Considerații teoretice privind stabilirea reputației unui domeniu Internet

Mihail DUMITRACHE[1,2,3], Ionuț-Eugen SANDU[1], Adriana-Meda UDROIU[1], Cristian-Alexandru GHEORGHIȚĂ[1,4]

[1] Institutul Național de Cercetare-Dezvoltare în Informatică – ICI București
[2] Facultatea de Litere, Universitatea din București
[3] Academia Oamenilor de Știință din România, București
[4] Facultatea de Automatică și Calculatoare, Universitatea Politehnica din București

mihail.dumitrache@ici.ro, ionut.sandu@ici.ro, meda.udroiu@ici.ro, alexandru.gheorghita@ici.ro

Rezumat: Lucrarea prezintă un demers științific privind tehnologiile și metodologiile existente pentru stabilirea reputației unui domeniu Internet, precum și câteva considerații legate de introducerea tehnicilor de inteligență artificială pentru monitorizarea dinamică a domeniilor. Lucrarea este parte a unui proiect de cercetare a cărui scop este de a asigura gradul de încredere acordat domeniilor împotriva activităților malițioase din spațiul Internet prin monitorizarea automată a acestora și oferirea unor soluții de păstrare a siguranței datelor. Numărul extrem de ridicat de domenii nou înregistrate, proliferarea domeniilor rău intenționate care reprezintă vectori de atac pentru serverele Domanin Name System (DNS) se constituie în cauzele principale care impun găsirea unei soluții automatizate pentru stabilirea nivelului de reputație a unui domeniu. Prin stabilirea nivelului de reputație a domeniului/domeniilor deținute, viitorii utilizatori ai acestora (autorități, instituții ale statului, companii private, persoane fizice etc.) vor avea o imagine corectă a gradului de încredere a domeniului utilizat, creându-se, astfel, un spațiu Internet mai sigur. Alături de tehnicile clasice de stabilire a domeniilor infestate (listele de tip blacklist) folosite în prezent de către administratorii DNS, instrumentele de detecție automată prezentate în lucrare au avantajul detecției rapide și actualizării automate a posibilelor domenii compromise.

Cuvinte cheie: reputație DNS, malware, inteligență artificială, machine learning, domenii Internet, securitate cibernetică, date pasive.

Theoretical considerations about establishing the Internet domain reputation

Abstract: This paper presents a scientific approach regarding the existing technologies and methodologies for establishing the reputation of an Internet domain, as well as some considerations related to the introduction of artificial intelligence techniques for the dynamic monitoring of the domains. The work is part of a research project whose purpose is to ensure the degree of trust given to domains against malicious activities in the Internet space by automatically monitoring them and offering solutions to keep data safe. The huge number of newly registered domains, the proliferation of malicious domains that represent attack vectors for DNS servers are the main causes that require finding an automated solution for establishing the reputation level of each domain. By establishing the level of reputation of the domain/domains owned, their future users (authorities, state institutions, private companies, individuals, etc.) will have a correct image of the degree of trust of the used domain, thus, creating a safer Internet space. Along with the classical techniques for establishing infested domains (blacklist) currently used by DNS administrators, the automatic detection tools presented in the paper have the advantage of fast detection and automatic updating of possible compromised domains.

Keywords: DNS reputation, malware, artificial intelligence, machine learning, Internet domains, cybersecurity, pasive data.

1. Introducere

Sistemul de nume de domenii (Domain Name System - DNS) are un rol important în funcționarea Internetului prin traducerea bidirecțională a numelor de domenii către IP-uri și invers, de la IP-uri la numele de domenii. Dată fiind importanța acestui sistem, încă de la începuturile dezvoltării Internetului, DNS-ul a devenit un instrument important pentru atacatori, exploatându-l, în special pentru a atrage venituri materiale. Rețelele de „botnet” sau campaniile de „phishing” prin

https://doi.org/10.33436/v33i1y202307

e-mail utilizează impersonarea site-urilor drept tehnică de atac. Acestea reprezintă câteva dintre metodele prin care atacatorii malițioși utilizează sistemul DNS folosind nume de domenii similare cu cele vizate de către atacator pentru a masca rețele de calculatoare compromise, folosite pe post de platforme de atac (Udroiu, 2020). Ca răspuns la această utilizare rău intenționată a DNS, listele negre de domenii statice care conțin domenii malware cunoscute au fost folosite de operatorii de rețea pentru a detecta interogările DNS care provin de la mașinile infectate cu malware și pentru a bloca comunicațiile acestora cu atacatorii.

Spațiul online necesită o modalitate rapidă, fiabilă și automată de a evalua și raporta nivelurile de risc ale domeniilor, astfel încât să poată fi adoptate măsuri de securitate adecvate, într-un timp cât mai scurt (Lockheed Martin, 2019). Conceptul de reputație, așa cum este aplicat domeniilor, adreselor IP și adreselor URL, ajută oamenii sau tehnologiile de securitate automatizate să ia decizii cu privire la posibilitatea de a permite, de a refuza sau de a permite condiționat diferite tipuri de conexiuni.

În investigațiile cibernetice, stabilirea reputației domeniului ajută la ghidarea anchetatorului către acele domenii care sunt cele mai susceptibile de a reprezenta risc. Spre deosebire de listele de autorizare/refuzare („blacklist”), scorul reputației permite adaptarea poziției de securitate la nivelul de risc (OWASP, 2020). Securitatea domeniilor este extrem de importantă în această perioadă în care milioane de dolari se pot pierde ajungând către domenii necunoscute prin atacuri sofisticate. Organizațiile au nevoie de controale stricte și sisteme automatizate pentru verificarea traficului către și dinspre Internet.

