STRATEGII DE DETECŢIE A PĂTRUNDERILOR NEAUTORIZATE ÎN SISTEME INFORMATICE

Teză destinată obţinerii
titlului ştiinţific de doctor inginer
la
Universitatea "Politehnica" din Timişoara
în domeniul INGINERIA SISTEMELOR
de către

Ing. Emanuel Ciprian Sasu

Conducător ştiinţific: prof.univ.dr.ing. Octavian Proştean
Referenţi ştiinţifici: prof.univ.dr.ing. Mihail Abrudean
prof.univ.dr.ing. Valentina Bălaş
prof.univ.dr.ing. Mircea Vlăduţiu

Ziua susţinerii tezei: 22.11.2013

Seriile Teze de doctorat ale UPT sunt:

1. Automatică
2. Chimie
3. Energetică
4. Ingineria Chimică
5. Inginerie Civilă
6. Inginerie Electrică
7. Inginerie Electronică şi Telecomunicaţii
8. Inginerie Industrială
9. Inginerie Mecanică
10. Ştiinţa Calculatoarelor
11. Ştiinţa şi Ingineria Materialelor
12. Ingineria sistemelor
13. Inginerie energetică
14. Calculatoare şi tehnologia informaţiei
15. Ingineria materialelor
16. Inginerie şi Management

Universitatea „Politehnica” din Timişoara a iniţiat seriile de mai sus în scopul diseminării expertizei, cunoştinţelor şi rezultatelor cercetărilor întreprinse în cadrul şcolii doctorale a universităţii. Seriile conţin, potrivit H.B.Ex.S Nr. 14 / 14.07.2006, tezele de doctorat susţinute în universitate începând cu 1 octombrie 2006.

România, 300159 Timişoara, Bd. Republicii 9,
tel. 0256 403823, fax. 0256 403221
e-mail: editura@edipol.upt.ro

Cuvânt înainte

Teza de doctorat a fost elaborată pe parcursul activităţii mele în cadrul Departamentului de Automatică şi Informatică Aplicată, din cadrul Universităţii „Politehnica” din Timişoara.

Lucrarea de faţă se adresează tuturor celor interesaţi de domeniul securizării reţelelor de calculatoare împotriva pătrunderilor neautorizate. Cercetarea s-a focalizat pe unul dintre posibilele atacuri, şi anume cel realizat prin intermediul falsificării adreselor fizice ale interfeţei de reţea, numite adrese MAC. Luându-se în considerare tendinţele actuale în acest domeniu, teza de faţă îşi aduce propria contribuţie prin dezvoltarea unei noi metode de detecţie a falsificării adreselor MAC, metodă care stabileşte identitatea unei staţii pe baza unei amprente alcătuite din destinaţiile cu care staţia comunică în mod constant.

Consider că lucrarea de faţă reprezintă un suport ştiinţific important în cercetările viitoare care au ca subiect detecţia adreselor MAC falsificate.

Timişoara, octombrie 2013 Emanuel Ciprian SASU

Alese mulţumiri şi profundă recunoştinţă se cuvin adresate conducătorului de doctorat **prof.dr.ing. Octavian PROŞTEAN** pentru consilierea permanentă şi îndrumarea atentă pe tot parcursul realizării lucrării.

Doresc să mulţumesc domnului **prof.dr.ing. Ioan FILIP** şi domnului **ş.l.dr.ing. Cristian VAŞAR** pentru sprijinul acordat pe durata stagiului doctoral şi pentru sugestiile importante referitoare la elaborarea acestui material.

Mulţumesc, de asemenea domnului **prof.dr.ing. Mihail ABRUDEAN**, doameni **prof.dr.ing. Valentina BĂLAŞ** şi domnului **prof.dr.ing. Mircea VLĂDUŢIU**, în calitate de referenţi ai lucrării şi domnului **prof.dr.ing. Radu Emil PRECUP** în calitate de preşedinte al comisiei.

Nu în ultimul rând, doresc să mulţumesc familiei mele pentru suportul şi încrederea acordată în toată perioada stagiului doctoral.

Sasu, Emanuel Ciprian

Strategii de detecţie a pătrunderilor neautorizate în sisteme informatice

Teze de doctorat ale UPT, Seria 12, Nr. 8, Editura Politehnica, 2013, 124 pagini, 70 figuri, 10 tabele.

ISSN: 2068-7990

ISBN: 978-606-554-730-8

Cuvinte cheie: securitate, adrese MAC, atacuri cibernetice, pătrunderi neautorizate, amprente de trafic validare identitate

Rezumat,

Prin subiectul abordat, teza de doctorat răspunde unor probleme de maximă actualitate privind securizarea reţelelor de calculatoare împotriva pătrunderilor neautorizate, atacuri realizate prin intermediul falsificării adreselor MAC. Luându-se în considerare tendinţele actuale, teza îşi aduce propriile contribuţii prin dezvoltarea unei metode noi prin care se detectează pătrunderile neautorizate prin falsificarea adreselor MAC. Metoda se numeşte *„Destination Traffic Fingerprint”* (DTF) şi validează adresele MAC întâlnite în trafic pe baza unei amprente care are în componenţa sa destinaţiile IP cu care staţia comunică în mod constant.

Concluzia lucrării este că utilizarea acestei metode de detecţie aduce o serie de avantaje importante şi oferă o bună recunoaştere a identităţii adreselor MAC întâlnite în trafic.

CUPRINS

NOTAŢII ŞI ABREVIERI

AP	Access Point
DHCP	Dynamic Host Control Protocol - reprezintă un protocol ce permite atribuirea automată a adresei IP pentru o staţie din reţea
DIPA	setul format din destinaţiile IP aferente amprentei actuale în cadrul metodei DTF
DIPR	setul format din destinaţiile IP aferente amprentei de referinţă în cadrul metodei DTF
FPW	puterea amprentei de referinţă în cadrul metodei DTF
HCF	reprezintă o metodă de detecţie a adreselor MAC falsificate prin filtrare hop-count
LP	limita minimă de prezenţă a unei destinaţii IP, pentru a putea fi inclusă într-o amprentă de referinţă
LPOS	limita minimă necesară pentru ca o adresă MAC să fie declarată ca posibil validă în cadrul metodei DTF
LREC	limita minimă necesară pentru ca o adresă MAC să fie declarată ca validă în cadrul metodei DTF
MAC	Media Access Controll - adresa interfeţei de reţea a unei staţii
MFPW	valoarea medie a puterii amprentei de referinţă în cadrul metodei DTF
MININTERV	numărul minim de subintervale necesar pentru a admite o destinaţie IP în amprenta de referinţă în cadrul metodei DTF
NIP	numărul total de adrese din spaţiul IPv4
NP	reprezintă un pachet captat din reţea
ODR	Overall Degree of Recognition - Gradul Global de Recunoaştere în cadrul metodei DTF
PA	procent de absenţă maximă în cadrul metodei DTF

PP	procent de prezenţă în cadrul metodei DTF
RD	gradul de recunoaştere al unei destinaţii IP în cadrul metodei DTF
RFC	reprezintă o publicaţie a Internet Engineering Task Force cu privire la date tehnice şi standarde în Internet
RSS	nivelul semnalului wireless recepţionat
TCP	reprezintă un protocol clasic de comunicare în reţea
TDIP	numărul total de destinaţii IP ce intră în componenţa amprentei în cadrul metodei DTF
TMAC	numărul maxim de minute consecutive, în care nu s-a identificat trafic către destinaţia IP evaluată, în cadrul metodei DTF
TME	numărul total de minute evaluate pentru extragerea amprentei de referinţă, în cadrul metodei DTF
TMP	numărul total de minute în care s-a identificat trafic către o destinaţie IP, în cadrul metodei DTF
TTL	Time-To-Live - este un parametru folosit în cadrul transmiterii informaţiilor în reţea
TU	reprezintă unitatea de tip pentru care se extrag datele în vederea obţinerii amprentei de referinţă în cadrul metodei DTF
UDP	reprezintă un protocol clasic de comunicare în reţea
VLAN	reprezintă o reţea virtuală

LISTA DE TABELE

LISTA DE FIGURI

1. INTRODUCERE

1.1. Oportunitate şi obiective.

Teza îşi propune să aducă o serie de contribuţii în domeniul detecţiei pătrunderilor neautorizate în reţelele de calculatoare. O dată cu dezvoltarea calculatoarelor, a început să apară tot mai frecvent nevoia de comunicare între acestea, pentru a permite transferul de informaţie şi o prelucrare mult mai facilă a datelor. Dacă două sau mai multe calculatoare pot schimba informaţie între ele, se spune că sunt legate într-o „reţea de calculatoare". Andrew S. Tanenbaum afirmă în [Tan-03] că *„diferenţele dintre colectarea, transportul, stocarea şi procesarea informaţiilor dispar rapid. Oraganizaţii cu sute de birouri sunt răspândite pe o mare arie geografică şi se aşteaptă să fie capabile oricând să examineze starea oricăruia, printr-o simplă apăsare de buton*".

Utilizarea reţelelor este astăzi larg răspândită şi oferă utilizatorilor o gamă variată de servicii, de la cele mai simple până la servicii extrem de complexe. Problemele de securitate însă cresc pe măsură ce se dezvoltă reţelele de calculatoare [Sar-13], [Chu-12], [Nat-12], [Gao-11], [Del-10], [Gup-10], [Bea-09], [Nis-05]. Unele dintre acestea sunt pur întâmplătoare dar majoritatea sunt intenţionate. Cele din urmă sunt cunoscute în literatură ca „atacuri" şi au ca scop întreruperea unor servicii, utilizarea neautorizată a unor servicii, sau chiar furtul de informaţie. Indiferent de cauza atacului, sistemul este afectat negativ într-un grad mai mic sau mare. Aceste efecte sunt nedorite şi trebuie detectate cât mai rapid şi eliminate pentru ca sistemul distribuit să poată funcţiona la parametrii normali.

Una dintre problemele de securitate majore este dată de pătrunderea în sistem a unor utilizatori neautorizaţi, fie cu scop distructiv, fie pentru a beneficia de anumite facilităţi pe care le au utilizatorii autorizaţi. Aceste pătrunderi neautorizate se folosesc de diverse vulnerabilităţi ale sistemului, care permit accesul unui intrus şi validarea lui ca utilizator autorizat.

Oportunitatea lucrării de faţă este definită în contextul detectării rapide a pătrunderilor neautorizate prin falsificarea adreselor fizice ale interfeţelor de reţea, numite MAC (Media Access Control).

Principalele obiective propuse în lucrarea de faţă sunt următoarele:

- Evaluarea stadiului actual în domeniul pătrunderilor neautorizate în reţelele de calculatoare prin falsificarea adreselor MAC, pentru determinarea gradului de acoperire pe care aceste metode le au în abordarea problemelor concrete apărute în practică.
- Dezvoltarea unei metode noi de detecţie a falsificării adreselor MAC, care să valideze o staţie prin compararea unei amprente de trafic de referinţă, stabilită într-o fază iniţială de evaluare, cu amprenta actuală identificată în traficul monitorizat.
- Modelarea, implementarea şi analizarea fiabilităţii sistemului de detecţie propus.
- Implementarea unui mediu de simulare adecvat problematicii studiate

1.2. Prezentarea conţinutului lucrării.

Conţinutul lucrării este dezvoltat pe parcursul a şase capitole.

În primul capitol, sunt prezentate principalele obiective şi modul în care acestea determină structurarea lucrării.

Capitolul 2 parcurge stadiul actual de dezvoltare a metodelor de detecţie a pătrunderilor neautorizate prin falsificarea adreselor fizice ale interfeţelor de reţea (adresele MAC). Se prezintă o analiză critică a metodelor evaluate, prin prisma răspunsului acestora la diverse situaţii specifice.

În capitolul 3 se dezvoltă o metodă originală de detecţie a adreselor MAC falsificate, numită *„Destination Traffic Fingerprint*" (DTF). Se demarează cu o descriere generală a metodei, şi stabilirea domeniului de aplicabilitate al acesteia. Se prezintă o serie de parametrii care au fost introduşi cu scopul de a fi folosiţi în procesul de determinare a traficului constant. În continuare se dezvoltă procesul de determinare al Gradului Global de Recunoaştere pentru o adresă MAC, analizând comparativ varianta de calcul standard cu varianta de calcul ponderat. Este realizată o modelare matematică a determinări amprentei de referinţă şi validarea acesteia în timp real. Modelarea Fuzzy este abordată în contextul determinării traficului constant, punându-se în evidenţă avantajele pe care le aduce în procesul determinării amprentei de referinţă. Capitolul se încheie cu prezentarea unor servicii şi tehnologii, care favorizează utilizarea metodei DTF în practică.

În capitolul 4 se prezintă Toolbox-ul Software care a fost conceput pentru studiul metodei DTF. Este alcătuit în esenţă din două aplicaţii software realizate de către autor pentru captarea pachetelor de date din reţea şi prelucrarea lor în vederea aplicării metodei DTF. Aplicaţiile permit vizualizarea modului de lucru, începând cu faza de extragere a amprentei de referinţă şi până la validarea acesteia în timp real pe baza unor module de simulare.

În capitolul 5 se prezintă rezultatele obţinute în urma unor experimente care au avut ca scop demonstrarea aplicabilităţii metodei DTF. În prima parte sunt descrise caracteristicile experimentului şi motivaţia alegerii cadrului de evaluare. În partea a doua sunt prezentate rezultatele experimentale cu privire la variaţia în timp a amprentelor de trafic, demonstrând faptul că există destinaţii IP care sunt prezente în traficul unei staţii pe durate lungi de timp, şi care pot forma în felul acesta o amprentă a staţiei, utilă în procesul de validare. Ultima parte prezintă câteva concluzii în urma experimentului realizat.

În finalul lucrării sunt prezentate concluziile, contribuţiile personale şi câteva dintre direcţiile de cercetare generate de studiile efectuate. Teza se întinde pe 122 pagini, conţine 70 figuri, 10 tabele şi 93 referinţe bibliografice. Validarea contribuţiilor s-a realizat prin publicarea a 7 lucrări ştiinţifice, la care autorul tezei este prim autor, după cum urmează:

- 3 lucrări publicate în volumele unor conferinţe indexate ISI Proceedings
- 4 lucrări publicate în volumele unor conferinţe indexate BDI

2. SECURIZAREA REŢELELOR DE CALCULATOARE ÎMPOTRIVA PĂTRUNDERILOR NEAUTORIZATE

Capitolul de faţă abordează câteva aspecte importante din domeniul securizării reţelelor de calculatoare. Se prezintă necesitatea securizării reţelelor şi se analizează un caz aparte de securizare împotriva pătrunderilor autorizate prin falsificarea adreselor MAC. Se prezintă o formalizare a detecţiei adreselor MAC falsificate şi o serie de metode actuale, care se aplică în acest context.

La final, se realizează o analiză critică a metodelor prezentate, prin prisma răspunsului acestora la diverse situaţii concrete. Analiza scoate în evidenţă atât aspectele „tari" cât şi cele „slabe", oferind o imagine de ansamblu extrem de utilă în vederea înţelegerii modului în care se abordează în momentul de faţă pătrunderile neautorizate prin falsificarea adreselor MAC.

2.1. Necesitatea securizării reţelelor împotriva pătrunderilor neautorizate.

Într-un sistem distribuit, securitatea comunicării dintre componente este absolut vitală. Funcţie de implementarea sistemului distribuit, se pot identifica o serie de probleme care pot afecta securitatea în ansamblul ei. Reţelele de calculatoare reprezintă un mediu de interconectare a modulelor unui sistem distribuit. Problemele de securitate care pot să apară la nivelul unei reţele, sunt foarte diverse, literatura de specialitate abordând pe larg acest domeniu şi oferind soluţii care pot fi implementate cu succes, în vederea creşterii fiabilităţii şi a gradului de încredere asociat, cum ar fi cele din [Ahm-13], [Kuf-13], [Gau-13], [Yan-11], [Dis-11], [Xin-11], [Bzo-11], [Pop-11], [Dik-10], [Mie-10].

Atacurile sunt frecvente şi pot avea consecinţe de la cele mai inofensive, până la cele mai grave [Wax-11]. S. Woo şi colaboratorii au dezvoltat în [Woo-13] o nouă metodă de detecţie a tiparelor folosite în atacuri. Aceste tipare, bazate pe ontologii de comportament, sunt organizate pe ierarhii şi reprezintă o nouă direcţie de abordare a problemelor de securitate. De asemenea, în [Raz-13], se prezintă cele mai cumune ameninţări, precum şi soluţiile aplicate cel mai frecvent.

Din marea varietate de problematici, lucrarea de faţă abordează un subdomeniu, şi anume accesul neautorizat în sistem. Dacă se face referire la o reţea de calculatoare, accesul neautorizat înseamnă de fapt ca un utilizator, voluntar sau involuntar, să intre în reţea şi să acceseze diverse resurse, la care în mod normal, nu ar avea acces. Accesul neautorizat este tratat pe larg in literatură prin diverse lucrări, cum ar fi [Pin-13], [Esw-13], [Zha-12], [Sha-12], [Som-10], [Cor-11], [Xia-10], scopul preocupărilor constând în principal în detectarea cât mai rapidă a pătrunderilor neautorizate, pentru a evita consecinţele ce pot deriva din aceste situaţii.

În [Avi-04], se defineşte *dependabilitatea* unui sistem ca fiind *„abilitatea de a furniza servicii, care în mod sigur pot fi considerate de încredere"*.

Pentru o reţea de calculatoare, este foarte important să se asigure faptul că utilizatorii reţelei sunt persoane autorizate, demne de încredere. Se defineşte de asemenea *securitatea* ca fiind „*un compus format din atributele de confidenţialitate, integritate şi disponibilitate, necesitând existenţa simultană a disponibilităţii numai pentru acţiunile autorizate, confidenţialităţii şi integrităţii în sens de autorizare*". Autorizarea are un rol esenţial în securitatea reţelelor de calculatoare.

Problema pătrunderilor neautorizate în sistem se încadrează în categoria atacurilor. Un utilizator extern poate să pătrundă într-un sistem restricţionat, fie pentru a produce pagube voite, fie pentru a beneficia de anumite drepturi pe care le au utilizatorii autorizaţi. De exemplu, se poate ca un utilizator să introducă în reţeaua companiei la care lucrează, laptop-ul personal, doar pentru a beneficia de o conexiune Internet de viteză foarte mare sau de alte servicii oferite. Astfel de utilizatori nu vor distruge informaţii, dar pot afecta utilizarea lăţimii de bandă. Cele mai mari probleme apar atunci când din rea intenţie se încearcă să se pătrundă în sistem pentru a deterimina pagube. Dar, fie că este vorba de o categorie sau alta, sistemul nu trebuie să permită astfel de „spargeri" ale securităţii sale.

Sunt multe variante prin care se poate încerca pătrunderea neautorizată. Lucrarea de faţă se ocupă în continuare de unul din acestea, şi anume falsificarea adreselor MAC. Aceste adrese, unice pentru fiecare placă de reţea, pot fi utilizate pentru autorizarea sau respingerea cererilor de acces. Dar, datorită faptului că un echipament poate falsifica adresa MAC, este posibil ca un utilizator să intre în sistem, dându-se drept o staţie autorizată. Orice vulnerabilitate este exploatată, ca de exemplu cadrele de tip „beacon" în reţelele wireless, după cum reiese şi din [Mar-08]. Literatura oferă o serie de soluţii pentru detectarea adreselor MAC falsificate, fiecare din ele având avantajele şi dezavantajele ei. În ceea ce urmează, se va prezenta o abordare originală, care completează spectrul soluţiilor oferite în domeniul detecţiei pătrunderii neautorizate într-o reţea de calculatoare.

2.2. Pătrunderi neautorizate prin falsificarea adreselor MAC.

O variantă clasică de detectare a pătrunderilor neautorizate are la bază un software de monitorizare a traficului, care verifică tot traficul şi adresele MAC folosite pentru transferul de informaţie. Orice flux determinat de o adresă MAC regăsită în tabelul cu adresele autorizate, este încadrat ca trafic autorizat. Dar, în momentul în care se sesizează prezenţa unui trafic care are la bază o adresă MAC neautorizată, software-ul de monitorizare va anunţa detectarea unui intrus şi va trece la etapa de localizare şi eliminare a intrusului.

Problema este că un intrus rău intenţionat, poate să deţină un grad ridicat de calificare în domeniul IT şi poate încerca să simuleze că ar fi o staţie autorizată, prin simpla clonare a adresei MAC a unui calculator autorizat. Fenomenul acesta, cunoscut în literatură sub numele de „*spoofed MAC addresses*" este şi azi supus cercetărilor şi diverse metode au fost propuse pentru rezolvare.

În literatură au apărut o serie de metode prin care se încearcă o cât mai bună identificare a falsificării adreselor MAC [Chu-12], [Oht-11], [Nag-10], [Gao-10], [Ara-10], [Jan-10], [Goe-09], [Ban-08], [Bra-08], [Bri-08], [Loh-08], [Wan-07], [Lig-06]. Unele dintre acestea se aplică în contexte particulare, cu ar fi metoda

prezentată în [Pua-11] pentru servere Egress NAC, dar în esenţă, ele s-ar putea împărţi în două categorii.

Prima categorie de metode extrage pachetele care circulă printr-un punct al reţelei şi urmăreşte anumiţi parametrii din cadrul acestora. Aceşti parametrii sunt verificaţi prin prisma anumitor reguli iar încadrarea, sau neîncadrarea lor în rezultatele aşteptate, reprezintă un indicator cu privire la validitatea sau falsitatea adreselor MAC.

Cea de-a doua categorie se aplică în mod particular reţelelor wireless. Algoritmii caută să determine anumite semnături pe care echipamente le păstrează în timpul funcţionării. Apariţia unui trafic care provine de la o sursă a cărei semnătură este diferită de cea aşteptată, este un indicator cu privire la falsificarea adreselor MAC. Folosind trasformări ale semnalului, se extrag caracteristicile de frecvenţă ale semnalului pentru diferite contexte date. Ulterior se compară semnalul actual cu cel de referinţă. Seriile Fourier sunt utilizate des în acest context pentru a evideţia cazurile în care apar atacatorii, văzuţi ca surse multiple. Alte metode încearcă determinarea anumitor amprente ale semnalului [Lan-12], [Jan-10], [Ara-10], [Gao-10], [Edm-09] , [Cha-09], [Loh-08], [Bra-08], [Bri-08], [Jan-08], [She-08].

În cele ce urmează, se realizează o formalizare a detecţiei pătrunderilor neautorizate prin falsificarea adreselor MAC şi se vor considera câteva din cele mai importante metode, care se adresează concret aspectelor prezentate anterior. Metodele au fost alese astfel încât să confere o imagine de ansamblu asupra abordării actuale a subiectului.

2.3. Abordarea sistemică a detecţiei adreselor MAC falsificate.

Abordarea sistemică presupune formalizarea problematicii studiate, în contextul dependenţei cauză-efect. În acest context, se definesc în primul rând datele problemei:

- se dispune de o reţea de calculatoare;
- accesul în reţea nu este permis decât unor staţii autorizate;
- se cere o monitorizare permanentă a traficului din reţea, pentru detectarea rapidă a încercărilor de pătrundere neautorizată.

Pentru a detecta pătrunderile neautorizate, este necesar ca în rețea să existe un modul de monitorizare, care să semnaleze eventualele atacuri. În Fig. 2.1 este reprezentată structura unui astfel de modul.

Traficul este captat şi trimis ca intrare către un subansamblu de validare a adreselor MAC identificate în trafic. Validarea însă, necesită o bază de date de unde să se obţină informaţii relevante în procesul de validare.

O implementare simplă a unei astfel de validări, se poate face pe baza unui tabel cu adresele MAC autorizate. Orice adresă care nu se regăseşte în baza de date, este imediat semnalată la exterior. Algoritmul este simplu, rapid, dar din păcate nu este eficient, tocmai datorită posibilităţii de falsificare a adreselor MAC.

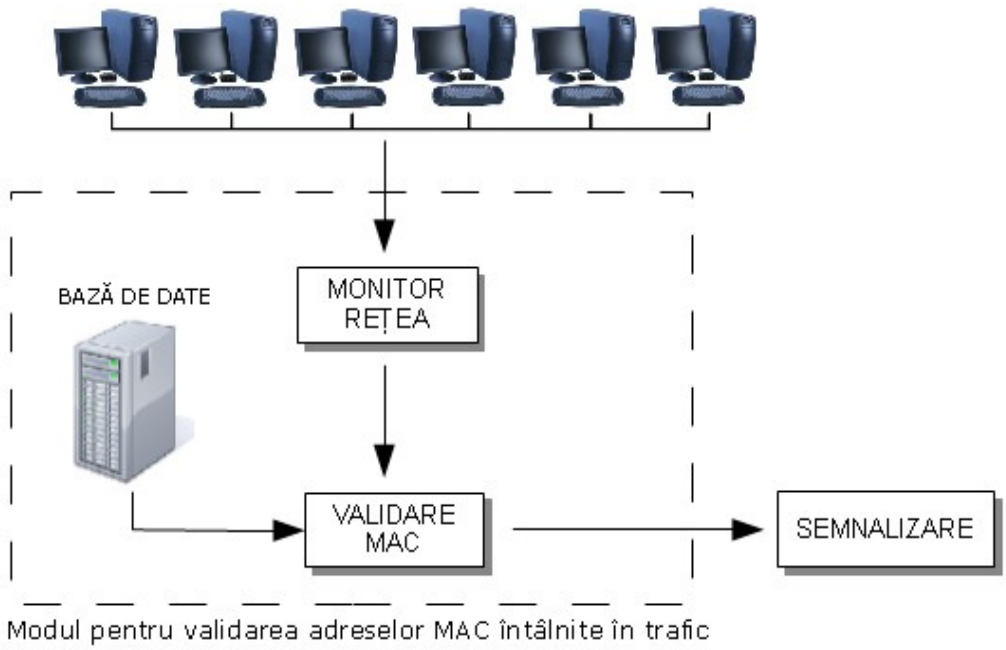


Fig. 2.1 Detectarea pătrunderilor neautorizate într-o reţea de calculatoare

Structura reprezentată în Fig. 2.1 se poate transpune într-o schemă bloc sistemică, reprezentată la rândul ei în Fig. 2.2:

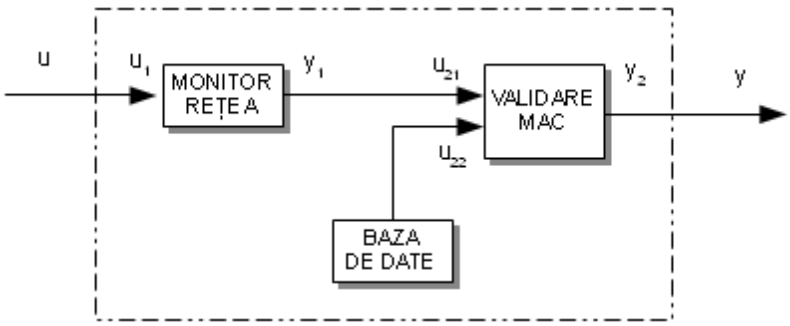


Fig. 2.2 Schemă bloc pentru validarea adreselor MAC

Modulul de validare poate fi descris prin intermediul unei „cutii negre”, având intrarea „u(t)” şi ieşirea „y(t)” (Fig. 2.2).

