Ilmiy tadqiqotlar va ularning yechimlari jurnali
Aoka guvohnoma № 323759 ISSN: 3060-4818

ZARARLI DASTURLARNING TARMOQ PROTOKOLLARIDAN FOYDALANISH USULLARI VA COMMAND AND CONTROL (C2) ALOQALARINI ANIQLASHNING ZAMONAVIY YONDASHUVLARI

Quyliyev Behro‘z Sherzod o‘g‘li
Toshkent axborot texnologiyalari universiteti,
Kiberxavfsizlik fakulteti
4-bosqich talabasi, Toshkent, O‘zbekiston
E-mail: behruzquliyev06@gmail.com, +998996655554

Olimov Iskandar Salimboyevich
lmiy rahbar,Toshkent axborot texnologiyalari universiteti,
Kriptologiya kafedrasi,
Katta o‘qituvchi, Toshkent, O‘zbekiston
E-mail:olimoviskandar89@gmail.com
https://doi.org/10.5281/zenodo.24189120

Annotatsiya. Mazkur maqolada zararli dasturlarning Command and Control (C2) aloqasini tashkil etishda foydalanadigan tarmoq protokollari (HTTP/HTTPS, DNS, ICMP, bulut xizmatlari) va bu aloqalarni aniqlashning zamonaviy yondashuvlari qiyosiy tahlil qilinadi. Imzoga asoslangan, statistik, mashinali o‘rganishga asoslangan va tahdidlar haqidagi razvedka (threat intelligence) usullari taqqoslanib, ularning shifrlangan trafik, DGA va DNS tunneling holatlaridagi cheklovlari ko‘rsatiladi. Tahlil asosida beaconing va DNS tunneling belgilarini ajratishga mo‘ljallangan ko‘p qatlamli aniqlash yondashuvi taklif etiladi. Natijalar korporativ tarmoqlarda C2 trafigini erta aniqlash tizimlarini loyihalashda foydalanilishi mumkin.
Kalit so‘zlar: zararli dastur, C2, beaconing, DNS tunneling, DGA, tarmoq trafigini tahlil qilish, mashinali o‘rganish, IOC.

СПОСОБЫ ИСПОЛЬЗОВАНИЯ СЕТЕВЫХ ПРОТОКОЛОВ ВРЕДОНОСНЫМИ ПРОГРАММАМИ И СОВРЕМЕННЫЕ ПОДХОДЫ К ОБНАРУЖЕНИЮ COMMAND AND CONTROL (C2) КОММУНИКАЦИЙ

Аннотация. В статье проводится сравнительный анализ сетевых протоколов (HTTP/HTTPS, DNS, ICMP, облачные сервисы), используемых вредоносными программами для организации каналов Command and Control (C2), и современных подходов к их обнаружению. Сопоставляются сигнатурные, статистические методы, методы машинного обучения и threat intelligence, выявляются их ограничения при шифрованном трафике, DGA и DNS-туннелировании. На основе анализа предлагается многоуровневый подход к обнаружению beaconing и DNS-туннелирования. Результаты могут быть использованы при проектировании систем раннего обнаружения C2-трафика в корпоративных сетях.
Ключевые слова: вредоносная программа, C2, beaconing, DNS-туннелирование, DGA, анализ сетевого трафика, машинное обучение, IOC.

METHODS OF NETWORK PROTOCOL ABUSE BY MALWARE AND MODERN APPROACHES TO DETECTING COMMAND AND CONTROL (C2) COMMUNICATIONS

Abstract. The article presents a comparative analysis of network protocols (HTTP/HTTPS, DNS, ICMP, cloud services) that malware uses to establish Command and Control (C2) channels and of modern approaches to detecting them. Signature-based, statistical, machine-learning and threat-intelligence methods are compared, and their limitations in the case of encrypted traffic, DGA and DNS tunneling are identified. Based on the analysis, a multi-layer approach for detecting beaconing and DNS tunneling is proposed. The results can be used when designing early detection systems for C2 traffic in corporate networks.

Keywords: malware, C2, beaconing, DNS tunneling, DGA, network traffic analysis, machine learning, IOC.

