich baholash sintetik dataset ustida o‘tkazildi. Sintetik dataset asosiy real-dunyo xulq-atvor mintaqalarini taqlid qiladi:

• Normal sessiyalar (n = 500): ish vaqti (8:00–18:00), normal hafta kunlari, 5–80 buyruq, kichik miqdorda sudo (0–1), xavfli yoki tarmoq buyruqlari yo‘q yoki minimal.

• Anomal sessiyalar (n = 50): tunda yoki dam olish kunlarida, ko‘p sudo (3–10), bir nechta dangerous_cmd, network_cmd, compression_cmd kombinatsiyasi.

Real korporativ dataset to‘plash va ground-truth labeling jarayoni hozirda davom etmoqda va keyingi maqolalarda taqdim etiladi.

6.3. Yagona algoritm vs. ansambl

Boshlang’ich tajribada to‘rttala algoritmning yagona ishlashi va ansamblning ishlashi solishtirildi. Kontaminatsiya darajasi 0.05 (5%) ga o‘rnatildi.

Yondashuv	False Positive (%)	True Positive (%)	O‘rta cha vaqt (ms)
Isolation Forest (yagona)	6.0	88.0	2
One-Class SVM (yagona)	8.4	82.0	3
LOF (yagona)	7.2	84.0	4
Elliptic Envelope (yagona)	9.6	78.0	6
Ensemble (average)	2.3	90.0	12
Ensemble (max)	11.2	94.0	12
Ensemble (majority)	3.1	87.0	12

Natijalar ko‘rsatishicha, ensemble (average agregatsiya) yagona Isolation Forest ga nisbatan false positive ni 6.0% dan 2.3% gacha kamaytiradi, true positive esa 88% dan 90% gacha oshadi. Hisoblash vaqti 4 algoritm parallel ishlasa ham 12 ms dan oshmaydi — bu real-vaqt monitoring uchun yetarli.

Ensemble (max) konservativ yondashuv sifatida true positive ni 94% gacha oshiradi, ammo false positive ni 11.2% ga ko‘taradi. Bu yuqori sezuvchanlik talab qilinadigan vaziyatlarda (masalan, mudofaa yoki moliyaviy sektorda) maqbul kompromis.

6.4. Algoritmlar o‘rtasidagi korrelyatsiya

Ensemble diversity’ni tasdiqlash uchun algoritmlar natijalari o‘rtasidagi Pearson korrelyatsiya koeffitsientlari hisoblandi:

	IF	OC-SVM	LOF	EE
IF	1.00	0.58	0.53	0.42
OC-SVM	0.58	1.00	0.61	0.49

LOF	0.53	0.61	1.00	0.46
EE	0.42	0.49	0.46	1.00

Korrelyatsiyalar 0.42 dan 0.61 gacha — bu o‘rtacha darajadagi mustaqillik degani. IF va EE eng past korrelyatsiya ko‘rsatkichiga ega (0.42), bu ulardan biri ushlay olmagan anomaliyalarni boshqasi ushlay olishini tasdiqlaydi. Agar barcha algoritmlar bir-biriga 0.95+ korrelyatsiyaga ega bo‘lsa, ensemble foyda keltirmasdi.

6.5. Agent performansi

PAM agent o‘lchovlarini olish uchun real foydalanuvchi simulyatsiyasi o‘tkazildi. 100 ta sessiya davomida har birida 50 ta buyruq bajarildi. Natijalar:

- Buyruq kechikishi (latency overhead): o‘rtacha 0.3 ms (background mode tufayli).
- Foydalanuvchi sezadigan ta’sir: yo‘q (sub-millisecond).
- Agent CPU iste’moli: 0.1% (oddiy yuk).
- Agent xotira iste’moli: 4 MB.
- Tarmoq trafigi: ~200 baytlik so‘rov har buyruq uchun (HTTPS-da JSON).

Agent foydalanuvchi shell’ning ish samaradorligiga sezilarli ta’sir ko‘rsatmaydi. Bu PROMPT_COMMAND ning nosinxron bajarilishi va curl ning 1 sekundlik timeout bilan ishlatilishidan kelib chiqadi.

7. MUHOKAMA VA CHEKLOVLAR

7.1. Tizimning afzalliklari

WRAP quyidagi asosiy afzalliklarga ega:

- Integratsiya: web-asosli PAM va ML anomaliya aniqlash yagona platformada birlashtirilgan.
- Universal monitoring: web terminal va tashqi SSH klientlarning ikkalasi ham bir xil ML tahlil olib boradi.
- Ensemble robustligi: 4 algoritm parallel — yagona algoritmning kuchsiz tomonlari kompensatsiya qilinadi.
- Reproducibility: to‘liq Docker stack, sintetik dataset, retraining API — har qanday tadqiqotchi natijalarni tekshirib ko‘rishi mumkin.
- Real-vaqt: 12 ms tahlil vaqti monitoring tizimi uchun yetarlicha tez.

7.2. Cheklovlar

Ushbu ish quyidagi cheklovlarga ega:

- Sintetik dataset: hozirgi natijalar sintetik ma’lumotlarga asoslangan. Real-dunyo korporativ datasetida natijalar farq qilishi mumkin.
- Cheklangan tahdid modellari: hozirgi xususiyatlar to‘plami buyruq-asosli statistik tahlilga e’tibor qaratadi. Yangi tahdid turlari (masalan, fileless attacks, supply chain) qo‘shimcha xususiyatlarni talab qilishi mumkin.
- Sequence semantics yo‘q: hozirgi yondashuv buyruqlar sonini hisobga oladi, lekin ularning ketma-ketligini emas. LSTM autoencoder yoki Transformer asosida ketma-ketlik modelining qo‘shilishi kelajakdagi yaxshilanish.
- Faqat bash: agent hozirgi versiyada faqat bash shellni qo‘llab-quvvatlaydi. Zsh, fish va PowerShell uchun moslik kelajakdagi ish.