Ca *intrare* se va considera lista pachetelor captate în reţea la un moment de timp „t”. Astfel, intrarea poate fi descrisă printr-o relaţie de forma:

$$u(t) = M_P\{P_k | k - variabil\ ,\ functie\ de\ t\} \quad (1)$$

unde „M_P” reprezintă mulţimea tuturor pachetelor captate în reţea la momentul de timp „t”.

Ieşirea sistemului la momentul de timp „t”, este o listă cu toate adresele MAC identificate în trafic, împreună cu rezultatul validării, dat fie de forma „adevărat” / „fals”, fie sub forma unui grad de recunoaştere, ca procent între 0% – 100%. Pentru un nivel mai mare de generalitate, se va considera identificarea pe baza gradului de recunoaştere. În aceste condiţii, adresa MAC_k, va avea asociat gradul de recunoaştere G_k, alcătuind astfel o pereche $P_{MAC}(MAC_k, G_k)$.

Ieşirea sistemului se poate atunci descrie matematic după cum urmează:

$$y(t) = M_{MAC}\{P_{MAC}(MAC_k, G_k) | G_k \in [0,100]\} \quad (2)$$

Dacă modulul de validare se defineşte foarte strict, atunci G_k va lua doar valorile 0 şi 100. Însă, aşa după cum s-a afirmat deja, nu toate metodele vor fi capabile de un răspuns atât de clar.

În această abordare, sistemul poate fi descris printr-o relaţie de forma:

$$f : M_P \rightarrow M_{MAC} \quad (3)$$

care permite ca pe baza mulţimii de intrare M_P, să se genereze mulţimea M_{MAC}, cât şi intrările şi ieşirile intermediare notate y_1, u_{21}, u_{22}, care depind de implementarea concretă a modulului de validare.

2.4. Validarea adreselor MAC prin analizarea unor parametrii din cadrul pachetelor.

În cadrul paragrafului sunt analizate câteva metode care identifică adresele MAC falsificate urmărind evoluţia unor parametrii care sunt prelevaţi din cadrul pachetelor de reţea.

2.4.1 Validarea adreselor MAC prin verificarea ID-ului de VLAN.

În [Ish-01], autorii dezvoltă o metodă proprie de control al accesului în reţea, metodă care este utilă chiar şi atunci când utilizatorii falsifică adresele IP sau MAC. Domeniul de aplicabilitate are în vedere reţelele publice, unde oricine poate intra cu computerul personal şi primeşte o adresă IP alocată dinamic.

Pentru validarea adreselor MAC, autorii folosesc un detaliu legat de switch-urile care permit crearea de reţele virtuale (VLAN), şi anume ID-ul de VLAN pe care switch-ul îl adaugă în frame-uri pentru identificare. Aceste ID-uri nu pot fi controlate de către utilizatori, şi în consecinţă pot fi luate ca element de siguranţă în autentificare.

Pentru implementarea metodei, autorii au dezvoltat un sistem numit LANA, prezentat în Fig. 2.3 şi care conectează împreună:

- un server de autentificare RADIUS (RFC 21380;
- un server pentru alocarea adreselor IP (server DHCP);
- un filtru de frame-uri;
- un număr oarecare de switch-uri dotate cu facilităţi de „VLAN tagging”;
- un server care comunică cu serverul de autentificare şi controlează filtrul şi switch-urile.

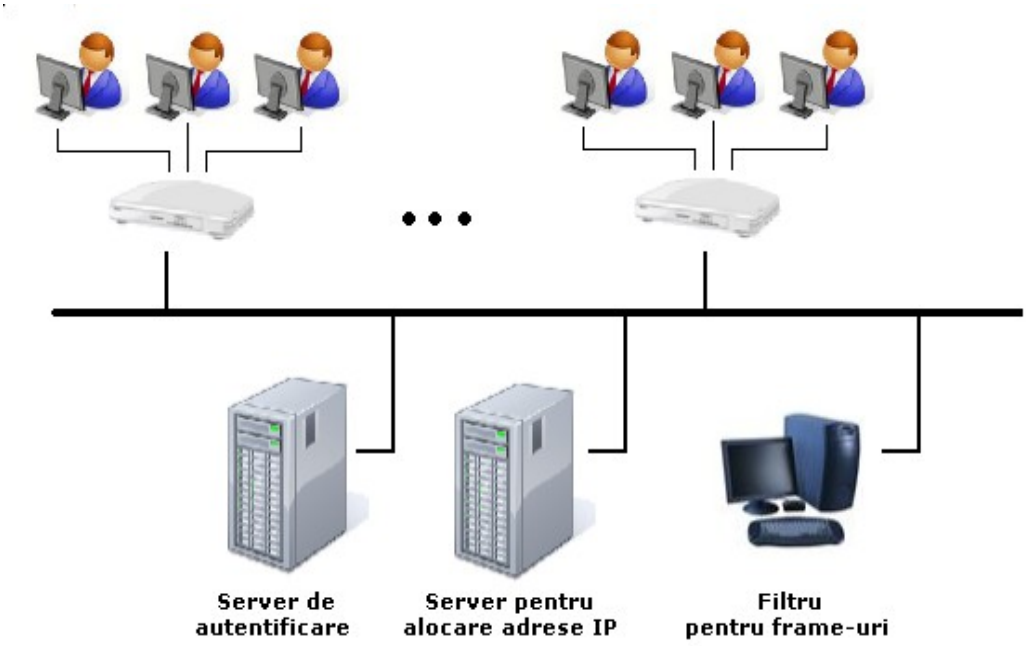


Fig. 2.3 LANA – Sistem pentru transfer bazat pe autentificare

Filtrul permite trecerea doar a frame-urilor care respectă regulile stabilite pe baza adreselor MAC ale sursei şi destinaţiei, adreselor de IP ale surselor şi a destinaţiilor, porturile TCP/UDP ale sursei şi destinaţiei şi în final identificatorul de VLAN. Aceste înregistrări permit păstrarea unui registru cu privire la activitatea utilizatorilor.

Când un client se conectează în reţea, el comunică direct cu serverul de DHCP pentru a i se aloca o adresă IP. Filtrul monitorizează dialogul şi află adresa care i se alocă. Transmite mai departe perechea (adresă IP, adresă MAC) către serverul din reţea. Apoi, serverul va încerca să se conecteze pe o aplicaţie care ar trebui să fie instalată pe computerul clientului. Dacă această aplicaţie lipseşte, atunci autentificarea este necesară prin alte modalităţi (cum ar fi acces prin browser). Indiferent de modul de autentificare, după ce aceasta a avut loc, serveul va comunica filtrului o pereche (adresă MAC, adresă IP, ID VLAN) pentru staţia

autentificată. Filtrul va permite atunci să treacă toate pachetele care conţin această pereche în interiorul mesajelor.

În urma experimentelor realizate, autorii au dovedit faptul că filtrarea pachetelor necesită o putere de calcul mai mare şi caracteristica „network throughput” are de suferit cu proximativ 33%. Pentru testare au folosit utilitarul „Netperf” [NetP].

Concluzii:

Metoda are avantajul faptului că fructifică o idee simplă, legată de identificatorul de VLAN „purtat” de către frame-uri. Acest identificator nu poate fi falsificat de către un atacator, şi de aceea folosirea lui permite o bună identificare şi validare.

Totuşi, metoda descrisă de autori are unele dezavantaje. În primul rând trebuie amintit faptul că este necesară instalarea unei aplicaţii software pe calculatoarele client, pentru autentificarea staţiei. Dacă această aplicaţie lipseşte, situaţia devine şi mai dificilă, întrucât se apelează la o metodă „by hand” cu un browser web.

Un al doilea dezavantaj apare atunci când adresa MAC falsificată apare în interiorul aceleiaşi reţele locale, adică vor folosi acelaşi identificator. Se pot face şi alte verificări, cum ar fi adresa de IP, dar, în sine, ideea de bază a folosirii identificatorului de VLAN nu are sens decât în reţele locale diferite.

O a treia problemă se poate întâlni în cazul unei reţele wireless, în care există o serie de Access Point-uri. Dacă la un moment dat, staţia mobilă trece din zona de acoperire a unui Access Point şi intră în zona deservită de altul, sistemul va recepţiona o cerere de conectare într-o altă subreţea, folosind un alt identificator de VLAN. Dacă aceste mesaje sunt recepţionate chiar înainte de finalizarea procesului de handover, s-ar putea ca pentru un scurt timp, sistemul să considere că a apărut un intrus, ceea ce de fapt nu este adevărat.

În final, metoda devine complicată, necesitând o colaborare continuuă între serverele de autentificare şi filtru. Chiar autorii afirmă că reţeaua este afectată negativ din punct de vedere al caracteristicii „throughput”, cu aproximativ o treime, ceea ce practic este mult.

2.4.2 Validarea adreselor MAC prin verificarea numărului de secvenţă şi a parametrului „interarrival time”.

În [Liq-06] s-a conceput o metodă de determinare a traficului falsificat, pe baza stabilirii unor reguli sau relaţii după care traficul este declarat ca normal sau falsificat, într-o modelare matematică adecvată.

Primele reguli stabilite se referă la numerele de secvenţă, numere care trebuie să urmărească o progresie liniară, crescătoare. Acest număr de secvenţă este compus din 12 biţi şi este adăugat în header de către echipamentul de transmisie. La fiecare pachet trimis, numărul de secvenţă se incrementează cu 1 până când ajunge la 4095, după care se resetează şi porneşte din nou de la 0. Chiar dacă anumite pachete sunt pierdute, iar numerele lor de secvenţă vor lipsi, traficul poate fi în continuare recunoscut ca „normal”. Dacă două sau mai multe emiţătoare folosesc aceeaşi adresă MAC, receptorul va putea sesiza faptul că numerele de secvenţă nu sunt liniar crescătoare ci variază puternic. Chiar dacă atacatorul poate citi numerele de secvenţă generate de către staţia adevărată şi îşi poate ajusta numerele proprii de secvenţă, receptorul va sesiza duplicate în lista numerelor de

secvenţă şi va interpreta situaţia tot ca anormală, Folosind un calcul matematic probabilistic, autorii au dovedit că aceste situaţii pot fi separate în mod evident şi se pot trage concluzii valide referitoare la starea traficului. Acelaşi mod de abordare este prezentat şi în [Wri-03], [Ban-08], [Goe-09]

A doua regulă folosită în detecţie se referă la parametrul „interarrival time" pentru pachete. Acest parametru se referă la timpul scurs între două pachete care provin de la aceeaşi sursă. În mod normal, o sursă va păstra aceleaşi caracterisitici de transmisie, aceeaşi distribuţie, care poate fi măsurată empiric. Dacă există două sau mai multe surse care se dau drept aceeaşi sursă, distribuţia înregistrată va fi diferită. Pe baza unor calcule matematice, se poate stabili gradul de asemănare între cele două distribuţii.

Ambele reguli sunt numite „reguli binare" în sensul că oferă doar două rezultate posibile referitoare la atacuri: prezent / absent. Pentru a măsura gradul de severitate al unui atac, nu se mai pot folosi variantele anterioare, ci trebuie utilizată o clasificare „multi-nivel".

Concluzii:

Metoda este eficientă, autorii demonstrând performanţele ei. Totuşi se poate menţiona faptul că parametrul „interarrival time" poate suferi modificări destul de semnificative într-o reţea wireless, datorită mediului de propagare şi chiar şi în reţelele cu fir, congestia va modifica valoarea parametrului. În aceste situaţii, sistemul de monitorizare şi validare a adreselor MAC ar putea semnala în mod eronat traficul care provine de la o sursă autorizată.

De asemenea, numerele de secvenţă se pot repeta chiar frecvent într-o reţea wireless perturbată de factori externi, cum ar fi vremea, factori care pot duce la pierderea unui mare număr de pachete şi implicit la retransmiterea lor. Sistemul poate identifica aceste cazuri ca fiind mai multe surse, ceea ce nu este adevărat.

Problema poate cea mai mare în cazul verificării numărului de secvenţă, este aceea că numărul de secvenţă poate fi folosit ca indicator numai dacă staţia autorizată funcţionează în acelaşi timp cu staţia intrus. Altfel, numerele de secvenţă generate de staţia intrus vor avea o distribuţie normală, şi în felul acesta intrusul va trece drept o staţie autorizată.

2.4.3 Validarea adreselor MAC prin verificarea parametrului „hop-count".

[Wan-07] dezvoltă o altă metodă de filtrare a traficului falsificat, pe baza unui contor numit „hop-count", actualizat de către rutere la fiecare trecere a pachetului printr-un ruter. Este foarte dificil, dacă nu chiar imposibil, ca un atacator să reuşească să injecteze un trafic falsificat, dar care să păstreze corect valorile acestui indicator. Mai mult, autorii au identificat faptul că se poate realiza o corespondenţă între hop-count şi câmpul TTL (Time-To-Live) din headerul IP. Pe baza acestor observaţii, autorii au dezvoltat o metodă numită HCF (Hop-Count-Filtering) care detectează traficul falsificat.

Sistemul conceput, reprezentat intuitiv în Fig. 2.4, are două moduri de lucru. Primul este modul de „învăţare", care nu elimină pachetele ci doar urmăreşte activitatea din reţea. Doar în momentul în care se detectează pachete falsificate, sistemul trece automat în modul de lucru „filtrare", mod care va elimina toate pachetele care prezintă contorul „hop-count" cu valoare greşită.

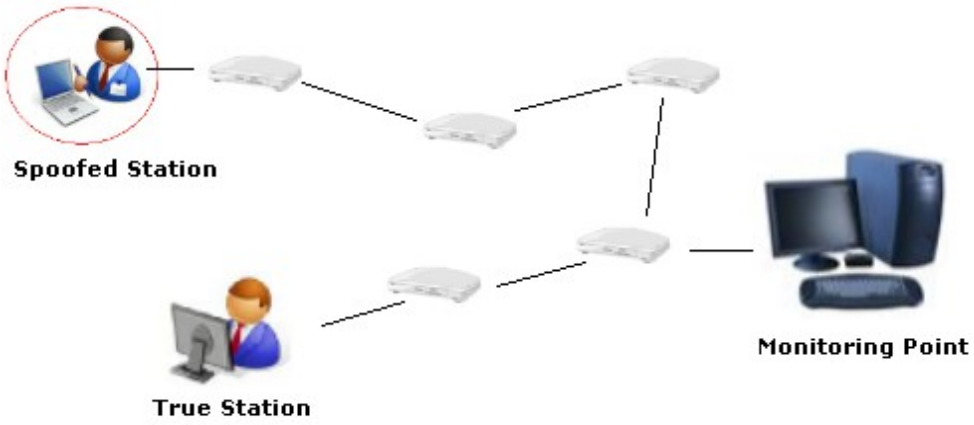


Fig. 2.4 Exemplu de structură de reţea în vederea determinării valorii „hop-count”

Prima problemă pe care autorii au trebuit să o rezolve a fost tocmai determinarea contorului „hop-count”. Acest contor de fapt nu se înregistrează ca o valoare dată în interiorul câmpurilor unui pachet, ci doar s-ar putea deduce din calcule asupra valorii TTL. Această valoare se stabileşte iniţial la trimiterea pachetului, şi fiecare ruter decrementează valoarea. Din păcate, sistemele de operare nu folosesc aceeaşi valoare iniţială, astfel că determinarea contorului „hop-count” poate fi numai aproximativă. Argumentele pe care autorii le aduc vin să dovedească faptul că „distanţele” în Internet de obicei sunt sub 30 de „hop”-uri şi datorită acestui aspect, estimează valoarea iniţială a valorii TTL ca fiind cea mai mică valoare standard TTL, care totuşi este mai mare decât valoarea TTL măsurată la receptor. După stabilirea valorii iniţiale se calculează contorul „hop-count” ca diferenţă dintre valoarea iniţială şi cea calculată la receptor.

Pe baza înregistrărilor efectuate, se poate păstra un tabel care asociază adresele de IP cu valorile „hop-count” corespunzătoare. Pentru ca atacatorii să nu poată compromite acest tabel, el se actualizează numai pe baza mesajelor „tree-way handshake” din cadrul conexiunilor TCP. Analizând traficul interceptat, se extrage adresa IP şi valoarea TTL din fiecare packet şi se calculează contorul „hop-count”, care mai apoi este comparat cu indicatorul disponibil în tabel. Dacă valoarea calculată diferă de cea înregistrată, pachetul este declarat ca falsificat.

Valorile TTL sunt folosite şi în alte abordări cu scopul de a determina traficul falsificat, cum ar fi de exemplu [Oht-11] şi [Moh-10].

Concluzii:

Metoda este eficientă, totuşi, rămâne problema dată de necunoaşterea exactă a valorii TTL. Aproximarea ei poate duce la semnale false, mai ales atunci când vorbim de o staţie mobilă, care se deplasează şi schimbă Access Point-ul.

2.5. Validarea adreselor MAC prin analizarea unor amprente ale echipamentelor wireless.

Paragraful consideră o altă categorie de metode de detecţie a adreselor MAC falsificate, categorie care se aplică strict echipamentelor wireless, şi care încearcă să determine identitatea echipamentului, pe baza unei amprente. În literatură, subiectul este tratat printr-o serie de articole, cum ar fi [Lan-12], [Jan-10], [Ara-10], [Gao-10], [Edm-09] , [Cha-09], [Loh-08], [Bra-08], [Bri-08], [Jan-08], [She-08]. Se vor prezenta modelele de amprentare wireless, analizându-se comparativ avantajele şi dezavantajele pe care le conferă.

2.5.1 Validarea adreselor MAC prin verificarea răspunsului echipamentelor wireless la mesaje deformate.

În [Bra-08] se demonstrează că se poate extrage o amprentă a echipamentelor wireless, amprentă care poate fi utilizată ulterior în a determina eventualele încercări de falsificare a lor. Această amprentă este extrasă folosind răspunsul pe care echipamentele îl dau la diverse mesaje deformate (alterate). La baza metodei stă observaţia că diferite implementări ale standardului 802.11b/g [STD-802.11] vor reacţiona diferit la evenimente care deviază de la formatul standard.

Testarea comportamentului echipamentelor în cazuri „deviate de la normal” implică de exemplu:

- setarea sau resetarea unor biţi în cotexte în care aceştia se aşteaptă să aibă o anumită valoare dată;
- cereri care nu ar trebui să fie fragmentate şi care totuşi sunt trimise în felul acesta;
- frame-uri care ar trebui să conţină anumite informaţii, dar acestea lipsesc;
- frame-uri care nu ar trebui să conţină anumite câmpuri, dar care tutuşi există;

Sistemul implementat de către autori utilizează două platforme, una „de scanare”, care se ocupă de trimiterea frame-urilor folosite în teste şi o platformă de monitorizare a răspunsurilor.

Concluzii:

Metoda permite caracterizarea unor echipamente wireless, nu neapărat a unor staţii din reţea, identice, şi din care numai una este autorizată. Din acest motiv, nu se poate aplica în cazul detectării adreselor MAC falsificate, mai ales dacă adresa MAC adevărată şi cea falsă au apărut în aceeaşi reţea locală.

2.5.2 Validarea adreselor MAC prin verificarea distribuţiei frame-urilor de tip „probe request”.

Autorii din [Loh-08] îşi aduc propria contribuţie la subiectul discutat, pintr-o metodă care crează o amprentă de identificare a echipamentelor wireless.

Identificarea se face pe baza unui grup de informaţii: echipament wireless, driver-ul plăcii de reţea wireless şi sistemul de operare. Metoda se foloseşte de o analiză a frame-urilor de tip „probe request”. Aceste frame-uri sunt folosite de către staţii pentru a se conecta la un Acces Point. În mod normal, acestea ar trebui să respecte o anumită ciclicitate. Funcţie de driverul interfeţei de reţea, intervalele acestea sunt diferite. Dar, studiile autorilor au demonstrat faptul că intervalele de timp sunt influenţate şi de sistemul de operare şi de maşina pe care rulează.

Prima fază este cea de captare a traficului. În această fază sunt colectate frame-urile de tip „probe request” emise de către staţiile client. A doua fază va genera amprenta folosind datele înregistrate. În final, pe baza amprentelor generate se verifică traficul.

Concluzii:

Metoda are două dezavantaje principale. Primul este dat de faptul că poate fi aplicată numai în reţelele wireless, iar al doilea constă în aceea că, dacă există două echipamente fabricate identic şi pe care s-a instalat aceeaşi versiune a sistemului de operare, amprenta lor va fi identică, ceea ce face imposibilă deosebirea lor în traficul captat din reţea.

2.5.3 Validarea adreselor MAC prin verificarea intensităţii semnalului recepţionat.

În [She-08] se dezvoltă o metodă de detecţie a pachetelor falsificate, bazată pe nivelul de semnal recepţionat. Deşi s-a considerat pentru mult timp că distribuţiile puterii semnalului receţionat (RSS) sunt de tip Gaussian, autorii demonstrează că acest lucru nu este totdeauna adevărat. În plus, ţinând cont de caracteristica „antena diversity” prin care un emiţător dispune de 2 sau chiar mai multe antene, autorii au stabilit o ipoteză prin care mostrele RSS extrase pentru o pereche de antene, din care una este pentru transmisie iar cealaltă este pentru recepţie, luate împreună, urmăresc o distribuţie Gaussiană. Caracterizarea tiparelor RSS este realizată cu ajutorul modelelor GMM (Gaussian Mixture Models), care sunt de fapt o combinaţie ponderată a mai multor distribuţii Gaussiene. Autorii realizează un profil GMM pentru fiecare pereche de antene ce aparţin unui emiţător. Acest profil poate fi actualizat periodic, colectând mostrele RSS.

În [Cha-09] se prezintă o altă metodă care se aplică tot domeniului wireless. Ideea care stă la baza algoritmului conceput de autori se referă tot la măsurarea nivelului de semnal RSSI (Received Signal Strength Indicator), dar combinat şi cu alţi factori. Acest nivel de semnal al pachetelor transferate prin mediul wireless poate fi măsurat în mai multe puncte, şi pe baza măsurătorilor, se poate alcătui o amprentă de trafic, care ţine cont de următorul grup de informaţii: adresa MAC a emiţătorului, numărul de secvenţă MAC, tipul pachetului şi momentul recepţiei. Pentru realizarea unei amprente „per pachet” este necesar ca primele trei informaţii menţionate mai sus să fie identice, iar momentele recepţiei să fie „suficient de apropiate”. Aplicând tehnici specifice, se poate obţine o localizare a surselor care au trimis pachetele. Urmărind distanţa euclidiană dintre pachetele care se presupune că aparţin aceleaşi surse, ar trebui să găsite diferenţe mari pentru situaţia în care există surse multiple (care se dau drept aceeaşi sursă). Dacă însă atacatorul este localizat la o distanţă foarte mică de victimă, metoda nu va da randament întrucât va considera cele două surse ca fiind de fapt una singură.

Modul de detecţie parcurge mai multe verificări. Este posibil ca identificarea să se realizeze în anumite cazuri chiar în primele faze, dar pentru alte situaţii, este necesară parcurgerea întregului algoritm.

Algoritmul propus de autori respectă următorii *paşi*:

- se extrag numerele de secvenţă din pachete şi se caută să se observe dacă acestea au o progresie lineară sau nu;
- se verifică dacă numerele de secvenţă cresc în mod linear, caz în care se trage concluzia că nu este vorba de un atac. În caz contrar, chiar dacă numerele de secvenţă nu sunt absolut linear crescătoare, totuşi se încearcă să se observe dacă nu cumva diferenţele se încadrează în limitele stabilite. Aceste limite au în vedere cazurile de retransmisii. Dacă testul dovedeşte încadrearea în limite, algoritmul realizează absenţa atacului;
- în caz contrar, următorul pas verifică tipul pachetelor înregistrate. Pentru pachete de tip „management" sau „regular data frames", numerele de secvnenţă trebuie sîă respecte creşterea lineară. În ca contrar, algoritmul identifică prezenţa unui atac;
- dacă pachetele nu sunt „management" sau „regular data", algoritmul ia în calcul „QoS priorities". Dacă acestea sunt identice, este vorba de un atac;
- dacă nu sunt identice, următorul pas aplică algoritmi de localizare pentru pachetele interceptate şi calculează distanţa Euclidiană. Dacă această distanţă depăşeşte limitele stabilite, algoritmul trage concluzia că este vorba de un atac. În caz contrar, este vorba de un trafic normal.

Concluzii:

Metoda nu va detecta situaţiile în care atacatorul se află aproape de staţia adevărată. În plus, deplasarea rapidă a staţiilor mobile s-ar putea să ducă la semnalizări greşite. Sau, dacă există două echipamente fabricate identic, ele vor produce aceeaşi amprentă.

2.5.4 Validarea adreselor MAC prin verificarea parametrului „clock skew".

În [Lan-12], [Ara-10], [Jan-08] şi [Jan-10] se prezintă o altă posibilitate de detecţie a unor Access Points (AP) falsificate, pe baza parametrului de timp numit „clock skew". Autorii au observat faptul că acest parametru este constant în timp pentru un Access Point, dar variază foarte mult de la un Access Point la altul. Deşi alte abordări care folosesc acest parametru îl calculează din marcajele de timp din TCP/ICMP, [Jan-08] s-au folosit marcajele de timp din cadrul TSF (Time Synchronization Function) din mesajele „beacon / probe response". Motivaţia acestei alegeri se referă în primul rând la faptul că aceste frame-uri au o rată mare de transmisie. Pe de altă parte, timpul măsurat prin funcţia TSF împarte timpul până la nivelul unei microsecunde. Mai mult, timpul memorat în TSF reprezintă efectiv timpul în care Access Point-ul a trimis frame-ul, spre deosebire de timpul „programat" pentru transmisie, care poate suferi întârzieri.

Calculul efectiv al „clock skew" este realizat prin două metode: LPM (Linear Programming Method) şi LSF (Least Square Fitting).

Prima metodă caută să determine o linie care se află deasupra tuturor punctelor care reprezintă offset-urile de timp dintre momentele de transmisie şi cele de recepţie. Parametrul „clock skew" este de fapt panta acestei drepte.

A doua metodă estimează parametrul tot ca pantă a unei drepte, numai că de data aceasta dreapta este stabilită diferit. Având în vedere că ambele metode sunt aplicabile numai dacă Access Point-ul adevărat şi cel fals funcţionează la momente de timp diferite, autorii au dezvoltat şi o metodă euristică prin care să diferenţieze traficul atunci când cele două funcţionează simultan.

Concluzii:

Metoda nu poate fi aplicată decât Access Point-urilor, deci nu calculatoarelor care accesează o reţea deservită de un Access Point. Aspectul acesta reprezintă o limitare în sensul că un intrus nu va fi detectat decât dacă încearcă să conecteze în reţea un echipament de tip Access Point.

2.5.5 Validarea adreselor MAC prin verificarea „Power Hopping".

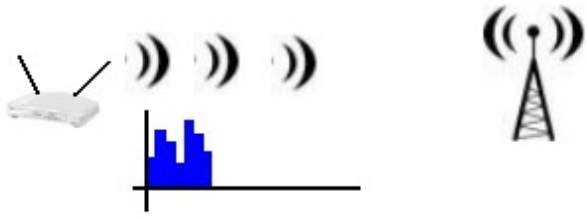


Fig. 2.5 Transmiterea pachetelor folosind puteri diferite ale semnalului

În [Nag-10] s-a propus o nouă metodă pentru detectarea traficului falsificat, ca în Fig. 2.5. Ideea care stă la baza metodei este aceea de a implementa in Access Point un algoritm prin care să accepte pachete numai daca se încadrează într-un anumit nivel de putere a semnalului.

