

Pannon Egyetem
Műszaki Informatikai Kar
Informatikai Tudományok Doktori Iskola

Zsarolóvírus-támadások időkritikus
felismerése gépi tanulás alkalmazási
lehetőségeinek vizsgálatával centralizált
fájltárolási környezetekben

Doktori (PhD) értekezés

DOI:10.18136/PE.2026.1011

Arányi Gábor

Témavezető: Dr. Szücs Veronika

2026

ZSAROLÓVÍRUS-TÁMADÁSOK IDŐKRITIKUS FELISMERÉSE GÉPI
TANULÁS ALKALMAZÁSI LEHETŐSÉGEINEK
VIZSGÁLATÁVAL CENTRALIZÁLT FÁJLTÁROLÁSI
KÖRNYEZETEK BEN

Az értekezés doktori (PhD) fokozat elnyerése érdekében készült a Pannon Egyetem Infor-
matikai Tudományok Doktori Iskolája keretében

informatikai tudományok tudományágban

Írta: Arányi Gábor

Témavezető: Dr. Szücs Veronika

Elfogadásra javaslom (igen / nem)

.....
témavezető

Az értekezés bírálatra bocsátható.

.....
TDHT elnök

A jelölt az értekezés nyilvános vitáján %-ot ért el.

A bíráló Bizottság tagjai:

elnök:.....

bírálok:.....

tagok:.....

Veszprém,

.....
Bíráló Bizottság elnök

A doktori (PhD) oklevél minősítése:.....

Veszprém,

.....
EDHT elnök

Tartalmi kivonat

A zsarolóvírus-támadások az elmúlt években az informatikai rendszerek egyik legsúlyosabb és leggyorsabban növekvő fenyegetésévé váltak, különösen az egészségügyi és oktatási szektorban, valamint a kis- és középvállalkozói (KKV-k) szegmensben, ahol a védekezés lehetőségei mind financiálisan, mind pedig szakmailag gyakran korlátozottak. A kutatás alapvető célja egy olyan gyakorlatban is alkalmazható védelmi modell kidolgozása volt, amely képes az egyéb védelmi megoldásoktól függetlenül ellátni a centralizáltan tárolt adatvagyon védelmét, ugyanakkor jelezni tud egy esetleges zsarolóvírus-támadást kvázi valós időben.

A kutatás első szakaszában a hazai KKV-k körében kérdőíves interjúkon keresztül tártam fel az informatikai biztonsági kockázatokat, az eddigi incidensek tapasztalatait, a tipikus támadási vektorokat, valamint a döntéshozó személye és az alkalmazott megoldások közötti összefüggéseket. Ezen eredményekre támaszkodva került kialakításra az ARDS (Anti-Ransomware Defense System) modell, amely megfigyelési, naplózási, riasztási és helyreállítási mechanizmusokat integrál. A kutatás második szakaszában az ARDS modell egyedi fájlkiszolgáló megoldásának teljesítményvizsgálata került fókuszba. A FreeBSD alapú kiszolgáló AES-256 GCM adattitkosítást alkalmazva tárolja az adatokat és időzítetten pillanatképeket készít az adatvagyonról. A mérések során különböző hálózati sávszélességek mellett (1 és 10 Gbit/s) elemeztém az olvasási és írási teljesítményt. A tesztek eredményeit összehasonlítottam a referencia-ként választott, alapértelmezett beállításokkal telepített MS Windows Server 2019 fájlkiszolgáló eredményeivel, ahol a BitLocker titkosítási szolgáltatás és az árnyékmásolatok (VSS) nincsenek engedélyezve. A két környezet teljesítményét a különböző tesztesetekre vonatkozóan egy újonnan bevezetett, kalkulált teljesítménymetrika segítségével tudtam egzakt módon összehasonlítani. Az eredmények igazolták, hogy a javasolt megoldás a legtöbb esetben jobb teljesítményt nyújtott a referenciakörnyezetnél. Az utolsó kutatási szakaszban a cél a zsarolóvírus-támadások hatékony és gyors detektálása volt szerveroldali fájlműveleti naplófájlok alapján, felügyelt gépi tanulási módszerek alkalmazásával. Egy tipikus KKV virtualizált informatikai infrastruktúrájának felépítését követően legitim, felhasználói, továbbá öt különböző zsarolóvírus által generált, másodpercre aggregáltan több, mint 2,5 millió fájlműveleti naplóbejegyzést tartalmazó adathalmazon klasszikus és ensemble (Random Forest, Support Vector Machine, Decision Tree, AdaBoost, XGBoost) gépi tanulási modellek teljesítmény-összehasonlítását végeztem el. A hiperparaméter-hangolást követően az XGBoost modell érte el a legjobb eredményt az észlelési pontosság, a késleltetés és a költség alapú értékelés tekintetében, miközben a teszt adatsorhoz (392 másodperc) tartozó, támadással érintett rekordok meghatározásához szükséges idő minden modell esetében a milliszekundumos tartományban maradt.

A kutatás eredményei egyértelműen megerősítették, hogy a zsarolóvírus-támadások jól azonosítható viselkedési mintákat hagynak a fájlkiszolgálók naplóiban és hogy a szerveroldali, napló alapú gépi tanulás hatékony, skálázható és gyakorlatban is integrálható megoldást kínál a kvázi valós idejű észlelésre, továbbá lehetővé teszi esetleges incidensválaszok kiváltását.

Kulcsszavak: ARDS, zsarolóvírus-észlelés, fájlkiszolgáló, fájlművelet-szekvencia, gépi tanulás, valós idő

Abstract

In recent years, ransomware attacks have become one of the most serious and fastest-growing threats to IT systems, particularly in healthcare and education, and in the small and medium-sized enterprise (SME) segment, where the means of defense are often limited, both financially and professionally. The fundamental goal of the research was to develop a practical protection model that can safeguard centrally stored data assets independently of other defense solutions, while it is also able to signal a potential ransomware attack in quasi real-time.

In the first phase of the research, I used questionnaire-based interviews conducted with Hungarian SMEs to identify IT security risks, experiences of incidents to date, typical attack vectors, and the relationships between decision-makers and the solutions used. Based on these results, the ARDS (Anti-Ransomware Defense System) model was developed, which integrates monitoring, logging, alerting, and recovery mechanisms. The second phase of the research focused on performance testing of the ARDS model's unique file server solution. The FreeBSD-based server stores data using AES-256 GCM encryption and takes scheduled snapshots of the data assets. During the measurements, I analyzed read and write performance at different network bandwidths (1 and 10 Gbit/s). I compared the test results with those of a reference MS Windows Server 2019 file server installed with default settings, where BitLocker encryption and shadow copies (VSS) were disabled. I was able to accurately compare the performance of the two environments for different test cases using a newly introduced, calculated performance metric. The results confirmed that the proposed solution delivered better performance than the reference environment in most cases. The goal of the final research phase was to detect ransomware attacks efficiently and quickly based on server-side file operation log files, using supervised machine learning methods. After setting up a typical SME virtualized IT infrastructure, I performed a performance comparison of classic and ensemble (Random Forest, Support Vector Machine, Decision Tree, AdaBoost, XGBoost) machine learning models on a dataset containing more than 2.5 million file operation log entries generated by legitimate users and five different ransomware viruses. After hyperparameter tuning, the XGBoost model achieved the best results in terms of detection accuracy, latency, and cost-based evaluation, while the time required to identify attack-affected records in the test dataset (392 seconds) remained in the millisecond range for all models.

The results of the research clearly confirmed that ransomware attacks leave easily identifiable behavioral patterns in file server logs and that server-side, log-based machine learning offers an effective, scalable, and practical solution for quasi real-time detection and also enables triggering of incident responses.

Keywords: ARDS, ransomware detection, file server, file operation sequence, machine learning, real-time

Abstract

Ransomwares sind in den letzten Jahren zu einer der schwerwiegendsten und am schnellsten wachsenden Bedrohungen für IT-Systeme geworden, insbesondere im Gesundheits-, und Bildungswesen und im Segment der kleinen und mittleren Unternehmen (KMU), wo die Abwehrmöglichkeiten sowohl finanziell als auch fachlich oft begrenzt sind. Das grundlegende Ziel der Forschung war die Entwicklung eines in der Praxis anwendbaren Schutzmodells, das in der Lage ist, zentral gespeicherte Daten unabhängig von anderen Schutzlösungen zu schützen und gleichzeitig einen möglichen Ransomware-Angriff quasi in Echtzeit zu melden.

In der ersten Phase der Forschung habe ich anhand von Fragebogeninterviews unter ungarischen KMU die IT-Sicherheitsrisiken, die bisherigen Erfahrungen mit Sicherheitsvorfällen, die typischen Angriffsvektoren sowie die Zusammenhänge zwischen den Entscheidungsträgern und den eingesetzten Lösungen untersucht. Auf der Grundlage dieser Ergebnisse wurde das ARDS-Modell (Anti-Ransomware Defense System) entwickelt, das gleichzeitig Überwachungs-, Protokollierungs-, Alarmierungs- und Wiederherstellungsmechanismen integriert. In der zweiten Phase der Forschung lag der Schwerpunkt auf der Leistungsmessung der einzigartigen Dateiserverlösung des ARDS-Modells. Der FreeBSD-basierte Server speichert die Daten mit AES-256-GCM-Datenverschlüsselung und erstellt regelmäßige Momentaufnahmen der Datenbestände. Während der Messungen habe ich die Lese- und Schreibleistung bei verschiedenen Netzwerkbandbreiten (1 und 10 Gbit/s) analysiert. Ich habe die Testergebnisse mit denen eines MS Windows Server 2019-Dateiservers verglichen, der mit den als Referenz ausgewählten Standardeinstellungen installiert wurde und auf dem die BitLocker-Verschlüsselungsfunktion und Schattenkopien (VSS) deaktiviert sind. Die Leistung der beiden Umgebungen konnte ich mithilfe einer neu eingeführten, kalkulierten Leistungsmetrik für die verschiedenen Testfälle exakt vergleichen. Die Ergebnisse bestätigten, dass die vorgeschlagene Lösung in den meisten Fällen eine bessere Leistung als die Referenzumgebung erzielte. In der letzten Forschungsphase war das Ziel die effektive und schnelle Erkennung von Ransomware-Angriffen aufgrund von Dateioperationsprotokollen auf der Serverseite unter Verwendung überwachter maschineller Lernmethoden. Nach dem Aufbau einer typischen KMU virtualisierten IT-Infrastruktur habe ich einen Datenbestand mit mehr als 2,5 Millionen Dateioperationsprotokolleinträgen erstellt (der sowohl von legitimen Benutzern als auch von fünf verschiedenen Ransomwares generierte Einträge enthält) und einen Leistungsvergleich zwischen klassischen und Ensemble-Modellen (Random Forest, Support Vector Machine, Decision Tree, AdaBoost, XGBoost) durchgeführt. Nach der Hyperparameter-Optimierung erzielte das XGBoost-Modell die besten Ergebnisse in Bezug auf Erkennungsgenauigkeit, Latenz und kostenbasierte Bewertung, während die zur Identifizierung der von Angriffen betroffenen Datensätze im Testdatensatz (392 Sekunden) benötigte Zeit bei allen Modellen im Millisekundenbereich blieb.

Die Forschungsergebnisse haben eindeutig bestätigt, dass Ransomware-Angriffe klar erkennbare Verhaltensmuster in den Protokollen der Dateiserver hinterlassen und serverseitiges, protokollbasiertes maschinelles Lernen eine effektive, skalierbare und in der Praxis integrierbare Lösung für eine quasi-Echtzeit-Erkennung bietet und darüber hinaus die Auslösung von möglichen Incident Response ermöglicht.

Schlagwörter: ARDS, Ransomware-Erkennung, Dateiserver, Dateioperationssequenz, maschinelles Lernen, Echtzeit

Köszönetnyilvánítás

Ezúton szeretnék köszönetet mondani témavezetőmnek, Dr. Szücs Veronikának amiért folyamatosan támogatott, szakmai tanácsokkal látott el és mindig a legmegfelelőbb módon motivált a képzés folyamán. Köszönöm Dr. Hartung Ferenc professzor úrnak a támogatást és a hasznos stratégiai tanácsokat, amik révén témámat a doktori iskola befogadta és a dolgozatom elkészülhetett, valamint Dr. Hangos Katalin professzor asszonynak, hogy mindig nyitottan állt a témámhoz és kérdéseivel mindig előre mozdított a beszámolók alkalmával.

Köszönettel tartozom Dr. Fogarassyné dr. Vathy Ágnes tanszékvezető asszonynak, amiért megismertette velem a mesterséges intelligencia világát és iránymutatásaival folyamatosan segítette a gépi tanulás módszereinek alkalmazását a kutatásaim során, továbbá Miseta Tamás hallgatótársamnak, amiért gyakorlati példákkal látott el. Szintén köszönet illeti Dr. Dávid Ákos adjunktus urat, amiért már a képzés kezdete előtt fantáziát látott a témámban és mindvégig mentorként támogatott, hogy elérhessem a céljaimat. Hálás vagyok Dr. Kő Andrea professzor asszonynak és Dr. Krasznay Csaba docens úrnak, akik opponensként tett javaslatokkal és iránymutatásaikkal hozzájárultak az értekezés színvonalának emeléséhez. Köszönöm a lehetőséget Dr. Kovács László és Dr. Haig Zsolt professzor uraknak, hogy lehetővé tették a témámhoz kapcsolódóan külső tárgyak teljesítését a Nemzeti Közszolgálati Egyetemen.

Köszönetemet szeretném kifejezni barátaimnak és kollégáimnak Markó Ferencnek, Tóth Szilveszternek és Funtek Imrének a pótolhatatlan szakmai segítségéért, amiket nyújtottak az évek során. Hálával tartozom a doktori iskola titkárainak, Dr. Dulai Tibornak és Dr. Görbe Péternek a tanulmányaim során nyújtott támogatásért, valamint Gazdag-Kéri Renátának és Schmidtné Lényi Szilviának az adminisztrációban nyújtott támogatásukért és hasznos tanácsaikért. Szeretném megköszönni feleségemnek a támogatást és a biztatást, amivel folyamatosan segítette a munkámat, fiaimnak pedig, hogy elfogadták az elhivatottságomat.

Édesanyám emlékének ajánlom ezt az értekezést, hálával mindazért a szeretetért és támogatásért, amelyet életem során kaptam tőle.

Tartalomjegyzék

Ábrajegyzék	7
Táblázatjegyzék	9
Rövidítésjegyzék	10
1. Bevezetés	11
1.1. A zsarolóvírus-támadások gazdasági vonatkozásai.....	13
1.2. Támadási vektorok, kiemelten veszélyeztetett területek	15
1.3. Motiváció.....	16
1.4. A kutatás célja, várható eredménye.....	16
1.5. Az értekezés tematikája.....	17
1.6. Publikációs teljesítmény és hivatkozások	18
2. Kutatásmódszertan.....	19
2.1. Szakirodalmfeldolgozás módszerének ismertetése	19
2.2. A kérdőíves felmérés módszertani bemutatása	23
2.3. A gépi tanulási modellek kiválasztásának módszertana.....	26
3. Szakirodalmi áttekintés.....	27
4. Hazai kis- és középvállalkozások (KKV-k) informatikai biztonságra vonatkozó felkészültségének felmérése kérdőíves kutatás alapján, az ARDS védelmi modell bemutatása	28
4.1. A kutatás menete, a vizsgálati módszer.....	31
4.1.1. A kutatás korlátai	33
4.2. A kérdőív felépítése.....	34
4.3. A tanulmány célcsoportja	35
4.4. Az eredmények összefoglalása, releváns kutatási kérdések megválaszolása.....	35
4.5. Az ARDS (Anti-Ransomware Defense System Model) védelmi modell	44
4.5.1. Támadási vektorok, jellemző belépési pontok a hálózatokon.....	44
4.5.2. A védelmi modell ismertetése.....	46
4.5.3. A védelmi modell és a hagyományos védelmi koncepciók összehasonlítása.....	49
5. Az ARDS modell fájl kiszolgálójának teljesítményértékelése	51

5.1. Adatvagyonvédelmi és -visszaállítási megoldások áttekintése	52
5.2. A zsarolóvírusok elleni védekezés lehetőségeinek rövid technológiai áttekintése	55
5.3. A kidolgozott fájlkiszolgáló védelmi megoldás jellemzői	56
5.4. A vizsgálati módszer	56
5.4.1. Mérési és vizsgálati módszer	56
5.4.2. A mérési bizonytalanság csökkentése és a mérési hibák kiszűrése	58
5.4.3. A reprodukálhatóság biztosítása	59
5.4.4. Az eredmények értékelésének módszertana, az összehasonlíthatóság biztosítása ..	60
5.5. A teljesítménymérés folyamatának és körülményeinek ismertetése	61
5.5.1. A tesztkörnyezetek kiválasztása, a tesztesetek meghatározása.....	61
5.5.2. Az adatátviteli közeg áteresztőképességének meghatározása.....	62
5.5.3. Méréshez használt adatok összetételének meghatározása	63
5.6. A teljesítménymérés eredményeinek ismertetése.....	64
5.7. Eredmények 1 Gbit/s-os kapcsolat esetén	65
5.7.1. Az olvasási teljesítményvizsgálat eredményei.....	65
5.7.2. Az írási teljesítménytesztek eredményei.....	66
5.8. Eredmények 10 Gbit/s-os kapcsolat esetén	67
5.8.1. Az olvasási teljesítménytesztek eredményei.....	67
5.8.2. Az írási teljesítménytesztek eredményei.....	69
5.9. Referencia alapú teljesítményértékelés	70
5.10. A releváns kutatási kérdések megválaszolása	72
6. Zsarolóvírus-tevékenység detektálása gépi tanulás segítségével szerveroldali, fájlműveleti naplófájlok alapján	73
6.1. A zsarolóvírusok elleni védelmi stratégiák áttekintése	74
6.1.1. Néhány kísérleti megoldás összehasonlítása.....	75
6.1.2. A mesterséges intelligencia alapú módszerek áttekintése.....	76
6.1.3. Egyéb releváns megközelítések	77
6.2. A fájlművelet-szekvenciákra alapuló gépi tanulás módszerének bemutatása	78
6.2.1. A szimulációs környezet felépítése és az adatsor előállítása	78
6.2.2. Adattisztítás és előfeldolgozás	80
6.2.3. Gépi tanulási modellek kiválasztása, a hiperparaméter hangolás célja	82
6.2.4. Értékelési metrikák, az összehasonlítás menete.....	83
6.2.5. ROC és pontosság-visszahívás pontok elemzése.....	83
6.2.6. Statisztikai validálás.....	84

6.2.7. Költség alapú összehasonlítás	84
6.2.8. A modellek tanítása, hiperparaméter hangolás és kiértékelés.....	84
6.3. Az eredmények kiértékelése.....	86
6.3.1. Hiperparaméter hangolás	86
6.3.2. A gépi tanulási modellek teljesítményének összehasonlítása	88
6.3.3. ROC és Precision-Recall pontelemzés.....	90
6.3.4. A költség alapú értékelés eredménye.....	92
6.4. Az eredmények összefoglalása.....	93
6.5. A megközelítés korlátai és továbbfejlesztési lehetőségei.....	96
6.6. A releváns kutatási kérdések megválaszolása	97
7. A doktori értekezés tézisei.....	98
8. Összegzés és kitekintés.....	103
8.1. Az eredmények értelmezése és tudományos jelentősége	103
8.2. Korlátok és értelmezési megfontolások.....	104
8.3. Az eredmények nemzetközi kutatásokkal való összevetése.....	104
8.3.1. A tudományos eredmények összevetése	105
8.3.2. A dolgozat tudományos pozicionálása.....	106
8.4. Jövőbeni kutatási irányok és feladatok.....	106

Ábrajegyzék

1. ábra Teljes populációra vonatkozó digitalizációs szám adatok [1]	11
2. ábra A Bitcoin árfolyam változása 5 éves intervallumra vonatkozóan - forrás: Google Finance [7].....	14
3. ábra Publikációs tevékenység összefoglaló táblázata.....	18
4. ábra Idézetség összefoglaló táblázata	18
5. ábra Publikációs tevékenység eloszlása a doktori képzés kezdetétől az értekezés benyújtásáig	19
6. ábra PRISMA flow diagram szakirodalom feldolgozáshoz [forrás: saját].....	20
7. ábra A kutatási folyamat vizualizációja [forrás: saját]	32
8. ábra Érzékeny, sérülékeny adatot kezelők aránya a cég méretéhez viszonyítva [forrás: saját]	36
9. ábra Informatikai infrastruktúrák jellemzői a vizsgált vállalatoknál [forrás: saját]	37
10. ábra A vizsgált vállalatok sebezhetőségi jellemzői [forrás: saját].....	37
12. ábra A döntéshozó informatikai besorolása és az informatikai rendszer védelme közötti kapcsolat [forrás: saját].....	38
11. ábra A döntéshozó informatikai besorolása és a fizikai védelem közötti kapcsolat [forrás: saját]	38
13. ábra A vizsgált vállalatok sebezhetőségi jellemzői [forrás: saját].....	40
14. ábra A jelenleg használt antivírus megoldások eloszlása a döntéshozók informatikai képzettségének függvényében [forrás: saját].....	40
15. ábra Vírus- és zsarolóvírus-támadások száma az interjú megelőző évben [forrás: saját] ..	41
16. ábra Kapcsolat a döntéshozó informatikai besorolása - az informatikai rendszerek védelmi szintje és az elszenvedett vírustámadások között [forrás: saját].....	42
17. ábra A home office-ra vonatkozó vállalati irányelvek, a vállalati e-mail használat és a zsarolóvírus-támadások közötti összefüggés [forrás: saját]	43
18. ábra A mikro- és kisvállalkozások által használt tipikus hálózati topológia [forrás: saját]	45
19. ábra Az ARDS-modell sematikus topológiája [forrás: saját]	47
20. ábra A kidolgozott védelmi modell funkcionális architektúrája [forrás: saját]	48
21. ábra Iperf3 mérési eredmények [forrás: saját]	63
22. ábra Olvasási teszteredmények összehasonlítása 1 Gbit/s sávszélesség esetén [forrás: saját]	65

23. ábra Írási teszteredmények összehasonlítása 1 Gbit/s sávszélesség esetén [forrás: saját] .	67
24. ábra Olvasási teszteredmények összehasonlítása 10 Gbit/s sávszélesség esetén [forrás: saját]	68
25. ábra Írási teszteredmények összehasonlítása 10 Gbit/s sávszélesség esetén [forrás: saját]	69
26. ábra A vizsgált környezet modellje és annak munkafolyamatai [forrás: saját]	78
27. ábra Támadási állapotok az idő függvényében [forrás: saját]	86
28. ábra A modellek késleltetésének összehasonlítása [forrás: saját]	88
29. ábra A modellek grafikus összehasonlítása [forrás: saját]	89
30. ábra Hamis előrejelzések a hangolt modellekben [forrás: saját]	90
31. ábra A modellek kezdeti precision-recall pontelemzése [forrás: saját]	90
32. ábra A modellek hiperparaméter hangolás utáni precision-recall pontelemzése [forrás: saját]	91
33. ábra A kezdeti modellek ROC pontjai [forrás: saját]	91
34. ábra A modellek hiperparaméter hangolás utáni ROC pontjai [forrás: saját]	92

Táblázatjegyzék

1. táblázat Támadási volumen 2024-ben	12
2. táblázat Szakirodalmi adatbázisokon futtaott keresések kulcsszavai, kombinációi	21
3. táblázat PRISMA elemzés	23
4. táblázat Az érdekelt felek eltérő prioritásai	29
5. táblázat A döntéshozó informatikai besorolása és a vállalat bemutatott védelmi szintje közötti kapcsolat	39
6. táblázat Az otthoni munkavégzés és a zsarolóvírus-támadások közötti összefüggés	42
7. táblázat Az ARDS és a hagyományos védelmi megoldások összehasonlítása	50
8. táblázat A korábban tárgyalt adatvisszaállítási megoldások jellemzői	54
9. táblázat Olvasási teszteredmények 1 Gbit/s sávszélesség esetén	65
10. táblázat Írási teszteredmények 1 Gbit/s sávszélesség esetén	66
11. táblázat Olvasási teszteredmények 10 Gbit/s sávszélesség esetén	68
12. táblázat Írási teszteredmények 10 Gbit/s sávszélesség esetén	69
13. táblázat Referencia alapú teljesítménymetrikák az 1 Gbit/s-os kapcsolaton elvégzett tesztek alapján.....	70
14. táblázat Referencia alapú teljesítménymetrikák a 10 Gbit/s-os kapcsolaton elvégzett tesztek alapján.....	71
15. táblázat A vizsgált zsarolóvírus-detekciós módszerek összehasonlítása.....	75
16. táblázat Fájlműveleti sablonok	79
17. táblázat Tesztadatsor minta	81
18. táblázat A költségmátrix szemléltetése.....	85
19. táblázat Hiperparaméter keresési terek és a legjobb paraméter értékek	86
20. táblázat A gépi tanulási modellek teljesítményértékelési metrikái	88
21. táblázat A modellek költségmátrix elemzése	92
22. táblázat A javasolt megközelítés erősségei és korlátai	93
23. táblázat Főbb megközelítések és azok zsarolóvírus-felismerési paradigmákra vonatkozó megállapításai	94

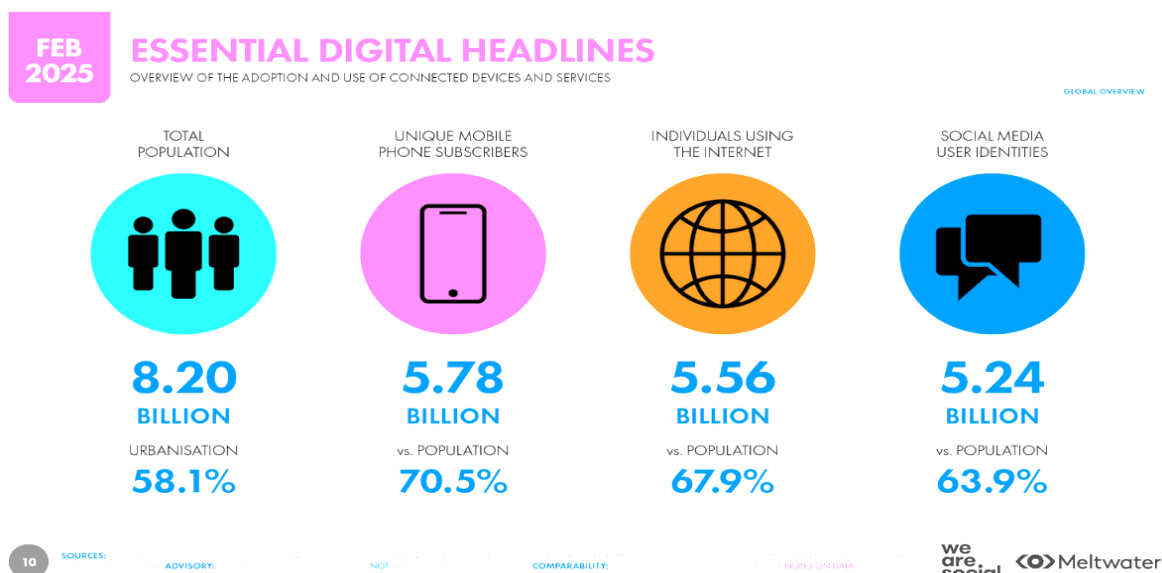
Rövidítésjegyzék

ADWIN Adaptive Windowing
AES Advanced Encryption Standard
AES-CTR Advanced Encryption Standard in Counter Mode
AES-GCM Advanced Encryption Standard in Galois/Counter Mode
AES-NI Advanced Encryption Standard New Instructions
AI Artificial Intelligence
API Application Programming Interface
APT Advanced Persistent Threat
AR Approximate Ranking number
ARC Adaptive Replacement Cache
ARDS Anti-Ransomware Defense System
BASH Bourne Again Shell
C2 Command and Control Infrastructure
CIS Cognitive Information System
CMADD Consistently Modifying Persistent Data of the Application
COTS Commercial-Off-The-Shelf
CPU Central Processing Unit
DAC Direct Attached Cable
DDM Drift Detection Method
DLP Data Loss Prevention
DNS Domain Name System
ECC-RAM Error Correction Code Random Access Memory
EDDM Early Drift Detection Method
EDR Endpoint Detection and Response
GDPR General Data Protection Regulation
GPO Group Policy Object
GPU Graphical Processing Unit
GUI Graphical User Interface
HCI Human Computer Interaction
HDFS Hadoop Distributed File System
ICN Information Centered Networking
IDS Intrusion Detection System
IPS Intrusion Prevention System
IoC Indicators of Compromise
IoT Internet of Things

iSCSI Internet Small Computer System Interface
KVM Kernel-based Virtual Machine
LIRS Low Inter-reference Recency Set
LLM Large Language Model
LotL Living off the Land
LXC Linux Containers
MITM Man in the middle
ML Machine Learning
NDN Named Data Networking
NIC Network Interface Card
OS Operating System
QEMU Quick Emulator
QoS Quality of Services
RaaS Ransomware-as-a-Service
RAID Redundant Array of Independent Disks
RDP Remote Desktop Protocol
RF Random Forest
ROC Receiver Operator Characteristic
RRNS Redundant Residual Number System
SATA Serial Advanced Technology Attachment
SFP Small Form-factor Pluggable network interface
SIEM Security Information and Event Management
SMB Server Message Block
SME Small and Medium-sized Enterprises
SOAR Security Orchestration, Automation, and Response
SSD Solid State Drive
SVM Support Vector Machine
VSS Volume Shadow Copy Service
XDR Extended Detection and Response
ZFS Zettabyte File System

1. Bevezetés

A digitális platformok és szolgáltatások elterjedése lehetővé tette az online kommunikáció széleskörű használatát és az online kereskedelem drasztikus mértékű növekedését. Az online oktatási platformok növelték a tanulás rugalmasságát, míg az online banki és számlázási megoldások egyszerűsítették az adminisztrációt és a pénzügyi műveleteket. Az e-közigazgatás bevezetése teljesen átalakította a polgárok és a kormányok közötti interakciót, míg a virtuális magánhálózatok (VPN) biztonságos távoli hozzáférést biztosítottak a mobilitást, vagy éppen az otthoni munkavégzést igénylők számára. A pandémiás időszak szintén erős katalizátorszerepet játszott a digitalizációs folyamat felgyorsulásában. A globális digitalizáció jelenlegi helyzetét jól szemlélteti az 1. ábra, amin leolvasható a teljes populáció arányában az internet hozzáféréssel, egyéni mobil előfizetéssel és közösségi média profillal rendelkezők száma.



1. ábra Teljes populációra vonatkozó digitalizációs szám adatok [1]

Mind a vállalkozások, mind a magánszemélyek egyre inkább a különböző digitális kommunikációs platformok kombinációját használják a saját igényeik kielégítésére. Ezek leginkább az e-mailt, az azonnali üzenetküldő alkalmazásokat és a videokonferenciákat foglalják magukba. A digitalizáció nem kerülte el az oktatási intézményeket sem. A koronavírus járvány időszaka óriási kihívások elé állította az oktatási rendszereket és radikális szemléletváltást követelt meg. Ezzel világossá vált tovább léptünk az innovatív e-learning módszerek, a távoktatás és az online oktatási platformok használatának irányába. A technológia fejlődésével párhuzamosan a felhasználói oldalról érkező igények is komoly kihívások elé állítják a fejlesztőket, mind a hardver-, mind a szoftverfejlesztés területén. A digitalizált tevékenységek köre változó, de jellemzően egyre több adat és annak kezelése kerül a digitális térbe, amely egyrészt információt hordoz és értéket képvisel, másrészt folyamatos védelmet és fokozott

elővigyázatosságot igényel. Napjaink digitális világa olyan mértékben befolyásolja a mindennapi életet, hogy alig van olyan gazdasági ágazat, ahol ne lennének informatikai rendszerek, legyen szó akár termelési folyamatirányításról, logisztikai rendszerekről, vagy a szervezet működéséhez elengedhetetlen adminisztrációs adatvagyon tárolásáról. A digitalizáció, a „dolgok internete” (IoT) és a digitális megoldások folyamatos térnyerése az élet szinte minden szegmensében megjelenik, azonban ez nemcsak előnyökkel, hanem kockázatokkal és felelősséggel is jár. A technológia fejlődésével párhuzamosan megjelent a kiberbűnözők egy magas szintű szakértelemmel rendelkező rétege, akik folyamatosan fejlődő technológiai eszközkészlettel használják ki a digitális sebezhetőségeket, illetve a felhasználók tájékozatlanságából, jóhiszeműségéből és tudatlanságából adódó lehetőségeket. Világszerte komoly kihívást jelent az információs és informatikai rendszerek komplex védelme. A személyes adatok védelmére ugyan külön jogi keretek és előírások vonatkoznak, de kérdéses egyrészt, hogy az érintettek mennyire tudnak megfelelni ezeknek a szabályozásoknak, másrészt pedig, hogy az előírások mennyire kivitelezhetőek a gyakorlatban. Az elmúlt néhány évben, ha valaki követte az informatikai biztonsággal kapcsolatos híreket, olvashatott a zsarolóvírus-támadások növekvő számáról is. Ezek a támadások sajnos szinte válogatás nélkül érintettek kis- és nagyvállalatokat egyaránt a világ minden tájáról. Ezekben az esetekben a hackerek általában rövid, 24-48 órás határidővel váltságdíjat követelnek, jellemzően Bitcoinban. 2024-ben továbbra is a zsarolóvírus volt a leggyakrabban előforduló kiberbűnözési forma világszerte, amit az 1. táblázat adatai is alátámasztanak [3, 4]. A Sophos „State of Ransomware 2024” adatai szerint 2024-ben a szervezetek 59%-át érte zsarolóvírus-támadás, ami sokkal nagyobb mértékű „penetrációt” mutat, mint amit más egyedi támadásoknál jellemzően láthatunk (pl.: DDoS, exploit alapú támadások) [2].

1. táblázat Támadási volumen 2024-ben

Mutató	Érték	Forrás
Zsarolóvírus-támadással érintett szervezetek aránya (5000 interjúalany)	59%	Sophos State of Ransomware 2024 (Sophos News) [2]
Publikált zsarolóvírus incidensek száma globálisan	5 414	Cyberint 2024 Annual Report (Cyberint) [3]
Éves növekedés 2023-hoz képest	+11%	Cyberint 2024 Annual Report (Cyberint) [3]
A zsarolóvírus kiemelt fenyegetés-e a kiberfenyegetések között?	igen	ENISA Threat Landscape 2024 (ENISA) [4]

A zsarolóvírus-támadások volumenének növekedését egyértelműen jelzi az okozott kár mértékének jelentős növekedése is. Az elmúlt néhány évben világszerte elszaporodtak a közintézmények, oktatási intézmények és sajnos az egészségügyi intézmények elleni kibertámadások is. Tovább nehezíti a támadások kivizsgálását és az adatok számszerűsítését az a tény,

hogyan az ilyen esetekben az érintettek gyakran nem jelentik az incidenst a hatóságoknak. Ennek oka egyrészt egyfajta rossz szokás, mivel az áldozatok általában szégyellik, hogy áldozatok, másrészt sok esetben valóban kapcsolatba léptek a kiberbűnözőkkel és a váltságdíjat ki is fizették, azonban tartva az esetleges bírságoktól (GDPR), reputációs veszteségektől és ügyféloldali bizalomvesztéstől inkább a hallgatás mellett döntenek. Az is többször előfordult fizetés esetén, hogy a támadók nem adtak hozzáférést később sem az adatokhoz, illetve az ellopott adatvagyon értékesítették a dark WEB-en, vagy publikusan elérhetővé tették. A zsarolóvírusok töretlen ütemű terjedése és technológiai fejlődése védelmi oldalon koncepcióváltást, az alternatív védelmi képességek fejlesztését és többfaktoros jelzőrendszerek implementálását teszik szükségessé. Az olyan hagyományos védelmi technológiák, mint például a tűzfalak, IDS/IPS rendszerek, SIEM/SOAR platformok, DLP rendszerek, vagy vírusvédelmi megoldások mellett a védelem másik kritikus pontja a tárolt adatok védelme, biztonságos kezelése, az adatok sérülése esetén a biztonságos és gyors helyreállítás feltételeinek megteremtése lett. A megbízható és gyors fájlkiszolgáló, illetve a hozzá kapcsolódó biztonsági mentési rendszer elengedhetetlen centralizált adattárolás esetén a produktív hálózati környezetekben. A leggyakrabban használt Microsoft Windows Server környezetek esetében azonban az árnyékmásolatok gyakran nem nyújtanak elegendő védelmet a zsarolóvírusok ellen, mert vagy eleve nincsen bekapcsolva a szolgáltatás, vagy a támadás során deaktiválják a támadók [5, 6]. Ezek a sokszor frissítések nélkül használt rendszerek fokozottan ki vannak téve széles körben kihasznált Windows specifikus, nulladik napi (zero day) sérülékenységeket (pl.:RDP, Exchange, OWA) kihasználó exploitoknak.

1.1. A zsarolóvírus-támadások gazdasági vonatkozásai

Az adatvagyon és a hozzá kapcsolódó szoftverkörnyezetek jelenleg a legnagyobb technológiai értéket képviselik. A zsarolóvírusok fejlesztői ezt kihasználva szinte minden olyan szervezetet megpróbálnak megtámadni, amely számára a fenti kettő elengedhetetlen a működéséhez. Napjainkban az informatikai rendszereket érintő biztonsági incidensek közel egyharmadáért ez a támadási forma a felelős. A kriptovaluták folyamatos térnyerése és árfolyamuk hosszú távú, tendenciális emelkedése (2. ábra) nagyon vonzóvá teszi a kiberbűnözők számára a váltságdíjként való bevonásukat. Tovább ront a helyzeten, hogy a technológiai ipar számos prominens szereplője is elfogadja ezeket a kriptovalutákat fizetőeszközként.

Bitcoin–USA-dollár

59 468,90 ↑ 84,60% +27 254,08 5 É

jún. 26., 12:39:05 UTC · Felelősségkizárás

1 N 5 N 1 H 6 H YTD 1 É 5 É MAX



2. ábra A Bitcoin árfolyam változása 5 éves intervallumra vonatkozóan -
forrás: Google Finance [7]

A Bitcoin felértékelődésével sajnos fokozódtak a váltságdíj-típusú támadások, ugyanakkor a támadók által Bitcoinban követelt váltságdíj ellenértéke is megsokszorozódott. Példaként lehet említeni, hogy amíg 2020 tavaszán a BTC 17 185 dollárért volt átváltható, addig ugyanez 2021 áprilisában 62 926 dolláros árfolyamot jelentett. Ez ugyanakkor azt is jelenti, hogy bár a zsarolóvírus-támadások során Bitcoinban meghatározott összeg nem változott, a vállalatok és intézmények mégis nagyságrendekkel többet fizettek ki a zsarolóknak az adatok visszaszerzéséért és az adatszivárgás megakadályozásáért. 2019-ben 10,1 milliárd euró váltságdíjat utaltak el az áldozatok az EU-n belül, ami 3,3 milliárd euróval több, mint 2018-ban. A felderített támadások száma 2018 és 2019 első 6 hónapja során 365%-os növekedést mutatott, miközben az áldozatok átlagosan az esetek 45%-ában fizettek váltságdíjat, azonban közel felük elveszítette adatait [8]. 2018-ban a támadók még főként multinacionális vállalatokat céloztak meg, 2019 óta azonban fokozatosan célponttá vált a közszféra és annak az oktatást és az egészségügyet biztosító szolgáltatásai is. Aggodalomra ad okot az is, hogy a 2019-es jelentések szerint az EU-ban az egészségügyi intézmények 66%-át érte ilyen jellegű támadás [9]. Ezenkívül a COVID-19 világjárvány okozta helyzet miatt az egészségügyi intézmények kitettsége tovább növekedett. Az első regisztrált zsarolóvírus okozta haláleset egy düsseldorfi klinikán történt 2020. szeptember 10-én, miközben széles körű támadási hullám söpört végig a német egészségügyi intézményeken [10]. A 2020-as év becslései szerint a zsarolóvírusok összesen 400 milliárd eurós kárt okoztak az EU-ban. Az amerikai statisztikák szerint 2020-ban közel 184 millió támadást észleltek [11]. Eközben a titkosításhoz használt algoritmusok növekvő komplexitása és a kulcsok hosszának növekedése is fokozatosan nehezítette a támadások során titkosított állományok dekriptálását. A támadók ekkorra már

gyakran nem csak azzal gyakoroltak nyomást, hogy túsul ejtették az adatvagyon, hanem azzal is fenyegették az érintetteket, hogy nyilvánosságra hozzák, vagy eladják az eltulajdonított adataikat. A fenyegetésekről készült riportok rendre szintén megerősítik, hogy ezek a kártevők jelentős kiberfenyegetést jelentenek. Az ENISA Threat Landscape 2024 jelentés a zsarolóvírusokat a legtöbb kárt okozó támadástípusok között sorolja fel [12], míg a Verizon DBIR 2025 jelentés szerint a jelentett biztonsági incidensek 44%-ában szerepet játszottak [13]. Az ágazatspecifikus elemzések, mint például az ENISA Finance Threat Landscape 2024 szintén kiemelik az ilyen típusú malware-ek hosszútávú hatását a kritikus iparágakra [4]. Az Europol Internet Organised Crime Threat Assessment (IOCTA) jelentése ugyancsak a zsarolóvírusokat azonosítja az elmúlt évek egyik legjelentősebb és legkárosabb formájaként a kiberbűnözés területén [14]. Ezek a független jelentések együttesen megerősítik a fenyegetés jelentős mértékét és tartós jellegét, ezzel is ösztönözve a hatékonyabb felderítési stratégiák kutatását. 2022-ben a támadások számának jelentős csökkenése miatt a fizetett váltságdíjak összege 416 millió dollárral (összesen 567 millió dollárra) csökkent 2021-hez képest. 2023-ban azonban a zsarolóvírus-támadások száma megugrott, és a fizetett váltságdíjak összege új rekordot döntött, így 1,1 milliárd dollárt tett ki [15].

1.2. Támadási vektorok, kiemelten veszélyeztetett területek

A támadási vektorokkal kapcsolatosan Natalie Paskoski, aki az RH-ISAC marketing kommunikációs vezetője ismertette, hogy a Unit42 incidensválaszok és az adatvédelmi incidensekről szóló jelentések szerint a Microsoft Remote Desktop Protocol (RDP) érintettsége a zsarolóvírus-támadásokkal kapcsolatosan több, mint 50%-os [16]. Ezt követi a sorban az e-mailes adathalászat, ami a fertőzések körülbelül negyedét teszi ki, majd pedig a szoftver sebezhetőségek kihasználása következik 12%-kal. Míg korábban egy zsarolóvírus-támadáshoz több éves fejlesztési, kriptográfiai, üzemeltetési és behatolásvizsgálati tapasztalatra volt szükség, addig mára már a Ransomware-as-a-Service (RaaS) platformok egyre elterjedtebbé és széles körben elérhetővé váltak, ezzel szakértelem nélküli személyek számára is lehetőség nyílt a támadások kivitelezésére. Az informatikai rendszerek sok helyen elavultak, ami magas kockázatot jelent. Például az egészségügyi rendszerek általánosan sebezhető és vonzó terület a támadók számára [6], hiszen, ha egy egészségügyi szolgáltató elveszíti az adataihoz való hozzáférést, akkor nincs más választása, mint kifizetni a váltságdíjat. Az alapvető probléma az, hogy a versenyképes védelmi megoldások fejlett hardvereszközöket, magas szintű szakmai ismereteket és folyamatos pénzügyi beruházásokat igényelnek, ezek pedig gyakran nem állnak rendelkezésre a kis- és középvállalkozások, valamint az oktatási és egészségügyi intézmények esetében. Így a gyakran elavult szoftverkörnyezet, a túlterhelt rendszergazdai csapatok és a kritikus szolgáltatások érzékenysége miatt ezek a rendszerek egyre inkább célpontokká válnak.

1.3. Motiváció

Az informatikai rendszerek komplexitásának növekedése, a felhasználói oldalon tapasztalható biztonságtudatossági hiányosságok, a szakmai kompetenciák hiánya, valamint az elavult vagy nem megfelelően karbantartott rendszerek egyaránt hozzájárulnak a kibertámadásoknak való kitettséghez. A zsarolóvírus-támadások napjaink talán legsúlyosabb kiberbiztonsági fenyegetését jelentik, amit a támadási módszerek folyamatos fejlődése és a Ransomware-as-a-Service (RaaS) platformok elterjedése még inkább fokoz. A korábban magas szintű szakértelmet igénylő támadások végrehajtása mára lényegesen egyszerűbbé vált, miközben az egészségügyi és oktatási intézmények, valamint a mikro-, kis- és középvállalkozások elavult informatikai infrastruktúrájuk, korlátozott erőforrásaik és a rendelkezésre álló szakértelem hiánya miatt egyre kiszolgáltatottabb célpontokká váltak. Az olyan védelmi megoldások, mint például a Next Generation (NG) tűzfalak, a behatolásjelző és megelőző rendszerek (IDS/IPS), a biztonsági információ- és eseménykezelő (SIEM) platformok, az adatszivárgás megelőző (DLP) szoftverek és az antivírus megoldások ugyan fontos szerepet töltenek be az informatikai infrastruktúrák védelmében, azonban önmagukban sokszor nem elegendők a zsarolóvírus-támadások hatékony kivédéséhez és a károk minimalizálásához. A hangsúly gyakran a határvédelmen és a kliens eszközök védelmén van, miközben a centralizáltan tárolt adatvagyon védelme háttérbe szorul, holott kritikus feladat lenne a tárolt adatokon végzett műveletek felügyelete és a gyors helyreállítás biztosítása. A kritikus adatok védelme érdekében olyan költséghatékony és könnyen integrálható megoldásokra van szükség, amelyek a végpontvédelmi rendszerektől függetlenül képesek biztosítani a zsarolóvírus-támadások hatékony felismerését, az adatok biztonságos tárolását és gyors helyreállítását alacsony erőforrásigény mellett. Az értekezésben ismertetett kutatások legfőbb célja egy szerveroldali, a végpontvédelmi rendszerektől független védelmi megközelítés gyakorlati alkalmazhatóságának vizsgálata, amely referenciaalapú teljesítményértékelésre és gépi tanulással támogatott zsarolóvírus-detektálásra épül.

1.4. A kutatás célja, várható eredménye

A bevezetésben bemutatott támadási trendek, valamint a mikro-, kis- és középvállalkozások körében végzett feltáró jellegű kérdőíves felmérés eredményei egyaránt arra utaltak, hogy igény mutatkozik egy, a végpontvédelmi rendszerektől független, könnyen integrálható és komplex védelmi képességekkel rendelkező szerveroldali védelmi megközelítésre. Ennek megfelelően a kutatás első célja az ARDS (Anti-Ransomware Defense System) koncepcionális modell kidolgozása, valamint a modell részét képező fájlkiszolgáló oldali védelmi megoldás gyakorlati alkalmazhatóságának igazolása referenciaalapú teljesítményértékeléssel. A kutatás második célja annak vizsgálata volt, hogy a szerveroldalon naplózott fájlműveleti szekvenciák felhasználhatók-e a zsarolóvírus-támadások közel valós idejű, alacsony erőforrásigényű felismerésére felügyelt gépi tanulási módszerek alkalmazásával. A kutatás várható eredményeként egyrészt egy referenciaalapú fájlkiszolgáló-teljesítményértékelési módszer, másrészt egy szerveroldali fájl-

műveleti szekvenciák elemzésére épülő, gépi tanulást alkalmazó detektálási módszer kerül bemutatásra, amely hozzájárulhat a centralizált adattárolási környezetek zsarolóvírusok elleni védelmének hatékonyabbá tételéhez.

1.5. Az értekezés tematikája

A bevezetésben először ismertetem a zsarolóvírus-támadásokkal kapcsolatos trendeket, az ilyen jellegű incidensek gazdasági vonatkozásait, majd összefoglalom a leggyakoribb támadási vektorokat, illetve a leginkább érintett területeket. Ezt követően bemutatom a kutatásaim motivációját, a célkitűzéseimet és a várható eredményeket. Az 1.6-os alfejezet összefoglalja az eddigi publikációs teljesítményemet és a hivatkozási statisztikáimat. A következő alfejezet módszertani áttekintést nyújt a szakirodalomfeldolgozással, a kérdőíves felméréssel és az mesterséges intelligencia eszköztár kiválasztásával kapcsolatosan. A 3. fejezet a fertőzési tendenciák, biztonságos fájlátviteli megoldások és a vírus felismerési technikák szakirodalmi áttekintését tartalmazza. A 4. fejezetben a hazai kis- és középvállalkozások (KKV-k) körében készített, informatikai biztonságra vonatkozó kérdőíves felmérést ismertetem részletesen, majd az ARDS védelmi modellt mutatom be. A következő fejezetben az ARDS modell részét képező egyedi fájlkezelő megoldás összehasonlító teljesítményértékelését és annak eredményeit ismertetem részletesen. A 6. fejezet az utolsó kutatási szakasz ismertetését foglalja magában. Ebben a fejezetben a zsarolóvírus-támadások felismerésének lehetőségét vizsgálom fájlműveleti naplófájlok alapján, gépi tanulási modellek segítségével. A következő fejezetben ismertetem a doktori értekezés téziseit, a tézisekkel kapcsolatos eredményeket, illetve a hozzájuk kapcsolódó publikációkat. Az utolsó fejezet összefoglalja az eredmények tudományos jelentőségét, az eredmények korlátait és értelmezési megfontolásait, illetve a jövőbeni kutatási irányokat és feladatokat.

1.6. Publikációs teljesítmény és hivatkozások

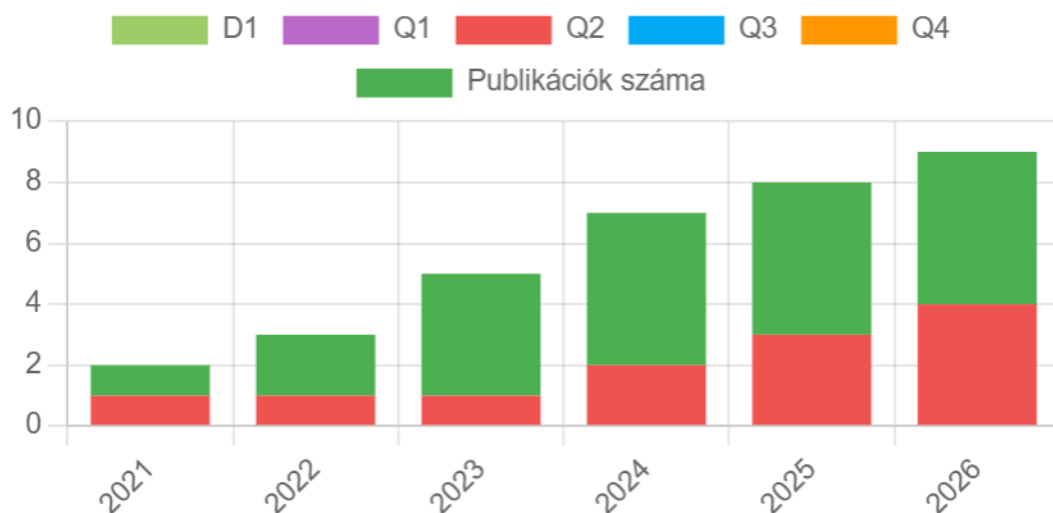
Publikációk	Külföldön	Magyarországon		pontszám
		idegen nyelven	magyarul	
	megjelent közlemények száma			
Tudományos folyóiratcikk				
-lektorált	3	1	0	3.446
-lektorált, WoS Q1 rangú	0	0	0	
-lektorált, IF-ral	3	1	0	
-ebből egyszerűs	0	0	0	
Konferenciák (min. 4 oldal) konferenciakötetben, folyóiratban, könyvrészletben	0	1	0	0.067
Tudományos cikk összesen				3.513
Tudományos könyv, könyvrészlet szerzőként				
-könyv	0	0	0	0.000
-könyvrészlet	0	0	0	0.000
Tudományos könyv összesen				0.000
Internetes adatbázisban megtalálható	3	1	0	
Teljes szöveggel elérhető a Weben	3 (?)	1 (?)	0 (?)	

3. ábra Publikációs tevékenység összefoglaló táblázata

Hivatkozások darabszáma	könyvben, fejezetben	folyó- iratban	konf. cikkben	érteke- zésben	eddig összesen	oltalmi formában	további köz- leményben
csak külföldi szerző, külföldi kiadás	0	1	0	0	1	0	0
csak külföldi szerző, hazai kiadás vagy nincs megjelölve	0	0	0	0	0	0	0
hazai szerző (is) itthon (vagy nincs megjelölve hol)							
-idegen nyelven	0	0	0	0	0	0	0
-magyarul	0	0	0	0	0	0	0
nincs megadva, hogy van-e hazai szerző az idézők között	0	6	0	0	6	0	0
Összesen	0	7	0	0	7	0	0

4. ábra Idézettség összefoglaló táblázata

A megjelent közleményeim 6 független WoS hivatkozással és 1 darab független Scopus hivatkozással rendelkeznek az értekezés benyújtásakor.



5. ábra Publikációs tevékenység eloszlása a doktori képzés kezdetétől az értekezés benyújtásáig

A 3., 4., és 5. ábra szemlélteti az MTMT adatbázisában nyilvántartott közleményeimet, valamint a publikációs tevékenységem folyamatát a doktori képzés megkezdése óta. A megjelent publikációk közül 4 darab WoS Q2-es, impakt faktoros publikációm jelent meg.