Din păcate, eficiența listelor negre de domenii statice este din ce în ce mai limitată, deoarece la momentul actual există un număr foarte mare de noi nume de domenii care apar pe Internet în fiecare zi, iar atacatorii trec frecvent la diferite domenii pentru a-și desfășura activitățile rău intenționate, ceea ce face dificilă actualizarea real-time a listelor negre (Fukushima et al., 2011). Pentru a depăși limitările listelor negre de domenii statice este necesar un sistem de detectare care poate determina, în mod dinamic, noi domenii legate de malware. Acest sistem de detectare ar trebui:

- *Să aibă vizibilitate globală* asupra mesajelor de solicitare și răspuns DNS legate de zonele DNS mari. Acest lucru permite „avertizare timpurie”, prin care domeniile malware pot fi detectate înainte ca infecțiile cu malware să ajungă la rețelele locale;
- *Să permită operatorilor DNS să implementeze în mod independent sistemul* și să detecteze domeniile legate de malware din zonele lor de autoritate, fără a fi nevoie de date din alte rețele sau alte coordonări interorganizaționale. Acest lucru va permite detectarea și răspunsul focusat, cu costuri reduse și eficient în timp;
- *Să detecteze cu acuratețe domeniile legate de malware* chiar și în absența datelor despre reputație pentru spațiul de adrese IP indicat de domenii. Datele despre reputația IP sunt adesea dificil de acumulat și sunt fragile. Această problemă poate deveni deosebit de importantă, datorită spațiului de adrese IP extins.

Introducerea conceptului de reputație a domeniilor permite identificarea facilă a numelor de domenii compromise, astfel încât atacatorilor să le fie din ce în ce mai dificil să opereze rețele de calculatoare compromise, iar în ce privește tehnicile de phishing, să se permită alertarea utilizatorului cu privire la adevărata identitate a domeniului pe care îl accesează și în care introduce date personale. De-a lungul anilor s-au studiat tehnici de detectare și identificare a numelor de domenii cu potențial malițios: Notos, Exposure, Kopis, dezvoltându-se algoritmi și unelte care pot identifica un astfel de domeniu cu o acuratețe de 97%, restul de 3% reprezentând rezultate “false pozitive” sau în termenii litearaturii de specialitate, victime colaterale. Rezultatele analizei unui domeniu sunt păstrate în cadrul unor liste numite “blacklist” sau “blocklist”. Aceste liste pot fi interogate de către părțile interesate pentru a verifica prezența unui domeniu, iar pe baza acestei informații să alerteze potențiala victimă cu referire la intențiile malițioase sau să blocheze traficul către acel site.

2. Domain Name Server (DNS). Infrastructură și securitate

2.1. Infrastructura DNS

Infrastructura DNS este formată din sisteme de calcul și de comunicare, care sunt distribuite world-wide-web. Organizarea numelor de domenii are forma unei ierarhii, în care la nivelul cel mai înalt se află root (domeniul rădăcină) care are ca reprezentare un punct („.”). Următorul nivel din ierarhie este denumit domeniu de nivel superior (Top Level Domain - TLD). Fiecare TLD, la rândul său, poate avea mai multe subdomenii.

DNS translatează numele de domenii pe care oamenii le pot memora, în numerele utilizate de computere pentru a căuta destinația. Acest proces este realizat în mai multe etape. Primul nivel în care se desfășoară „căutarea” este nivelul superior al serviciului director (directory service) - sau zona de root (root zone). Deci, pentru a accesa www.google.com spre exemplu, sistemul „întreabă” directorul zonei de root unde poate găsi informații despre domeniul de nivel superior (TLD) „.com”. După ce primește un răspuns, întreabă apoi serviciul de directoare „.com”(primul nivel), identificat de root, unde poate găsi informații despre .google.com (al doilea nivel), și în cele din urmă interoghează serviciul director google.com care este adresa pentru www.google.com (al treilea nivel). După acest proces - care este aproape instantaneu - adresa completă este furnizată computerului care a inițiat cererea.

Diferite entități gestionează fiecare dintre aceste directoare de servicii: Google administrează google.com, VeriSign Corporation administrează „.com”, alte domenii de nivel superior sunt gestionate de către alte organizații, iar zona root de către ICANN - Internet Corporation for Assigned Names and Numbers (Corporația Internet pentru Nume și Numere Alocate).

Vulnerabilitățile descoperite recent referitoare la DNS, combinate cu progresele tehnologice au redus foarte mult timpul necesar pentru ca un atacator să deturneze orice etapă a procesului de căutare DNS și, prin urmare, să preia controlul unei sesiuni sau să redirecționeze utilizatorii către propriile site-uri web capcană, pentru colectarea datelor personale sau de acces ale acestora. Singura soluție pe termen lung pentru această vulnerabilitate este implementarea unor măsuri de securitate de către toți participanții ICANN, Registre, Registrari și Registranți sau utilizatorii de domenii.

2.2. Domeniile .ro

Pentru ca un nume de domeniu de nivel înalt (Top Level Domain) cum este .ro să poată fi accesat în Internet este nevoie ca el să fie introdus în nameserverele root de către IANA - Internet Assigned Numbers Authority (Autoritatea de Numere Alocate pe Internet), autoritatea care aprobă introducerea unui nume de domeniu de nivel înalt în zona root.