Întrucât puterea semnalului variază în timp, s-a definit o mulţine P, compusă din totalitatea nivelelor de putere acceptate, cu menţiunea că toate sunt mai mari decât o putere minimă stabilită. Pe baza unei secvenţe de numere pseudoaleatoare, generate cu ajutorul unei singure valori „de amestecare", se pot alege o listă de puteri, din mulţimea P. Metoda este implementată în doi paşi: iniţializare şi „power hop mode".

Faza de iniţializare este reprezentată de o suită de mesaje între Access Point şi staţia locală:

- AP trimite periodic prin broadcast un „beacon", care conţine nivelul de semnal folosit la transmisie.
- Staţia recepţionează beacon-ul şi înregistrează atât nivelul de semnal calculat la recepţie cât şi cel înregistrat în beacon.
- Staţia trimite o cerere de conexiune către AP.
- După primirea cererii, AP-ul iniţiază o conexiune HTTPS cu staţia.
- Staţia se autentifică pe baza unei parole şi comunică prin HTTPS.
- Staţia generează o „valoare de amestecare" (seed). Pe baza ei se aleg valori din mulţime P, valori ce reprezintă puteri cu care AP-ul ar trebui să recepţioneze pachetele.
- Valorile sunt recepţionate de AP şi trimite înapoi un ACK.

A doua fază este cea de filtrare a pachetelor. Pe baza datelor stabilite în faza de iniţializare, AP-ul verifică nivelul de semnal primit de la staţie în fiecare pachet şi îl compară cu nivelul aşteptat. Dacă nivelele corespund, se înţelege că este vorba de staţia adevărată. Dacă nivelele nu corespund, se înţelege că este vorba de un atac.

Concluzii:

Metoda este complexă, având nevoie de implementări la nivelul protocolului de comunicare wireless. Atât Access Point-urile cât şi staţiile trebuie să fie echipate corespunzător pentru a permite metodei să funcţioneze.

Puterea semnalului poate fi afectată de mediul de transmisie, ceea ce reprezintă un neajuns al metodei.

2.6. O analiză critică a metodelor prezentate, prin prisma răspunsului acestora la anumite situaţii specifice.

Tabelul 4 prezintă o analiză comparativă a metodelor discutate anterior, din punct de vedere al modului în care aceste metode pot sau nu pot fi folosite atunci când în practică apar anumite situaţii particulare. Fiecare metodă are performanţe evidenţiate printr-o serie de teste şi verificări, însă ceea ce interesează în lucrarea de faţă este să se analizeze cât de extins sau cât de restrâns este domeniul de aplicabilitate al acestor metode, ţinând cont de faptul că în practică se întâlnesc o serie de situaţii specifice. Tabelul îşi propune să evidenţieze cele mai importante situaţii particulare, şi modul în care metodele descrise anterior, pot gestiona astfel de cazuri.

2.6.1 Detectarea unui intrus care vine în reţeaua locală, chiar în locul staţiei autorizate.

Aspectul acesta are o importanţă deosebit de mare, deoarece de cele mai multe ori se încearcă securizarea reţelelor faţă de exterior, şi nu se pune un accent foarte mare pe posibilitatea ca atacatorul să înceapă să lucreze prin conectare directă în reţeaua locală.

Mai mult, atacatorul s-ar putea să se conecteze chiar în locul staţiei autorizate. Pentru a-şi ascunde cât mai bine identitatea, dacă atacatorul are acces la staţia autorizată, atunci ar putea copia adresa MAC, după care să deconecteze din reţea staţia autorizată şi să conecteze un calculator personal.

Cât de bine sunt pregătite metodele de detecţie a adreselor MAC falsificate, ca să facă faţă unor astfel de situaţii? Ca să se poată răspunde la această întrebare, trebuie luat pe rând fiecare caz în parte. Tabelul 4 sugerează faptul că din cele nouă metode prezentate, cinci sigur nu vor detecta intrusul, iar celelalte patru îl vor detecta numai dacă foloseşte un dispozitiv diferit de cel autorizat.

Detecţia prin ID-ul de VLAN încearcă să urmărească tentativele de pătrundere prin falsificarea adreselor MAC, urmărind ID-ul de VLAN alocat de către switch. Dacă atacatorul este în aceaşi reţea, va avea acelaşi ID şi astfel va trece drept staţie autorizată.

Urmărirea numerelor de secvenţă va conduce la acelaşi rezultat. Faptul că atacatorul vine în locul staţiei autorizate, va determina ca numerele de secvenţă să aibă o distribuţie normală.

Verificarea parametrului „interarrival time" se bazează foarte mult pe echipamentele folosite. În aceste condiţii, dacă atacatorul dispune de un calculator identic cu cel autorizat, atunci calculul „interarrival time" va da rezultate identice cu cele ale staţiei autorizate. Totuşi, dacă atacatorul nu dispune de un calculator de acelaşi tip, detecţia prin „interarrival time" va semnala corespunzător intrusul, întrucât parametrul va fi cu totul diferit de cel al staţiei autorizate.

Calculul şi verificarea contorului „hop-count" nu permite deloc detectarea intrusului. El face parte din reţeaua locală, deci numărul de „salturi" până la punctul de monitorizare rămâne acelaşi.

Din cele cinci metode de amprentare wireless, două dintre ele vor ignora complet apariţia intrusului. Este vorba despre detecţia prin măsurarea nivelului de semnal şi „Power Hopping". Cele două metode identifică o staţie după intensitatea cu care se recepţionează semnalul într-un punct. Faptul că intrusul vine în locul staţiei autorizate, înseamnă că nivelul de semnal va corespunde cu cel al staţiei autorizate. Celelalte trei metode vor semnala corect intrusul, dacă acesta foloseşte un echipament diferit. În caz contrar, intrusul va trece drept staţie autorizată. Rezultatul este eronat deoarece răspunsul la mesaje deformate, parametrul „clock-skew" şi „Power Hopping" vor fi identice cu cele ale staţiei autorizate.

În *concluzie*, apariţia unui intrus în reţeaua locală, în locul staţiei autorizate, reprezintă o reală problemă, care este puţin acoperită prin metodele folosite în prezent. Aspectul acesta conduce la necesitatea dezvoltării unor metode care să permită identificarea unui intrus care foloseşte adresa MAC a unei staţii autorizate, chiar şi atunci când reuşeşte să intre în reţeaua locală, în locul staţiei autorizate.

2.6.2 Limitarea metodelor la reţele wired sau wireless.

Un alt aspect important care trebuie luat în discuţie atunci când sew consideră în ansamblu metodele analizate, este cel referitor la aplicabilitatea metodelor în orice tip de reţea, sau numai în reţele de un anumit tip (wired sau wireless).

Din Tabelul 4 reiese faptul că cele cinci metode de amprentare wireless nu pot fi aplicate decât în domeniul wireless. Aceasta înseamnă că nu vor putea fi utilizate în contextul reţelelor cu fir.

Metodele care folosesc numerele de secvenţă, „interarrival time" şi „hop-count", vor putea fi aplicate indiferent de tipul reţelei, ceea ce automat le face să fie general valabile, crescând astfel domeniul de aplicabilitate.

Metoda care ia în calcul ID-ul de VLAN este limitată la reţele care folosesc switch-uri cu VLAN.

În *concluzie*, o serie de metode sunt aplicabile doar în reţelele wireless, ceea ce reduce mult domeniul de utilizare. Totuşi ele au o importanţă reală, în contextul respectiv, dar pentru o detecţie rapidă şi eficientă a încercârilor de pătrundere prin falsificarea adreselor MAC, sunt necesare metode care să nu depindă de tipul reţelei, şi să poată fi folosite chiar in reţele care conţin atât comunicare wireless cât şi comunicare wired.

Tabelul 4 – Răspunsul algoritmilor la probleme concrete

Caracteristica urmărită	*Verificare VLAN ID*	*Verificare numere secvenţă*	*Verificare interarrival time*	*Verificare „hop-count”*	*Amprentare wireless prin răspuns la mesaje deformate*	*Amprentare wireless prin distribuţia „probe-request”*	*Amprentare wireless prin nivelul semnalului recepţionat*	*Amprentare wireless prin parametrul clock-skew*	*Amprentare wireless prin „Power Hopping”*
Detectează un intrus care vine în reţeaua locală, în locul staţiei autorizate.	NU	NU	LIMITAT	NU	LIMITAT	LIMITAT	NU	LIMITAT	NU
Suportă orice tip de reţea.	LIMITAT	DA	DA	DA	NU	NU	NU	NU	NU
Necesită instalarea unui software pe staţia client.	DA	NU	NU	NU	NU	NU	NU	NU	DA
Aplicabilă în absenţa suprapunerii traficului de la staţia autorizată cu cel provenit de la staţia intrus.	NU	NU	LIMITAT	LIMITAT	LIMITAT	LIMITAT	LIMITAT	LIMITAT	LIMITAT
Permite staţiei client să-şi schimbe localizarea geografică şi să treacă în altă subreţea.	NU	?	LIMITAT	NU	-	-	-	-	-
Aplicabilă chiar dacă intrusul foloseşte acelaşi tip de echipament.	LIMITAT	LIMITAT	LIMITAT	LIMITAT	NU	NU	LIMITAT (dacă poziţionarea intrusului este diferită)	NU	DA
Aplicabilă pentru computere.	DA	DA	DA	DA	NU	NU	NU	NU	NU

2.6.3 Necesitatea instalării unui software pe staţia client.

O soluţie convenabilă atunci când se doreşte să creşterea gradului de recunoaştere a unei adrese MAC, este dată de instalarea unui software pe fiecare staţie client, software care să fie responsabil de autentificarea şi validarea staţiei în reţea. Dar, abordarea aceasta nu este de dorit, în sensul că necesită o serie de operaţii pe fiecare staţie client. Mai mult, nu în orice situaţie există permisiunea de a instala programe pe staţiile care trebuie să se conecteze în reţea.

Din metodele discutate, doar două intră în categoria celor care necesită software adiţional, şi anume verificarea ID-ului de VLAN şi amprentarea prin „Power Hopping".

În *concluzie*, se caută evitarea instalărilor de software adiţional, şi se preferă metodele care nu necesită un astfel de demers.

2.6.4 Aplicabilitate în absenţa suprapunerii traficului provenit de la sursa autorizată cu cel provenit de la sursa neautorizată.

Identificarea unui intrus este abordată în anumite situaţii prin observarea unei devieri în trafic faţă de comportamentul aşteptat, al unei staţii care transmite date în reţea. Metodele de acest tip depind în totalitate de suprapunerea, sau nesuprapunerea traficului autorizat cu cel neautorizat. Dacă cel autorizat lipseşte, atunci traficul care provine de la o adresă MAC falsificată va fi greşit interpretat ca trafic autorizat.

Din această categorie fac parte metodele care se bazează pe ID-ul de VLAN şi numerele de secvenţă. Aceşti parametrii vor avea o distribuţie normală în timp, dacă funcţionează numai o singură sursă pentru o adresă MAC. Deşi identifică extrem de rapid un intrus, atunci cât traficul produs de acesta se suprapune cu traficul autorizat, metodele nu au eficienţa dorită atunci când intrusul lucrează într-o perioadă de timp în care staţia validă este oprită.

Parametrul „interarrival time" calculat pentru staţia autorizată va fi diferit de cel calculat pentru staţia cu adresă MAC falsificată, dacă modelul echipamentului este diferit. Totuşi, în situaţia în care atacatorul poate veni în reţea cu un echipament identic cu cel autorizat, „interarrival time" va conduce la interpretare eronată, întrucât valoarea va corespunde cu cea a echipamentului autorizat.

Metodele de amprentare wireless nu sunt afectate negativ de absenţa suprapunerii traficului de la cele două surse. Fiecare tip de amprentă caută să identifice anumite caracteristici ale echipamentelor wireless. Totuşi, şi în cazul acesta trebuie remarcat faptul că în absenţa traficului autorizat, folosirea de către intrus a unor echipamente wireless de acelaşi tip cu cel autorizat, va putea ascunde identitatea reală a sursei.

În *concluzie*, în absenţa suprapunerii traficului ce provine de la sursa autorizată cu cel ce provine de la sursa neautorizată, metodele prezentate nu oferă totdeauna o certitudine în ceea ce înseamnă identificarea corectă a intrusului. Limitarea provine de obicei datorită faptului că un atacator poate veni cu un echipament similar celui pe care vrea să-l copieze.

2.6.5 Aplicabilitate în pofida mobilităţii staţiei client, şi trecerii acesteia dintr-o subreţea în alta.

O altă problematică de care trebuie ţinut cont este cea legată de mobilitatea staţiilor în timpul funcţionării. Această mobilitate va determina ca în punctul de monitorizare, o anumită adresă MAC, care era asociată cu un IP sursă, de la un moment dat va fi asociată cu alt IP. Reasocierea aceasta se va produce foarte rapid, pentru ca staţia mobilă să nu-şi piardă conectivitatea [Sas-09]. Schimbări rapide în datele asociate cu o adresă MAC pot surveni şi în urma apariţiei unui intrus. De aceea, metodele de detecţie trebuie să fie în stare să facă deosebire între cazurile de falsificare şi cazurile de trecere a staţiei mobile într-o altă subreţea.

Nu doar trecerea dintr-o subreţea în altă subreţea poate determina schimbarea IP-ului sursă. Sunt şi alte cazuri, cum ar fi de exemplu deconectarea scurtă şi reconectarea unei staţii în reţea. Serviciul de DHCP poate să aloce o altă adresă IP.

Metodele prezentate nu stau deloc bine la capitolul acesta. În primul rând trebuie remarcat faptul că metodele de amprentare wireless se referă de obicei nu la amprentarea unor calculatoare, ci la amprentarea unor dispozitive de reţea, cum ar fi Acces Point-urile. Din acest motiv, cele cinci metode de amprentare nu pot fi utilizate în contextul identificării unor adrese MAC ce au ca surse calculatoare mobile.

Metodele rămase sunt influenţate de trecerea staţiei mobile într-o altă subreţea. ID-ul de VLAN, numerele de secvenţă, valoarea „hop-count” sunt direct influenţate de trecerea în noua subreţea. Doar parametrul „interarrival time” ar putea să fie folosit, dar şi el este limitat deoarece trecerea dintr-o reţea în alta schimbă ratele de transfer, deci implicit se va modifica şi timpul mediu între două pachete.

În *concluzie*, metodele existente nu oferă un suport adecvat pentru tratarea corectă a falsificării adreselor MAC, în cazul în care staţiile sunt mobile şi îşi schimbă periodic apartenenţa de la o subreţea la alta.

2.6.6 Aplicabilitate în pofida utilizării de către intrus a unui echipament identic cu cel autorizat.

Deşi în discuţiile anterioare s-au abordat diverse aspecte legate de posibilitatea ca un intrus să folosească un echipament identic cu cel autorizat, se impun unele detalieri.

Riscul ca un atacator să fie în stare să aducă un echipament identic cu cel autorizat depinde de echipamentul în cauză. Dacă se face referire la calculatoare sau echipamente de reţea cu configuraţii speciale, atunci probabil că va fi dificil ca cineva să aduce un alt echipament identic. Dar, dacă se ţine cont de faptul că în multe reţele se folosesc ca staţii de lucru calculatoare sau laptop-uri de valoare mică sau medie, este foarte posibil ca în anumite situaţii, atacatorul să copieze nu numai adresa MAC, ci să aducă efectiv un calculator sau un laptop identic cu cel de la care a clonat adresa MAC.

Metoda de determinare pe baza ID-ului de VLAN dă rezultate cu condiţia ca intrusul să facă parte dintr-o reţea cu alt ID. În caz contrar, el va fi considerat ca autorizat.

Folosirea numerelor de secvenţă este la rândul ei influenţată de prezenţa sau absenţa traficului de la sursa autorizată, împreună cu cel de la sursa falsificată.

Dacă cele două surse funcţionează în paralel, numerele de secvenţă vor avea o distribuţie compromisă şi sistemul va identifica rapid situaţia. Dacă însă traficul înregistrat pe o adresă MAC provine numai de la sursa falsificată, atunci sistemul o va identifica drept autorizată.

Parametrul „interarrival time" va da valori identice pentru două echipamente identice. În consecinţă, singurul caz de aplicare a metodei este atunci când se suprapune traficul de la cele două surse. În rest, nu se poate face diferenţa între ele.

Parametrul „hop-count" poate şi el să fie folosit în regim limitat. Dacă intrusul provine dintr-o zonă diferită, este foarte probabil ca valoarea parametrului să fie diferită de cea a staţiei autorizate. Dacă însă vine în locul staţiei autorizate, sistemul nu o va identifica în mod corespunzător.

Despre metodele de amprentare wireless care folosesc răspunsul la mesajele deformate, distribuţia „probe-request" şi parametrul „clock skew", se poate afirma de la bun început faptul că nu pot fi aplicate în contextul unui intrus care vine cu acelaşi tip de echipament. Singurul caz posibil ar fi la metoda de identificare pe bază de intensitate a semnalului, dacă intrusul se află într-un alt loc decât staţia autorizată.

2.6.7 Aplicabilitate pentru echipamente de tip Desktop/Laptop.

Gama de echipamente prezente într-o reţea variază destul de mult. Orice echipament care poate trimite pachete în reţea, va avea o adresă MAC, pe care cineva ar putea încerca să o falsifice.

Cu toate acestea, interesul particular care va fi urmărit în ceea ce urmează se referă strict la calculatoare, fie că sunt în variantă desktop sau lapop. Se va studia în ce măsură se pot folosi metodele prezentate pe astfel de echipamente.

Metodele de amprentare wireless au în vedere echipamentele de reţea. De aceea, nu sunt aplicabile pentru calculatoare decât eventual dacă se face referire la laptop-uri care au interfaţă de reţea wireless. Totuşi, după cum reiese din literatura de specialitate [Lan-12], [Jan-10], [Ara-10], [Gao-10], [Edm-09] , [Cha-09], [Loh-08], [Bra-08], [Bri-08], [Jan-08], [She-08], metodele de amprentare wireless au fost concepute în mod special pentru echipamentele de reţea (rutere, switch-uri). Din acest motiv, se vor ignora în contextul calculatoarelor.

Toate celelalte patru metode discutate sunt aplicabile calculatoarelor. De aici rezultă că securizarea reţelelor împotriva pătrunderilor neautorizate prin falsificarea adreselor MAC de către atacatori care vin în sistem prin intermediul unui calculator, va putea fi realizată numai cu metode care permit aplicabilitatea pentru calculatoare.

În *concluzie*, limitările metodelor de amprentare wireless reduc foarte mult numărul metodelor care pot fi folosite în contextul identificării unui intrus.

2.7. Concluzii.

În urma considerării metodelor folosite în prezent pentru identificarea adreselor MAC falsificate, sau fost sintetizate atât avantajele, cât şi dezavantajele lor. Ca avantaj principal, fiecare metodă caută să determine cât mai rapid şi cât mai exact dacă adresa MAC descoperită în trafic, provine sau nu de la sursa care deţine în mod normal adresa respectivă, sau este vorba de un fals. Dezavantajele se referă în principiu la limitări ale domeniului de aplicabilitate.

Metodele au în general anumite caracteristici, dar ele sunt valide decât în anumite condiţii. Sau, metodele pot fi aplicate în general, dar cu anumite excepţii.

Drept urmare, se pune problema existenţei unor metode cu grad general de aplicabilitate, cost redus de implementare şi eficienţă sporită. Combinarea metodelor poate mări domeniul de aplicabilitate şi poate reduce astfel dezavantajele, însă nu pe toate.

Dacă se consideră detecţia unui intrus care a pătruns în reţeaua locală, abordarea pe care o au metodele prezentate nu este deloc încurajatoare. Cu alte cuvinte, un intrus care începe atacul său direct din interior, din locul unde funcţiona anterior staţia autorizată, va putea uşor să treacă peste o serie de alarme şi să fie considerat ca staţie autorizată. Cazurile acestea nu pot fi ignorate, întrucât pot fi produse chiar de către angajaţii instituţiei atacate. În contextul actual, atacurile cibernetice din interiorul reţelelor reprezintă o problemă reală.

De asemenea, s-a putut observa că unii algoritmi sunt limitaţi la reţele wireless, sau alţi algoritmi au nevoie să instaleze aplicaţii software pe calculatoarele clienţilor, ceea nu este posibil totdeauna. De asemenea, faptul că unii algoritmi identifică intrusul numai în prezenţa simultană a acestuia împreună cu staţia autorizată, reprezintă din nou o limitare importantă.

Posibilitatea ca o staţie să-şi modifice localizarea geografică şi să treacă dintr-o subreţea în alta, este o problemă care apare frecvent. Drept urmare, detecţia intruşilor trebuie să fie capabilă să ţină seama de acest aspect şi să nu interpreteze mobilitatea staţiei ca şi cum ar fi vorba de apariţia unui intrus. Totuşi, după cum s-a putut observa, abordările curente nu pot face faţă unei astfel de cerinţe.

În concluzie, trebuie să remarcăm necesitatea unor algoritmi care să fie capabili să depăşească neajunsurile metodelor prezentate şi să poată oferi o soluţie viabilă, şi care să facă faţă situaţiilor concrete din practică.

3. DETECŢIA ADRESELOR MAC FALSIFICATE PRIN METODA „DESTINATION TRAFFIC FINGERPRINT”

Capitolul descrie contribuţiile autorului în domeniul detecţiei adreselor MAC falsificate. Se prezintă o metodă originală de amprentare a unei staţii de lucru, bazată pe traficul generat către destinaţiile IP cu care aceasta comunică, şi care a fost denumită „*Destination Traffic Fingerprint*”, sau prescurtat *DTF*.

Capitolul debutează cu o descriere generală a metodei şi a contextului în care se aplică, continuând cu o caracterizare a „traficului constant”, folosit pentru amprentarea staţiei şi a parametrilor care definesc acest concept. Punctul focal al capitolului este dat de definirea Gradului Global de Recunoaştere, parametru calculat în faza de amprentare a staţiei, şi utilizat apoi în timp real pentru validarea staţiei.

O secţiune aparte în cadrul capitolului este alocată modelării matematice a metodei DTF. Această formalizare permite un studiu aprofundat al fazelor care compun stabilirea amprentei de referinţă şi validarea ulterioară, în timp real, pe baza amprentei stabilite iniţial. Etapa de amprentare este extrem de importantă întrucât afectează direct recunoaşterea staţiei. Logica fuzzy este adusă în discuţie în acest context, cu scopul de a determina caracterul constant al traficului.

La final, sunt prezentate o serie de servicii/tehnologii care favorizează aplicarea metodei DTF. Concluziile subliniază avantajele clare ale metodei şi beneficiile obţinute prin utilizarea ei.

3.1. Descrierea generală a metodei.

Pentru a stabili dacă identitatea unei staţii este reală sau falsificată, se urmăreşte alcătuirea unei amprente, sau semnături a staţiei originale. Deşi în literatură se întâlnesc o serie de metode care se bazează pe semnături [Lan-12], [Ara-10], [Gao-10], [Bri-08], [Bra-08], [Loh-06], modul de abordare al metodei propuse în teza de faţă, este unul original, şi încearcă să stabilească amprenta de trafic, urmărind adresele de IP cu care staţia comunică în mod constant [Sas-10a]. Metoda doar semnalează intrările neautorizate, dar nu intervine pentru blocarea sau eliminarea intrusului. Metoda a fost numită „*Destination Traffic Fingerprint*”, şi o vom nota în continuare *DTF*.

La baza dezvoltării metodei DTF stau câteva elemente de observaţie:

- un calculator este de cele mai multe ori folosit de către acelaşi utilizator;
- programele instalate pe calculator necesită în cele mai multe cazuri un trafic de date în reţeaua locală sau Internet;
- datorită faptului că utilizatorul foloseşte regulat aplicaţiile instalate pe calculator, apare probabilitatea ca anumite destinaţii IP să fie accesate periodic;

- pe calculator pot exista aplicaţii utilitare, cum ar fi de exemplu antivirusul, care rulează în memoria calculatorului şi verifică periodic anumite servere pentru update sau noutăţi;
- în reţelele companiilor, de multe ori se folosesc aplicaţii ERP ale companiei, servere de mail, baze de date sau alte resurse care necesită un trafic de date în reţea.

Toate aceste observaţii au condus la ideea că, din totalitatea adreselor IP cu care o staţie comunică într-un anumit timp, unele adrese apar frecvent în trafic şi ar putea defini un „trafic constant". Prin *„trafic constant"* nu se înţelege o rată de transfer constantă, ci o apariţie constantă a adreselor IP în trafic.

O amprentă de trafic, stabilită pe baza adreselor de IP către care o staţie emite constant pachete de date, este formată dintr-o mulţime de perechi:

$$M = \{P_1, P_2, \ldots P_n\}, \tag{4}$$

unde fiecare pereche conţine o adresă de IP şi procentul de prezenţă cu care apare în trafic. O astfel de pereche poate fi reprezentată ca:

$$P_i = (IP_i, PP_i) \tag{5}$$

unde PP_i reprezintă de fapt procentul de prezenţă al adresei IP în traficul pe care staţia îl generează.

Verificarea traficului se face pe o anumită perioadă, dar calculele sunt realizate la nivel de minut. Pentru fiecare adresă MAC care trimite pachete în unitatea de timp evaluată, interesează să se extragă toate adresele de IP spre care sursa (adresa MAC) a emis pachete şi să se calculeze numărul total de minute în care există trafic pentru fiecare IP destinaţie.

În continuare, pentru fiecare adresă IP destinaţie se calculează procentul de prezenţă, ca fiind raportul dintre numărul de minute în care s-a identificat trafic şi numărul total de minute evaluate. Nu interesează cantitatea de informaţie vehiculată, ci interesează dacă s-a vehiculat sau nu informaţie şi care este procentul de prezenţă al adresei în totalul de timp evaluat.

Procentul de prezenţă poate fi scris ca:

$$PP_i = \frac{TMP_i}{TME} \ \% \tag{6}$$

unde:

- TMP_i – reprezintă numărul de minute în care s-a identificat trafic către IP_i,
- TME – reprezintă numărul total de minute evaluate

Astfel, pentru o unitate de timp evaluată, notată „TU" (Time Unit), amprenta de trafic a unei staţii este alcătuită dintr-un număr de adrese IP „TDIP" (*Total Destination IPs*), fiecare cu procentul de prezenţă aferent, şi se poate reprezenta astfel:

$$M = \left\{ \left(IP_1, \frac{TMP_1}{TME} \right), \left(IP_2, \frac{TMP_2}{TME} \right), \ldots, \left(IP_{TDIP}, \frac{TMP_{TDIP}}{TME} \right) \right\} \quad (7)$$

Pentru o reprezentare mai elocventă, se descrie amprenta de trafic sub forma unui tabel, în care fiecare element al mulţimii M devine o linie a tabelului.