KIRISH

Zamonaviy zararli dasturlarning aksariyati zararlangan tizim bilan hujumchi o‘rtasida boshqaruv kanalini talab qiladi: buyruqlarni qabul qilish, ma’lumotlarni chiqarib yuborish, qo‘shimcha modullarni yuklash va hujumni muvofiqlashtirish aynan shu kanal orqali amalga oshadi. Bunday kanal Command and Control (C2) deb ataladi va MITRE ATT&CK bilimlar bazasida alohida taktika sifatida tasniflangan [8]. Botnet, ransomware va APT hujumlarining turli-tumanligiga qaramay, C2 aloqasi ko‘plab hujumlarda muhim boshqaruv komponentidir; biroq uning uzilishi barcha hujum turlarida nazoratning darhol yo‘qolishini anglatmaydi.

Hujumchilar C2 trafikini tarmoq chegarasidagi himoya vositalaridan yashirish uchun ruxsat etilgan protokollardan, birinchi navbatda HTTP(S) va DNS dan foydalanadi [14, 15]. TLS 1.3 ning keng tarqalishi shifrlangan sessiyalarda paket yuklamasini bevosita tekshirish imkoniyatlarini cheklaydi [16]. Natijada e’tibor trafik mazmunidan uning xulq-atvor xususiyatlariga — aloqalar davriyligi, hajmi, so‘rovlar tuzilmasi va ketma-ketligiga — ko‘chmoqda.

Maqolaning maqsadi — zararli dasturlar C2 aloqasi uchun ishlatadigan tarmoq protokollari va bu aloqalarni aniqlashning zamonaviy yondashuvlarini yagona tizimda tahlil qilish hamda aniqlash samaradorligini oshirishga yo‘naltirilgan ko‘p qatlamli yondashuvni taklif etishdir. Buning uchun quyidagi vazifalar hal etildi: (1) C2 aloqasida ishlatiladigan protokollarni tasniflash; (2) C2 arxitekturalari va beaconing xususiyatlarini ko‘rib chiqish; (3) aniqlash usullarini qiyosiy tahlil qilish; (4) aniqlash uchun belgilar to‘plami va ko‘p qatlamli sxemani asoslash.

MATERIALLAR VA USULLAR

Tadqiqot analitik-qiyosiy xarakterga ega. Birinchidan, botnet va C2 trafikini aniqlash, DGA va DNS tunneling, shifrlangan trafikni tasniflash mavzularidagi ilmiy ishlar [1–7, 9–13, 17–19] hamda DNS, HTTP va TLS protokollari standartlari (RFC) [14–16] o‘rganildi. Ikkinchidan, protokollar C2 aloqasini tashkil etish usuliga ko‘ra tasniflandi va MITRE ATT&CK taktikasi bilan solishtirildi [8]. Uchinchidan, aniqlash usullari to‘rtta mezon bo‘yicha taqqoslandi: shifrlangan trafikda ishlay olishi, yangi (noma’lum) C2 infratuzilmasiga chidamliligi, yolg‘on ijobiy signallar xavfi va joriy etish murakkabligi. Baholashda adabiyotda keng ishlatiladigan ochiq ma’lumotlar to‘plamlari — CTU-13 [1] va CICIDS2017 [19] — xususiyatlari ham inobatga olindi.

Maqoladagi qiyosiy xulosalar mavjud tadqiqotlar va texnik manbalar tahliliga tayanadi; taklif etilgan yondashuvning amaliy sinovi keyingi tadqiqot bosqichida o‘tkaziladi.

NATIJALAR VA MUHOKAMA

1. C2 aloqasida ishlatiladigan protokollar

Tahlil natijalariga ko‘ra, zararli dasturlar C2 kanalini asosan beshta guruh protokollar yordamida tashkil etadi (1-jadval).