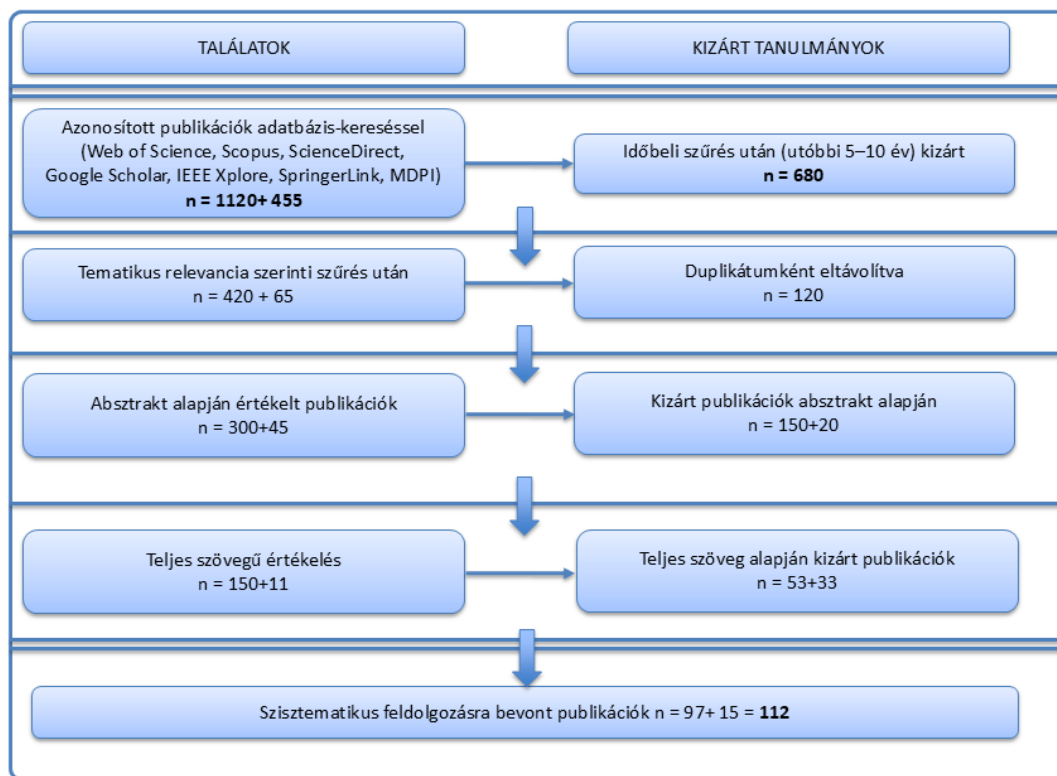
2. Kutatásmódszertan

Ebben a fejezetben a kutatási módszereket ismertetem, amelyek alapján a szakirodalom feldolgozását, a kérdőíves vizsgálatot és a fájlserver-védelmi megoldás kidolgozása során alkalmazott mesterséges intelligencia eszközök kiválasztását végeztem.

2.1. Szakirodalomfeldolgozás módszerének ismertetése

A feldolgozás PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) módszertani keretrendszer szerint történt, amely biztosítja az irodalomkeresés átláthatóságát, reprodukálhatóságát és módszertani megalapozottságát (6. ábra).

PRISMA folyamatábra – szakirodalmi feldolgozáshoz



6. ábra PRISMA flow diagram szakirodalom feldolgozáshoz [forrás: saját]

Keresési stratégia és adatforrások

A PRISMA módszertan alkalmazása során a kezdeti kulcsszavas keresések eredményeként összesen mintegy 1 120 publikáció került azonosításra (3. táblázat) a Web of Science, Scopus, ScienceDirect, Google Scholar, valamint a releváns kiadói adatbázisok (IEEE Xplore, SpringerLink, MDPI) bevonásával. A kutatás tématerületekhez illeszkedően több szakaszban végeztem az megfelelő szakirodalmi háttér kutatását, a módszertan bemutatásakor ezek összesített értékei kerültek ismertetésre. A keresések strukturált kulcsszókombinációk alkalmazásával történtek (2. táblázat), amelyek lefedték a kutatási téma fő dimenzióit (pl. ransomware detection, file system monitoring, machine learning cybersecurity, log-based detection, SME cybersecurity), ezeket a 2. táblázatban foglaltam össze. A kezdeti kulcsszavas keresések eredményeként összesen 1 120 publikáció került azonosításra az egyes adatbázisokban, majd a kutatási eredmények kiértékelésekor további, aktuális szakirodalom forrásokat kerestem, amelyeknek az eredményeivel a saját eredményemet összevettem.

2. táblázat Szakirodalmi adatbázisokon futtatt keresések kulcsszavai, kombinációi

Kategória	Keresőkifejezés
Ransomware	„ransomware” VAGY „crypto-ransomware” VAGY „file-encrypting malware” VAGY „ransomware detection”
Detection	„log-based detection” VAGY „identification” VAGY „classification”
Monitoring	„file system monitoring” VAGY „file activity monitoring” VAGY „file operation logs”
AI/ML	„machine learning” VAGY „anomaly detection” VAGY „log analysis”
Environment	„server-side” VAGY „file server” VAGY „SME cybersecurity”
Security awareness	„cybersecurity awareness” VAGY „online safety” VAGY „digital activities”

Időbeli és tematikus szűrés

A PRISMA módszertan ajánlásait követve az első szűrési lépésben az időbeli relevancia került alkalmazásra. A publikációk körét a friss tanulmányokra szűkítettem, figyelembe véve a zsarolóvírusok és a gépi tanulási módszerek gyors technológiai fejlődését. Az időbeli szűrés eredményeként – az elmúlt 5–10 évben megjelent publikációkra korlátozva – a vizsgált publikációk száma 1 120-ról 680 darabra csökkent. Ezt követően szakterületi relevancia szerinti szűrést végeztem. Kizárásra kerültek azok a publikációk, amelyek kizárólag kriptográfiai algoritmusok elméleti elemzésére fókuszáltak, nem informatikai biztonsági kontextusban alkalmaztak gépi tanulást, nem érintették sem a zsarolóvírus-fenyegetést, sem a fájlrendszer- vagy naplóalapú megközelítéseket. A szakterületi relevancia alapján további szűrés során a nem informatikai biztonsági, illetve nem ransomware-fókuszú publikációk is kizárásra kerültek, így 420 tanulmány maradt a további feldolgozásra. A kutatás záró szakaszában az eredményeim nemzetközi szakirodalomban publikált módszerekkel történő összehasonlításához ismételt keresést futtattam a WoS adatbázisán, és kizárólag a 2025-2026 évre korlátoztam a keresést annak érdekében, hogy a legfrissebb publikált eredményeket tudjam felhasználni. A kulcsszavak tekintetében célzottan a „ransomware detection (308 találat)”, „ransomware detection AND SME (3 találat)”, „ransomware detection AND file IO (3 találat)”, „ransomware detection AND machine learning (137 találat)” keresőkifejezésekre fókuszáltam, amelyek a saját eredmények összehasonlíthatóságához voltak szükségesek.

Duplikátumok eltávolítása

Az adatbázisok közötti átfedések miatt a címek, szerzők és DOI-azonosítók összevetésével 120 duplikált rekord került eltávolításra. A duplikátumok eltávolítását követően 300 egyedi publikáció képezte az absztrakt-alapú elemzés alapját. Külön figyelmet fordítottam a pre-print megjelenésű, DOI-val rendelkező találatok egyeztetésére és a végleges megjelenéskori

verziók duplikációként való kezelésére. A saját eredmények összehasonlítására irányuló aktuális keresés esetén csak a kombinált kulcsszavakat tartalmazó kereséseket szűrtem, önmagában a ransomware detection minden további kulcsszavas kereséséi találatot magában foglalt, a kutatási eredmények szempontjából azonban elegendő volt a kombinált kulcsszavas eredmények alapján a további szűréseket elvégezni.

Absztrakt-alapú szűrés

Az absztraktok értékelése során kizárásra kerültek azok a tanulmányok, amelyek nem tartalmaztak módszertani vagy empirikus hozzájárulást, illetve nem voltak relevánsak a fájlkihasználó-alapú vagy naplóalapú detektálás szempontjából. Kizárásra kerültek továbbá azok a tanulmányok, amelyek kizárólag kliensoldali vagy végpontvédelmi megoldásokkal foglalkoztak szerveroldali kontextus nélkül, illetve nem tárgyalták a detektálás időbeliségét vagy gyakorlati alkalmazhatóságát. Az absztrakt-alapú értékelést követően mintegy 150 publikáció került kizárásra.

Teljes szövegű értékelés és végső kiválasztás

A fennmaradó 150 publikáció esetében teljes szövegű elemzés történt, amelynek eredményeként 97 publikáció bizonyult relevánsnak a kutatási terület szempontjából. Ezek a tanulmányok szolgálták a kutatási háttér, a kapcsolódó tudományos eredmények és a módszertani megközelítések átfogó feltárásának alapjául, és megalapozták a kutatási kérdések megfogalmazását, a saját módszertani megközelítés kialakítását. A disszertáció szakirodalmi áttekintésébe és hivatkozási rendszerébe 68, majd további 15 olyan szakirodalmi forrás került beépítésre, amelyek közvetlenül kapcsolódtak a kutatás célkitűzéseéhez, a vizsgált problématerülethez, valamint a saját módszertani megoldásom és az eredmények értelmezéséhez. A fennmaradó publikációk a kutatási terület szélesebb körű megértését, a fejlődési irányok és megközelítések feltérképezését támogatták, azonban tartalmuk nem kapcsolódott olyan mértékben a disszertáció konkrét kutatási kérdéseéhez, hogy közvetlen hivatkozásként szerepeljenek a dolgozatban, valamint bekerültek a források közé olyan webes referencia oldalak, amelyek a tématerület nem kizárólag tudományos jellegű alátámasztását szolgálták, pl. nemzetközi szervezetek hivatalos oldalai, szabvány és technológiai ismertetéshez háttér hivatkozások (további 29 forrás). A kiválasztás során elsőbbséget élveztek azok a publikációk, amelyek részletesen ismertették az alkalmazott módszertant, napló- vagy fájlrendszer-alapú detekciós megközelítést alkalmaztak, gépi tanulási vagy viselkedéselemzési módszerekre épültek, illetve gyakorlati implementációs vagy teljesítményértékelési szempontokat is tartalmaztak. A kutatási eredmények kiértékeléséhez a feldolgozott irodalomforrások alapján az értekezés diszkussziós fejezetében kerül ismertetésre a jelenlegi legfrissebb, nemzetközi kutatási eredményekkel való összehasonlító tárgyalása az elért eredményeimnek.

3. táblázat PRISMA elemzés

PRISMA folyamat lépései	Publikációk száma (-2024)	Publikációk száma (2025-2026 WoS adatbázisán szűkített kereséssel)
Kezdeti kulcsszavas keresés (összes adatbázis)	~1 120	
Időbeli szűrés	~680	451
Tematikus relevancia szerinti szűrés	~420	65
Duplikátumok eltávolítása	~300	--
Absztrakt-alapú kizárás	~150	45
Teljes szövegű elemzés	150	23
Feldolgozott források	97	15
Felhasznált források száma összesítve:		112

2.2. A kérdőíves felmérés módszertani bemutatása

A kutatás célja

A kérdőíves felmérés célja a hazai kis- és középvállalkozások informatikai biztonsági felkészültségének feltáró jellegű feltérképezése volt, különös tekintettel a malware- és zsarolóvírus-támadásokkal kapcsolatos tapasztalatokra, a védelmi gyakorlatokra, valamint az informatikai döntéshozók kompetenciaszintjére. A vizsgálat célja nem statisztikailag reprezentatív országos kép kialakítása volt, hanem olyan gyakorlati mintázatok, hiányosságok és összefüggések azonosítása, amelyek közvetlenül hasznosíthatók egy KKV-k számára is alkalmazható védelmi modell tervezésénél. A felmérés módszertani szerepe az értekezésben kettős: egyrészt empirikus alapot biztosít a kutatási probléma relevanciájának igazolásához, másrészt bemeneti információkat szolgáltat a javasolt védelmi modell követelményrendszerének kialakításához.

A célcsoport és a mintavételi stratégia

A kérdőív célcsoportját magyarországi kis- és középvállalkozások alkották, amelyek működésük során informatikai rendszereket, digitális adatvagyonot és hálózati infrastruktúrát használnak. A vizsgálat során olyan vállalatok kerültek bevonásra, amelyek esetében az informatikai üzemeltetés jellemzően saját erőforrásokra vagy kis létszámú külső szolgáltatóra támaszkodik, ami tipikusnak tekinthető a KKV-szektorban. A mintavétel célzott (purposive) megközelítéssel történt. A vizsgálat során nem került sor a mikro-, kis- és középvállalkozá-

sok előzetes kizárására, mivel a kutatás szempontjából valamennyi vállalati kategória releváns lehet a malware- és zsarolóvírus-fenyegetettség vizsgálata során. A támadások nem kizárólag a nagyobb szervezeteket érintik, hanem a korlátozott erőforrásokkal rendelkező mikro- és kisvállalkozásokat is, amelyek gyakran fokozottan kitettek az információbiztonsági kockázatoknak. A vizsgált szervezetek bevonása elsősorban szakmai kapcsolatrendszeren keresztül történt. Ennek oka az volt, hogy a kutatás érzékeny tématerületet érintett, hiszen a kérdőív többek között korábbi informatikai incidensekre, malware- és zsarolóvírustámadásokra, valamint a szervezetek információbiztonsági hiányosságaira vonatkozó kérdéseket is tartalmazott. A személyes szakmai kapcsolat jelentősen növelte a válaszadási hajlandóságot és a válaszok őszinteségét, különösen olyan kérdések esetében, amelyek a vállalatok számára reputációs vagy üzleti kockázatot hordozhatnak. A kutatás célja nem statisztikailag reprezentatív minta kialakítása volt, hanem különböző méretű és eltérő informatikai érettségi szinttel rendelkező vállalkozások gyakorlati tapasztalatainak feltárása. A zsarolóvírus-incidensekkel kapcsolatos adatok jelentős része rejtett információ. A kutatás egyik kihívása éppen az volt, hogy olyan szervezeteket érjek el, amelyek hajlandóak valós tapasztalataikról beszámolni. A zsarolóvírus-incidensekkel kapcsolatos tapasztalatok jelentős része nem nyilvános, ezért a bizalmi alapon történő célzott mintavétel a kutatás feltáró jellegéből következően indokoltnak tekinthető. Emiatt a bizalmi alapú célzott mintavétel nem a kutatás korlátja, hanem bizonyos értelemben annak előfeltétele volt. A célzott mintavétel ezért a kutatás feltáró jellegével összhangban került alkalmazásra. A kérdőívet összesen 46 vállalat töltötte ki, ami lehetővé tette az alapvető összefüggések vizsgálatát.

A kérdőív felépítése és kérdéscsoportjai

A kérdőív strukturált felépítésű és összesen 86 kérdést tartalmaz (ld.: melléklet), amelyek logikusan egymásra épülő kérdéscsoportokba rendeződtek. A kérdésblokkok kialakítása során törekedtem arra, hogy a megkérdezett vállalatok működésének általános jellemzőitől fokozatosan haladjak az informatikai biztonság specifikus területei felé. A kérdőív nem egyetlen korábban publikált kérdőív közvetlen adaptációjaként készült. A kérdésblokkok kialakítása a kiberbiztonsági érettség, az információbiztonsági kontrollok, a szervezeti felkészültség és az incidenskezelés témakörében megjelent szakirodalom szintetizálásával történt. A kérdőív felépítése emellett igazodik az ISO/IEC 27001 szabvány, a NIST Cybersecurity Framework, valamint az ENISA KKV-kra vonatkozó kiberbiztonsági ajánlásaiban megjelenő fő vizsgálati dimenziókhoz (szervezeti jellemzők, technológiai környezet, humán tényezők, biztonsági kontrollok, incidensek, kockázati tényezők, fejlesztési szándék). A kérdések végleges kialakításában a szerző információ-biztonsági és rendszerüzemeltetési tapasztalatai is szerepet játszottak. A kérdőív elsődleges célja nem egy validált mérőskála alkalmazása vagy új mérőeszköz létrehozása volt, hanem a KKV-k információbiztonsági helyzetének feltárása és a kutatás további szakaszait megalapozó gyakorlati tapasztalatok összegyűjtése.

A kérdőív főbb kérdéscsoportjai az alábbiak voltak:

1. A vállalat szervezeti adatai
(tevékenységi kör, méret, telephelyek száma, alkalmazottak létszáma)

2. Informatikai infrastruktúra jellemzői
(munkaállomások és szerverek száma, használt operációs rendszerek, hálózati felépítés, megosztott erőforrások)
3. Informatikai döntéshozók kompetenciái
(informatikai végzettség, képzettségi szint, szerepkör)
4. Biztonsági gyakorlatok és védelmi megoldások
(vírusirtók, hálózatbiztonsági eszközök, mentési megoldások, belső szabályzatok)
5. Korábbi biztonsági incidensek
(vírustámadások előfordulása, zsarolóvírus-tapasztalatok, incidensek következményei)
6. Munkavégzési szokások és kockázati tényezők
(táv munka, saját eszközök használata, felhasználói jogosultságok)
7. Nyitottság új védelmi megoldások iránt
(költségérzékenység, automatizáltság iránti igény, szakértői támogatás szükségessége)

Alkalmazott kérdéstípusok

A kérdőívben többféle kérdéstípus került alkalmazásra annak érdekében, hogy a válaszok statisztikai feldolgozásra alkalmasak legyenek, ugyanakkor a kitöltők számára egyszerűen értelmezhetők maradjanak.

A leggyakrabban alkalmazott kérdéstípusok a következők voltak:

- Igen-nem típusú kérdések, amelyek elsősorban a biztonsági megoldások meglétére vagy hiányára, illetve bizonyos gyakorlatok alkalmazására vonatkoztak;
- Többszörös választásos kérdések, ahol a válaszadók előre definiált válaszlehetőségek közül választhattak (pl.: használt védelmi technológiák, végzettségi kategóriák);
- Kategorizáló kérdések, amelyek a vállalatokat vagy döntéshozókat előre meghatározott csoportokba sorolták.

A kérdések egy része ellenőrző jellegű volt, vagyis korábbi válaszokra visszakérdezve segítette a bizonytalan vagy ellentmondásos válaszok kiszűrését.

Az adatelemzés célja és a vizsgált összefüggések

A kérdőíves felmérés feldolgozása során a cél nem egyetlen változó önálló értékelése volt, hanem összefüggések feltárása a vállalati jellemzők, a döntéshozói kompetenciák és az informatikai biztonsági kitettség között.

Az elemzés során az alábbi típusú információk kinyerésére törekedtem:

- a biztonsági incidensek előfordulásának gyakorisága;
- a védelmi megoldások alkalmazásának szintje;
- a döntéshozók informatikai képzettsége és a védelmi megoldások közötti kapcsolat;
- a munkavégzési szokások és a kockázati szint összefüggései;

- a vállalatok nyitottsága új, automatizált védelmi megoldások bevezetésére.

A módszertani megközelítés célja olyan korrelációs és csoportos összehasonlítási alapok megteremtése volt, amelyek alátámasztják a disszertáció későbbi fejezeteiben bemutatott védelmi modell szükségességét és gyakorlati relevanciáját.

2.3. A gépi tanulási modellek kiválasztásának módszertana

A szerveroldali fájlműveleti naplók alapján történő zsarolóvírus-detektálás sajátos osztályozási feladatot jelent, amelynek célja a normál felhasználói aktivitás és a rosszindulatú titkosítási folyamatok elkülönítése. A kutatási probléma jellemzői alapvetően meghatározzák az alkalmazható gépi tanulási modellekkel szemben támasztott követelményeket, ezért a modellek kiválasztásakor a teljesítmény mellett a vizsgált alkalmazási környezet sajátosságai is meghatározóak voltak. A kutatás célja nem egy adott gépi tanulási algoritmus fejlesztése volt, hanem annak vizsgálata, hogy a szerveroldalon megfigyelhető fájlműveleti események alkalmasak-e a zsarolóvírus-tevékenység korai felismerésére. Ennek megfelelően a módszertani megközelítés középpontjában az adatrepresentáció, a modellek összehasonlíthatósága és az eredmények objektív értékelése állt. A vizsgált probléma jellemzőiből következően olyan felügyelt tanulási módszerek alkalmazása volt indokolt, amelyek strukturált adatokon hatékonyan működnek, képesek bináris osztályozási feladatok megoldására, valamint alkalmasak közel valós idejű osztályozásra. Fontos további módszertani követelmény volt, hogy a modellek számítási igénye, tanítási adatigénye és döntési késleltetése összhangban maradjon a KKV-környezetben történő gyakorlati alkalmazhatóság követelményeivel. Ennek megfelelően a vizsgálat klasszikus gépi tanulási algoritmusokra (Decision Tree, Support Vector Machine, Random Forest), valamint ensemble megközelítésekre (AdaBoost, XGBoost) terjedt ki. A neurális hálózatokon alapuló megközelítések szakirodalmi áttekintése ugyan megtörtént, azonban a rendelkezésre álló strukturált bemeneti adatok, a korlátozott tanítási adatmennyiség, valamint a közel valós idejű működés és az alacsony erőforrásigény követelményei miatt a kutatás a klasszikus és ensemble gépi tanulási modellekre fókuszált. A tanulási folyamatot megelőzően a szerveroldali fájlműveleti eseményeket előre definiált időablakokban aggregált jellemzőkké alakítottam és előszűrtem. A jellemzők kialakításának célja nem a bemeneti dimenzió maximalizálása volt, hanem olyan informatív reprezentáció létrehozása, amely alkalmas a normál működés és a támadási aktivitás közötti különbségek megragadására. Ez egyben hozzájárult a modellek általánosíthatóságának javításához és a túltanulás kockázatának csökkentéséhez is. A modellek összehasonlítása egy-egy tanítási és értékelési eljárás szerint történt. Ennek megfelelően minden modell ugyanazon adatahalmazon, azonos előfeldolgozási lépések alkalmazásával került értékelésre. A teljesítményértékelés során alkalmazott metrikák kiválasztását a kutatási probléma jellege indokolta. A zsarolóvírus-detektálásban nem elegendő kizárólag az általános osztályozási pontosság vizsgálata, mivel a téves negatív döntések jelentős biztonsági kockázatot hordoznak, míg a túlzott számú téves riasztás a rendszer gyakorlati alkalmazhatóságát csökkenti. Ennek

megfelelően a modellek értékelése több, egymást kiegészítő teljesítménymutató alapján történt annak érdekében, hogy a felismerési képesség, a téves riasztások aránya és a gyakorlati alkalmazhatóság egyaránt objektíven összehasonlítható legyen. Az így kialakított módszertani keretrendszer biztosította, hogy a különböző gépi tanulási modellek összehasonlítása egységes feltételek mellett történjen, és lehetővé tette annak vizsgálatát, hogy a szerveroldali fájlműveleti naplók milyen mértékben alkalmasak a zsarolóvírus-támadások közel valós idejű felismerésére.

3. Szakirodalmi áttekintés

A szakirodalmi feldolgozás célja a zsarolóvírus-támadások felismerésével, megelőzésével és hatáscsökkentésével kapcsolatos, mesterséges intelligenciát alkalmazó megközelítések átfogó, szisztematikus feltérképezése volt, különös tekintettel a fájlkiszolgáló- és naplóalapú detekciós módszerekre, valamint a kis- és középvállalkozások működési környezetében alkalmazható megoldásokra. A szakirodalmi áttekintést tematikusan három különálló részre osztottam. Ebben a fejezetben ismertetem az általános módszertani megközelítéseket. Az 5.1 fejezet tartalmazza az adatvagyon védelmét és visszaállíthatóságát biztosító kísérleti megoldások áttekintését, míg a 6.1-es fejezetben kifejezetten a zsarolóvírusok észlelésére, a honeypot-alapú megoldásokra, illetve a mesterséges intelligencia segítségével történő detektálásra helyezem a hangsúlyt. A RAID architektúrák elsődlegesen a hardverhibákból eredő adatvesztés ellen nyújtanak védelmet, azonban a zsarolóvírusok által végrehajtott titkosítás a fájlrendszer szintjén legitim írási műveletként jelenik meg, ezért a módosítások a RAID tömb minden lemezére replikálódnak. Hasonló korlát érvényes a meghajtó- és fájlrendszer-szintű titkosítási megoldásokra (pl. BitLocker, eCryptFS, LUKS, ZFS) is. Ezek elsősorban az offline jogosulatlan hozzáférés ellen védenek, de egy már felcsatolt fájlrendszer esetén nem akadályozzák meg a zsarolóvírusok adat-titkosítását [109]. A szakirodalom ezért továbbra is alapvető védekezési stratégiának tekinti a 3-2-1 mentési szabály alkalmazását, a pillanatfelvételek (snapshotok) rendszeres készítését, valamint a legkisebb jogosultság elvének érvényesítését [98]. A modern védekezési megközelítések egyre nagyobb hangsúlyt helyeznek a naplóalapú eseményfeldolgozásra és a viselkedésalapú detektálásra. A szerveroldali naplóállományokból felépíthető referencia-viselkedési profilok lehetővé teszik az anomáliák automatikus felismerését, ugyanakkor a nagy mennyiségű naplóadat közel valós idejű feldolgozása jelentős számítási kihívást jelenthet. Számos kutatás igazolta, hogy a gépi tanulási és mélytanulási módszerek ezen a területen hatékonyabbak lehetnek a hagyományos statisztikai megközelítéseknél [17], illetve alkalmasak a zsarolóvírusok és más rosszindulatú programok elkülönítésére is [18]. A szakirodalom emellett kiemeli a C2 infrastruktúrák felismerésének [18], a kliensoldali fertőzések megelőzésének [19], valamint a különböző ransomware-családok viselkedési sajátosságainak jelentőségét [20]. Mivel a világjárvány időszaka tovább növelte a célzott kibertámadások számát (különösen az egészségügyi szek-

torban [21][22]), erre válaszként a zsarolóvírusok elleni védekezés területén számos új kutatási irány körvonalazódott. Ide sorolható többek között a Windows-alapú ransomware-ek viselkedéselemzése [23], a rendszerfájl-felügyeleten alapuló detektálás [24], az API-hívások elemzése [25], az új ransomware-taxonómiák kialakítása [26], a hálózati kommunikáció és reverse engineering alapú felismerés [27], a WannaCry-specifikus védelmi megoldások [28], az IoT- és felhőkörnyezetek biztonsági kérdései [29][30], valamint az IDS/IPS rendszerek és honeypot-alapú megközelítések [31]–[34] fejlesztése. A feldolgozott szakirodalmak egybehangzóan arra a következtetésre jutnak, hogy a hagyományos szignatúra-alapú védelem önmagában már nem elegendő, ezért a korszerű megoldások a viselkedés-elemzésre, a gépi tanulásra és a több rétegű védelemre épülnek. A korszerű végpontvédelmi rendszerek ennek megfelelően a klasszikus vírusvédelmi funkciókat viselkedésalapú ransomware-védelemmel, sandboxinggal, URL-szűréssel, fejlett tűzfallal, alkalmazás- és eszközfelügyelettel, identitásvédelemmel, valamint automatikus rend-szer-helyreállítási és mentési funkciókkal egészítik ki [26]. Ugyanakkor megállapítható, hogy a kliensoldali védelem továbbfejlesztése önmagában nem elegendő. A hatékony védekezési startégia megköveteli a hálózatban történő oldalirányú (laterális) mozgás figyelését, a központi fájlkiszolgálók napló alapú monitorozását és a megfelelő mentési stratégia alkalmazását is. A feldolgozott szakirodalom alapján megállapítható, hogy bár a korszerű védelmi megoldások egyre nagyobb hangsúlyt helyeznek a viselkedésalapú felismerésre és a gépi tanulási módszerek alkalmazására, a központi fájlkiszolgálók szerveroldali fájlműveleteinek elemzésére épülő, alacsony erőforrás-igényű és közel valós idejű detektálási megközelítések jellemzően nem kerülnek fókuszba. Ez alátámasztja jelen kutatás relevanciáját és a javasolt ARDS-modell kialakításának szükségességét..

4. Hazai kis- és középvállalkozások (KKV-k) informatikai biztonságra vonatkozó felkészültségének felmérése kérdőíves kutatás alapján, az ARDS védelmi modell bemutatása

A digitálisan végezhető tevékenységek köre folyamatosan bővül, így egyre több információ és ezáltal digitálisan tárolt adat kerül a digitális térbe. Ezek az adatvagyonok egyrészt értéket képviselnek, másrészt folyamatos védelmet és biztonságos tárolást igényelnek. Világszerte komoly problémát jelent az információs rendszerek megfelelő védelme. A gyors ütemű digitalizáció elkerülhetetlenül a különböző érdekeltsgű szereplők elvárásainak és igényeinek változásához vezetett. Egy tipikus KKV-n belül az érdekelt felek elkülönítése például a következőképpen lehetséges: végfelhasználók, tulajdonos, IT-biztonsági szakemberek, rendszergazdák, külső szakértők. Amint az az 5. táblázatban látható, az informatikai rendszerekkel kapcsolatosan a különböző érdekelt felek céljai és preferenciái eltérőek. A táblázatban szereplő eltérő prioritások következtében az érdekelt felek érdekei folyamatosan ütköznek, de biztonsági szempontból olyan egyensúlyt kell találni, amely minden résztvevő számára elfogadható.

4. táblázat Az érdekelt felek eltérő prioritásai

Végfelhasználói célok	Könnyen használható rendszer
	Kevésbé bonyolult hitelesítés
	Könnyen megjegyezhető jelszavak
Tulajdonosi célok	Olcsó beszerzés
	Minimális erőforrásigény
	Üzemeltetés szakértő nélkül (ha lehetséges)
Rendszergazdák céljai	Minimális változás a működésben
	Automatizált folyamatok
	Minimális időigény, könnyű adminisztráció
	Kevés új ismeretre legyen szükség
IT biztonsági szakemberek céljai	Alapvető IT-ismeretekkel rendelkezzenek a végfelhasználók
	Biztonságtudatos végfelhasználók
	Magas szintű ismeretekkel rendelkező rendszergazdák
A kiberbűnözők céljai	Könnyű hozzáférés a zárt rendszerekhez („low hanging fruits”)
	Gyenge jelszavak, gyenge autentikációs mechanizmusok
	"Lusta" rendszergazdák
	Hiányos, vagy hiányzó biztonsági megoldások

Ahhoz, hogy megközelítő képet kaphassak a hazai KKV szektor informatikai biztonsággal kapcsolatos helyzetéről, illetve az igényeikről, kérdőíves kutatás keretében kerestem a választ olyan kérdésekre, mint például, hogy milyen gyakran fordulnak elő vírustámadások ebben a vállalkozói rétegben, mennyire sikeresek ezek a támadások, milyen összefüggések fedezhetők fel a vezetők informatikai képzettsége és az informatikai biztonság között, illetve, hogy ezek a gazdasági szereplők a jelenlegi helyzetben hajlandóak lennének-e egy új, komplex, a felmerülő igények alapján kialakított és modellezett védelmi rendszer bevezetésére. A felmérés eredményei alapján indokolt volt egy olyan költséghatékony védelmi modell megtervezése, amely képes együttműködni a jelenleg elérhető hálózati és

vírusvédelmi megoldásokkal, lehetővé teszi a zsarolóvírus-támadások korai felismerését, hatékonyan védi a vállalati adatvagyon, valamint minden esetben lehetővé teszi az adatvagyon helyreállítását incidensek esetén.

Mivel a KKV-k méretük és korlátozott erőforrásaik miatt gyakran egyedi kihívásokkal szembesülnek, fő célom a kutatás során az volt, hogy általános képet kapjak a kis- és középvállalkozások informatikai biztonság tudatosságának helyzetéről. Az áttekintés fő fókuszpontjai a következők voltak:

- a döntéshozó informatikai végzettségének szerepe informatikai biztonsági kérdésekben;
- a használt IT biztonsági megoldások;
- a munkavállalók biztonság tudatos viselkedésének szerepe;
- az online feladatokkal kapcsolatos előírások;
- az előírások és a zsarolóvírus-támadások közötti kapcsolat;
- milyen gyakran fordulnak elő ilyen támadások ebben a vállalkozói rétegben, és ezek mennyire sikeresek;
- milyen összefüggések fedezhetők fel a vállalati informatikai kultúra és az informatikai biztonság elemei között;
- a gazdasági szereplők a jelenlegi helyzetben hajlandók-e egy új, komplex, önálló, a felmerülő igények alapján kialakított és modellezett védelmi rendszer bevezetésére.

A kérdésekre egy interjún alapuló felmérés alapján adtak választ a résztvevők. A kérdőíves felmérés 2021/2022-ben, személyes interjúk formájában zajlott. Az interjúalanyok kiválasztásának egyik szempontja a vállalkozás mérete volt. A felmérés kifejezetten a kis- és középvállalkozásokra fókuszált, mivel általánosságban megállapítható, hogy a vállalkozások többsége (közel 85%-a) Magyarországon és az Európai Unióban jelenleg ebbe a kategóriába tartozik. Ez azt is jelenti, hogy a megkérdezett cégek száma, bár nem jelentős, de arányosan reprezentatív a szektorra nézve. A kérdőív összeállítása során igyekeztem elemezni és feltérképezni a vállalatok informatikai rendszereit és biztonsági helyzetét, ugyanakkor felmérni a korábbi, rosszindulatú szoftverekkel kapcsolatos tapasztalataikat. A következő alfejezetekben részletezett kérdésekkel összességében arra kerestem a választ, hogy a megkérdezett vállalatok jelenleg milyen felkészültségi szintet képviselnek a lehetséges informatikai biztonsági incidensekre vonatkozóan, illetve milyen biztonsági gyakorlatokat alkalmaznak az ilyen jellegű fenyegetésekkel szemben. Az interjúk során megvizsgáltam továbbá például az üzleti tevékenységek nagyságrendjét, az informatikai infrastruktúrát és a döntéshozók informatikai kompetenciaszintjét. Foglalkoztam továbbá a nemzetközi üzleti kapcsolatokkal, a belső IT-folyamatokkal, a munkavállalók képzésével és a zsarolóvírus-támadásokkal kapcsolatos konkrét tapasztalatokkal is. Ezzel a megközelítéssel igyekeztem betekintést nyerni abba, hogy a KKV-k hogyan kezelik az informatikai biztonsággal kapcsolatos kockázatokat és milyen tényezők befolyásolják a kiberfenyegetettségük mértékét. A felmérés szándékaim szerint feltárja továbbá a jelenlegi gyakorlatok hatékonyságát, és rávilágít azokra a területekre, amelyek fejlesztésre szorulnak az informatikai biztonsági incidensekkel kapcsolatos ellenállóképesség javítása érdekében.

A tanulmány elsődleges célja az volt, hogy megállapítsam, a vállalatok mennyire fordítanak figyelmet az adataik védelmére, milyen belső szabályozásokkal segítik a védelmet, és hogy a jelenlegi világjárvány hogyan befolyásolta a távmunkára és az otthoni munkavégzésre való nyitottságukat. Emellett megvizsgáltam a vállalatok személyi összetételét is, az informatikai feladatra való alkalmasság, az informatikai végzettség, a képzettségi szint, valamint a vállalatokon belül az informatikai kérdésekben döntést hozó személy kompetenciái szempontjából. Egy külön kérdéscsoport az interneten keresztül elvégezhető napi adminisztratív feladatok, például az internetes banki ügyintézés és az online számlázás feltérképezésére irányult. Kiemelten fontos volt az is, hogy információkat kapjak a vállalati informatikai infrastruktúrákat ért korábbi vírus- és kifejezetten zsarolóvírus-támadások tapasztalatairól, illetve az incidensek kezelésének gyakorlatairól. Ezekkel kapcsolatosan a bevezető kérdés az volt, hogy az elmúlt egy évben történt-e olyan vírustámadás a vállalati hálózaton (akár szerver, akár munkaállomás oldalon), amelynek észrevehető jelei és következményei voltak. Az ezt követő kérdések segítségével megvizsgálhattam az incidenssel érintett vállalkozások támadásokra adott reakcióit, a helyreállítás sikerességét, az üzletmenet-folytonosságot érintő időtartományt, a teljes helyreállítás időtartamát, a támadók által követelt váltságdíj mértékét és a váltságdíjfizetési hajlandóságot. Ugyancsak fontos információhoz jutottam azáltal, hogy megtudtam, a vállalatok milyen összeget fordítottak a támadások utáni változtatásokra az informatikai infrastruktúra biztonságának javítása érdekében.

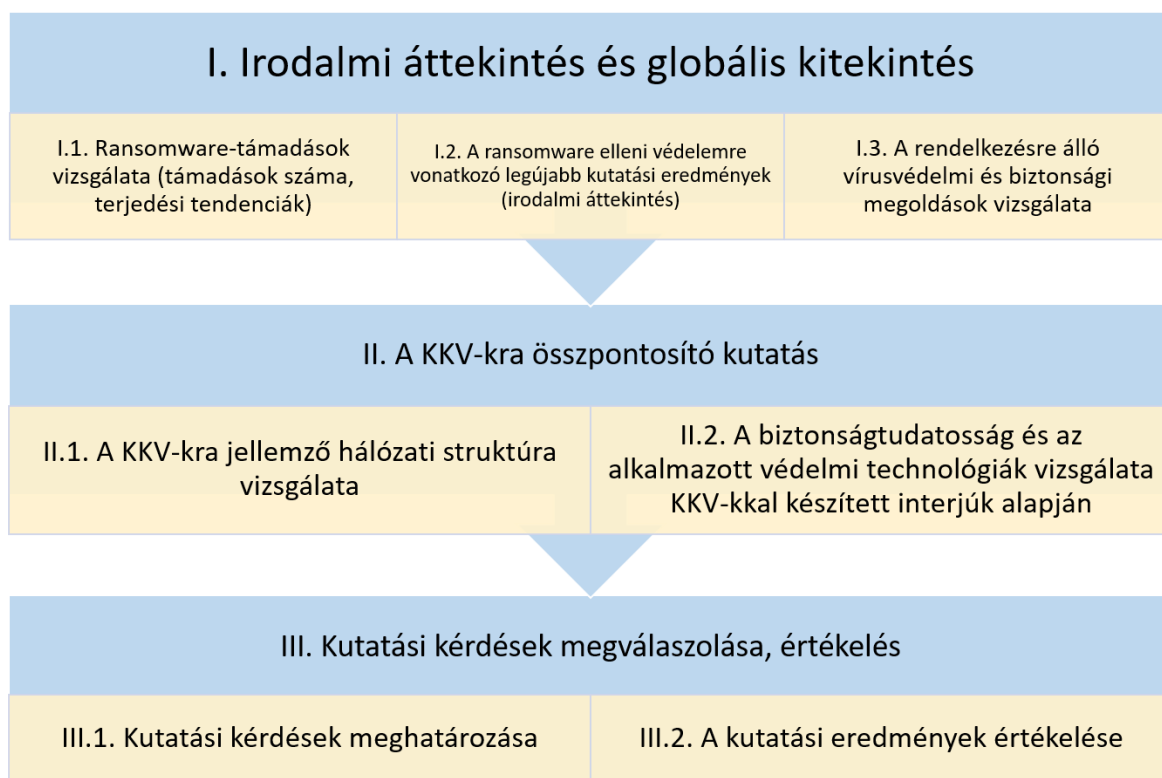
Kutatási kérdések:

- K-1: Van-e szignifikáns összefüggés a vállalatok informatikai döntéshozóinak képzettségi szintje és a hálózatban jelen lévő informatikai biztonsági megoldások között?
- K-2: Minden vállalat rendelkezik-e telepített antivírus megoldással a munkaállomásain és a kiszolgálóin, és ha igen, akkor ez egy ingyenes, korlátozott funkciókkal rendelkező változat, vagy kereskedelmi forgalomban kapható, teljes körű védelmet nyújtó változatot használtak?
- K-3: Milyen szerepet játszik az informatikai döntéshozó a vírusvédelmi megoldás kiválasztásában és ez milyen kapcsolatban áll az informatikai területen megszerzett szakmai képzettségükkel?
- K-4: Milyen hajlandóságot mutatnak a mikro-, kis- és középvállalkozások arra, hogy egy új, költséghatékony, komplex védelmi rendszert bevezessenek?

4.1. A kutatás menete, a vizsgálati módszer

A kutatás első szakasza a médiában megjelent zsarolóvírus-támadásokkal kapcsolatos aktuális hírekre és a szakértői statisztikákra összpontosított. A következő szakaszban számba

vettem a kereskedelmi termékeknél alkalmazott zsarolóvírus-detektálás és védekezés módszereit. Ezt követően kérdőíves interjúk segítségével mértem fel a kutatásban résztvevő vállalkozások hálózati informatikai infrastruktúráját, a biztonságtudatosság szintjét, illetve az alkalmazott védelmi megoldásokat. A tanulmány következő fázisában a zsarolóvírus-támadással kapcsolatos interjúalanyok és válaszadók körében felmérést végeztem azzal kapcsolatosan, hogy milyen károk keletkeztek a támadás után, és milyen lépéseket tettek egy esetleges következő támadás megelőzése érdekében. A kérdőívek feldolgozását követően összegeztem a tapasztalatokat. A teljes kutatási folyamatot a 7. ábra szemlélteti. A kutatás eredményei alapján egy olyan hálózatvédelmi rendszermodell került kidolgozásra, amelyben az azonosított hozzáférési pontokat, támadási vektorokat, a hálózati eszközök forgalmát, illetve bizonyos kritikus szolgáltatásokat folyamatosan monitorozhatunk. A javasolt koncepció egy hozzáadott védelmi réteggént jelenik meg a hálózaton, ahol a beavatkozás lehetőségét pedig egy ún. MITM-ügynök biztosítja.



7. ábra A kutatási folyamat vizualizációja [forrás: saját]

A felmérés során irányított, személyes interjú technika került alkalmazásra. Az interjú során minden interjúalany 86 kérdésre válaszolt, amelyeket közvetlen és közvetett kérdések formájában kaptak.

4.1.1. A kutatás korlátai

Mivel az adatfelvétel időpontja (COVID-19 időszak) jelentős hatással lehetett a válaszokra és azóta az EU szabályozási, technológiai és fenyegetettségi környezete is jelentősen megváltozott, ezért fontos hangsúlyozni az eredmények korlátozott jelenlegi relevanciáját. Az említett időszakban a vállalkozások informatikai környezete jelentős átalakuláson ment keresztül, egyfajta kényszerű digitalizáció zajlott. Emiatt a kapott eredmények értelmezése során figyelembe kell venni, hogy a válaszok nem kizárólag a vállalatok hosszú távú információbiztonsági gyakorlatát tükrözhetik, hanem részben a pandémia által előidézett rendkívüli működési környezet hatásait is. A járványhelyzet következtében a kis- és középvállalkozások jelentős része rövid idő alatt vezette be vagy bővítette a távoli munkavégzést, az online ügyintézkedést és a felhőalapú szolgáltatások használatát. Az új működési modell számos olyan sérülékenységet eredményezett, amelyek korábban nem, vagy csak kisebb mértékben jelentek meg. Megnőtt a dolgozók saját eszközeinek használati aránya, a VPN-hozzáférések, illetve a távoli adminisztrációs eszközök száma. Ezzel párhuzamosan a kibertámadások, különösen a zsarolóvírusok, az adathalászat és az üzleti e-mail kompromittálás (BEC) elleni támadások is jelentősen megsaporodtak. Az European Union Agency for Cybersecurity több fenyegetettségi jelentése szerint a pandémia éveiben a zsarolóvírusok váltak az európai szervezetek egyik legjelentősebb kiberkockázatává. Mindezek következtében a kutatás több vizsgált változóját is biztosan befolyásolhatta a járványhelyzet. Elsősorban a K-2 és K-4 kutatási kérdések esetében feltételezhető torzító hatás. A vírusvédelmi megoldások alkalmazása, valamint egy új védelmi rendszer bevezetésére mutatott hajlandóság valószínűleg magasabb volt, mint egy stabilabb gazdasági időszakban, mivel a vállalatok közvetlenül tapasztalták a megnövekedett kockázatokat. A K-1 és K-3 kutatási kérdések esetében a döntéshozók szerepe szintén felértékelődhetett. Számos KKV-ban korábban nem jelentett stratégiai kérdést az informatikai biztonság, azonban a távmunkára történő gyors átalás során az IT-felelősök vagy a vezetők rövid idő alatt kénytelenek voltak biztonsági döntéseket meghozni. Ez átmenetileg erősebb kapcsolatot eredményezhetett az informatikai kompetencia és a bevezetett biztonsági megoldások között. A vírus- és zsarolóvírus-támadások gyakoriságára vonatkozó válaszok szintén nem tekinthetők időfüggetlennek. A pandémia alatt az európai incidensstatisztikák jelentős növekedést mutattak [110, 111, 112], így a felmérésben szereplő incidensarányok várhatóan magasabbak lehettek, mint egy korábbi vagy egy későbbi időszakban. A kutatás elkészítése óta az európai információbiztonsági környezet több szempontból is átalakult. A szabályozási környezet jelentősen szigorodott, mivel 2023/2024 során életbe lépett és végrehajtás alá került a NIS2 irányelv, ami nagyságrendekkel több szervezet számára ír elő kötelező kiberbiztonsági intézkedéseket, kockázatkezelést és kötelező incidensbejelentést. Emellett elfogadták a Cyber Resilience Act rendeletet is, amely a digitális termékek biztonságát szabályozza, tovább növelve a beszállítói láncok biztonsági követelményeit. Ennek következtében ma valószínűleg magasabb lenne azoknak a KKV-knak az aránya, amelyek valamilyen formális információbiztonsági intézkedést alkalmaznak. Azóta a zsarolóvírus fenyegetettség sokkal összetettebb jelenség lett és lényegesen magasabb reputációs és likviditási kockázatot hordoz, mint korábban. Ez főleg annak köszönhető, hogy a klasszikus adattitkosítást

alkalmazó vírusok mellett mára általánossá stratégiává vált az egyszerű adatlopás, a kettős zsarolás (nyilvánosságra hozás, titkosított adatok visszaállítása), a hármás zsarolás (ügyfelek, partnerek bevonása), valamint az ellátási láncokon keresztüli támadás. Ez azt jelenti, hogy a jelenlegi környezetben készülő felmérésben már nem lenne elegendő kizárólag a titkosító vírusok vizsgálata, hanem az adatlopási incidensekre is mindenképpen ki kellene térni. Sok helyen a hibrid munkavégzés tartósan megmaradt és bár a COVID-időszak lezárult, a hibrid munkavégzés számos ágazatban állandóvá vált. Ez azt jelenti, hogy a végpontvédelem, a többtényezős hitelesítés, a felhőbiztonság, valamint a felhasználói tudatosság ma még fontosabb szerepet játszik, mint 2021-ben. Fontos megemlíteni, hogy 2022 végétől a generatív mesterséges intelligencia fokozatos elterjedése új kihívásokat eredményezett. Ma már egy hasonló kutatásban cél-szerű lenne vizsgálni például az AI használati arányát, az AI által támogatott adathalász táma-dásokat, a lehetséges AI-szabályozást, valamint a munkavállalók AI-biztonsági ismereteit is. Szintén lényeges változás, hogy az elmúlt néhány évben az EU számos tudatosságnövelő prog-ramot indított, miközben a biztosítók, a beszállítók és az ügyfelek is egyre gyakrabban várnak el dokumentált kiberbiztonsági intézkedéseket. Emiatt ma valószínűleg magasabb arányban je-lennének meg a következők:

- rendszeres biztonsági mentés (pl.: felhőalapú megoldások, „immutable backups”);
- többfaktoros hitelesítés (TFA, MFA);
- végpontvédelmi rendszerek (pl.: EDR, esetleg MDR szolgáltatással kibővítve);
- biztonságtudatossági képzések;
- incidenskezelési eljárások.

4.2. A kérdőív felépítése

A felmérés során használt kérdőív több kérdéscsoport köré épült. E kérdések egy része a vállalat alapadataira (tevékenység, telephely, méret), más részük a szervezet informatikai infrastruktúrájára (szerverek, hálózat, végpontok száma, használt operációs rendszerek, fel-hőszolgáltatások, hálózatbiztonsági megoldások), külön kérdéscsoport a jelenlegi biztonsági felkészültségre (fizikai védelem, informatikai biztonsági irányelvek, mentési megoldások), illetve az alkalmazott informatikai biztonsági irányelvekre vonatkozott. Kiemelt kérdéscso-port fókuszált az eddigi biztonsági incidensekre, illetve azok következményeire. Az interjúk alapjául szolgáló kérdőív összesen 86 kérdést tartalmazott, amelyek közül az első kérdések a vállalatok tevékenységi köreire, az alkalmazottak számára és az informatikai döntéshozók informatikai ismereteire fókuszáltak. Egy másik kérdéscsoport a nemzetközi üzleti kapcsola-tokra összpontosított. A következő kérdéscsoport a vállalat informatikai infrastruktúrájá-nak felépítésére vonatkozott (pl.: hány számítógépet használnak, van-e belső hálózatuk, üze-meltetnek-e szervereket, használ-e a vállalat megosztott erőforrásokat (pl.: fájlserver, háló-zati nyomtató, szkennerek stb.)). További kérdések kerültek a kérdőívbe az informatikai házi-

rendekkel, az alkalmazottak belső képzéseivel, az informatikai biztonságtudatossággal kapcsolatosan. Az interjúk zárásaként a múltban elszenvedett malware-támadásokkal, azon belül is különösen a zsarolóvírus-támadásokkal, valamint az informatikai rendszerek aktív védelmével kapcsolatosan érdeklődtem. A kérdések között olyan ellenőrző kérdések is szerepeltek, amelyek már a korábban adott válaszokra vonatkoztak, hogy ki tudjuk küszöbölni a bizonytalan válaszok okozta torzítást az eredményekben. Az egymást követő kérdéseknél több esetben is szándékos témaváltások történtek. Ezzel el szerettem volna kerülni azt a jelenséget, hogy az interjúalanyok a válaszadás során a legpozitívabb képet fessék magukról, ahelyett, hogy valós információkat adnának meg.

4.3. A tanulmány célcsoportja

A vizsgálatok elsődleges célcsoportja a mikro-, kis- és középvállalkozások voltak, ahol az alkalmazottak száma nem haladja meg a 250 főt. A vizsgált vállalatok között 2-250 főt foglalkoztató vállalatok szerepeltek, ami a kis- és középvállalkozásokra jellemző. A hasonló létszámú vállalkozások száma ma Magyarországon (a legfeljebb kilenc főt foglalkoztató mikrovállalkozások alkotják a magyarországi KKV-k többségét) mintegy 549 ezer, az Európai Unióban pedig a kis- és középvállalkozások (KKV-k) becsült száma 2018-ban mintegy 25,1 millió volt [35, 36]. Számukat tekintve jelenleg ez a vállalati kategória van jelen legnagyobb számban a magyar gazdasági szférában, ami egyben azt is jelenti, hogy bár éves bevételük nem közelíti meg a nagyvállalatokét, az éves GDP-termelés jelentős részét képezik. A vizsgálatba 46 kis- és középvállalat vezetőit, tulajdonosait vagy informatikai vezetőit vontam be. A minta alacsony elemszáma miatt a vizsgálat nem nevezhető reprezentatívnak, de betekintést adhat a nem informatikai biztonsággal kapcsolatos profillal rendelkező kis- és középvállalkozások helyzetébe. A felmérésben résztvevő vállalatok különböző ágazatokban tevékenykedtek (pl.: agrár-élelmiszeripari termelés, ipari termelés, egészségügyi szolgáltatások, vendéglátás, szépségipari szolgáltatások). A kis- és középvállalkozások kiválasztásának kritériumai a következők voltak:

- a vizsgálatban érintett vállalat alkalmazottainak száma nem haladja meg a 250 főt;
- a vállalatok székhelyei, illetve telephelyei nem a fővárosban vannak;
- a vállalat működési ideje legalább 2 év;
- a vállalat hajlandóságot mutat a névtelen, személyes interjún való részvételre.

4.4. Az eredmények összefoglalása, releváns kutatási kérdések megválaszolása

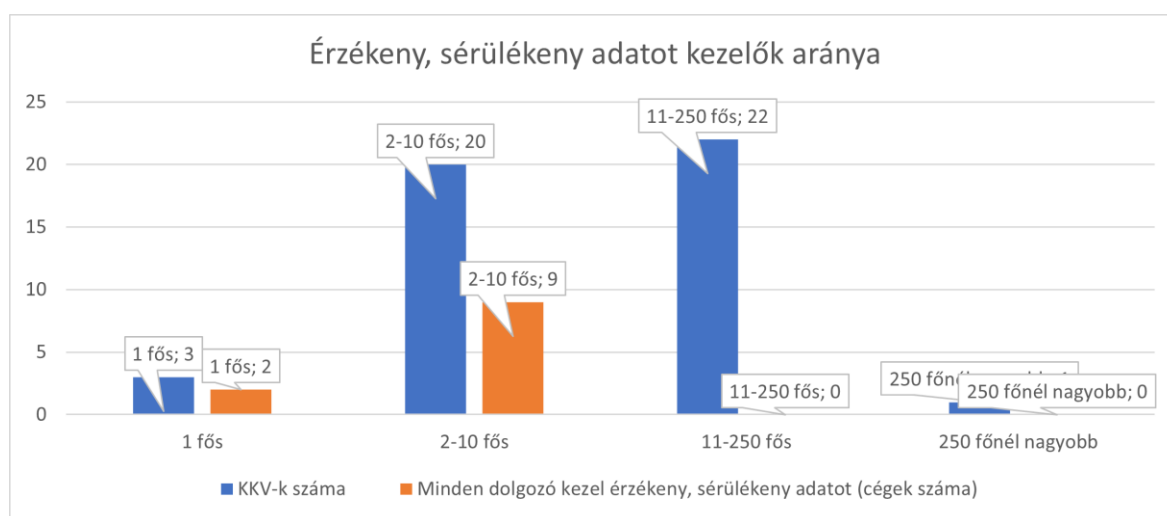
Az interjúk eredményei alapján általánosan megállapítható, hogy a KKV-k esetében vannak olyan sajátosságok, amelyek megnehezítik az informatikai rendszerek üzemeltetését és védelmét. A vizsgált minta alapján megállapítható volt, hogy a felmérésben résztvevő vállalkozások nem az informatika területén működnek, hanem a kereskedelmi, a termelés és a

szolgáltatási ágazatok széles skáláját képviselik. Ezek a kiskereskedelem, ipar, szolgáltatás, állattenyésztés, növénytermesztés, élelmiszertermelés, magán- és közegészségügyi szolgáltatások, biztonsági berendezések telepítése, fodrászat, kozmetika, tetováló stúdió, manikűrös, állateledel-kereskedelem, oktatás, közigazgatási szolgáltatások.