Începând cu data de 26 februarie 1993 autoritatea delegată de către IANA pentru înregistrarea numelor de domenii .ro este Institutul Național de Cercetare-Dezvoltare în Informatică - ICI București (IANA, 2020). În acea perioadă, Institutul opera prima linie închiriată cu Universitatea din Viena și furniza conexiune Internet și servicii EARN (European Academic and Research Network) comunității de cercetare din România (Stăicuț, 1995).

Institutul Național de Cercetare-Dezvoltare în Informatică - ICI București, ca Registru național de domenii .ro, este implicat în activitatea unor organizații internaționale precum: ICANN (Internet Corporation for Assigned Names and Numbers), CENTR (Council for European National Top Level Domain Registries) și RIPE (Reseaux Européens IP).

Pentru fiecare zonă DNS există un singur Registru delegat, în consecință, există un singur Registru pentru orice ccTLD (country code Top-Level Domain) sau gTLD (generic Top-Level Domains). Registrele trebuie să se supună regulamentelor adoptate la nivel internațional.

Persoana fizică sau juridică care solicită înregistrarea unui domeniu se numește „Registrant” și poate fi considerat că are un drept de utilizare al respectivului nume de domeniu. Pentru a înregistra un nume de domeniu, un solicitant (Registrant) contactează un Registrar care, în numele Registrantului, interacționează cu Registrul care gestionează baza de date a numelor de domenii înregistrate.

Registrarii sunt companii partenere cu Registrul, companii care oferă în unele cazuri și alte servicii, precum găzduire web și servicii de email. Registrul propagă modificările către baza de date zonă (scrie zona) iar, ulterior, aceasta este folosită pentru a popula serverele DNS și serverul whois.

În general, numele de domenii sunt înregistrate pentru un număr de ani, în funcție de regulile Registrului și de taxele plătite. La sfârșitul perioadei de înregistrare a unui nume de domeniu, deținătorul poate solicita reînnoirea acestuia iar, în caz contrar, acesta poate deveni disponibil pentru ca altcineva să îl înregistreze. Un domeniu poate fi transferat de la un Registrant la altul pe baza unei solicitări.

O caracteristică importantă a Registrului .ro este structura bazei sale de date conform modelului Registru-Registrar-Registrant. Registrarii comunică cu Registrul în numele Registranților (utilizatori de domenii). Această abordare asigură faptul că fiecare obiect care este înregistrat în baza de date este alocat unui Registrar (denumit „Registrar desemnat"), partener cu Registrul, și nici un alt Registrar nu îl poate modifica. Un solicitant al înregistrării are permisiunea de a schimba Registrarul desemnat al unui obiect prin intermediul procesului de transfer.

2.3. Securitate DNS

În ultima decadă, numărul atacurilor cibernetice a crescut exponențial, în special numărul atacurilor de tip distributed denial-of-service (DDoS), phishing și ransomware. Pentru a împiedica astfel de atacuri sau pentru a minimiza impactul acestora este necesar ca Registrul să implementeze o serie de măsuri de securitate la nivel fizic, organizațional, de rețea și de aplicație. Detectarea proactivă a incidentelor de securitate este necesară întrucât permite descoperirea activităților malițioase prin utilizarea unor instrumente de monitorizare. Monitorizarea poate fi internă, prin sistemele proprii sau externă, prin contractarea unor servicii specializate care semnalează intruziunile înainte ca entitățile afectate să devină conștiente de problemă. Securitatea proactivă presupune utilizarea anumitor măsuri pentru prevenirea unui atac sau identificarea unui atac pe măsură ce se produce.

Pentru un operator TLD, integritatea Registrului este esențială pentru operațiunile sale. În funcție de modelul său operațional, un TLD poate furniza servicii partenerilor săi (Registrari) și/sau utilizatorilor de nume de domenii (Registranților). Integritatea Registrului depinde atât de capacitatea Registrarului, cât și a Registrantului de a-și proteja credențialele de acces la sistem. La fel ca alți furnizori de servicii care permit utilizatorilor accesul la sistemele online, operatorii TLD aplică mecanisme de securitate sporite, precum autentificarea în doi factori și accesul pe bază de adrese IP declarate de parteneri (whitelisting).

În cazul unui atac cibernetic care afectează sistemele sau informațiile, Registrul trebuie să notifice clienții despre intruziunea detectată și consecințele acesteia. O amenințare potențială pentru Registre o reprezintă compromiterea conturilor Registrarilor sau Registranților, având drept consecință infiltrarea atacatorilor în sistemul de gestiune a domeniilor. Acest tip de intruziune poate avea ca efect redirecționarea utilizatorului care accesează un nume de domeniu către un site compromis. Atacatorii tind să vizeze nume de domenii cu trafic ridicat și să redirecționeze utilizatorii către site-uri clonă utilizate pentru a colecta informații personale sau pentru implementarea și răspândirea de programe malware.

Alți vectori de atac la securitatea Registrelor de domenii sunt phishing-ul și distribuția de malware prin intermediul domeniilor Doppelgänger. Phishing-ul este o metodă populară de atac de tip social engineering, constând în furtul datelor utilizatorilor, precum credențialele de acces, date personale și informații bancare. Site-urile Doppelgänger sunt utilizate în general pentru distribuirea programelor malware și pentru executarea schemelor de phishing, de obicei prin imitarea instituțiilor financiare sau a agențiilor guvernamentale, pentru a colecta informații personale valoroase care pot fi folosite pentru a fura identități și a devaliza conturile bancare. Aceste tipuri de atac implică înregistrări de domenii noi care imită domenii existente, de obicei cu trafic intens, cu ortografii similare sau permutări ușor de greșit.