1-jadval. C2 aloqasida ishlatiladigan protokollar va ularni aniqlash belgilari

Protokol	Foydalanish usuli	Aniqlash belgilari	Qarshi chora

Protokol	Foydalanish usuli	Aniqlash belgilari	Qarshi chora
HTTP/HTTPS	Oddiy veb-so‘rovlar ko‘rinishida buyruq olish (polling); TLS bilan yuklamani shifrlash; CDN orqali yo‘naltirish	Davriy so‘rovlar (beaconing), nostandart User-Agent, doimiy so‘rov/javob hajmi, nodir TLS parametrlari (JA3)	Proksi va TLS-inspeksiya, JA3/JA3S, davriylik tahlili
DNS	TXT/CNAME/NULL yozuvlari orqali buyruq va ma’lumot uzatish (tunneling); DGA orqali domenlarni dinamik yaratish	Uzun, yuqori entropiyali subdomenlar; ko‘p NXDOMAIN javoblari; bitta domenga ko‘p noyob so‘rov	Rekursiv resolver monitoringi, DGA klassifikatori, DNS RPZ
ICMP	Echo so‘rov/javob yuklamasiga ma’lumot yashirish	Standartdan katta yoki o‘zgaruvchan yuklama, ICMP hajmining keskin o‘sishi	ICMP yuklamasini cheklash, chiquvchi ICMP filtri
Bulut va ijtimoiy xizmatlar (API)	Qonuniy xizmatlar (fayl va xabar almashish) API orqali buyruq almashish	Xizmatga nostandart mijozdan muntazam murojaat, ish vaqtidan tashqari faollik	Ilova darajasida nazorat (CASB), API kalitlari monitoringi
P2P va maxsus protokollar	Markazsiz tarmoq, tugunlar o‘zaro aloqasi	Tugunlar orasida o‘xshash trafik naqshlari, ko‘p manzillar bilan bir xil aloqa	Oqim (flow) tahlili va klasterlash

HTTP(S) keng uchraydigan C2 kanallaridan biri hisoblanadi, chunki uning trafigi oddiy veb-so‘rovlar orasida yo‘qolib ketadi [15]. Zararlangan host davriy ravishda serverga murojaat qilib (polling), buyruq bor-yo‘qligini so‘raydi. TLS yordamida yuklama shifrlangani sababli imzoga asoslangan qoidalar kuchsizlanadi, ammo TLS metama’lumotlari, DNS va oqim xususiyatlari orqali zararli trafikni ajratish mumkinligi ko‘rsatilgan [18].

DNS kanali monitoring uchun muhim: ko‘plab tarmoqlarda DNS so‘rovlariga ruxsat berilgan, protokol tuzilmasi esa [14] subdomen va yozuvlar maydoniga ixtiyoriy ma’lumotni joylashtirish imkonini beradi. Shu tarzda buyruqlar va chiqarib yuboriladigan ma’lumotlar DNS tunneling orqali uzatiladi; past tezlikdagi, kam seziladigan chiqarib yuborish holatlarini aniqlash usullari ham o‘rganilgan [5]. Bundan tashqari, DGA (domain generation algorithm) dasturlari ko‘plab domen nomlarini avtomatik hosil qilib, domenni bloklash orqali himoyalanishni chetlab o‘tadi. Bunday oilalarning o‘lchovlarga asoslangan keng tahlili [11] ishida keltirilgan, DGA botlarini mavjud bo‘lmagan domenlarga (NXDOMAIN) ko‘p murojaat qilishi bo‘yicha topish usuli esa [4] da taklif etilgan.

2. C2 arxitekturalari va beaconing

C2 arxitekturasi markazlashgan, markazsiz (P2P) va gibrid turlarga bo‘linadi. Markazlashgan modelda barcha zararlangan hostlar bitta yoki bir nechta serverga ulanadi; bu oddiy va boshqarishga qulay, ammo server manzili aniqlansa, kanal bir zumda uziladi. Shu sababli hujumchilar DGA va tez almashinuvchi manzillardan (fast-flux) foydalanadi. Markazsiz modelda buyruqlar tugunlar orasida uzatiladi, markaziy nuqta bo‘lmaydi va tarmoq bardoshliroq bo‘ladi; bunday botnetlarni aniqlashda hostlarning statistik oqim xususiyatlari va o‘zaro aloqa naqshlari ishlatiladi [13]. Gibrid model ikkala yondashuvni birlashtiradi.