A kis- és középvállalkozások kivétel nélkül olyan szolgáltatásokat vesznek igénybe a napi adminisztratív és ügyviteli feladataik ellátásához, amelyek csak online internetkapcsolaton keresztül végezhetők el. Ilyen szolgáltatások például az elektronikus banki szolgáltatások, az adatszolgáltatási kötelezettségek elektronikus formában történő teljesítése, az online számlázás és az online vásárlás.

A hálózati eszközök kiválasztásánál fontos tényező volt az alacsony beszerzési ár és a szakértelmet nem igénylő telepítés lehetősége, éppen ezért hálózataik általában nem rendelkeznek nagyvállalati környezetben megszokott szolgáltatásplatformmal, illetve fejlett védelmi képességekkel.

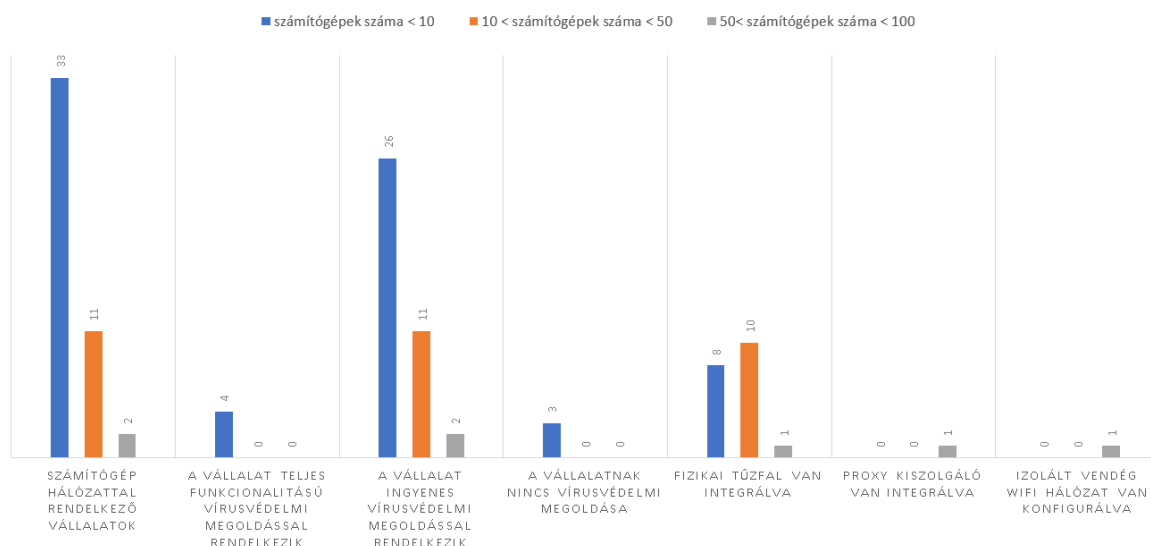
Általános kérdések vonatkoztak a vállalatok alkalmazotti létszámára, illetve a személyes adatokkal dolgozó személyek számára. A 8. ábra a vizsgált vállalatoknál foglalkoztatottak, illetve a személyes adatokkal (pl.: a munkavállalók személyes adatai, szerződések, béradatok, pénzügyek, céges bankszámlák) dolgozó alkalmazottak számát szemlélteti. A gyors osztályozás alkalmazásával a vállalatokat négy kategóriába soroltam az alkalmazottak száma alapján.



8. ábra Érzékeny, sérülékeny adatot kezelők aránya a cég méretéhez viszonyítva [forrás: saját]

Lényeges volt olyan vállalatspecifikus adatokat is rögzíteni a felmérés során, mint például a szerverek és munkaállomások száma (későbbiekben "számítógépek száma"), a szoftveres és hardveres védelmi eszközök megléte az informatikai infrastruktúrában, valamint az informatikai döntéshozó szakmai ismereteinek foka (9. ábra).

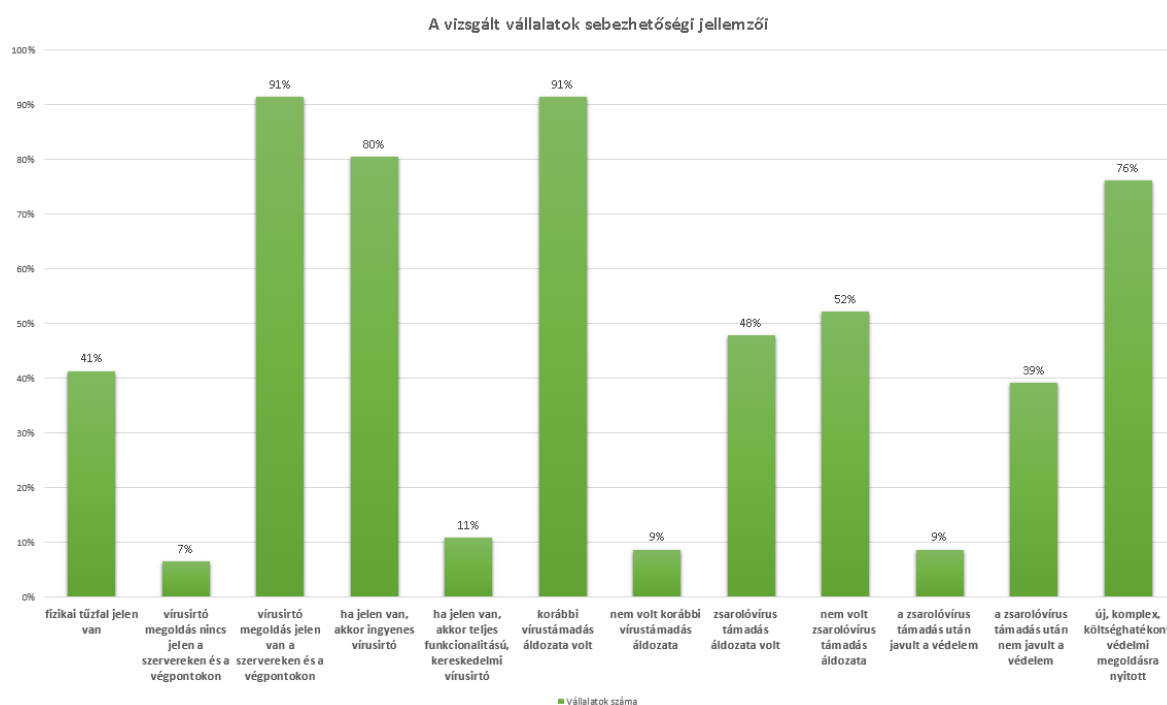
INFORMATIKAI INFRASTRUKTÚRÁK JELLEMZŐI (SZÁMÍTÓGÉPEK SZÁMA- RELEVÁNS SZOLGÁLTATÁSOK)



9. ábra Informatikai infrastruktúrák jellemzői a vizsgált vállalatoknál [forrás: saját]

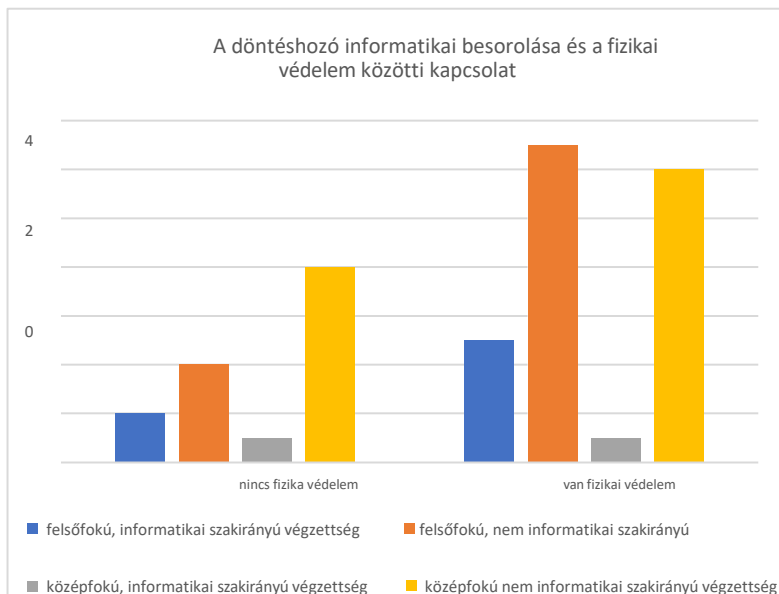
Ezen kutatási szakasz első kutatási kérdése (K-1) a következő volt: van-e szignifikáns összefüggés a vállalatok informatikai döntéshozóinak képzettségi szintje és a hálózatban jelen lévő informatikai biztonsági megoldások között?

Mindemellett a kutatás során választ kerestem arra a kérdésre is, hogy a döntéshozó informatikai ismereteinek volt-e szerepe az informatikai biztonsággal kapcsolatos kérdések meghozatalában, illetve van-e kapcsolat az ismeretek és az elszenvedett vírustámadások között. Ehhez elsőként a vállalatok sérülékenységi karakterisztikáját kellett megvizsgálni, melyet a 10. ábra összegez.

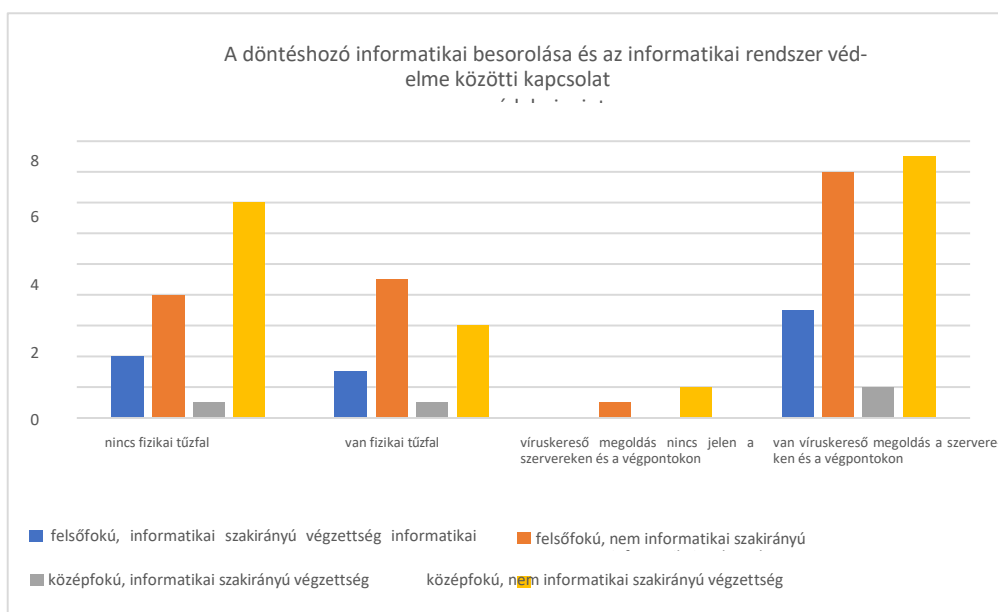


10. ábra A vizsgált vállalatok sebezhetőségi jellemzői [forrás: saját]

A megkérdezett KKV-k egyike sem rendelkezett olyan döntéshozóval, aki kizárólag az informatikai rendszerrel kapcsolatos kérdésekben való döntéshozatalért felelt volna. Ezt a szerepet kivétel nélkül a megkérdezett vállalkozások tulajdonosa töltötte be, függetlenül attól, hogy rendelkezett-e informatikai képesítéssel, vagy sem. A kapott válaszokból megállapítható volt, hogy a döntéshozó informatikai kérdésekben való jártassága és képzettsége nem áll szoros összefüggésben az alkalmazott biztonsági technológiákkal és rosszindulatú szoftverek elleni megoldásokkal (11-12. ábra).



12. ábra A döntéshozó informatikai besorolása és a fizikai védelem közötti kapcsolat [forrás: saját]



11. ábra A döntéshozó informatikai besorolása és az informatikai rendszer védelme közötti kapcsolat [forrás: saját]

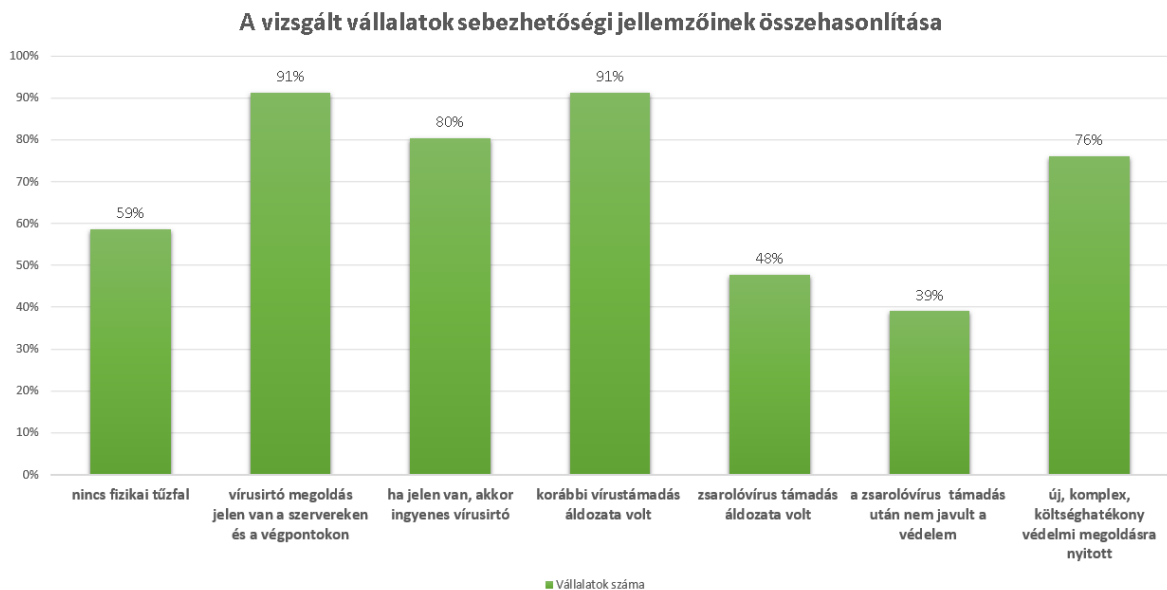
A 6. táblázat a vállalatok fizikai behatolásérzékelés és a fizikai tűzfalak meglétével kapcsolatos válaszait foglalja össze. Az eredmények azt mutatják, hogy 67%-uk rendelkezik fizikai behatolásérzékelővel a telephelyein, de csak a vállalatok 41%-ánál található fizikai tűzfal. A vállalatok nagyobb arányban rendelkeznek tűzfallal azokon a hálózatokon, melyeknél az informatikai döntéshozó nem rendelkezik informatikai végzettséggel. Ez az állítás egyben bizonyítja, hogy az informatikai infrastruktúrák fizikai és hardveres védelme nem függ szorosan a döntéshozó informatikai ismereteitől.

5. táblázat A döntéshozó informatikai besorolása és a vállalat bemutatott védelmi szintje közötti kapcsolat

Képesítés	Behatolásvédelem		Fizikai tűzfal	
	nincs	igazoltan van	nincs	igazoltan van
felsőfokú, szakirányú	2	5	4	3
felsőfokú, nem szakirányú	4	13	8	9
középfokú szakirányú	1	1	1	1
középfokú, nem szakirányú	8	12	14	6
A vállalatok száma	15	31	27	19
%-os megoszlás	33%	67%	59%	41%

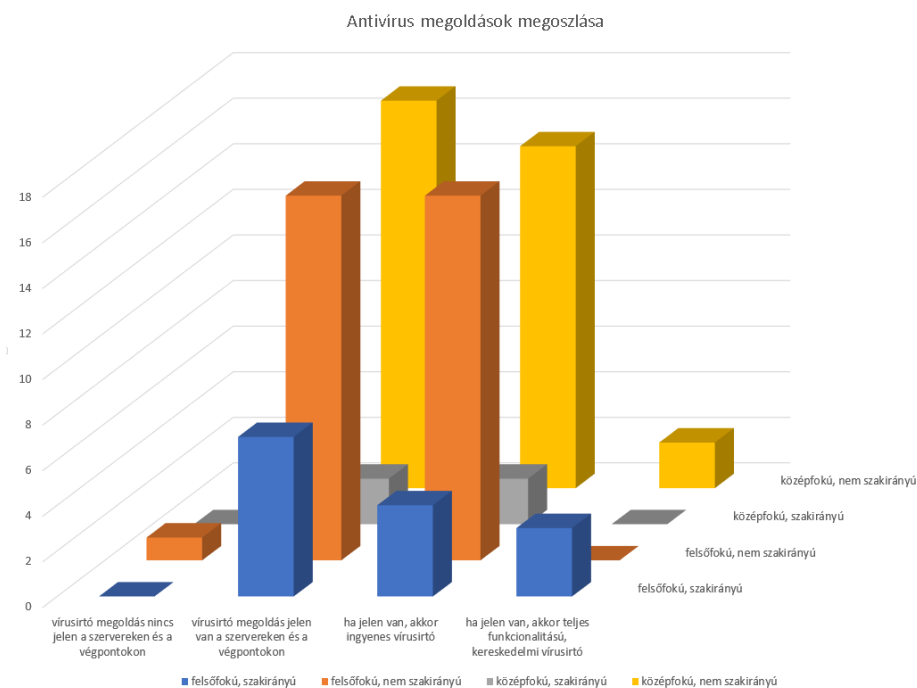
A második kutatási kérdés (K-2) arra vonatkozott, hogy minden vállalat rendelkezik-e telepített antivírus megoldással a munkaállomásain és a kiszolgálóin, és ha igen, akkor ez egy ingyenes, korlátozott funkciókkal rendelkező változat, vagy kereskedelmi forgalomban kapható, teljes körű védelmet nyújtó változatot használtak?

Az alkalmazott vírusirtó megoldások mellett az interjúk kérdései arra is fókuszáltak, hogy a megkérdezett vállalatok estek-e már áldozatul bármilyen vírustámadásnak, kifejezetten zsarolóvírus-támadásnak és ha igen, akkor hogyan kezelték ezt, és hogyan hatott a támadás a vállalat üzletmenetére és működésére. A megkérdezett vállalatok mindegyike végez olyan online tevékenységeket, mint például a saját tevékenységének reklámozása, webshop működtetése, online banki ügyintézés, online számlázás az egyéb elektronikus ügyintézési feladatok mellett. Ennek ellenére a megkérdezettek 59%-a nem rendelkezik fizikai tűzfallal. Bár 91%-uk használ szoftveres vírusirtó megoldást, de 80 %-uk valamilyen ingyenes vírusirtó szoftver használata mellett döntött. A válaszadók 91%-a volt már korábban vírustámadás áldozata, 48% pedig zsarolóvírus-támadás áldozata is. A támadásokat követően a vállalatok 39%-a nem változtatott korábbi hardveres vagy szoftveres védelmén, 76%-uk azonban azt mondta, hogy hajlandó lenne költséghatékony, komplex IT-biztonsági rendszert bevezetni a vállalatánál. A 13. ábra a fent említett adatok összehasonlítását szemlélteti.



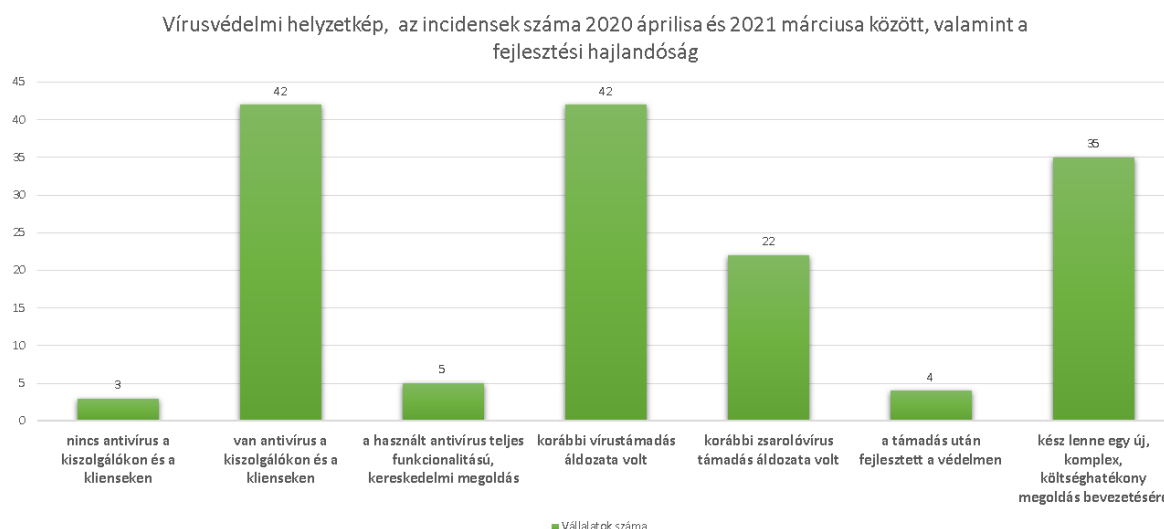
13. ábra A vizsgált vállalatok sebezhetőségi jellemzői [forrás: saját]

A második kutatási kérdéshez szorosan kapcsolódott a következő (K-3), vagyis, hogy milyen szerepet játszik az informatikai döntéshozó a vírusvédelmi megoldás kiválasztásában és ez milyen kapcsolatban áll az informatikai területen megszerzett szakmai képzettségükkel? A döntéshozók informatikai készségeit megvizsgálva arra kerestem a választ, hogy az informatikai ismeretek megléte, vagy éppen hiánya milyen összefüggésekben jelenik meg a döntéshozatalon keresztül az alkalmazott antivírus megoldásokkal kapcsolatosan. A 14. ábrán jól látható, hogy a vírusvédelmi megoldások alkalmazása ugyan nincsen összefüggésben a képzettség szintjével, azonban a nem specifikusan informatikai végzettséggel rendelkező döntéshozók sokkal gyakrabban döntenek teljeskörű antivírus megoldás használatára.



14. ábra A jelenleg használt antivírus megoldások eloszlása a döntéshozók informatikai képzettségének függvényében [forrás: saját]

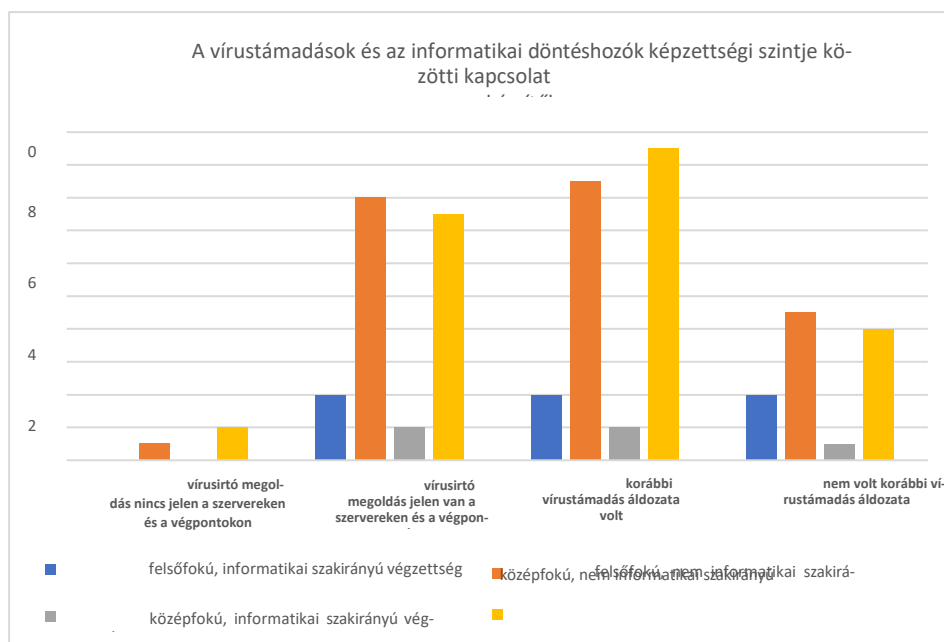
A következő kérdések az elszenvedett vírustámadások számát, illetve az interjút megelőző évben alkalmazott vírusvédelmi megoldásokat mérték fel. Ennek eredményét a 15. ábra mutatja be. A megkérdezett vállalkozások nagy része 2020 áprilisa és 2021 márciusa között valamilyen rosszindulatú szoftveres támadás, egyes esetekben zsarolóprogram áldozata lett.



15. ábra Vírus- és zsarolóvírus-támadások száma az interjút megelőző évben [forrás: saját]

Az interjúk alapján a megkérdezett 46 vállalkozás közül egy kivételével mindegyiket érte már korábban rosszindulatú szoftveres támadás az informatikai rendszereikben, 22 pedig zsarolóvírus-támadást szenvedett el. A kérdőív eredményeiből egyértelműen kitűnik az a tendencia, hogy minél több ismerettel rendelkeznek a vállalat alkalmazottai - különösen a döntéshozók -- (informatikai oktatás vagy biztonságtudatossági képzés révén), annál kisebb valószínűséggel válnak egy esetleges jövőbeli zsarolóvírus-támadás áldozatává.

A megkérdezett vállalatok közül csak ötnél volt professzionális, kereskedelmi forgalomban lévő vírusvédelmi megoldás telepítve a számítógépekre, és bár nagy részük (46-ból 42) már volt vírustámadás áldozata, csak négy vállalat hajtott végre változtatásokat az informatikai rendszerében a biztonság növelése érdekében. A válaszokból kiderült, hogy a döntéshozó iskolai végzettsége és informatikai ismeretei nem mutattak szoros összefüggést azzal, hogy az informatikai rendszereket ingyenes, vagy kereskedelmi forgalomban kapható professzionális vírusvédelmi szoftverrel védték-e (16. ábra).



16. ábra Kapcsolat a döntéshozó informatikai besorolása - az informatikai rendszerek védelmi szintje és az elszenvedett vírustámadások között [forrás: saját]

A megkérdezett 46 vállalat közül 35 engedélyezte a dolgozóknak, hogy otthonról dolgozzanak és ezeknél a cégeknél az alkalmazottak 10-50%-ának biztosított ez a lehetőség a munkavégzésre. Az otthoni munkavégzéshez minden esetben saját számítógépet használnak a dolgozók (7. táblázat). A felmérés során az alábbi eseteket különítettük el:

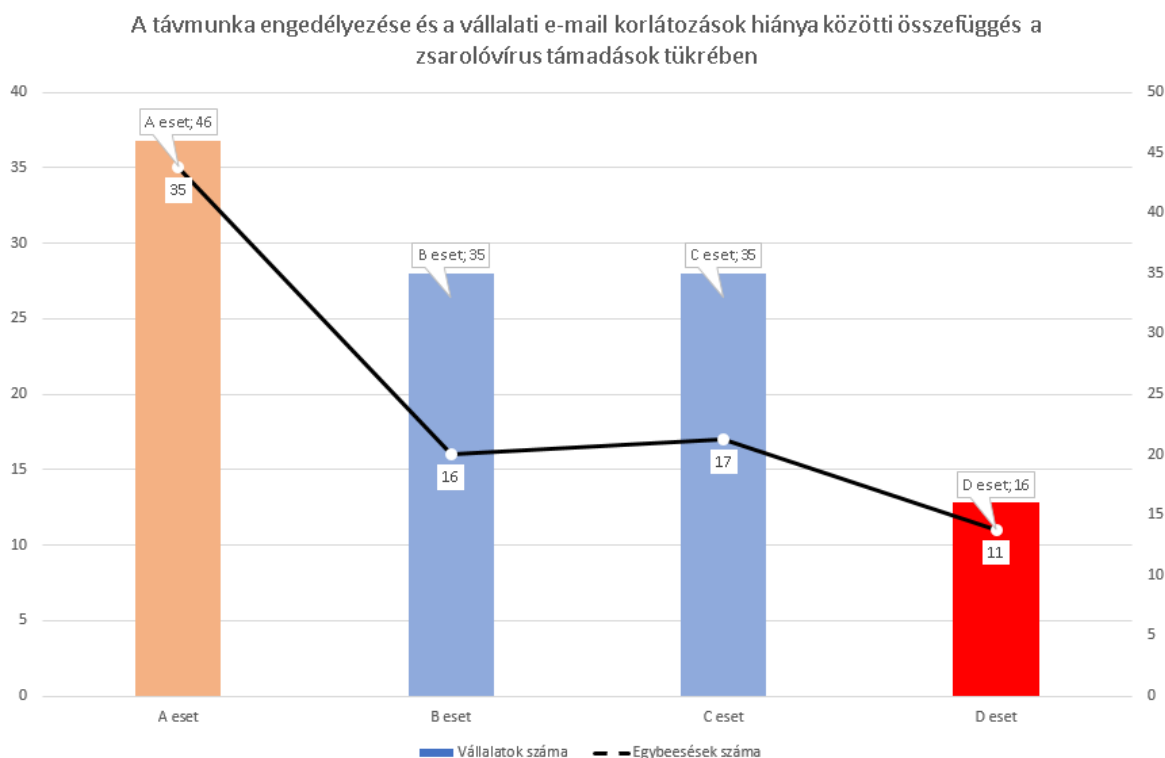
- A eset: engedélyezett az otthoni munkavégzés az alkalmazottak számára
- B eset: az otthoni munkavégzés engedélyezve van és a munkahelyi e-mail címeiken magánjellegű levelezés is folytatható
- C eset: az otthoni munkavégzés engedélyezve van és a vállalat zsarolóvírus-támadás áldozata lett
- D eset: az otthoni munkavégzés engedélyezve van, a munkahelyi e-mail címeiken magánjellegű levelezés is folytatható és a vállalat zsarolóvírus-támadás áldozata lett

6. táblázat Az otthoni munkavégzés és a zsarolóvírus-támadások közötti összefüggés

A eset	B eset	C eset	D eset
35	25	17	11
76.10%	71.42%	48.60%	68.75%

A fenti 35 vállalat közül 25-nél a munkavállalók saját vállalati e-mail címmel rendelkeznek, és a 25 vállalat közül 24-nél a munkavállalók minden szabályozás nélkül használhatják a vállalati e-mail fiókjukat magáncélokra. Érdemes megjegyezni, hogy ugyan a megkérdezett vállalkozások mindegyike használ online szolgáltatásokat, de egyrészt nincsen a vállalatokon belül kötelező informatikai biztonsági képzés, másrészt a vállalati infrastruktúra nem

rendelkezik a manapság elvárható biztonsági technológiákkal. A megkérdezett vállalatok egyikénél sincsen proxykiszolgáló, illetve a vendéghálózat sincsen elválasztva a produktív vezeték nélküli hálózattól, ami komoly kockázatot rejt magában. A felmérésben résztvevő KKV-k esetében átlagosan 2-10 alkalmazott dolgozik napi rendszerességgel online platformokon (pl.: számlázás, banki ügyintézés, kötelező adatszolgáltatás, vállalkozásspecifikus online rendszerek). Az otthoni munkavégzésre vonatkozó vállalati irányelvek, a vállalati e-mail használati szabályok, valamint a zsarolóvírus-támadások közötti összefüggéseket a 17. ábra szemlélteti. Jól látható, hogy ha az otthoni munkavégzés engedélyezve van és a munkahelyi e-mail címeken magánjellegű levelezés is folytatható, akkor az esetek közel felében zsarolóvírus-támadás áldozataivá váltak a vállalatok. Hasonló az arány az otthoni munkavégzés engedélyezése és a támadások valószínűsége kapcsán is.



17. ábra A home office-ra vonatkozó vállalati irányelvek, a vállalati e-mail használat és a zsarolóvírus-támadások közötti összefüggés [forrás: saját]

A negyedik kutatási kérdés (K-4) arra vonatkozott, hogy milyen hajlandóságot mutatnak a mikro-, kis- és középvállalkozások arra, hogy egy új, költséghatékony, komplex védelmi rendszert bevezessenek. A megkérdezett cégeknél megfigyelhető volt, hogy sok esetben nincsenek tisztában egy-egy támadás jelentőségével, a kár mértékével, valamint legtöbbször a vállalatoknak sincsen informatikai ismeretekkel rendelkező munkatársa (rendszergazda, vagy rendszerüzemeltető), aki szakszerűen tudná kezelni ezeket az incidenseket. A felmérésből kiderült, hogy a vállalatok jelentős (76%) része hajlandó lenne a már meglévő vírusvédelmi és egyéb IT-biztonsági megoldások mellett egy költséghatékony, automatizált és

komplex megoldásra fordítani annak érdekében, hogy tovább növelhessék a védelem hatékonyságát. A fenti összefüggéseket és az interjúk eredményeit vizsgálva világossá vált, hogy jelenleg nincs olyan teljes értékű koncepcionális megoldás a piacon, amely a szabályozási hibák kompenzálásával és azok negatív hatásainak hatékony megelőzésével képes megfelelő védelmet biztosítani. A kutatás során azonosítottam a hálózatok azon kritikus pontjait, ahol a vírusok jelenlétét jelezni lehet, illetve ahol a károkat az esetlegesen fertőzött kliensektől és sérülékeny kiszolgálóktól teljesen függetlenül meg lehetne előzni.

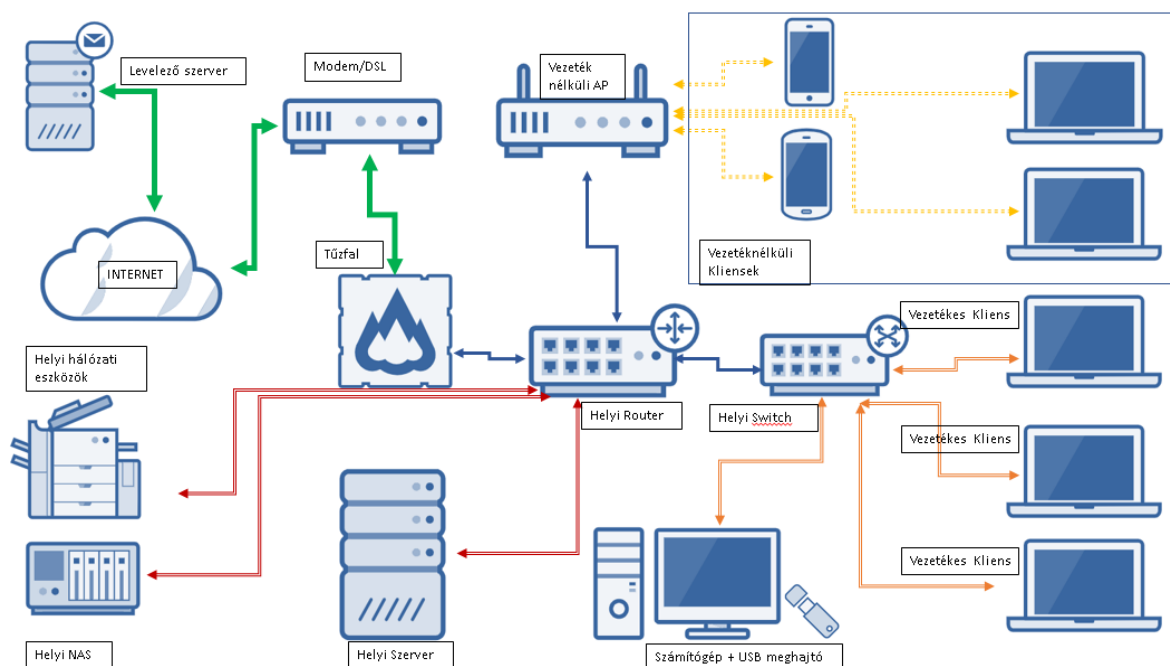
4.5. Az ARDS (Anti-Ransomware Defense System Model) védelmi modell

4.5.1. Támadási vektorok, jellemző belépési pontok a hálózatokon

Az interjúkból kiderült, hogy a megkérdezett vállalatok általában decentralizált peer-to-peer hálózati architektúrát, illetve egyszerű kis irodai-otthoni (SOHO) hálózati eszközöket használnak, melyeket sok esetben a megfelelő szakértelem hiányában nem megfelelően konfigurálnak. A LAN hálózatban lévő munkaállomások ilyenkor jellemzően egy vezeték nélküli AP funkcióval is rendelkező útválasztón (router) és néhány kizárólag layer 2-es funkciókkal rendelkező kapcsolón (switch) keresztül érik el az internetes erőforrásokat. A hálózati fájl-megosztásokat legtöbb esetben néhány kitüntetett kliens, vagy ritkább esetben valamilyen dedikált kiszolgáló, esetleg NAS biztosítja a dolgozók számára. A vállalatok sajátosságait, a zsarolóvírus-támadások jellegét és gyakoriságát elemezve világossá vált, hogy bár léteznek fejlett, hatékony vírusirtó megoldások, a felhasználók figyelmetlensége még mindig gyakran teret enged a zsarolóvírus-támadásoknak. Éppen ezért szükséges a hagyományos stratégiáktól eltérően működő védelmi modell kidolgozása és bevezetése. A kidolgozásra kerülő védelmi megoldásnak figyelembe kell vennie, hogy a felhasználói műveletek önmagukban is potenciális veszélyforrások lehetnek, amelyet többek között a következők okozhatnak:

- sebezhető külső eszközhöz férhet hozzá a munkavállaló a napi munka során a vállalati hálózathoz;
- az antivírust szándékosan, vagy figyelmetlenségből (beleértve a tudatlanságot is) kikapcsolják;
- gyenge jelszavak használata kényelmi vagy egyéb okokból;
- a biztonsági frissítéseket a felhasználó időhiány- és sávszélesség problémák miatt nem telepíti;
- egyes vállalatoknál elavult technológiákon alapuló rendszereket használnak;
- a felhasználó nem ismeri fel időben a rosszindulatú szoftverek jelenlétét a számítógépen;
- a vállalatnak nincsen szakképzett informatikusa, aki menedzselné az informatikai infrastruktúrát.

A 18. ábra egy átlagos KKV lehetséges hálózati topológiáját mutatja be. A felmérés során az is kiderült, hogy a legtöbb esetben a munkaállomások hitelesítés nélkül csatlakoznak a hálózati erőforrásokhoz (pl.: hálózati nyomtatókhoz, másolókhöz, szkennerekhez megosztott mappákhoz). A következő megfigyelés, hogy ha van is dedikált kiszolgáló a hálózaton, a szerver általában egyszerű fájlkiszolgáló, illetve adatbázisszerver, a legtöbbször nincs jelen például tartományvezérlő, központi hitelesítő szolgáltatás, tartományi csoportházirend, centralizált frissítéskezelés vagy tanúsítványtár. Mindezek mellett gyakran nincsen megfelelő vírusvédelem, és a távoli elérés (RDP, SSH, Telnet, WEBUI) nem kellően biztonságos.



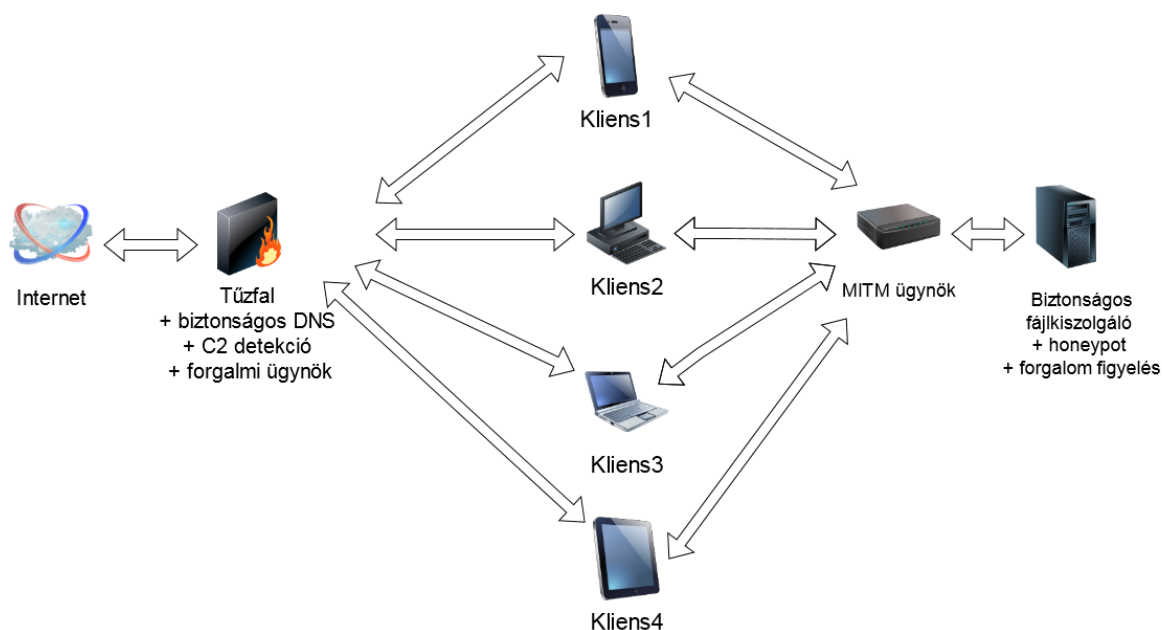
18. ábra A mikro- és kisvállalkozások által használt tipikus hálózati topológia [forrás: saját]

A támadási vektorokat számba véve megállapítható, hogy míg 2019-ben a támadások többnyire fertőzött fájlok (alkalmazástelepítők, dokumentumok, módosított patch-ek) segítségével zajlottak, addig 2020-ra a támadási stratégia gyökeresen megváltozott. A zsarolóvírusok innentől kezdve többnyire adathalász e-mailek segítségével terjedtek. Jó példa volt erre a megközelítésre a Zeppelin zsarolóvírus [37], amelyben egy MS Word dokumentumba képként beágyazott, látszólag ártalmatlan számla mellett a háttérben VBA makrók futnak. Az ilyen típusú támadások természetesen hatékonyan megelőzhetők megfelelő vírusirtó megoldással, spamszűrővel, makróvédelemmel és a felhasználók információbiztonsági tudatosításra nevelésével. A zsarolóvírus-támadások másik gyakori formája az, amikor az infrastruktúra sérülékeny, publikusan elérhető szolgáltatásait támadják exploitok segítségével. Ez jellemzően akkor fordul elő, ha a rendszer elavult, vagy hiányosan konfigurált. Sok esetben a hozzáférést a Darkweb-en vásárolt bejelentkezési adatokkal (körülbelül 15 milliárd RDP bejelentkezési adat áll rendelkezésre, 20 USD/fiók áron), vagy brute force támadások során

megszerzett hitelesítési adatokkal hajtják végre. Például a Microsoft RDP szolgáltatás alapértelmezett porton való közzététele 37 %-kal növeli a támadás valószínűségét, ugyanakkor a kétfaktoros hitelesítés, a szigorú jelszószabályzat betartatása és a sikertelen bejelentkezési próbálkozások számának korlátozása drasztikusan csökkentheti ezt a kockázatot. Egy további széles körben alkalmazott technika az, amikor a támadó sebezhető VPN-szolgáltatásokon keresztül szerez hozzáférést (pl.: a Pulse Secure [38] és a Fortinet korábbi verziói [39]). Ezekben az esetekben megoldást jelenthet a szolgáltatások rendszeres frissítése és a korábbi bejelentkezési adatok megváltoztatása.

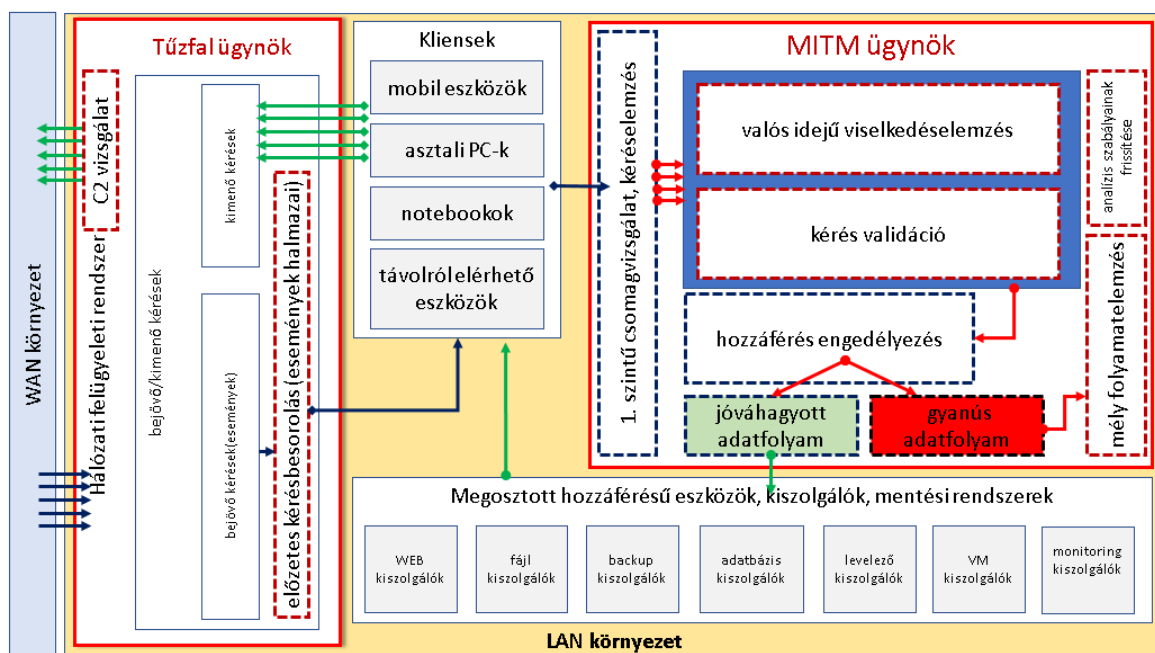
4.5.2. A védelmi modell ismertetése

Ahhoz, hogy értékes adatokhoz jussanak a támadók a zsarolóprogramok segítségével, vagy meg kell fertőzniük a hálózat egyik kliensét (pl.: adathalász e-mailek, kártékony weboldalak, fertőzött állományok letöltési linkjei), és onnan tovább haladva hozzáférhetnek az érzékeny adatokhoz, vagy közvetlenül kompromittálniuk kell a kiszolgálót (pl.: RDP, SSH, EAC, OWA). A fertőzött kliens munkaállomáshoz, vagy kiszolgálóhoz való távoli hozzáférést, az érzékeny adatok elérését és a távoli parancsvégrehajtást a támadó "command and control" (C2) infrastruktúrája biztosítja. Ez a kommunikációs folyamat a fertőzött kliensek esetében vagy DNS-lekérdezéssel, vagy közvetlen IP (hard coded IP) eléréssel kezdődik, így folyamatosan nyomon lehet követni a helyi DNS kiszolgálókon a gyanús DNS-lekérdezéseket (pl.: OpenDNS, Cisco Umbrella, Spamcop), illetve a tűzfalon a nem megfelelő IP reputációval rendelkező közvetlen IP-hozzáféréseket (pl.: Spamhaus, Cisco Talos, Emerging Threats, Alienvault). Szintén jól követhető metrika a kimenő adatmozgást jelző TX érték a központi router, vagy a WAN felőli átjáró megfelelő interfészein, mivel a szokatlanul nagy mennyiségű adat kifelé történő mozgását jelezhetik. A támadás részeként az ún. laterális mozgás során a támadó további felderítést végez, felkutatja a sebezhető eszközöket, illetve átvizsgálja a rendelkezésre álló hálózati erőforrásokat. Mivel az adatok távoli helyekre történő továbbítása, azok módosítása, titkosítása, törlése jelentős adatforgalmat és nagyszámú fájlműveletet eredményez, ezek egyértelmű jelzései lehetnek egy incidensnek (IoC). A piacon elérhető vírusirtó megoldások fejlesztése során elsősorban a fertőzések elhárítására és a rosszindulatú kódok végrehajtásának megakadályozására összpontosítanak a fejlesztők. Az általam kidolgozandó védelmi modell célja az, hogy a hálózatba már korábban beszivárgott, aktivizált zsarolóvírus a végpontvédelmeiktől és a kliensplatformoktól függetlenül felfedezhető legyenek (honeypotting, integritás-ellenőrzés, naplóelemzés, forgalom monitoring), miközben az adatok biztonságos tárolása (snapshot alapú technikák) és automatikus mentése (FC-P2P, iSCSI) biztosított. A fejlett zsarolóvírusok általában "Living off the Land" (LotL) technikát használnak, így például a fájlszervereken tárolt adatokhoz általánosan az operációs rendszer által natívan használt protokollokon keresztül férnek hozzá. A kliensek és a fájlkiszolgálók közötti kommunikáció a tanúsítványok megfelelő menedzsmentje mellett MitM (Man-in-the-Middle) technikák használatával követhetővé válik, így lehetőség nyílik az anomáliák, gyanús fájlművelet-szekvenciák, vagy az adatforgalmi kiugrások jelzésére, illetve incidensválaszok kiváltására. A javasolt modellem sematikus topológiája a 19. ábrán látható.



19. ábra Az ARDS-modell sematikus topológiája [forrás: saját]

Az MS Windows platformokon futó zsarolóvírusok jelentős része ábécésorrendben próbálja elérni a hálózati meghajtókat. Erre alapozva létre lehet hozni honeypot hálózati megosztásokat (az ábécé elejéről, illetve végéről választott meghajtóbetűkkel), amelyek látszólag érvényes, frissen generált, az adott szervezet profiljába illeszkedő adatokat tartalmaznak. Ezen fájlok tartalma azonban nem valós és egyetlen céljuk, hogy a honeypot meghajtókat figyelő szolgáltatások naplózzanak minden olyan fájlműveletet (hozzáférés, olvasás, írás, átnevezés, törlés), ami ezeket a meghajtókat érinti, valamint a művelet végrehajtó klienst le tudjuk választani a hálózatról. Természetesen a felhasználókat előre tájékoztatni kell arról, hogy a "hamis hálózati meghajtókat" semmilyen módon nem használhatják. Egy másik lehetőség, amit ez a módszer kínál, hogy a fájlműveletek sebességének különböző módon történő korlátozásával (pl.: I/O limitek, régi IDE-meghajtók használata, nagy mennyiségű kis méretű fájl) korlátozható a rosszindulatú programok aktivitása, így időt lehet nyerni az incidensek elhárítására. A vázolt működési mechanizmus független a többi védelmi megoldástól és alkalmas lehet a veszélyes folyamatok korai észlelésére a hálózat számos pontján, különböző szolgáltatási szinteken. A javasolt védelmi rendszer működési modellje a 20. ábrán látható. A rendszer kialakításának lényeges jellemzője, hogy többlépcsős szűrést, osztályozást és elemzést végezhet közel valós időben.



20. ábra A kidolgozott védelmi modell funkcionális architektúrája [forrás: saját]

A belső hálózaton használt fájlhozzáférési protokollokhoz tartozik egy-egy MitM-ügynök, melyek a hálózaton elérhető megosztott erőforrások védelméért felelősek. Az első lépés az általános csomagvizsgálat, amelyet egy valós idejű viselkedésvizsgálat és egy kérés-esemény érvényesítés követ. Az ügynök által megbízhatónak minősített események és kérések továbbításra kerülnek az eredeti kérésnek megfelelő kiszolgálóhoz. A megbízhatatlannak minősített kéréseket az MitM ügynök további elemzésre naplózza, így a későbbi vizsgálatok eredményeképpen az ügynök szabályrendszere frissülhet. A biztonsági mentések, valamint a fájl- és adatbázis-kiszolgálók védelme a javasolt ARDS modell másik kiemelt feladata. A kiszolgálói rétegben az adatintegritásért felelős ügynök egy újabb védelmi vonalat képvisel. Mivel az Extx, NTFS, HFS+, vagy FATx fájlrendszerek nem biztosítják a helyreállíthatóság alapját képező, időbélyeggel ellátott lemezkép-pillanatfelvétel készítésének lehetőségét, a választás a ZFS-re esett. A ZFS fájlrendszer segítségével az adatmentési ügynök csak olvasható pillanatfelvételeket (snapshot) készít a mentési házirendnek (gyakoriság, megtartási időkorlátok) megfelelően, majd igény esetén képes az adatok helyreállítására az időbélyeggel ellátott pillanatfelvételekből. A pillanatfelvételek akár teljes ZFS pool-okat, de akár csak dedikált adathalmazokat (datasets) is tárolhatnak. Lehetőségünk van ezen kívül klónozásra, ami egyszerűen egy duplikált adathalmaz létrehozására szolgál, vagy replikációra, ami pedig adatok másolására használható (akár egyik tárolóból egy másik tárolóba ugyanazon a gépen, vagy egy másik tárolóba is). A javasolt modell következő védelmi rétege a tűzfalügynök, ami egyrészt a belső hálózatról az internet irányába történő adatok mennyiségét monitorozza, másrészt ezen komponens segítségével a kimenő hálózati forgalom elemzése segítheti a lehetséges C2 tevékenység felismerését. A bemutatott ARDS modell egy olyan komplex rendszer, amely több lépésben, speciális beavatkozási és hozzáférési pontokon, az alkalma-

zási és szolgáltatási rétegekben képes lehet autonóm védelmet megvalósítani, és kiegészítheti a hálózatban már rendelkezésre álló szoftveres és hardveres védelmi technológiákat, ugyanakkor autonóm eszközként, fizikailag a belső hálózatra telepítve, automatikusan képes lehet biztosítani a kritikus üzleti adatok védelmét. A kutatás jelen szakaszában kidolgozott kliens-szerver architektúra modell publikálásra került [40].