Atacurile DDoS reprezintă cea mai mare amenințare la adresa infrastructurii Internet, inclusiv a infrastructurii DNS. Obiectivul unui atac de tip denial of service este de a epuiza

resursele de calcul sau de lățime de bandă ale site-ului țintă sau ale serviciului digital, blocând infrastructura care îl susține. Într-un atac DDoS, traficul utilizat pentru atac provine dintr-o rețea distribuită de sisteme compromise recrutate pentru a trimite cereri simultane către țintă.

Un botnet cuprinde o colecție de dispozitive conectate la Internet, infectate de programe malware, care permit hackerilor să le controleze. Infractorii cibernetici folosesc rețele bot pentru a declanșa atacuri botnet, pentru a obține credențiale, acces neautorizat la sisteme, furt de date etc.

Un atac DDoS poate recruta orice dispozitiv conectat la Internet. Vulnerabilitățile dispozitivelor IoT nesecurizate, inclusiv dispozitivele de tip „smart home", sunt frecvent vizate de atacatori. Utilizatorii nu știu, de cele mai multe ori, că dispozitivele lor sunt implicate într-un atac botnet din cauza programelor malware instalate pe mașinile lor. Atacurile DDoS necesită, de obicei, mii de dispozitive care funcționează concomitent formând un botnet.

Atacatorii care folosesc rețele botnet pot avea acces la câteva mii de computere/dispozitive simultan și le pot da comenzi pentru a derula atacuri cibernetice. În prima fază, aceștia obțin acces la aceste dispozitive utilizând viruși troieni speciali ce atacă sistemele de securitate, iar în faza a doua implementează software de comandă și control ce le permite să desfășoare activități ilicite la scară largă (Dumitrache et al., 2021).

Infrastructura DNS a unui TLD este alcătuită dintr-o rețea de servere DNS publice, situate într-o serie de locații geografice strategic selectate. Aceste servere sunt frecvent ținta atacurilor DDoS. Tipurile principale de atacuri DDoS sunt ,,brute force,, și amplificarea.

Atacurile DDoS de amplificare sunt deosebit de eficiente împotriva infrastructurii DNS și implică trei elemente: spoofing, reflexie și apoi amplificare. Obiectivul unui atacator este să trimită un număr foarte mare de cereri către un anumit server, astfel încât acesta să devină indisponibil, împiedicând totodată trecerea interogărilor legitime. Operarea serverelor DNS publice face ca operatorii de Registre DNS să fie o țintă principală pentru atacuri de amplificare la scară largă (Udroiu et al., 2022).

În cazul atacurilor care valorifică infrastructura DNS, o interogare de 64 de octeți DNS creată poate genera un răspuns de mii de octeți, suprasolicitând astfel ținta cu un volum mare de trafic nedorit.

Un alt tip de atac de amplificare implică interogarea a mii de servere deschise Memcached, care sunt de obicei utilizate pentru a îmbunătăți performanța site-urilor web (Scalzo, 2017). Spre exemplu, în februarie 2018 a avut loc un atac împotriva GitHub cel mai mare serviciu de control al versiunilor software din lume (Kottler, 2018).

DNS flood este un tip de atac DDoS în care un atacator inundă serverele DNS ale unui anumit domeniu în încercarea de a împiedica rezoluția DNS pentru acel domeniu. Astfel, un atac DNS flood compromite capacitatea unui site web, API sau aplicație web să răspundă traficului legitim. Atacurile de acest tip pot fi dificil de identificat, deoarece volumul mare de trafic provine adesea dintr-o multitudine de locații unice.

O altă amenințare cu impact asupra sistemului DNS este DNS Hijacking și Border Gateway Protocol Hijacking (Holland, 2019). Acest lucru se întâmplă atunci când atacatorii se declară proprietari ai unor resurse Internet, precum nameservere sau adrese IP deținute de alte persoane/organizații.

Trecerea de la securitatea cibernetică reactivă la cea proactivă implică îmbunătățirea strategiei existente prin măsuri precum:

- Dezvoltarea unui sistem de reputație a domeniilor pentru identificarea numelor de domenii compromise;
- Instruirea personalului cu privire la cele mai noi practici în materie de securitate și la amenințările de securitate care trebuie evitate;
- Autentificarea multi-factor crește nivelul de securitate al datelor și asigură un control adecvat al accesului;

- Evaluarea riscului de securitate cibernetică și elaborarea unui plan pentru gestionarea tuturor riscurilor de securitate cibernetică;
- Scanarea vulnerabilităților utilizând soluții automatizate ce permite identificarea zonelor vulnerabile înainte de apariția unei amenințări reale;
- Colaborarea cu un centru de operațiuni de securitate SOC (Security Operations Center) care furnizează servicii esențiale de monitorizare și răspuns la incidente. Un SOC include o echipă de experți care poate răspunde instant incidentelor de securitate anulând sau minimizând efectele acestora.

3. Sisteme de monitorizare

Odată cu complexitatea tot mai mare a sistemelor IT, apare nevoia de monitorizare și analiză a sistemelor care se dezvoltă continuu (Banciu et al., 2016), precum sistemele de nume de domenii. Una dintre problemele care apar când se operează sistemele de date de dimensiuni mari vizează monitorizarea lor într-o manieră eficientă, problemă apărută datorită complexității sistemelor. Pentru a maximiza beneficiile monitorizării datelor, acestea trebuie să fie stocate pentru o perioadă extinsă de timp, pentru analize ulterioare, într-un sistem care poate susține rate mari de citire/scriere date. Un prim pas al procesului de monitorizare este achiziționarea de date din surse predefinite, cum ar fi, de exemplu, informații despre numele de domenii și filtrarea acestora astfel încât să nu fie eliminată nici o informație utilă.