Markazlashgan C2 ning asosiy xulq-atvor belgisi — beaconing, ya’ni zararlangan hostning serverga ma’lum interval (T) bilan murojaat qilishi. Aniqlashdan qochish uchun hujumchi intervalga tasodifiy og‘ish (jitter) qo‘shadi, ammo bunday aloqaning vaqt taqsimoti baribir inson faoliyatidan hosil bo‘lgan trafikka nisbatan bir xilroq bo‘ladi. Buni murojaatlar orasidagi vaqt farqlarining o‘zgaruvchanlik koeffitsiyenti bilan baholash mumkin:

$CV = \sigma(\Delta t) / \mu(\Delta t)$,

bu yerda Δt — bir juft (host, domen) orasidagi ketma-ket murojaatlar oralig‘i, σ va μ — uning standart og‘ishi va o‘rtacha qiymati. CV qiymati past bo‘lgan juftliklar beaconing nomzodi hisoblanadi. BotSniffer zararlangan hostlarning o‘xshash vaqt va fazoviy xulq-atvor ko‘rsatishiga asoslanadi [3], BotMiner esa aloqa va faoliyat trafigini klasterlash orqali protokol va tuzilmadan mustaqil aniqlashni taklif qiladi [2].

3. Aniqlash yondashuvlarining qiyosiy tahlili

Adabiyotdagi yondashuvlar to‘rt guruhga ajratildi va tanlangan mezonlar bo‘yicha taqqoslandi (2-jadval).

2-jadval. C2 aloqasini aniqlash yondashuvlarining qiyosiy tahlili

Yondashuv	Afzalliklari	Cheklovlari	Manba
Imzoga asoslangan (Snort, Suricata)	Tez; yolg‘on ijobiy signal kam; ma’lum C2 uchun aniq	Yangi va o‘zgartirilgan C2 ni topa olmaydi; shifrlangan yuklamada kuchsiz	Snort, Suricata qoidalari
Statistik va xulq-atvorga asoslangan	Shifrlangan trafikda ishlaydi (metama’lumot asosida); imzo talab qilmaydi	Jitter va tasodifiy kechikishlarga sezgir; chegaralarni sozlash kerak	[2, 3, 13]
DNS tahlili	DGA va tunnelingni erta aniqlaydi; yuklama shifrlanishiga bog‘liq emas	DoH/DoT orqali ko‘rinmay qolishi mumkin; qonuniy CDN domenlari yolg‘on signal beradi	[4, 5, 17]
Mashinali o‘rganish (Random Forest, Isolation Forest, LSTM)	Murakkab naqshlarni topadi; oqim belgilari bilan ishlaydi	Sifatli yorliqli ma’lumot kerak; taqsimot o‘zgarishi; real tarmoqda aniqlik pasayishi mumkin	[1, 6, 9, 10, 18, 20]
Threat intelligence (IOC, STIX/TAXII)	Ma’lum IP, domen va xeshlar bo‘yicha tez bloklash	Infratuzilma tez almashadi; IOC muddati qisqa	STIX/TAXII, MISP

Imzoga asoslangan usullar ma’lum C2 oilalari uchun aniq va tez ishlaydi, biroq o‘zgartirilgan yoki yangi C2 ga nisbatan, ayniqsa shifrlangan trafikda, kuchsiz. Shu sababli tadqiqotlar xulq-atvor va statistik belgilarga yo‘naldi [2, 3, 18].

Mashinali o‘rganish usullari murakkab naqshlarni topish imkonini beradi. Nazoratsiz usullardan biri — Isolation Forest — anomaliyalarni tasodifiy bo‘lish orqali ajratadi va yorliqlangan ma’lumot talab qilmaydi [9]; ketma-ketlikli belgilar uchun LSTM tarmoqlari qo‘llanadi [10]. Biroq Sommer va Paxson ko‘rsatganidek, tarmoq buzilishlarini aniqlashda mashinali o‘rganishni qo‘llash boshqa sohalarga qaraganda murakkabroq: xatolik narxi yuqori, oddiy trafik juda xilma-xil, shuning uchun har qanday anomaliyani emas, ma’lum hujum turlarining o‘zgargan ko‘rinishlarini aniqlashga e’tibor qaratish maqsadga muvofiq [6]. Anomaliyalarni aniqlash usullarining umumiy tasnifi [7] da keltirilgan. Garcia va hammualliflar turli botnet aniqlash usullarini yagona CTU-13 ma’lumotlar to‘plamida taqqoslab, natijalar to‘plam va baholash mezonlariga bog‘liqligini ko‘rsatgan [1].