Ca exemplu, considerăm următoarea amprentă de trafic, calculată pentru o perioadă de 8 ore (TME = 480 minute), reprezentată în Tabelul 1

Tabelul 1 - Exemplu amprentă de trafic

Nr.	***Adresă IP***	***Procent Prezenţă***
1	193.252.115.186	27.41 %
2	200.36.0.0	18.21 %
3	192.168.200.255	8.38 %
4	224.0.0.251	2.67 %
5	224.0.0.252	1.66 %
6	239.255.255.250	1.25 %

Liniile din Tabelul 1 reprezintă perechile din mulţimea M. Procentul de prezenţă de 27.41% aferent adresei 193.252.115.186 reprezintă faptul că din cele 480 de minute evaluate, 131 minute conţin trafic către această destinaţie. Nu interesează cantitatea de informaţie transmisă, ci numai cât la sută din totalul minutelor se identifică trafic spre destinaţia dată. Analog, pentru adresa 200.36.0.0 s-a găsit trafic în 87 din cele 480 minute.

Ca observaţie, trebuie menţionat faptul că, deşi sunt de preferat destinaţiile pentru care procentul de prezenţă este foarte mare, chiar şi cele cu doar câteva procente nu sunt de neglijat, atâta timp cât traficul este păstrat la rata respectivă.

O dată ce s-a stabilit amprenta de trafic, ea poate fi înregistrată şi folosită ulterior pentru recunoaşterea staţiei. Recunoaşterea presupune calcularea amprentei de trafic actuală şi compararea ei cu amprenta nominală, înregistrată în procesul de evaluare iniţială.

3.2. Domeniul de aplicabilitate al metodei.

Datorită faptului că metoda DTF se foloseşte exclusiv de adrese IP ale unor destinaţii care apar constant în traficul unei staţii din reţea, domeniul de aplicabilitate al metodei se limitează la calculatoarele pentru care se poate determina amprenta de trafic. Nu la orice calculator se poate extrage amprenta de trafic, dar, după cum va rezulta în continuare, metoda poate fi aplicată în foarte multe cazuri.

Generarea alarmelor cu privire la pătrunderile neautorizate sunt destinate în mod special reţelelor de calculatoare ale unor companii, care nu doresc să permită unor utilizatori străini să beneficieze de resursele sau serviciile puse la dispoziţie prin intermediul reţelei locale. De obicei, companiile folosesc sisteme ERP, ale căror

module comunică între ele prin intermediul reţelei, astfel încât se creează premisele necesare pentru a putea genera semnăturile. În plus, multe companii au instalat reţele virtuale, care creează în mod obligatoriu un trafic între serverul de VPN şi staţiile locale. Alte aplicaţii folosesc servere de baze de date, care de asemenea crează trafic constant.

Segmentul calculatoarelor personale oferă la rândul lui o mare probabilitate de aplicabilitate pentru metoda DTF, întrucât pentru mulţi utilizatori, aplicaţiile software instalate pe calculator nu suferă modificări semnificative decât la intervale mai mari de timp şi astfel există şanse ridicate să se poată genera amprentele de trafic.

Practic, imposibilitatea de aplicare a metodei survine atunci când calculatoarele sunt folosite de către mai mulţi utilizatori, fiecare având propriile preferinţe la aplicaţiile software rulate. Exemplele cele mai reprezentative ar fi laboratoarele universităţilor sau liceelor, unde pe acelaşi calculator lucrează un număr foarte mare de persoane, fiecare având scopuri diferite şi implicit aplicaţii software diferite. De asemenea, un alt exemplu ar fi reţelele de tip „Internet Cafe", care sunt folosite fără reguli şi care crează un trafic „haotic". Toate cazurile amintite reduc şansele de a determina o amprentă de trafic deoarece este foarte dificil să se găsească adrese IP cu care staţia comunică în mod constant. Totuşi, există soluţii chiar şi în situaţiile acestea, prin forţarea unui trafic constant cu ajutorul unor aplicaţii software instalate pe calculatoarele respective, aplicaţii care să creeze trafic constant, precum şi utilizarea reţelelor VPN şi a serverelor de baze de date.

3.3. Determinarea traficului constant.

Traficul generat de o staţie din reţea, către destinaţiile IP cu care aceasta comunică, se poate încadra în următoarele categorii:

- trafic constant
- trafic temporar
- trafic punctual

Traficul constant, se defineşte ca fiind un transfer de date între o staţie din reţea şi o destinaţie IP, care, fiind monitorizat la nivel de minut pentru o perioadă lungă de timp, prezintă o frecvenţă relativ constantă a minutelor în care sursa a trimis pachete de date spre destinaţia IP pe întreaga perioadă evaluată. Nu se urmăreşte o evaluare cantitativă a traficului, ci doar prezenţa sau absenţa sa la nivel de minut.

Traficul temporar, se defineşte ca fiind un transfer de date între o staţie din reţea şi o destinaţie IP, care, fiind monitorizat la nivel de minut pentru o perioadă lungă de timp, prezintă o frecvenţă relativ constantă a minutelor în care sursa a trimis pachete de date spre destinaţia IP, doar pentru anumite subintervale de timp din întregul interval evaluat.

Traficul punctual, se defineşte ca fiind un transfer de date între o staţie din reţea şi o destinaţie IP, care, fiind monitorizat la nivel de minut pentru o perioadă lungă de timp, prezintă minute în care sursa a trimis pachete de date spre destinaţia IP, repartizate aleatoriu şi izolat pe axa timpului, pe durata intervalului de timp evaluat.

În studiul de faţă, scopul propus este acela de a detecta cu precizie traficul constant. Toate IP-urile care nu se încadrează în acest caz, vor fi eliminate din

procesul generării amprentei de trafic. Stabilirea amprentelor reprezintă partea cea mai importantă a metodei. Cu cât se stabileşte mai corect amprenta, cu atât creşte gradul de recunoaştere a staţiei şi scade rata alarmelor false. Prin „stabilire corectă” a amprentei se înţelege de fapt identificarea corectă a adreselor IP pentru care există trafic constant.

Pentru o detecţie corectă, este necesară introducerea unor parametrii, care să fie folosiţi în filtrarea destinaţiilor şi păstrarea doar a celor care au un trafic constant. Parametrii propuşi sunt:

- Procentul de Prezenţă
- Procentul de Absenţă Maximă
- Criteriul de prezenţă pe subintervale egale
- Puterea amprentei de referinţă

În continuare, fiecare parametru va fi prezentat în detaliu, urmărindu-se locul şi rolul pe care îl are în determinarea traficului constant.

3.3.1. Procentul de Prezenţă.

Primul parametru este *„Procentul de Prezenţă”* PP_i, stabilit ca fiind raportul dintre numărul de minute în care există trafic către o destinaţie anume, şi numărul total de minute aferente perioadei de evaluare:

$$PP_i = \frac{TMP_i}{TME} \quad \% \tag{8}$$

Folosirea procentului de prezenţă reprezintă o condiţie necesară, dar nu este suficientă pentru a caracteriza traficul constant. Aşa cum s-a menţionat deja, deşi se doreşte ca în componenţa amprentelor de trafic să intre destinaţii cu un procent mare de prezenţă, totuşi, ceea ce contează în primul rând nu este valoarea în sine, ci asigurarea că această valoare este păstrată la valori aproximativ constante. S-ar putea să se identifice destinaţii la care procentul de prezenţă să fie de doar câteva procente, dar care să fie mult mai bune decât destinaţii cu procent de prezenţă ridicat, dar cu variaţii mari ale prezenţei în trafic.

Pentru o înţelegere mai bună a aspectelor prezentate, se vor prezenta în continuare câteva exemple reprezentative. Graficele ilustrează prezenţa / absenţa unor destinaţii IP, în traficul înregistrat. Axa OY va marca timpul în minute, iar axa OX va marca prezenţa (valoarea „1”) sau absenţa (valoarea „0”) adresei IP evaluate, în traficul staţiei. Perioadele de oprire sunt marcate prin dreptunghiuri haşurate, cu notaţii de genul „Stop *i*”, unde „i” reprezintă numărul opririi, cu începere de la „0”. Urmărind datele extrase din traficul real al mai multor calculatoare, se pot identifica următoarele categorii:

- prezenţă foarte ridicată (chiar aproape de 100%);
- prezenţă ridicată sau moderată, dar constantă;
- prezenţă scăzută, dar totuşi constantă;
- prezenţă ridicată, dar concentrată pe anumite perioade scurte;
- prezenţă scăzută sau chiar punctuală;

Datele au fost preluate dintr-un test realizat pe aproximativ 100 de calculatoare, test despre care se discută în detaliu într-un capitol următor.

Prezenţă foarte ridicată, chiar aproape de 100%

Cazul acesta corespunde unor adrese IP care apar aproape permanent în traficul curent. Când se face referire la „aproape permanent”, se înţelege procente de prezenţă de peste 99% sau chiar de 100%. Identificarea lor este foarte importantă, întrucât conduc la stabilirea unor amprente de trafic cu grad de încredere ridicat.

În Fig. 3.1 este reprezentată prezenţa destinaţiei 30.24.0.0 în traficul înregistrat într-o reţea, pe o durată de 1600 minute. Pentru fiecare minut s-a reprezentat pe grafic valoarea „1” dacă în acel minut a existat trafic către destinaţia în cauză, sau „0” în caz contrar. Zona haşurată în culoare albastru reprezintă o perioadă de timp în care staţia evaluată a fost oprită.

Se poate observa faptul că, în afară de două minute izolate, staţia evaluată a trimis cel puţin un pachet către destinaţia IP menţionată. În condiţiile acestea, adresa IP destinaţie poate fi considerată o adresă cu care staţia evaluată comunică în mod constant.

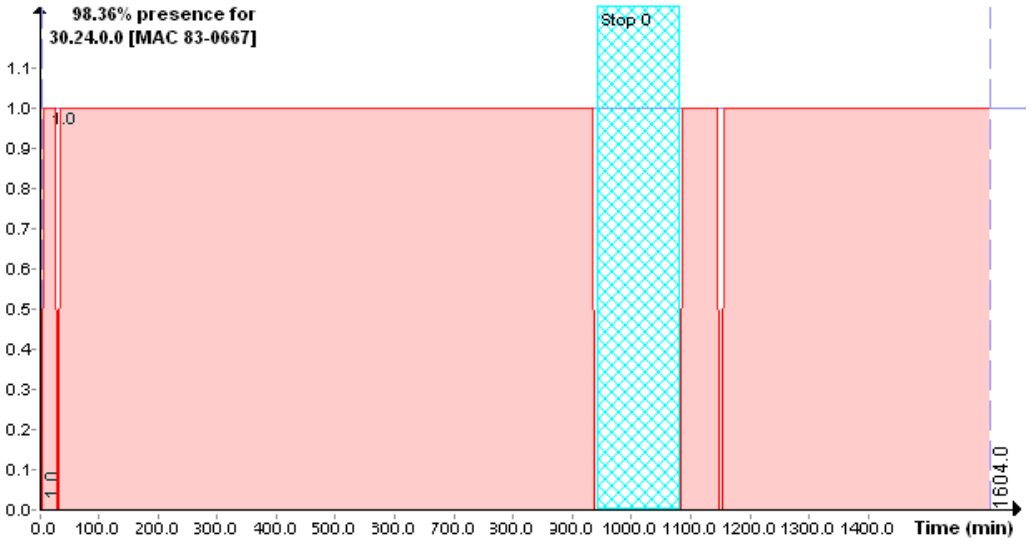


Fig. 3.1 Prezenţă foarte ridicată pentru destinaţia 30.24.0.0

Un alt exemplu este prezentat în Fig. 3.2, procentul de prezenţă fiind de 99.11% pentru un interval de 1450 minute. Pe durata evaluată staţia a fost oprită într-un interval foarte scurt, în jurul minutului 350. În rest, staţia a fost permanent funcţională, iar graficul afişează doar unul sau două minute în care nu s-a detectat trafic către adresa IP 94.178.105.3. Destinaţia aceasta poate intra în categoria destinaţiilor cu trafic constant.

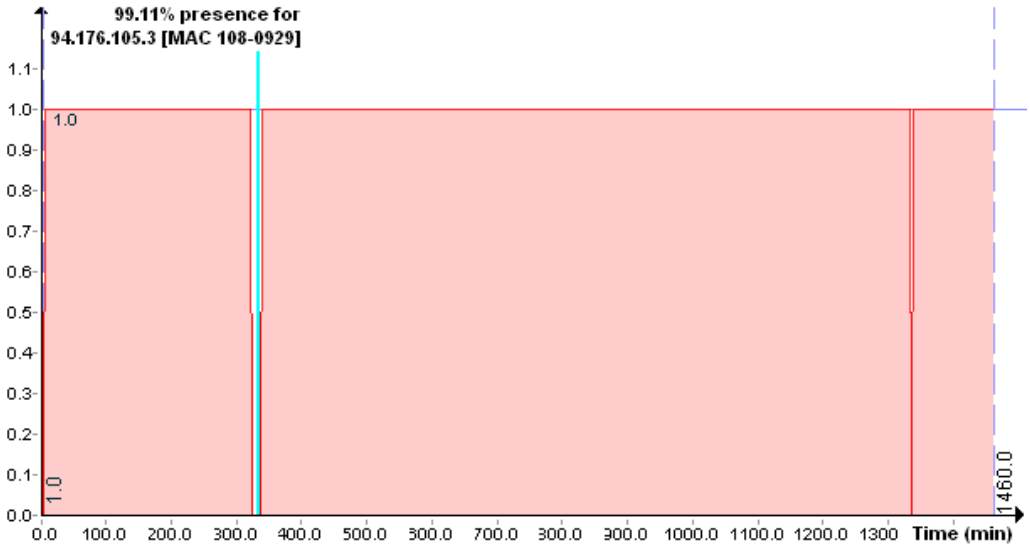


Fig. 3.2 Prezenţă foarte ridicată pentru destinaţia 94.176.105.3

Un alt exemplu este prezentat în Fig. 3.3, în care destinaţia 157.130.89.170 a fost urmărită pe o staţie timp de 1550 minute. În două rânduri staţia s-a oprit, momentele fiind marcate pe grafic prin două zone haşurate cu albastru. În timpul de funcţionare se pot identifica un mic număr de minute în care nu s-a identificat trafic către adresa IP menţionată anterior. În rest, fiecare minut conţine pachete spre IP-ul urmărit. Prezenţa totală este de 97.10%, foarte bună pentru ca destinaţia IP să fie considerată cu trafic constant.

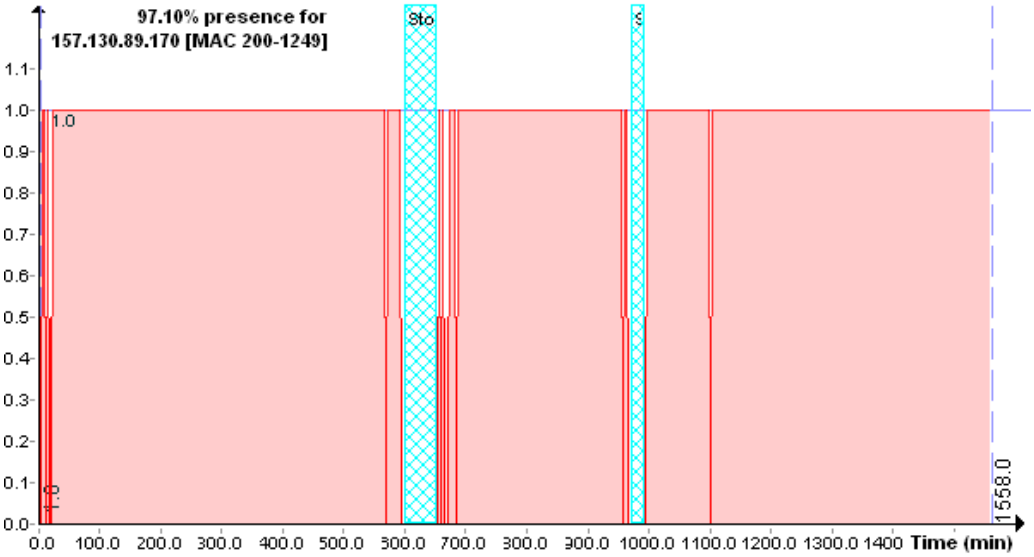


Fig. 3.3 Prezenţă foarte ridicată pentru destinaţia 157.130.89.170

Fig. 3.4 prezintă un caz cu prezenţă ridicată, dar cu o distribuţie diferită de cele observate în figurile anterioare. În cazul acesta, destinaţia 200.4.0.0 prezintă mult mai multe minute de absenţă în trafic, rezultând până la final un procent de prezenţă în valoare de 81.77%. Totuşi, comunicarea staţiei cu destinaţia menţionată este consistentă, astfel încât se poate încadra în categoria de trafic constant.

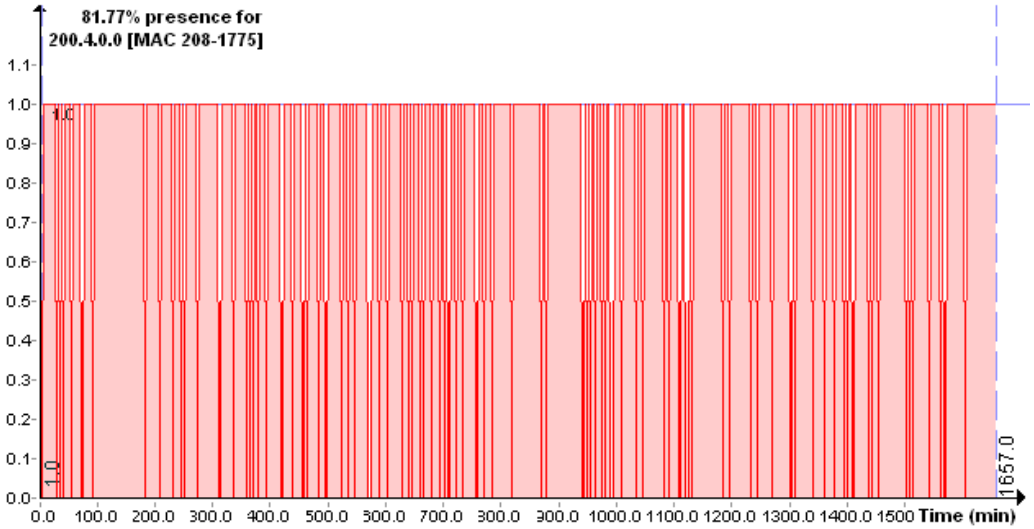


Fig. 3.4 Prezenţă foarte ridicată pentru destinaţia 200.4.0.0

În Fig. 3.5 prezenţa destinaţiei 239.192.152.143 este şi mai mică, întrucât minutele de pauză sunt mai numeroase. Distribuţia minutelor cu trafic ne conduce însă la concluzia că traficul poate fi considerat constant.

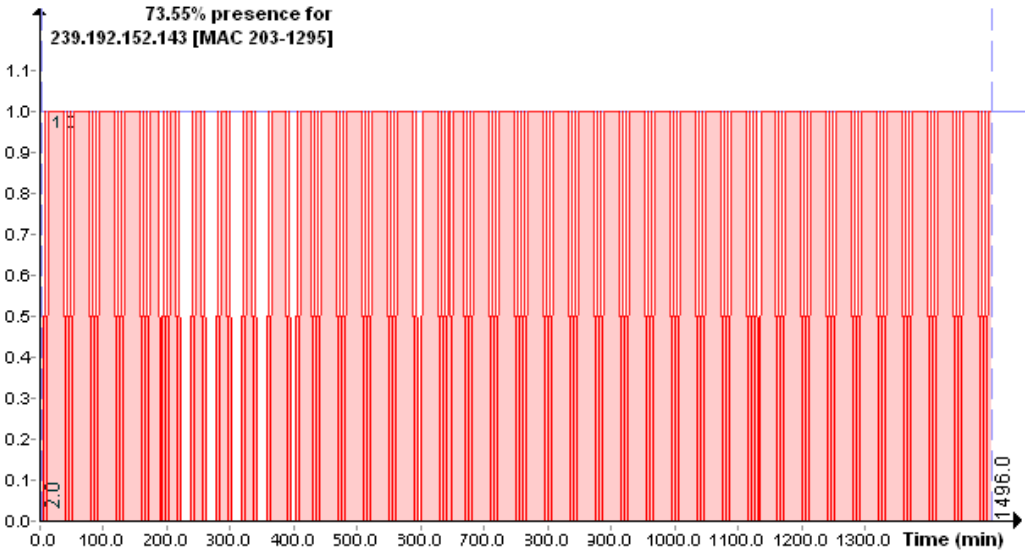


Fig. 3.5 Prezenţă foarte ridicată pentru destinaţia 239.192.152.143

Prezenţă ridicată sau moderată, dar constantă

Chiar dacă procentele de prezenţă nu tind spre 100%, destinaţiile îşi păstrează traficul în timp, şi pot fi incluse în amprente. Apar mai des perioade de „pauză”, către destinaţia luată în considerare, dar pauzele sunt de scurtă durată.

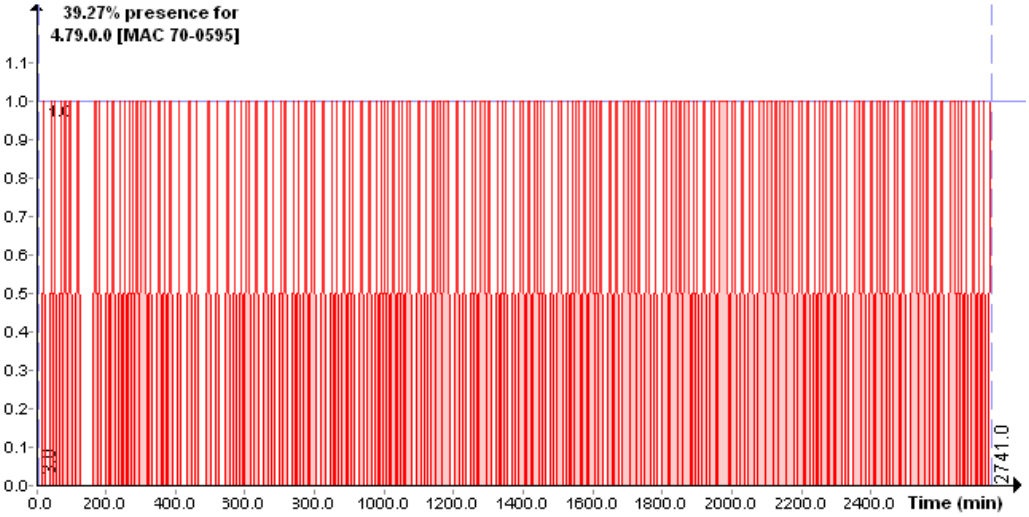


Fig. 3.6 Prezenţă moderată dar constantă pentru destinaţia 4.79.0.0

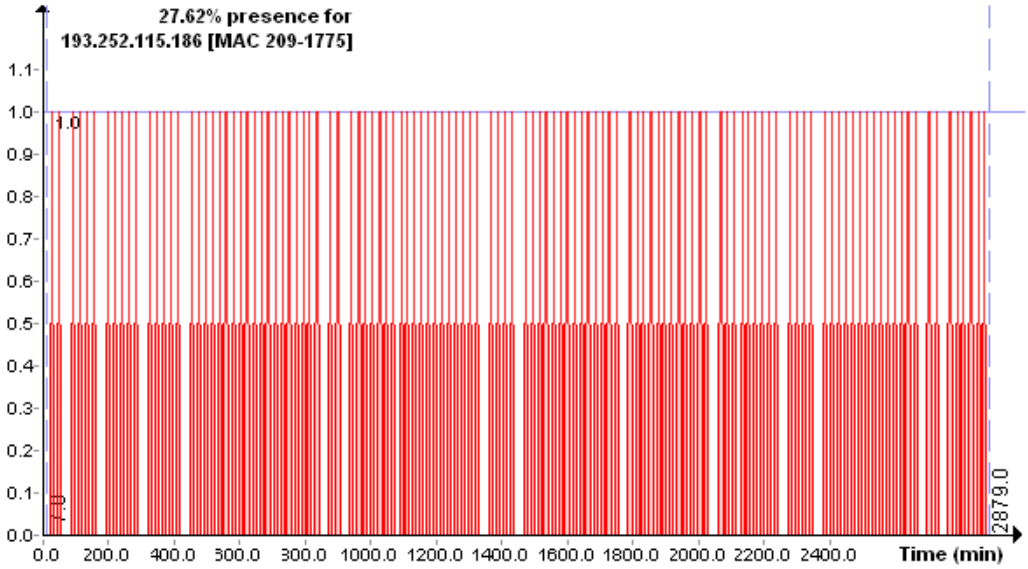


Fig. 3.7 Prezenţă moderată dar constantă pentru destinaţia 193.252.115.186

Fig. 3.6 prezintă identificarea traficului către destinaţia 4.79.0.0. Prezenţa calculată global pentru cele 2740 de minute evaluate, a condus la un rezultat de 39.27%. Prezenţa aceasta este mult mai mică decât valorile observate la punctul anterior. Totuşi, ceea ce se caută este nu neapărat o cantitate mare de trafic, cât o reluare a traficului într-o manieră cât mai constantă.

Aceasta înseamnă că, deşi se preferă ca procentul de prezenţă să tindă spre 100%, destinaţiile cu prezenţă mai scăzută dar totuşi constantă, pot fi luate în considerare atunci când interesează identificarea traficului constant.

Asemănător este cazul din Fig. 3.7, unde procentul de prezenţă este şi mai mic, dar păstrează o frecvenţă relativ constantă pe toată durata de aproximativ 2880 minute.

Fig. 3.8 prezintă o situaţie asemănătoare. Pauzele în traficul către destinaţia IP 200.13.0.0 sunt dese, dar cu frecvenţă relativ constantă. Adresa menţionată poate fi inclusă în amprentele de trafic.

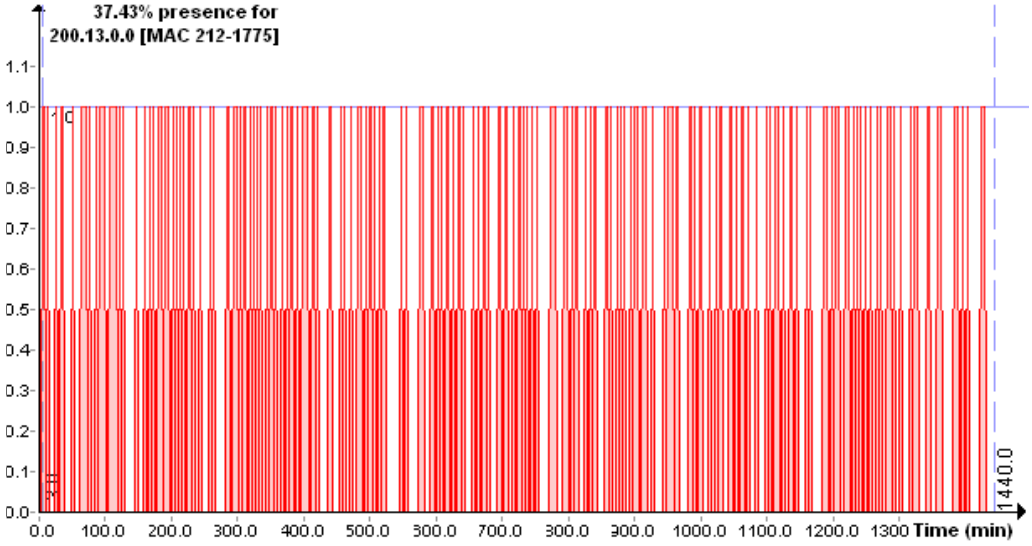


Fig. 3.8 Prezenţă moderată dar constantă pentru destinaţia 200.13.0.0

Prezenţă scăzută, dar totuşi constantă

De multe ori, procentul de prezenţă este foarte mic, de doar câteva procente. Cu toate acestea, s-au întâlnit situaţii frecvente în care traficul este reluat la intervale aproximativ egale, chiar dacă pauzele sunt mult mai mari. Cu toate acestea, faptul că destinaţia IP revine în trafic, este suficient pentru a considera IP-ul ca o destinaţie constantă şi a se putea include în amprentele de trafic.