4.5.3. A védelmi modell és a hagyományos védelmi koncepciók összehasonlítása

Fontos kiemelni, hogy az ARDS (Autonomous Ransomware Detection System) nem egyszerűen egy újabb észlelési megoldás, hanem egy többrétegű, incidens-megelőzési és helyreállítási architektúra, amelynek felépítése több ponton eltér a hagyományos SIEM, XDR, illetve behavior-based rendszerektől. A védelmi modell alapötlete, hogy a hálózatban már meglévő hardvereszközöket felhasználva, jól bevált, nyílt forráskódú szoftvermegoldásokat alkalmazva nyújthatunk komplex védelmi és helyreállítási funkciókat. A modell legfontosabb tudományos hozzájárulása nem egyetlen új detektálási algoritmus, hanem egy integrált, többrétegű architektúra, amely a zsarolóvírusok elleni védekezés során egyesíti a hálózati anomáliaészlelést, a „deception” (honeypot) technikákat, a fájlműveletek valós idejű monitorozását, a C2-kommunikáció felismerését, valamint a snapshot-alapú adat-helyreállítást. Ezzel az ARDS túlmutat a hagyományos SIEM vagy XDR rendszerek eseményközpontú megközelítésén, mivel nemcsak az incidensek észlelését támogatja, hanem azok hatásának mérséklését és az üzletmenet gyors helyreállítását is. Ez különösen a kis- és középvállalkozások számára jelenthet előnyt, ahol korlátozottak a kiberbiztonsági kockázatok csökkentésének lehetőségei. A „best practice” megoldások és a bevált technológiák integrációja inkább fejlesztési feladatnak tekinthető, míg a későbbiekben részletezett egyedi fájlkiszolgáló megoldás teljesítményértékelése és a gépi tanulással megvalósított fájlművelet-szekvenciákon alapuló zsarolóvírus-támadás felismerése mindenképpen tudományos hozzájárulást ad. A modell alapján épített architektúra nem kizárólag detektál, hanem a teljes incidens-életciklust igyekszik kezelni. A legtöbb SIEM és XDR megoldás elsődleges célja az események gyűjtése, a korreláció, a riasztások generálása, illetve az incidensek kivizsgálásának támogatása. Az ARDS ezzel szemben három funkciót egyesít. Ezek a korai felismerés, az automatikus beavatkozás és az adat-helyreállítás. Ezek különösen fontosak a zsarolóvírusok esetén, mivel ezeknél a támadásoknál nem elegendő a támadás észlelése, hanem biztosítani kell az üzletmenet gyors helyreállítását is. A hagyományos EDR/XDR rendszerek jellemzően végpontvédelmi ügynökre (endpoint agent), az operációs rendszer eseménynaplóira, kernel driverre, vagy antivírus motorra támaszkodnak. Ezzel szemben az ARDS modell felépítéséből adódóan független a végpontvédelmi megoldásoktól, illetve fontos kiemelni, hogy a támadást a végpont kompromittálása után is képes észlelni. A modell alapján felépített architektúra a hálózati forgalmat, a fájlműveleteket, a DNS-forgalmat, a C2 kommunikációt és a fájlserver-hozzáféréseket is figyeli. Ez különösen előnyös a Living off the Land (LotL) támadások esetében, ahol nincs klasszikus malware-folyamat. A legtöbb SIEM vagy XDR a kompromittált eszközt védi. Az ARDS modell ezzel szemben elsősorban az adatvagyon tekintetében védendő objektumnak, ami jelentős szemléletbeli különbség, hiszen támadás esetén lényegesebb lehet, mint önmagában a kliens védelme. A modell továbbá aktív honeypot mechanizmust

alkalmaz és csapdát állít, míg a legtöbb behavior-based rendszer passzívan figyeli a folyamatokat, az API-hívásokat, illetve a registry-műveleteket. A módszer előnye, hogy jól alkalmazva rendkívül kevés a hamis pozitív riasztás, nagyon gyors az észlelés, könnyen azonosítható a fertőzött kliens, ami akár automatikusan is leválasztható a hálózatról. A legtöbb XDR a végponton monitoroz, míg az ARDS modell MitM-alapú megközelítést javasol, aminek eredményeként az SMB és az NFS protokollok műveletei vizsgálhatóvá válnak. Ez lehetővé teszi a szokatlan fájl-művelet-szekvenciák nagyon korai felismerését. Számos modern XDR AI, ML, behavior-based megoldásokat alkalmaz amik ugyan rugalmasak, ugyanakkor esetükben a hatékonyság megőrzéséhez nagy mennyiségű tanítóadat szükséges, míg a fals pozitív arány magas lehet és gyakran hoznak "black box" döntéseket. A javasolt megközelítés ehelyett a könnyebb auditáció érdekében több, jól értelmezhető indikátort kombinál:

- DNS lekérdezések anomáliája;
- C2 kapcsolat;
- fájl-műveletek;
- forgalmi metrikák;
- honeypot események.

Talán a legjelentősebb különbség az említett megközelítésekhez képest, hogy az ARDS modell beépített helyreállítási mechanizmust tartalmaz, míg a legtöbb SIEM nem tárol produktív adatokat, az XDR megoldások pedig általában nem készítenek biztonsági mentést. Mivel az ARDS koncepciója közvetlenül integrálja a ZFS snapshot architektúrát, így egy incidens után nem kell külső backup rendszerre támaszkodni. Összefoglalva az ARDS modell talán legnagyobb erőssége, hogy a detektálást, az automatikus reagálást és az adat-helyreállítást egyetlen, egymást erősítő architektúrában egyesíti, miközben a támadást több, egymástól független megfigyelési ponton követi nyomon. A hagyományos megközelítések és az ARDS funkcióinak összehasonlítását a 8. táblázat foglalja össze.

7. táblázat Az ARDS és a hagyományos védelmi megoldások összehasonlítása

	SIEM	XDR/EDR	Behavior-based	ARDS
Eseménykorreláció	✓	✓	részben	✓
Végpontfigyelés	korlátozott	✓	✓	✗
Hálózati forgalom elemzése	korlátozott	részben	részben	részben
C2-kommunikáció felismerése	részben	✓	részben	✓
Honeypot/deception	ritka	ritka	nem jellemző	✓
Fájlszerver monitorozás	ritka	korlátozott	korlátozott	✓
Automatikus hálózati izoláció	részben	✓	részben	✓

Snapshot-alapú helyreállítás	×	×	×	✓
KKV-k költséghatékony támogatása	korlátozott	korlátozott	korlátozott	✓

5. Az ARDS modell fájlkiszolgálójának teljesítményértékelése

A kibertámadásokkal szembeni védelem felépítéséhez számos eszközzel rendelkezünk (pl.: protective DNS, fejlett tűzfalrendszerek, IDS/IPS, SIEM, vírusírtók, stb.) [41, 42]. Ezeket a szoftveres, illetve hardveres technológiákat párhuzamosan alkalmazva korszerű és hatékony védelmi architektúrák építhetők a helyi hálózaton kívülről érkező támadások megelőzésére, azonban a védelem egy másik kritikus pontja a tárolt adatok védelme, biztonságos kezelése, az adatok sérülése esetén a biztonságos és gyors helyreállítás feltételeinek megteremtése. Éppen ezért fontos, hogy hangsúly kerüljön a centralizáltan tárolt adatvagyron védelmének fejlesztésére, illetve fontos, hogy megvizsgáljuk a gyakorlatban is a fájlkiszolgálók esetében alkalmazható behatolás-, és a károkozást megelőző módszerek hatékonyságát.

Mivel a zsarolóvírus-támadások egyre kifinomultabbak, ezért a védelmi megoldásoknak is szükségszerűen fejlődniük kell. Sajnos a jelenleg alkalmazott stratégiák esetében nem fajsúlyos terület a támadás utáni adatvisszaállítás szempontjából kritikus fontosságú, validáción alapuló, automatizált biztonsági mentések menedzsmentje, a felhasználók fájlműveleteinek monitorozása, vagy a fájlkiszolgálók védelmének megerősítése. A fájlserverek védelmi megoldásainak funkcionális megfelelősége mellett a teljesítménye szintén fontos, vagyis, hogy a felhasználók ne érzékeljenek jelentős sebességbeli csökkenést a védelmi funkciók terheléstöbblete miatt. A vizsgálatok során a fő cél a ARDS modell fájlkiszolgálójának és a leggyakrabban alkalmazott megoldásnak az összehasonlítása volt teljesítménymérések segítségével. Ehhez kétirányú adatátviteli tesztek kerültek végrehajtásra különböző kliens-server architektúrákon alapértelmezett telepítési beállítások mellett.

Kutatási kérdések:

- K-5: Az előzőekben felvázolt ARDS védelmi modell atipikus fájlkiszolgálója képes-e az addicionális biztonsági mechanizmusok ellenére is hasonló teljesítményt nyújtani a felhasználóknak sebesség tekintetében a fájlok írása és olvasása esetén, mint a jelenleg leggyakrabban használt MS Windows 2019 Server megoldás?
- K-6: Összehasonlítható-e egzakt módon a különböző operációs rendszereken felépített és eltérő szolgáltatási környezeteket használó fájlkiszolgálók teljesítménye?
- K-7: Amennyiben lehetséges az összehasonlíthatóság, létezik-e olyan általánosan elfogadható metrika, ami képes kvantitatív módon jellemezni a heterogén rendszerek egymáshoz képesti teljesítményét?

5.1. Adatvagyonvédelmi és -visszaállítási megoldások áttekintése

A kutatások korai időszakában (~2006) az elsődleges megoldások közé tartozott a fájlrendszer helyreállításának folyamatszintű implementációja. A fájlrendszer helyreállítás (FSR) egyfajta helyreállítási lehetőség, amely lehetővé teszi a felhasználók számára, hogy a teljes fájlrendszer állapotát egy korábbi állapotra visszaállítsák. Liang és mtsai [43] tanulmányukban egy olyan virtuális lemezkörnyezetet (VDE) mutattak be, amely lehetővé teszi a merevlemezén végzett korábbi írási műveletek visszavonását és a fájlok korábbi verziójának visszaállítását. Ez a módszer a fájlrendszer gyors helyreállítására használható, még akkor is, ha a számítógépes rendszeren rendszerösszeomlás következik be, vagy rendszerindítási hiba lép fel. A VDE bizonyos szempontból megegyezik a virtualizált környezetekben használt virtuális lemezzel (vDisk), de alkalmazható VM támogatás nélküli környezetben is. Később a VDE teljesítményének kísérleti eredményeit elemezték a szerzők, és arra az eredményre jutottak, hogy más FSR-ekkel összehasonlítva a VDE fő előnye az alacsony teljesítményigény és a nagy helyreállítási sebesség [43]. Szintén a korai munkák egyike Gaonkar és munkatársainak kutatása [44], melyben megállapították, hogy az adatvesztés és az adatok elérhetetlenségének következtében fellépő költségek nagyok lehetnek, ezért a vállalkozások számos adatvédelmi technikát, például távoli tükrözést, pillanatfelvételeket és biztonsági mentéseket alkalmaznak a hibák elleni védelem érdekében. A szerzők továbbá megállapították azt is, hogy a technikák megfelelő kombinációjának kiválasztása nehéz feladat, mivel számos megközelítés létezik az adatok védelmére és az erőforrások elosztására. A tárolórendszerek tervezői jellemzően ad hoc technikákat alkalmaznak, ami gyakran túlméretezett, drága, vagy éppen, hogy alultervezett, nem megfelelő megoldásokat eredményez. Emellett a szerzők tanulmányukban egy elvi, automatizált megközelítést mutattak be a megbízható tárolórendszerek tervezéséhez. Életszerű tárolórendszer-követelményeket használva mutatták meg, hogy az általuk fejlesztett tervezőeszköz olyan javaslatokat dolgoz ki, amelyeknél a bekerülési költség akár a harmadára csökken az emulált emberi tervezési folyamat által létrehozott javaslatokhoz képest [44]. Oorschot és kollégái [45] kutatásukban a digitális objektumok rosszindulatú módosításának problémáját vizsgálták. Eredményképpen egy olyan védelmi mechanizmust mutattak be, amely a digitális objektumok védelmét szolgálja, megakadályozva azok jogosulatlan cseréjét, illetve módosítását, miközben az engedélyezett frissítéseket átlátható módon lehetővé teszi. A bemutatott megoldásukban digitális aláírásokat használtak anélkül, hogy központosított, nyilvános kulcsú infrastruktúrát igényelt volna a használata. A javasolt megoldásuk megvalósíthatóságának vizsgálatához egy Linux alapú prototípust implementáltak, ami a fájlrendszer bináris állományait, valamint az operációs rendszer és az alkalmazások bináris állományait védte a háttértáron. A tesztelések során megmutatták, hogy a javasolt védelmi megoldás védelmet nyújt a kernel módosításán alapuló - a fejlesztés idejében elérhető - különböző rootkitek ellen, miközben a megvalósítása minimális költséggel jár [45]. Az IoT technológiák és az edge, fog, illetve a cloud computing a bizalmas adatkezelés, az integritás és a rendelkezésre állás kihívásaival járnak együtt. Ezeknél a megoldásoknál kiemelkedő kockázatot jelent az információvesztés, az adatok hosszú ideig tartó elérhetetlensége vagy az információ kiszivárgása. Chervyakov és mtsai

tanulmányukban [46] egy testreszabható, robosztus és redundáns adattárolási rendszert javasoltak, amely képes titkosított adatok és a műveletek validálására. Rendszerük a redundáns maradékszámrendszert (RRNS) használja a hibajavító kódok generálásához és a kulcsmegosztáshoz. Munkájukban bevezették a szám rangjának közelítő értékét (AR), amely lehetővé teszi, hogy csökkentsék az RNS-ről bináris ábrázolásra történő dekódolás számítási bonyolultságát és az együtthatók méretét. A közelítő értéket és az RNS aritmetikai tulajdonságait alapul véve kidolgozták az AR-RRNS módszert a hibák felismerésére, javítására és a számítási eredmények ellenőrzésére. Emellett elméleti alapot adtak az információvesztés valószínűségének, az adatredundanciának, a kódolás és dekódolás sebességének konfigurálásához, azért, hogy számíthatók legyenek a különböző objektív preferenciák, működési terhelések és tárolási tulajdonságok. Az elemzésük azt mutatta, hogy az RRNS paraméterek megfelelő kiválasztásával a javasolt módszer nemcsak az adattárolás biztonságát és a megbízhatóságát növelte, hanem a titkosított adatok feldolgozásának idejét és az erőforrásigény csökkentését is lehetővé tette. Verma és kollégái [47] egy olyan fájlrendszer-mechanizmus tervét mutatták be 2015-ös tanulmányukban, amely megvédi az alkalmazásadatok integritását az olyan hibáktól, mint a folyamatok összeomlása, a kernel pánik és az áramkimaradás. Megoldásukban egy egyszerű interfész garantálja az alkalmazások számára, hogy a fájlban lévő alkalmazásadatok mindig a fájlon végzett legutóbbi sikeres fsync, vagy msync művelet eredményét tükrözzék. A javasolt fájlrendszerük ezenfelül egy új szinkronizációs mechanizmust kínál, amely hiba esetén a változásokat több fájlban is rögzíti. Hibainjekciós tesztekkel igazolták, hogy fájlrendszerük megvédi az alkalmazásadatok integritását az összeomlástól, a teljesítménymérések pedig megerősítették, hogy a megvalósításuk egyúttal hatékony is. A tesztekhez hagyományos hardverelemeket és módosítás nélküli Linux kernelt használtak, és később a termék kereskedelmi forgalomba is került. Verma és mtsainak [47] megoldása bármilyen olyan fájlrendszeren implementálható, amely támogatja a fájlonkénti, írható pillanatfelvételeket. Később azt is megmutatták, hogy az alkalmazás tartós adatainak konzisztens módosítására szolgáló mechanizmus (CMADD) egyszerűen megvalósítható a fájlonkénti klónozással, amely funkció már elérhető az AdvFS-ben és számos más modern fájlrendszerben. Az O_ATOMIC névre keresztelt kernel flag egyszerű felületet biztosít az alkalmazások számára, a fájl módosításait írás és leképezés során is hibamentessé teszi, seperi-metergít elkerülni az ismételt írásokat, és támogatja az alkalmazásadatok nagyobb volumenű és gyakori tranzakciós frissítéseit. Empirikus eredményeik azt mutatták, hogy az AdvFS O_ATOMIC implementációja megőrzi az alkalmazási adatok integritását a frissítések során mind a külső beavatkozást követő összeomlások, mind a véletlen összeomlások (pl.: áramkimaradások) esetén. Teljesítményméréseik alapján az O_ATOMIC használata tolerálható többletterhelést jelent, különösen gyakori, nagy adattömeget érintő frissítések esetén. A CMADD-mechanizmus fájlrendszerbe történő integrálását megkönnyíti, hogy nem igényel sem speciális hardvert, sem módosított operációs rendszer kernelt. Newberry és munkatársai [48] a HDFS fájlrendszer sávzélesség-optimalizálásával foglalkoztak. A HDFS egy olyan hálózati fájlrendszer, amelyet széles körben használnak Big Data adatfeldolgozási keretrendszerek-nél, mivel nagy klaszterek esetében is hatékonyan skálázható. 2019-es tanulmányukban HDFS-t alkalmazó klaszterrendszeren a hálózaton belüli gyorsítótárazás hatékonyságát vizsgálták. A

cél a hálózaton belüli sávszélesség-felhasználás csökkentése volt úgy, hogy a switch-eket tehermentesítsék. A tesztek során felismerték, hogy néhány alkalmazásban nagy mennyiségű, több ügyfél által használt adat szerepelt, ezáltal a hálózaton belüli olvasott adatok gyorsítótárazásával az olvasási műveletek átlagos, linkenkénti sávszélesség-használata ezekben az esetekben több, mint felére csökkenthető. Azt is megállapították, hogy a gyorsítótár-használati stratégia megválasztása jelentős hatással lehet a gyorsítótár-hatékonyságra ebben a környezetben, így a LIRS és az ARC általában a legnagyobb, illetve a kisebb méretű gyorsítótárak esetében teljesített a legjobban. A HDFS írási műveletek struktúráját figyelembe véve, kidolgoztak egy mechanizmust az írási műveletek linkenkénti teljes sávszélesség-felhasználásának csökkentésére, aminél a hagyományos írási metódust (pipelining) többsküldéssel (multicast) helyettesítették. Makris és mtsai [49] a nagy mennyiségű és nagy méretű adatok kezelésével kapcsolatos kutatásaikban a korlátozott erőforrások kérdését vizsgálták. Úgy találták, hogy az edge computing ígéretes paradigma például az IoT eszközök által generált hatalmas mennyiségű adat kezelésére és feldolgozására. Ha az adatok és a számítások közelebb kerülnek az ügyfélhez, azzal lehetővé válik, hogy a késleltetésre és sávszélességre érzékeny alkalmazások használata hatékonyabbá váljon. Azt is megjegyezték, hogy a periméter elosztott és heterogén jellege, valamint korlátozott erőforrás-kapacitása számos kihívást jelent a megfelelő tárolórendszer kiválasztása, illetve integrációja során. Tárolórendszerek széles skáláját vizsgálták, ami egyrészt változatos eredményekhez vezetett, másrészt átfogóbb képet adott a különböző megoldásokról. A teljesítményértékelést erőforrás-kihasználtsági és szolgáltatásminőségi (QoS) mérőszámok alapján végezték. Minden egyes tárolórendszert egy Raspberry Pi-re telepítettek, ami periméter eszközként szolgált és képes volt a hatékonyságot minimális energiafelhasználás és költség mellett optimalizálni. A kísérleti eredményeik azt mutatták, hogy a MinIO rendelkezik a legjobb általános teljesítménnyel a lekérdezési válaszidők, a memória felhasználás, a lemez I/O ideje és a tranzakciók arányának tekintetében [49]. A kutatások fő iránya esetükben magában foglalta a hálózaton mozgatott adatok átvitelének hatékonyságnövelését, valamint az adattárolás biztonságára irányuló fejlesztéseket is, mind az operációs rendszer kernelének szintjén, mind pedig a fájlstruktúra szintjén. A fájlkiszolgálók méretezése és a működési paraméterek optimalizálása a heterogén hálózati környezetben komoly kihívást jelent. Sokszor új módszerek, új technológiák segítenek a megvalósításban, de gyakori, hogy a már jól ismert és bevált technológiák újraszervezése hozza meg a kívánt eredményt. A fájlkiszolgálókra vonatkozó elvárásoknak való megfelelés, az adatok elérhetőségének, illetve védelmének biztosítása nagy mennyiségű adat esetén egy összetett feladat, amit a megnövekedett számú és egyre szofisztikáltabb módszerekkel kivitelezett zsarolóvírus-támadások tovább nehezítenek. A 9. táblázat a korábban tárgyalt adatvisszaállítási megoldások jellemzőit foglalja össze.

8. táblázat A korábban tárgyalt adatvisszaállítási megoldások jellemzői

Kutató-csoport	Megoldás/ Működési környezet	A javasolt megoldások képességei/szolgáltatásai					
		Teljes fájl-	Egyet- len fájl	Mű- velet	Hiba- kor- rekció/	Tel- jesít- mén	Hely- reállítási

		rend- szer helyre- állítás	helyre- állítás	érvé- nyesi- tés	észle- lés	yigé- ny	sebes- ség
Liang, J. és mtsai [43]	FSR VDE	igen igen	nincs igen	nincs nincs	nincs nincs	nagy kicsi	lassú gyors
Gaonkar, S. és mtsai [44]	Automati- zált megköze- lítés	nincs	igen	nincs	nincs	átla- gos	konf. függő
Van Oorschot, P.C. és mtsai [45]	Rootkit ellenálló mechaniz- mus	nincs	nincs	igen	van	átla- gos	nincs adat
Chervya- kov N. és mtsai [46]	RRNS - AR- RRNS	nincs	igen	igen (fájl műv./ titk.)	van	átla- gos	gyors
Verma, R. és mtsai [47]	AdvFS - CMADD megoldás	nincs	igen	igen	csak észle- lés	átla- gos	átla- gos

5.2. A zsarolóvírusok elleni védekezés lehetőségeinek rövid technológiai áttekintése

A zsarolóvírusok elleni korszerű védelmi rendszerek több, egymást kiegészítő technológiai megközelítést alkalmaznak a támadások megelőzésére, korai felismerésére és a bekövetkezett károk mérséklésére. Az egyes megoldások funkcionális felépítése jelentős eltéréseket mutat, ugyanakkor közös céljuk a rosszindulatú tevékenységek minél korábbi felismerése, illetve a helyreállítás támogatása. A jelen alfejezet célja nem az egyes kereskedelmi termékek részletes ismertetése, hanem a legfontosabb technológiai irányok áttekintése, amelyekhez a dolgozatban bemutatott megközelítés viszonyítható. A korszerű vállalati védelmi rendszerek jellemzően az alábbi funkciócsoportokat kombinálják:

- detekciós mechanizmusok (szignatúra- és viselkedés alapú felismerés);
- folyamatos monitorozás (fájl-, memória-, folyamat- és hálózati események megfigyelése);
- incidenskezelési és helyreállítási mechanizmusok (biztonsági mentések, verziókezelés, automatikus visszaállítás);

- központi menedzsment és környezeti integráció (felhőszolgáltatások, végpontok és egyéb infrastruktúra-elemek központi menedzsmentje).

A korszerű vállalati XDR-, EDR- és SIEM-platformok több adatforrás együttes elemzésével végzik a támadások felismerését. A végponti események mellett hálózati kommunikációt, rendszerfolyamatokat, memóriaműveleteket és felhasználói aktivitást is monitoroznak, majd az eltéréseket szabályalapú vagy mesterséges intelligenciát alkalmazó módszerekkel értékelik [15, 50–52]. Az ismertett megoldások közös jellemzője, hogy működésük több adatforrás együttes elemzésére, kliensoldali komponensekre, vagy komplex vállalati infrastruktúrára épül. Ezzel szemben a jelen kutatás annak vizsgálatára irányul, hogy kizárólag a központi fájlkiszolgálón keletkező fájlműveleti naplók felhasználásával megvalósítható-e a zsarolóvírus-tevékenység közel valós idejű felismerése. Ez a megközelítés egyszerűbb infrastruktúrák használatával, alacsonyabb erőforrásigény mellett biztosíthat alkalmazható detektálási képességet..

5.3. A kidolgozott fájlkiszolgáló védelmi megoldás jellemzői

A 4.5 alfejezetben ismertett ARDS modell [40] fájlkiszolgáló komponensének vizsgálata abból a feltételezésből indult ki, hogy a rosszindulatú kód már jelen van a hálózati környezetben, így a központi fájlkiszolgáló jelenti az utolsó védelmi vonalat. Ennek megfelelően a vizsgálat nem a kliensoldali vagy infrastruktúra-szintű sérülékenységek (pl. Active Directory, IIS, RDP vagy Living-off-the-Land technikák) kezelésére irányult, hanem annak elemzésére, hogy a fájlkiszolgálón alkalmazott védelmi mechanizmusok milyen teljesítményhatást gyakorolnak a rendszer működésére. A kutatás módszertani követelménye ezért egy olyan vizsgálati környezet kialakítása volt, amelyben az ütemezett pillanatképek, a honeypot alapú monitorozás, az IDS/IPS mechanizmusok és a vírusvédelmi funkciók együttes teljesítményhatása objektíven mérhető. A vizsgálat hipotézise szerint megfelelő architektúra alkalmazásával az említett védelmi mechanizmusok úgy integrálhatók, hogy a referenciaként választott fájlkiszolgálóhoz viszonyítva azok nem eredményeznek jelentős teljesítménycsökkenést. A hipotézis igazolására reprodukálható tesztkörnyezetben összehasonlító teljesítményvizsgálatot végeztem, amelynek módszertanát és eredményeit a következő alfejezetek ismertetik..

5.4. A vizsgálati módszer

5.4.1. Mérési és vizsgálati módszer

A vizsgálat célja annak meghatározása volt, hogy a javasolt fájlkiszolgáló-architektúrában alkalmazott adatvédelmi mechanizmusok milyen teljesítményhatást gyakorolnak egy tipikus kis- és középvállalati (KKV) környezetet reprezentáló referencia-rendszerhez viszonyítva. Ennek

érdekében összehasonlító fájlátviteli méréseket végeztem egy alapértelmezett konfigurációval telepített Windows Server 2019 alapú fájlkiszolgáló és a javasolt FreeBSD-alapú megoldás között. A Windows Server 2019 referencia-rendszerként történő kiválasztását az indokolta, hogy a vizsgálat időpontjában ez képviselte a KKV-k körében legelterjedtebb vállalati fájlkiszolgáló-platformok egyikét, ezért alkalmasnak tekinthető egy tipikus üzemeltetési környezet modellezésére. A Windows Server 2019 Standard operációs rendszer telepítése önmagában nem biztosít automatikusan RAID konfigurációt, BitLocker-alapú adattitkosítást és rendszeres Volume Shadow Copy snapshotokat. Mindhárom funkció külön konfigurációt, adminisztrációt és rendszeres felügyeletet igényel. A KKV-k korlátozott informatikai erőforrásai miatt ezek a szolgáltatások a gyakorlatban gyakran nem kerülnek bevezetésre, vagy csak részlegesen alkalmazzák őket. A referencia-rendszer kiválasztásakor fontos szempont volt továbbá, hogy mivel a BitLocker titkosítás, a VSS snapshot-kezelés és a RAID réteg egyaránt többletterhelést jelent, így ezek nélkül a szolgáltatások nélkül a lehető leggyorsabb, tipikusnak tekinthető referencia-környezet álljon elő, így a fájlműveletek sebességének tekintetében magasabb elvárásokat definiálhatunk a javasolt, egyedileg konfigurált FreeBSD 13.0 megoldással szemben. Ez végül konzervatív összehasonlítási alapot eredményezett, mivel a FreeBSD-alapú vizsgált konfiguráció ezzel szemben aktív ZFS tükrözést, AES-256-GCM titkosítást és folyamatos snapshot-kezelést alkalmazott. Ennek következtében a FreeBSD rendszer teljesítményelőnye bizonyos tesztesetekben olyan környezetben került igazolásra, ahol a vizsgált rendszernek nagyobb I/O terhelést kellett kiszolgáltatnia, ami tovább növeli az eredmények hitelességét és gyakorlati jelentőségét. A FreeBSD-alapú vizsgálati környezetben az adatvédelmi mechanizmusok a következő konfiguráció szerint kerültek kialakításra:

- OpenZFS alapú szoftveres RAID1 konfiguráció (tükrözés), ami a tesztadatokat tároló két fizikai lemezt (HDD) foglalta magába;
- automatikus, 2 percre ütemezett pillanatfelvétel készítése az adatokat tároló dataset-ről (snapshots) 30 napos megőrzési idővel (TTL);
- AES-256GCM kötetszintű titkosítás.

Azért esett a választás a FreeBSD operációs rendszerre, mert kimagasló stabilitást, megbízhatóságot és erőforráshatékonyságot nyújt, miközben a biztonság és a sebesség (szolgáltatások, I/O műveletek, hálózat) kiemelkedő marad. Ezt mi sem bizonyítja jobban, minthogy NAS rendszereknek éppúgy alapját képezi (pl.: TrueNAS CORE), mint tűzfalaknak (pl.: OpenSense, PfSense). A vizsgálati platform kialakításához emellett olyan operációs rendszerre volt szükség, amely natív módon támogatja az OpenZFS fájlrendszert, és lehetővé teszi a tükrözés, a kötetszintű titkosítás, valamint az időzített pillanatfelvételek egyidejű alkalmazását további külső komponensek nélkül. Az OpenZFS alkalmazását elsősorban olyan tulajdonságai indokolták, amelyek közvetlenül kapcsolódnak a kutatás célkitűzéseéhez. A fájlrendszer Copy-on-Write működési elve miatt a pillanatfelvételek létrehozása metaadat-alapú művelet, ezért normál üzemi körülmények között csak elhanyagolható CPU-, memória- és I/O-terhelést eredményez, mindemellett biztosítja a végpontok közötti adatintegritást ellenőrző összegek és önjavító mechanizmusok alkalmazásával, valamint támogatja a kötetszintű titkosítást, a tömörítést és az

igény szerinti deduplikációt. Ezek a tulajdonságok alkalmassá tették arra, hogy a javasolt adatvédelmi architektúra alapját képezze anélkül, hogy a vizsgálatok során önmagukban jelentős teljesítménytörzítést okoznának. Fontos megemlíteni ugyanakkor, hogy a snapshotok törlése OpenZFS esetén már számottevő I/O- és CPU-terhelést idézhet elő, mivel ilyenkor a ZFS felszabadítja a korábban hivatkozott blokkokat és módosítja a kapcsolódó metaadatokat. A pillanatfelvételek időköze a produktív környezetekben általánosan alkalmazott 10–15 perc helyett 2 percben került meghatározásra. Ez módszertani döntés volt, amely biztosította, hogy a mérési sorozatok során végrehajtott ismételt, 4 GiB méretű fájlátvitel nagy valószínűséggel legalább egy snapshot-ciklussal átfedésben történjenek, így annak a teljesítményhatása mérhetővé válik. A mérési eredmények torzításának elkerülése érdekében a vizsgált környezetekben nem alkalmaztam további, harmadik féltől származó vírusvédelmi megoldásokat, a Windows Server referencia-rendszer esetében pedig a Microsoft Defender valós idejű védelme is kikapcsolásra került. A vírusvédelmi mechanizmusok teljesítményhatásának vizsgálata nem képezte a kutatás részét, ezért ezeket kontrollváltozóként kezeltem. A mérési scenáriók kialakításakor a vizsgált fájlméreteket a kis- és középvállalkozások tipikus gyakorlatának megfelelően választottam ki. A szerver- és kliensoldalon minden mérés során azonos hardverkonfiguráció került alkalmazásra a hardvereredetű torzítások minimalizálása érdekében. Az adatátviteli időket kliensoldalon mértem mind olvasási, mind írási műveletek esetén. A vizsgálatokat egyaránt elvégeztem 1 Gbps és 10 Gbps névleges sávszélességű összeköttetéseken, mivel ezek reprezentálják a KKV-k környezetében jelenleg leggyakrabban alkalmazott Ethernet infrastruktúrákat.

5.4.2. A mérési bizonytalanság csökkentése és a mérési hibák kiszűrése

A teljesítménymérések megbízhatóságának biztosítása érdekében kiemelt cél volt a mérési bizonytalanság csökkentése és a mérési eredményeket torzító tényezők minimalizálása. A mérések során felléphetnek véletlen és szisztematikus hibák. A véletlen hibák kiszámíthatatlanok, előfordulhatnak például a mérés során használt elektronikai berendezések és a közöttük lévő fizikai kapcsolatot biztosító eszközök működési anomáliájaként. A szisztematikus hibák ezzel szemben a mérési módszer hiányosságából, a mérés során használt eszközök rendellenes működéséből, vagy esetleg a vizsgált rendszer működéséből fakadó sajátosságként jelenhetnek meg. A vizsgálatok során a célom ezeknek a hibáknak a felderítése, illetve a mérési eredményekre gyakorolt torzító hatásaik minimalizálása volt. A tesztelt szerverkörnyezetek teljesítményének pontosabb méréséhez a lehető legnagyobb mértékben csökkenteni kellett a mérési bizonytalanságot azáltal, hogy csökkentjük a bizonytalanságot okozó külső tényezők számát. E cél elérése érdekében pont-pont hálózati összeköttetést használtam (crossover CAT6 UTP, SFP+DAC), ami kizárta a hagyományos kapcsolt hálózati infrastruktúra késleltetéséből és terheléséből származó bizonytalanságokat. valamint kikapcsoltam a valós idejű vírusvédelmet mindkét oldalon a mérések idejére. A kontrollált mérési környezet ellenére továbbra is fennmaradtak olyan rendszerjellemzők, amelyek teljes mértékben nem voltak kiküszöbölhetők, ugyanakkor hatásuk ismert és értelmezhető volt. Ilyen tényezőknek tekinthetők:

- az operációs rendszerek gyorsítótárazási és pufferkezelési mechanizmusai, valamint háttérfolyamatai;
- a fájlrendszerek metaadat-kezeléséből és a célterület allokációjából adódó többletműveletek;
- az SSD-meghajtók belső vezérlési mechanizmusai és azok kisebb teljesítményingadozásai, illetve a HDD-k működési mechanizmusából adódó korlátok;
- a háttértárolók SATA III interfészéből adódó korlátok (6 Gbit/s vs. 10 Gbit/s), amely egyes scenáriókban a gyorsítótárazási mechanizmusokon keresztül kisebb mérési ingadozásokat eredményezhetett.

A mérési bizonytalanság és a mérési hibák hatásainak csökkentése érdekében minden olvasási és írási tesztet egymás után tízszer végeztem el mindegyik kliens-szerver konfiguráció esetében. Az ismételt mérések lehetővé tették az esetleges kiugró értékek azonosítását, valamint a mérési eredmények statisztikai jellemzését. Ehhez megvizsgáltam a minimális és maximális átviteli sebességeket, az átlagos átviteli sebességet, illetve a szórás értékeket az átviteli idők és az átviteli sebességek esetén. Az átviteli sebességek kiszámítása az átvitelhez szükséges idők alapján történt. A mérési folyamat során folyamatosan ellenőriztem a hálózati kapcsolat stabilitását annak érdekében, hogy az esetleges kommunikációs anomáliák ne befolyásolják az eredményeket. Az adatátvitel megbízhatóságának igazolása érdekében minden mérési ciklus során hash-alapú integritásellenőrzés történt. A fájlok hash-értékeinek összehasonlítása igazolta, hogy a mérési eredményeket adatátviteli hiba vagy sérült állomány nem befolyásolta, így a teljesítménymérések kizárólag a vizsgált architektúrák működéséből származó különbségeket tükrözik.

5.4.3. A reprodukálhatóság biztosítása

A vizsgálatok megismételhetősége szintén segít kiküszöbölni a mérési bizonytalanságot. A mérések a mérési terv alapján kerültek végrehajtásra és minden teszt során külön naplózásra kerültek az átviteli idők, amibe természetesen nem számítottam bele az ideiglenesen tárolt fájlok törléséhez szükséges időt. Mind a hardver-, mind pedig a szoftver környezetek specifikációja dokumentált. A mérésekhez használt fájlok generálása a következő BASH szkript segítségével, a fájlméretek módosításával történt:

```
for i in {1..2}; do head -c 256M < /dev/urandom > demo256M_${i}.bin; done
```

A fájlok átvitele Windows kliensek esetén a következő batch szkripttel történt:

```
@echo off
setlocal enabledelayedexpansion
for /l %%x in (1,1,10) do (
```

```
del /s /q c:\temp\*. * > NUL
@echo !time: =!
copy r:\temp\256M c:\temp\ > NUL
@echo !time: =!
)
```

```
A fájlok átvitelét Ubuntu kliens esetén a következő BASH szkript végezte:
for ((n=0;n<10;n++)); do rm -f /temp/256M/*;
time cp -f /home/user/Desktop/Bench_demo/256M/* /temp/256M;
done
```

5.4.4. Az eredmények értékelésének módszertana, az összehasonlíthatóság biztosítása

A mérési eredmények feldolgozása során a mért értékeket statisztikai módszerekkel elemeztem. Az ismételt mérések során meghatározásra és kizárásra kerültek a tranzien্স állapotot mutató átviteli sebesség értékek, az értékelhető mérési adatokból pedig a méréseket normális eloszlású valószínűségi változóként kezelve meghatározásra kerültek az egyes architektúrákra jellemző várható értékek, a mérésenkénti eltérések alapján pedig a statisztikai eloszlást jellemző szórás értékek. Mivel a különböző szoftverkonfigurációk esetében az átlagos átviteli idők egyedi jellemzők, így önmagukban nem alkalmasak a rendszerek szemléletes és egzakt összehasonlítására. Éppen ezért fontos, hogy egy kalkulált metrika segítségével arányszámként álljon elő a referenciakörnyezet és a vizsgált környezet egymáshoz képesti teljesítménye. Ezért kidolgoztam és bevezettem egy kvantitatív mérőszámot és egy egyszerű kiértékelés rendszert, amely alkalmassá teszi az inhomogén rendszerek referencia alapú, számszerű összehasonlítását. Ez az arányszám a következőképpen állt elő:

T_v : a vizsgált környezet átlagos átviteli ideje
 T_r : a referencia környezet átlagos átviteli ideje
 P_{rv} : teljesítmény indikátor

$$P_{rv} = \frac{T_r}{T_v}$$

Amennyiben $P_{rv} > 1$, akkor a vizsgált környezet teljesítménye jobb, mint a referencia környezet teljesítménye; ha $P_{rv} \approx 1$ ($\pm 0,1$), akkor a környezetek teljesítményét közel azonosnak tekintjük, ha pedig $P_{rv} < 1$, akkor a referencia környezet teljesítménye jobb, mint az aktuálisan tesztelt környezeté.

5.5. A teljesítménymérés folyamatának és körülményeinek ismertetése

5.5.1. A tesztkörnyezetek kiválasztása, a tesztesetek meghatározása

Az előzetes felmérések alapján a kisvállalati környezetben alkalmazott kliens-szerver architektúrákban a fájlkiszolgálók a leggyakrabban Windows Server operációs rendszerrel, vagy valamilyen Linux alapú operációs rendszerrel (pl.: Synology DSM (Debian alapú), TrueNAS SCALE, Ubuntu (Debian alapú)) vannak telepítve. A kliensek összetételét vizsgálva azt az eredményt kaptam a kutatás korábbi szakaszában [40], hogy a kérdőíves felmérés időpontjában elsősorban Windows 7, Windows 10, illetve Ubuntu operációs rendszerrel telepített munkaállomások voltak jelen a hálózatokban. Emiatt a vizsgálati környezet kialakítása során kliensoldali munkaállomáson up-to-date Windows 7 x64 SP1-et, Windows 10 x64 Pro 21H2-t és Ubuntu 22.04.4 LTS-t telepítettem alapértelmezett beállításokkal, míg kiszolgálóoldalon egy egyedileg konfigurált FreeBSD 13.0 alapú szoftverkörnyezetet, illetve Windows 2019 Server LTSC-t hasonlítottam össze. Így a következő tesztesetek kerültek meghatározásra a teljesítményértékeléshez:

- {TE-01}: Windows 7 x64 SP1 (6.1.7601) kliens - ARDS fájlkiszolgáló
- {TE-02}: Windows 7 x64 SP1 (6.1.7601) kliens - Windows 2019 Server (LTSC)
- {TE-03}: Windows 10 x64 Pro 21H2 kliens - ARDS fájlkiszolgáló
- {TE-04}: Windows 10 x64 Pro 21H2 - Windows 2019 Server (LTSC)
- {TE-05}: Ubuntu 22.04.4 LTS - ARDS fájlkiszolgáló
- {TE-06}: Ubuntu 22.04.4 LTS - Windows 2019 Server (LTSC)

A mérések a következő, jellemzően KKV produktív környezetekre méretezett kiszolgálón kerültek végrehajtásra:

DELL PowerEdge T140

- CPU: Intel Xeon 4Core E-2224 3,4 GHz
- RAM: 8GB DDR4 ECC RAM
- rendszer háttértár: Apacer AS350X 128 GB SSD
- adat háttértár: Gigabyte GP-GSTFS31240GNTD 240 GB SSD
- LAN: Intel X520-DA2 SFP+ 10GBE PCI-E Dual port NIC

Az 1 Gbit/s sebességű hálózati kapcsolat esetén a kliens a következő volt:

DELL Latitude 6230

- CPU: Intel Core i7-3520M 2.9 GHz
- RAM: 16 GB DDR3 non-ECC RAM
- rendszer és adat háttértár: Kingston SSD now V300 120 GB
- LAN: Intel 82579LM Gigabit NIC

A mérések során 10 Gbps-os aggregált sávszélesség esetén is vizsgáltam a fájlátvitelt, ezért fontos szempont volt, hogy a hálózati kártya natív szinten támogatott legyen (BIOS, OS,

driver) a kliens oldalon is, ezért ebben az esetben nem hagyományos munkaállomás volt a kliens. A munkaállomás ebben az esetben a következő volt:

DELL PowerEdge Mini T40

- CPU: Intel Xeon 4Core E-2224G 3.5GHz
- RAM: 8GB DDR4 ECC RAM
- rendszer és adat háttértár: Apacer AS350X 128 GB SSD
- LAN: Intel X520-DA2 SFP+ 10GBE PCI-E Dual port NIC

5.5.2. Az adatátviteli közeg áteresztőképességének meghatározása

Ahhoz, hogy össze tudjam hasonlítani a különböző kiszolgálói környezetek teljesítményét, gondoskodnom kellett az azonos hardver-, és identikus hálózati kapcsolatok mellett a megfelelő mintaadatokról is. Az adatátviteli tesztek 1 Gbit/s-os, illetve 10 Gbit/s-os réz alapú, pont-pont összeköttetéseket használva hajtottam végre. Ahhoz, hogy egy KKV, vagy akár egy több száz munkaállomással rendelkező multinacionális vállalat hálózatában előforduló, reprezentatív forgalom létrejöjjön (egyetlen ügyfél teljes sávszélességét kihasználva), fontos a megfelelő hálózati átviteli teljesítmény. Az első mérési sorozatban kereszttekő CAT6 UTP-kábelt használtam a hálózati kapcsolat közvetlen létrehozásához (minimalizálva a hálózati eszközök késleltetési hatását) 1 Gbit/s névleges sávszélességgel. A második mérési sorozatban 10G SFP+ DAC kábellel kötöttem össze a két Intel X520-DA2 hálózati kártyát közvetlenül, ami közel 10 Gbit/s tényleges sávszélességet biztosított. A hálózati kapcsolatok üresjáratú áteresztő képességét az átviteli tesztek megelőzően Iperf3 nyílt forráskódú szoftver segítségével vizsgáltam meg, ami "de facto" szabványnak tekinthető ezen a területen. Mind az 1 Gbit/s-os, mind a 10 Gbit/s-os kapcsolatok esetében mindkét irányba teszteltem a hálózati átviteli sebességet TCP-alapú forgalom esetén. A mérések eredményeit a 18. ábra szemlélteti. Fontosnak tartom megjegyezni, hogy az itt mért értékek a hálózati átviteli sebességet jelentik, és nem veszik figyelembe például a különböző protokollok jellemzőit (pl.: kódolás, nyugtázás, signing, egyéb protokollspecifikus kommunikáció), a lemezműveletekből eredő késleltetést, a különböző szoftverkörnyezetekben alkalmazott fájlkezelési és fájlrendszerbeli megoldásokat, vagy az eszköz-illesztőprogramok eltérő hatékonyságát.

```

root@user-Latitude-E6230:/home/user# lperf3 -c 192.168.2.119 -p 5201
Connecting to host 192.168.2.119, port 5201
[ S] local 192.168.2.1 port 60248 connected to 192.168.2.119 port 5201
[ ID] Interval      Transfer    Bitrate    Retr    Cwnd
[ S] 0.00-1.00 sec  91.7 MBytes 769 Mb/s    0       266 KBytes
[ S] 1.00-2.00 sec  90.3 MBytes 758 Mb/s    0       330 KBytes
[ S] 2.00-3.00 sec  89.2 MBytes 748 Mb/s    0       330 KBytes
[ S] 3.00-4.00 sec  90.2 MBytes 756 Mb/s    0       330 KBytes
[ S] 4.00-5.00 sec  90.2 MBytes 757 Mb/s    0       330 KBytes
[ S] 5.00-6.00 sec  89.1 MBytes 747 Mb/s    0       330 KBytes
[ S] 6.00-7.00 sec  90.2 MBytes 757 Mb/s    0       330 KBytes
[ S] 7.00-8.00 sec  89.2 MBytes 748 Mb/s    0       330 KBytes
[ S] 8.00-9.00 sec  90.1 MBytes 756 Mb/s    0       330 KBytes
[ S] 9.00-10.00 sec 90.2 MBytes 757 Mb/s    0       330 KBytes
-----
[ ID] Interval      Transfer    Bitrate    Retr
[ S] 0.00-10.00 sec 900 MBytes 755 Mb/s    0
[ S] 0.00-10.00 sec 898 MBytes 753 Mb/s
sender receiver

lperf Done.
root@user-Latitude-E6230:/home/user# lperf3 -c 192.168.2.119 -p 5201 -R
Connecting to host 192.168.2.119, port 5201
Reverse mode, remote host 192.168.2.119 is sending
[ S] local 192.168.2.1 port 60254 connected to 192.168.2.119 port 5201
[ ID] Interval      Transfer    Bitrate    Retr    Cwnd
[ S] 0.00-1.00 sec  101 MBytes 846 Mb/s    0
[ S] 1.00-2.00 sec  101 MBytes 846 Mb/s    0
[ S] 2.00-3.00 sec  101 MBytes 846 Mb/s    0
[ S] 3.00-4.00 sec  101 MBytes 846 Mb/s    0
[ S] 4.00-5.00 sec  101 MBytes 846 Mb/s    0
[ S] 5.00-6.00 sec  101 MBytes 846 Mb/s    0
[ S] 6.00-7.00 sec  101 MBytes 846 Mb/s    0
[ S] 7.00-8.00 sec  101 MBytes 846 Mb/s    0
[ S] 8.00-9.00 sec  101 MBytes 846 Mb/s    0
[ S] 9.00-10.00 sec 101 MBytes 846 Mb/s    0
-----
[ ID] Interval      Transfer    Bitrate    Retr
[ S] 0.00-10.00 sec 1010 MBytes 847 Mb/s    266
[ S] 0.00-10.00 sec 1008 MBytes 846 Mb/s
sender receiver

lperf Done.
root@user-Latitude-E6230:/home/user# lperf3 -c 192.168.3.120 -p 5201
Connecting to host 192.168.3.120, port 5201
Reverse mode, remote host 192.168.3.120 is sending
[ S] local 192.168.3.117 port 57616 connected to 192.168.3.120 port 5201
[ ID] Interval      Transfer    Bitrate    Retr    Cwnd
[ S] 0.00-1.00 sec  1.09 GBytes 9.40 Gb/s    0
[ S] 1.00-2.00 sec  1.09 GBytes 9.38 Gb/s    0
[ S] 2.00-3.00 sec  1.09 GBytes 9.40 Gb/s    0
[ S] 3.00-4.00 sec  1.09 GBytes 9.38 Gb/s    0
[ S] 4.00-5.00 sec  1.09 GBytes 9.40 Gb/s    0
[ S] 5.00-6.00 sec  1.09 GBytes 9.35 Gb/s    0
[ S] 6.00-7.00 sec  1.09 GBytes 9.37 Gb/s    0
[ S] 7.00-8.00 sec  1.08 GBytes 9.26 Gb/s    0
[ S] 8.00-9.00 sec  1.08 GBytes 9.31 Gb/s    0
[ S] 9.00-10.00 sec 1.09 GBytes 9.33 Gb/s
-----
[ ID] Interval      Transfer    Bitrate    Retr
[ S] 0.00-10.00 sec 10.9 GBytes 9.36 Gb/s    62
[ S] 0.00-10.00 sec 10.9 GBytes 9.36 Gb/s
sender receiver

lperf Done.
root@user-Latitude-E6230:/home/user# lperf3 -c 192.168.3.120 -p 5201 -R
Connecting to host 192.168.3.120, port 5201
Reverse mode, remote host 192.168.3.120 is sending
[ S] local 192.168.3.117 port 57620 connected to 192.168.3.120 port 5201
[ ID] Interval      Transfer    Bitrate    Retr    Cwnd
[ S] 0.00-1.00 sec  1.10 GBytes 9.44 Gb/s    0 1.51 MBytes
[ S] 1.00-2.00 sec  1.09 GBytes 9.40 Gb/s    0 1.51 MBytes
[ S] 2.00-3.00 sec  1.09 GBytes 9.36 Gb/s    0 1.51 MBytes
[ S] 3.00-4.00 sec  1.09 GBytes 9.41 Gb/s    0 1.51 MBytes
[ S] 4.00-5.00 sec  1.09 GBytes 9.38 Gb/s    0 1.51 MBytes
[ S] 5.00-6.00 sec  1.09 GBytes 9.41 Gb/s    0 1.51 MBytes
[ S] 6.00-7.00 sec  1.07 GBytes 9.23 Gb/s    0 1.51 MBytes
[ S] 7.00-8.00 sec  1.09 GBytes 9.40 Gb/s    0 1.60 MBytes
[ S] 8.00-9.00 sec  1.09 GBytes 9.40 Gb/s    0 1.60 MBytes
[ S] 9.00-10.00 sec 1.09 GBytes 9.41 Gb/s    0 1.60 MBytes
-----
[ ID] Interval      Transfer    Bitrate    Retr
[ S] 0.00-10.00 sec 10.9 GBytes 9.38 Gb/s    0
[ S] 0.00-10.00 sec 10.9 GBytes 9.38 Gb/s
sender receiver

```

21. ábra Iperf3 mérési eredmények [forrás: saját]

A mérési adatokból jól látszik az elméleti átviteli sebesség maximuma mind a két irány és mindkét technológia esetében (755 Mbit/s - 847 Mbit/s; 9,36 Gbit/s - 9,38 Gbit/s).

5.5.3. Méréshez használt adatok összetételének meghatározása

A mérési terv kialakítása során fontos volt figyelembe venni, hogy különböző produktív környezetekben jellemzően eltérő méretű fájlokkal dolgoznak a felhasználók. Annak érdekében, hogy a vizsgált fájlátviteli protokollok, CIFS/SMB verziók és konfigurációk által használt tömörítési eljárások ne befolyásolhassák számottevően az eredményeket, véletlenszerű bitmintázat jellemezte a tesztekhez használt fájlokat. Mind az olvasási, mind az írási tesztek esetén összesen 1 GiB adatmennyiség került átvitelre, ami a különböző fájlméreteknek megfelelő számú fájlból tevődött össze. Ez azt jelenti, hogy 500 kiB-os, 4MiB-os, 256 MiB-os és 1 GiB-os fájlok formájában állt elő az 1 GiB adatmennyiség a következőképpen:

- 1 x 1 GiB
- 4 x 256 MiB
- 256 x 4 MiB
- 2097 x 500 kiB

A fájlműveleteket szándékosan a normál felhasználói gyakorlathoz igazítva, natív parancs-sori megoldásokat használva (copy, del, cp, rm) hajtottam végre, mivel az olyan eszközök, mint például a robocopy, illetve az rsync nagymértékben befolyásolhatták volna a tesztek eredményeit, mindamelllett, hogy nem tükrözték volna az életszerű használatot. Minden olvasási, illetve írási teszt után egy rövid törlési folyamat következett, amikor a célkönyvtár

kiürítésre került. Ez persze nem tényleges és biztonságos felülírást (wipe) jelentett, hanem logikai törlést és nem számított bele az átviteli időbe.

5.6. A teljesítménymérés eredményeinek ismertetése

Elsőként 1 Gbit/s sebességű pont-pont alapú hálózati kapcsolatot használva vizsgáltam a különböző kliens-szerver platformok átviteli sebességét. Ez elsősorban azért volt fontos, mert ez a sávszélesség még aggregált terhelést modellezve is tipikusnak mondható napjainkban a non-profit szervezetek, illetve KKV-k esetében. Másodsorban pedig azért, mert kapunk egy átviteli karakterisztikát az egyes rendszerekre vonatkozólag, ami összehasonlítási alapot ad. Ezt követően 10 Gbit/s sebességű, szintén pont-pont alapú hálózati összeköttetésen keresztül zajlott az adatátviteli teszt, ami már a nagyvállalati hálózati infrastruktúrákat tekintve is releváns léptékű. Itt a Microsoft Windows 7 x64 operációs rendszert már nem vizsgáltam, mivel ez a rendszer nem támogatja az Intel X520-DA2 hálózati kártyát, valamint életszerűtlen lett volna, ha egy biztonsági frissítésekkel nem támogatott operációs rendszert használunk, miközben modern hálózati infrastruktúrát biztosítunk a mérésekhez. Az 1 Gbit/s-os hálózati kapcsolat esetén a tíz mérés átlaga 846 Mbit/s (100,85 MiB/s) volt, míg a 10 Gbit/s-os összeköttetésnél 9,38 Gbit/s-os (1118,18 MiB/s) értéket mértünk. Azért is volt fontos az 1 Gbit/s kapcsolati sebességen elvégzett méréseket követően jelentősen megnövelni a hálózat áteresztőképességét, mert ezt a technológiát használva a legtöbb esetben megközelítettem az elméleti átviteli sebesség felső határát (125 MiB/s), ami miatt a teljesítménykülönbségek már kevésbé voltak szembe ötlőek, valamint a munkaállomások és a kiszolgálók szignifikáns terhelése nem valósult meg. A mérések során sem kliens, sem pedig szerver oldalon nem haladta meg a CPU terheltsége a 47, a rendszermemória kihasználtsága pedig az 52%-ot. Az SSD meghajtók azonban többször is 100%-os kihasználtságot mutattak a gyorsabb, 10 Gbit/s sávszélességű teszteknel. Az eredmények feldolgozása során azt tapasztaltam, hogy az első mérések eredményei jelentősen eltértek átviteli idő tekintetében a további kilenc mérés eredményétől, főleg a kisebb fájlok (500 KiB, 4 MiB) átvitelénél, jellemzően Windows operációs rendszerek esetén. Mivel azonban a következő kilenc mérés eredményeiből számított szórásértékek sehol nem mutattak jelentős eltérést, így a mérések átlagát alapul véve reprezentatívnak tekinthettem az eredményeket. A tesztek eredményeit a két különböző hálózati kapcsolat esetében elemeztem, olvasási és írási műveletekre bontva. Annak érdekében, hogy könnyen össze lehessen hasonlítani a különböző szerverkörnyezetek teljesítményét ugyanazon ügyfelek esetében, azokat egymás mellett elhelyezett diagramoszlopokban jelenítjük meg. A mérések során tárolt nyers teljesítményadatok és szórás értékek a mellékletben találhatóak. A kapcsolódó szakirodalom áttekintését és a rendelkezésre álló megoldások számbavételét követően látható, hogy nincsen hasonló koncepciójú, integrált megoldás, ezért a kidolgozott modell hiánypótló lehet ezen a területen.