Threat intelligence (IOC) ma’lum IP, domen va fayl xeshlari bo‘yicha zudlik bilan bloklash imkonini beradi, ammo hujumchi infratuzilmani tez almashtirgani uchun IOC ning amal qilish muddati qisqa. Sandbox (masalan, Cuckoo) hisobotlaridan IOC ajratish C2 infratuzilmasi haqida ma’lumot manbai bo‘lishi mumkin, lekin sandbox tajribalarini loyihalashda xulq-atvorning real muhitdan farq qilishi va izolyatsiya talablarini hisobga olish zarur [12].

4. Ko‘p qatlamli aniqlash yondashuvi

Tahlil shuni ko‘rsatdiki, birorta usul yolg‘iz yetarli emas: imzo tez, ammo cheklangan; statistik va ML usullari kengroq qamrovga ega, ammo yolg‘on signal xavfi yuqori. Shu sababli beshta qatlamdan iborat yondashuv taklif etiladi.

1-qatlam — ma’lumot yig‘ish. Chegara tarmog‘idan PCAP yoki oqim (NetFlow, Zeek) yozuvlari, DNS loglari va TLS metama’lumotlari yig‘iladi; sandbox hisobotlaridan IOC ajratiladi.

2-qatlam — tezkor filtr. Ma’lum IOC va imzo qoidalari bo‘yicha bloklash amalga oshiriladi; bu qatlam ma’lum indikatorlarga mos kelgan hodisalarni ajratib, keyingi tahlil uchun ustuvorlik belgilaydi.

3-qatlam — belgilar ajratish. Har bir (host, domen) juftligi uchun belgilar hisoblanadi. HTTP(S) va oqim bo‘yicha: murojaatlar oralig‘ining o‘rtacha qiymati va CV, so‘rov va javob hajmining dispersiyasi, aloqalar soni, TLS parametrlari. DNS bo‘yicha: subdomen uzunligi, belgilar entropiyasi, NXDOMAIN ulushi, noyob subdomenlar soni, TXT/NULL so‘rovlar ulushi. Belgilar entropiyasi Shennon formulasi bilan hisoblanadi:

$$H(s) = -\Sigma\, p(c) \cdot \log_2 p(c),$$

bu yerda $p(c)$ — c belgisining satrdagi nisbiy chastotasi. Kodlangan ko‘rinishdagi ayrim subdomenlarda H yuqori bo‘lishi mumkin, lekin entropiya yakka holda zararli faoliyatni isbotlamaydi.

4-qatlam — tasniflash. Dastlabki bosqichda yorliqsiz ma’lumotda Isolation Forest [9] yordamida chetga chiqishlar ajratiladi; yorliqlangan ma’lumot (masalan, CTU-13 [1]) to‘plangach, nazorat ostida o‘qitiladigan model (Random Forest [20]) qo‘shiladi; ketma-ketlik belgilari uchun LSTM [10] ishlatilishi mumkin.

5-qatlam — boyitish va ulashish. Tasdiqlangan yoki tahlilchi tomonidan tekshirilgan shubhali manzillar va domenlar IOC sifatida STIX formatida ifodalanib, TAXII orqali ulashilishi mumkin, signal MITRE ATT&CK taktikasiga bog‘lanadi [8] va tahlilchiga qaror qabul qilish uchun uzatiladi.

5. Cheklovlar va muhokama

Birinchi cheklov — shifrlangan DNS (DoH, DoT) va TLS 1.3: bunda nazorat nuqtasidan ko‘rinadigan ma’lumot kamayadi, tahlil metama’lumot va oqim belgilariga tayanishga majbur bo‘ladi [16, 18]. Ikkinchidan, hujumchi jitter va tasodifiy hajmlar bilan davriylik belgilarini yashirishi mumkin, shuning uchun CV kabi yagona belgiga tayanish xavfli. Uchinchidan, qonuniy xizmatlar (dasturiy yangilanishlar, monitoring agentlari, CDN) ham davriy va bir xil trafik hosil qiladi, bu yolg‘on ijobiy signallarni oshiradi; oq ro‘yxatlar va kontekst (jarayon, foydalanuvchi, ish vaqti) bilan boyitish zarur. To‘rtinchidan, ochiq ma’lumotlar to‘plamlari [1, 19] nazorat qilinadigan sharoitda shakllangan, real

tarmoqdagi taqsimot o‘zgarishi tufayli modellar aniqligi pasayishi mumkin [6]. Shu sababli taklif etilgan sxemani yangi trafik bilan muntazam qayta baholash lozim.