În categoria aceasta vor intra destinaţii IP la care traficul apare cu pauze de durată mai lungă decât minutele în care identificăm trafic. Totuşi aceste minute apar la intervale aproximativ egale, ducând în felul acesta la obţinerea unei frecvenţe relativ constante.

De exemplu, Fig. 3.9 prezintă traficul către destinaţia 65.55.17.39. Traficul a fost urmărit pentru aproximativ 1750 minute, timp în care se observă din grafic că a existat o perioadă de oprire a staţiei pentru aproximativ 120 minute. În rest, frecvent apar minute în care se descoperă adresa IP în traficul staţiei. Deşi per ansamblu se obţine doar 5.16% prezenţă, acestă prezenţă este distribuită relativ constant.

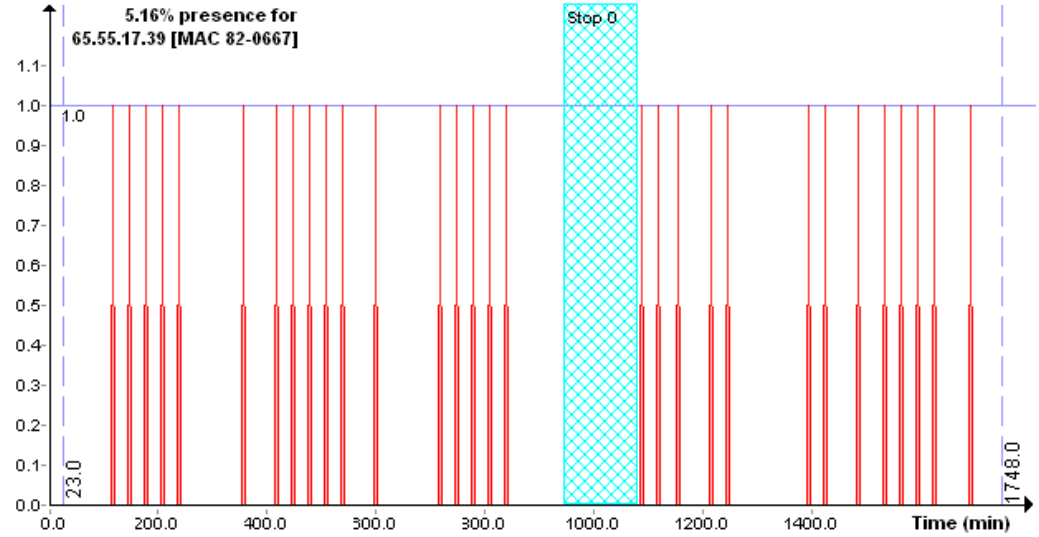


Fig. 3.9 Prezenţă scăzută dar constantă pentru destinaţia 65.55.17.39

Un alt exemplu este în Fig. 3.10, unde prezenţa traficului spre adresa IP 93.113.235.93 are o valoare foarte mică, de 3.61% pe întreg intervalul de 1440 minute. Apariţia traficului este practic de forma unor „impulsuri”, care apar frecvent în timp.

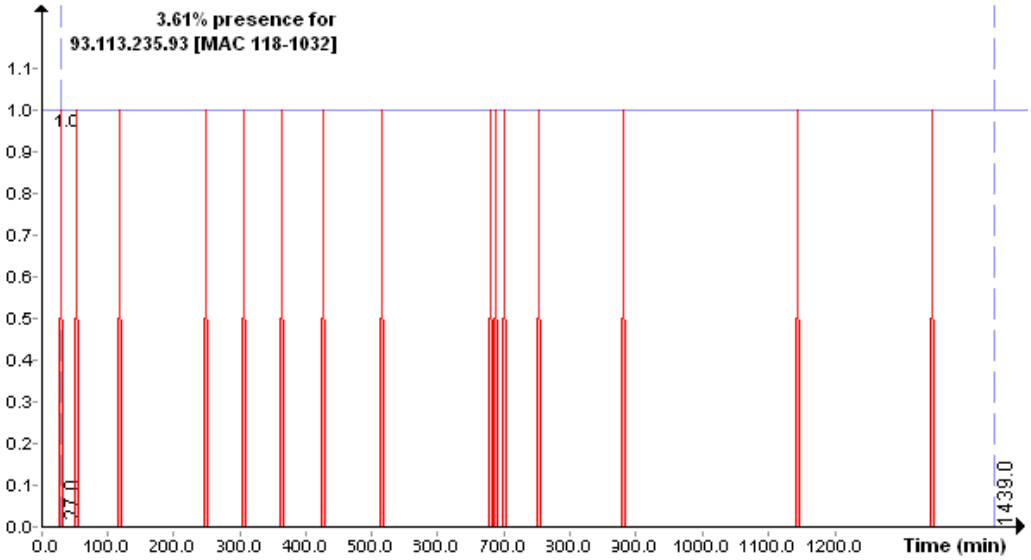


Fig. 3.10 Prezenţă scăzută dar constantă pentru destinaţia 93.113.235.93

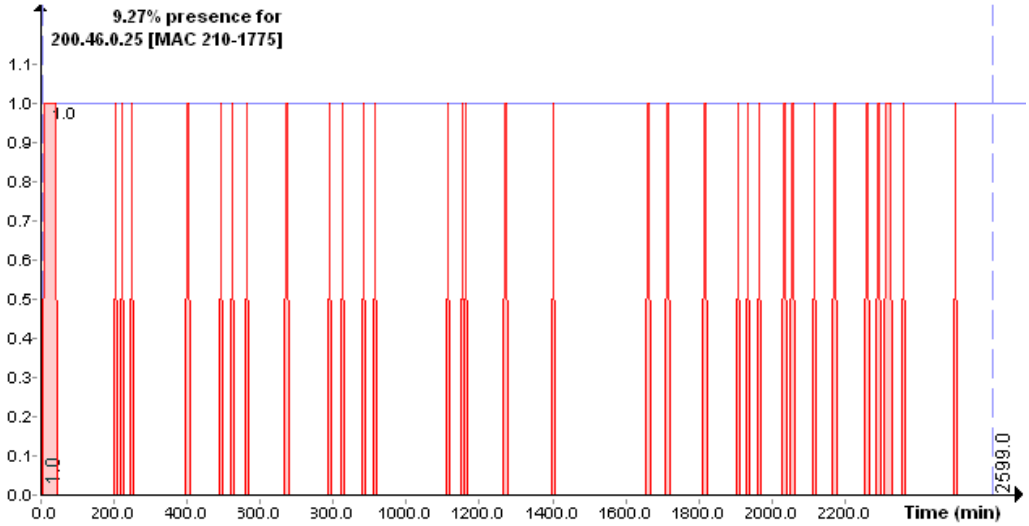


Fig. 3.11 Prezenţă scăzută dar constantă pentru destinaţia 200.46.0.25

În Fig. 3.10 şi Fig. 3.11, deşi este o prezenţă redusă, frecvenţa „impulsurilor" de trafic este mai bună. Repartizarea minutelor în care descoperim trafic are o acoperire mai avantajoasă, cu un grad mai ridicat de constnanţă.

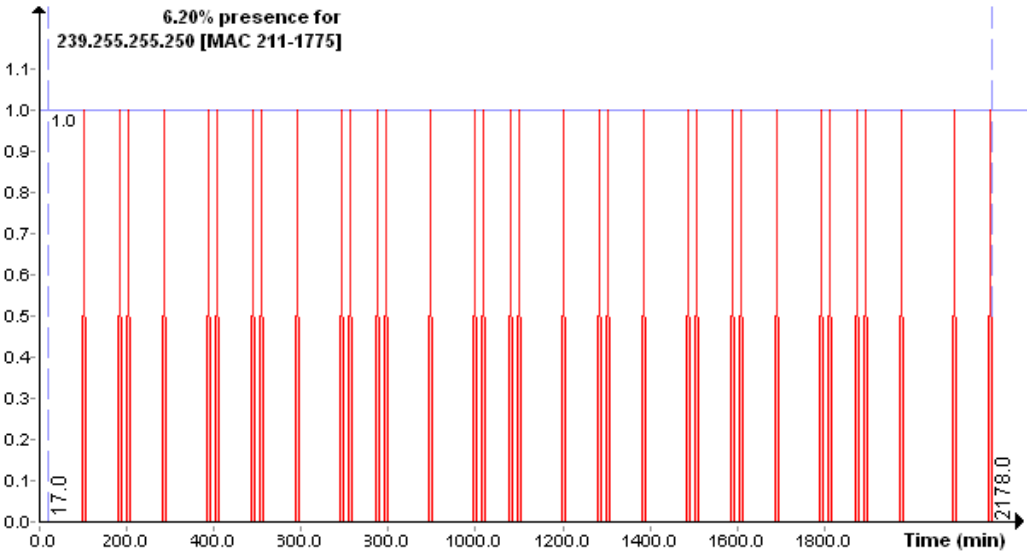


Fig. 3.12 Prezenţă scăzută dar constantă pentru destinaţia 239.255.255.250

Prezenţă ridicată, dar concentrată pe anumite perioade scurte

S-a afirmat faptul că procentul de prezenţă este un bun început, dar nu este întotdeauna suficient pentru a descrie realitatea din punct de vedere al „traficului constant". Dacă traficul este concentrat puternic pe anumite zone a axei timpului, iar

în rest este absent sau sporadic, per ansamblu s-ar putea ca procentul de prezemță să dea valori semnificative, dar destinaţia să nu aibă deloc trafic constant. În situaţia aceasta se află şi cazurile descrise în continuare.

Considerând cazul ipotetic în care o destinaţie IP este prezentă 100% pe jumătate de interval, iar apoi dispare complet, din punct de vedere al procentului de prezenţă, el are o valoare considerabilă: 50%. Totuşi, faptul că traficul către destinaţia respectivă dispare complet, este un indicator clar că nu se poate accepta traficul ca fiind constant.

Pentru evidenţierea traficului ridicat, dar neconstant, trebuie efectuate verificări suplimentare, pe baza altor parametrii.

Fig. 3.11 este un exemplu de trafic concentrat doar pe anumite perioade. Din totalul de aproximativ 3100 minute, staţia a fost oprită două perioade mai lungi, şi două perioade foarte scurte. Între aceste perioade de oprire, staţia a trimis pachete către destinaţia IP 67.195.186.249 aproape în fiecare minut al perioadei de dinaintea primei opriri, după care nu se poate identifica nici măcar un singur minut cu trafic spre destinaţia menţionată.

În mod evident, procentul de prezenţă calculat global la valoarea de 18.94% nu poate fi luat în considerare întrucât tot traficul este limitat în timp. Pentru primele 350 de minute, există o prezenţă de aproape 100% a destinaţiei în traficul staţiei, dar, după oprirea staţiei, nu se mai reia dialogul cu IP-ul menţionat.

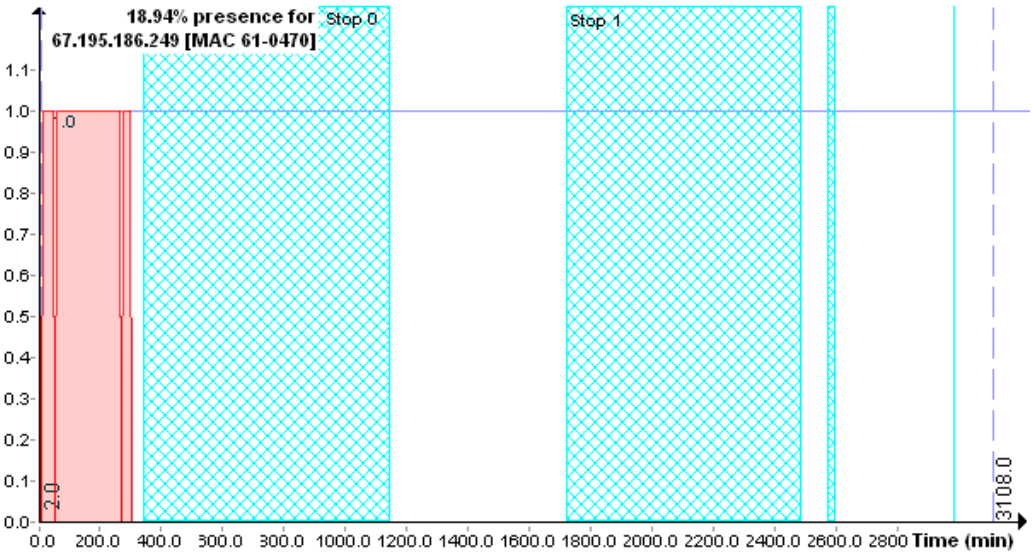


Fig. 3.13 Prezenţă concentrată pe perioadă scurtă pentru destinaţia 67.195.186.249

De asemenea, în Fig. 3.14, procentul de prezenţă este aproape dublu faţă de cel din Fig. 3.13. Destinaţia 98.138.26.127 are trafic concentrat în perioada dintre primele două opriri ale staţiei.

Nu se poate folosi o astfel de adresă pentru amprenta de trafic a staţiei, întrucât după momentul de prezenţă aproape 100%, traficul dispare complet. Dacă o astfel de destinaţie intră în componenţa amprentei de trafic a staţiei, rezultatele ar fi alterate puternic.

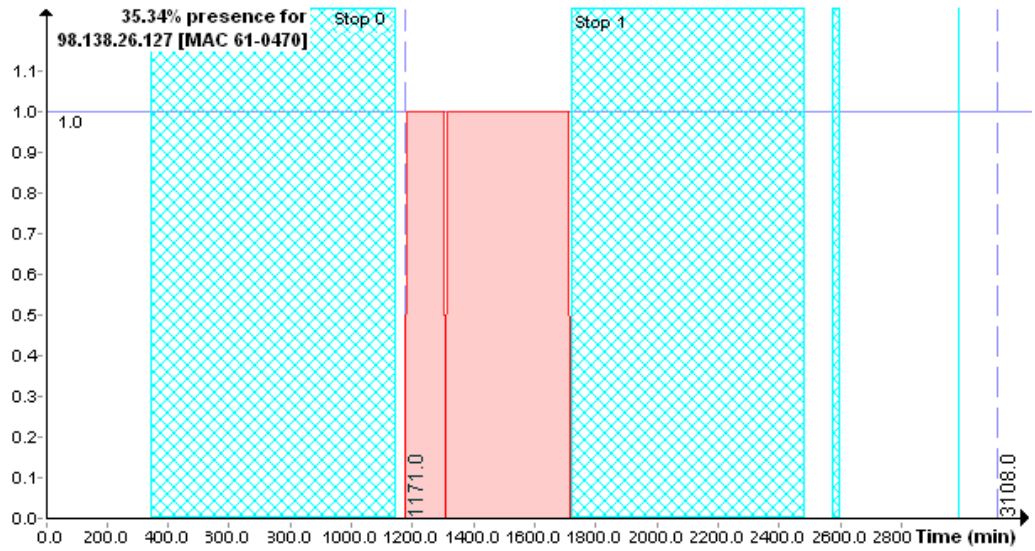


Fig. 3.14 Prezenţă concentrată pe perioadă scurtă pentru destinaţia 98.138.26.127

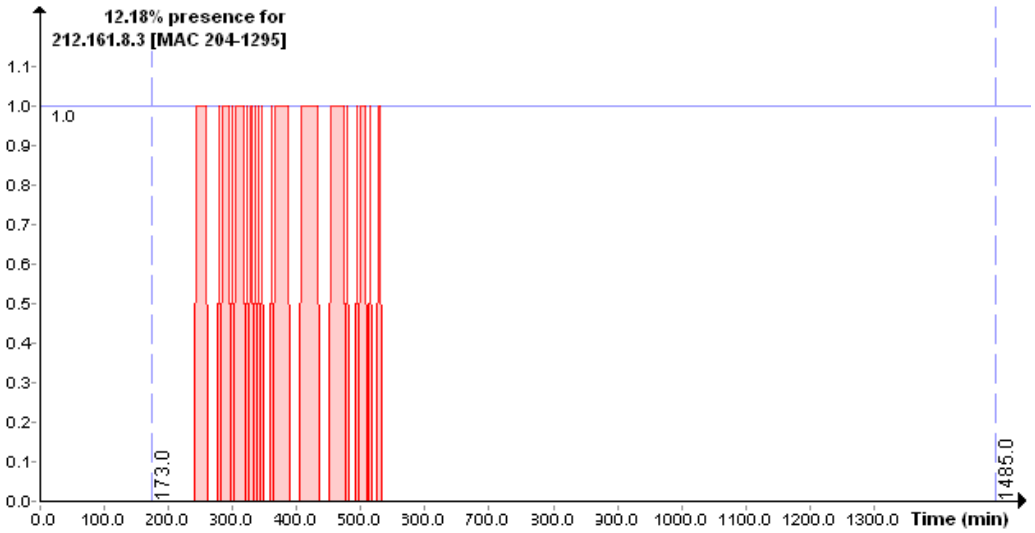


Fig. 3.15 Prezenţă concentrată pe perioadă scurtă pentru destinaţia 212.161.8.3

Un alt exemplu este în Fig. 3.15. Nici în cazul acesta nu se poate vorbi despre un trafic constant întrucât este concentrat în zona delimitată de minutele 250 şi 550.

Prezenţă scăzută sau chiar punctuală

Cazul acesta este cel mai frecvent. Este şi normal, întrucât traficul de reţea depinde de cerinţele de moment ale utilizatorilor, care au nevoie de informaţii aflate pe servere disparate. Nu se poate vorbi de pauze în trafic, ci, mai degrabă de prezenţă limitată sau chiar punctuală. Rezultă în mod clar că nu se poate permite ca în amprentele de trafic să apară astfel de situaţii.

Pentru calculatoarele care au instalate aplicaţii software care generează trafic către un număr foarte mare de IP-uri, destinaţiile punctuale vor trebui eliminate rapid din procesul de verificare.

În continuare se prezintă două situaţii concrete identificate pe aceeaşi staţie. Fig. 3.16 şi Fig. 3.17 se referă la acelaşi calculator, pentru care se observă o serie de opriri marcate pe grafic prin zone haşurate cu albastru.

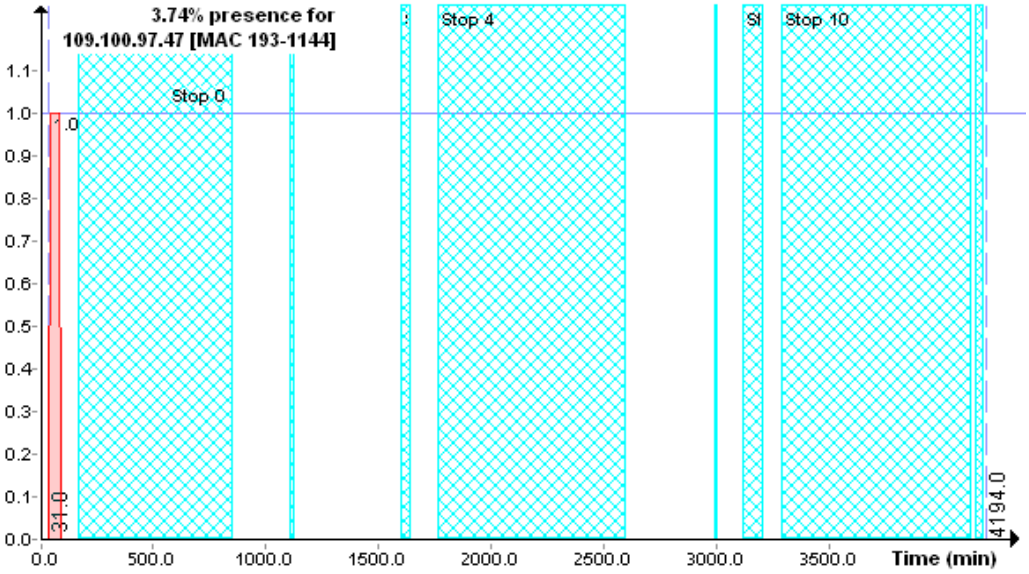


Fig. 3.16 Prezenţă punctuală pentru destinaţia 109.100.97.47

Atât în Fig. 3.16 cât şi în Fig. 3.17, minutele în care există trafic sunt punctuale. Nu pot fi incluse astfel de adrese în amprentele de trafic întrucât ar produce rezultate cu totul eronate.

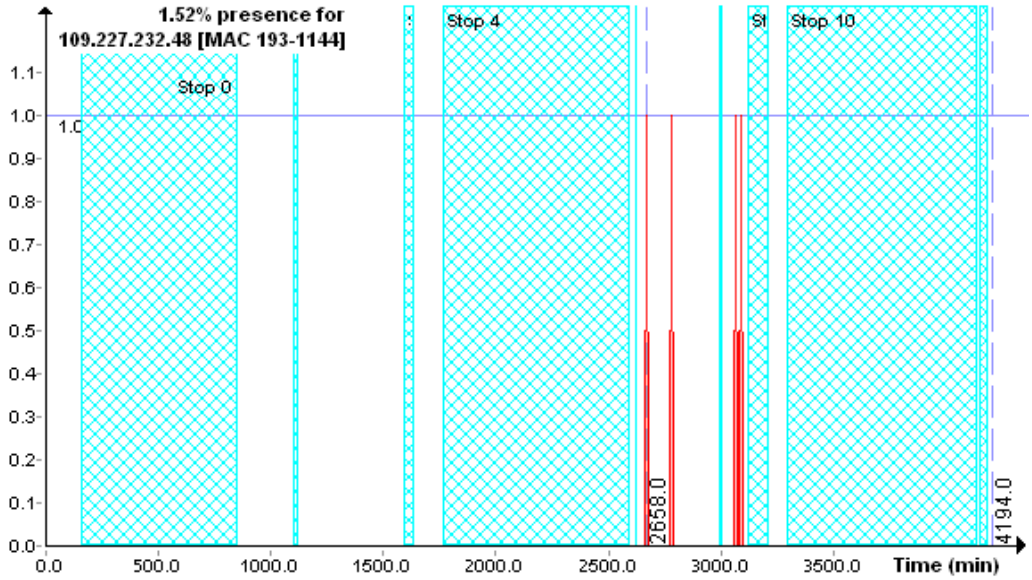


Fig. 3.17 Prezenţă punctuală pentru destinaţia 109.227.232.48

Fig. 3.18 prezintă o situaţie de pe o staţie care a funcţionat continuu 2600 minute. Din totalul minutelor, doar în 3 minute izolate se poate observa trafic către adresa de IP 209.85.149.100.

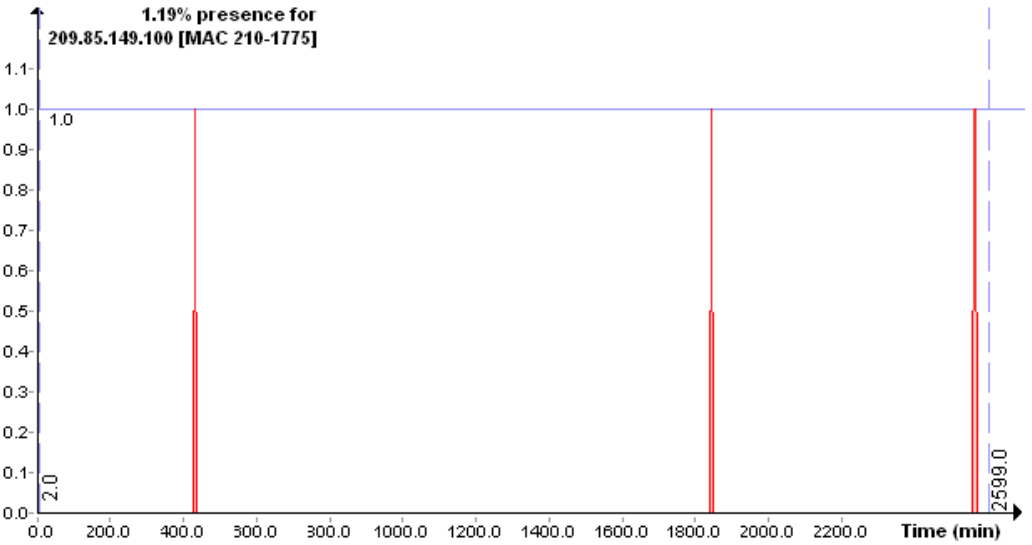


Fig. 3.18 Prezenţă punctuală pentru destinaţia 209.85.149.100

În Fig. 3.19, traficul este monitorizat pe parcursul a 1485 minute, urmărindu-se pachetele care pleacă spre destinaţia cu IP 212.233.167.81. În pimele 250 de minute, staţia a comunicat frecvent cu destinaţia în cauză, dar ulterior nu.

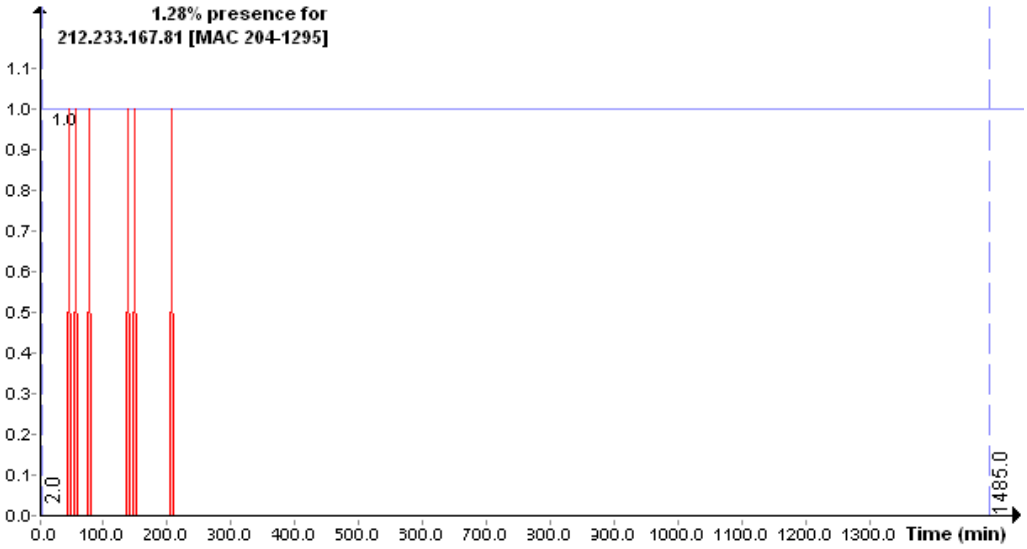


Fig. 3.19 Prezenţă punctuală pentru destinaţia 212.233.167.81

3.3.2. Procentul de Absenţă Maximă.

Procentul de absenţă maximă reprezintă numărul maxim de minute consecutive în care nu a avut loc trafic către o destinaţie, raportat la numărul total de minute de evaluare:

$$PA_i = \frac{TMAC_i}{TME} \ \% \tag{9}$$

unde:

- PA – procent de absenţă maximă;
- TMAC – numărul total de minute de absenţă, consecutive;
- TME – numărul total de minute de evaluare.