Un alt aspect important este înțelegerea și adnotarea comportamentului normal pentru fiecare componentă din ecosistem, iar apoi, pe baza acestuia, identificarea abaterilor și anomaliilor operaționale. Acest proces, combinat cu analiza tranzacției, va evidenția dacă există o problemă (de exemplu, fluctuații nejustificate), iar dacă se descoperă o astfel de problemă, se va identifica cauza și se vor lua măsuri pentru rezolvarea acesteia (Banciu & Fodorean, 2022).

În literatura de specialitate, sistemele de stabilire a reputației și de monitorizare a traficului se bazează pe algoritmi clasici. Totuși, se caută soluții de inteligență artificială (machine learning) care să realizeze această monitorizare într-un mod inteligent și să poată face predicții cât mai corecte cu putință. O inițiativă în acest sens se desfășoară în cadrul consorțiului format din Centrul de Calcul și Departamentul de Informatică din cadrul Universității din Oslo (Lison & Mavroeidis, 2017). Autorii studiului au folosit o rețea neuronală pentru implementarea unui model de reputație, rețea care învață să recunoască rezultatele pe un set de date DNS pasive. Abordarea se bazează pe o arhitectură neuronală profundă care combină o gamă largă de caracteristici legate de proprietățile și structura relațională a înregistrărilor DNS. Modelul este antrenat pe un set mare de date extras din monitorizarea pasivă DNS.

Algoritmii Machine Learning sunt antrenați să facă clasificări și/sau predicții, prin găsirea de informații cheie în proiecte care implică seturi mari de date (Almashhadani et al., 2020). Informațiile preluate sunt utilizate de părțile interesate pentru a lua decizii în cadrul proceselor business.

Un proces de învățare automată implică crearea de algoritmi matematici și statistici care pot accepta date de intrare și pot utiliza un proces de analiză a datelor pentru a face o predicție:

1. Primul pas este colectarea datelor pentru setul de date care urmează să fie analizat;
2. Odată colectate datele, se selectează tipul de algoritm de utilizat, apoi se construiește un model;
3. Modelul este antrenat cu setul de date de testare și utilizat pentru luarea deciziilor viitoare.

În Machine Learning este necesar un set de date consistent și divers pentru a dezvolta o soluție robustă.

Procesarea unui volum mare de date, precum cele asociate numelor de domenii necesită timp și este dificil de realizat conform standardelor umane. Astfel de seturi de date sunt cea mai bună

sursă pentru antrenarea algoritmilor de învățare automată. Cu cât există mai multe date utilizabile și mai citite de mașină într-un set mare de date, cu atât mai eficient va fi antrenamentul algoritmului de învățare automată (Rotună et al., 2022).

3.1. Sisteme de reputație studiate

Mai multe studii anterioare au investigat utilizarea monitorizării DNS pasive pentru identificarea domeniilor rău intenționate.

3.1.1. Notos

Notos este un sistem de reputație dinamic bazat pe observația că utilizările rău intenționate ale DNS au caracteristici unice, care se disting de serviciile DNS legitime furnizate profesional (Antonakakis et al., 2010).

Notos reprezintă o soluție de stabilire a reputației care se bazează pe analiza pasivă a datelor din cadrul serverelor DNS. Aceasta are capacitatea de a detecta un domeniu malițios înaintea sistemelor clasice de "blacklist", cu o acuratețe de 96.8% true-positive și doar 0.38% false-positive.

Notos folosește o gamă largă de caracteristici care pot fi bazate pe rețea (numărul de IP-uri totale asociate istoric cu un domeniu, diversitatea locațiilor geografice, numărul de sisteme autonome distincte în care își au reședința etc.), bazate pe zone de nume de domenii în domenii înrudite, numărul de domenii de nivel superior distincte, frecvențele caracterelor etc.) și bazate pe dovezi (numărul de mostre de malware care au contactat domeniul sau care sunt conectate la o adresă IP indicată de domeniu).

Sursa principală a sistemului o reprezintă informațiile colectate pasiv din cadrul serverelor DNS prin analizarea istorică a IP-urilor rezolvate de către un domeniu. Această bază de date, numită de către autori pDNS, preia date de la doi mari provideri de Internet (ISP) din Atlanta și San Jose, având acces la aproximativ 30.000 de interogări DNS pe secundă. O a doua sursă de informații o reprezintă listele de tipul "blacklist" ce conțin domenii cunoscute ca fiind malițioase precum și o sursă de domenii cunoscute ca fiind legitime, Alexa.com.

Scorul de reputație este calculat în funcție de activitățile pe care domeniul vizat le găzduia. În cazul în care sunt descoperite acțiuni catalogate ca fiind malițioase, scorul domeniul era coborât, iar dacă se constată că domeniul este asociat unor activități legitime scorul este crescut.

O limitare importantă a sistemului Notos o reprezintă incapacitatea de a stabili nivelul de reputație pentru un domeniu "tânăr" deoarece nu există destule informații istorice aflate în bazele de date pasive ale serverelor DNS. Astfel, dacă un atacator continuă să achiziționeze domenii noi pentru activități ilegale, acesta poate eluda foarte ușor sistemul de detecție.