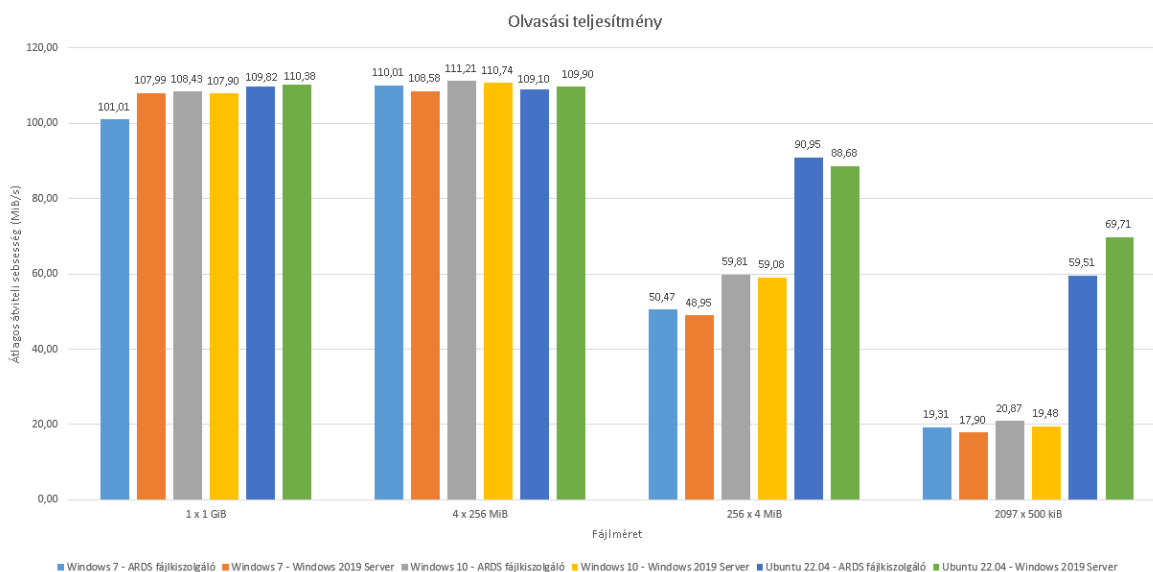
5.7. Eredmények 1 Gbit/s-os kapcsolat esetén

5.7.1. Az olvasási teljesítményvizsgálat eredményei

Az 1 Gbit/s sebességű hálózaton végzett olvasási tesztek részletes eredményeit a 10. táblázat és a 22. ábra szemlélteti. Az alkalmazott jelölési konvenció a fájlátviteli műveletek során a következő: $a \text{ fájl} \text{ok száma} \times \text{fájl méret}$.

9. táblázat Olvasási teszteredmények 1 Gbit/s sávszélesség esetén

Teszteset	Fájlok darabszáma és mérete			
	1x1 GiB	4x256 MiB	256x4 MiB	2097x500 kiB
Átlagos átviteli sebesség MiB/s-ban				
Windows 7 - ARDS fájlkiszolgáló {TE-01}	101,01	110,01	50,47	19,31
Windows 7 - Windows 2019 Server {TE-02}	107,99	108,58	48,95	17,90
Windows 10 - ARDS fájlkiszolgáló {TE-03}	108,43	111,21	59,81	20,87
Windows 10 - Windows 2019 Server {TE-04}	107,90	110,74	59,08	19,48
Ubuntu 22.04 - ARDS fájlkiszolgáló {TE-05}	109,82	109,10	90,95	59,51
Ubuntu 22.04 - Windows 2019 Server {TE-06}	110,38	109,90	88,68	69,71



22. ábra Olvasási teszteredmények összehasonlítása 1 Gbit/s sávszélesség esetén
[forrás: saját]

A 256 MiB és az 1 GiB méretű fájlok esetében mind a hat kliens-kiszolgáló párosítás nagyon hasonló teljesítményt nyújtott (min. 101,01 MiB/s, max. 111,21 MiB/s), amint a 7. táblázatban látható, de a 4 MiB méretű fájlok esetében jelentős teljesítménycsökkenés volt megfigyelhető a Windows ügyfelek esetében. Az eredmények azt is megmutatták, hogy az átviteli sebesség mindkét szerverplatform esetében közel azonos volt mind a Windows 7, mind pedig a Windows 10 kliensek tesztéseit figyelembe véve. Az is szembeötlő volt, hogy a fájl-kiszolgálótól függetlenül az Ubuntu 22.04 LTS kliens 30-40 %-kal gyorsabb átviteli sebességet tudott elérni. Az 500 kiB méretű fájlok olvasásakor a sebesség a Windows 7 és a Windows 10 esetében a szerveroldali platformtól függetlenül majdnem ötödével csökkent, míg az Ubuntu kliens mindkét esetben nagyjából háromszor gyorsabb volt.

5.7.2. Az írási teljesítménytesztek eredményei

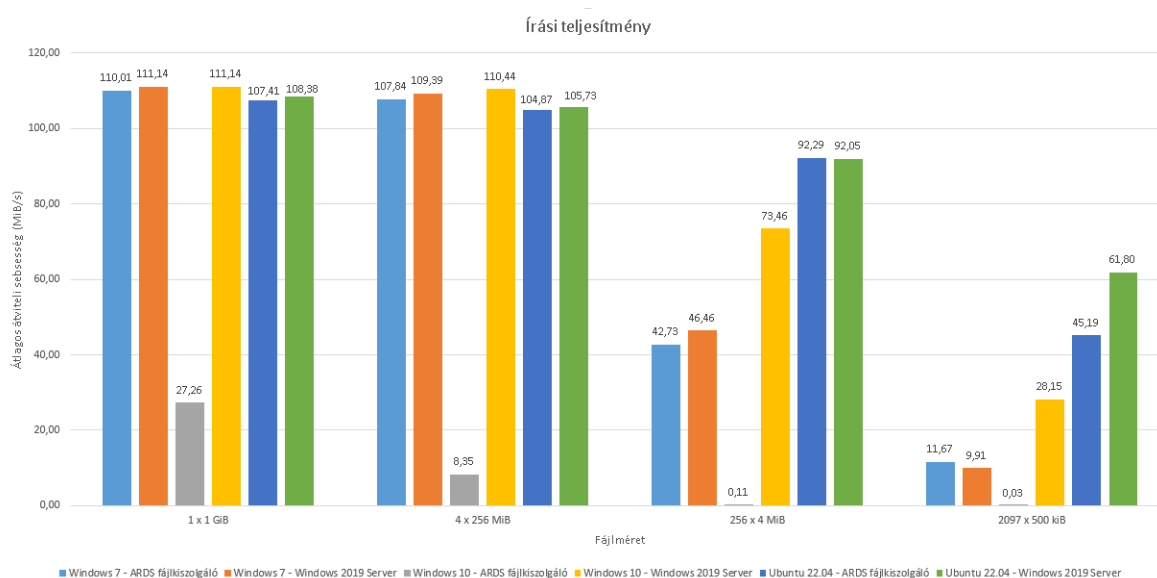
Az 1 Gbit/s sebességű hálózaton végzett írási tesztek részletes eredményeit a 11. táblázat és a 23. ábra szemlélteti.

10. táblázat Írási teszteredmények 1 Gbit/s sávszélesség esetén

Teszteset	Fájlok darabszáma és mérete			
	1x1 GiB	4x256 MiB	256x4 MiB	2097x500 kiB
	Átlagos átviteli sebesség MiB/s-ban			
Windows 7 - ARDS fájlkiszolgáló {TE-01}	110,01	107,84	42,73	11,67
Windows 7 - Windows 2019 Server {TE-02}	111,14	109,39	46,46	9,91
Windows 10 - ARDS fájlkiszolgáló {TE-03}	27,26	8,35	0,11	0,03
Windows 10 - Windows 2019 Server {TE-04}	111,14	110,44	73,46	28,15
Ubuntu 22.04 - ARDS fájlkiszolgáló {TE-05}	107,45	104,87	92,29	45,15
Ubuntu 22.04 - Windows 2019 Server {TE-06}	108,38	105,73	92,05	61,80

Az eredmények azt mutatják, hogy a Windows 10 - ARDS fájlkiszolgáló teszt esetében minden fájl méret tekintetében látványos teljesítménycsökkenés következett be, a két legkisebb fájl méretnél (4 MiB, 500 KiB) pedig a nullához közelített a sebesség. Az is jól látható, hogy ugyan a kisebb fájlok átviteli sebessége alacsonyabb volt az összes többi kliens-kiszolgáló párosítás esetében is, de a két kiszolgáló ezekben az esetekben nagyon hasonló teljesítményt nyújtott. A fájlkiszolgálókra történő írás során a nagyobb méretű fájlok (1 GiB, 256 MiB) eredményei nagyon hasonlóak voltak a kliens-kiszolgáló kombinációtól függetlenül, egy

eset kivéve. Az eredmények arra utalnak, hogy a Windows 10 kliens és az ARDS architektúra kombinációjában olyan implementáció- vagy protokollszerű kölcsönhatás jelentkezett, amely különösen a kis méretű fájlok írási műveleteit érintette. Mivel ez a jelenség a többi kliens–szerver konfigurációban nem volt megfigyelhető, feltételezhető, hogy nem az alkalmazott adatvédelmi mechanizmus általános teljesítményhatását tükrözi, hanem konfigurációs-specifikus sajátosság. A Windows 10 - ARDS fájlkiszolgáló ügyfél-kiszolgáló konfiguráció mindkét fájl méret esetében látványos teljesítménycsökkenést mutatott, a kisebb fájlok (4 MiB, 500 KiB) esetében a sebesség megközelítette a nullát. Az is megfigyelhető volt, hogy a fájl méret csökkenésével az átviteli sebesség minden más kliens-kiszolgáló párosítás esetében is csökkent, de a két kiszolgáló szinte azonos teljesítményt nyújtott. A Windows 10 átviteli sebességének drasztikus csökkenése “Unix-like” környezetekben egy széles körben ismert és jól dokumentált jelenség [16, 53, 54, 55, 56], ami a kiszolgáló oldali konfigurációs paraméterek finomhangolásával, kliensoldali rendszerbeállítások megváltoztatásával, illetve rendszerleíró adatbázis módosításokkal kerülhető el. Ezek a változtatások kliensoldalon tartományi környezetben akár központilag, csoportházirendeken keresztül is elvégezhetők.



23. ábra Írási teszteredmények összehasonlítása 1 Gbit/s sávszélesség esetén [forrás: saját]

5.8. Eredmények 10 Gbit/s-os kapcsolat esetén

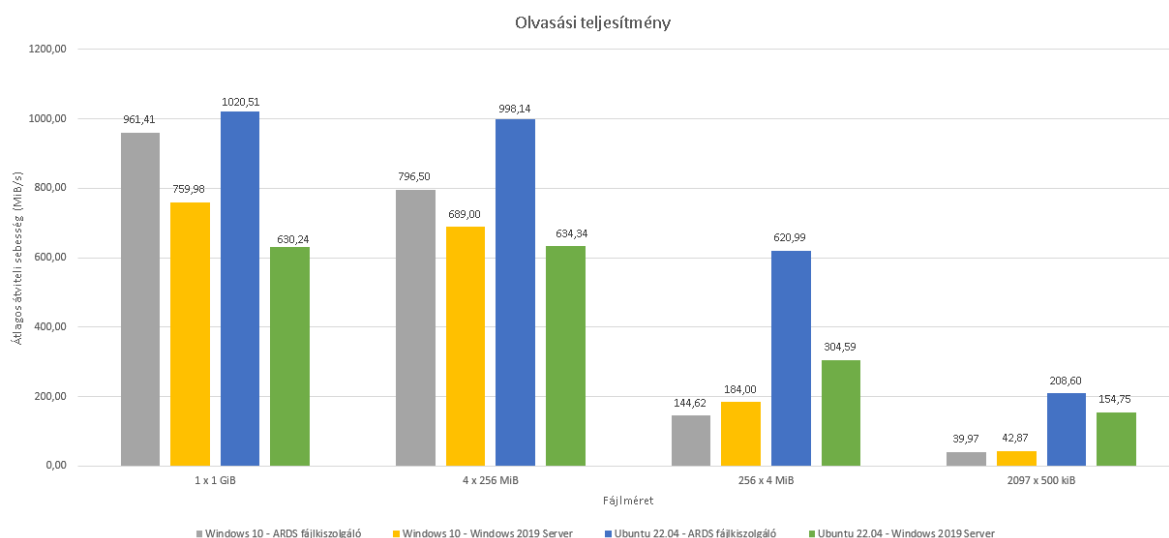
5.8.1. Az olvasási teljesítménytesztek eredményei

A 10 Gbit/s sebességű hálózaton végzett olvasási tesztek részletes eredményeit a 12. táblázat és a 24. ábra szemlélteti.

11. táblázat Olvasási teszteredmények 10 Gbit/s sávszélesség esetén

Teszteset	Fájlok darabszáma és mérete			
	1x1 GiB	4x256 MiB	256x4 MiB	2097x500 kiB
Átlagos átviteli sebesség MiB/s-ban				
Windows 10 - ARDS fájlkiszolgáló {TE-03}	961,41	796,50	144,62	39,97
Windows 10 - Windows 2019 Server {TE-04}	759,98	689,00	184,00	42,87
Ubuntu 22.04 - ARDS fájlkiszolgáló {TE-05}	1020,51	998,14	620,99	208,60
Ubuntu 22.04 - Windows 2019 Server {TE-06}	630,24	634,34	304,59	154,75

Az olvasási tesztek során az ARDS fájlkiszolgáló jellemzően gyorsabb volt. Olvasáskor azonban 4 MiB és 500 kiB méretű fájlok esetében a Windows 10 kliens a Windows 2019 kiszolgálóval az előbbi fájl méretnél 27,3 %-kal múlta felül a nyílt forráskódú fájlkiszolgálót, utóbbinál pedig 7,2 %-kal nyújtott jobb teljesítményt. Fontos megjegyezni, hogy a teljesen nyílt forráskódú kliens-szerver kombinációnál az átvitel gyorsabb volt, mint bármely más szcenárió esetén. Ilyen hálózati sebességnél a kliens jelentős I/O terhelésnek van kitéve, ami különösen a két kisebb fájl méret esetében a lemezre történő intenzív írási tevékenység miatt látványos teljesítménycsökkenést jelentett [57]. Megfigyelhető volt, hogy a fájlátviteli sebesség meghaladta a SATA III interfész elméleti maximális áteresztőképességét (750 MiB/s) több esetben is, ami a különböző gyorsítótárazási mechanizmusok miatt (SSD belső gyorsítótár, RAM, CPU adat gyorsítótár, NIC gyorsítótár, stb.) volt lehetséges, melyek különböző módon és eltérő sebességgel működtek.



24. ábra Olvasási teszteredmények összehasonlítása 10 Gbit/s sávszélesség esetén
[forrás: saját]

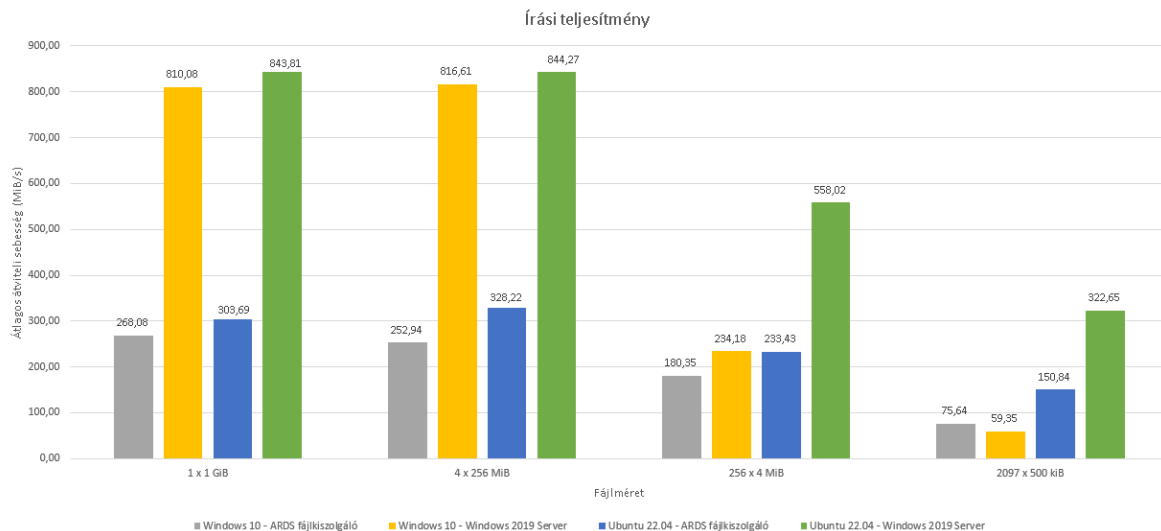
5.8.2. Az írási teljesítménytesztek eredményei

A 10 Gbit/s sebességű hálózaton végzett írási tesztek részletes eredményeit a 13. táblázat és a 25. ábra szemlélteti.

12. táblázat Írási teszteredmények 10 Gbit/s sávszélesség esetén

Teszteset	Fájlok darabszáma és mérete			
	1x1 GiB	4x256 MiB	256x4 MiB	2097x500 kiB
Átlagos átviteli sebesség MiB/s-ban				
Windows 10 - ARDS fájlkiszolgáló {TE-03}	268,08	252,94	180,35	75,64
Windows 10 - Windows 2019 Server {TE-04}	810,08	816,61	234,18	59,35
Ubuntu 22.04 - ARDS fájlkiszolgáló {TE-05}	303,69	328,22	233,43	150,84
Ubuntu 22.04 - Windows 2019 Server {TE-06}	843,81	844,27	558,02	322,65

Az Ubuntu 22.04 kliens - Windows 2019 kiszolgáló párosítás teljesített a legjobban a mérések során mind a négy fájl méret esetében. Az eredményekből kitűnik, hogy az ARDS fájlkiszolgáló Windows 10 klienssel párosítva hozzávetőlegesen háromszor lassabb, mint a Windows-kiszolgáló az 1 GiB és 256 MiB méretű fájlok esetében. A 4 MiB-os fájlok átvitelét tekintve a sebességkülönbség már sokkal kisebb volt. Az 500 kiB-os fájlokat tekintve pedig már több, mint 20 %-kal gyorsabb volt az átvitel a nyílt forráskódú fájlkiszolgáló vizsgálatánál. Fontos továbbá azt is megjegyezni, hogy a Windows szerveren nem volt engedélyezve az árnyékmásolat-szolgáltatás (VSS), ami hasonló többletterhelést jelentett volna a szerverre való íráskor, mint az ARDS fájlkiszolgálón konfigurált automatikus pillanatfelvétel készítése 2 percenként.



25. ábra Írási teszteredmények összehasonlítása 10 Gbit/s sávszélesség esetén [forrás: saját]

5.9. Referencia alapú teljesítményértékelés

A kiszolgálóoldali szoftverkörnyezetek teljesítményének egy egzaktabb és átláthatóbb összehasonlítási módja a korábban az 5.4.4 alfejezetben bemutatott és bevezetett referencia alapú teljesítménymutató. Ez a kalkulált metrika úgy áll elő, hogy elosztjuk egymással az átviteli idők kiszolgálónkénti átlagát ugyanazon kliensek és ugyanazon fájlméretek esetén. Mivel az ARDS fájlkiszolgáló teljesítményét szeretnénk értékelni a referencia környezethez képest, ezért az erre a kiszolgálóra vonatkozó átlagos átviteli időt osztjuk el a Windows 2019 kiszolgáló eredményeivel. Ebből természetesen az következik, hogy ha a hányados kisebb, mint 1, akkor a vizsgált megoldás jobban teljesített, ha pedig nagyobb, mint 1, akkor az rosszabb teljesítményt jelez. A számított referencia alapú teljesítménymetrikákat olvasás és írás esetén a 14. és a 15. táblázat foglalja össze.

13. táblázat Referencia alapú teljesítménymetrikák az 1 Gbit/s-os kapcsolaton elvégzett tesztek alapján

	1 x 1 GiB	4 x 256 MB	Fájlok 256 x 4 MB	2097 x 500 kiB
Olvasási teszteredmények				
Windows 7 - ARDS fájlkiszolgáló / Windows 7 - Windows 2019 Server	1.0723(≈)	0.9872 (+)	0.9703 (+)	0.9078 (+)
Windows 10 - ARDS fájlkiszolgáló / Windows 10 - Windows 2019 Server	0.9958(+)	0.9957 (+)	0.9873 (+)	0.9363 (+)
Ubuntu 22.04 - ARDS fájlkiszolgáló / Ubuntu 22.04 - Windows 2019 Server	1.0043(≈)	1.0075 (–)	0.9749 (+)	1,1715 (–)
Írási teszteredmények				
Windows 7 - ARDS fájlkiszolgáló / Windows 7 - Windows 2019 Server	1.0109 (≈)	1.0150 (≈)	1.0947 (≈)	0.8491 (+)
Windows 10 - ARDS fájlkiszolgáló / Windows 10 - Windows 2019 Server	4.0782 (–)	13.2265 (–)	683.8709 (–)	878.7343 (–)
Ubuntu 22.04 - ARDS fájlkiszolgáló / Ubuntu 22.04 - Windows 2019 Server	1.0085 (≈)	1.0072 (≈)	0.9973 (+)	1,3681 (–)
(+) Az ARDS fájlkiszolgáló jobban teljesített (≈) A Windows 2019 Server jobban teljesített, de a sebességkülönbség 10% alatti volt (–) A Windows 2019 Server jobban teljesített, és a sebességkülönbség meghaladta a 10%-ot				

14. táblázat Referencia alapú teljesítménymetrikák a 10 Gbit/s-os kapcsolaton elvégzett tesztek alapján

	1 x 1 GiB	4 x 256 MB	Fájlok 256 x 4 MB	2097 x 500 kiB
Olvasási teszteredmények				
Windows 10 - ARDS fájlkiszolgáló / Windows 10 - Windows 2019 Server	0.7809 (+)	0.8672 (+)	1,2733 (-)	1.0728 (≈)
Ubuntu 22.04 - ARDS fájlkiszolgáló / Ubuntu 22.04 - Windows 2019 Server	0.6134 (+)	0.6199 (+)	0.4947 (+)	0.7620(+)
Írási teszteredmények				
Windows 10 - ARDS fájlkiszolgáló / Windows 10 - Windows 2019 Server	3.0263 (-)	3.2113 (-)	1,2903 (-)	0.7893 (+)
Ubuntu 22.04 - ARDS fájlkiszolgáló / Ubuntu 22.04 - Windows 2019 Server	2.8602 (-)	2.7480 (-)	2.4110 (-)	2.1825 (-)
⁽⁺⁾ Az ARDS fájlkiszolgáló jobban teljesített ^(≈) A Windows 2019 Server jobban teljesített, de a sebességkülönbség 10% alatti volt. ⁽⁻⁾ A Windows 2019 Server jobban teljesített, és a sebességkülönbség meghaladta a 10%-ot.				

Az átviteli tesztek átfogó képet adtak a javasolt szerverkörnyezet teljesítményéről egy általánosan használt megoldáshoz viszonyítva. Az olvasási tesztekéből kapott 20 különböző átviteli időarányt megvizsgálva jól látszik, hogy az esetek túlnyomó többségében (14) a javasolt megoldás eltérő mértékben ugyan, de jobban teljesített, 3 esetben 10% alatti mértékben alulmaradt és csupán 3 esetben volt lényegesen lassabb az átvitel során. Az olvasás sebességének növelésére szolgáló módszerek Samba fájlkiszolgáló esetén jól dokumentáltak a vonatkozó szakirodalomban (pl.: socket options, sendfile, aio, aio read size, stb.). A gyakran inhomogén munkaállomások miatt, legfőképp a szerveroldali szolgáltatási paramétereket kell hangolni, vagy testreszabottan meghatározni minden környezetben az adott kliens operációs rendszerektől függően. Az írási tesztek során a legtöbb esetben (12) az átviteli sebesség jelentősen alacsonyabb volt a Windows kiszolgálóhoz képest, 5 esetben hasonló, de gyengébb eredmények születtek, míg 3 esetben gyorsabb volt a javasolt megoldás. Ez főként a kötetszintű titkosításnak volt köszönhető, ami a fizikai adatlopás megakadályozása érdekében lett engedélyezve, kis mértékben pedig a kötetszintű automatikus pillanatfelvételek készítésének, ami csak az adatok írása során jelentkezik. E két szolgáltatás integrálása természetesen hátrányt jelentett a Microsoft rendszeréhez képest, de mivel a legtöbb produktív Windows szerveren alapértelmezetten nincs bekapcsolva a fájltitkosítás (BitLocker) és az árnyékmásolatok készítése (VSS) csak a rendszerpartícióra engedélyezett, az volt a célom, hogy a vizsgálat a szokásos rendszergazdai gyakorlaton alapuljon. Az ARDS fájlkiszolgáló az olvasási tesztek 83,3%-ában jobban, vagy hasonlóan teljesített, míg az írási tesztek értékelve ez az arány 40% volt az 1 Gbit/s-os hálózati kapcsolat esetén. A 10 Gbit/s-os sáv szélesség esetén ugyanez az érték olvasásnál 87,5% volt, írási tesztek esetén pedig 12,5%. A vizsgálatok azt mutatták, hogy az ARDS fájlkiszolgáló olvasási teljesítménye általánosan jobb volt, mint a Windows 2019 környezeté, azonban utóbbi a 10 Gbit/s-os hálózati kapcsolat esetén íráskor jobban teljesített. Ugyanakkor az egyes írási tesztesetek során tapasztalt átviteli sebességcsökkenés a legjobb gyakorlatokra épülő megoldások használatával, illetve

a hivatalos dokumentáció iránymutatásai szerint eljárva hatékonyan kiküszöbölhető. Az eredményeket összegezve kijelenthető, hogy míg az olvasási műveletekhez csak minimális paraméterkorrekciókra van szükség, addig az írási műveleteknél jelentősebb sebességnövekedés szükséges ahhoz, hogy ellensúlyozzuk a terheléstöbbletet. Ebben a következő "best practice" megoldások segíthetnek:

- konfigurációs beállítások, gyorsítótárazási paraméterek finomhangolása szerveroldalon (pl.: socket options, write cache size, strict allocate, strict locking, stb.);
- rendszermemória növelése;
- IPv6 támogatásának eliminálása, amennyiben lehetséges;
- a pufferek méretének növelése;
- addicionális SSD/NVMe meghajtó telepítése a gyorsítótárazás támogatására;
- kliensoldali optimalizálás (registry, GPO, SMB paraméterek).

A javasolt megoldásom esetében fontos biztosítani, hogy az olyan szolgáltatások, mint például az ütemezett pillanatfelvételek készítése, a honeypot hálózati meghajtófigyelés, az IDS/IPS mechanizmusok, a vírusvédelem (valós idejű vagy ütemezett víruskeresés a hardverkapacitástól függően) ne okozzon jelentős teljesítménycsökkenést az általánosan használt megoldásokhoz képest. Az eredmények azt mutatják, hogy bár minimális paraméterkorrekcióra van szükség az olvasási teljesítmény javításához, az írási műveleteknél jelentős beavatkozás szükséges.

5.10. A releváns kutatási kérdések megválaszolása

Az első kutatási kérdésre (K-5) adott válasz az, hogy a jelenlegi konfigurációs szinten a mérési eredmények értékelése megerősíti, hogy a javasolt modell alapú rendszer az adatok olvasása esetén jellemzően nem marad el a referencia környezetként használt MS Windows 2019 Server megoldástól, azonban az írási műveletek vizsgálata alapján jelentős performancia növekedésre van szükség. Ehhez szerencsére jól dokumentált, gyakorlatban széles körben alkalmazott konfigurációs beállítások állnak rendelkezésre. Az elvégzett vizsgálatok alapján kijelenthető a második kutatási kérdés (K-6) kapcsán, hogy a különböző szolgáltatási platformok teljesítménye mért, illetve kalkulált metrikák alapján a jelentősebb teljesítménybeli különbségek szintjén összehasonlíthatók, azonban a részletes teljesítménykarakterisztikák elemzéséhez több metrikára, lokális teljesítménynaplókra és nagyobb időbeli felbontásra is szükség van. A harmadik kutatási kérdésre (K-7) határozott igennel válaszolhatunk, mivel a bevezetett referencia alapú teljesítménymetrika tetszőleges mérőszámok esetén alkalmazható és mindig megadja a vizsgált rendszer teljesítményét a referenciakörnyezethez képest. Ezzel a módszerrel egyszerre több heterogén rendszer teljesítménye is összehasonlítható arányszámokon keresztül.

6. Zsarolóvírus-tevékenység detektálása gépi tanulás segítségével szerveroldali, fájlműveleti naplófájlok alapján

A kutatás következő szakaszában a zsarolóvírusok detektálására irányuló vizsgálatokat végeztem el. Napjainkban számos antivírus megoldás és végpontvédelmi technológia áll rendelkezésre, azonban sok esetben ezeket megkerülve, illetve kiiktatva válik sikeressé a támadás. Mivel a zsarolóvírusok elsősorban a mindennapi működéshez elengedhetetlen adatvagyon titkosítják és teszik elérhetetlenné, a szervezetek napi működése a helyreállítás idejére részben vagy teljesen ellehetetlenedik. Gyors terjedésük és folyamatos technológiai fejlődésük miatt fontos, hogy koncepcionális változásokat vezessünk be a védelmi oldalon, új biztonsági mechanizmusokat fejlesszünk ki és többfaktoros észlelési rendszereket használjunk. A károk minimalizálása érdekében elengedhetetlen, hogy az ilyen jellegű támadásokat a lehető leghamarabb felismerjük. Szemléletes példa erre a LockBit [58] nevű kártevő, ami 4 perc 9 másodperc alatt 53,83 GB adatot (100 000 fájlt) titkosított különböző MS Windows munkaállomásokon, ami átlagosan 216,18 MB/s-os sebességnek felel meg. Bár a titkosítási sebességet erősen befolyásolják az olyan tényezők, mint a hálózati sávszélesség, a tárolási technológia, az interfészek típusa, a titkosítási algoritmus vagy a fájlrendszer, ezek a számok egyértelműen azt mutatják, hogy a hatékony védelemhez elengedhetetlen a gyors, lehetőleg közel valós idejű észlelés. A támadás fókuszában lévő adatvagyon jellemzően centralizáltan kerül tárolásra és a kliensek legtöbbször hálózati megosztás formájában érik el. Amíg a kliensoldali védelem sokszor kérdésesnek tekinthető, addig a szerveroldali fájlműveleti naplófájlok közvetlen és megbízható információt nyújtanak a fájlműveletekről. Ebből adódóan a szerveroldali naplófelügyelet is jó alapul szolgálhat a támadások észleléséhez. A különböző zsarolóvírus családok általában eltérő gyakorisággal hoznak létre, neveznek át és törölnek fájlokat típusuktól függően. Megfelelően tervezett és implementált naplózás esetén ezek a fájlművelet-szekvenciák akár nagy időbeli felbontás mellett (ns, nanosecundum) is hatékonyan tárolhatók.

Kutatási kérdések:

- K-8: Megvalósítható-e olyan MI alapú védelmi megoldás, ami kvázi valós időben képes reagálni egy zsarolóvírus-támadásra fájlműveleti naplók elemzési eredményei alapján?
- K-9: Létezik-e olyan naplózási módszer, amely lehetővé teszi az AI alapú védelem számára a hatékony felügyelt tanulást?
- K-10: Implementálható olyan fájlkiszolgáló védelmi mechanizmus, ami kvázi valós időben, a zsarolóvírus típusától függetlenül képes az adatvesztést minimalizálni a védett kiszolgálón?

6.1. A zsarolóvírusok elleni védelmi stratégiák áttekintése

Néhány releváns tanulmányt megvizsgálva kijelenthető, hogy a terület egyik fő kihívása a zero-day támadások felismerése, mivel a hagyományos, aláírás alapú módszerek nem képesek megbízhatóan azonosítani az új, vagy korábban más variánsként előforduló zsarolóvírus változatokat. A viselkedésfigyelési megközelítések -különösen azok, amelyek API (alkalmazásprogramozási interfész) híváskövetésre és rendszerfolyamat-elemzésre támaszkodnak, gyakran jelentős számítási erőforrásokat igényelnek, ami korlátozza alkalmazhatóságukat, ha a cél a közel valós idejű észlelés. Korábbi kutatások rámutattak a honeypot alapú rendszerek hiányosságaira is, mivel a legtöbb megoldás statikus honeypot fájlokra támaszkodik, amelyek hatástalanok a kifinomultabb kártevőkkel szemben. A támadók egyre gyakrabban alkalmaznak fejlett kriptográfiai technikákat (pl.: hibrid titkosítás), ami rendkívül megnehezíti a fájlok helyreállítására fókuszáló megoldások implementálását. Hasonlóképpen a hálózati forgalom folyamatos elemzése is korlátozottan lehet eredményes, mivel a támadók fejlett anonimizálási és pénzmosási technikákat (pl.: mixer szolgáltatásokat) alkalmaznak a forgalom és a Bitcoin tranzakciók elrejtésére. Éppen ezért a kutatók alternatív megközelítéseket igyekeztek találni a fenti korlátok leküzdésére. A gépi tanuláson (ML) alapuló módszerek, amelyek a viselkedési minták elemzésén alapulnak, ígéretesnek bizonyultak a zsarolóvírusok észlelési arányának javításánál. A viselkedéselemzés gépi tanulási algoritmusokkal történő támogatása (Guo, 2020 [59]), mint például a Long Short-Term Memory (LSTM), a Support Vector Machine (SVM), a Random Forest (RF) és a Feed Forward Deep Neural Network (Singh és mtsai, 2023, [60]) kombinációja felgyorsíthatja a rosszindulatú szoftverek azonosítását. Különösen a dinamikusan – akár MI által – generált honeypot fájlok és a mozgó célpont alapú védelmi (MTD) technikák szintén fontos szerepet játszhatnak a jövőben, mivel ezek folyamatosan módosítják a fájlneveket és formátumokat, hogy megzavarják a zsarolóvírus működését. A zero-day támadások elleni védelem megerősítése érdekében a Kaur és mtsai [61], Ahmad és mtsai [62], Singh és mtsai [63], Guisao és mtsai [64] kutatásai szerint javasolt a fájlok entrópiájának és hash-értékeinek folyamatos figyelemmel kísérése, valamint a Windows eseménynaplók és API-hívások mintázatának elemzése, ami segít az új támadási stratégiák gyors azonosításában. A fájlok helyreállításához továbbra is elengedhetetlen a titkosítási kulcsok megszerzése és a titkosítás feloldása, amelyhez olyan rendszerek nyújthatnak segítséget, mint a PayBreak és a Cryptographic Service Provider. A blokklánc alapú tranzakciókövetés pedig a zsarolóvírus operátorok azonosítására és pénzügyi tranzakcióik elemzésére jelenthet megoldást AlEroud és mtsai [65] korai munkája szerint. Az automatizált, gépi tanuláson alapuló megközelítések, mint például a Federated Learning, lehetővé teszik a szervezetek számára, hogy érzékeny adatok közvetlen megosztása nélkül fejlesszenek ki észlelési modelleket. A hibrid kriptográfiai rendszerek értékelése szintén ígéretes kutatási irányként merült fel, amely potenciálisan lehetővé teszi a dekódolást anélkül, hogy teljes mértékben ismernénk a támadók módszereit. Az online tranzakciók MI-alapú viselkedéselemzése szintén lehetőséget kínál a megfigyelés és az észlelés javítására.

6.1.1. Néhány kísérleti megoldás összehasonlítása

A legfrissebb kísérleti módszerek a védelmi mechanizmusok széles skáláját vonultatják fel. Az Early Ransomware Detection (2023) című tanulmány [66] az API-hívások és a fájlrendszer-monitoring kombinálásával foglalkozik, míg a ProtectNIC (2023) névre keresztelt megközelítés (Xu és mtsai, [67]) a hálózati forgalom elemzését kombinálta a SmartNIC technológiával, hogy minimális késleltetéssel kvázi valós idejű észlelést tegyen lehetővé. Az Applying Staged Event-Driven Access Control (2023) módszer, McIntosh és mtsai [68] egy automatizált, szabály alapú rendszert javasoltak a zsarolóvírus tevékenység minimalizálására, bár ez lényeges erőforrástöbbletet igényel. Denham és mtsai Analysis of Decoy Strategies (2023) [69] című tanulmányukban a honeypot alapú megoldások korlátait vizsgálták és arra a következtetésre jutottak, hogy a statikus tartalmak gyakran hatástalanok, míg a dinamikusan generált honeypot fájlok akár 75 %-kal is csökkenthetik a téves riasztások számát. A WannaLaugh (2024) nevű eljárás során (Diamantopoulos és mtsai, 2024, [70]) egy olyan kártevő emulátort fejlesztettek ki, amely valósághű szimulációkat biztosít, és lehetővé teszi a modellek gyorsabb tanítását. Az összehasonlító elemzés azt mutatja, hogy a különböző kutatások eltérő fókuszokkal rendelkeznek, de közös céljuk a ransomware támadások korai felismerése és hatékony megelőzése. A gépi tanulás alapú és viselkedéselemzésre épülő megoldások különösen ígéretesek, míg a SmartNIC technológia a hálózati szintű védelmet erősíti. A csali rendszerek fejlesztése és a dinamikusan generált honeypotok szintén hatékonyabbá tehetik a detekciós mechanizmusokat. A jövőbeli kutatások során az automatizált ML-modellek, a SmartNIC-alapú védelmek és a hibrid csali stratégiák kombinációja jelentheti a legjobb megoldást a zsarolóvírus-támadások elleni küzdelemben. A vizsgált módszerek összehasonlítását a 16. táblázat tartalmazza.

15. táblázat A vizsgált zsarolóvírus-detekciós módszerek összehasonlítása

Tanulmány	Fő fókusz	Fő hiányosság	Javasolt megoldás	Hatékonyság
Early ransomware detection (2023) [66]	API-hívások, fájlrendszerfigyelés	Zero-day támadások észlelése nehézkes	Viselkedéselemzés + ML kombinációja	90% feletti felismerés
ProtectNIC (2023) [67]	SmartNIC alapú detekció	Hálózati forgalom-elemzés időigényes	SmartNIC + ML kombinációja	Késleltetés nélküli valós idejű detekció
Applying Staged Event-Driven Access Control (2023) [68]	Hozzáférésvezérlés	Felhasználói interakció lassítja a folyamatot	Automatikus policy alapú védelem	Jobb döntési hierarchia, de erőforrás igényes
Analysis of Decoy Strategies (2023) [69]	Honeypot és csali rendszerek	Statikus csali fájlok hatástalansága	Dinamikus csali fájlok	75%-kal kevesebb hamis pozitív

WannaLaugh (2024) [70]	Ransomware emulátor	Reális teszt- környezet hiá- nya	Valóság-hű szimulációk	Modellek gyorsabb tanítása
---------------------------	------------------------	--	---------------------------	-------------------------------

Az eredmények azt mutatják, hogy a gépi tanulás, a viselkedéselemzés és a dinamikus decoy rendszerek együttes alkalmazása lehet a következő mérföldkő a hatékonyabb védelmi stratégiák irányába. A jövőbeli kutatások pedig az automatizált ML-modellekre, a SmartNIC-alapú védelmekre és a hibrid csali stratégiákra összpontosíthatnak a támadások észlelési hatékonyságának növelése érdekében.

6.1.2. A mesterséges intelligencia alapú módszerek áttekintése

A gépi tanulás (ML) és a mélytanulás (DL) [71, 72, 73] egyre fontosabb szerepet játszik a zsarolóvírusok elleni védekezésben, mivel ezek a módszerek képesek azonosítani az LLM alapú (pl.: Promptlock) és a zero-day sérülékenységeken alapuló támadásokat, amelyek ki-játsszák a hagyományos vírusirtó rendszereket [59]. A dinamikus észlelés során a gépi tanuló algoritmusok rosszindulatú mintákat keresnek [74] többek között a hálózati forgalomban [75, 76] (pl.: NetConverse). Hasonlóképpen az API-hívások figyelése is fontos eleme lehet a védelemnek. Az LSTM és az SVM modellek képesek előre jelezni a gyanús titkosítási folyamatokat. A Windows eseménynapló elemzése szintén segítheti a korai észlelést. Ilyen megközelítés például a DeepRanBiLSTM. A statikus elemzés további lehetőségeket is kínál [77], mivel az olyan gépi tanulási modellek, mint a Random Forest vagy a Gradient Boosting osztályozhatják a bináris struktúrákat (pl.: opkódok, bájtminák). A mélytanulási modellek közül például a konvolúciós neurális hálózatokkal (CNN) végzett Portable Executable (PE) fejléc elemzés szintén képes új zsarolóvírus családokat felismerni. A fájlaktivitás entrópia alapú elemzése gépi tanulással kombinálva, egy másik hasznos eszköz a titkosítási folyama-tok kezdeti stádiumban történő észlelésére [78]. Manzor és mtsai a pontosság további javí-tása érdekében a hibrid módszerek statikus és dinamikus elemzési módszereket kombinálnak [79]. Az E2E-RDS (End-to-End Ransomware Detection System) ezeket a megközelítéseket ötvözi a felismerés megbízhatóságának javítása érdekében. Kim és mtsai megoldása hasonló hibrid megoldás, amikor a hozzáférés-vezérlési megoldásokat és a gépi tanulási algoritmu-sokkal történő monitorozást kombináljuk, ami lehetővé teszi a jogosulatlan fájlműveletek azonosítását [80]. Zahoor és mtsai [81] egy költségérzékeny „pareto ensemble” mélytanu-lási modellt fejlesztettek ki a zero-day alapú zsarolóvírusok észlelésére, bemutatva ezzel az ensemble architektúrákban rejlő potenciált. Baldwin és Dehghantanha [82] SVM alapú mo-delleken keresztül vizsgálták a statikus opkód sűrűség elemzésének lehetőségét, ezzel is ki-emelve a statikus megközelítések folyamatos relevanciáját. A csalifájlok és a honeypotok használatánál az MI alapú csapdák elterelhetik a támadókat és megvédhetik a valódi adatokat a titkosítástól. Az előbbieken tárgyalt megközelítések nemcsak felgyorsítják az észlelést, hanem javítják az új fenyegetésekkel, például a zero-day támadásokkal szembeni ellenálló képességet is. Az adaptív modellek - különösen azok, amelyek a hálózati forgalmat elemzik

és az API hívásokat figyelik - szintén jelentősen javíthatják a védelmet. Ugyanakkor az ilyen modellek folyamatos tanítása és hangolása továbbra is elengedhetetlen a megfelelő hatékonyság fenntartásához.

6.1.3. Egyéb releváns megközelítések

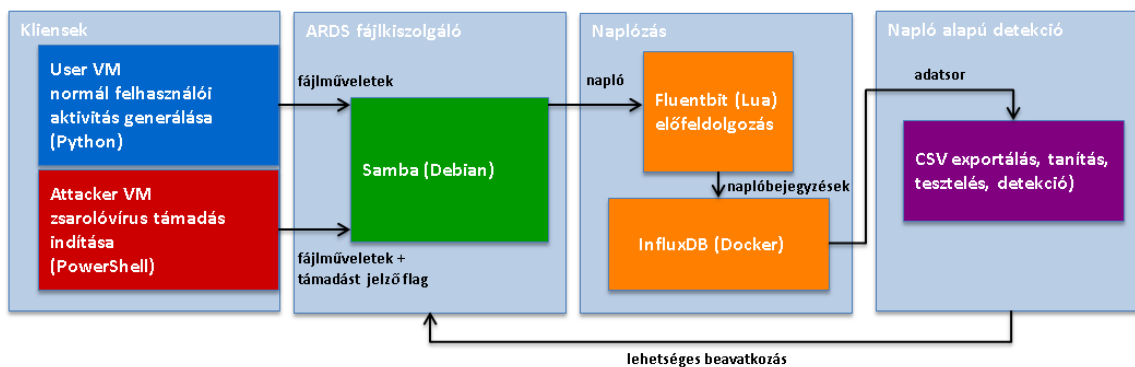
Egyes tanulmányok más megközelítéssel vizsgálták a zsarolóvírusok észlelésének lehetőségeit. Ilyen módszer például a kernel és I/O szintű monitorozás, vagy a hypervisor szintű viselkedéselemzés. Peeler és mtsai [83] egy gyors reakcióidővel és alacsony erőforrás-fogyasztással rendelkező rendszert javasoltak, amelynek alapvető jellemzői összhangban vannak az ARDS modell kapcsán javasolt megoldással. Hitaj és mtsai „Minerva” [84] elnevezésű rendszerének az észlelést megakadályozó technikákkal szembeni ellenállóképessége inspirálta a robusztusság-orientált megközelítést. Herrera-Silva és mtsai [85] dinamikus funkciókészletet vezettek be a folyamatosan fejlődő zsarolóvírus minták, például a fájl I/O és a folyamatindítások gépi tanuláson alapuló észleléséhez, ami szintén illeszkedik a saját kutatásomban vizsgált fájlműveleti napló alapján történő megközelítéshez. Lee és mtsai [86] a formátummegőrző titkosítás (FPE) esetén végzett észlelési lehetőségeket vizsgálták és egy olyan módszert mutattak be, amely kiegészíti a napló alapú megközelítéseket. Zhou és mtsai [87] bemutatták az ANIMAGUS nevű rendszerüket, ami az I/O alapú észlelési mechanizmus robusztusságát teszteli támadási körülmények között, ezzel is tovább hangsúlyozva az éles körülmények között történő rendszertervezés szükségességét. Számos referenciatanulmány következetesen bizonyította, hogy a modern AES implementációk rendkívül magas átviteli sebességet érnek el a különböző platformokon, ezzel megerősítve azt a tényt, hogy a titkosítási sebesség önmagában nem alkalmas a zsarolóvírus tevékenységek megkülönböztetésére. Például az AES-NI-vel optimalizált AES-GCM a mai CPU-kon 1,5–2,5 ciklust ér el bájtonként, amint azt Kampanakis [88], illetve Gueron és szerzőtársai [89] is megerősítették. A többmagos processzorok Barnes és mtsai mérései [90] szerint közel 7 Gb/s-os teljesítményt biztosítanak, valamint még az alacsony kategóriájú mikrokontrollerek is több száz ciklust érnek el bájtonként Kim és mtsai kutatásai alapján [91]. A GPU-val gyorsított AES-CTR implementációk tovább növelik a titkosítás sebességét, ahogy ezt Assa-Agyei [92] 2023-ban kutatásai eredményeként publikálta. Ezek az eredmények mind alátámasztják, hogy a titkosítási sebesség önmagában valóban nem szolgálhat megbízható megkülönböztető jellemzőként a zsarolóvírusok felismerésénél. A fent említett megközelítések fényében az ARDS modell részeként javasolt észlelési/detekciós eljárás illeszkedik a nemzetközi kutatási irányokhoz és azok eredményeihez, mivel szerveroldali, naplóalapú gépi tanulási megközelítésként pozicionálható, amely hatékonyságot, robusztusságot és jelentős előremutató eredményként a fájlkiszolgáló-védelem időkritikus összetevőjében a kvázi valós idejű észlelés gyakorlati megvalósíthatóságát támasztja alá.

6.2. A fájlművelet-szekvenciákon alapuló gépi tanulás módszerének bemutatása

A zsarolóvírus-támadások fájlművelet-szekvencia alapú detekciójának problémája tulajdonképpen egy bináris osztályozási feladatnak tekinthető. Ezt a bináris osztályozási feladatot oldom meg felügyelt tanulás segítségével. A hiperparaméter hangolást követően hasonlítom össze a három klasszikus (Random Forest, Support Vector Machine, Decision Tree) és a kettő ensemble (Adaptive Boosting, eXtreme Gradient Boosting) gépi tanulási modellt különböző metrikák (osztályozáshoz szükséges idő (s, sec), pontosság, F1-pontszám, specificitás, érzékenység) és a keveredési mátrixok alapján. Fontos kritérium, hogy a feladat megoldása során nem azonosítjuk a vírus pontos típusát, kizárólag a támadás tényét jelezzük. A kitűzött cél, hogy minél nagyobb hatékonysággal azonosítsuk a támadással érintett időszakokat kvázi valós időben klasszikus gépi tanuló, illetve ensemble modellek segítségével. A felügyelt tanulás végrehajtásához szükséges adatsort szimulációs környezet segítségével hoztam létre, melynek részletes módszertanát a következő alfejezet tartalmazza.

6.2.1. A szimulációs környezet felépítése és az adatsor előállítása

Mivel produktív környezetekből a megfelelő formátumú, zsarolóvírus-támadások nyomait is tartalmazó fájlműveleti naplófájlok beszerzése egyrészt korlátozottan lehetséges, másrészt a különböző környezetekből származó naplók eltérő fájlhasználati szokásokat és csupán egyetlen támadás részleges fájlműveleteit tartalmaznák, így a reprodukálhatóságot és a hangsúlyozhatóságot szem előtt tartva virtualizált infrastruktúra felépítése mellett döntöttem. Ahhoz, hogy megközelíthessem a valós működési körülményeket, illetve a gépi tanuláshoz szükséges adatsort kinyerhessem, egy tipikusnak mondható (10-50 dolgozó), KKV léptékű infrastruktúrára volt szükség. A kialakított környezet modelljét és működési folyamatait a 26. ábra szemlélteti.



26. ábra A vizsgált környezet modellje és annak munkafolyamatai [forrás: saját]

A szimuláció egy fizikai kiszolgálón (DELL Precision T3610-Intel Xeon E5-1620, 32 GB RAM, Apacer AS350x 512GB SSD), Proxmox VE 8.1 (QEMU, KVM, LXC) virtualizációs platform használatával történt. Négy különböző felhasználó fájlműveleti sablonja (eltérő on-site és home office aktivitással) eredményezi a normál működés szimulációját, ami tulajdonképpen véletlen időpillanatokban végrehajtott fájl létrehozás, -átnevezés és -törölés műveleteken keresztül történik. Ezekkel a műveletekkel párhuzamosan, minden 1 és 6 óra közötti időintervallumban egy véletlen időpillanatban indul el az öt különböző, teljes titkosítást alkalmazó zsarolóvírus (Ryuk, WannaCry, NotPetya, Lockbit, TeslaCrypt) egyike. Ezzel egyrészt elkerülöm a torzítást, másrészt igyekszem tükrözni a valós támadások véletlen időbeli lefutását. Az infrastruktúrát három virtuális gép és az őket összekötő szegmentált hálózat alkotta. Egy Debian 12.4.0 Bookworm operációs rendszerre telepített Samba 4.19.3 fájlkiszolgáló biztosította a centralizáltan tárolt fájlok elérését a két Windows 10 Professional 22H2 kliens számára hálózati meghajtókon keresztül. A megfelelő formátumú naplózás unix socket-en keresztül, Fluentbit (Lua modul) felhasználásával InfluxDB idősorfolytonos (nanoszekundumos pontosság) adatbázisba történt, melyet Docker konténer biztosít. Az adatbázisban minden rekordhoz (egy másodpercre aggregált értékek) a következő attribútumokat tároljuk:

- időbélyeg (UTC RFC3339)
- nc (egész típusú mező, a létrehozott fájlok száma)
- nr (egész típusú mező, az átnevezett fájlok száma)
- nu (egész típusú mező, a törölt fájlok száma)
- att (egész típusú mező, ami a támadás ideje alatt 1-et, minden más időpontban 0 értéket vesz fel)

Az egyik Windows munkaállomás (user) a korábban említett négy különböző fájlhasználati séma alapján (17. táblázat) jelenléti és otthoni munkavégzést szimulált a felhasználói profiloknak megfelelően. A felhasználói profilok működésének szimulációja során négy, folyamatosan futó Python szkript időzítette és hajtotta végre a fájlok létrehozását, átnevezését és törlését.

16. táblázat Fájlműveleti sablonok

munkakör/munkavégzés	időintervallum	fájl létrehozás	fájl átnevezés	fájl törlés
dolgozó jelenléti	8:00 – 16:00	30 - 60	10 - 30	5 - 20
dolgozó otthoni	16:00 – 8:00	20 - 60	10 - 20	5 - 10
titkárság jelenléti	8:00 – 16:00	30 - 60	10 - 20	5 - 20
titkárság otthoni	16:00 – 8:00	20 - 30	5 - 10	5 - 10
rendszergazda jelenléti	8:00 – 16:00	30 - 50	10 - 30	10 - 20
rendszergazda otthoni	16:00 – 8:00	0 - 20	10 - 20	0 - 20

ügyvezető jelenléti	8:00 – 16:00	0 - 10	0 - 5	0 - 5
ügyvezető otthoni	16:00 – 8:00	0 - 5	0 - 2	0 - 2

A támadást szimuláló Windows 10 hoszt (attacker) Powershell szkript segítségével 1 és 6 óra közötti véletlen késleltetést alkalmazva az öt zsarolóvírus egyikét indította el. Ezzel egyidejűleg a Proxmox VE kiszolgálón konténerizált formában futó InfluxDB "attack" bucket-jének "att" mezőjébe HTTP API-n keresztül, natívan időbélyeggel ellátva 1-es értéket inzerál a szkript. Mivel a fájlkiszolgálón minden támadást követően hat különböző típusú fájlt (docx, pdf, ppt, pptx, xlsx, zip), összesen 45 534 809 bájt méretben állítunk helyre és a felhasználói szkriptek által létrehozott docx formátumú fájlok mérete is elhanyagolható (fájlönként 30 162 byte), a titkosítási folyamat gyorsan lezajlik. A Proxmox VE kiszolgálón háttérben futó BASH szkript periodikusan, másodpercenként, NFS protokollt használva megvizsgálja, hogy a monitorozásra használt "b.docx" nevű fájl létezik-e. Ez a fájl a felhasználói szkriptek átnevezési és törlési függvényeiben kivételként van megjelölve, így csak vírus-támadás során módosulhat. Amennyiben a monitorozott állomány nem található, a monitorozó szkript a fent említett bucket "att" mezőjébe HTTP API-n keresztül 0 értéket inzerál, vírusmentes pillanatképre állítja vissza az "attacker" hosztot, majd törli a megosztás teljes tartalmát és visszaállítja az eredeti állományokat. A felügyelt tanulást úgy tesszük lehetővé, hogy az „att” mező 1-0 átmenetei a támadás elejét, míg a 0-1 átmenetek a támadás végét jelzik. Ez azt jelenti, hogy az "att" mező 0-1 és 1-0 változásai közötti időszakban zsarolóvírus-támadás zajlik, míg a többi időpillanat támadással nem érintett. Végül egy olyan adatsor állt elő, amely egyaránt tartalmazott normál felhasználói és rosszindulatú kód által végrehajtott fájlműveleteket is.