El reprezintă un alt mod de a caracteriza traficul, totuşi cu aplicabilităţi reduse. Procentul de absenţă maximă poate să dea o indicaţie despre constanţa traficului, dar el nu poate caracteriza complet traficul. O valoare mare a procentului se traduce într-un număr mare de minute în care traficul a lipsit cu desăvârşire. Dacă timpul evaluat (măsurat) devine comparabil cu gradul de absenţă, înseamnă că s-ar prea putea ca aceasta să dovedească un trafic temporar sau punctual.

Problema principală a acestui indicator, este că el este puternic influenţat de timpul de evaluare. Timpul maxim de absenţă consecutivă îşi poate păstra valoarea de-a lungul timpului, pe când numărul total de minute de evaluare creşte o dată cu fiecare minut. De aici rezultă că PA va furniza valori mari pentru timp de evaluare mic, şi va scădea în timp, pe măsură ce timpul de evaluare creşte. Din cauza aceasta, folosirea PA pentru caracterizarea traficului, nu este concludentă.

3.3.3. Criteriul de prezenţă pe subintervale egale.

Pentru a detecta cât de „constant" este traficul către o anumită destinaţie IP, se împarte intervalul de timp evaluat în subintervale, apoi, fiecare subinterval este verificat individual, pentru a afla dacă apare sau nu trafic în fiecare subinterval. Cu cât este mai mare numărul de subintervale, cu atât creşte precizia de detecţie a traficului constant.

Din punct de vedere teoretic, pentru ca o destinaţie să fie acceptată ca având trafic constant, ar trebui să se regăsească în toate subintervalele, dar, din punct de vedere practic, se pot accepta chiar şi destinaţii care nu se regăsesc în toate subintervalele. Ipoteza este realistă datorită faptului că, pe măsură ce se creşte numărul de subintervale, scade timpul unui subinterval, şi este posibil ca unele subintervale să nu conţină trafic datorită unor pauze temporare.

Folosirea criteriului de prezenţă pe subintervale este mai relevant decât procentul de absenţă maximă şi poate fi utilizat pentru eliminarea din amprenta de trafic, a tuturor adreselor IP care au trafic periodic.

3.3.4. Puterea amprentei de referinţă FPW (Fingerprint Power).

Puterea amprentei de referinţă, notată FPW (Fingerprint Power), este dată în principiu de doi factori:

- numărul de destinaţii IP care intră în componenţa amprentei de referinţă;
- valoarea procentelor de prezenţă corespunzătoare acestor destinaţii.

Legat de numărul destinaţiilor IP, se remarcă faptul că este de dorit ca în amprentele de referinţă să intre cât mai multe adrese IP spre care staţia trimite în mod constant pachete.

În ce priveşte valoarea procentului de prezenţă, după cum s-a menţionat deja în capitolele precedente, nu se urmăresc în primul rând valori mari ale prezenţei în trafic, cât faptul că acest trafic să fie constant. Referitor la puterea unei amprente, cu cât valorile procentelor de prezenţă sunt mai ridicate, cu atât amprentele sunt mai puternice.

Puterea amprentei de referinţă este o valoare numerică, ce dă o măsură asupra calităţii amprentei. Cu cât valoarea numerică este mai mare, cu atât calitatea este mai bună. Această calitate se evidenţiază prin următoarele caracteristici:

- stabilitate;
- credibilitate;
- viteză de validare.

Prin *stabilitate* se va înţelege capacitatea amprentei de referinţă de a valida o adresă MAC, pe o durată mare de funcţionare în timp, fără a fi puternic afectată de variaţiile temporare ale prezenţei în trafic a destinaţiilor care intră în componenţa amprentei. Cu cât numărul de destinaţii IP este mai mare, cu atât efectul perturbaţiilor care apar în dreptul uneia, este mai mic.

Efectul procentelor de prezenţă asupra stabiltăţii se reflectă prin faptul că, pentru valori mici de prezenţă, este mult mai probabil să apară momente de „pauză", în care destinaţiile respective să nu se regăsească în trafic. Dacă procentele

de prezenţă sunt mari, înseamnă că staţia evaluată are un dialog sistematic şi consistent cu destinaţiile respective, fapt care reduce probabilitatea de întrerupere a traficului.

Prin *credibilitate* se înţelege gradul de încredere a unei amprentă de referinţă în procesul de validare a unei adrese MAC.

Valori mari ale procentelor de prezenţă sporesc mult gradul de încredere, prin aceea că este evident că destinaţiile respective prezintă un trafic constant. De asemenea, numărul mare de adrese IP din componenţa amprentelor asigură că traficul constant este consistent şi variat.

Viteza de validare reprezintă timpul necesar sistemului ca să ofere un răspuns corect cu privire la validitatea adresei MAC.

Dat fiind faptul că metoda DTF realizează toate calculele la nivel de minut, este necesar ca numărul de minute până la validarea sau invalidarea staţiei, să fie cât mai mic. Perioada „tranzitorie” poate da rezultate false, sau poate varia brusc între validare şi invalidare. De aceea, se doreşte ca tranziţia să fie cât mai scurtă.

Viteza de validare creşte cu atât mai mult cu cât valorile procentelor de referinţă sunt mai mari, deoarece acestea sunt identificate din primul sau primele două, trei minute. Adresele care au prezenţă redusă necesită un timp mai îndelungat până să apară în trafic. De asemenea, cu cât numărul de destinaţii din amprentă este mai mare, cu atât recunoaşterea se realizează mai rapid.

Se va defini *puterea amprentei de referinţă* (FPW), ca fiind suma procentelor de prezenţă ale destinaţiilor IP care formează amprenta. Definiţia tine cont atât de numărul componentelor cât şi de valorile individuale ale acestora.

$$FPW = \sum_{i=1}^{TDIPref} PP_i \qquad (10)$$

Pentru o măsură mai elocventă a cuantificării şi calităţii amprentei, se poate introduce un alt parametru şi anume valoarea medie a procentelor de prezenţă în cadrul amprentei, notat cu MFPW. Cu cât valorile MFPW sunt mai mari, cu atât calitatea este mai mare.

3.4. Validarea adreselor MAC. Gradul Global de Recunoaştere ODR (Overall Degree of Recognition).

În secţiunea de faţă se va prezenta modul în care se poate valida o adresă MAC, care se pretinde a fi o staţie locală „S”. Validarea se face pe baza unei aplicaţii de monitorizare, care analizează traficul la intervale de 1 minut şi calculează *„amprenta de trafic actuală”* a fiecărei adrese MAC. Amprenta de trafic actuală se actualizează permanent, pentru toată durata de timp scursă din momentul apariţiei adresei MAC în trafic, şi până în momentul curent. Ea va fi comparată cu „amprenta de trafic de referinţă”, înregistrată în timpul de evaluare.

Cele două amprente, vor fi caracterizate pe baza mulţimilor de perechi de forma „adresă IP – procent prezenţă”.

Se notează cu „DIPR” setul de IP-uri destinaţie aferente amprentei de referinţă şi cu IPR, destinaţiile IP care intră în componenţa setului.

Cu „DIPA” se notează setul de IP-uri destinaţie aferente amprentei actuale, iar cu IPA, destinaţiile IP care intră în componenţa setului. În aceste condiţii, se pot defini matematic cele două seturi astfel:

$$DIPR = \{IPR_1, IPR_2, \ldots IPR_{TDIPref}\}$$
$$DIPA = \{IPA_1, IPA_2, \ldots IPA_{TDIPact}\}$$

Procentele de prezenţă se notează cu „PPR” pentru amprenta de referinţă şi respectiv cu „PPA” pentru amprenta actuală.

Folosind aceste notaţii, amprenta de referinţă va fi definită de perechi de forma *(IP_i, PPR_i)*, iar amprenta actuală va fi definită de mulţimea de perechi *(IP_i, PPA_i)*.

Validarea presupune compararea celor două mulţimi de perechi şi stabilirea unui grad de similaritate între cele două. Cu cât gradul de recunoaştere este mai mare, cu atât creşte probabilitatea ca adresa MAC verificată să fie chiar ceea ce se pretinde a fi.

Notând cu „DTF” amprenta de trafic *(„destination traffic fingerprint”)*, cele două amprente pot fi reprezentate astfel:

$$DTF_{referinta} = \left\{ \left(IP_1, PPR_1 \right), \left(IP_2, PPR_2 \right), \ldots, \left(IP_{TDIPref}, PPR_{TDIPref} \right) \right\} \quad (11)$$

$$DTF_{actuala} = \left\{ \left(IP_1, PPA_1 \right), \left(IP_2, PPA_2 \right), \ldots, \left(IP_{TDIPact}, PPA_{TDIPact} \right) \right\} \quad (12)$$

Calculul gradului global de recunoaştere (ODR) se poate realiza prin câteva metode. Fiecare variantă de calcul oferă performanţe diferite, aşa cum se va discuta în paragrafele 3.4.1 şi 3.4.2.

Ideea care stă la baza calculului gradului global de recunoaştere, este aceea că fiecare destinaţie IP_i din setul DIPR al amprentei de referinţă, trebuie căutată în setul DIPA al amprentei actuale. Vor fi astfel două procente de prezenţă, PPR_i şi PPA_i, aferente amprentei de referinţă şi respectiv amprentei actuale. Problema care se pune este cum se calculează un grad de recunoaştere al adresei IP_i in traficul actual? Răspunsul la această întrebare este dat în paragrafele următoare.

3.4.1. Calculul standard al Gradului Global de Recunoaştere.

În varianta standard, pentru fiecare din cele $TDIP_{ref}$ adrese IP din amprenta de referinţă, se va calcula gradul de recunoaştere RD_i (Recognition Degree) ca fiind raportul dintre procentul de prezenţă actual PPA_i şi procentul de prezenţă de referinţă PPR_i, dacă primul este mai mare ca al doilea, sau se consideră 100% în caz contrar.

Deci, *gradul de recunoaştere* al unei destinaţii, în varianta standard, poate fi scris ca:

$$RD_i = \begin{cases} \dfrac{PPA_i}{PPR_i}, daca\ PPA_i < PPR_i \\ 100, daca\ PPA_i \geq PPR_i \end{cases} \quad (13)$$

Se poate considera RD_i = 100% dacă PPA_i este mai mare decât PPR_i, deoarece gradul de recunoaştere relevă cât de mult este recunoscut un IP. Ori, dacă prezenţa actuală este mai mare decât cea de referinţă, înseamnă că adresa este recunoscută, deci se poate considera recunoaştere 100%, fără să fie nevoie să se mărească procentul de recunoaştere.

Pentru a determina Gradul Global de Recunoaştere (ODR), se va calcula media gradelor de recunoaştere individuale, ţinând cont că în amprenta de referinţă există un număr de $TDIP_{ref}$ destinaţii IP. Amprenta actuală conţine $TDIP_{act}$ perechi, cele care nu se regăsesc în amprenta de referinţă sunt ignorate în procesul de determinare a ODR.

Gradul Global de Recunoaştere standard este atunci:

$$ODR_{standard} = \frac{\sum_{i=1}^{TDIP_{ref}} RD_i}{TDIP_{ref}} \quad (14)$$

În Fig. 3.20, este reprezentată schema bloc MATLAB a algoritmului de calcul pentru Gradul Global de Recunoaştere, pentru o amprentă cu cinci componente, se află. S-a folosit Toolbox-ul Matlab [MatW]. Se observă în partea superioară a figurii, setul DIPR format dintr-o adresă IPR_1 cu procent de prezenţă 53%, o adresă IPR_2 cu procent de prezenţă 40%, o adresă IPR_3 cu procent de prezenţă 95%, o adresă IPR_4 cu procent de prezenţă 60% şi o adresă IPR_5 cu procent de prezenţă 3%.

Setul DIPA aferent amprentei actuale este prezentat în partea inferioară a figurii. Valorile de prezenţă pentru traficul actual sunt diferite de cele stabilite în referinţă. Tabelul 2 prezintă valorile procentului de prezenţă pentru amprenta de referinţă şi pentru amprenta actuală.

Tabelul 2 – Valori pentru procentele de prezenţă în amprenta de referinţă şi cea actuală pentru calculul Gradului Global de Recunoaştere standard

	PP_1	PP_2	PP_3	PP_4	PP_5
Amprenta de referinţă	53	40	95	60	3
Amprenta actuală	49	43	90	55	0

Se poate observa că valorile actuale sunt mai mici decât cele de referinţă, iar ultima destinaţie IP lipseşte complet din amprenta actuală.

Schema MATLAB permite calculul valorilor implicate în procesul determinării Gradului Global de Recunoaştere. Folosind valorile amprentei de referinţă, se poate calcula puterea amprentei, ca fiind FPW = 251. Impărţind valorile procentelor de prezenţă actuale la cele de referinţă, se obţine rapoartele notate pe diagramă cu „Raport IP_i”. Ulterior, suma rapoartelor se împarte la numărul total de componente

ale amprentei de referinţă ($TDIP_{ref}$ = 5) şi se înmulţeşte cu 100, pentru a afla valoarea Gradului Global de Recunoaştere: $ODR_{standard}$ = 76,89%.

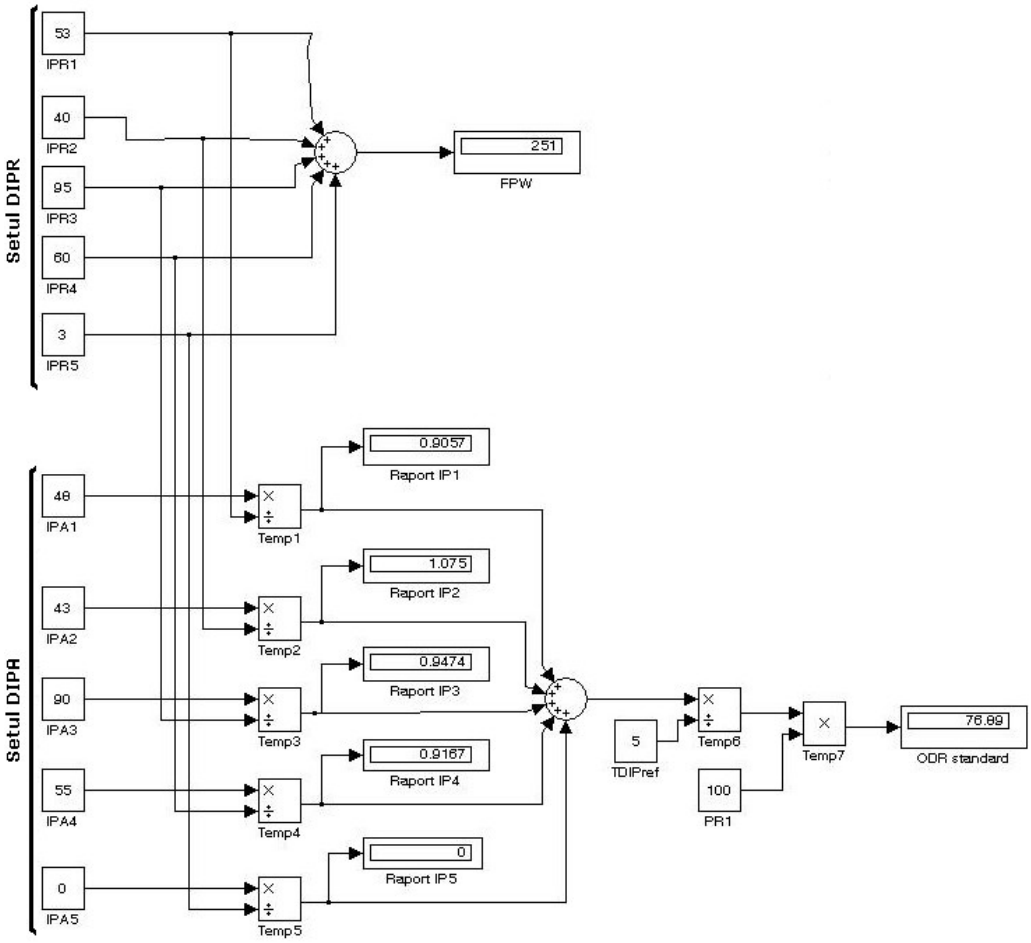


Fig. 3.20 Schema bloc MATLAB pentru determinarea ODR standard pentru o amprentă cu cinci componente

3.4.2. Calculul ponderat al Gradului Global de Recunoaştere.

Calculul standard este cel mai simplu, dar oferă performanţe mai reduse, având în vedere că în acest caz fiecare destinaţie IP are aceeaşi importanţă, indiferent cât de mare sau cât de mic este procentul de prezenţă. În felul acesta, modificări temporare ale procentului de prezenţă poate conduce la modificări drastice în valoarea Gradului Global de Recunoaştere.

Ca exemplu, se consideră o amprentă de referinţă compusă din patru adrese IP şi se presupune că la un moment dat, una din cele patru destinaţii dispare din trafic. Calculul standard determină un grad de recunoaştere egal cu 0 pentru

această destinaţie, afectându-se astfel media aritmetică. Chiar dacă celelalte trei destinaţii sunt recunoscte 100%, faptul ca a patra lipseşte va determina ca media aritmetică să fie doar 75% pentru Gradul Global de Recunoaştere, ceea ce nu este tocmai convenabil pentru ca sistemul să declare ca autentică adresa MAC.

Problema însă apare atunci când destinaţia IP care lipseşte, sau care scade mult sub valoarea de referinţă, are un procent de prezenţă mic sau foarte mic. Variaţiile acestui procent sunt mici, dar produc diferenţe mari în calculul Gradului Global de Recunoaştere.

De aceea, se impune o altă modalitate de calcul, care să ţină cont de procentele de prezenţă ale fiecărei destinaţii IP ce intră în componenţa amprentei de referinţă.

În calculul ponderat, înainte de a stabili gradul de recunoaştere al fiecărei destinaţii IP, se stabileşte o pondere cu care fiecare destinaţie IP va afecta Gradul Global de Recunoaştere.

Ponderea P_i, a destinaţiei IP_i, se calculează ca fiind raportul dintre procentul de prezenţă PP_i al destinaţiei IP_i, şi valoarea medie a puterii amprentei MFPW.

$$P_i = \frac{PP_i}{MFPW} \tag{15}$$

Gradul de recunoaştere al destinaţiei IP_i se calculează înmulţind raportul dintre procentul de prezenţă actual PPA_i şi procentul de prezenţă de referinţă, PPR_i cu ponderea P_i, cu care destinaţia IP_i intră în calculul Gradului Global de Recunoaştere.

$$RD_i = P_i * \left(\frac{PPA_i}{PPR_i}\right) \tag{16}$$

Calculul ponderat oferă avantajul că limitează efectele diferenţelor dintre valorile actuale şi cele de referinţă, astfel încât să se ţină cont de “importanţa” fiecărei destinaţii IP. Astfel, dacă o amprentă de referinţă conţine adrese IP cu procente de prezenţă mici şi adrese IP cu procente de referinţă mari, atunci variaţia Gradului Global de Recunoaştere va fi afectată mai mult sau mai puţin, funcţie de nivelul procentului de prezenţă.

Schema bloc MATLAB reprezentată în Fig. 3.21 implementează calculul Gradului Global de Recunoaştere prin ambele metode.

Exemplul este acelaşi cu cel din Fig. 3.20, astfel încât datele referitoare la procentele de prezenţă din Tabelul 2 sunt valabile şi în cazul acesta. Suplimentar, au fost incluse blocuri necesare calcului ponderat.

Pe lângă determinarea puterii amprentei de referinţă FPW=251, s-a calculat şi media MFPW=50,2, care este folosită ulterior pentru determinarea ponderilor cu care sunt luate în calcul destinaţiile IP din componenţa amprentei de referinţă. Tabelul 3 prezintă succint aceste valori.

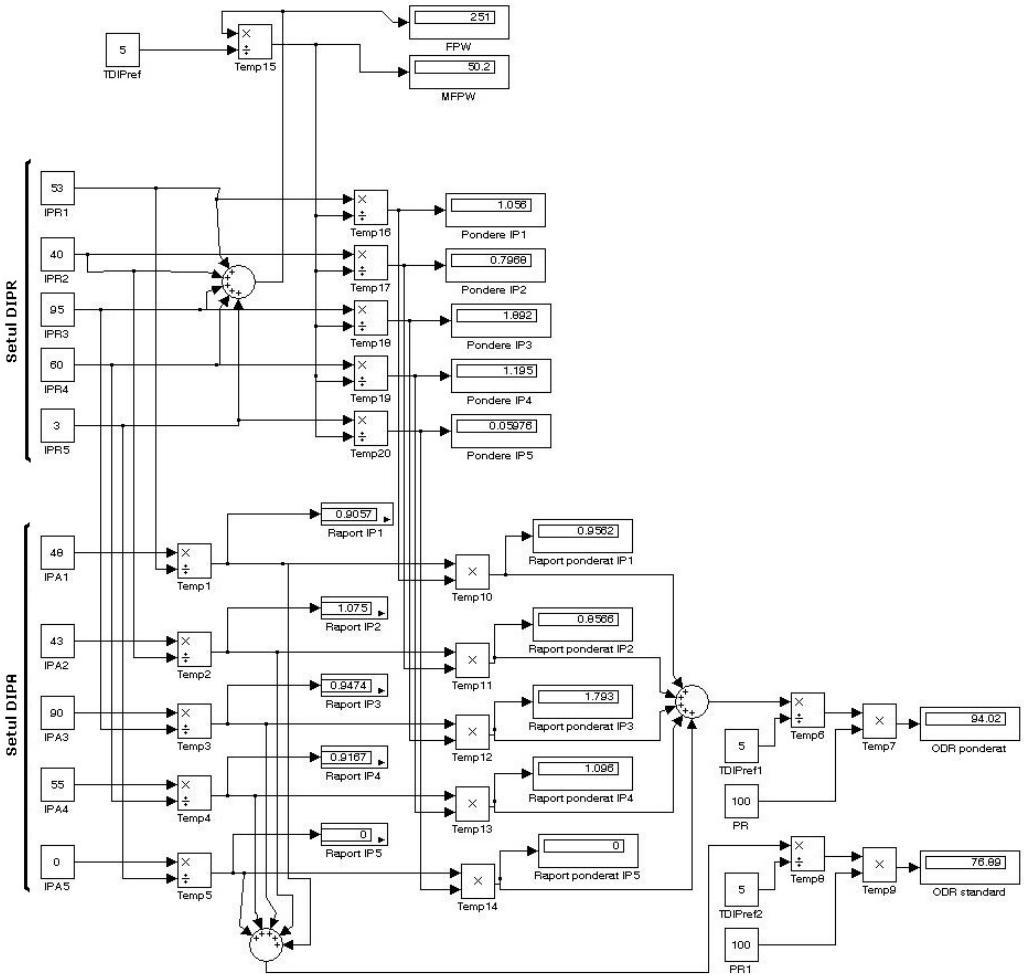


Fig. 3.21 Schema bloc MATLAB pentru determinarea ODR ponderat pentru o amprentă cu cinci componente

Tabelul 3 – Valori pentru procentele de prezenţă în amprenta de referinţă şi cea actuală pentru calculul Gradului Global de Recunoaştere ponderat

	IP_1	IP_2	IP_3	IP_4	IP_5
Procent de prezenţă (%)	53	40	95	60	3
Pondere în calculul ODR	1,056	0,7968	1,892	1,195	0,05976

Se observă faptul că destinaţia IP_3 va afecta cel mai mult Gradul Global de Recunoaştere, întrucât factorul de multiplicare este aproximativ 1.9, adică aproape dublu. Cel mai mic efect îl va produce destinaţia IP_5, care are un factor de multiplicare sub 0.1.

Faptul că destinaţia IP_5 lipseşte din amprenta actuală, a dus în calculul standard la un Grad Global de Recunoaştere de numai 76.89%, valoare relativ mică pentru ca sistemul să poată afirma cu certitudine că adresa MAC este validă.

Destinaţia IP_5 are un procent de prezenţă de numai 3%, mult mai mic decât cel al destinaţiei IP_3, care are un procent de prezenţă de 95%. Variaţiile în trafic ale destinaţiei IP_5 nu ar trebui să afecteze prea mult Gradul Global de Recunoaştere. Calculul ponderat ţine cont de acest aspect, şi asigură destinaţiei IP_5 o contribuţie la determinarea ODR numai cu un factor de 0.05976, care nu afectează semnificativ rezultatul final.

Valoarea Gradului Global de Recunoaştere în variantă ponderată este de 94.02%, mult mai corectă decât valoarea rezultată din calculul standard. Sistemul va putea recunoaşte ca validă adresa MAC chiar dacă destinaţia IP_5 lipseşte din traficul actual.

3.4.3. Concluzii asupra Gradului Global de Recunoaştere.

Gradul Global de Recunoaştere este un procent care dă o măsură a gradului de asemănare dintre amprenta de referinţă şi cea actuală. Pentru a fi siguri că adresa MAC este reală, ar fi de dorit ca valorile Gradului Global de Recunoaştere să tindă spre 100%. Dar, în practică, se poate alege o limită de recunoaştere „LREC”, care să indice faptul că adresa MAC este validată pentru ODR > LREC. De asemenea, dacă ODR scade sub LREC, chiar dacă nu mai există siguranţa că adresa MAC este reală, totuşi încă se poate afirma că „probabil” este reală, dacă ODR > LPOS, unde „LPOS” este o limită stabilită. Cu aceste notaţii, *adresa MAC evaluată* este validată după cum urmează:

$$Adresa\ MAC \rightarrow \left\{ \begin{array}{cc} REALA & ODR \geq LREC \\ PROBABIL\ REALA & LPOS \leq ODR < LREC \\ FALSIFICATA & ODR < LPOS \end{array} \right\} \quad (17)$$

Gradul Global de Recunoaştere se dă sub forma unui procent, funcţie de care se validează adresa MAC. Pentru o acurateţe sporită, este de dorit ca $TDIP_{ref}$ (numărul destinaţiilor ce intră în componenţa amprentei de referinţă) să fie cât mai mare. Fluctuaţiile prezenţei adreselor de IP în trafic pot influenţa mai puternic gradul de recunoaştere dacă numărul destinaţiilor din amprenta de referinţă este mic.

3.5. Model Matematic pentru descrierea metodei DTF.

În cadrul paragrafului, metoda DTF se formalizează printr-un model matematic ce descrie atât etapa de generare iniţială a amprentei, cât şi etapa de recunoaştere şi validare în timp real [Sas-12d].

3.5.1. Definiţii şi notaţii.

Se defineşte M_{IP}, ca fiind mulţimea tuturor adreselor IP v.4 astfel:

$$M_{IP} = \{IP_k, k \in [0, NIP]\} \quad (18)$$

unde:

- IP_k reprezintă o adresă IP din spaţiul IP v.4.
- NIP reprezintă numărul total de adrese IP din spaţiul IP v.4.