3.1.2. Exposure

Exposure este un sistem similar cu Notos, dar care necesită mai puțin timp de antrenament și date (Bilge et al., 2011). Exposure este capabil să depășească unele dintre limitările Notos, deoarece este capabil să identifice domenii și adrese rău intenționate care nu au fost niciodată văzute în activitățile rău intenționate de mai înainte. Sistemul folosește 15 caracteristici extrase din traficul DNS, permițându-le să caracterizeze diferite proprietăți ale numelor de domenii și modul în care au fost acestea interogate. Mai exact, caracteristicile Exposure sunt fie bazate pe timp (viață scurtă, similaritate zilnică, modele repetate, raport de acces), pe răspunsuri DNS (număr de adrese IP distincte, număr de țări distincte, număr de domenii care partajează adresa IP, rezultate inverse de interogare DNS), pe valoarea TTL - Time To Live (TTL mediu, abatere standard a TTL, numărul de valori TTL distincte, numărul de modificări TTL) cât și bazate pe nume de domeniu (procentul de caractere numerice și lungimea normalizată a celui mai lung subșir semnificativ). Într-o evaluare a lumii reale de două săptămâni, sistemul a identificat 3000 de domenii rău intenționate anterior necunoscute, fără a genera niciun fals pozitiv.

Diferența majoră o reprezintă capacitatea sistemului Exposure de a detecta inclusiv domenii „tinere" folosind metode diferite de analiză, în speță, analiza zonei în care este înregistrat domeniul

și nu neapărat activitatea IP-urilor rezolvate de către acesta în rețea. Exposure a reușit să identifice domenii care au avut durată de viață în medie de una, două zile, atenuând tehnica atacatorilor de a înregistra un număr mare de domenii zilnic pentru a fi folosite drept paravan pentru acțiunile ilicite ale acestora.

Un alt avantaj deținut de către soluția Exposure este acela de a nu necesita un volum mare de date pentru a demara analiza unui domeniu, fiind necesare doar câteva zile de analiză asupra traficului acestuia.

Limitarea principală a soluției Exposure o reprezintă publicarea metodologiei de analiză prin care este determinată reputația unui domeniu, fapt ce poate permite unui atacator să schimbe strategiile de utilizare a domeniilor malițioase și, astfel, să eludeze sistemul de detectare.

3.1.3. Kopis

Spre deosebire de Notos și Exposure, unde ambele se bazează pe monitorizarea traficului de la serverele DNS recursive locale, sistemul Kopis (Antonakakis et al., 2011) utilizează date DNS pasive agregate la nivelurile superioare ale ierarhiei DNS și pot detecta domeniile malware chiar și atunci când nu sunt prezente informații despre reputația IP, analizând modelele globale de rezoluție a interogărilor DNS. Kopis împarte fluxurile de date monitorizate în epoci și rezumă traficul DNS pentru un anumit nume de domeniu la sfârșitul fiecărei epoci, calculând o serie de caracteristici statistice, cum ar fi: diversitatea adreselor IP asociate cu serverele DNS recursive care au interogat un anumit domeniu, volumul relativ de interogări de la serverele DNS recursive de interogare și informațiile istorice legate de spațiul IP la care este indicat domeniul.

Kopis reprezintă o soluție de stabilire a reputației care vizează analiza traficului DNS la cel mai înalt nivel al ierarhiei serverelor DNS. Astfel, soluția poate accesa la nivel global informații referitoare la activitățile unui domeniu, spre deosebire de soluțiile prezentate anterior Notos și Exposure care abordează analiza traficului pe plan local, în cadrul providerilor de Internet (ISP).

Analiza este efectuată strict pe informațiile disponibile din cadrul serverelor DNS și nu sunt luate în calcul reputațiile adreselor IP precum în cazul soluțiilor prezentate anterior. Acest lucru decuplează notarea unui domeniu de notarea adreselor IP permițând aplicarea unui scor chiar și în cazul în care nu există date despre IP-urile rezolvate de către domeniu.

Rata de succes a soluției este de 98.4% true-positive și doar 0.3 - 0.5% false-positive. Suplimentar, sistemul Kopis are capacitatea de a detecta domenii malițioase cu zile sau săptămâni înainte ca acestea să fie înregistrare în listele publice de tip “blacklist”.

4. Concluzii

Stabilirea eficace și eficientă, în timp real, a compromiterii unui domeniu este un obiectiv major al oricărei autorități care gestionează nume de domenii. Numărul mare de domenii care se înregistrează zilnic, descentralizarea lor și explozia de atacuri cibernetice determină ca acest obiectiv major să fie greu de atins. Astfel, dezvoltarea unor tehnologii și instrumente rapide de detecție a apărut din nevoia de a contrabalansa atacurile cibernetice și numărul mare de domenii compromise. Instrumentele care folosesc tehnologii clasice se dovedesc folositoare, dar greu de actualizat și nu acoperă tot spectrul de cerințe. De aceea, introducerea tehnicilor de inteligență artificială de tip Machine Learning s-ar putea dovedi eficientă pentru detecția domeniilor compromise. Cercetarea acestor metode inteligente este încă la început, dar reprezintă o soluție bună de a elimina domeniile compromise sau de a alerta asupra acestora, printr-o învățare artificială și o clusterizare a acestora. Aceste direcții de cercetare reprezintă ariile viitoare pe care dorim sa le abordăm în cadrul proiectului științific elaborat.

Confirmare

Acest articol a fost realizat în cadrul proiectului „Platforma de monitorizare automată a domeniilor Internet prin dezvoltarea unui sistem dinamic de stabilire a reputației (TLDRep)”

finanțat de către Ministerul Cercetării, Inovării și Digitalizării (MCID), prin Programul Nucleu PN 2338 02 01 și al proiectului „Instrumente de transformare digitală pentru eGuvernare prin utilizarea domeniilor .ro” finanțat de Academia Oamenilor de Știință din România prin competiția „AOSR-TEAMS-II” EDIȚIA 2023-2024 – „Transformarea digitală în științe”. Mulțumim colegilor participanți pentru colaborare.