6.2.2. Adattisztítás és előfeldolgozás

Ahhoz, hogy a felügyelt tanulás folyamata hatékony lehessen, célszerűen megválasztott attribútumokra (features) volt szükségem. Mivel a célom, hogy fájlművelet-szekvenciák alapján jelezzem a támadást, elsőként a jellemző fájlműveleteket kellett kiválasztani. A zsarolóvírusok ugyan különböző stratégiákat, fertőzési mechanizmusokat és titkosítási eljárásokat alkalmaznak, azonban a hagyományos kriptó vírusok esetében elengedhetetlen -típustól független- a fájlok eltérő mértékben történő létrehozása, átnevezése, illetve törlése rövid időn belül. Ezen műveletek monitorozásához megfelelő pontosságú - esetünkben nanoszekundum - időbélyegre is szükségünk van, hogy később az alkalmazott időablakok tetszőlegesen módosíthatóak legyenek. A szimuláció 2024.06.06 17:00-tól 2024.07.06 17:00-ig tartott és mivel nem tettem különbséget a munkanapok, illetve a hétvégék között, így végül pontosan 30 munkanapra vonatkozó fájlműveleti napló állt rendelkezésünkre. Úgy gondoltam, hogy ez az időtartam elegendő arra, hogy szemléltethető legyen a felhasználói profilokból adódó bizonyos mértékű ciklikusság, ugyanakkor a véletlenszerű műveletek és a támadások is megfelelő számban lesznek jelen az adatsorokban. A fájlkiszolgálón a naplóbejegyzéseket "single node" InfluxDB adatbázis backend tárolta, ami hatékonyan használja a háttértárat

(SSD) és a memóriát a rekordok tárolására és így elméletben 250 000 értéket képes tárolni másodpercenként. Ez a mi esetünkben nagyságrendileg 50 000 rekordot jelent másodpercenként, ami kielégíti a kvázi valós idejű észlelés adatbázis követelményét. Lekérdezések (Influx Query Language) segítségével a fájlműveleti naplóbejegyzéseket és a támadásokat tartalmazó bucket-ek adatait felhasználva, másodperces időablakra vonatkozóan számláltam meg a műveleteket, illetve generáltam le az eseménnyel nem érintett másodpercekhez tartozó rekordokat. Ezt követően a lekérdezés eredményét hagyományos CSV fájlformátumba exportáltam. A nagyszámú, kizárólag nullákat tartalmazó rekordok („0,0,0,0”) a vizsgált gépi tanulási modellek esetében nem hordoznak számunkra hasznos információt, azonban a tanulási folyamatot lassítják, illetve kedvezőtlenül befolyásolhatják, így a közvetlenül egymás után előforduló, kizárólag nullákat tartalmazó rekordból mindig legfeljebb csak egy került bele az adatsorokba. Egy-egy ilyen rekord minden esetben egy összefüggő, fájlműveleti szempontból eseménytelen időszakot jelöl. Mivel a zsarolóvírus-támadások elindítása és a támadás következtében a fájlkiszolgálón észlelt fájlműveletek kezdete között néhány másodpercnyi különbséget észleltem (a vírus először inicializálja a folyamatait, felderíti a helyi és hálózati erőforrásokat, megpróbálja felvenni a kapcsolatot a "C2" szerverrel, stb.), ami félrevezetheti a gépi tanuló algoritmusokat, így eliminálnom kellett az adatsorból azokat a rekordokat, amiknél a támadás állapota aktív volt, azonban közben semmilyen fájlművelet nem történt ("0,0,0,1"). Az abszolút időbélyegeket nem használtam közvetlenül a felügyelt tanítás során, mivel nem memóriával bővített neurális hálókkal dolgoztunk. Az időbélyegek kizárólag az idősorfolytonosság szemléltetésére szolgáltak. Az aggregált naplóadatokból álló adatkészletet 80-10-10 arányban osztottam fel tréning, teszt és validálási adathalmazokra a modellek tanítása, tesztelése és validálása céljából. A 18. táblázat a tesztadatsorból („test_onezero_nozeroatt.csv”) kiragadott, indexekkel kiegészített mintán keresztül szemlélteti az adattisztítás eredményét.

17. táblázat Tesztadatsor minta

In-dex	Időbélyeg	nc	nr	nu	att
169	2024-06-07 05:46:24+00:00	2	0	0	0
170	2024-06-07 05:46:25+00:00	0	0	0	0
171	2024-06-07 06:00:10+00:00	0	0	2	0
172	2024-06-07 06:00:11+00:00	0	0	0	0
173	2024-06-07 06:02:01+00:00	0	0	1	0
174	2024-06-07 06:02:02+00:00	0	0	0	0
175	2024-06-07 06:02:31+00:00	1	0	0	0
176	2024-06-07 06:02:32+00:00	0	0	0	0
177	2024-06-07 06:10:50+00:00	5	2	1	1
178	2024-06-07 06:10:59+00:00	1	2	0	1
179	2024-06-07 06:11:00+00:00	19	17	0	1
180	2024-06-07 06:11:46+00:00	0	0	21	1
181	2024-06-07 06:11:59+00:00	0	0	0	0
182	2024-06-07 06:15:07+00:00	0	1	0	0
183	2024-06-07 06:15:08+00:00	0	0	0	0

184	2024-06-07 06:20:09+00:00	0	0	1	0
185	2024-06-07 06:20:10+00:00	0	0	0	0

6.2.3. Gépi tanulási modellek kiválasztása, a hiperparaméter hangolás célja

A korábbi tapasztalatokra építve és a kvázi valós idejű észlelés követelményének fontossága miatt a klasszikus és az ensemble gépi tanulási modellek kerültek fókuszba a vizsgálatok során. A mélytanulási módszerek teljesítményértékelése nem volt cél, mivel az adatsor egyszerű attribútumokból állt és a célfüggvény szintén nem volt komplex, azonban ezeknek a neurális hálózatoknak mind az erőforrás-szükséglete, mind pedig az összetett, többrétegű architektúrából adódó tanítási ideje nagyobb. A kiválasztott modellek jól alkalmazhatók közepes összetettségű bináris osztályozási problémákra, ahol mind a számítási hatékonyság, mind a prediktív megbízhatóság elengedhetetlen. Tekintettel az attribútumok kompakt reprezentációjára - melyet az adatsor megfelelő előállításával, az adatszűréssel és az előtisztítással együttesen sikerült elérni az 6.2.1 és az 6.2.2 fejezetekben bemutatott módon - a magas dimenziójú adatokat igénylő mélytanulási megközelítések helyett a robusztusságukról és hatékonyságukról ismert modelleket választottam. A klasszikus gépi tanulási modellek, mint például az SVM, illetve a DT (Decision Tree) ugyan számításigényesek, ellenben hatékonyan kezelik a kis adathalmazokat és kevesebb hiperparamétert használnak. Az ensemble módszerek, ideértve értve az RF, az ADA (Adaptive Boosting) és az XGB (Extreme Gradient Boosting) modelleket, gyakran kiváló pontosságot érnek el több gyenge tanuló algoritmus előrejelzéseinek összesítésével. Különösen a faalapú ensemble modellek (RF, XGB) hatékonyak a heterogén jellemzőeloszlások kezelésében, míg a boosting típusú megközelítések (ADA, XGB) erősek a variancia és a torzítás csökkentésében. Ezek a módszerek általában jobban tolerálják a zajt és a kiugró értékeket, valamint kifejezetten kiegyensúlyozatlan adatkészletek kezelésére hangolhatók. Mindezeket figyelembe véve esett a választás az öt fent említett gépi tanulási modellre a teljesítmény összehasonlítása során. Minden modell tanítása rétegzett mintavétellel történt, hogy a normál működési és a támadással érintett időtartamok aránya mind a tréning, mind a teszt adatsoroknál megmaradjon. A hiperparaméterek optimalizálása során a GridSearchCV esetében alapértelmezett keresztvalidációt alkalmaztam. A GridSearchCV, mint automatizált hangolási folyamat biztosította, hogy a hangolt modellek optimális hiperparaméter konfigurációt érjenek el, miközben minden alkalommal másik adathalmazt használ validációs adatsorként. Ez a módszer robusztus és megbízható becslést ad a modell teljesítményéről, csökkenti a túlillesztés kockázatát, ugyanakkor növeli a számítási igényt. A hiperparaméterek hangolásának célja elsősorban a “true positive” felismerési arány maximalizálása volt, miközben minimalizáltam a “false positive” eseteket, ezáltal javítva az érzékenységet anélkül, hogy romlott volna az általános pontosság.

6.2.4. Értékelési metrikák, az összehasonlítás menete

Az egyes modellek teljesítményértékeléséhez hagyományos osztályozási teljesítménymetrikákat alkalmaztam. Minden modell esetében a teljesítményt a következő mutatók alapján értékeltem:

- **Pontosság:** a helyesen osztályozott minták aránya
- **F1-pontszám:** a pontosság és a visszahívás (recall) harmonikus átlaga, amely egyensúlyt teremt a hamis pozitív és hamis negatív eredmények között
- **Specifitás (igaz negatív arány):**

$$\frac{TN}{TN+TP}$$

- **Érzékenység (igaz pozitív arány / visszahívás):**

$$\frac{TP}{TP+FN}$$

Ezek a metrikák együttesen átfogó képet adnak a teljesítményről. Különösen fontos közülük az érzékenység, mivel egy fel nem fedezett zsarolóvírus-támadás (hamis negatív támadási állapot) súlyos következményekkel járhat. A specifitás kiegészíti ezt azzal, hogy a megbízhatóságot értékeli a helyes osztályozásra vonatkozóan, míg az F1-pontszám egyensúlyt teremt a pontosság és a visszahívás között. Ezek a mutatók kiegészítő perspektívákat nyújtanak az osztályozási teljesítményre vonatkozóan, figyelembe véve mind a támadások felismerésének képességét, mind a különböző kimeneteli eloszlások mellett a támadási állapot megállapítására vonatkozó megbízhatóságát. A bináris osztályozáshoz szükséges idő mindemellett tükrözte a modellek sebességét, azaz egy adott minta osztályozásához szükséges késleltetést. Ez a metrika jól kifejezi, hogy az adott modell alkalmas-e a kvázi valós idejű osztályozásra vagy sem. A keveredési mátrixok segítségével az előző eredmények tükrében az eredmények (TP, FP, TN, FN) részletes megoszlását lehetett megfelelően megjeleníteni. Az említett mérőszámok átfogó képet adtak az észlelés hatékonyságáról, figyelembe véve nemcsak az általános pontosságot, hanem az érzékenység, a specifikusság, valamint a hamis pozitív és hamis negatív eredmények közötti egyensúly közötti kompromisszumokat is, amelyek kritikus fontosságúak a zsarolóvírusok észlelésének szempontjából.

6.2.5. ROC és pontosság-visszahívás pontok elemzése

A bemutatott metrikákon túlmenően ROC-pont és Precision-Recall (PR) pontelemzéseket is elvégeztem az osztályozások teljesítményeinek vizualizálása érdekében. Mivel ez egy bináris osztályozási feladat, az összes modellt az alapértelmezett 0,5-ös döntési küszöbértékkel értékeltem, ami biztosította az osztályozók közötti konzisztens és összehasonlítható eredményeket. Mivel minden modellt fix döntési küszöbértékkel értékeltem, ezért az eredmények

nem teljes görbét, hanem egyetlen működési pontot adnak. A ROC-térben a valódi pozitív arányt (TPR) a hamis pozitív arányhoz (FPR) viszonyítva ábrázoltam, míg a PR-térben a pontosságot a visszahíváshoz viszonyítva ábrázoltam. Ezek a pontok szemléltetik az érzékenység, a specificitás, a pontosság és a visszahívás közötti kompromisszumokat.

6.2.6. Statisztikai validálás

A kísérleti eredmények statisztikailag validálhatóak, mivel az adatgyűjtési folyamat kialakítása során biztosított volt, hogy az öt különböző zsarolóvírus (Ryuk, WannaCry, NotPetya, Lockbit, TeslaCrypt) véletlenszerűen fut le a 30 napos időszak alatt egy és hat óra közötti időintervallumokban. Ez az eljárás garantálta a támadások (fájlművelet-szekvenciák mintázata) és a támadások időpontjai közötti elfogulatlan mintavételt. Ennek a kísérleti stratégiának az elsődleges célja nem az volt, hogy a különböző vírusokat egymással összehasonlítsa hanem, hogy biztosítsa a felügyelt tanuláshoz, a hiperparaméter hangoláshoz és a teljesítmény értékeléséhez szükséges adatsort. Természetesen meg kell jegyezni, hogy valós forgatókönyvek szerint csak egyetlen zsarolóvírus futna le egyszer, ezzel elérhetetlenné téve a felhasználói fájlokat. Az eredmények következetesen azt mutatták, hogy függetlenül attól, hogy melyik zsarolóvírus típus indult el, vagy milyen gyakorisággal történt a támadás, a gépi tanulási modellek a ransomware tevékenységet hasonló arányú igaz és hamis előrejelzésekkel észlelték, mint azt a keveredési mátrixok is mutatják.

6.2.7. Költség alapú összehasonlítás

Költség alapú elemzésre is szükségem volt, hogy figyelembe tudjam venni a hamis pozitív és hamis negatív eredmények egyenlőtlen hatását. A keveredési mátrix eredményei alapján egy költségmátrix került meghatározásra, amelyben $C_{FP} = 1$ és $C_{FN} = 10$, míg $C_{TP} = C_{TN} = 0$. Ez tükrözi azt a racionális megközelítést, hogy egy fel nem fedezett zsarolóvírustámadás (FN) sokkal költségesebb, mint egy téves riasztás (FP). Minden modell teljes költsége kiszámításra és összehasonlításra került. Ez a megközelítés további döntési kritériumot biztosított, kiegészítve a metrikákat azzal, hogy prioritást adott azoknak a modelleknek, amelyek minimalizálják az általános működési kockázatot, mivel a fel nem fedezett zsarolóvírustevékenység (FN) lényegesen költségesebb, mint egy téves riasztás (FP).

6.2.8. A modellek tanítása, hiperparaméter hangolás és kiértékelés

A referenciaértékek megállapítása érdekében minden modell tanítása először az alapértelmezett hiperparaméterekkel történt. Az optimalizáláshoz meghatároztam a célfüggvényt, a validációs technikát és a hiperparaméter keresési stratégiát. A hiperparaméterek hangolását GridSearchCV segítségével végeztem, amely alapértelmezés szerint ötszörös keresztellenőrzéssel, kimerítően feltárja a hiperparaméterek terét. Ez a stratégia robusztus és elfogulatlan

teljesítménybecslést biztosít, bár a számítási költségek megnőnek. Az optimalizálás célfüggvénye az érzékenység (igaz pozitív arány maximalizálása) volt, tükrözve a zsarolóvírus-felismerés során a hamis negatív esetek magasabb költségét. A hiperparaméterek keresési terei feltáró futtatások alapján lettek meghatározva, hogy elkerülhetőek legyenek a szélsőséges értékek. A végleges konfigurációkat annak alapján választottam ki, hogy mennyire képesek javítani az érzékenységet anélkül, hogy túlzottan rontanák a pontosságot. Az optimális paraméter értékek az 6.3.1 alfejezetben (20. táblázat) kerülnek ismertetésre. A kiértékelés során a metrikák között szerepelt a pontosság, az F1-pontszám, az érzékenység (visszahívás), a specifitás és az összehasonlítást megkönnyítő keveredési mátrix. Ezeket a standard mutatókat kiegészítve, két további értékelési módszert is alkalmaztam:

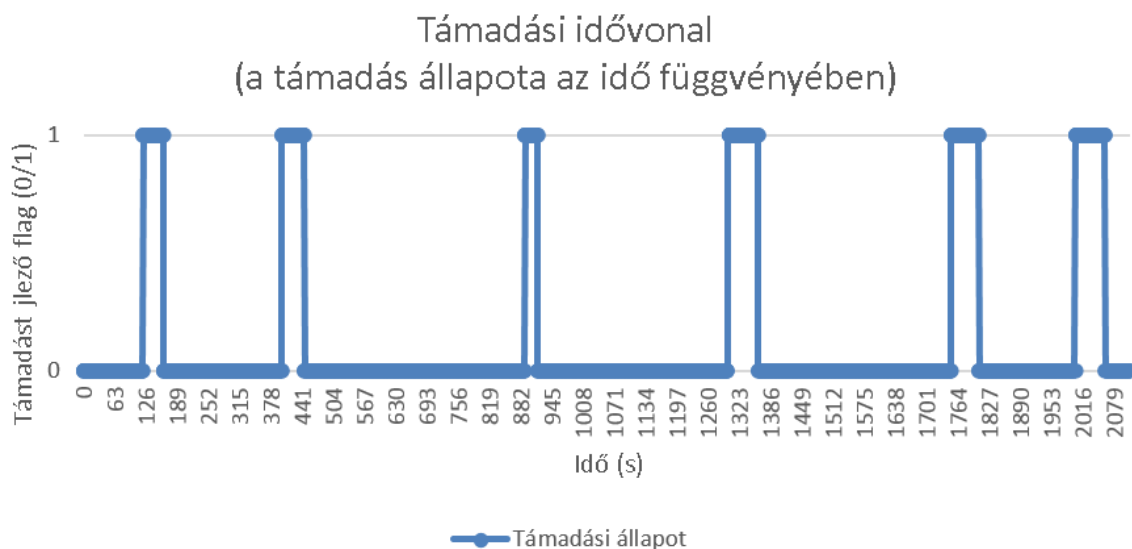
- **ROC és Precision-Recall pontelemzés:** mivel az összes modellt bináris osztályozóként értékeltem, 0,5-ös rögzített döntési küszöbértékkel, mindegyik egyetlen működési pontként jelent meg a ROC- és PR-terekben. Ezek a pontok szemléltetik az érzékenység, a specifikusság, a pontosság és a visszahívás közötti kompromisszumokat. Minden modellt egyetlen működési pont képviselt a megfelelő térben, a keveredési mátrixból származtatott TPR, FPR, Precision és Recall értékek alapján.
- **Költség alapú értékelés:** bevezettem egy költségmátrixot, amivel figyelembe vettem a téves osztályozások aszimmetrikus hatását. Míg az igaz pozitív és az igaz negatív esetek nem járnak költségekkel, addig a hamis pozitív mérsékelt költségeket okoz a téves riasztások miatt, a hamis negatív esetek pedig magas költségeket okoznak a fel nem fedezett zsarolóvírus tevékenységek miatt (22. táblázat). Az elemzés eredményeit a 6.3.4 alfejezetben (19. táblázat) mutatom be.

18. táblázat A költségmátrix szemléltetése

	Előre jelzett pozitív	Előre jelzett negatív
Tényleges pozitív	TP: 0 (helyes felismerés)	FN: C_{FN} (magas költség,- fel nem ismert támadás)
Tényleges negatív	FP: C_{FP} (közepes költség - téves riasztás)	TN: 0 (helyes felismerés)

Ez az átfogó keretrendszer biztosította, hogy a modelleket ne csak a standard mutatók alapján értékeljem, hanem a gyakorlati alkalmazás szempontjait is figyelembe véve, ahol a téves besorolások relatív hatása nem egyenértékű. A klasszikus mutatók, a statisztikai indikátorok

és a költségérzékeny értékelés integrálásával lehetővé vált az osztályozók összetettebb összehasonlítása. A 24. ábra az idő függvényében ábrázolja a támadási állapotokat (0 = nincs támadás, 1 = aktív támadás) egy adott időszakra vonatkozóan úgy, hogy az eseménytelen (0,0,0,0) időszeletek eliminálva lettek az adatsorból. Jól látható, hogy a támadások gyakorisága és hossza (különböző típusú zsarolóvírusok eltérő adatmennyiségen hajtják végre a módosításokat) eltérő.



27. ábra Támadási állapotok az idő függvényében [forrás: saját]

6.3. Az eredmények kiértékelése

6.3.1. Hiperparaméter hangolás

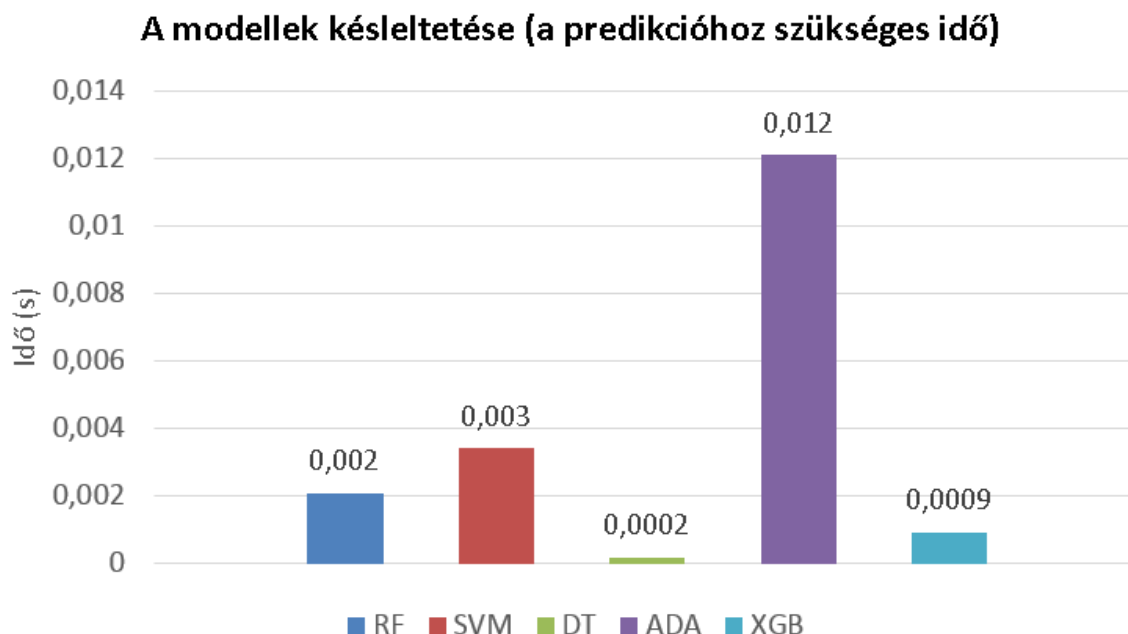
A gépi tanulási modellek először alapértelmezett hiperparaméterekkel kerültek kiértékelésre, majd a GridSearchCV segítségével kimerítő hangolás történt. A hiperparaméterek keresési tereinek és az optimalizált paraméter értékek összefoglalása a 20. táblázatban látható. Az optimalizálás célja az érzékenység (igaz pozitív arány) maximalizálása volt, miközben a pontosság és az érzékenység közötti egyensúlyt is meg kellett őrizni.

19. táblázat Hiperparaméter keresési terek és a legjobb paraméter értékek

	RF	SVM	DT	ADA	XGB
--	----	-----	----	-----	-----

Hiperparaméter keresési tér	'n_estimators': [10, 25, 50, 100, 150], 'criterion': ['gini', 'entropy'], 'max_depth': range(2, 6), 'min_samples_split': range(2, 6), 'min_samples_leaf': range(1, 6)	'C': [0.1, 1, 10], 'kernel': ['linear', 'rbf', 'poly'], 'gamma': [0.01, 0.1, 1, 'scale', 'auto']	'criterion': ['gini', 'entropy'], 'max_depth': range(2, 10), 'min_samples_split': range(2, 10), 'min_samples_leaf': range(1, 10)	'estimator_criterion': ['gini', 'entropy'], 'n_estimators': [50, 100, 150, 200], 'learning_rate': [0.1, 1, 10], 'estimator_max_depth': range(2, 5), 'estimator_min_samples_split': range(2, 5), 'estimator_min_samples_leaf': range(1, 5)	'n_estimators': [50, 100, 150, 200], 'learning_rate': [0.2, 0.3, 0.5, 1], 'subsample': [0.3, 0.4, 0.5, 0.7, 1], 'gamma': [0, 1, 3, 5], 'max_depth': [4, 5, 6, 10], 'tree_method': ['hist'], 'grow_policy': ['lossguide'], 'min_child_weight': [0, 3, 5], 'max_delta_step': [0, 1, 2, 5, 10], 'lambda': [0, 1, 2], 'alpha': [0, 1, 2]
Legjobb paraméter értékek	'criterion': 'entropy', 'max_depth': 5, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 25	'C': 1, 'gamma': 0.1, 'kernel': 'poly'	'criterion': 'entropy', 'max_depth': 9, 'min_samples_leaf': 2, 'min_samples_split': 9	'estimator_criterion': 'gini', 'estimator_max_depth': 4, 'estimator_min_samples_leaf': 2, 'estimator_min_samples_split': 4, 'learning_rate': 1, 'n_estimators': 100	'alpha': 0, 'gamma': 5, 'grow_policy': 'lossguide', 'lambda': 0, 'learning_rate': 1, 'max_delta_step': 0, 'max_depth': 5, 'min_child_weight': 0, 'n_estimators': 50, 'subsample': 1, 'tree_method': 'hist'

A hiperparaméter hangolás után az egyes modellek teljesítményének összehasonlítása a teszt adathalmazon elvégzett – 392 másodpercre vonatkozó – osztályozási eredmények alapján történt. A támadási állapot meghatározásához szükséges időigényt szintén figyelembe kellett venni a modellek értékelésénél. Ezeket a késleltetési időket a 28. ábra szemlélteti.



28. ábra A modellek késleltetésének összehasonlítása [forrás: saját]

Amint látható, a bináris osztályozáshoz szükséges idő az említett 392 másodpercre vonatkozóan gyakorlatilag minden esetben a milliszekundumos tartományban maradt (a legnagyobb késleltetés 12 századmásodperc volt), ami megerősíti, hogy a kiválasztott modellek mindegyike a kvázi valós idejű észlelést lehetővé teszi.

6.3.2. A gépi tanulási modellek teljesítményének összehasonlítása

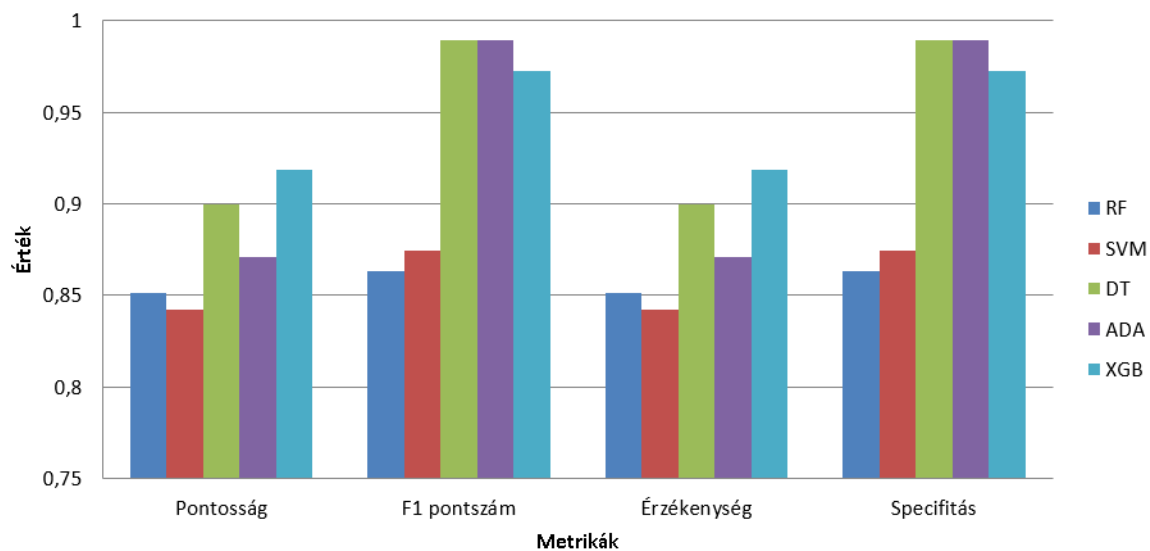
A modellek vizsgált teljesítmény metrikáit a hiperparaméter hangolás előtt, illetve után a 21. táblázat foglalja össze. Minden modell esetében az érzékenység a hangolás után meghaladta a 84%-ot, az XGBoost pedig a legmagasabb, 91,9%-os észlelési arányt érte el. Ez számszerűsítve azt jelenti, hogy a tesztadatkészletben a 209 támadással érintett másodpercből 192-t észlelt. A 29. ábra grafikusán szemlélteti a hiperparaméter hangolás után a modellek teljesítménybeli különbségeit, míg a 30. ábra a hamis előrejelzéseket mutatja meg a hangolt modelleknél. Az XGBoost és a döntési fa következetesen jó teljesítményt nyújtott, emellett a leggyorsabb osztályozók is voltak. Mivel az XGBoost a legjobb igaz pozitív arányt érte el megfelelő sebesség mellett, így a lenti metrikák alapján ideális választás lehet a bináris osztályozási feladatra.

20. táblázat A gépi tanulási modellek teljesítményértékelési metrikái

		RF	SVM	DT	ADA	XGB
Alapértelmezett paraméterekkel	Késleltetés (sec)	0,00208	0,00341	0,00019	0,01208	0,00093
	Pontosság	0,903	0,737	0,878	0,750	0,941

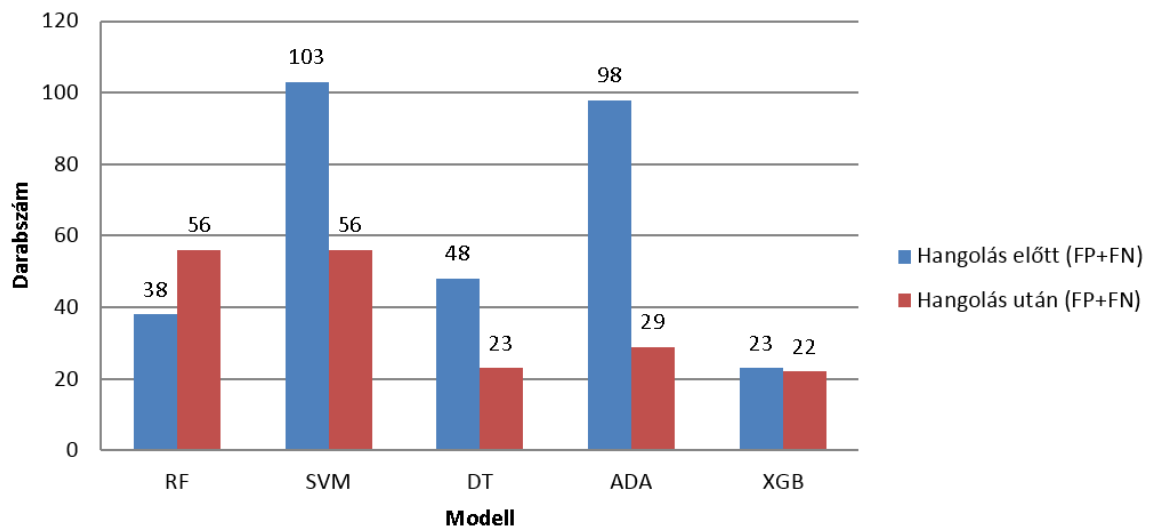
	F1 pontszám	0,903	0,681	0,879	0,750	0,942
	Érzékenység	0,842	0,526	0,833	0,703	0,900
	Specificitás	0,973	0,978	0,929	0,803	0,989
	Keveredési mátrix	TP:176 FP:5 TN:178 FN:33	TP:110 FP:4 TN:179 FN:99	TP:174 FP:13 TN:170 FN:35	TP:147 FP:36 TN:147 FN:62	TP:188 FP:2 TN:181 FN:21
Hangolt paraméterekkel	Pontosság	0,857	0,857	0,941	0,926	0,944
	F1 pontszám	0,864	0,863	0,942	0,926	0,946
	Érzékenység	0,852	0,842	0,900	0,871	0,919
	Specificitás	0,863	0,874	0,989	0,989	0,973
	Keveredési mátrix	TP:178 FP:25 TN:158 FN:31	TP:176 FP:23 TN:160 FN:33	TP:188 FP:2 TN:181 FN:21	TP:182 FP:2 TN:181 FN:27	TP:192 FP:5 TN:178 FN:17

A hangolt modellek teljesítményösszehasonlítása



29. ábra A modellek grafikus összehasonlítása [forrás: saját]

Hamis előrejelzések száma (FP+FN) a modellek hangolása előtt és után

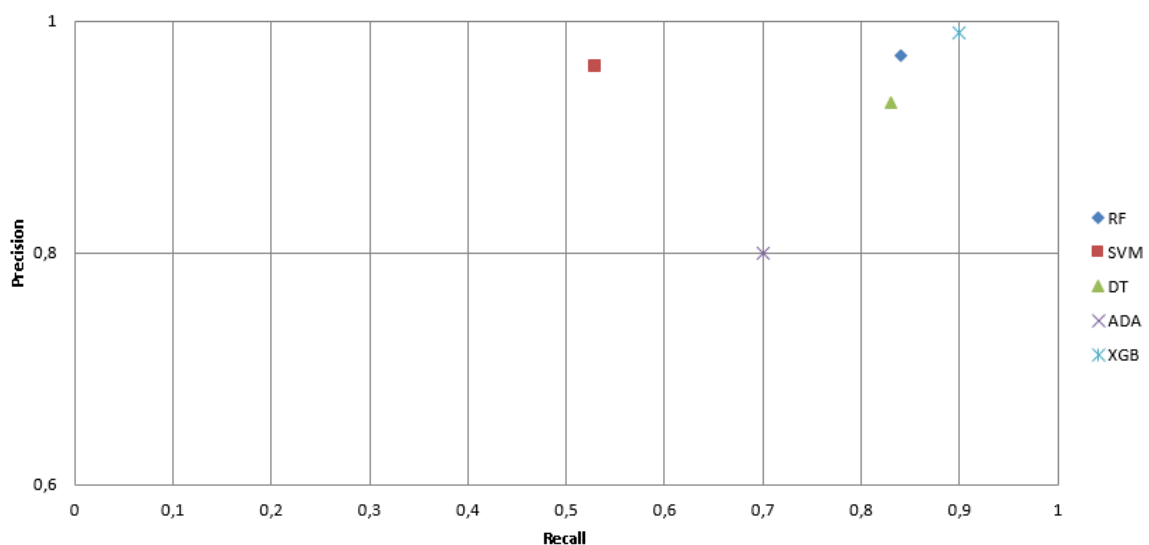


30. ábra Hamis előrejelzések a hangolt modellekben [forrás: saját]

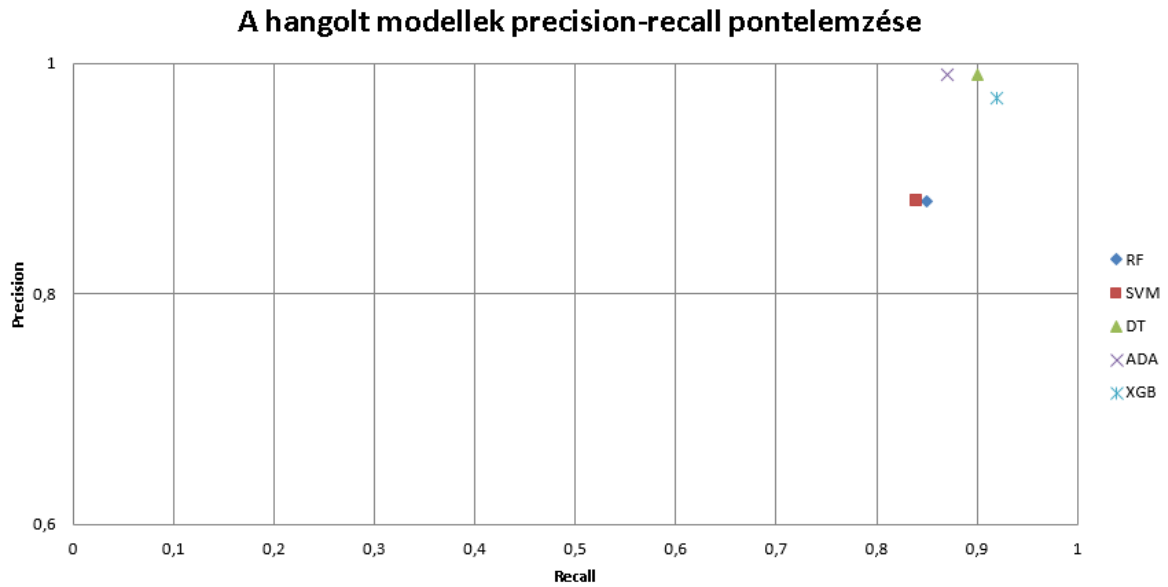
6.3.3. ROC és Precision-Recall pontelemzés

A fenti mutatók kiegészítéseként ROC és precision-recall pont elemzéseket végeztem, melyeket a 31., 32., 33., illetve a 34. ábra szemléltet. A paraméterek finomhangolása általánoságban javította a modell elhelyezkedését a ROC- és PR-térben. Az XGBoost elérte a legkedvezőbb kompromisszumot a recall, a precision és az FPR között. Ezzel szemben például az SVM korlátozott javulást mutatott, és egyes esetekben csökkent a Recall értéke.

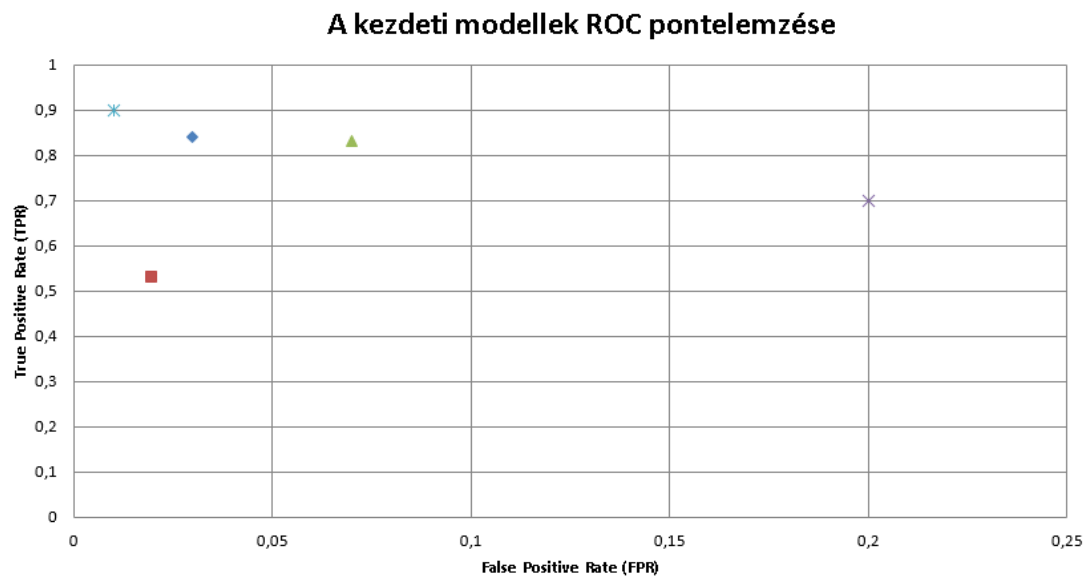
A kezdeti modellek precision-recall pontelemzése



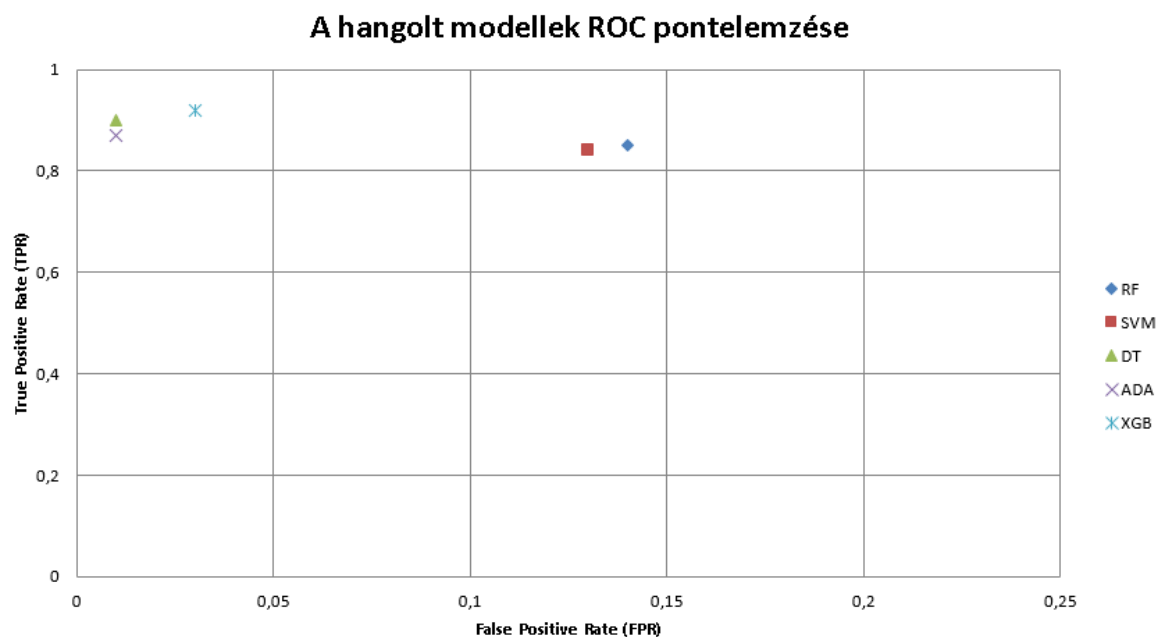
31. ábra A modellek kezdeti precision-recall pontelemzése [forrás: saját]



32. ábra A modellek hiperparaméter hangulás utáni precision-recall pontelemzése
[forrás: saját]



33. ábra A kezdeti modellek ROC pontjai [forrás: saját]



34. ábra A modellek hiperparaméter hangolás utáni ROC pontjai [forrás: saját]

6.3.4. A költség alapú értékelés eredménye

A téves besorolások egyenlőtlen következményeinek számszerűsítése érdekében költség-mátrix elemzést végeztem (22. táblázat) a következő költségek mellett:

$$C_{TP} = 0, C_{TN} = 0, C_{FP} = 1, C_{FN} = 10$$

Az XGBoost modell eredményezte a legalacsonyabb összköltséget, megerősítve ezzel a korábbi eredményt, így kijelenthető, hogy a felismerés és a megbízhatóság közötti egyensúly tekintetében valóban kimagasló. Ezzel szemben például az SVM a számos hamis negatív eredmény miatt a legmagasabb költséget eredményezte, ezzel is igazolva a zsarolóvírus-felismerésre való korlátozott alkalmazhatóságot.

21. táblázat A modellek költségmátrix elemzése

	RF	SVM	DT	ADA	XGB
FP	25	23	2	2	5
FN	310	330	210	270	170
Összköltség	335	353	212	272	<u>175</u>

6.4. Az eredmények összefoglalása

Összesen több mint 2,5 millió naplóbejegyzés került rögzítésre a vizsgált 30 napos időszak alatt. A következetes adattisztítás után a kapott adatkészleteket a gépi tanulási modellek felügyelt tanítására, tesztelésére és validálására használtam. Az értékelési keretrendszer támogatásán túl a nyers naplóbejegyzések elemzése a naplózási módszertan fejlesztéséhez is hozzájárult, valamint a finomított adatábrázolások bevezetésével jelentősen csökkenteni tudtam az adatsor méretét, ezáltal is gyorsítva és javítva a feldolgozás hatékonyságát. Mindeközben megőriztem az összes releváns viselkedési mintát és szem előtt tartottam a jövőbeni skálázhatóságot. Az eredmények megerősítik, hogy a zsarolóvírus-támadások észlelhető viselkedési nyomokat hagynak centralizált fájlátrolás esetén a fájlkiszolgáló naplóiban. A teljesítménymérések megerősítették, hogy a bináris osztályozás miatti késleltetés minden modell esetén a milliszekundumos tartományban maradt. A védelmi modell természeténél fogva robusztus a természetesen kiegyensúlyozatlan adatfolyamok esetén, amelyek pontosan tükrözik a valós felhasználási környezetet. A klasszikus modellek, mint például a döntési fa, ugyan megfelelő teljesítményt ért el, azonban az ensemble típusú modellek, és különösen az XGBoost, következetesen a legjobb kompromisszumot nyújtották az észlelési pontosság, a késleltetés és a költségérzékenység tekintetében. Az eredmények igazolják továbbá, hogy kiszolgálói oldalon a gépi tanulási modellekkel történő, kvázi valós idejű észlelés megvalósítható. A javasolt naplóalapú észlelési módszer nagy potenciállal rendelkezik a már meglévő biztonsági infrastruktúrák esetében is, például a biztonsággal kapcsolatos információk és események monitorozására szolgáló SIEM platformokba, vagy a végponti eszközök védelmére szolgáló (EDR) rendszerekbe való integrálás terén. A gépi tanulási modellek egyszerű adatstruktúrákon való használata bináris osztályozási feladatok esetén a mai számítási és memóriakapacitások mellett minimálisnak tekinthető CPU-terhelést és memóriaszükségletet igényel, így a módszer kvázi valós idejű telepítésre ad lehetőséget anélkül, hogy kernel-szintű, vagy hypervisor szintű módosításokra lenne szükség. A módszer hatékony lehet a “zero day” sérülékenységeket kihasználó, illetve polimorf zsarolóvírus változatoknak, mivel a felismerés a fájlművelet-szekvenciákon alapul, nem pedig víruskód mintákon vagy általános heurisztikákon. A javasolt megoldásom fő korlátja a kellően reprezentatív tanítási adathalmaztól való függőség, ami befolyásolhatja a különböző környezetek közötti hordozhatóságot. Ez a korlát azonban nem rontja a módszer általános megvalósíthatóságát, hanem kiemeli az adathalmazok strukturális tervezésének és a folyamatos adatgyűjtési stratégiáknak a fontosságát. A javasolt megközelítés erősségeit és korlátait a 23. táblázat foglalja össze.

22. táblázat A javasolt megközelítés erősségei és korlátai

Erősségek	Korlátok
-----------	----------

• minimális CPU- és memória-igény, alkalmas kvázi valós idejű használatra • nem igényel kiterjedt adatkészlet előkészítést a különböző zsaroló-vírus-támadások azonosításához • könnyen integrálható SIEM/EDR rendszerekbe, kernel módosítások nélkül • ellenálló a kiegyensúlyozatlan adatfolyamokkal szemben, biztosítva ezzel a valós körülmények közötti alkalmazhatóságot • nem szignatúra alapú, így hatékony lehet az ismeretlen, új vírus-variánsokkal szemben	• a kellően reprezentatív tanítási adathalmaztól függ a hatékonyság • egyedi naplózási rendszert igényel
--	---

A jelenleg elterjedt felismerési paradigmákkal történő összehasonlítás szintén fontos ahhoz, hogy lássam a modell erősségeit és gyengeségeit a többi megoldáshoz képest. A szignatúra alapú víruskereső rendszerekről köztudott, hogy nem képesek következetesen felismerni a “zero day”, vagy a polimorf változatokat, ahogy ezt Kothamali és mtsai vizsgálati eredményei is alátámasztják [93]. Az entrópia alapú detektorok viszont sebezhetőek a kijátszási technikákkal szemben (pl.: Base64 kódolás, részleges titkosítás), és általában jelentős számítási terheléssel járnak Kim és mtsai [94], illetve McIntosh és mtsai [95] 2023-ban publikált eredményei alapján. Hasonló aggályok merülnek fel a legújabb tanulmányokban is, amelyek kiemelik az I/O-alapú megközelítések pontossága és a futási hatékonyság közötti kompromisszumot, például Van Sloun és mtsai 2025-ös publikációja [96]. A magas entrópiájú titkosított adatfolyamokra alkalmazott fejlett ML-osztályozók pedig szintén jelentős erőforrásigényűek (Casino és mtsai, 2025, [97]). Ezzel szemben a szerveroldali, fájlműveleti napló alapú megközelítés magasabb felismerési arányt és alacsonyabb terhelést mutat ezekhez a módszerekhez képest. A 21. táblázatban összefoglaltam néhány korábbi szakirodalomban publikált tanulmány megállapításait.

23. táblázat Főbb megközelítések és azok zsarolóvírus-felismerési paradigmákra vonatkozó megállapításai

Forrás	Megállapítás
Kothamali és Banik (2025) Az aláírásalapú fenyegetésérzékelés korlátai [93]	A szignatúra-alapú rendszerek nehezen reagálnak az új (zero day) fenyegetésekre és a rejtett rosszindulatú programváltozatokra.

McIntosh, Jang-Jaccard, Watters, Susnjak (2023) Az entrópia-alapú zsarolóvírus-felismerés elégtelensége [95]	Az entrópia-alapú detektorok könnyen ki- játszhatók (pl.: Base64 kódolás, részleges titkosítás), az entrópiaértékek manipulál- hatók.
Van Sloun és mtsai (2025) Zsarolóvírus észlelése az I/O-terhelés da- cára: egy gyakorlati, többfázisú megköze- lítés [96]	Az I/O-alapú monitorozás hatékony, de mérhető többletterhelést jelent, pontosság és a hatékonyság közötti kompromisszum- kényszert mutat.
Kim (2022) Bájtfrekvencia-alapú indikátorok kriptó- gráfiai zsarolóvírusokhoz [94]	A bájtfrekvencia- és entrópia-alapú mód- szerek nagy fájlok esetén jelentős többlet- terhelést okoznak; hangsúlyozzák az erő- forrás-gazdálkodás fontosságát.
Casino és mtsai (2025) „Not on My Watch”: zsarolóvírus-felis- merés osztályozás segítségével [97]	Módszerük megvizsgálja a titkosított bit- folyamokat és megmutatja, hogy a gépi ta- nulási megközelítések képesek hatéko- nyan működni, de ez gyakran jelentős szá- mítási költségekkel jár.

Összefoglalva az eredményeket kijelenthetem, hogy a zsarolóvírus észlelése hatékonyan megvalósítható a szerver oldali fájlműveleti naplófájlok egyszerű, összesített jellemzőinek kihasználásával. Klasszikus és ensemble gépi tanulási modelleket (RF, SVM, DT, ADA, XGB) hasonlítottam össze szisztematikusan, és kimutattam, hogy az ensemble alapú megközelítések közül az XGBoost biztosítja a legkedvezőbb kompromisszumot az érzékenység, a specifitás és a számítási költség között. Az értékelési keretrendszer tartalmazta a késleltetés mérését, a ROC és a Precision–Recall pontok elemzését, valamint a költségérzékeny értékelést, ezáltal formális és reprodukálható módszertant kínálva a zsarolóvírus-felismerési megközelítések összehasonlításához. Megállapíthatjuk továbbá, hogy a felismerési teljesítmény esetünkben független a támadás során használt zsarolóvírus variánstól, ami lehetővé teszi a módszer általánosíthatóságát. A 21. táblázatban szereplő megközelítések korlátaira vonatkozóan a következő megállapításokat tehetjük a bemutatott módszerrel kapcsolatban:

- nem szignatúra alapú, így könnyebben reagál az új vírusmutánsokra, zero day alapú fenyegetésekre;
- nem alkalmaz entrópia alapú detekciót, hanem kizárólag fájlművelet szekvenciákat vizsgál, ami irrelevánsá teszi a titkosítás fajtáját, illetve az entrópia manipulációját
- az I/O monitorozás lényegesen nagyobb terhelést jelent, mint az alkalmazott bináris osztályozás, miközben a legitim folyamatok figyelmen kívül hagyása további kihívást jelentő feladatot jelent az előbbinél;
- a magas entrópiájú titkosított adatfolyamokra alkalmazott fejlett ML-osztályozókhoz képest a szerveroldali, fájlműveleti napló alapú megközelítés magasabb felismerési arányt és alacsonyabb terhelést mutatott.

6.5. A megközelítés korlátai és továbbfejlesztési lehetőségei

A javasolt modell teljesítménye elsősorban azon a feltételezésen alapul, hogy a tanítóadatokban szereplő vírusok által kiváltott fájlműveleti viselkedés (fájlok létrehozása, átnevezése, törlése) reprezentálja a később előforduló támadásokat. A vizsgálat során a modell öt különböző zsarolóvírus (Ryuk, WannaCry, NotPetya, LockBit és TeslaCrypt) mintázata alapján tanult, ezért az ezekhez hasonló támadási és titkosítási stratégiák esetén várhatóan továbbra is magas felismerési arányt biztosít. Azonban a modern zsarolóvírusok egyre gyakrabban alkalmaznak olyan technikákat, amik eltérő fájlműveleti mintákat eredményeznek. Ilyenek például a következők:

- szakaszos/ritkított titkosítás (intermittent encryption/sparse encryption): a titkosítás szabályos vagy szabálytalan elrendezésű szegmensekben történik a fájlban belül;
- fájlfejléc/metaadat titkosítása (header-only/metadata encryption): a fájl fejlécének vagy metaadatainak titkosítása, így a fájl használhatatlanná válik;
- alacsony intenzitású, fokozatos titkosítás (low-and-slow encryption): a titkosítás lassan, hosszú idő alatt zajlik;
- késleltetett/többlépcsős titkosítás (delayed/staged encryption): a titkosítás csak egy késleltetett időpontban vagy több fázisban indul el;
- részleges titkosítás (partial encryption): csak a fájlok egy része kerül titkosításra (pl.: a 10-20 százaléka, vagy az első 500 kB);
- titkosítás nélküli zsarolás (encryption-less/data theft-only extortion): nincs titkosítás, az adatok kiszivárogtatásával zsarolják az áldozatot.