Se defineşte un pachet de date NP, captat într-un punct al reţelei, la momentul de timp T, ca fiind o mulţime de variabile ce caracterizează pachetul, de forma:

$$NP = \{T, IP_{sursa}, MAC_{sursa}, PORT_{sursa}, IP_{destinatie}, MAC_{destinatie}, PORT_{destinatie}, LUNGIME, PROTOCOL\} \quad (19)$$

unde:

- IP_{sursa} reprezintă adresa IP a staţiei care a trimis pachetul
- $IP_{destinaţie}$ reprezintă adresa IP a staţiei căreia îi este destinat pachetul
- MAC_{sursa} reprezintă adresa fizică a plăcii de reţea care a trimis pachetul
- $MAC_{destinaţie}$ reprezintă adresa fizică a plăcii de reţea căreia îi este destinat pachetul
- $PORT_{sursa}$ reprezintă portul folosit la transmiterea pachetului
- $PORT_{destinaţie}$ reprezintă portul spre care se va direcţiona pachetul pe staţia destinaţie
- LUNGIME reprezintă lungimea pachetului, exprimată în biţi
- PROTOCOL reprezintă protocolul folosit la codificarea pachetului

Se notează:

- PP_k – procentul de prezenţă a traficului către destinaţia IP_k, în traficul aferent unei adrese MAC
- LP – limita minimă de prezenţă necesară ca o destinaţie IP să intre în amprenta de trafic.
- S_{INTERV} – numărul de subintervale egale, în care se împarte intervalul de determinare a amprentei de trafic pentru a stabili gradul de constanţă în timp a traficului către o destinaţie IP
- NSI – numărul de subintervale în care există trafic catre destinaţia IP observată
- MININTERV – numărul minim de subintervale necesar pentru a admite o destinaţie IP în amprenta de trafic

3.5.2. Sistem pentru determinarea amprentei de trafic a unei adrese MAC precizate.

În Fig. 3.22 este reprezentată schema bloc a sistemului conceput pentru determinarea amprentei de trafic.

Toate blocurile vor fi analizate sistemic, prin prisma relaţiilor după care se generează ieşirile, funcţie de intrări şi de rolul blocului.

- *Intrarea sistemului: s*e defineşte ca fiind mulţimea tuturor pachetelor de reţea, care au fost captate în punctul de observare al sistemului la un moment , mulţime notată cu M_{NP} (reprezintă o submulţime a mulţimii M_{IP}).

$$u(t)=M_{NP}\{NP_k|k-variabil\ ,\ functie\ de\ t\},\ M_{NP}\subset M_{IP}. \tag{20}$$

- *Ieşirea sistemului: s*e defineşte ca fiind mulţimea tuturor perechilor de tip {Adresă_IP, Procent_Prezenţă}, notată M_{DTF}, şi care formează amprenta de trafic a adresei MAC urmărite.

$$y(t)=M_{DTF}\{P(IP_k, PP_k)|IP_k-destinaţie\ cu\ trafic\ constant\ PP_k\} \tag{21}$$

- *Subsistemul „FILTRU MAC”: p*rimeşte la intrare un set de pachete din mulţimea M_{NP}, şi le filtrează, lasând la ieşire doar acele pachete care sunt emise de către sursa cu adresa MAC urmărită. Mulţimea de pachete filtrată va fi notată cu M_{NPM}.

$$u_1(t)=u(t)=M_{NP}\{NP_k|k-variabil\ ,\ functie\ de\ t\},\ M_{NP}\subset M_{IP}. \tag{22}$$

$$y_1(t)=M_{NPM}\{NP_{kM}|MAC_k=MAC\},\ M_{NPM}\subset M_{NP}\subset M_{IP}. \tag{23}$$

- *Subsistemul „LISTA IP”:* primeşte la intrare un set de pachete din mulţimea M_{NPM}, şi generează o listă cu toate adresele de IP ale destinaţiilor.

$$u_2(t)=y_1(t)=M_{NPM}\{NP_{kM}|MAC_k=MAC\},\ M_{NPM}\subset M_{NP}\subset M_{IP}. \tag{24}$$

$$y_2(t)=M_2\{IP_k|\forall\ NP_{kM}\in M_{NPM},\ IP_k=adresa\ IP\ destinatie\ din\ NP_{kM}\} \tag{25}$$

- *Subsistemul „SUMATOR IP”:* primeşte la intrare o listă de adrese IP destinaţie şi crează o altă listă în care se memorează atât adresa IP cât şi numărul total de apariţii.

$$u_3(t)=y_2(t)=M_2\{IP_k|\forall\ NP_{kM}\in M_{NPM},\ IP_k=adr.\ IP\ dest.\ din\ NP_{kM}\} \tag{26}$$

$$y_3(t)=y_2(t-1)+u_3(t)=M_3\{a_k\ IP_k|a_k-numar\ de\ aparitii\ a\ lui\ IP_k\} \tag{27}$$

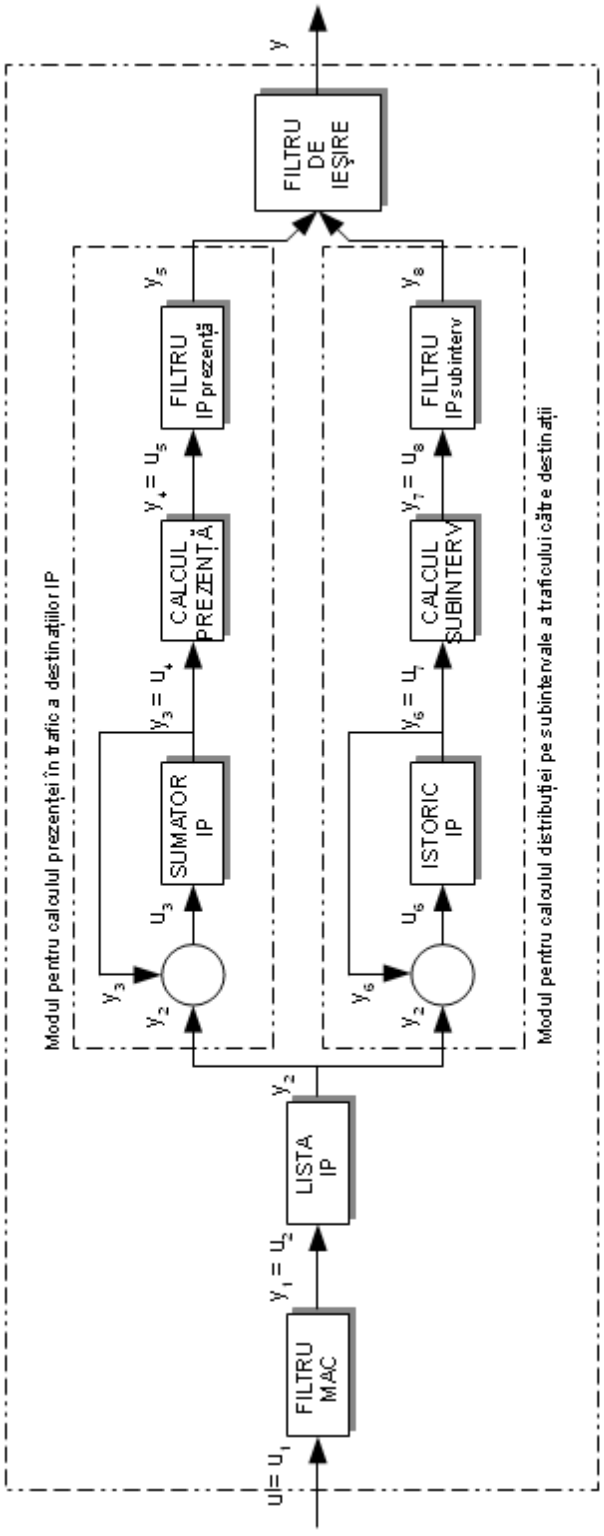


Figura 3.22 – Schema bloc a sistemului destinat determinării amprentei de trafic, prin metoda DTF

- *Subsistemul „CALCUL PREZENŢĂ”: p*rimeşte la intrare o lista adreselor IP destinaţie, împreună cu numărul lor de apariţii, şi calculează procentul de prezenţă în trafic a fiecărei destinaţii. Rezultatul este furnizat sub forma unor perechi de forma P(IP_k, PP_k), unde PP_k reprezintă procentul de prezenţă al destinaţiei IP_k.

$$u_4(t)=y_3(t)=M_3\{a_k\, IP_k | a_k - numar\, de\, aparitii\, a\, lui\, IP_k\} \tag{28}$$

$$y_4(t)=M_4\{P(IP_k, PP_k) | PP_k=\frac{a_k}{t}\} \tag{29}$$

- *Subsistemul „FILTRU $IP_{prezenţă}$”:* primeşte la intrare mulţimea M_4 a perechilor IP-prezenţă şi filtrează la ieşire doar acele perechi pentru care procentul de prezenţă depăşeşte limita minimă LP.

$$u_5(t)=y_4(t)=M_4\{P(IP_k, PP_k) | PP_k=\frac{a_k}{t}\} \tag{30}$$

$$y_5(t)=M_5\{P(IP_k, PP_k) | P(IP_k, PP_k)\in M_4\, si\, PP_k \geqslant LP\} \tag{31}$$

- *Subsistemul „ISTORIC IP”: p*rimeşte la intrare o listă de adrese IP destinaţie şi, împreună cu înregistrările anterioare, păstrează toate momentele de timp când a existat trafic către fiecare destinaţie.

$$u_6(t)=y_2(t)=M_2\{IP_k | \forall\, NP_{kM}\in M_{NPM}, IP_k=adr.\, IP\, dest.\, din\, NP_{kM}\} \tag{32}$$

$$y_6(t)=M_6\{P(IP_k, M_{ISTORIC}) | M_{ISTORIC}=\{t_k | t_k - min.\, cu\, trafic\, catre\, IP_k\}\} \tag{33}$$

- *Subsistemul „CALCUL SUBINTERVALE”: p*rimeşte la intrare înregistrările din mulţimea M_6 şi calculează numărul de subintervale în care se regăseşte trafic către fiecare destinaţie IP.

$$u_7(t)=y_6(t)=M_6\{P(IP_k, M_{ISTORIC}) | M_{ISTORIC}=\{t_k | t_k - min.trafic\}\} \tag{34}$$

$$y_7(t)=M_7\{P(IP_k, NSI_k) | NSI_k - nr\, de\, subinterv.\, pentru\, IP_k, 0\leqslant NSI_k \tag{35}$$

- *Subsistemul „FILTRU $Ip_{subinterv}$”: p*rimeşte la intrare mulţimea M_7 a perechilor IP-subintervale şi filtrează la ieşire doar acele perechi pentru care numărul de subintervale depăşeşte limita minimă MIN.

$$u_8(t)=y_7(t)=M_7\{P(IP_k, NSI_k) | NSI_k - nr\, subinterv.\, pt.\, IP_k, 0\leqslant NSI_k \tag{36}$$

$$y_8(t) = M_8\{P(IP_k, NSI_k) | NSI_k \in M_7 si\ NSI_k \geq MIN_{INTERV}\} \quad (37)$$

- *Subsistemul „FILTRU DE IEȘIRE":* primeşte la intrare două mulţimi M_5 şi M_8, şi pe baza lor crează la ieşire mulţimea M_{DTF} care este de forma mulţimii M_5, dar cu proprietatea că toate destinaţiile IP din M_{DTF}, se regăsesc atât în M_5 cât şi în M_8.

$$u_9(t) = y_5(t)\ si\ y_8(t) \quad (38)$$

$$y(t) = y_9(t) = y_5(t) \cap y_8(t) = M_{DTF}\{P(IP_k, PP_k) | IP_k \in M_5 si\ IP_k \in M_8\} \quad (39)$$

3.5.3. Sistem pentru validarea unei adrese MAC întâlnite în trafic.

În Fig. 3.23 este prezentată schema bloc a sistemului conceput pentru validarea unei adrese MAC în timp real, bazat pe metoda DTF În esenţă, schema este compusă din două module, unul reprezentând o bază de date, care stochează şi furnizează la cerere amprentele de referinţă, şi celălalt format din câteva blocuri, care au rolul de a extrage din traficul real amprenta actuală. Cele două module sunt conectate împreună pentru a compara amprenta actuală cu amprenta de referinţă, furnizându-se la ieşire Gradul Global de Recunoaştere (ODR).

Sistemul conceput realizează astfel trei funcţionalităţi distincte şi anume:

- ➔ determinarea amprentei de trafic actuale, pe baza transferului de pachete captat la intrarea sistemului
- ➔ realizarea unei baze de date care stochează amprentele de trafic ale tuturor adreselor MAC din sistem
- ➔ determinarea Gradului Global de Recunoaştere.

Subsistemele primei ramuri, care implementează prima funcţie, au fost explicitate şi formalizate în paragraful 3.6.2.

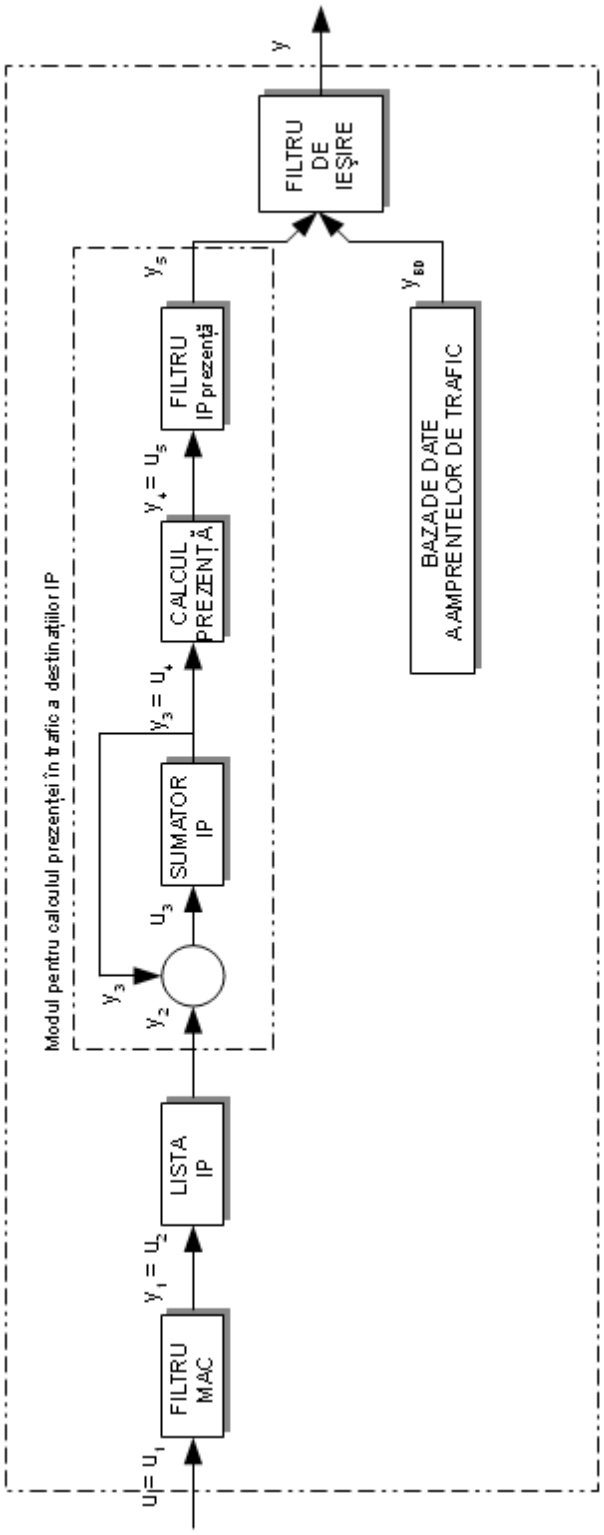


Figura 3.23 – Schema bloc a sistemului destinat pentru validarea unei adrese MAC în timp real, prin metoda DTF

Schema bloc a subsistemului care implementează determinarea Gradului Global de Recunoaştere, utilizând metoda DTF, este prezentată în Fig. 3.24.

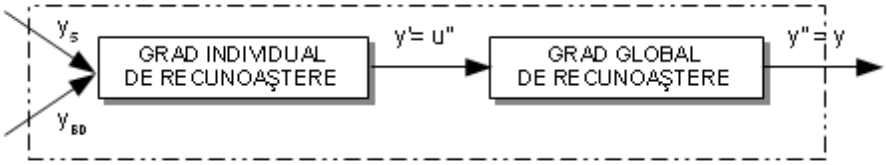


Fig. 3.24 Schema bloc pentru determinarea Gradului Global de Recunoaştere în metoda DTF

- *Subsistemul "GRAD INDIVIDUAL DE RECUNOAŞTERE":* primeşte la intrare , ieşirile y_5 şi y_{BD}, conform schemei bloc din Fig. 3.24. Cele două mărimi reprezină de fapt amprenta de trafic actuală M_{ACT} (y_5) şi amprenta de trafic de referinţă M_{REF} (y_{BD}), definite conform relaţiilor (40) şi (41). Funcţie de prezenţa actuală a componentelor din amprenta de trafic de referinţă, se stabileşte un *grad de recunoaştere individual* G_k.

$$y_5(t) = M_{ACT}\{P(IP_{kACT}, PP_{kACT})\} - amprenta\ de\ trafic\ actuala \tag{40}$$

$$y_{BD}(t) = M_{REF}\{P(IP_{kREF}, PP_{kREF})\} - amprenta\ de\ trafic\ de\ referinta \tag{41}$$

$$y'(t) = M'\{G_k | \forall\ IP_{kREF} \in M_{REF}\ si \forall\ IP_{kACT} \in M_{ACT}, cu\ IP_{kREF} = IP_{kACT}\} \tag{42}$$

$$G_k = \left\{ \begin{array}{l} 100\ daca\ PP_{kACT} \geqslant PP_{kREF} \\ \dfrac{PP_{kACT}}{PP_{kREF}}\ daca\ PP_{kACT} < PP_{kREF} \end{array} \right\} \tag{43}$$

unde notaţiile sunt cele introduse în paragraful 3.6.1

- *Subsistemul "GRAD GLOBAL DE RECUNOAŞTERE":* primeşte la intrare mulţimea M' a gradelor individuale de recunoaştere şi calculează valoarea Gradului Global de Recunoaştere, aferent celor NP_{REF} perechi din M_{REF}.

$$y(t) = y''(t) = \frac{\sum_{k=0}^{NP_{REF}} G_k}{NP_{REF}} \tag{44}$$

3.6. Dezvoltarea unui model Fuzzy pentru determinarea traficului constant.

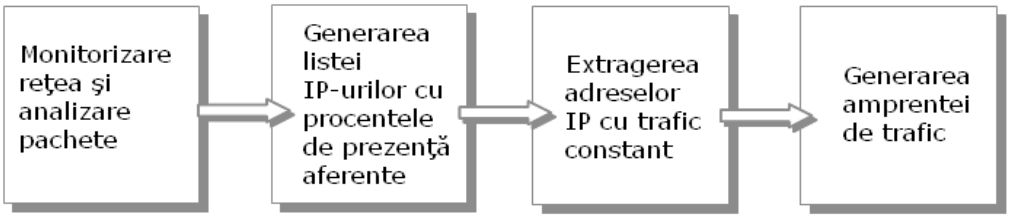


Fig. 3.25. Schema bloc pentru validarea unei adrese MAC prin metoda DTF

Funcţionarea de ansamblu a metodei DTF propuse, poate fi schematizată ca în Fig. 3.25. Traficul vehiculat prin reţea este trecut prin primul modul, care analizează pachetele şi le stochează într-o bază de date special pregătită. În a doua parte se extrage o listă cu destinţiile IP întânite în trafic şi se calculează pentru fiecare procentul de prezenţă, ca fiind raportul dintre numărul de minute în care a existat trafic către o destinaţie IP, şi numărul total de minute ale intervalului evaluat.

Următorul modul are cea mai importantă sarcină, şi anume să stabilească dacă traficul către o destinaţie IP este sau nu un „trafic constant”. La ieşire, modulul trebuie să furnizeze numai IP-urile care prezintă un trafic constant. Importanţa stabilirii corecte a traficuliu constant, derivă din aceea că rezultatele furnizate în partea aceasta vor fi folosite în procesul de validare a adreselor MAC întâlnite în trafic. Dacă sistemul permite „trecerea” spre ieşire a unei destinaţii IP care nu prezintă trafic constant, atunci procesul de identificare în timp real va da semnale greşite.

Ultimul modul preia lista destinaţiilor IP cu trafic constant şi le pune împreună sub forma unei amprente de trafic. Amprenta aceasta va purta numele de „amprentă de referinţă” şi va fi comparată cu „amprenta actuală” în procesul de identificare în timp real a adresei MAC.

Logica Fuzzy [Zad-07], [Zad-05], [Zad-96], [Zad-94], [Zad-62], poate reprezenta o abordare utilă în procesul de determinare a traficului constant.

3.6.1. Descrierea sistemului.

Fig. 3.26 prezintă un sistem Fuzzy de tip Mandami [Mam-77], [Lee-90], [Elk-94], cu patru intrări şi o ieşire [Sas-12c]. Rolul sistemului constă în evaluarea traficului spre o destinaţie IP, într-un interval de timp T.

Cele patru variabile de intrare semnifică de fapt patru diviziuni egale ale intervalului T. Pe fiecare subinterval se calculează procentul de prezenţă al traficului către destinaţia IP, valoarea rezultată fiind transmisă spre intrarea corespunzătoare a sistemului mandami.

Datorită faptului că fiecare intrare este un subinterval al intervalului T, rezultă că toate cele patru intrări vor avea aceeaşi funcţionalitate. Singura diferenţă este dată de faptul că fiecare intrare se referă la un interval de timp diferit.

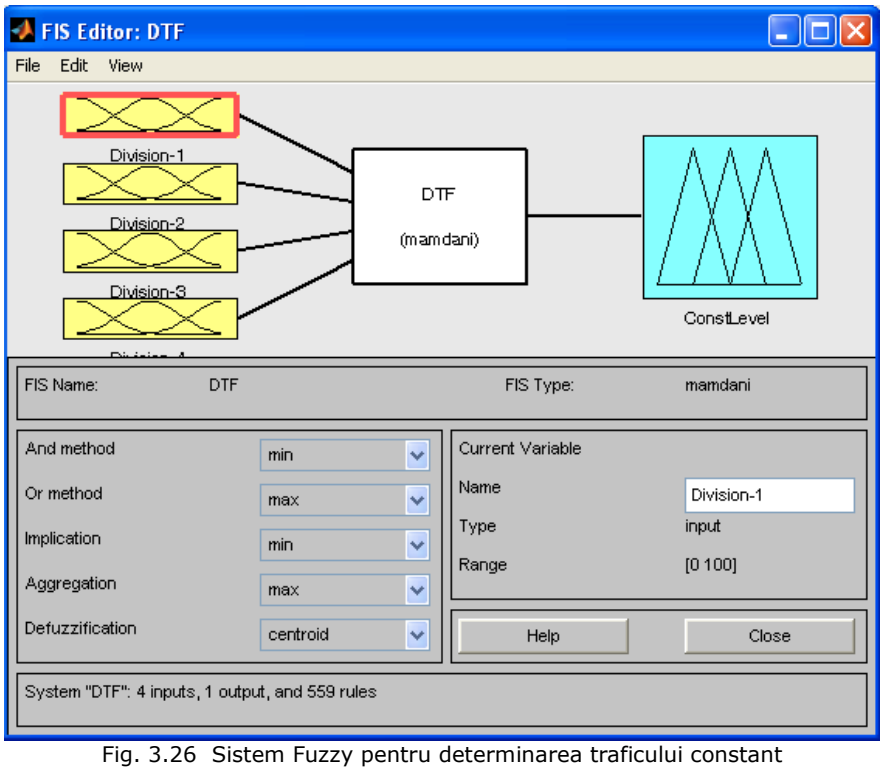


Fig. 3.26 Sistem Fuzzy pentru determinarea traficului constant

Ieşirea sistemului este dată sub forma unei valori, care dă o măsură a gradului de constanţă a traficului spre destinaţia IP evaluată. Cu cât valorile de ieşire sunt mai mari, cu atât certitudinea este mai mare cu privire la traficul constant.

Pentru fiecare subinterval din cele 4, metoda DTF calculează procentul de prezenţă şi îl trimite ca intrare în sistemul mandami. Acesta evaluează intrările şi generează ieşirea corespunzătoare acestora.

3.6.2. Structura intrărilor.

Domeniul de valori pentru variabilele de intrare, prezentate grafic în Fig. 3.27, este considerat ca fiind [0..100], întrucât fiecare variabilă reprezintă procentul de prezenţă în trafic a adresei IP evaluate.

Valorile funcţiilor de apartenenţă au fost stabilite astfel:

- „Continuu” - o funcţie de tip „trapmf” caracterizată prin următorii parametrii: [85 95 100 100]. Se încadrează în categoria aceasta traficul care prezintă o rată de prezenţă atât de ridicată încât poate fi considerată continuuă.

- „Prezenţă-Ridicată” - o funcţie de tip „trimf” caracterizată prin următorii parametrii: [55 75 95]. Se încadrează în categoria aceasta traficul care prezintă o rată de prezenţă ridicată.
- „Prezenţă-Medie” - o funcţie de tip „trimf” caracterizată prin următorii parametrii: [30 50 70]. Se încadrează în categoria aceasta traficul care prezintă o rată de prezenţă medie.
- „Prezenţă-Mică” - o funcţie de tip „trimf” caracterizată prin următorii parametrii: [5 25 45]. Se încadrează în categoria aceasta traficul care prezintă o rată de prezenţă mică.
- „Absent” - o funcţie de tip „trapmf” caracterizată prin următorii parametrii: [0 0 5 15]. Se încadrează în categoria aceasta traficul care prezintă o rată de prezenţă atât de mică încât poate fi considerat absent.

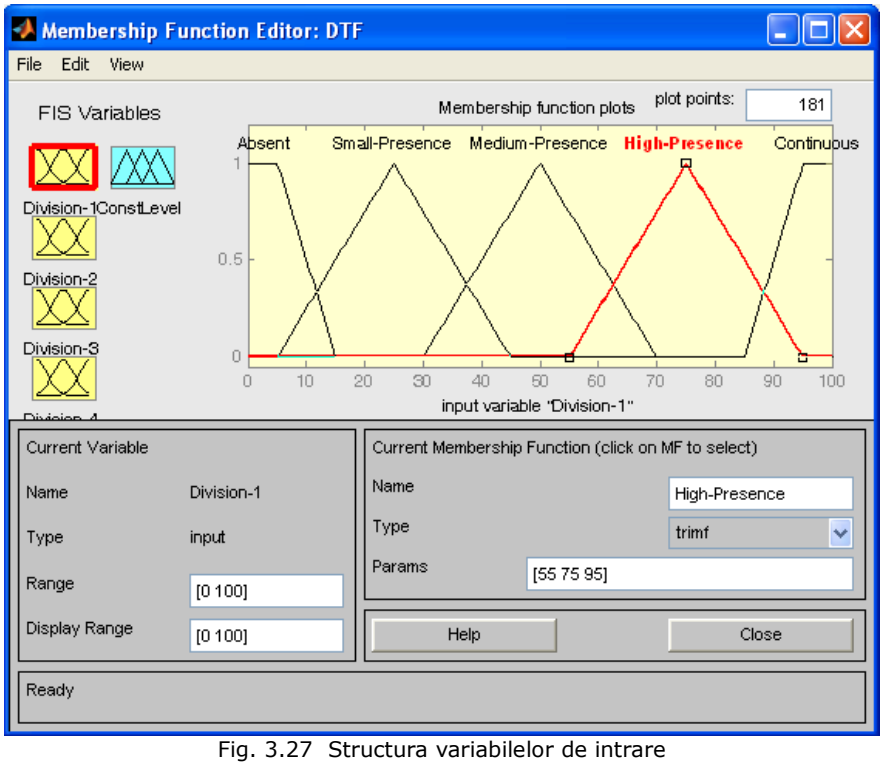


Fig. 3.27 Structura variabilelor de intrare

Cu alte cuvinte, ca să se poată afirma despre un trafic că este „continuu”, ar trebui să prezinte rate de prezenţă în jurul valorii de 100%. Dacă rata scade in jurul valorii de 75%, traficul va avea „prezenţă ridicată”. La 50% deja se vorbeşte despre „prezenţă medie”, iar la 25% despre „prezenţă mică”. Valorile foarte mici vor fi considerate ca „absent”.