6. Ilmiy tadqiqotlar o‘rtasidagi farqlar va tadqiqot bo‘shlig‘i

BotMiner [2] trafikni aloqa va zararli faoliyat guruhlari bo‘yicha klasterlashga tayanadi, BotSniffer [3] esa xostlar o‘rtasidagi vaqt va fazoviy o‘xshashliklarni tekshiradi. Bu ikki yondashuv o‘rtasidagi asosiy farq shundaki, birinchisi bir necha xulq-atvor guruhining mos kelishini, ikkinchisi esa muvofiqlashgan faoliyatning vaqtga oid belgilarini ustuvor ko‘radi. Shu bois ularni faqat umumiy aniqlik foizi orqali solishtirish metodik jihatdan yetarli emas: test trafikining tarkibi, kuzatuv muddati va baholash birligi ham bir xil bo‘lishi zarur.

DGA bilan bog‘liq tadqiqotlarda [4, 11] domen nomlarining generatsiya xususiyatlari va DNS javoblarining taqsimoti muhim bo‘lsa, DNS orqali ma’lumot chiqishini tekshiruvchi ishlarda [5] kam hajmli, uzoq muddatli faoliyat ham e’tiborga olinadi. Shu sababli faqat NXDOMAIN ulushi yoki faqat domen entropiyasiga asoslangan detektor barcha holatlarni bir xil darajada qamrab ololmaydi. Xususiyatlarning bir-birini to‘ldirishini sinovdan o‘tkazish alohida tadqiqot vazifasidir.

Anderson va McGrew [18] shifrlangan zararli trafikni mashinali o‘rganish yordamida tasniflashda noto‘g‘ri yorliqlar va ma’lumotlar taqsimotining vaqt bo‘yicha o‘zgarishini alohida muammo sifatida ko‘radi. Sommer va Paxson [6] ham tarmoqdagi real hodisalar taqsimoti laboratoriya sharoitidan farq qilishini ta’kidlaydi. Demak, modelni faqat bitta tarixiy to‘plamdagi tasodifiy ajratilgan test yozuvlari bilan baholash yetarli emas; vaqt va ssenariy bo‘yicha ajratilgan tekshiruv ham kerak.

Mazkur sharhda aniqlangan tadqiqot bo‘shlig‘i — DNS tuzilishi, oqim statistikasi, davriylik va tahdid kontekstini bir xil baholash protokolida, noto‘g‘ri ijobiy ogohlantirishlar va muhitlararo moslashuvni alohida o‘lchagan holda taqqoslash zaruratidir. Besh qatlamli sxema yangi algoritm sifatida emas, aynan shu bo‘shliqni sinovdan o‘tkazishga xizmat qiluvchi konseptual integratsiya taklifi sifatida talqin qilinadi.

7. Taklif etilgan sxemani tekshirish metodikasi

Birinchi bosqichda tadqiqot savollari oldindan belgilanishi kerak: (RQ1) DNS va oqim belgilarini birlashtirish bitta belgilar guruhiga qaraganda qanday farq qiladi? (RQ2) beaconingni aniqlashda qonuniy avtomatik xizmatlar qanday noto‘g‘ri ogohlantirishlar hosil qiladi? (RQ3) yangi vaqt davri yoki yangi tarmoq muhitida model natijalari qanchalik barqaror qoladi? Har bir savol uchun alohida tajriba va hisobot ko‘rsatkichi ko‘zda tutiladi.

Ma’lumotlar sifatida CTU-13 [1] va CICIDS2017 [19] ochiq to‘plamlari ko‘rib chiqilishi mumkin, lekin ularning sinflari, trafik davri va C2ga tegishli yorliqlari bir xil emas. Shuning uchun bu to‘plamlarni bevosita birlashtirishdan oldin yorliq xaritasi, mavjud xususiyatlar va ssenariy mosligi tekshiriladi. CICIDS2017 umumiy intrusion-detection to‘plami bo‘lib, undagi har bir hujum yozuvini C2 deb belgilash ilmiy jihatdan noto‘g‘ri bo‘ladi.