Ezek közös jellemzője, hogy jelentősen csökkentik azokat a fájlműveleti aktivitásokat, amelyekre a modell épít. Ennek következtében a fájlművelet-szekvenciák kevésbé különülhetnek el a normál felhasználói aktivitástól, ami növelheti a hamis pozitív riasztások számát és csökkentheti a sikeres felismerések arányát. Ilyenkor a modell bemeneti attribútumai és a célparaméter közötti statisztikai összefüggés az idő múlásával változik. Ez a jelenség a gépi tanulásban „concept drift” problémaként ismert.

A concept drift kezelésére több lehetőség is rendelkezésre áll:

- a modell rendszeres újratanítása aktuális és nagyobb számú zsarolóvírus-mintával;
- az írási műveletek bevonása az adatsorba;
- az 1 másodperces időablak csökkentése (pl.: századmásodperc);
- online vagy inkrementális tanulási módszerek alkalmazása;
- drift-detektáló algoritmusok (pl.: DDM, EDDM, ADWIN) használata;

- több adatforrás együttes felhasználása (fájlműveletek mellett folyamat-, memória-, hálózati és rendszerhívási események);
- anomáliadetektálási vagy félig felügyelt tanulási módszerek bevonása;
- a fájlentropia változások figyelése;

A jelenlegi modell elsősorban ismert vagy hasonló viselkedésű támadások detektálására alkalmas. A modern, alacsony intenzitású vagy adaptív titkosítási technikák kezelése érdekében a modellt folyamatos adatfrissítéssel, concept drift kezelésével és többforrású viselkedéselemzéssel célszerű továbbfejleszteni. A titkosítás nélküli zsarolás viszont már alapvetően más támadási startégia. Ebben az esetben nincs tömeges fájlmodosítás vagy titkosítás, így egy kizárólag a felsorolt fájlműveletekre épülő detektáló modell nem képes az ilyen támadások felismerésére. Ezekben az esetekben azonban az olvasási műveletek bevonása az adatsorba, a hálózati kommunikáció és a megfelelő hálózati interfészek adatforgalmi karakterisztikáinak figyelése (pl.: periméter átjáró és fájlkiszolgáló) alternatívát jelenthet.

6.6. A releváns kutatási kérdések megválaszolása

Az első kutatási kérdés (K-8) tekintetében, bár a kísérleti környezet egy KKV léptékű infrastruktúra viselkedését modelleztem egyetlen munkaállomással (aggregált fájlműveletek), a mérések azt mutatják, hogy az XGBoost ensemble modell használatával a támadási állapot megállapításának sebessége több száz kliens esetén is kvázi valós idejű marad. A második kutatási kérdéssel (K-9) kapcsolatosan kijelenthetem, hogy az alkalmazott naplózási infrastruktúra (unix socket, FluentBit, InfluxDB, Lua), a naplózott attribútumok célszerűen minimalizált száma és típusa, valamint az előfeldolgozási lépések valóban biztosították, hogy a gépi tanulási modellek felügyelt tanítása hatékony legyen. Ez azt jelenti, hogy a kezdetben 2 577 145 rekordot tartalmazó adathalmaz a redundáns, illetve irreleváns sorok eltávolítása után 10278 rekordot tartalmazott, ami a tréning és a tesztadatsor együttes méretét adta. A vizsgált modellek felügyelt tanítása a hiperparaméter hangolással együttesen is csupán másodperces léptékű volt. Az utolsó kutatási kérdést (K-10) illetően, bár az egy másodperces aggregációs intervallumok és a kanári fájl alapú helyreállítás miatt az adatvesztés pontos számszerűsítése nem lehetséges, a kísérletek azt mutatják, hogy a jelentős fájlkárosodás bekövetkezte előtt elérhető a felismerés az öt zsarolóvírus működési mechanizmusától függetlenül. Az eredmények tükrében az egy másodperces aggregációs időablak tizedmásodperces intervallumra biztosan csökkenthető.

7. A doktori értekezés tézisei

1. Tézis

Igazolható, hogy a heterogén kliens–szerver architektúrák fájlátviteli teljesítménye páronként definiált, dinamikus referencia-alapú relatív teljesítménymutató alkalmazásával objektív módon összevethető, ahol az összehasonlíthatóságot az azonos művelet sor mellett mért átviteli sebességek egymáshoz viszonyított aránya biztosítja, nem pedig abszolút teljesítményértékek.

1. Tézis alátámasztása

A heterogén kliens–szerver architektúrák vizsgálata során megállapítható volt, hogy az abszolút fájlátviteli teljesítményértékek még azonos hardverkörnyezet és fájl tartalmak mellett is jelentős eltéréseket mutathatnak a szoftverkörnyezetek függvényében. Ez a jelenség megnehezíti az eltérő konfigurációk objektív összehasonlítását, mivel az abszolút mérési eredmények nem különíthetők el egyértelműen a környezeti sajátosságoktól. A heterogén környezetekben végzett teljesítménymérések összehasonlíthatósági problémájának kezelésére a kutatás során egy új, általam kidolgozott, páronként definiált, dinamikus referencia-alapú relatív teljesítménymutatót vezettem be, amely platformfüggetlen módon alkalmazható, reprodukálható mérési eljárást biztosít, és megfelelő referencia-rendszer definiálása mellett általánosítható. A mutató a mérések során az azonos fájlátviteli irányok (olvasás, írás) mellett mért átviteli sebességek egymáshoz viszonyított arányán alapult, így nem az abszolút átviteli sebességeket tekinti elsődleges összehasonlítási alapnak, hanem minden vizsgált konfiguráció teljesítményét egy előre definiált referencia-rendszerhez viszonyítja. Ez mérsékli a platformfüggő sajátosságok összehasonlítás torzító hatását, és lehetővé teszi különböző operációs rendszerek, fájlrendszerek és adatvédelmi mechanizmusok által nyújtott fájlátviteli teljesítmény egységes skálán történő értékelését. Fontos kiemelni továbbá, hogy a kidolgozott módszerrel az egyes konfigurációk közötti teljesítményváltozások – például rendszerhangolás vagy védelmi funkciók bevezetése – következetesen nyomon követhetővé válnak. A kutatás tudományos hozzájárulása nem kizárólag egy FreeBSD-alapú konfiguráció teljesítményének meghatározásában rejlik, hanem egy olyan általánosítható referencia-alapú értékelési módszer kidolgozásában, amely heterogén szolgáltatáskörnyezetek objektív összehasonlítását teszi lehetővé. A módszer alkalmazásával különböző operációs rendszerek, fájlrendszerek, titkosítási eljárások és adattárolási architektúrák teljesítménye egységes mérőszám segítségével értékelhető. A megközelítés ezért nemcsak a jelen kutatásban vizsgált konfigurációkra alkalmazható, hanem általános módszertani keretet biztosít jövőbeli rendszerösszehasonlító vizsgálatok számára is. A vizsgálatok során meghatározott relatív teljesítménymutató valamennyi vizsgált konfiguráció esetében konzisztensen értelmezhető eredményeket szolgáltatott, és lehetővé tette az eltérő operációs rendszerek, fájlrendszerek és adatvédelmi konfigurációk objektív összehasonlítását. A különböző mérési sorozatokban kapott eredmények igazolták, hogy a referencia-alapú mutató alkalmas a teljesítményváltozások számszerű követésére, ezáltal alátámasztja a tézisben megfogalmazott módszertani állítást.

Az 1. tézist alátámasztó saját publikációk:

[S1] Szücs, V., Arányi, G., & Dávid, Á. (2021). Introduction of the ARDS-Anti-Ransomware Defense System Model-Based on the Systematic Review of Worldwide Ransomware Attacks. *Applied Sciences-Basel*, 11(13). <http://doi.org/10.3390/app11136070>, SJR Q2, WoS Q2 IF: 2,838

[S2] Arányi, G.; Vathy-Fogarassy, Á.; Szücs, V.; Evaluation of a New-Concept Secure File Server Solution. *Future Internet* 2024, MDPI, 16, 306. <https://doi.org/10.3390/fi16090306>, SJR Q2, WoS Q2 IF: 3,6

2. Tézis

Igazolható, hogy centralizált fájlkiszolgáló környezet esetén, virtualizált kliens–szerver infrastruktúrában a szerveroldali időbélyeggel ellátott, idősorfolytonos fájlműveleti naplóadatok megfelelő előszűréssel, normalizálással és egy másodperces időablakra történő aggregációval olyan strukturált idősoros reprezentációvá alakíthatók, amely alapján bizonyos gépi tanulási modellek alkalmasak kliensoldali információ nélkül is a teljes titkosítást alkalmazó zsarolóvírusok által generált intenzív fájlműveleti aktivitás kvázi valós idejű hatékony detektálására.

2. Tézis alátámasztása és korlátai

A vizsgálat olyan klasszikus, teljes titkosítást alkalmazó zsarolóvírusokra (Ryuk, WannaCry, NotPetya, LockBit, TeslaCrypt) terjedt ki, amelyek működésük során intenzív létrehozási, átnevezési és törlési fájlműveleteket generálnak. A mérések centralizált fájlkiszolgálást megvalósító, virtualizált kliens–szerver infrastruktúrában történtek. Az elemzés idősorfolytonos, egy másodperces időablakokra aggregált adatsoron zajlott, ahol a referenciaállapotot tipikus irodai felhasználói fájlhasználati minták alapján definiáltam. A mérések igazolták, hogy a fájlkiszolgáló oldalon keletkező, időbélyeggel ellátott fájlműveleti naplóadatok önmagukban elegendő információt hordoznak a vizsgált zsarolóvírus-tevékenységekhez kapcsolódó intenzitás- és mintázatváltozások kimutatásához. Normál működés során a létrehozási, átnevezési és törlési műveletek száma és időbeli eloszlása stabil referenciaállapotot alkotott, míg zsarolóvírus-tevékenység esetén rövid idő alatt nagyságrendi intenzitásváltozás és jellegzetes mintázateltérés jelent meg. Az előszűrés és az időablakos aggregáció eredményeként a naplóadatok nem nyers eseményhalmazként, hanem strukturált, idősorfolytonos numerikus reprezentációként álltak rendelkezésre. Ez az adatrepresentáció közvetlenül felhasználható bemenetet biztosított a felügyelt gépi tanulási modellek számára, megalapozva a zsarolóvírus-tevékenység automatikus detektálását. Az egymást követő időablakokban aggregált eseménysorok alapján a fájlműveleti aktivitás időbeli lefutása folyamatosan nyomon követhetővé vált, így az észlelési folyamat nem utólagos elemzésként, hanem kvázi valós idejű megfigyelésként volt értelmezhető. A kutatás módszertani hozzájárulása egy olyan szerveroldali, időablakos adatrepresentáció kidolgozása, amely kizárólag fájlkiszolgálói naplóadatokra támaszkodva biztosítja a zsarolóvírus-tevékenység gépi tanulással történő detektálásához szükséges bemeneti reprezentációt. A módszer nem kötődik konkrét

gépi tanulási algoritmushoz, ezért különböző osztályozási modellek bemeneteként is alkalmazható. A következtetések olyan centralizált fájlkiszolgáló környezetekre érvényesek, ahol a fájlműveletek időbélyeggel naplózhatók, és a támadók intenzív fájlrendszer-szintű titkosítási folyamatot hajtanak végre. A vizsgálat nem terjedt ki olyan zsarolóvírus-változatokra, amelyek minimális fájlműveleti aktivitással dolgoznak, vagy nem alkalmaznak titkosítást. A referenciaállapot tipikus irodai munkaterhelés mellett került kialakításra, ezért a hosszú távú üzemeltetés során jelentkező felhasználói viselkedésváltozás, valamint az újabb ransomware-családokra (concept drift) történő általánosíthatóság vizsgálata további kutatási irányt jelent. Emellett a jelen kutatás időkerete nem tette lehetővé a módszer hosszabb időtartamra vonatkozó validálását, ezért annak stabilitása és adaptivitása hosszabb üzemeltetési periódusokban további empirikus vizsgálatot igényel.

3. Tézis

Igazolható, hogy a normál felhasználói aktivitás és a teljes titkosítást alkalmazó zsarolóvírusok által kiváltott intenzív fájlműveleti mintázatok a vizsgált környezetben bináris osztályozási problémaként kezelhetők és klasszikus, valamint ensemble gépi tanulási módszerekkel a támadás időkritikus lefutásához illeszkedő döntési késleltetés mellett hatékonyan detektálhatók.

3. Tézis alátámasztása

A kutatás során a fájlkiszolgáló oldalon keletkező, időbélyeggel ellátott fájlműveleti naplóadatokra vonatkozóan olyan előfeldolgozási eljárást dolgoztam ki, amely előszűrést és egy másodperces időablakok szerinti aggregációt alkalmazott. Az adatelőkészítés eredményeként a létrehozási, átnevezési és törlési fájlműveletek nem nyers eseménysorként, hanem strukturált, idősorfolytanos numerikus reprezentációként álltak rendelkezésre, amely közvetlenül felhasználható bemenetet biztosított a felügyelt gépi tanulási modellek számára. A referenciaállapotot különböző munkakörökhöz tartozó, tipikus irodai fájlhasználati minták alapján definiáltam. Normál működés során a fájlműveletek intenzitása és időbeli eloszlása stabil tartományban maradt, míg a vizsgált zsarolóvírusok (Ryuk, WannaCry, NotPetya, LockBit, TeslaCrypt) működése során rövid idő alatt nagyságrendi intenzitásnövekedés, valamint jól elkülöníthető időbeli mintázatváltozás jelentkezett. Ezek az eltérések nem elszigetelt eseményekként, hanem egymást követő időablakokon keresztül fennálló viselkedési mintázatként jelentek meg, amely lehetővé tette a normál és a támadással érintett időszakok egyértelmű elkülönítését. Mivel a kutatás célja a zsarolóvírus-tevékenység jelenlétének felismerése volt, nem pedig az egyes kártevőcsaládok azonosítása, a detektálási feladat bináris osztályozási problémaként volt kezelhető. A strukturált adatrepresentáción végzett vizsgálatok igazolták, hogy mind a klasszikus, mind az ensemble gépi tanulási modellek alkalmasak a normál és a támadással érintett időablakok elkülönítésére. A hiperparaméter-optimalizálást követően valamennyi vizsgált modell 0,84 feletti érzékenységet ért el, míg az XGBoost modell 0,919-es értékkel kiemelkedett a többi modell közül. Ez a tesztadatkészlet 209 támadással érintett másodpercére vetítve 192 esetben helyes eredményt jelent. Az XGBoost emellett a legkedvezőbb kompromisszumot nyújtotta az érzékenység, a precision,

és a hamis pozitív arány között, valamint a költségalapú értékelésnél is a legjobbnak bizonyult. A mérések azt is igazolták, hogy a döntési folyamat időigénye nagyságrendileg megfelelő a zsarolóvírus-támadások időkritikus észleléséhez. Az osztályozás végrehajtásához szükséges idő a legjobb teljesítményt nyújtó XGBoost modell esetében 392 időablakhoz tartozó bemeneti adatsor feldolgozásánál 0,93 ms volt, míg a többi modellnél 0,19 ms és 3,41 ms közötti értékeket mértem. Az időablakos feldolgozás és a modellek alacsony számítási igénye lehetővé tette, hogy az osztályozás eredménye az adatgyűjtést követően gyakorlatilag azonnal rendelkezésre álljon, miközben az erőforrásigény nem haladta meg a kis- és középvállalati környezetekben alkalmazott általános célú szerverek kapacitását. A vizsgálatok alapján így igazolható, hogy a normál felhasználói aktivitás és a fájlrendszer-szintű ransomware-ek által kiváltott intenzív fájlműveleti mintázatok a vizsgált környezetben bináris osztályozási feladatként kezelhetők, és megfelelő adatelőkészítés esetén klasszikus, illetve ensemble gépi tanulási módszerekkel a támadások időkritikus lefutásához illeszkedő döntési késleltetés mellett hatékonyan detektálhatók. A következtetések a vizsgálatban szereplő teljes titkosítást alkalmazó zsarolóvírusokra, valamint centralizált, virtualizált fájlkiszolgáló infrastruktúrára érvényesek. A referencia-állapot egy tipikus irodai munkaterhelés mellett került kialakításra. A kutatás nem terjedt ki a hosszú távú üzemeltetés során fellépő felhasználói viselkedésváltozás kezelésére, illetve a módszer több hónapos vagy éves időtávú validálására. Nem került vizsgálatra továbbá a később megjelenő, jelentősen eltérő működési karakterisztikájú ransomware-változatokra történő általánosíthatóság sem (concept drift vizsgálata), ezért ezek a kérdések további kutatási irányt jelentenek.

A 2. és 3. téziseket alátámasztó saját publikációk:

[S1] Szücs, V., Arányi, G., & Dávid, Á. (2021). Introduction of the ARDS-Anti-Ransomware Defense System Model-Based on the Systematic Review of Worldwide Ransomware Attacks. *Applied Sciences-Basel*, 11(13). <http://doi.org/10.3390/app11136070> SJR Q2, WoS Q2 IF: 2,838

[S3] Arányi, G; Miseta, T; Szücs V., Ransomware Detection Based on Server-side File Operation Logs using Machine Learning. *Journal on Information Security, Advances in Information Forensics and Security*; Springer, 2025. <https://doi.org/10.1186/s13635-026-00229-7> SJR Q2, WoS Q2 IF: 2,1

4. Tézis

Igazolható, hogy a vizsgált klasszikus, valamint ensemble gépi tanulási modellek gyakorlati használhatóságának megítéléséhez önmagában egyetlen teljesítménymutató nem elegendő, és a pontosság, az F1-pontszám, az érzékenység, a specificitás, a keveredési mátrix, valamint a költség-alapú értékelés együttes alkalmazása átfogó és gyakorlati szempontokat szem előtt tartó összehasonlíthatóságot biztosít.

4. Tézis alátámasztása

A kutatás során végzett vizsgálatok rámutattak arra, hogy a zsarolóvírus-tevékenység detektálása esetén a gyakorlati használhatóság megítélése nem alapozható egyetlen teljesítménymutatóra. Különösen kiegyensúlyozatlan osztályeloszlás mellett a pontosság önmagában félrevezető képet adhat a modell teljesítményéről, mivel nem tükrözi a különböző típusú osztályozási hibák eltérő súlyú következményeit. A keveredési mátrix elemzése lehetővé tette a hamis pozitív és hamis negatív döntések arányának feltárását, míg az érzékenység és a specificitás vizsgálata rávilágított a detektálási képesség és a téves riasztások közötti kompromisszumokra. Az F1-pontszám ezek együttes figyelembevételével szolgált összegzett mutatóként, azonban önmagában ez sem bizonyult elegendőnek a gyakorlati alkalmazhatóság megítéléséhez. A költség alapú értékelés alkalmazása lehetővé tette, hogy a téves osztályozási eredményeket súlyozva tudjuk összehasonlítani a modellek teljesítményét. Így szemléltethető volt, hogy az egyes osztályozási hibák milyen mértékű károkozást jelentenének egy valós környezetben. Az eredmények alapján megállapítható, hogy a pontosság, az F1-pontszám, az érzékenység, a specificitás, a keveredési mátrix és a költség alapú értékelés együttes alkalmazása egy alkalmas megközelítés ahhoz, hogy a különböző gépi tanulási modellek teljesítményét összehasonlíthassuk és a gyakorlati használat szempontjából értelmezhető módon megítélhessük. A kutatás tudományos újdonsága annak igazolása, hogy a zsarolóvírusok detektálásához használt modellek értékelése többkritériumos döntési problémaként kezelhető, amelyben a modellek gyakorlati alkalmassága a klasszikus teljesítménymutatók és a költség alapú értékelés integrált alkalmazásával megalapozottan megítélhető.

A 4. tézist alátámasztó saját publikáció:

[S3] Arányi, G; Miseta, T; Szücs V., Ransomware Detection Based on Server-side File Operation Logs using Machine Learning. Journal on Information Security, Advances in Information Forensics and Security; Springer, 2025. <https://doi.org/10.1186/s13635-026-00229-7> SJR Q2, WoS Q2 IF: 2,1

8. Összegzés és kitekintés

8.1. Az eredmények értelmezése és tudományos jelentősége

A nem reprezentatív kérdőíves interjú eredményei igazolták, hogy a megkérdezett kis- és középvállalkozások informatikai környezetét tekintve valóban jellemző a korlátozott szakértelem, a védelmi technológiák jelentős hiánya és a vírustámadásoknak való fokozott kitettség, amik együttesen indokoltá teszik az olyan automatizált, alacsony üzemeltetési igényű technikai megoldások gyakorlati relevanciáját, mint aminek az alapjait a dolgozat bemutatja. A kiértékelés során világossá vált, hogy a technológiai megvalósíthatóság és a szervezeti elfogadottság fontos szempont az ilyen rendszerek sikeres bevezetéséhez és a résztvevők többsége nyitott lenne egy hasonló adatbiztonsági megoldás bevezetésére. A heterogén kliens-szerver architektúrák fájlátviteli sebességtesztjei (hálózati átvitel és írás/olvasás) rávilágítottak arra, hogy szükség van az abszolút teljesítménymutatók mellett olyan kalkulált metrika használatára is, amelyek önmagukban alkalmasak eltérő környezetek teljesítményének egzakt, számszerű összehasonlítására. A referencia-alapú teljesítményértékelés bevezetése lehetővé tette a különböző hardver- és szoftverkombinációk közötti összehasonlíthatóságot. Ez a megközelítés nem csupán a konkrét vizsgálati környezetben bizonyult működőképesnek, hanem általánosítható módszertani eszközként is értelmezhető bármilyen teljesítménymetrika esetén, ha heterogén rendszerek viszonylagos teljesítményét szeretnénk jellemezni. Az értekezés eredményei alapján megállapítható, hogy centralizáltan tárolt adatvagyron esetén a fájlműveleti naplókra épülő zsarolóvírus-tevékenység észlelése életképes alternatívát jelenthet a hagyományosan kliensoldali védelmi megközelítésekkel szemben. A mérések és kísérletek igazolták, hogy a fájlműveleti naplóbejegyzések olyan, a normál működéstől szignifikánsan eltérő mintázatokat mutatnak valós zsarolóvírus-támadások esetén, amelyek önállóan is alkalmasak detekciós célokra. Ez az eredmény különösen fontos azon környezetekre nézve, ahol a kliensoldali, komplex védelmi megoldások teljeskörű alkalmazása technikai vagy szervezeti okokból nem megoldható. A gépi tanuláson alapuló zsarolóvírus-tevékenység detektálására irányuló kísérletek eredményei azt mutatták, hogy a klasszikus gépi tanulási módszerek közül a döntési fa, míg az ensemble módszerek közül az XGBoost alkalmas leginkább a megfelelően strukturált bemeneti adatok esetén hatékony felismerési teljesítményt nyújtani. A vizsgált modellek közül összességében az XGBoost teljesített a legkifejezettebben a felismerési pontosság, a hamis negatív arány és a számítási igény tekintetében. Különösen fontos eredmény, hogy a detektálás időigénye összhangban van a kvázi valós idejű működés követelményeivel. Ez azt jelenti, hogy a javasolt megoldás nem kizárólag utólagos elemzésre alkalmas, hanem képes a támadási folyamat során releváns jelzést adni, adott esetben incidensválaszt kiváltani. Ez a tulajdonság a zsarolóvírus-támadások elleni védekezésben kulcsfontosságú, mivel az időben történő beavatkozás közvetlenül csökkentheti az adatvesztés mértékét.

8.2. Korlátok és értelmezési megfontolások

A kutatás során alkalmazott megközelítés egyik korlátja, hogy a vizsgálatok során az egyedileg kialakított fájlkiszolgáló környezet esetén a mentési technológia, a naplóadatok szerkezete és részletessége nem általánosan alkalmazott, ezért jelen formájában nem integrálható könnyen meglévő infrastruktúrákba. További korlátozó tényező, hogy a mesterséges intelligencia alapú modellek tanítása és értékelése kontrollált kísérleti környezetben történt. Bár ez ugyan biztosította a reprodukálhatóságot és az összehasonlíthatóságot, valós üzemi környezetben további zajforrások és atipikus használati minták jelenhetnek meg, amelyek befolyásolhatják a detektálási teljesítményt. Az értekezésben bemutatott fájlkiszolgáló alapú zsarolóvírus-tevékenység detektálási megközelítés nem önálló védelmi mechanizmusként, hanem egy többrétegű védekezési koncepció részeként értelmezhető. A dolgozatban röviden, de egyértelműen kitértem a mentési és helyreállítási megoldások szerepére, amelyek a zsarolóvírus-támadások elleni védekezés elengedhetetlen, ugyanakkor önmagukban nem elegendő elemei. A mentési és detektálási megoldások alapvetően reaktív jellegű védelmet biztosítanak, amelyek elsősorban az adatvesztés következményeinek mérséklésére szolgálnak. A dolgozatban foglalt eredményeim ugyanakkor rámutatnak arra, hogy a kizárólag mentési stratégiára épülő védekezés nem képes kezelni a támadások időbeli lefolyásából adódó kockázatokat, különösen akkor, ha a zsarolóvírus-tevékenység hosszabb ideig észrevétlen marad. Ebben az esetben a mentések is kompromittálódhatnak a mentési láncban, ami az adatvagyon részleges vagy teljes elvesztését is jelentheti. A fájlkiszolgáló oldali, kvázi valós idejű detektálás ebben a kontextusban a mentési és helyreállítási réteg hatékonyságát erősítő komponensként értelmezhető. A korai felismerés lehetőséget teremt arra, hogy a fertőzött kliensek időben izolálásra kerüljenek, az adattitkosítási folyamat megállítható legyen, és a helyreállítás kontrollált módon történjen meg. A dolgozatban bemutatott mesterséges intelligencia alapú megközelítés így nem helyettesíti, hanem funkcionálisan kiegészíti a mentési megoldásokat. A bemutatott eredmények alapján megállapítható, hogy a fájlkiszolgáló-alapú detektálás és a mentési-helyreállítási megoldás együttese olyan védelmi kombinációt alkot, amely a kis- és középvállalkozások erőforráskorlátai mellett is képes csökkenteni a zsarolóvírus-támadások által okozott károkat. Fontos, hogy a hangsúly nem a mentési megoldások kiváltásán, hanem azok hatékonyabb alkalmazásán és a támadások korai felismerésén van.

8.3. Az eredmények nemzetközi kutatásokkal való összevetése

A kutatás eredményeinek értelmezését és kontextusba helyezését segíti a Sophos által 2025-ben publikált nemzetközi ransomware-felmérés eredményeinek áttekintése. A vizsgálat 3400 olyan szervezet tapasztalatait dolgozta fel 17 országban, amelyek az előző év során zsarolóvírus-támadás áldozataivá váltak. A kutatásba 100–5000 fő közötti szervezetek kerültek bevonásra, így a vizsgálat elsősorban a közép- és nagyobb méretű vállalatok tapasztalatait tükrözi. A Sophos jelentése szerint a sikeres támadások leggyakoribb technikai oka

az ismert sérülékenységek kihasználása volt, amely az esetek 32%-ában jelent meg. Emellett továbbra is jelentős szerepet játszottak a kompromittált hitelesítő adatok, a rosszindulatú e-mail üzenetek és az adathalász támadások. A szervezeti háttértényezők vizsgálata során a kutatás arra a következtetésre jutott, hogy a támadások jelentős részében a szakértelem hiánya, a fel nem ismert biztonsági rések, valamint a rendelkezésre álló humán erőforrások korlátozottsága is hozzájárult a sikeres kompromittációhoz. A válaszadók több mint 60%-a jelölt meg olyan humán vagy szervezeti tényezőt, amely közvetlenül szerepet játszott az incidens bekövetkezésében. Ezek az eredmények összhangban állnak a jelen kutatás alapfeltevéseivel, amelyek szerint a technológiai védelmi intézkedések önmagukban nem elegendőek a ransomware-fenyegetések hatékony kezeléséhez. A kérdőíves felmérés során vizsgált informatikai döntéshozói kompetenciák, a szervezeti biztonsági gyakorlatok és a védelmi érettség szintje egyaránt olyan tényezők, amelyek jelentős hatással lehetnek a szervezetek ellenállóképességére. A pénzügyi következmények összehasonlítása ugyanakkor csak korlátozott mértékben lehetséges. A Sophos kutatása szerint a medián váltságdíj-kifizetés elérte az 1 millió USD-t, míg a helyreállítás átlagos költsége 1,53 millió USD volt. A jelen kutatás fókuszában ugyanakkor a magyar mikro-, kis- és középvállalkozások álltak, amelyek gazdasági lehetőségei jelentősen eltérnek a nemzetközi felmérésben szereplő szervezetekétől. A KKV-szektorban már egy jóval kisebb volumenű incidens is súlyos működési problémákat okozhat, mivel a helyreállítás költségei, az üzletmenet-kiesés és az adatvesztés következményei sok esetben a vállalkozás több éves nyereségét vagy akár működőképességét is veszélyeztethetik [108]. A két kutatás eredményei alapján megállapítható, hogy a ransomware-fenyegetések mögött meghúzódó technikai és szervezeti okok nemzetközi szinten és a hazai KKV-k körében is hasonló mintázatokat mutatnak. Ugyanakkor a korlátozott pénzügyi, humán és szakmai erőforrások miatt a KKV-k sérülékenysége és a támadások következményeinek súlya arányosan nagyobb lehet. Ez alátámasztja olyan költséghatékony, automatizált és minimális szakértői beavatkozást igénylő védelmi megoldások fejlesztésének szükségességét, amelyek kifejezetten a kis- és középvállalkozások sajátosságaihoz igazodnak. A Sophos eredményei összhangban állnak a Microsoft Digital Defense Report, a CrowdStrike Global Threat Report, valamint az ENISA és Verizon legfrissebb jelentéseiben megfogalmazott trendekkel, amelyek egyaránt arra utalnak, hogy a zsarolóvírusok továbbra is a legjelentősebb kiberfenyegetések közé tartoznak, miközben egyre nagyobb szerepet kapnak a viselkedésalapú felismerési módszerek [4,13,14,105–108].

8.3.1. A tudományos eredmények összevetése

Az IEEE Access áttekintő tanulmány [99] és Alraizza és Algarni összefoglaló munkája [100] egyaránt arra a következtetésre jut, hogy a ransomware-felismerés kutatásában egyre nagyobb jelentősége van a viselkedésalapú jellemzők és a felügyelt gépi tanulási algoritmusok alkalmazásának. E kutatások szerint a hagyományos szignatúraalapú felismerési módszerek önmagukban már nem biztosítanak megfelelő védelmet, ezért a korszerű kutatások a fájlrendszeri események, naplóállományok és egyéb telemetriai adatok elemzésére helyezik a hangsúlyt [99,100]. A dolgozatban alkalmazott XGBoost modell választását Al-Hashmi és

munkatársainak eredményei is alátámasztják, akik több klasszikus és ensemble gépi tanulási algoritmus összehasonlítása során szintén az XGBoost modellt az egyik legeredményesebb megoldásnak találták korai ransomware-detektálási feladatokban [102].

8.3.2. A dolgozat tudományos pozicionálása

A legújabb publikációk ugyanakkor jól érzékeltetik azt is, hogy a jelen értekezés milyen szempontból képvisel önálló kutatási irányt. A legtöbb korszerű megoldás végponti eseményeket, memóriahasználatot, rendszerhívásokat, hardveres teljesítményszámlálókát vagy blokkszintű tárolóműveleteket elemez [101,103,104]. Figyelemre méltó ugyanakkor, hogy a legújabb áttekintő tanulmányok [99,100] is rámutatnak arra, hogy jelenleg kevés olyan publikált megoldás áll rendelkezésre, amely kizárólag szerveroldali fájlműveleti naplók alapján végez ransomware-detektálást. Ez tovább erősíti a jelen értekezésben bemutatott megközelítés újdonságát, jelen kutatásom kizárólag szerveroldali fájlműveleti naplófájlokra épül, centralizált fájl tárolási környezetben, amely lehetővé teszi a kliensoldali operációs rendszerektől és végpontvédelmi megoldásoktól nagyrészt független, kvázi valós idejű ransomware-detektálást. Ez a megközelítés egyszerre biztosítja a jó skálázhatóságot és a meglévő vállalati infrastruktúrába történő integrálhatóságot. Összességében megállapítható, hogy a doktori kutatás lezárását követően megjelent tudományos publikációk és iparági fenyegetettségi jelentések nem módosítják a dolgozat fő következtetéseit, hanem azok érvényességét és aktualitását támasztják alá. A viselkedésalapú, szerveroldali fájlműveleti naplók elemzésére épülő gépi tanulási megközelítés a legfrissebb kutatások tükrében is korszerű és ígéretes irányt képvisel, amely különösen a korlátozott erőforrásokkal rendelkező kis- és középvállalkozások számára jelenthet költséghatékony védelmi lehetőséget [99–108].

8.4. Jövőbeni kutatási irányok és feladatok

A kutatás eredményei több további kutatási irányt is kijelölnek. Egyik fontos terület a referencia-alapú teljesítményértékelési módszer kiterjesztése más típusú kiszolgálói környezetekre és szolgáltatásplatformokra. Ez lehetővé tenné a heterogén infrastruktúrák még szélesebb körű összehasonlítását. További kutatási lehetőséget jelent a mélytanulási modellek (pl.: LSTM, RNN) adaptív tanulási képességének vizsgálata. A környezet változásaihoz automatikusan alkalmazkodó modellek csökkenthetnék a manuális újra tanítás szükségességét, növelhetnék a hosszútávú működési stabilitást és csökkenthetnék a fals pozitív jelzések számát. Érdemes lenne vizsgálni továbbá a fájlkiszolgáló-alapú detektálás integrálását más védelmi rétegekkel, például monitoring rendszerekkel vagy incidenskezelő platformokkal. A jövőbeni kutatások másik lehetséges iránya az új mentési technológiák kidolgozása helyett inkább a már meglévő mentési és helyreállítási megoldások kombinációjának és integrációjának vizsgálata a fájlkiszolgáló-alapú detektálási mechanizmusokkal. Különösen indokolt lehet annak az elemzése, hogy a detektálási események miként használhatók fel automatikus

mentésekkel kapcsolatos döntések támogatására, mentési ütemezések módosítására, a pillanatképek élettartamának finomhangolására vagy a helyreállítási folyamatok automatizált indítására. Mivel a zsarolóvírusok jelentős része az adatvagyonát a titkosítás előtt a támadó tárolórendszerére replikálja, célszerű lenne a későbbiekben a fájlműveletek közül az olvasási műveleteket is figyelemmel kísérni akár tisztán statisztikai módszerekkel, akár mesterséges intelligencia alapú megoldásokkal. További kutatások szükségesek az üzemeltetők integrációra való nyitottságának faktorai és a döntéshozók döntési szempontjai közötti összefüggések mélyebb megismeréséhez. A különböző iparágak és vállalatméretek összehasonlító vizsgálata hozzájárulhat ahhoz, hogy a technikai megoldások jobban illeszkedjenek az eltérő iparági elvárásokhoz, illetve a különböző szervezeti struktúrákhoz.

Irodalomjegyzék

- [1] „Digital 2025 Global Overview Report”, Datareportal, 2025, URL: {<https://data-reportal.com/reports/digital-2025-global-overview-report>} (Utolsó hozzáférés: 2025. december 22.)
- [2] „The State of Ransomware 2024”, Sophos News, 2024, URL: {<https://news.sophos.com/en-us/2024/04/30/the-state-of-ransomware-2024>} (Utolsó hozzáférés: 2025. december 22.)
- [3] „Ransomware Annual Report 2024”, Cyberint, 2024, URL: {<https://cyberint.com/blog/research/ransomware-annual-report-2024>} (Utolsó hozzáférés: 2025. december 22.)
- [4] ENISA, „ENISA Threat Landscape: Finance Sector 2024”, 2025, URL: {https://www.enisa.europa.eu/sites/default/files/2025-02/Finance%20TL%202024_Final.pdf} (Utolsó hozzáférés: 2025. december 22.)
- [5] FBI, USA, „Ransomware”, URL: {<https://www.fbi.gov/how-we-can-help-you/safety-resources/scams-and-safety/common-scams-and-crimes/ransomware>} (Utolsó hozzáférés: 2025. december 22.)
- [6] Adams, Delaney, „Healthcare Ransomware Attacks: Understanding the Problem and How to Protect Your Organization”, LogRhythm, 2024, URL: {<https://securityboulevard.com/2024/03/healthcare-ransomware-attacks-understanding-the-problem-and-how-to-protect-your-organization>}
- [7] BITCOIN -USD rate, Google finance, 2025. URL: {<https://www.google.com/finance/quote/BTC-USD>} (Utolsó hozzáférés: 2025. december 22.)
- [8] Jan Lemnitzer, „Ransomware gangs are running riot – paying them off doesn’t help”, URL: {<https://theconversation.com/ransomware-gangs-are-running-riot-paying-them-off-doesnt-help-155254>} (Utolsó hozzáférés: 2025. december 22.)
- [9] Vumetric, „More than 66% of healthcare organizations experienced a ransomware attack in 2019”, URL: {<https://www.vumetric.com/statistics/more-than-66-of-healthcare-organizations-experienced-a-ransomware-attack-in-2019/>} (Utolsó hozzáférés: 2025. december 22.)
- [10] Juliana De Groot, „A History of Ransomware Attacks: The Biggest and Worst Ransomware Attacks of All Time”, URL: {<https://digitalguardian.com/blog/history-ransomware-attacks-biggest-and-worst-ransomware-attacks-all-time>} (Utolsó hozzáférés: 2025. december 22.)
- [11] Daniela Ruckwied, „Cyber-Angriff mit Todesfolge an der Uniklinik Düsseldorf”, URL: {<https://www.dsin-blog.de/2020/10/14/cyber-angriff-mit-todesfolge-an-der-uniklinik-duesseldorf/>} (Utolsó hozzáférés: 2025. december 22.)
- [12] European Union Agency for Cybersecurity (ENISA), „ENISA Threat Landscape 2024”, 2024, URL: {<https://www.enisa.europa.eu/publications/enisa-threat-landscape-2024>} (Utolsó hozzáférés: 2025. december 22.)

- [13] Verizon, „Verizon Data Breach Investigations Report 2025”, 2025, URL: {<https://www.verizon.com/business/resources/reports/dbir/>} (Utolsó hozzáférés: 2025. december 22.)
- [14] Europol, „Internet Organised Crime Threat Assessment (IOCTA) 2024”, 2024, DOI: 10.2813/442713, URL: {<https://www.europol.europa.eu/cms/sites/default/files/documents/Internet%20Organised%20Crime%20Threat%20Assessment%20IOCTA%202024.pdf>} (Utolsó hozzáférés: 2025. december 22.)
- [15] IBM, USA, „IBM Security QRadar XDR”, URL: {<https://www.ibm.com/qradar>} (Utolsó hozzáférés: 2025. december 22.)
- [16] Natalie, Paskoski, „Remote Desktop Protocol Use in Ransomware Attacks”, URL: {<https://learn.microsoft.com/en-us/troubleshoot/windows-client/performance/troubleshooting-slow-file-copying-in-windows>} (Utolsó hozzáférés: 2025. december 22.)
- [17] Huang C, Deep A, Zhou S, Veeramani D. A deep learning approach for predicting critical events using event logs. Qual Reliab Eng Int. 2021; 37: 2214–2234. <https://doi.org/10.1002/qre.2853>
- [18] Bae SI, Lee GB, Im EG. Ransomware detection using machine learning algorithms. Concurrency Computat Practice and Experience. Wiley, 2020; 32:e5422. DOI: <https://doi.org/10.1002/cpe.5422>
- [19] Yousaf, Bux Khuda; Yousaf, Muhammad; Jalbani, Akhtar Hussain; Batool, Komal, „Detection of Malicious Servers for Preventing Client-Side Attacks”, MEHRAN UNIVERSITY RESEARCH JOURNAL OF ENGINEERING AND TECHNOLOGY, vol. 40, no. 1, pp. 230–240, 2021, DOI: 10.22581/muet1982.2101.20
- [20] Lemmou, Yassine; Lanet, Jean-Louis; Souidi, El Mamoun, „A behavioural in-depth analysis of ransomware infection”, IET INFORMATION SECURITY, vol. 15, no. 1, pp. 38–58, 2021, DOI: 10.1049/ise2.12004
- [21] P.A., Shemitha; Dhas, Julia Punitha Malar, „Research perceptions on ransomware attack: a complete analysis on conventional authentication protocols in network”, EVOLUTIONARY INTELLIGENCE, pp. 16, 2020, DOI: 10.1007/s12065-020-00502-9
- [22] Croke, Lisa, „Protecting your organization from e-mail phishing and ransomware attacks”, AORN JOURNAL, vol. 112, no. 4, pp. 10–12, 2020, DOI: 10.1002/aorn.13229
- [23] Naseer, Ayesha; Mir, Riffat; Mir, Azmat; Aleem, Muhammad, „Windows-based Ransomware: A Survey”, JOURNAL OF INFORMATION ASSURANCE AND SECURITY, vol. 15, no. 3, pp. 107–125, 2020
- [24] Mohammad, Adel Hamdan, „Analysis of Ransomware on Windows platform”, INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND NETWORK SECURITY, vol. 20, no. 6, pp. 21–27, 2020. DOI: 10.13140/RG.2.2.11150.59202
- [25] Hampton, Nikolai; Baig, Zubair; Zeadally, Sherali, „Ransomware behavioural analysis on windows platforms”, JOURNAL OF INFORMATION SECURITY AND APPLICATIONS, vol. 40, pp. 44–51, 2018, DOI: 10.1016/j.jisa.2018.02.008

- [26] Al-rimy, Bander Ali Saleh; Maarof, Mohd Aizaini; Shaid, Syed Zainudeen Mohd, „Ransomware threat success factors, taxonomy, and countermeasures: A survey and research directions”, Elsevier, COMPUTERS & SECURITY, vol. 74, pp. 144–166, 2018, DOI: 10.1016/j.cose.2018.01.001
- [27] Zimba, Aaron; Mulenga, Mwenge, „A DIVE INTO THE DEEP: DEMYSTIFYING WANNACRY CRYPTO RANSOMWARE NETWORK ATTACKS VIA DIGITAL FORENSICS”, INTERNATIONAL JOURNAL ON INFORMATION TECHNOLOGIES AND SECURITY, vol. 10, no. 2, pp. 57–68, 2018, URL: {https://www.researchgate.net/publication/325678210_A_Dive_into_the_Deep_Demystifying_WannaCry_Crypto_Ransomware_Network_Attacks_Via_Digital_Forensics}
- [28] Lee, Suhyeon; Kim, Huy Kang; Kim, Kyounggon, „Ransomware protection using the moving target defense perspective”, COMPUTERS & ELECTRICAL ENGINEERING, vol. 78, pp. 288–299, 2019, DOI: 10.1016/j.compeleceng.2019.07.014
- [29] N. Chaabouni, M. Mosbah, A. Zemmari, C. Sauvignac and P. Faruki, "Network Intrusion Detection for IoT Security Based on Learning Techniques," in *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2671-2701, thirdquarter 2019, DOI: 10.1109/COMST.2019.2896380.
- [30] Castiglione, Jason; Pavlovic, Dusko, „Dynamic Distributed Secure Storage Against Ransomware”, IEEE TRANSACTIONS ON COMPUTATIONAL SOCIAL SYSTEMS, vol. 7, no. 6, pp. 1469–1475, 2020, DOI: 10.1109/TCSS.2019.2924650
- [31] Alomary, Fahad Omar, „Defend Against Ransomware Detection Using Intrusion Detection System (IDS)”, INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND NETWORK SECURITY, vol. 20, no. 7, pp. 11–16, 2020, URL: {http://paper.ijcsns.org/07_book/202007/20200702.pdf}
- [32] Keong Ng, Chee; Rajasegarar, Sutharshan; Pan, Lei; Jiang, Frank; Zhang, Leo Yu, „VoterChoice: A ransomware detection honeypot with multiple voting framework”, CONCURRENCY AND COMPUTATION-PRACTICE & EXPERIENCE, vol. 32, no. 14, pp. 29, 2019, DOI: 10.1002/cpe.5726
- [33] Kim, DaeYoub; Lee, Jihoon, „Blacklist vs. Whitelist-Based Ransomware Solutions”, IEEE CONSUMER ELECTRONICS MAGAZINE, vol. 9, no. 3, pp. 22–28, 2020, DOI: 10.1109/MCE.2019.2956192
- [34] Cabaj, Krzysztof; Gregorczyk, Marcin; Mazurczyk, Wojciech, „Software-defined networking-based crypto ransomware detection using HTTP traffic characteristics”, COMPUTERS & ELECTRICAL ENGINEERING, vol. 66, pp. 353–368, 2018, DOI: 10.1016/j.compeleceng.2017.10.012
- [35] An official website of the European Union, „Internal Market, Industry, Entrepreneurship and SMEs”, URL: {https://ec.europa.eu/growth/smes/sme-definition_en} (Utolsó hozzáférés: 2025. december 22.)
- [36] D. Clark, „Number of small and medium-sized enterprises (SMEs) the European Union in 2018”, URL: {<https://www.statista.com/statistics/878412/number-of-smes-in-europe-by-size/>} (Utolsó hozzáférés: 2025. december 22.)

- [37] Cybersecurity and Infrastructure Security Agency, #StopRansomware: Zeppelin Ransomware URL: {<https://www.cisa.gov/news-events/cybersecurity-advisories/aa22-223a>} (Utolsó hozzáférés: 2025. december 22.)
- [38] IVANTI Pulse Secure, URL: {https://forums.ivantipulse.com/s/pulse-secure-desktop-client?language=en_US} (Utolsó hozzáférés: 2025. december 22.)
- [39] Fortinet Global Security, URL: {<https://www.fortinet.com/>} (Utolsó hozzáférés: 2025. december 22.)
- [40] Szücs, Veronika; Arányi, Gábor; Dávid, Ákos, „Introduction of the ARDS—Anti-Ransomware Defense System Model—Based on the Systematic Review of Worldwide Ransomware Attacks”, *Applied Sciences*, vol. 11, no. 6070, 2021, DOI: 10.3390/app11136070, URL: {<https://www.mdpi.com/2076-3417/11/13/6070>} (Utolsó hozzáférés: 2025. december 22.)
- [41] Ngo, Fawn T.; Agarwal, Anurag; Govindu, Ramakrishna; MacDonald, Calen, „Malicious Software Threats”, *The Palgrave Handbook of International Cybercrime and Cyberdeviance*, ; Holt, T.J.; Bossler, A.M., Eds.; Springer International Publishing: Cham, 2020; pp. 793–813. DOI: 10.1007/978-3-319-78440-3_35
- [42] Vanness, Robert; Chowdhury, Md Minhaz; Rifat, Nafiz, „Malware: A Software for Cybercrime”, 2022 IEEE International Conference on Electro Information Technology (eIT), pp. 513-518, 2022, DOI: 10.1109/eIT53891.2022.9813970
- [43] Jinqian Liang; Xiaohong Guan, „A virtual disk environment for providing file system recovery”, *Computers & Security*, vol. 25, no. 8, pp. 589-599, 2006, DOI: <https://doi.org/10.1016/j.cose.2006.08.002>, URL: {<https://www.sciencedirect.com/science/article/pii/S0167404806001209>} (Utolsó hozzáférés: 2025. december 22.)
- [44] Gaonkar, S.; Keeton, K.; Merchant, A.; Sanders, W.H., „Designing dependable storage solutions for shared application environments”, *International Conference on Dependable Systems; Networks (DSN'06)*, pp. 371-382, 2006, DOI: 10.1109/DSN.2006.27
- [45] Van Oorschot, Paul C.; Wurster, Glenn, „Reducing Unauthorized Modification of Digital Objects”, *IEEE Transactions on Software Engineering*, vol. 38, no. 1, pp. 191-204, 2012, DOI: 10.1109/TSE.2011.7
- [46] Nikolay Chervyakov; Mikhail Babenko; Andrei Tchernykh; Nikolay Kucherov; Vanessa Miranda-López; Jorge M. Cortés-Mendoza, „AR-RRNS: Configurable reliable distributed data storage systems for Internet of Things to ensure security”, *Future Generation Computer Systems*, vol. 92, pp. 1080-1092, 2019, DOI: <https://doi.org/10.1016/j.future.2017.09.061>, URL: {<https://www.sciencedirect.com/science/article/pii/S0167739X17306015>} (Utolsó hozzáférés: 2025. december 22.)
- [47] Rajat Verma; Ajay Mendez; Stan Park; Sandya Mannarswamy; Terence Kelly; Charles B Morrey III, „Failure-Atomic Updates of Application Data in a Linux File System”, *Proceedings of the 13th USENIX Conference on File and Storage Technologies (FAST '15)*, pp. 203-212, 2015
- [48] Newberry, Eric; Zhang, Beichuan, „On the Power of In-Network Caching in the Hadoop Distributed File System”, *Proceedings of the 6th ACM Conference on Information-Centric Networking*, pp. 89–99, 2019, DOI: 10.1145/3357150.3357392,

URL: {<https://doi.org/10.1145/3357150.3357392>} (Utolsó hozzáférés: 2025. december 22.)

- [49] Makris, Antonios; Kontopoulos, Ioannis; Psomakelis, Evangelos; Xyalis, Stylianos Nektarios; Theodoropoulos, Theodoros; Tserpes, Konstantinos, „Performance Analysis of Storage Systems in Edge Computing Infrastructures”, *Applied Sciences*, vol. 12, no. 8923, 2022, DOI: 10.3390/app12178923, URL: {<https://www.mdpi.com/2076-3417/12/17/8923>} (Utolsó hozzáférés: 2025. december 22.)
- [50] S.C. BITDEFENDER S.R.L., Romania, „GravityZone Business Security Premium”, URL: {<https://www.bitdefender.com/business/products/gravityzone-premium-security.html>} (Utolsó hozzáférés: 2025. december 22.)
- [51] Sophos Limited, United Kingdom, „The World’s Best Endpoint Protection XDR, EDR, ZTNA, MDR Services”, URL: {<https://www.sophos.com/en-us/products/endpoint-anti-virus>} (Utolsó hozzáférés: 2025. december 22.)
- [52] Trend Micro Incorporated, USA, „propel Security Operations Forward”, URL: {https://www.trendmicro.com/en_us/business/products/security-operations.html} (Utolsó hozzáférés: 2025. december 22.)
- [53] Natalie Paskoski, RH-ISAC Manager of Marketing & Communications, „Troubleshooting slow file copying in Windows”, URL: {<https://rhisac.org/ransomware/remote-desktop-protocol-use-in-ransomware-attacks/>} (Utolsó hozzáférés: 2025. december 22.)
- [54] GeekPage, Inc, „Fix Slow file copy speed in Windows 10 / 11”, URL: {<https://thegeek-page.com/fix-slow-file-copy-speed-in-windows-10/>} (Utolsó hozzáférés: 2025. december 22.)
- [55] AOMEI, Inc, „Windows 10 File Copy/Transfer Very Slow”, URL: {<https://www.ubackup.com/windows-10/windows-10-file-copy-slow-1021.html>} (Utolsó hozzáférés: 2025. december 22.)
- [56] Synology, Inc, „What can I do when the file transfer via Windows (SMB/CIFS) is slow?”, URL: {https://kb.synology.com/hu-hu/DSM/tutorial/What_can_I_do_when_the_file_transfer_via_Windows_SMB_CIFS_is_slow} (Utolsó hozzáférés: 2025. december 22.)
- [57] Ye, Xin; Zhai, Zhengjun; Li, Xiaochang, „ZDC: A Zone Data Compression Method for Solid State Drive Based Flash Memory”, *Symmetry*, vol. 12, pp. 623, 2020, DOI: 10.3390/sym12040623
- [58] LockBit ransomware, Nemzeti Kiberbiztonsági Intézet, URL: {<https://nki.gov.hu/figyelmezteteselek/karos-kod/lockbit-zsarolovirus/>} (Utolsó hozzáférés: 2025. december 22.)
- [59] Guo, Yang, „A review of Machine Learning-based zero-day attack detection: Challenges and future directions”, *COMPUTER COMMUNICATIONS*, vol. 198, pp. 175-185, 2023, DOI: 10.1016/j.comcom.2022.11.001
- [60] Singh, Priyanka; Borgohain, Samir Kumar; Sarkar, Achintya Kumar; Kumar, Jayendra; Sharma, Lakhan Dev, „Feed-Forward Deep Neural Network (FFDNN)-Based Deep Features for Static Malware Detection”, *International Journal of Intelligent Systems*, vol. 2023, no. 1, pp. 9544481, 2023, DOI: <https://doi.org/10.1155/2023/9544481>,

URL: {<https://onlinelibrary.wiley.com/doi/abs/10.1155/2023/9544481>} (Utolsó hozzáférés: 2025. december 22.)