7.3. Etika va xavfsizlik mulohazalari

Imtiyozli foydalanuvchilarni kuzatish va ularning xulq-atvorini tahlil qilish bir qancha etika muammolarini keltirib chiqaradi. WRAP quyidagi tamoyillarga rioya qiladi:

- Shaffoflik: foydalanuvchilar tizim ularni kuzatayotganidan xabardor bo‘lishi shart (kompaniya siyosatida e’lon qilinadi).

• Maqsadli foydalanish: ma’lumotlar faqat xavfsizlik tahlili uchun ishlatiladi, foydalanuvchini reyting yoki KPI uchun emas.

• Saqlash muddati: yozuvlar va audit loglari belgilangan retention policy (default 90 kun) ga muvofiq saqlanadi va keyin avtomatik o‘chiriladi.

• Kirish nazorati: anomaliya alertlarini faqat tegishli xavfsizlik xodimlari ko‘ra oladi (RBAC orqali).

8. XULOSA VA KELAJAK ISHLARI

Ushbu maqolada WRAP — to‘liq web-asosli imtiyozli kirish boshqaruvi platformasi va ko‘p algoritmli ML asosida anomaliya aniqlash tizimini birlashtirilgan yagona yechim taqdim etildi. Asosiy hissalar:

• To‘rt xil oilaga (tree, boundary, density, statistical) mansub algoritmlarni birlashtirgan SSH-mahsus ensemble arxitekturasi.

• PAM va PROMPT_COMMAND mexanizmlariga asoslangan tashqi SSH klientlarini kuzatish uchun kam resurs talab qiluvchi agent yondashuvi.

• Sigmoid asosidagi skor normallashtirish, kelishuv koeffitsienti tushunchasi va uchta agregatsiya strategiyasi.

• To‘liq reproducible research stack: ochiq kod, Docker stack, sintetik dataset.

Boshlang’ich tajribalar ko‘rsatishicha, taklif qilingan yondashuv yagona Isolation Forest ga nisbatan false positive ni 6.0% dan 2.3% gacha kamaytiradi va true positive ni 88% dan 90% gacha oshiradi, real-vaqt tahlil uchun yetarli tezlikni saqlab qoladi.

8.1. Kelajak ishlari

Kelajakdagi tadqiqotlar quyidagi yo‘nalishlarda davom etadi:

• 1000+ sessiyadan iborat real korporativ ma’lumotlar to‘plamini yig‘ish va ekspert tomonidan asosiy haqiqat (ground-truth) bo‘yicha belgilash.

• Precision, Recall, F1, AUC kabi standart metrikalar bilan kengaytirilgan baholash.

• LSTM autoencoder va Transformer asosida ketma-ketlikni hisobga oluvchi modellarni ansamblga qo‘shish.

• Moslashuvchan og‘irliklash: algoritmlarni vaznli tarzda birlashtirish — har bir algoritmning ishonchliligini dinamik ravishda baholash asosida agregatsiya qilish.

• Zsh, Fish va PowerShell shelllari bilan agentning mosligini ta’minlash.

• Real-vaqt alerting: kritik sessiya davomida (sessiya tugashini kutmasdan) anomaliyani aniqlash.

• RDP sessiya videosini OCR yoki hodisalarni kuzatish (event tracking) orqali tahlil qilish.

ADABIYOTLAR

1. Aggarwal, C. C., & Sathe, S. (2017). Outlier Ensembles: An Introduction. Springer International Publishing.
2. Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. (2000). LOF: Identifying density-based local outliers. In Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data (pp. 93–104).
3. Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys, 41(3), Article 15.
4. Cui, S., Asghar, M. R., Galbraith, S. D., & Russello, G. (2014). Privacy in big data: A review. IEEE Communications Surveys & Tutorials.
5. Goldstein, M., & Uchida, S. (2016). A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. PLOS ONE, 11(4), e0152173.
6. IBM Security. (2023). Cost of a Data Breach Report 2023. IBM Corporation.

7. Khan, M. A., Karim, M. R., & Kim, Y. (2019). A scalable and hybrid intrusion detection system based on the convolutional-LSTM network. Symmetry, 11(4), 583.
8. Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation Forest. In 2008 Eighth IEEE International Conference on Data Mining (pp. 413–422). IEEE.
9. Morgan, A. G., & Kukuk, T. (2010). The Linux-PAM System Administrators' Guide. Linux-PAM Documentation Project.
10. NIST Special Publication 800-53 Rev. 5. (2020). Security and Privacy Controls for Information Systems and Organizations. National Institute of Standards and Technology.
11. Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825–2830.
12. Ramey, C. (2024). Bash Reference Manual, Version 5.2. GNU Project.
13. Rousseeuw, P. J., & Driessen, K. V. (1999). A fast algorithm for the minimum covariance determinant estimator. Technometrics, 41(3), 212–223.
14. Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., & Williamson, R. C. (2001). Estimating the support of a high-dimensional distribution. Neural Computation, 13(7), 1443–1471.
15. Verizon. (2023). 2023 Data Breach Investigations Report. Verizon Business.
16. Ylonen, T., & Lonvick, C. (Eds.). (2006). RFC 4251: The Secure Shell (SSH) Protocol Architecture. IETF.