REFERINȚE BIBLIOGRAFICE

Almashhadani, A. O., Kaiiali, M., Carlin, D. & Sezer, S. (2020) Maldom Detector: A system for detecting algorithmically generated domain names with machine learning. *Computers & Security*. 93, 101787. doi: 10.1016/j.cose.2020.101787.

Antonakakis, M., Perdisci, R., Dagon, D., Lee, W. & Feamster, N. (2010) Building a Dynamic Reputation System for DNS. In: *Proceedings of 19th USENIX Security Symposium (USENIX Security'10)*. pp. 273-290.

Antonakakis, M., Perdisci, R., Lee, W., Vasiloglou, N. & Dagon, D. (2011) Detecting Malware Domains at the Upper DNS Hierarchy. In: *Proceedings of 20th USENIX Security Symposium (USENIX Security '11), San Francisco, CA.* USA, USENIX Association https://www.usenix.org/legacy/events/sec11/tech/full_papers/Antonakakis.pdf [Accessed 22 February 2023]. doi: 10.5555/2028067.2028094.

Banciu, D. & Fodorean, D. (2022) E-Learning developer: specialized training program for online teaching. *Romanian Journal of Information Technology and Automatic Control [Revista Română de Informatică şi Automatică]*. 32(2) 117-126. doi: 10.33436/v32i2y202209.

Banciu, D., Petre, I., Smada, D. & Gheorghiță, A. (2016) Sistem de măsurare şi monitorizare a indicatorilor privind societatea informaţională. *Revista Română de Informatică şi Automatică [Romanian Journal of Information Technology and Automatic Control]*. 26(3), 17-24.

Bilge, L., Kirda, E., Kruegel, C. & Balduzzi, M. (2011) EXPOSURE: Finding Malicious Domain Using Passive DNS Analysis. In: *Proceedings of the Network and Distributed System Security Symposium (NDSS), 6 – 9 February 2011, San Diego, California, USA*. pp. 1-17. https://sites.cs.ucsb.edu/~chris/research/doc/ndss11_exposure.pdf [Accessed 17 January 2023].

Dumitrache, M., Sandu, I. E. & Petre, I. (2021) Soluții pentru implementarea funcțiilor de securitate în cazul aplicațiilor tipice în medii SMART. *Revista Română de Informatică şi Automatică [Romanian Journal of Information Technology and Automatic Control]*. 31(2), 111-124. doi: 10.33436/v31i2y202109.

Fukushima, Y., Hori, Y. & Sakurai, K. (2011) Proactive Blacklisting for Malicious Web Sites by Reputation Evaluation Based on Domain and IP Address Registration. In: *Proceedings of International Joint Conference of IEEE TrustCom-11/IEEE ICESS11/FCST-11*. pp. 352-361. doi: 10.1109/TrustCom.2011.46.

Holland, B. (2019) TLD Operator Perspective on the Changing Cyber Security Landscape. *Centre for International Governance Innovation.* https://www.cigionline.org/articles/tld-operator-perspective-changing-cyber-security-landscape/ [Accessed 20 January 2023].

IANA. (2020) *Root Zone Database*. https://www.iana.org/domains/root/db [Accessed 22.01.2023].

Kottler, S. (2018) DDoS Incident Report. *GitHub Blog.* https://github.blog/2018-03-01-ddos-incident-report [Accessed 23.01.2023].

Lison, P. & Mavroeidis, V. (2017) Neural reputation models learned from passive DNS data. In *2017 IEEE International Conference on Big Data (Big Data), 11-14 Decembre 2017*. pp. 3662-3671. doi: 10.1109/BigData.2017.8258361.

Lockheed Martin. (2019) *Cyber Kill Chain®*. https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html [Accessed 24 January 2023].

OWASP. (2020) *Open Source Foundation for Application Security*. https://www.owasp.org. [Accesat 26 January 2023].

Rotună, C. I, Dumitrache, M. & Sandu, I. E. (2022) Assessment of Machine Learning algorithms for automated monitoring. *Romanian Journal of Information Technology and Automatic Control [Revista Română de Informatică şi Automatică]*. 32(3), 73-84. doi: 10.33436/v32i3y202206.

Scalzo, F. (2017) DNS-based threats: DNS reflection and amplification attacks. *Verisign blog*. https://blog.verisign.com/security/dns-based-threats-dns-reflection-amplification-attacks/ [Accessed 10 February 2023].

Stăicuț, E. (1995) Domain Name Systems, InterNic and RIPE Procedures. In: *Proceedings of the NATO Advanced Networking Workshop, The First CEENet Workshop on Networks Technology, The road to Global connectivity, A CEEet Publication, 15-24 September 1995, Warsaw, Poland*.

Udroiu, A. M., Sandu, I. & Dumitrache, M. (2022) Integrated Information System for the Management of Activities in the Organization. *Studies in Informatics and Control*. 31(2), 25-35. doi: 10.24846/v31i2y202203.

Udroiu, A. M. (2020) Designing Cybersecurity Solution Using Blockchain Technology. In: Cîrciumaru, F & Potîrniche, M (eds.) *Proceedings of the International Conference Strategies XXI Proceedings.* Bucharest, Romania, "Carol I" National Defence University Publishing House. pp. 210-215.