- [61] Kaur, Ratinder; Singh, Maninder, „A Survey on Zero-Day Polymorphic Worm Detection Techniques”, IEEE COMMUNICATIONS SURVEYS AND TUTORIALS, vol. 16, no. 3, pp. 1520-1549, 2014, DOI: 10.1109/SURV.2014.022714.00160
- [62] Ahmad, Rasheed; Alsmadi, Izzat; Alhamdani, Wasim; Tawalbeh, Lo'ai, „Zero-day attack detection: a systematic literature review”, ARTIFICIAL INTELLIGENCE REVIEW, vol. 56, no. 10, pp. 10733-10811, 2023, DOI: 10.1007/s10462-023-10437-z
- [63] Singh, Umesh Kumar; Joshi, Chanchala; Kanellopoulos, Dimitris, „A framework for zero-day vulnerabilities detection and prioritization”, JOURNAL OF INFORMATION SECURITY AND APPLICATIONS, vol. 46, pp. 164-172, 2019, DOI: 10.1016/j.jisa.2019.03.011
- [64] Sebastian Guisao, Juan; Toro Rendon, Juan Carlos, „Mitigation and Detection of Zero-Day Vulnerabilities”, CUADERNO ACTIVA, no. 6, pp. 63-67, 2014
- [65] A. AlEroud, G. Karabatis „A Contextual Anomaly Detection Approach to Discover Zero-Day Attacks”, 2012 ASE INTERNATIONAL CONFERENCE ON CYBER SECURITY (CYBERSECURITY), (Acad Sci & Engn; IEEE Comp Soc, 2012), pp. 40–45. DOI: <https://doi.org/10.1109/CyberSecurity.2012.12.>, URL: {https://www.researchgate.net/publication/233942490_A_Contextual_Anomaly_Detection_Approach_to_Discover_Zero-Day_Attacks}
- [66] Cen, Mingcan; Jiang, Frank; Qin, Xingsheng; Jiang, Qinghong; Doss, Robin, „Ransomware early detection: A survey”, Comput. Networks, vol. 239, pp. 110138, 2024, DOI: 10.1016/j.comnet.2023.110138
- [67] Xu, Anson; Choudhury, Arnav; Liu, Eason; Choi, Sean, „ProtectNIC: SmartNIC-based Ransomware Detection”, 2024 Silicon Valley Cybersecurity Conference (SVCC), pp. 1-3, 2024, DOI: 10.1109/SVCC61185.2024.10637308
- [68] McIntosh, Timothy ; Kayes, A. S. M.; Chen, Yi-Ping; Ng, Alex; Watters, Paul. (2023). Applying Staged Event-Driven Access Control to Combat Ransomware. Computers & Security. DOI: 128. 103160. 10.1016/j.cose.2023.103160.
- [69] Denham, Byron; Thompson, Dale R., „Analysis of Decoy Strategies for Detecting Ransomware”, 2023 IEEE Conference on Communications and Network Security (CNS), pp. 1-6, 2023, DOI: 10.1109/CNS59707.2023.10288691
- [70] Diamantopoulos, Dionysios; Pletka, Roman; Sarafijanovic, Slavisa; Reddy, A. L. Narasimha; Pozidis, Haris, „WannaLaugh: A Configurable Ransomware Emulator - Learning to Mimic Malicious Storage Traces”, ACM Conferences, pp. 118–131, 2024, DOI: 10.1145/3688351.3689163
- [71] Hindy, Hanan; Atkinson, Robert; Tachtatzis, Christos; Colin, Jean-Noel; Bayne, Ethan; Bellekens, Xavier, „Utilising Deep Learning Techniques for Effective Zero-Day Attack Detection”, ELECTRONICS, vol. 9, no. 1684, 2020, DOI: 10.3390/electronics9101684
- [72] Deldar, Fatemeh; Abadi, Mahdi, „Deep Learning for Zero-day Malware Detection; Classification: A Survey”, ACM COMPUTING SURVEYS, vol. 56, no. 36, 2024, DOI: 10.1145/3605775

- [73] Millar, Stuart; McLaughlin, Niall; del Rincon, Jesus Martinez; Miller, Paul, „Multi-view deep learning for zero-day Android malware detection”, JOURNAL OF INFORMATION SECURITY AND APPLICATIONS, vol. 58, no. 102718, 2021, DOI: 10.1016/j.jisa.2020.102718
- [74] Ganame, Karim; Allaire, Marc Andre; Zagdene, Ghassen; Boudar, Oussama, „Network Behavioral Analysis for Zero-Day Malware Detection - A Case Study”, INTELLIGENT, SECURE, AND DEPENDABLE SYSTEMS IN DISTRIBUTED AND CLOUD ENVIRONMENTS (ISDDC 2017), vol. 10618, pp. 169-181, 2017, DOI: 10.1007/978-3-319-69155-8_13
- [75] Wangde, I, Fawaz; Mulay, Shivam P.; Adhao, Rahul B.; Pachghare, Vinod K., „Zero-Day Attack Detection using Ensemble Technique”, INTERNATIONAL JOURNAL OF NEXT-GENERATION COMPUTING, vol. 12, no. 5, SI, pp. 551-557, 2021, DOI: 10.47164/ijngc.v12i5.423
- [76] Rushdan, Huthifh; Shurman, Mohammad; Alnabelsi, Sharhabeel; Qutaibah, Althebyan. (2019). Zero-Day Attack Detection and Prevention in Software-Defined Networks. DOI: 10.1109/ACIT47987.2019.8991124.
- [77] Zhangying He, Amin Rezaei, Houman Homayoun, and Hossein Sayadi. 2022. Deep Neural Network and Transfer Learning for Accurate Hardware-Based Zero-Day Malware Detection. In Proceedings of the Great Lakes Symposium on VLSI 2022 (GLSVLSI '22). Association for Computing Machinery, New York, NY, USA, 27–32. <https://doi.org/10.1145/3526241.3530326>
- [78] Bu, Seok-Jun; Cho, Sung-Bae, „Deep Character-Level Anomaly Detection Based on a Convolutional Autoencoder for Zero-Day Phishing URL Detection”, ELECTRONICS, vol. 10, no. 1492, 2021, DOI: 10.3390/electronics10121492
- [79] F. Manzoor, V. Khattar, C. -C. Liu and M. Jin, "Zero-day Attack Detection in Digital Substations using In-Context Learning," 2024 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), Oslo, Norway, 2024, pp. 220-225, DOI: 10.1109/SmartGridComm60555.2024.10738025
- [80] Kim, Chiho; Chang, Sang-Yoon; Kim, Jonghyun; Lee, Dongeun; Kim, Jinoh, „Automated, Reliable Zero-Day Malware Detection Based on Autoencoding Architecture”, IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, vol. 20, no. 3, pp. 3900-3914, 2023, DOI: 10.1109/TNSM.2023.3251282
- [81] Zahoor, Umme; Aziz, Abdul; Ashraf, Arslan; Rajarajan, Muttukrishnan; Khan, Asifullah, „Ransomware detection using deep learning based unsupervised feature extraction and a cost sensitive Pareto Ensemble classifier” Scientific Reports, vol. 12, art.num.15647, 2022, DOI: 10.1038/s41598-022-19443-7
- [82] Baldwin, John; Dehghantanha, Ali, „Efficacy of Opcode Density via SVM for Crypto-Ransomware Detection”, Cyberthreat Intelligence, Springer, arXiv preprint arXiv:1807.10442, 2018, URL: {<https://arxiv.org/abs/1807.10442>} (Utolsó hozzáférés: 2025. december 22.)

- [83] Ahmed, Muhammad Ejaz; Kim, Hyounghshick; Camtepe, Seyit; Nepal, Surya, „Peeler: Profiling Kernel-Level Events to Detect Ransomware”, ESORICS (Lecture Notes in Computer Science, vol. 12972), pp. 240–260, 2021, DOI: 10.1007/978-3-030-88418-5_12
- [84] Hitaj, Dorjan; Pagnotta, Giulio; De Gaspari, Fabio; De Carli, Lorenzo; Mancini, Luigi V., „Minerva: A File-Based Ransomware Detector”, arXiv preprint arXiv:2301.11050, 2023, ASIA CCS '25: Proceedings of the 20th ACM Asia Conference on Computer and Communications Security, pages 576-590, URL: {<https://doi.org/10.48550/arXiv.2301.11050>}
- [85] Herrera-Silva, J.A.; Hernández-Álvarez, M.; „Dynamic Feature Dataset for Ransomware Detection Using Behavioral Patterns”, Sensors, vol. 23, no. 3, 2023, DOI: 10.3390/s23031053
- [86] Lee, J.; Kim, J.; Jeong, H.; Lee, K. A Machine Learning-Based Ransomware Detection Method for Attackers’ Neutralization Techniques Using Format-Preserving Encryption. Sensors 2025, 25, 2406. <https://doi.org/10.3390/s25082406>
- [87] Zhou, Chijin; Guo, Lihua; Hou, Yiwei; Ma, Zhenya; Zhang, Quan; Wang, Mingzhe; Liu, Zhe; Jiang, Yu, „Limits of I/O Based Ransomware Detection: An Imitation Based Attack”, Proceedings of IEEE Symposium on Security and Privacy (S&P), pp. 2584–2601, 2023, DOI: 10.1109/SP46215.2023.10179372
- [88] Panos Kampanakis, „Practical Challenges with AES-GCM: Performance, Side Channels, Misuse Resistance”, NIST Workshop on Block Cipher Modes of Operation, 2024, URL: {<https://csrc.nist.gov/csrc/media/Events/2023/third-workshop-on-block-cipher-modes-of-operation/documents/accepted-papers/Practical%20Challenges%20with%20AES-GCM.pdf>} (Utolsó hozzáférés: 2025. december 22.)
- [89] Shay Gueron; Vlad Krasnov, „AES-GCM Software Performance on Intel Architecture Processors”, Proceedings of the 12th International Conference on Cryptology and Network Security (CANS), pp. 65–93, 2013, DOI: 10.1007/978-3-319-02937-5_5
- [90] Andrew Barnes; Jeremy Smith; Daniel Greg, „Improving the throughput of the AES algorithm with multicore processors”, arXiv preprint arXiv:1403.7295, 2014, URL: {<https://arxiv.org/abs/1403.7295>} (Utolsó hozzáférés: 2025. december 22.)
- [91] Kyoungwoo Kim; Hojin Lee; Jongin Lim, „Practical AES-GCM Encryption for Low-End Microcontrollers”, Applied Sciences, vol. 10, no. 9, pp. 3131, 2020, DOI: 10.3390/app10093131
- [92] Kwame Assa-Agyei, Funminiyl Olajide, Temitope Alade, „Optimizing the Performance of the Advanced Encryption Standard Techniques for Secured Data Transmission”, International Journal of Computer Applications, vol. 18, no. 21, pp. 31–36, 2023, DOI: 10.5120/ijca2023922941 (Utolsó hozzáférés: 2025. december 22.)
- [93] Kothamali, Praveen; Banik, Subhra, „Limitations of Signature-Based Threat Detection”, International Journal of Computer Applications, 2025, URL: {https://www.researchgate.net/publication/388494583_Limitations_of_Signature-Based_Threat_Detection} (Utolsó hozzáférés: 2025. december 22.)

- [94] Kim, G.Y., „Byte Frequency Based Indicators for Crypto-Ransomware”, *Journal of Computer Science and Technology*, vol. 37, no. 2, pp. 423-442, 2022, DOI: 10.1007/s11390-022-0263-4, URL: {<https://jcst.ict.ac.cn/fileup/1000-9000/PDF/2022-2-10-0263.pdf>} DOI: 10.1007/s11390-021-0263-x (Utolsó hozzáférés: 2025. december 22.)
- [95] McIntosh, Shane; Jang-Jaccard, Julianne; Watters, Paul; Susnjak, Teo, „The Inadequacy of Entropy-Based Ransomware Detection”, *Communications in Computer and Information Science*, 2019, URL: {https://www.researchgate.net/publication/337776801_The_Inadequacy_of_Entropy-Based_Ransomware_Detection} DOI: 10.1007/978-3-030-36802-9_20 (Utolsó hozzáférés: 2025. december 22.)
- [96] Van Sloun, Christian; Woeste, Vincent; Wolsing, Konrad; Pennekamp, Jan; Wehrle, Klaus, „Detecting Ransomware Despite I/O Overhead: A Practical Multi-Staged Approach”, *Proceedings of the 22nd Annual Network and Distributed System Security Symposium (NDSS '25)*, 2025, DOI: 10.14722/ndss.2025.240764
- [97] Casino, Fran; Patsakis, Constantinos; Batista, Edgar; Fayyaz, Muhammad; Garcia, Pedro, „Not on My Watch: Ransomware Detection through Classification of Encrypted Streams”, *Journal of Cybersecurity*, vol. 11, no. 1, pp. tyaf009, 2025, DOI: 10.1093/cybsec/tyaf009, URL: {<https://academic.oup.com/cybersecurity/article/11/1/tyaf009/8109429>} (Utolsó hozzáférés: 2025. december 22.)
- [98] [CISA Cybersecurity and Infrastructure Security Agency https://www.cisa.gov/sites/default/files/publications/data_backup_options.pdf]
- [99] Ispahany, Jamil & Islam, Rafiqul & Khan, M. & Islam, Zahidul. (2024). Ransomware Detection Using Machine Learning: A Review, Research Limitations and Future Directions. *IEEE Access*. PP. 1-1. 10.1109/ACCESS.2024.3397921.
- [100] Alraizza, Amjad, and Abdulmohsen Algarni. 2023. "Ransomware Detection Using Machine Learning: A Survey" *Big Data and Cognitive Computing* 7, no. 3: 143. <https://doi.org/10.3390/bdcc7030143>
- [101] Hill, Jennie & Walker, T. & Blanco, Justin & Ives, Robert & Rakvic, Ryan & Jacob, Bruce. (2024). Ransomware Classification Using Hardware Performance Counters on a Non-Virtualized System. *IEEE Access*. PP. 1-1. 10.1109/ACCESS.2024.3395491.
- [102] Al-Hashmi, Asma & Darem, Abdulbasit & Alshammari, Ahmed & Darem, Laith & Sheatah, Huda & Effghi, Rachid. (2024). Ransomware Early Detection Techniques. *Engineering, Technology & Applied Science Research*. 14. 14497-14503. 10.48084/etasr.6915.
- [103] DOI: 10.48550/arXiv.2412.21084 Reategui, N., Pletka, R., Diamantopoulos, D. On the Generalizability of Machine Learning-based Ransomware Detection in Block Storage. *arXiv*, 2024
- [104] G. O. Ganfure, C. -F. Wu, Y. -H. Chang and W. -K. Shih, "RTrap: Trapping and Containing Ransomware With Machine Learning," in *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 1433-1448, 2023, DOI: 10.1109/TIFS.2023.3240025.

keywords: {Ransomware;Feature extraction;Cryptography;Codes;Organizations;Encryption;Behavioral sciences;Deception-based detection;ransomware detection;affinity propagation;machine learning;adaptive decoy files},

- [105] Microsoft Digital Defense Report 2024. <https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/final/en-us/microsoft-brand/documents/Microsoft%20Digital%20Defense%20Report%202024%20%281%29.pdf>
- [106] Amy Hogan-Burney: Extortion and ransomware drive over half of cyberattacks <https://blogs.microsoft.com/on-the-issues/2025/10/16/mddr-2025/>
- [107] CrowdStrike Global Threat Report <https://www.crowdstrike.com/en-us/resources/reports/global-threat-report-executive-summary-2025/>
- [108] The State of Ransomware 2025, Sophos News, 2024, URL: {<https://assets.sophos.com/X24WTUEQ/at/9brgj5n44hqvgsp5f5bqcps/sophos-state-of-ransomware-2025.pdf>} (Utolsó hozzáférés: 2026. június 22.)
- [109] Dargahi, T., Dehghantanha, A., Bahrami, P. N., Conti, M., Bianchi, G., & Benedetto, L. (2019). A Cyber-Kill-Chain based taxonomy of crypto-ransomware features. *Journal of Computer Virology and Hacking Techniques*, 15, 277–305. DOI: <https://doi.org/10.1007/s11416-019-00338-7>
- [110] European Union Agency for Cybersecurity: ENISA Threat Landscape 2024, URL: {<https://www.enisa.europa.eu/publications/enisa-threat-landscape-2024>} (Utolsó hozzáférés: 2026. július 19.)
- [111] European Commission: Digital Economy and Society Index (DESI) 2014–2022, URL: {<https://ec.europa.eu/newsroom/dae/redirection/document/88764>} (Utolsó hozzáférés: 2026. július 19.)
- [112] European Union Agency for Cybersecurity: NIS360 Report - May 2026, URL: {2026. július 19.)<https://www.enisa.europa.eu/enisa-nis360-2026>} (Utolsó hozzáférés: 2026. július 19.)

Saját publikációk listája

- [S1] Arányi, G., Kocsis, T., Marsi, T., & Szarvák, A. (2020). KIBERTÉRI FENYEGETÉSEK. Budapest: Nemzeti Közsolgálati Egyetem, Közigazgatási Továbbképzési Intézet.
- [S2] Szücs, V., Arányi, G., & Dávid, Á. (2021). Introduction of the ARDS-Anti-Ransomware Defense System Model-Based on the Systematic Review of Worldwide Ransomware Attacks. APPLIED SCIENCES-BASEL, 11(13).
<http://doi.org/10.3390/app11136070> - **SJR Q2, WoS Q2 IF: 2,838**
- [S3] Szücs, V., Arányi, G., & Dávid, Á. (2021). Security awareness of HCI in DigitAll Reality. In 12th IEEE International Conference on Cognitive Infocommunications (CogInfoCom 2021) (pp. 1095–1096).
- [S4] Veronika, Szücs, Gábor, Arányi, Ágnes, Vathy-Fogarassy (2022). Security of IT systems in healthcare institutions. Az egészségügyi informatika COVID előtt és COVID után - A XXXV. Neumann Kollokvium konferencia kiadványa (pp. 45–50).
- [S5] Arányi, G., Szücs, V. (2023). IoT related security issues and black-box testing experiences in a smart lighting system. INTERNATIONAL SCIENTIFIC JOURNAL SECURITY & FUTURE, VII(2), pp. 61–64.
- [S6] Arányi, G.; Vathy-Fogarassy, Á.; Szücs, V.; Evaluation of a New-Concept Secure File Server Solution. Future Internet 2024, 16, 306. <https://doi.org/10.3390/fi16090306>, **SJR Q2, WoS Q2 IF: 3,6**
- [S7] Szücs, Veronika ; Arányi, Gábor ; Dávid, Ákos; Security Awareness of HCI in DigitAll Reality, ACTA POLYTECHNICA HUNGARICA 22 : 6 pp. 9-23. , 15 p. (2025).DOI: 10.12700/APH.22.6.2025.6.2 , **SJR Q2, WoS Q2 IF: 1,8**
- [S8] Arányi, G; Miseta, T; Szücs V., Ransomware Detection Based on Server-side File Operation Logs using Machine Learning. Journal on Information Security, Advances in Information Forensics and Security; Springer, 2025. <https://doi.org/10.1186/s13635-026-00229-7>, **SJR Q2, WoS Q2 IF: 2,1**

Mellékletek

A kérdőíves interjú kérdései

1. A vállalkozás központi telephelye, ahol az adminisztratív ügyintézkést végzik, melyik országban található?	szabad szöveges válasz
2. A cég tevékenysége melyik ágazathoz tartozik leginkább?	szabad szöveges válasz
3. Van-e további ágazati terület, ahol a cég bevételeinek legalább 20%-a keletkezik?	igen/nem
4. Van-e külföldi irodája, leányvállalata, telephelye a vállalkozásnak?	igen/nem
5. Vannak-e külföldi üzleti partnerei?	igen/nem
6. Hány fő alkalmazottja van a vállalkozásnak? (1, 2-10, 11-250, 250 felett)	1, 2-10, 11-250, 250 felett
7. A tulajdonosok közvetlenül részt vesznek-e az operatív tevékenységben?	igen/nem
8. Van-e a vállalkozás IT infrastruktúrájáért felelős döntéshozó a vállalatnál?	igen/nem
9. Milyen végzettséggel rendelkezik az IT döntésekért felelős személy?	felsőfokú, szakirányú
10. Van-e dedikált IT üzemeltetési, fejlesztési csoportja?	igen/nem
11. Van-e saját szervere a vállalkozásnak?	igen/nem
12. Milyen szerverek vannak a vállalkozáson belül? (mail, web, DB, streaming, stb.)	Webszerver, Mail szerver,LDAP kiszolgáló, VPN szerver, Adatbázis szerver, Proxy szerver, File kiszolgáló, Stream szerver
13. Mennyi időksek átlagosan a kiszolgálók?	szabad szöveges válasz
14. Hálózatban dolgoznak-e a vállalat kollégái?	igen/nem
15. Van-e home office-ban dolgozó kollégája?	igen/nem
16. A kollégák mekkora hányada dolgozik otthonról?	szabad szöveges válasz
17. Saját gépüket használják-e home office-ból?	igen/nem
18. Van-e saját céges email fiókjuk a dolgozóknak?	igen/nem
19. Van-e szabályzat a céges email fiók használatára vonatkozóan?	igen/nem
20. Korlátozzák-e a céges email fiók használatát (pl.: mágnjellegű használat tiltása)?	igen/nem
21. Hány gép tartozik az ügyviteli, üzemeltetési feladatok ellátásért felelős IT infrastruktúrához?	szabad szöveges válasz
22. Van-e olyan tevékenysége, amit offline gépről végez?	igen/nem
23. Milyen hálózati végponti eszközök csatlakoznak a céges hálózathoz? (pl.: nyomtatók, szkennerek, NAS, IP kamerák, beléptető kapu)	pl.:NAS, IP kamerák, NVR, Beléptető rendszer, Nyomtató, Szkenner
24. Van-e saját domainje?	igen/nem
25. Van-e rendszergazdája?	igen/nem
26. A rendszergazdájának milyen végzettsége van IT területen?	felsőfokú, szakirányú

27. Van-e a vállalatnál felsőfokú IT végzettségű kollégája, aki ezen a területen dolgozik?	igen/nem
28. Van-e belső IT képzés a belépő dolgozóknak?	igen/nem
29. Az IT területen dolgozó alkalmazottai, rendszergazdái figyelemmel kísérik-e a technológia (szoftveres és hardveres) fejlesztéseket a világban?	igen/nem
30. Hány kollégája használ kritikus, érzékeny adatokat a napi munkavégzés során? (pénzügy, bérszámfejtés stb.)	szabad szöveges válasz
31. Van-e fizikai védelem a telephelyre, irodába történő bejutás szempontjából?	igen/nem
32. Van-e biztonsági szolgálata, portaszolgálata?	igen/nem
33. Van-e olyan IT üzemeltetési szolgáltatás, amit szakirányú profillal rendelkező vállalkozástól, szolgáltatótól vesz igénybe?	igen/nem
34. A saját szerverein milyen operációs rendszer fut?	pl.Windows és Linux
35. A munkaállomásokon milyen operációs rendszer fut?	pl. Windows10
36. A hálózati infrastruktúra alkalmas-e az internetes és intranetes forgalom monitorozására?	igen/nem
37. Milyen típusú internet kapcsolattal rendelkezik a vállalkozása?	szabad szöveges válasz
38. Van-e hozzáférés szabályozás a vállalati infrastruktúrában?	igen/nem
39. Használnak-e proxy szervert?	igen/nem
40. Van-e felhő alapú szolgáltatás, amit igénybe vesz a vállalkozása?	igen/nem
41. Használnak-e online számlázó rendszert?	igen/nem
42. Használnak-e internet bankot?	igen/nem
43. Engedélyezett-e a céges hálózaton belül az eszközök vezeték nélküli csatlakozása?	igen/nem
44. Van-e szeparált vezeték nélküli vendéghálózat a rendszerében?	igen/nem
45. Van-e hardveres tűzfal a hálózaton?	igen/nem
46. Van-e vírusvédelmi és biztonsági megoldás telepítve a szerverein?	igen/nem
47. Van-e szoftveres munkaállomás védelem a számítógépeken?	igen/nem
48. Ingyenes vagy előfizetéses vírusvédelmi, illetve végpontvédelmi megoldást használnak?	ingyenes megoldás/előfizetéses
49. Van-e korlátozás a munkaállomásokon ezeknek a védelmi szoftvereknek a beállítás módosításaival kapcsolatban?	igen/nem
50. MS Windows rendszerű hálózaton használnak-e központi hitelesítést, címtárat? (pl. LDAP)?	igen/nem
51. Van-e tartományi kiszolgáló telepítve a hálózaton?	igen/nem
52. Van-e VPN kiszolgáló telepítve a hálózaton?	igen/nem
53. Van-e felelőse a biztonsági frissítések telepítésének?	igen/nem
54. Végeznek-e rendszeres vírus ellenőrzést a munkaállomásokon?	igen/nem
55. Van-e biztonsági mentés a vállalkozás működése szempontjából kritikus adatokról?	igen/nem

56. A biztonsági mentést hány helyen tárolja?	pl.: Kiszolgálón több példányban és privát online tárhelyen, offline meghajtón
57. Van-e olyan mentése, amit a rendszerről leválasztva, fizikailag elkülönítve tart?	igen/nem
58. Van-e biztonsági mentésekért felelős kollégája a vállalatnak?	igen/nem
59. Végeznek-e ellenőrzést a biztonsági mentésekből történő sikeres helyreállítás lehetőségével kapcsolatban?	igen/nem
60. Van-e domain korlátozás az interneten elérhető tartalmakra vonatkozóan?	igen/nem
61. Volt-e már vírustámadás a rendszerében?	igen/nem
62. Áldozata volt-e korábban zsarolóvírus-támadásnak?	igen/nem
63. Mekkora összeget követeltek vállalkozásától váltságdíjként?	szabad szöveges válasz
64. Milyen módon kérték a váltsádját? (pl.: Bitcoin, más cryptovaluta)	szabad szöveges válasz
65. Jelentette-e a támadást a megfelelő szervezetnél?	igen/nem
66. Mi volt az oka, ha nem jelentette?	szabad szöveges válasz
67. Fizetett-e váltságdíjat a támadóknak?	igen/nem
68. Váltsádij fizetés nélkül sikerült-e dekódolni a titkosított állományokat?	igen/nem
69. Igénybe vette-e ehhez szakember segítségét?	igen/nem
70. Veszett-e el adata, ami a későbbiekben gondot okozhat hivatalos adatszolgáltatás teljesítésnél, vagy más szempontból?	igen/nem
71. Mennyi kiesést okozott időben a támadás?	szabad szöveges válasz
72. Mennyi ideig tartott a teljes helyreállítás?	szabad szöveges válasz
73. Mennyibe került a helyreállítás? (ezer Ft-ban)	szabad szöveges válasz
74. Volt-e második vagy esetleg további alkalom, amikor zsarolóvírus-támadás áldozata lett?	igen/nem
75. Ez a másodszori támadás ugyanazt a rendszerét érintette?	igen/nem
76. Végzett-e hálózati biztonság célú fejlesztést az IT infrastruktúráján a támadást követően?	igen/nem
77. Változtatott-e az informatikai eszközök használatára vonatkozó szabályzatán az eset után?	igen/nem
78. Szerzett-e be komplexebb védelmet biztosító megoldást az eset után?	igen/nem
79. Ha nem, mi volt az oka?	szabad szöveges válasz
80. Biztonságban érzi-e adatait, online tevékenységeit, privát adatait az internet használatakor?	igen/nem
81. Tudja-e, milyen jelei lehetnek egy lehetséges támadásnak?	igen/nem
82. Tudja-e követni, és megérteni-e a jelenlegi veszélyforrások mibenlétét, működési mechanizmusát?	igen/nem
83. Hajlandó lenne-e áldozni egy olyan költséghatékony, komplex rendszerre, ami átfogó, de mégis könnyen integrálható védelmet kínál egyszerre?	igen/nem

84. Tudja-e mi az a sérülékenység vizsgálat?	igen/nem
85. Végeztek-e már a vállalkozásánál információ biztonsági auditot?	igen/nem
86. Ha nem, mi az oka?	szabad szöveges válasz

A fájlkiszolgálók teljesítményadatai és a mérésekhez tartozó statisztikai értékek

1 Gbit/s

W7-ARDS olvasás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájl méret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	9,64	106,22	9,35	109,52	26,26	38,99	117,51	8,71
2. teszt	11,50	89,04	9,24	110,82	20,06	51,05	53,12	19,28
3. teszt	10,20	100,39	9,31	109,99	20,78	49,28	53,27	19,22
4. teszt	10,86	94,29	9,41	108,82	19,86	51,56	52,85	19,38
5. teszt	10,07	101,69	9,34	109,64	21,55	47,52	52,18	19,62
6. teszt	9,74	105,13	9,30	110,11	20,08	51,00	52,90	19,36
7. teszt	9,64	106,22	9,35	109,52	20,07	51,02	52,44	19,53
8. teszt	10,02	102,20	9,28	110,34	19,95	51,33	53,65	19,09
9. teszt	9,79	104,60	9,28	110,34	20,42	50,15	53,96	18,98
10. teszt	9,70	105,57	9,27	110,46	19,96	51,30	52,94	19,34
Minimum	9,64	89,04	9,24	108,82	19,86	47,52	52,18	18,98
Maximum	11,50	106,22	9,41	110,82	21,55	51,56	53,96	19,62
Átlag	10,17	101,01	9,31	110,01	20,30	50,47	53,03	19,31
Szórás	0,623447	5,792334	0,000001	0,601719	0,000006	1,312260	0,000006	0,200933
W7-ARDS írás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájl méret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	9,60	106,67	9,72	105,35	29,14	35,14	111,75	9,16
2. teszt	9,67	105,89	9,34	109,64	21,77	47,04	86,63	11,82
3. teszt	9,32	109,87	9,36	109,40	28,98	35,33	87,79	11,66

4. teszt	9,28	110,34	9,35	109,52	25,21	40,62	85,10	12,03
5. teszt	9,22	111,06	9,66	106,00	24,29	42,16	87,16	11,75
6. teszt	9,23	110,94	9,36	109,40	25,56	40,06	88,39	11,59
7. teszt	9,22	111,06	9,78	104,70	24,74	41,39	86,20	11,88
8. teszt	9,28	110,34	9,74	105,13	23,18	44,18	89,89	11,39
9. teszt	9,31	109,99	9,51	107,68	22,26	46,00	88,61	11,56
10. teszt	9,26	110,58	9,39	109,05	21,41	47,83	89,97	11,38
Minimum	9,22	105,89	9,34	104,70	21,41	35,33	85,10	11,38
Maximum	9,67	111,06	9,78	109,64	28,98	47,83	89,97	12,03
Átlag	9,31	110,01	9,50	107,84	24,16	42,73	87,75	11,67
Szórás	0,000002	1,604545	0,000002	2,029503	0,000027	3,964264	0,000019	0,219385

W7/W201 9 olvasás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájl méret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	11,31	90,54	11,37	90,06	29,36	34,88	98,43	10,40
2. teszt	9,44	108,47	9,44	108,47	20,82	49,18	53,53	19,13
3. teszt	9,44	108,47	9,50	107,79	21,66	47,28	59,83	17,12
4. teszt	9,67	105,89	9,33	109,75	20,78	49,28	85,39	11,99
5. teszt	9,35	109,52	9,52	107,56	20,84	49,14	57,75	17,73
6. teszt	9,36	109,40	9,44	108,47	20,85	49,11	54,18	18,90
7. teszt	9,42	108,70	9,50	107,79	20,86	49,09	54,68	18,73
8. teszt	9,49	107,90	9,36	109,40	20,85	49,11	53,88	19,01
9. teszt	9,76	104,92	9,41	108,82	20,83	49,16	53,35	19,19
10. teszt	9,43	108,59	9,38	109,17	20,80	49,23	53,15	19,27
Minimum	9,35	104,92	9,33	107,56	20,78	47,28	53,15	11,99
Maximum	9,76	109,52	9,52	109,75	21,66	49,28	85,39	19,27

Átlag	9,48	107,99	9,43	108,58	20,92	48,95	58,42	17,90
Szórás	0,139204	1,561473	0,066978	0,771576	0,278273	0,631861	10,366921	2,332014
W7/W2019 írás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájl méret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	9,22	111,06	9,47	108,13	28,66	35,73	131,74	7,77
2. teszt	9,23	110,94	9,36	109,40	21,48	47,67	103,71	9,87
3. teszt	9,25	110,70	9,38	109,17	21,34	47,99	106,54	9,61
4. teszt	9,20	111,30	9,36	109,40	21,48	47,67	102,52	9,99
5. teszt	9,20	111,30	9,36	109,40	21,65	47,30	103,32	9,91
6. teszt	9,21	111,18	9,36	109,40	22,90	44,72	102,94	9,95
7. teszt	9,22	111,06	9,34	109,64	21,77	47,04	102,71	9,97
8. teszt	9,21	111,18	9,36	109,40	23,92	42,81	102,41	10,00
9. teszt	9,20	111,30	9,37	109,28	21,82	46,93	102,85	9,96
10. teszt	9,20	111,30	9,36	109,40	22,26	46,00	103,03	9,94
Minimum	9,20	110,70	9,34	109,17	21,34	42,81	102,41	9,61
Maximum	9,25	111,30	9,38	109,64	23,92	47,99	106,54	10,00
Átlag	9,21	111,14	9,36	109,39	22,07	46,46	103,34	9,91
Szórás	0,017321	0,208523	0,010541	0,123205	0,84450 3	1,699207	1,26528 7	0,118577

W10- ARDS olva- sás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájl méret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	9,96	102,81	9,79	104,60	23,21	44,12	94,81	10,80
2. teszt	9,68	105,79	9,20	111,30	17,14	59,74	57,22	17,90

3. teszt	9,67	105,89	9,20	111,30	17,12	59,81	48,55	21,09
4. teszt	9,20	111,30	9,22	111,06	17,09	59,92	48,55	21,09
5. teszt	9,39	109,05	9,20	111,30	17,21	59,50	47,81	21,42
6. teszt	9,36	109,40	9,21	111,18	17,04	60,09	48,59	21,07
7. teszt	9,40	108,94	9,21	111,18	17,04	60,09	48,03	21,32
8. teszt	9,46	108,25	9,20	111,30	17,20	59,53	48,00	21,33
9. teszt	9,43	108,59	9,22	111,06	17,08	59,95	48,11	21,28
10. teszt	9,42	108,70	9,21	111,18	17,16	59,67	47,99	21,34
Minimum	9,20	105,79	9,20	111,06	17,04	59,50	47,81	17,90
Maximum	9,68	111,30	9,22	111,30	17,21	60,09	57,22	21,42
Átlag	9,45	108,43	9,21	111,21	17,12	59,81	49,21	20,87
Szórás	0,000002	1,710570	0,000000	0,100614	0,000000 1	0,220198	0,000003 5	1,123270
W10- ARDS írás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájl méret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	23,30	43,95	65,93	15,53	2455,79	0,42	31816,66	0,03218
2. teszt	37,84	27,06	122,54	8,36	9152,87	0,11	31849,55	0,03215
3. teszt	37,58	27,25	122,56	8,36	9337,82	0,11	32003,27	0,03200
4. teszt	37,58	27,25	122,79	8,34	9690,30	0,11	31974,70	0,03203
5. teszt	37,49	27,31	122,53	8,36	9877,30	0,10	32009,64	0,03199
6. teszt	37,38	27,39	122,53	8,36	9736,21	0,11	32029,56	0,03197
7. teszt	37,59	27,24	122,86	8,33	9779,69	0,10	31901,37	0,03210
8. teszt	37,57	27,26	122,44	8,36	8978,06	0,11	31998,13	0,03200
9. teszt	37,40	27,38	122,65	8,35	9596,80	0,11	31994,22	0,03201
10. teszt	37,65	27,20	122,59	8,35	9649,42	0,11	32033,80	0,03197
Minimum	37,38	27,06	122,44	8,33	8978,06	0,10	31849,55	0,03

Maximum	37,84	27,39	122,86	8,36	9877,30	0,11	32033,80	0,03
Átlag	37,56	27,26	122,61	8,35	9533,16	0,11	31977,14	0,03
Szórás	0,000002	0,099527	0,000002	0,009194	306,920000	0,003545	61,700109	0,000062

W10- W2019 ol- vasás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájl méret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	9,16	111,79	11,23	91,18	24,47	41,85	90,05	11,37
2. teszt	9,42	108,70	9,73	105,24	17,12	59,81	52,33	19,57
3. teszt	9,42	108,70	9,19	111,43	17,46	58,65	53,14	19,27
4. teszt	9,47	108,13	9,19	111,43	17,37	58,95	53,68	19,08
5. teszt	9,45	108,36	9,19	111,43	17,77	57,63	51,95	19,71
6. teszt	9,95	102,91	9,19	111,43	17,22	59,47	51,88	19,74
7. teszt	9,41	108,82	9,18	111,55	17,21	59,50	52,60	19,47
8. teszt	9,60	106,67	9,19	111,43	17,28	59,26	52,65	19,45
9. teszt	9,40	108,94	9,19	111,43	17,36	58,99	52,79	19,40
10. teszt	9,32	109,87	9,20	111,30	17,23	59,43	52,02	19,68
Minimum	9,32	102,91	9,18	105,24	17,12	57,63	51,88	19,08
Maximum	9,95	109,87	9,73	111,55	17,77	59,81	53,68	19,74
Átlag	9,49	107,90	9,25	110,74	17,34	59,08	52,56	19,48
Szórás	0,186682	2,053734	0,180069	2,062216	0,192296	0,646728	0,595021	0,219293
W10- W2019 Írás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájl méret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	9,29	110,23	9,56	107,11	23,14	44,25	68,87	14,87
2. teszt	9,50	107,79	9,27	110,46	14,42	71,01	35,81	28,60

3. teszt	9,18	111,55	9,27	110,46	13,83	74,04	36,62	27,96
4. teszt	9,18	111,55	9,31	109,99	13,85	73,94	37,02	27,66
5. teszt	9,17	111,67	9,26	110,58	13,92	73,56	36,27	28,23
6. teszt	9,19	111,43	9,28	110,34	13,90	73,67	36,86	27,78
7. teszt	9,18	111,55	9,27	110,46	13,99	73,20	34,74	29,48
8. teszt	9,18	111,55	9,27	110,46	14,03	72,99	36,92	27,74
9. teszt	9,17	111,67	9,26	110,58	13,76	74,42	36,20	28,29
10. teszt	9,18	111,55	9,26	110,58	13,78	74,31	37,03	27,65
Minimum	9,17	107,79	9,26	109,99	13,76	71,01	34,74	27,65
Maximum	9,50	111,67	9,31	110,58	14,42	74,42	37,03	29,48
Átlag	9,21	111,14	9,27	110,44	13,94	73,46	36,39	28,15
Szórás	0,107251	1,259644	0,015635	0,185719	0,200361	1,033551	0,747531	0,592836

U2204- ARDS olva- sás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájl méret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	9,45	108,37	9,25	110,65	11,15	91,83	17,93	57,13
2. teszt	9,32	109,93	9,39	109,01	11,21	91,34	17,28	59,26
3. teszt	9,32	109,91	9,39	109,05	11,25	91,00	17,17	59,64
4. teszt	9,32	109,91	9,39	109,02	11,35	90,21	17,32	59,13
5. teszt	9,32	109,82	9,35	109,47	11,20	91,43	17,27	59,31
6. teszt	9,37	109,27	9,36	109,39	11,29	90,72	17,00	60,25
7. teszt	9,32	109,92	9,35	109,51	11,21	91,31	17,09	59,93
8. teszt	9,31	109,95	9,36	109,39	11,23	91,18	17,24	59,41
9. teszt	9,31	109,97	9,42	108,70	11,26	90,93	17,45	58,67
10. teszt	9,33	109,73	9,45	108,41	11,33	90,42	17,07	59,98
Minimum	9,31	109,27	9,35	108,41	11,20	90,21	17,00	58,67

Maximum	9,37	109,97	9,45	109,51	11,35	91,43	17,45	60,25
Átlag	9,32	109,82	9,39	109,10	11,26	90,95	17,21	59,51
Szórás	0,018644	0,218771	0,032369	0,375526	0,052791	0,425349	0,141830	0,490072
U2204- ARDS írás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájl méret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	10,11	101,25	11,88	86,17	13,46	76,07	22,63	45,25
2. teszt	9,53	107,45	9,80	104,52	11,12	92,08	22,83	44,85
3. teszt	9,55	107,28	9,76	104,92	11,22	91,26	22,10	46,34
4. teszt	9,54	107,37	9,78	104,75	11,11	92,19	22,68	45,16
5. teszt	9,53	107,46	9,74	105,14	11,27	90,88	22,92	44,68
6. teszt	9,54	107,39	9,77	104,81	11,06	92,62	22,05	46,44
7. teszt	9,53	107,51	9,76	104,94	10,97	93,35	23,20	44,14
8. teszt	9,53	107,41	9,77	104,79	11,02	92,91	22,71	45,09
9. teszt	9,53	107,51	9,76	104,89	11,01	93,05	22,62	45,27
10. teszt	9,54	107,34	9,74	105,11	11,09	92,34	22,91	44,71
Minimum	9,53	107,28	9,74	104,52	10,97	90,88	22,05	44,14
Maximum	9,55	107,51	9,80	105,14	11,27	93,35	23,20	46,44
Átlag	9,53	107,41	9,76	104,87	11,10	92,29	22,67	45,19
Szórás	0,006764	0,076195	0,017660	0,189593	0,098264	0,813730	0,376844	0,758382

U2204- W2019 olvasás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájl méret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	9,29	110,24	9,20	111,36	11,47	89,26	15,25	67,15
2. teszt	9,30	110,06	9,32	109,89	11,47	89,25	14,39	71,14

3. teszt	9,28	110,36	9,33	109,79	11,58	88,45	14,66	69,85
4. teszt	9,28	110,39	9,31	109,98	11,60	88,31	14,81	69,15
5. teszt	9,27	110,49	9,32	109,89	11,53	88,84	14,44	70,93
6. teszt	9,28	110,39	9,31	109,95	11,61	88,21	14,66	69,85
7. teszt	9,27	110,45	9,31	110,04	11,52	88,92	14,80	69,21
8. teszt	9,27	110,46	9,31	109,98	11,48	89,17	14,81	69,13
9. teszt	9,28	110,40	9,33	109,80	11,54	88,74	14,71	69,61
10. teszt	9,28	110,39	9,33	109,79	11,61	88,22	14,94	68,54
Minimum	9,27	110,06	9,31	109,79	11,47	88,21	14,39	68,54
Maximum	9,30	110,49	9,33	110,04	11,61	89,25	14,94	71,14
Átlag	9,28	110,38	9,32	109,90	11,55	88,68	14,69	69,71
Szórás	0,010639	0,126316	0,007828	0,092328	0,052056	0,399962	0,179378	0,856476
U2204-W2019 írás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájl méret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	10,95	93,53	12,21	83,86	13,68	74,85	21,75	47,09
2. teszt	9,48	108,05	9,73	105,25	11,08	92,39	17,11	59,84
3. teszt	9,46	108,21	9,71	105,50	11,76	87,06	16,57	61,79
4. teszt	9,44	108,50	9,71	105,46	11,20	91,47	16,39	62,48
5. teszt	9,44	108,43	9,67	105,86	11,12	92,09	16,41	62,40
6. teszt	9,44	108,43	9,67	105,94	11,04	92,80	16,40	62,44
7. teszt	9,44	108,50	9,67	105,94	11,00	93,12	17,01	60,20
8. teszt	9,46	108,23	9,68	105,81	10,95	93,49	16,47	62,18
9. teszt	9,43	108,56	9,67	105,94	10,92	93,74	16,40	62,46
10. teszt	9,44	108,52	9,67	105,85	11,10	92,29	16,41	62,40
Minimum	9,43	108,05	9,67	105,25	10,92	87,06	16,39	59,84
Maximum	9,48	108,56	9,73	105,94	11,76	93,74	17,11	62,48

Átlag	9,45	108,38	9,69	105,73	11,13	92,05	16,57	61,80
Szórás	0,015147	0,173541	0,023447	0,255461	0,251956	2,001528	0,283561	1,036727

10 Gbit/s

W10- ARDS olva- sás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájl méret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	1,14	898,25	5,70	179,65	16,51	62,02	65,45	15,65
2. teszt	1,11	922,52	1,29	793,80	7,08	144,63	25,82	39,66
3. teszt	1,06	966,04	1,08	948,15	7,93	129,13	25,44	40,25
4. teszt	1,08	948,15	1,28	800,00	6,87	149,05	25,41	40,30
5. teszt	1,03	994,17	1,35	758,52	6,97	146,92	25,48	40,19
6. teszt	1,06	966,04	1,36	752,94	7,17	142,82	25,41	40,30
7. teszt	1,04	984,62	1,29	793,80	7,39	138,57	26,36	38,85
8. teszt	1,07	957,01	1,92	533,33	6,76	151,48	25,67	39,89
9. teszt	1,08	948,15	1,26	812,70	6,85	149,49	25,64	39,94
10. teszt	1,06	966,04	1,05	975,24	6,85	149,49	25,38	40,35
Minimum	1,03	922,52	1,05	533,33	6,76	129,13	25,38	38,85
Maximum	1,11	994,17	1,92	975,24	7,93	151,48	26,36	40,35
Átlag	1,07	961,41	1,32	796,50	7,10	144,62	25,62	39,97
Szórás	0,000000	21,093903	0,000003	126,407259	0,000004	7,064363	0,000004	0,481031
W10- ARDS írás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájl méret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB

1. teszt	7,00	146,29	9,77	104,81	9,92	103,23	12,53	81,72386
2. teszt	4,53	226,05	4,00	256,00	5,39	189,98	15,66	65,38953
3. teszt	3,76	272,34	3,94	259,90	5,94	172,39	14,50	70,62069
4. teszt	3,64	281,32	4,14	247,34	5,78	177,16	12,40	82,58065
5. teszt	3,95	259,24	4,05	252,84	5,83	175,64	13,06	78,40735
6. teszt	3,60	284,44	3,81	268,77	5,61	182,53	14,40	71,11111
7. teszt	3,61	283,66	4,22	242,65	5,64	181,56	15,49	66,10717
8. teszt	3,79	270,18	3,98	257,29	5,70	179,65	13,86	73,88167
9. teszt	3,86	265,28	4,03	254,09	5,90	173,56	11,61	88,19983
10. teszt	3,79	270,18	4,31	237,59	5,37	190,69	12,12	84,48845
Minimum	3,60	226,05	3,81	237,59	5,37	172,39	11,61	65,39
Maximum	4,53	284,44	4,31	268,77	5,94	190,69	15,66	88,20
Átlag	3,84	268,08	4,05	252,94	5,68	180,35	13,68	75,64
Szórás	0,000003	17,907004	0,000002	9,363060	0,000000	6,596897	0,000017	8,190071

W10-W2019 olvasás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájl méret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	1,25	819,20	10,61	96,51	22,46	45,59	46,30	22,12
2. teszt	1,26	812,70	1,81	565,75	5,39	189,98	23,90	42,85
3. teszt	1,73	591,91	1,25	819,20	5,98	171,24	23,96	42,74
4. teszt	1,34	764,18	1,98	517,17	5,50	186,18	23,76	43,10
5. teszt	1,32	775,76	1,32	775,76	5,41	189,28	23,90	42,85
6. teszt	1,21	846,28	1,39	736,69	5,85	175,04	23,60	43,39
7. teszt	1,42	721,13	1,44	711,11	5,76	177,78	24,28	42,17
8. teszt	1,25	819,20	1,72	595,35	5,41	189,28	23,73	43,15
9. teszt	1,22	839,34	1,52	673,68	5,39	189,98	23,81	43,01
10. teszt	1,53	669,28	1,27	806,30	5,47	187,20	24,03	42,61

Minimum	1,21	591,91	1,25	517,17	5,39	171,24	23,60	42,17
Maximum	1,73	846,28	1,98	819,20	5,98	189,98	24,28	43,39
Átlag	1,36	759,98	1,52	689,00	5,57	184,00	23,89	42,87
Szórás	0,000002	85,354837	0,000003	108,813474	0,000003	7,281342	0,000002	0,351496
W10-W2019 írás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájlméret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	2,44	419,67	2,54	403,15	8,09	126,58	30,34	33,75
2. teszt	1,17	875,21	1,21	846,28	4,49	228,06	16,38	62,52
3. teszt	1,20	853,33	1,20	853,33	4,10	249,76	18,27	56,05
4. teszt	1,24	825,81	1,45	706,21	4,04	253,47	16,37	62,55
5. teszt	1,37	747,45	1,44	711,11	4,24	241,51	19,09	53,64
6. teszt	1,19	860,50	1,30	787,69	4,07	251,60	17,47	58,61
7. teszt	1,36	752,94	1,17	875,21	4,93	207,71	15,07	67,95
8. teszt	1,28	800,00	1,17	875,21	4,40	232,73	17,86	57,33
9. teszt	1,29	793,80	1,17	875,21	4,14	247,34	17,93	57,11
10. teszt	1,31	781,68	1,25	819,20	5,24	195,42	17,53	58,41
Minimum	1,17	747,45	1,17	706,21	4,04	195,42	15,07	53,64
Maximum	1,37	875,21	1,45	875,21	5,24	253,47	19,09	67,95
Átlag	1,27	810,08	1,26	816,61	4,41	234,18	17,33	59,35
Szórás	0,000001	46,454836	0,000001	67,725498	0,000005	20,577703	0,000014	4,304398

J2204-ARDS olvasás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájlméret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	1,03	999,02	5,56	184,21	1,59	642,81	5,65	181,37

2. teszt	1,03	998,05	1,01	1010,86	1,56	655,99	5,87	174,36
3. teszt	0,99	1033,30	1,01	1010,86	2,04	500,98	4,44	230,63
4. teszt	1,00	1020,94	1,04	986,51	1,92	532,22	5,52	185,57
5. teszt	1,00	1027,08	1,03	990,33	1,69	605,20	4,75	215,58
6. teszt	0,98	1045,97	1,03	993,21	1,55	659,79	5,96	171,78
7. teszt	1,04	981,78	1,05	973,38	1,54	666,67	4,68	218,62
8. teszt	1,01	1013,86	1,03	999,02	1,54	663,21	3,79	269,97
9. teszt	1,00	1022,98	1,04	983,67	1,54	663,64	6,21	164,95
10. teszt	0,98	1040,65	0,99	1035,39	1,60	641,20	4,16	245,98
Minimum	0,98	981,78	0,99	973,38	1,54	500,98	3,79	164,95
Maximum	1,04	1045,97	1,05	1035,39	2,04	666,67	6,21	269,97
Átlag	1,00	1020,51	1,03	998,14	1,67	620,99	5,04	208,60
Szórás	0,02029	20,35625	0,01885	18,556	0,18902	62,60883	0,86817	36,67672
J2204-ARDS írás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájlméret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	4,90	208,85	3,59	285,40	4,38	234,00	8,21	124,73
2. teszt	2,55	401,10	3,75	273,43	3,55	288,69	5,60	182,86
3. teszt	4,23	241,97	1,96	522,18	4,76	215,13	5,78	177,16
4. teszt	3,73	274,83	3,27	313,15	4,61	222,22	7,35	139,38
5. teszt	3,92	261,49	3,92	261,22	4,47	229,13	7,73	132,47
6. teszt	3,64	281,01	4,00	255,74	4,98	205,66	6,41	159,65
7. teszt	3,78	271,19	4,12	248,54	4,31	237,59	9,00	113,82
8. teszt	3,82	268,13	2,30	445,02	4,48	228,57	7,71	132,81
9. teszt	2,83	361,71	2,73	375,23	4,07	251,91	6,55	156,36
10. teszt	2,75	371,82	3,95	259,44	4,61	221,93	6,28	163,03
Minimum	2,55	241,97	1,96	248,54	3,55	205,66	5,60	113,82

Maximum	4,23	401,10	4,12	522,18	4,98	288,69	9,00	182,86
Átlag	3,47	303,69	3,33	328,22	4,43	233,43	6,93	150,84
Szórás	0,59714	57,83094	0,81218	98,4403	0,41938	24,54255	1,09651	22,74657

U2204- W2019 olva- sás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájlméret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB
1. teszt	1,67	613,17	1,85	554,71	3,60	284,84	6,56	156,05
2. teszt	1,56	656,83	1,36	750,73	3,28	312,29	6,62	154,80
3. teszt	1,62	633,27	1,83	560,48	3,63	282,40	6,58	155,62
4. teszt	1,67	612,44	1,29	795,65	3,49	293,24	6,43	159,38
5. teszt	1,51	678,60	1,94	528,93	3,36	304,49	6,72	152,31
6. teszt	1,94	529,20	1,84	558,04	3,45	296,64	6,58	155,69
7. teszt	1,69	606,64	1,51	676,80	3,41	300,73	6,78	151,08
8. teszt	1,40	731,43	1,42	721,63	3,28	312,58	6,63	154,52
9. teszt	1,68	609,52	1,64	624,01	3,30	310,02	6,70	152,74
10. teszt	1,67	614,28	2,08	492,78	3,11	328,94	6,54	156,60
Minimum	1,40	529,20	1,29	492,78	3,11	282,40	6,43	151,08
Maximum	1,94	731,43	2,08	795,65	3,63	328,94	6,78	159,38
Átlag	1,64	630,24	1,66	634,34	3,37	304,59	6,62	154,75
Szórás	0,14765	55,96977	0,2775	106,961	0,14833	13,43828	0,10649	2,499805
U2204- W2019 írás	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)	Átv. idő (sec)	Átv. seb. (MiB/s)
Darabszám x fájlméret	1 x 1 GiB	1 x 1 GiB	4 x 256 MiB	4 x 256 MiB	256 x 4 MiB	256 x 4 MiB	2097 x 500 kiB	2097 x 500 kiB

1. teszt	1,21	846,98	1,33	771,08	3,80	269,83	4,07	251,84
2. teszt	1,20	850,50	1,22	842,80	1,82	563,88	3,03	338,51
3. teszt	1,21	849,09	1,21	843,49	1,86	552,02	3,10	330,11
4. teszt	1,21	849,79	1,22	842,80	1,84	555,62	3,16	324,36
5. teszt	1,21	848,38	1,21	846,28	1,83	559,87	3,26	314,21
6. teszt	1,21	849,09	1,21	843,49	1,81	566,06	3,24	316,24
7. teszt	1,20	850,50	1,21	845,58	1,85	553,81	3,10	329,90
8. teszt	1,21	848,38	1,21	843,49	1,92	534,73	3,36	305,03
9. teszt	1,21	849,79	1,21	844,19	1,81	565,12	3,28	312,00
10. teszt	1,28	798,75	1,21	846,28	1,79	571,11	3,07	333,44
Minimum	1,20	798,75	1,21	842,80	1,79	534,73	3,03	305,03
Maximum	1,28	850,50	1,22	846,28	1,92	571,11	3,36	338,51
Átlag	1,21	843,81	1,21	844,27	1,84	558,02	3,18	322,65
Szórás	0,025524	16,914639	0,002028	1,412499	0,036094	10,759825	0,111559	11,264504