Tajriba ma’lumotlarini tayyorlashda takroriy yozuvlar, bo‘sh qiymatlar va bir xostning juda o‘xshash sessiyalari nazorat qilinadi. Trening va test ssenariy yoki vaqt bo‘yicha ajratiladi. Normalizatsiya parametrlari faqat trening qismida hisoblanadi; aks holda test haqidagi ma’lumot modelga oldindan sizib o‘tishi mumkin. Ma’lumotlar sinfi nomutanosib bo‘lsa, umumiy accuracy ko‘rsatkichi noto‘g‘ri tasavvur berishi mumkin.

Natijalarni taqqoslash uchun

1)Precision – aniqlik

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

2) Recall – aniqlab olish darajasi

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

3) F1-score – uyg‘unlashtiruvchi o‘rtacha ko‘rsatkich

$$F1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision}+\text{recall}}$$

4) False Positive Rate (FPR) – noto‘g‘ri ijobiy natijalar ulushi

$FPR = \frac{FP}{FP+TN}$ qo‘llanadi. Bunda TP — to‘g‘ri aniqlangan shubhali holatlar, FP — noto‘g‘ri shubhali deb belgilangan normal holatlar, FN — o‘tkazib yuborilgan shubhali holatlar va TN — to‘g‘ri normal deb belgilangan holatlardir. Ko‘rsatkichlar alohida sinf va ssenariy kesimida ham taqdim etilishi kerak.

Baholash kamida uch bazaviy variantni qamrab olishi maqsadga muvofiq: (a) faqat davriylik va oqim statistikasi; (b) faqat DNS xususiyatlari; (c) ularning birlashgan ko‘rinishi. Zarur hollarda ma’lum IOC bilan moslik alohida bazaviy filtr sifatida olinadi. Har bir variantda bir xil test bo‘linishi, bir xil baholash birligi va bir xil ogohlantirish chegarasini tanlash tamoyillari saqlanishi kerak.

Taklif etilgan yondashuvning ilmiy jihatdan tekshirilishi uchun faqat o‘rtacha natijalar emas, xatolar namunasi ham tahlil qilinadi. Masalan, yangilanish xizmati yoki monitoring agenti nega beaconing deb belgilanganini izohlash, DNS so‘rovi qonuniy kodlangan nom bilan bog‘liqmi yoki yo‘qmi tekshirish zarur. Bunday tahlil noto‘g‘ri ijobiy signallarning amaliy qiymatini tushunishga yordam beradi.

8. Qo‘llash shartlari, xavfsizlik va cheklovlar

Monitoring faqat vakolatli tarmoq va ruxsat etilgan ma’lumotlar bilan olib borilishi lozim. DNS jurnallari, xost identifikatorlari va ulanish metama’lumotlari shaxsiy yoki tashkilotga tegishli axborotni o‘z ichiga olishi mumkin. Shuning uchun kirish huquqi, ma’lumotlarni saqlash muddati, anonimlashtirish va audit tartibi tadqiqot boshlanishidan oldin belgilanishi kerak.

Modelning natijasi yakuniy hukm emas, tahlilchi uchun tekshirilishi zarur bo‘lgan signal bo‘lishi kerak. NAT, proksi, umumiy DNS rezolveri, DoH/DoT va bulut xizmatlari xostga tegishlilikni aniqlashni murakkablashtirishi mumkin. Xost, domen va vaqt oralig‘i o‘rtasidagi korrelyatsiyada ma’lumot manbalarining sinxronligi tekshirilishi zarur.

Maqola yangi tajriba o‘tkazilganini da’vo qilmaydi. Keltirilgan taqqoslashlar manbalar tahliliga, besh qatlamli sxema esa konseptual taklifga asoslangan. Uni samaraliroq deb baholash faqat takrorlanuvchi eksperiment, mustaqil test va natijalarni oshkora e’lon qilishdan keyin mumkin bo‘ladi.

XULOSA

Tadqiqot natijasida quyidagi xulosalar olindi:

1) Zararli dasturlar C2 aloqasi uchun HTTP(S), DNS, ICMP, bulut xizmatlari va P2P protokollardan foydalanadi; asosiy tendensiya — kanalni ruxsat etilgan protokollar ichiga yashirish.