Mihail DUMITRACHE este absolvent al Facultății de Electrotehnică, Universitatea Politehnica din București, specializarea „Inginerie Asistată de Calculator", inginer și doctor în Inginerie Electrică. Deține studii masterale în specializarea „Inginerie Electrică", Universitatea Politehnica din București și în specializarea „Administrație Publică Electronică", Universitatea din București. Și-a început activitatea profesională în cadrul Institutului Național de Cercetare-Dezvoltare în Informatică – ICI București în anul 2002, ca programator. În prezent este Cercetător Științific gradul III, Șef la Departamentul „Administrare domenii RoTLD" – ICI București și Lector Universitar la Universitatea din București. Este autor și coautor al unor studii și articole de specialitate.

Mihail DUMITRACHE graduated from Politehnica University of Bucharest, the Faculty of Electrical Engineering with the specialization "Computer Assisted Engineering", he is an engineer and holds a PhD degree in Electrical Engineering. In between, he obtained two Master's Degrees, one in Electrical Engineering at Politehnica University of Bucharest and one in Electronic Public Administration, at the University of Bucharest. His professional career started at the National

Institute for Research and Development in Informatics – ICI Bucharest in 2002 as a computer programmer. Currently, he is a Scientific Researcher III and Head of the .ro Domain Administration Department (RoTLD) – ICI Bucharest and also a Lecturer at the University of Bucharest. He is the author and co-author of several scientific studies and articles.

Ionuț-Eugen SANDU este licențiat în Știința Sistemelor și a Calculatoarelor (2006), obține master în Administrație Publică Electronică în anul 2007. Din anul 2010 devine cercetător științific la Departamentul de Administrare Domenii .ro din cadrul Institutului Național de Cercetare-Dezvoltare în Informatică – ICI București, iar începând cu anul 2015 devine Șef Serviciu Tehnic RoTLD și Cercetător Științific gradul III în cadrul aceluiași Institut. Domeniile sale principale de interes sunt: administrare sisteme, dezvoltare de noi servicii, dezvoltare a infrastructurii de comunicații, precum și relația cu partenerii. În prezent este Director Tehnic al Institutului Național de Cercetare-Dezvoltare în Informatică – ICI București.

Ionuț-Eugen SANDU graduated university with a BS în Computer and Systems' Science (2006) and obtained a Master's Degree in Electronic Public Administration in 2007. In 2010, he became Scientific Researcher within the .ro Domain Administration Department (RoTLD) of the National Institute for Research and Development in Informatics - ICI Bucharest, and since 2015 is Scientific Researcher grade III and Head of the Technical Division of RoTLD, with responsibilities in systems' administration, development of new services, development of communication infrastructures. He is also in charge with maintaining a close relationship with partners. Currently, he is Technical Director of National Institute for Research & Development in Informatics – ICI Bucharest.

Adriana-Meda UDROIU activează ca Manager securitatea informației (CISO - Chief Information Security Officer) în cadrul Institului Național de Cercetare-Dezvoltare în Informatică – ICI București. Are o activitate didactică și de cercetare de peste 20 de ani în domeniul TIC. A obținut titlul de doctor în ,,Sisteme automate" la Universitatea Politehnica din București în anul 2003. Este autor a peste 10 volume în domeniul tehnologiei informației și comunicațiilor și a peste 50 de articole publicate în reviste naționale și internaționale indexate BDI și ISI, precum și director/responsabil a numeroase proiecte de cercetare din domeniul TIC. Este conferențiar universitar asociat la Facultatea de Automatică și Calculatoare din cadrul Universității Politehnica din București și la Facultatea de Sisteme electronice și informatice militare din cadrul Academiei

Tehnice Militare „Ferdinand I”. Principalele domenii de interes pentru activitatea de cercetare sunt: TIC, securitate cibernetică, infrastructuri critice, securitatea informației, e-learning, formare continuă.

Adriana-Meda UDROIU works as Chief Information Security Officer at the National Institute for Research and Development in Informatics. She has a teaching and research activity for over 20 years in the field of ICT. She obtained his PhD degree in Automated Systems at Politehnica Bucureşti in 2003. She is an author of more than 10 volumes in ICT field and more than 50 articles published in national and international journals BDI and ISI indexed. Also, she is director/manager in numerous research ICT projects. She is associate professor at Faculty of Automatic Control and Computers, University Politehnica of Bucharest and at the Faculty of Military Electronic and Information Systems, Military Technical Academy „Ferdinand I”. The main areas of interest to research are: ICT, cybersecurity, critical infrastructures, information security, elearning, lifelong learning.

Cristian-Alexandru GHEORGHIȚĂ a absolvit Facultatea de Informatică din cadrul Universității din București, în anul 2013. Este doctorand la Facultatea de Automatică și Calculatoare, Universitatea Politehnica din București. În prezent lucrează ca cercetător în cadrul Institutului Național de Cercetare-Dezvoltare în Informatică – ICI București. Principalele sale domenii de interes sunt: Cyber Security, Big Data, Cloud Computing, Cloud-Native, DevOps. Este implicat în proiecte de cercetare specifice societății informaționale. Cercetările sale au fost publicate în articolele revistelor de specialitate și în lucrările conferințelor științifice.

Cristian-Alexandru GHEORGHIȚĂ graduated the Faculty of Informatics, University Bucharest in 2013. He is a Ph.D. student at Faculty of Computer Science and Automatics, Politehnica University of Bucharest. Currently he works as Researcher at I.C.I Bucharest. His main areas of interest are Cyber Security, Big Data, Cloud Computing, Cloud-Native, DevOps. He is involved in research projects specific to the Information Society. His research was published in journal articles and proceedings of conferences.