2) Imzoga asoslangan usullar ma’lum C2 ga nisbatan samarali, ammo shifrlangan va yangi C2 ni aniqlashda statistik, DNS tahlili va mashinali o‘rganish usullari bilan to‘ldirilishi kerak.

3) Beaconing (CV) va DNS (entropiya, NXDOMAIN ulushi, subdomen uzunligi) belgilariga asoslangan beshta qatlamli aniqlash yondashuvi taklif etildi.

4) Keyingi bosqichda taklif etilgan sxemani CTU-13 va CICIDS2017 ma’lumotlarida eksperimental baholash hamda prototipni Cuckoo Sandbox va STIX/TAXII bilan integratsiya qilish maqsadga muvofiq.

FOYDALANILGAN ADABIYOTLAR

1. Garcia S., Grill M., Stiborek J., Zunino A. An empirical comparison of botnet detection methods // Computers & Security. 2014. Vol. 45. P. 100–123. DOI: 10.1016/j.cose.2014.05.011.
2. Gu G., Perdisci R., Zhang J., Lee W. BotMiner: Clustering analysis of network traffic for protocol- and structure-independent botnet detection // Proc. 17th USENIX Security Symposium. 2008.
3. Gu G., Zhang J., Lee W. BotSniffer: Detecting botnet command and control channels in network traffic // Proc. Network and Distributed System Security Symposium (NDSS). 2008.
4. Antonakakis M., Perdisci R., Nadji Y. et al. From throw-away traffic to bots: Detecting the rise of DGA-based malware // Proc. 21st USENIX Security Symposium. 2012.

5. Nadler A., Aminov A., Shabtai A. Detection of malicious and low throughput data exfiltration over the DNS protocol // Computers & Security. 2019. Vol. 80. P. 36–53. DOI: 10.1016/j.cose.2018.09.006.
6. Sommer R., Paxson V. Outside the closed world: On using machine learning for network intrusion detection // Proc. IEEE Symposium on Security and Privacy. 2010. P. 305–316.
7. Chandola V., Banerjee A., Kumar V. Anomaly detection: A survey // ACM Computing Surveys. 2009. Vol. 41, No. 3. Article 15.
8. MITRE ATT&CK®. Command and Control, Tactic TA0011. URL: https://attack.mitre.org/tactics/TA0011/
9. Liu F.T., Ting K.M., Zhou Z.-H. Isolation forest // Proc. 8th IEEE International Conference on Data Mining (ICDM). 2008. P. 413–422.
10. Hochreiter S., Schmidhuber J. Long short-term memory // Neural Computation. 1997. Vol. 9, No. 8. P. 1735–1780.
11. Plohmann D., Yakdan K., Klatt M., Bader J., Gerhards-Padilla E. A comprehensive measurement study of domain generating malware // Proc. 25th USENIX Security Symposium. 2016.
12. Rossow C., Dietrich C.J., Grier C. et al. Prudent practices for designing malware experiments: Status quo and outlook // Proc. IEEE Symposium on Security and Privacy. 2012.
13. Zhang J., Perdisci R., Lee W., Sarfraz U., Luo X. Detecting stealthy P2P botnets using statistical traffic fingerprints // Proc. IEEE/IFIP International Conference on Dependable Systems and Networks (DSN). 2011.
14. Mockapetris P. Domain names — implementation and specification: RFC 1035. IETF, 1987.
15. Fielding R., Nottingham M., Reschke J. HTTP Semantics: RFC 9110. IETF, 2022.
16. Rescorla E. The Transport Layer Security (TLS) Protocol Version 1.3: RFC 8446. IETF, 2018.
17. Bilge L., Kirda E., Kruegel C., Balduzzi M. EXPOSURE: Finding malicious domains using passive DNS analysis // Proc. NDSS. 2011.
18. Anderson B., McGrew D. Machine learning for encrypted malware traffic classification: Accounting for noisy labels and non-stationarity // Proc. 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2017.
19. Sharafaldin I., Lashkari A.H., Ghorbani A.A. Toward generating a new intrusion detection dataset and intrusion traffic characterization // Proc. 4th International Conference on Information Systems Security and Privacy (ICISSP). 2018.
20. Breiman L. Random forests // Machine Learning. 2001. Vol. 45, No. 1. P. 5–32.