În practică, procentul de prezenţă nu apare sub forma unor valori discrete de tipul 25% / 50% / 75% / 100%, ci mai degrabă sub forma unor valori repartizate pe tot intervalul între 0% şi 100%. Din cauza aceasta, pentru a putea caracteriza în mod deplin traficul, au fost alese funcţii de tip „trimf” şi „trapmf”.

Este important de menţionat faptul că metoda DTF nu este interesată numai de ratele foarte ridicate ale procentelor de prezenţă. Sigur, este de dorit ca amprentele de trafic să conţină cât mai multe adrese care prezintă trafic cu rate mari de prezenţă, însă nu este absolut necesar. Cel mai important este ca traficul să fie constant, fără să fie obligatorie o rată mare de prezenţă.

Chiar dacă metoda DTF nu are nevoie neapărat de adrese IP cu rate mari de prezenţă în trafic, totuşi, „puterea" unei amprente de trafic este influenţată de ratele de prezenţă.

Prin *„putere"* se înţelege cât de rapid şi cât de sigur este procesul de identificare în timp real. Componentele cu rate mari de prezenţă conduc la o detecţie rapidă şi cu un grad foarte mare de încredere. Componentele cu rate mici de prezenţă vor conduce şi ele la răspunsul final, dar vor avea nevoie de un timp mai mare.

3.6.3. Structura ieşirii.

Ieşirea sistemului, prezentată grafic în Fig. 3.28 trebuie să reflecte natura constantă, sau non-constantă a traficului către o anumită destinaţie IP.

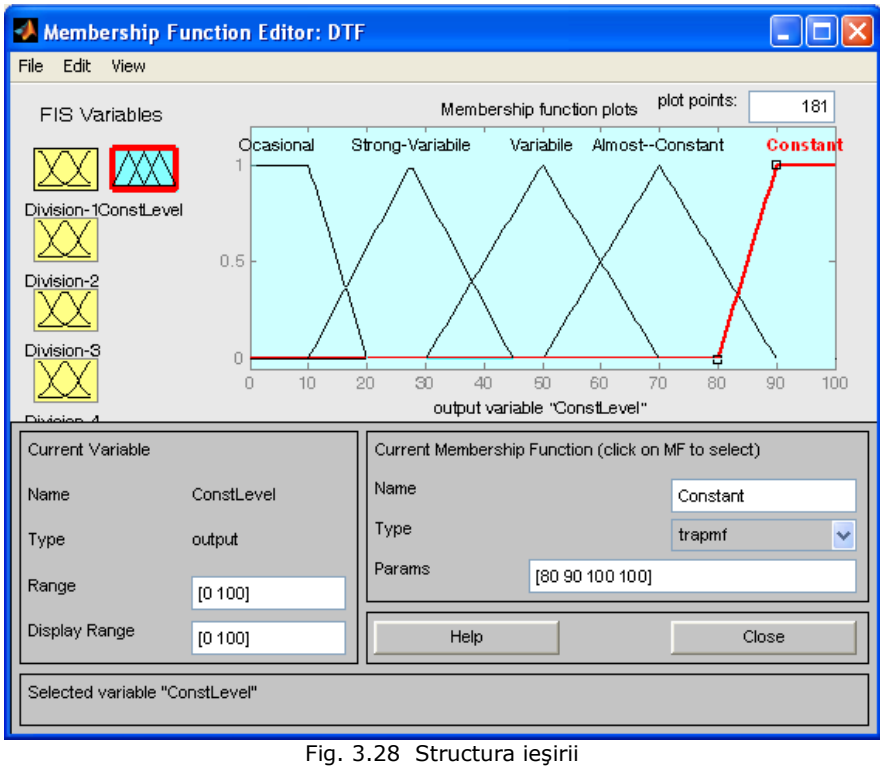


Fig. 3.28 Structura ieşirii

Din acest motiv, funcţiile de apartenenţă au fost definite în felul următor:

- „Constant” - o funcţie de tip „trapmf” caracterizată prin următorii parametrii: [80 90 100 100], reprezintă un trafic care este constant în toate cele 4 subintervale de intrare.
- „Aproape Constant” - o funcţie de tip „trimf” caracterizată prin următorii parametrii: [50 70 90], reprezintă un trafic care prezintă anumite variaţii pe parcursul celor 4 subintervale de intrare, dar care poate fi totuşi considerat aproape constant.
- „Variabil” - o funcţie de tip „trimf” caracterizată prin următorii parametrii: [30 50 70], reprezintă un trafic care deşi apare în toate cele 4 subintervale, prezintă variaţii mari pe parcursul celor 4 subintervale şi drept urmare nu poate fi folosit ca parte din amprentele de trafic.
- „Puternic Variabil” - o funcţie de tip „trimf” caracterizată prin următorii parametrii: [10 27.5 45], reprezintă un trafic care prezintă variaţii foarte mari pe parcursul celor 4 subintervale de intrare.
- „Ocazional” - o funcţie de tip „trapmf” caracterizată prin următorii parametrii: [0 0 10 20], reprezintă un trafic care prezintă variaţii foarte mari pe parcursul celor 4 subintervale de intrare

3.6.4. Definirea regulilor.

Regulile definite trebuie să genereze ieşirea funcţie de intrari (Fig. 3.29). Fiecare intrare reprezintă un procent de prezenţă în trafic a unei adrese IP. Sunt patru intrări deoarece intervalul evaluat a fost împărţit în patru subintervale egale.

Este important faptul că „trafic constant” nu implică în mod obligatoriu să existe rate mari de prezenţă, ci mai degrabă să fie variaţii mici de-a lungul celor patru subintervale egale. În aceste condiţii, regulile au fost generate pe baza unor convenţii, de tipul celor menţionate mai jos.

Sistemul va da la ieşire „Constant” dacă se situează într-unul din următoarele cazuri:

- toate cele 4 intrări sunt la fel;
- trei intrări sunt de nivel „L” iar una este de nivel „L-1” sau „L+1”;
- două intrări sunt de nivel „L” iar două sunt de nivel „L-1” sau „L+1”;
- alte câteva cazuri particulare.

Dacă nu s-au îndeplinit condiţiile anterioare, dar valorile sunt aproape de acestea, atunci ieşirea va fi „Aproape Constant”. De exemplu, trei intrări sunt pe nivel „L” şi o intrare este pe nivel „L-2” sau „L+2”. (Multe alte exemple ar putea fi menţioante la această categorie).

Pentru o ieşire de tip „Variabil”, diferenţele dintre nivele sunt mult mari. Tot aici se includ şi cazurile în care există trafic cu rate mari în 3 nivele, dar nu există în al patrulea. Pentru ca ieşirea să fie declarată „Puternic Variabilă”, diferenţele sunt mult mai mari, iar pentru „Ocazional” sunt necesare două sau trei intervale cu intrări declarate absent.

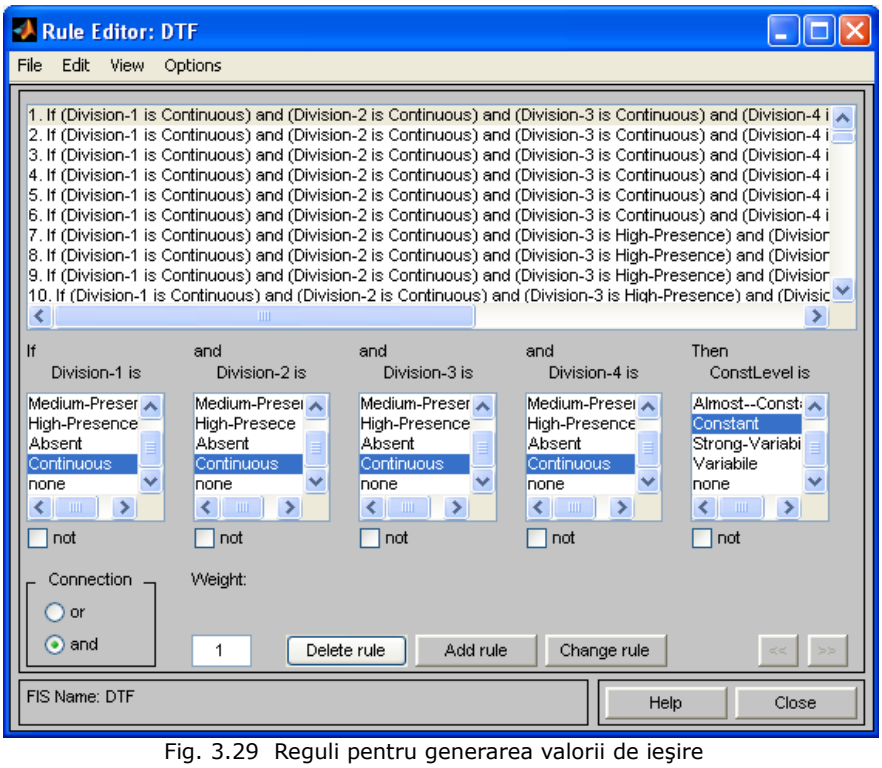


Fig. 3.29 Reguli pentru generarea valorii de ieşire

Toate regulile date ca exemple, au ca scop o înţelegere a modelului în sine. Fiecare caz a fost analizat în mod individual şi i s-au atribuit regulile potrivite.

3.6.5. Utilizarea modelului fuzzy în determinarea amprentei de referinţă a metodei DTF.

Generarea amprentei de referinţă afectează puternic performanţele procesului de recunoaştere în timp real a staţiilor din reţea. De aceea este foarte important să se identifice cu exactitate adresele IP care prezintă în timp un trafic constant.

Printr-un calcul standard, calculul procentului de prezenţă a unei adrese IP se realizează prin împărţirea numărului total de minute în care a existat trafic spre destinaţia IP, la numărul total de minute evaluate. Chiar dacă relaţia este una simplă, rezultatul nu reflectă neapărat cât de „constant” este traficul pe perioada luată în considerare.

Aplicarea logicii fuzzy în procesul de generare a amprentei de referinţă, creşte performanţele procesului de identificare în timp real, şi evidenţiază mai elocvent adresele IP care prezintă un trafic constant, faţă de cele cu trafic variabil. Pentru a înţelege şi mai bine fenomenul, ne vom ocupa în continuare de câteva cazuri în care modelarea fuzzy conduce la oţinerea unor rezultate rapide şi de calitate.

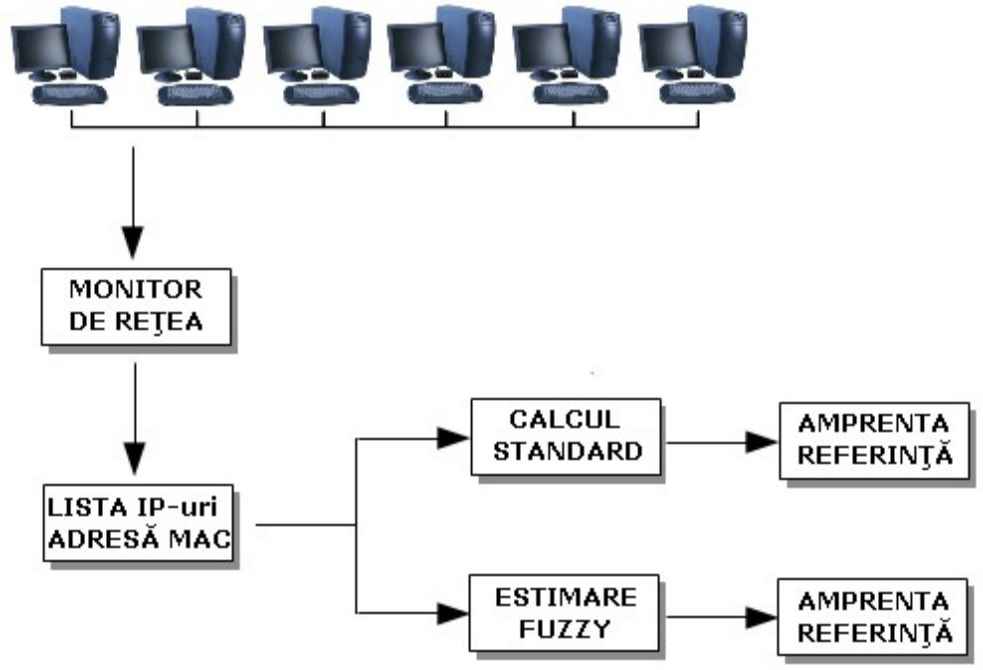


Fig. 3.30 Generarea amprentei de referinţă folosid calculul standard, respectiv modelarea Fuzzy

Fig. 3.30 prezintă grafic generarea amprentei de referinţă în două variante: calcul standard şi modelare fuzzy. Reţeaua este monitorizată permanent într-un anumit punct, extrăgându-se lista adreselor IP spre care au plecat pachetele. Această listă are ataşat în dreptul fiecărei adrese IP, numărul de minute în care a existat trafic spre ea.

Lista este ulterior trimisă ca intrare spre cele două module diferite de estimare a traficului constant. Estimatorul standard utilizează ca relaţie de calcul raportul menţionat anterior, în timp ce estimatorul fuzzy aplică regulile definite în sistemul mandami. La final se obţin două amprente de referinţă, fiecare generată din modulele respective.

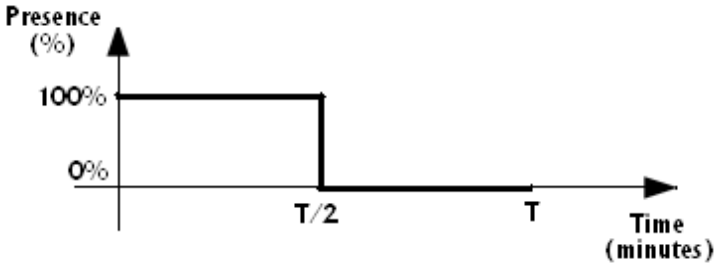


Fig. 3.31 Prezenţă 100% în prima parte a intervalului şi 0% în cea de-a doua parte

În Fig. 3.31 s-a descris unul din cazurile „înşelătoare” pentru calculul standard. După cum se poate vedea, destinaţia IP evaluată prezintă trafic în fiecare

minut din prima jumătate a intervalului, după care se opreşte complet şi este total absent în cea de-a doua parte. Printr-o estimare standard, procentul de prezenţă rezultat va fi de 50%, valoare care va duce la concluzia eronată, că această destinaţie IP ar avea trafic constant. În contrast, modelul Fuzzy identifică absenţa din cea de-a doua parte şi generează ca răspuns un trafic „variabil”.

Traficul „izolat”, caracterizat prin apariţii sporadice de-a lungul axei timpului, va fi identificat de asemenea corect de către modelarea fuzzy, prin împărţirea intervalului în cele patru subintervale şi verificarea fiecăruia în parte. Rezultatul va fi dat sub forma unui trafic „variabil”, „puternic variabil” sau „ocazional”.

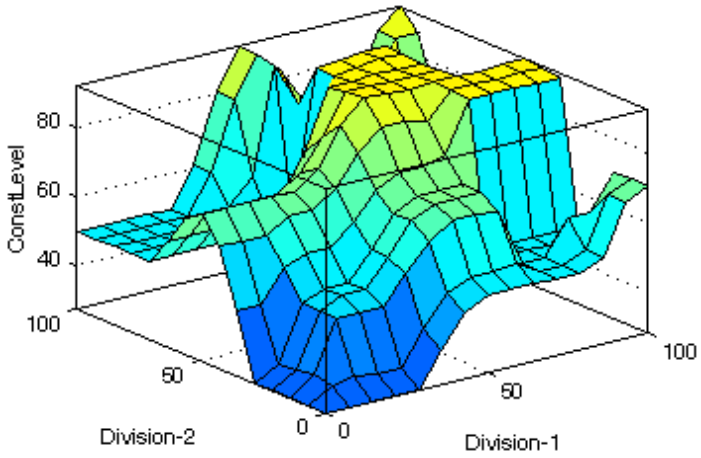


Fig. 3.32 Perspectiva de suprafaţă a modelului fuzzy

Fig. 3.32 prezintă perspectiva de suprafaţă a modelului fuzzy, utilizând Fuzzy Logic Toolbox.

Rezultatele obţinute prin modelarea fuzzy sunt estimate mai corect decât printr-un calcul standard, în principal, datorită faptului că sunt evaluate 4 subintervale în locul intervalului întreg.

Amprentele de referinţă generate printr-o estimare standard pot accepta destinaţii IP care nu prezintă în mod real un trafic constant. Aceasta înseamnă că în procesul de validare a adreselor MAC, sistemul va căuta adrese IP care vor lipsi din amprentele actuale extrase din traficul curent. Lipsa lor va conduce la semnalarea adresei MAC ca fiind falsificată, deşi acest lucru poate să nu fie adevărat. Chiar dacă totuşi traficul conţine adresele IP din amprenta de referinţă, stabilitatea procesului de recunoaştere prin estimare standard este mai slabă şi poate conduce la alarme false cu privire la integritatea adresei MAC avute în vedere. Alarmele false sunt procesate de către nivelele superioare decizionale şi s-ar prea putea ca pe baza lor să se aplice eronat anumite măsuri.

Cooperarea dintre regulile modelului fuzzy pentru generarea răspunsului final, poate fi observată cu ajutorul „Rule Viewer” din Fig. 3.33. Se poate observa cum se aplică diverse reguli funcţie de cazul concret. Mai mult, se poate folosi „Rule Viewer” pentru a testa manual modelul pentru diverse valori de intrare. Putem seta printr-un simplu click, valoarea pe care o dorim la fie care intrare, iar la ieşire vom observa valoarea rezultată.

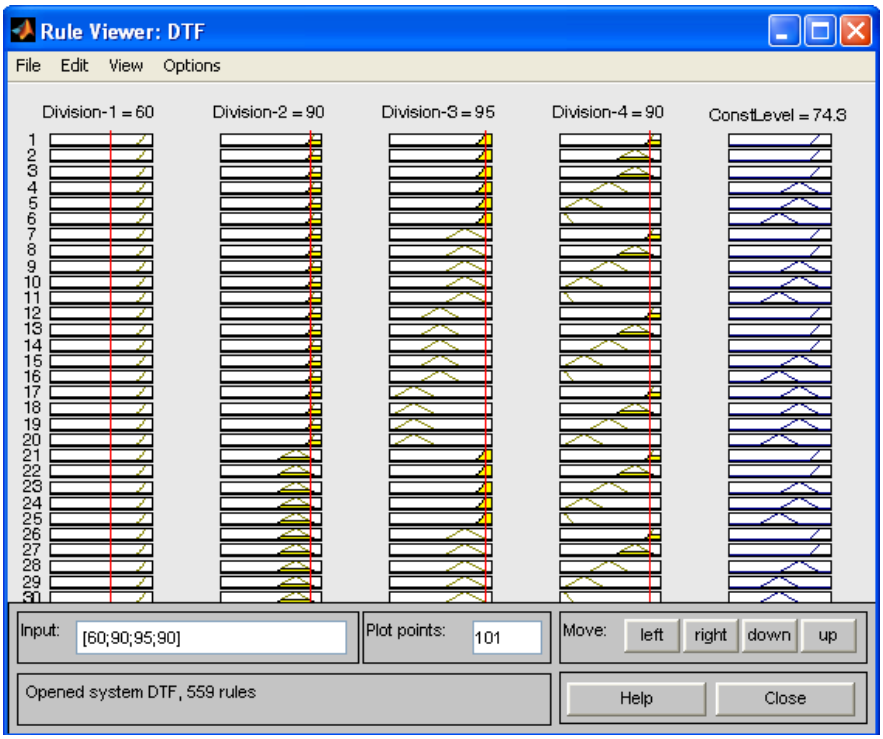


Fig. 3.33 „Rule Viewer” pentru modelul fuzzy

3.6.6. Concluzii asupra modelării fuzzy.

Utilizarea modelării logicii fuzzy în cadrul metodei DTF reprezintă o abordare atractivă în momentul în care se doreşte să se afle dacă traficul către o destinaţie IP poate sau nu să fie considerat constant. Rezultatul este dat sub forma unei valori procentuale, şi de asemenea sub forma unui răspuns de genul „constant”, „aproape constant”, „variabil”, „puternic variabil”, „ocazional”.

Cel mai important segment al metodei DTF este generarea amprentelor de referinţă. Dacă o amprentă conţine adrese IP care nu prezintă trafic constant, atunci procesul de recunoaştere va fi alterat şi se vor genera alarme false. De aceea este foarte important să existe certitudinea că în amprentele de referinţă nu intră decât adrese IP cu trafic constant.

Varianta de estimare standard utilizează în calcul raportul dintre numărul total de minute în care avem trafic către o anumită adresă IP şi numărul total de minute. Modelul fuzzy împarte intervalul în patru subintervale egale şi estimează procentul de prezenţă pe fiecare subinterval în parte. Ieşirea este generată cu ajutorul setului de reguli definit în sistem. Simulările în Matlab oferă o serie de informaţii utile în procesul de monitorizare şi vizualizare a diverselor aspecte prezente în sistem. Fiecare bloc din diagramele MATLAB poate fi uşor configurat.

Se poate deci afirma, în concluzie, că aplicarea logicii fuzzy în metoda DTF creşte performanţele acesteia, oferind o soluţie rapidă şi eficientă pentru diferenţierea dintre traficul constant şi cel variabil. „Puterea" unei amprente de trafic este dată de gradul de „constanţă" pe care îl au componentele sale, adică traficul către diverse adrese IP. Amprentele de trafic mai „puternice" conduc la rezultate mai bune. Logica fuzzy s-a dovedit a fi un instrument important în acest proces.

3.7. Servicii/tehnologii care favorizează utilizarea metodei DTF.

După cum s-a prezentat în lucrarea de faţă, metoda DTF se bazează pe traficul emis de o staţie către diverse destinaţii. Cu cât se reuşeşte să se genereze amprente de trafic mai puternice, cu atât creşte gradul de recunoaştere globală şi viteza de identificare a staţiei.

În practică, există o serie de aspecte care favorizează aplicarea metodei DTF întrucât crează un cadru propice pentru obţinerea amprentelor de trafic de mare putere. Metoda DTF este necesară cu predilecţie companiilor, mai ales a celor mari. Persoanele individuale, sau companiile mici, cu număr redus de calculatoare, nu au nevoie de sisteme avansate de detecţie a pătrunderilor neautorizate. Companiile mari sunt cele care vor folosi în primul rând metode de securizare.

Modul în care se desfăşoară activităţile în companiile mari, dezvoltă un cadru foarte prielnic pentru ca metoda DTF să poată fi aplicată cu succes. În continuare, se vor prezenta câteva astfel de cazuri, care favorizează aplicarea metodei DTF.

Utilizarea aplcaţiilor software de tip ERP.

Companiile mari utilizează în mod curent aplicaţii software de tip ERP, pentru gestionarea întregii activităţi prestate de către firmă, de la actele contabile până la urmărirea producţiei, calitate şi managementul resurselor umane. Literatura prezintă diverse implementări, cu avantajele şi dezavantajele aferente [Kal-12], [Rit-08], [Esf-10], [Zha-09].

Fiecare persoană care lucrează în firmă, se identifică în sistem pe baza numelui de utilizator şi a parolei, după care accesează resursele puse la dispoziţie de sistemul ERP.

Un astfel de sistem ERP crează un trafic permanent între staţiile client şi server. Traficul nu poate fi realizat însă decât de utilizatori autentificaţi în sistemul ERP. În felul acesta, se crează premisele necesare unei amprente de trafic care să aibă în componenţa ei traficul generat prin intermediul sistemului ERP.

Comunicarea între angajaţi prin intermediul unor servere de e-mail proprii.

Serviciile de e-mail sunt vitale în comunicarea dintre angajaţii unei companii. Procesul tehnologic se bazează pe dialogul purtat pe scară ierarhică între diverse persoane implicate în activitatea zilnică.

Pentru gestionarea comunicării prin e-mail, companiile nu folosesc adresele de mail private ale angajaţilor, ci adresele de mail din cadrul companiei, gestionate de serverul de mail. Orice comunicare prin e-mail va însemna un trafic între staţiile din reţea, şi serverul de mail. Şi în cazul acesta, autentificarea pe serverele de mail împiedică un intrus să formeze un trafic asemănător cu staţiile autentificate.

Metoda DTF va identifica traficul generat prin intermediul serverelor de mail şi va folosi destinaţiile acestea pentru a crea amprente de trafic de mare calitate.

Utilizarea reţelelor VPN.

Un alt aspect important care favorizează aplicarea metodei DTF, este dat de utilizarea reţelelor de tip VPN. Interconectarea diverselor staţii pentru crearea unor reţele virtuale, pune din nou la dispoziţie un trafic constant care se poate folosi pentru realizarea amprentelor de trafic.

Reţelele VPN se folosesc de certificate de securitate instalate pe calculatoarele client, certificate care nu permit accesul decât pentru utilizatorii autorizaţi. Pătrunderea în sistem a unui calculator care nu deţine un certificat valid, va împiedica staţia respectivă să formeze trafic asemănător cu staţia autorizată.

Virtualizare şi Cloud.

Noul trend care revoluţionează din nou domeniul IT este cel legat de virtualizare şi cloud. Costuri enorme pentru achiziţionarea şi întreţinerea echipamentelor, sunt înlocuite cu variante reduse de abonamente lunare, plătite unor companii care oferă sisteme de calcul virtualizate, de la staţii de lucru şi până la servere superperformante.

Prin virtualizare, se crează un trafic permanent între staţiile de lucru şi serverele de virtualizare. Traficul acesta este foarte util în procesul de generare a amprentei de referinţă din cadrul metodei DTF.

3.8. Concluzii

În continuare sunt punctate câteva dintre elementele importante care caracterizează metoda DTF.

În primul rând, metoda DTF detectează un intrus care vine în reţeaua locală, în locul staţiei autorizate, ceea ce este important întrucât, aşa cum reiese din Tabelul 4 şi respectiv din analiza critică a metodelor prezentate în capitolul precedent, metodele clasice din literatură nu acoperă, sau acoperă doar parţial problematica aceasta.

Metoda DTF oferă o soluţie practică şi eficientă în detecţia pătrunderilor neautorizate, care sunt iniţiate din interiorul reţelei. Faptul că metoda îşi bazează calculele pe amprentarea traficului emis de o anumită staţie, înseamnă că nu contează locaţia intrusului întrucât el va fi depistat după activitatea pe care o generează în reţea şi nu după alţi parametrii care să fie influenţaţi de locaţie.

Chiar dacă intrusul îşi lansează atacul înlocuind staţia autorizată cu alta proprie, nu va putea să recreeze amprenta de trafic aferentă staţiei autorizate, ceea ce va determina o asemenea valoare pentru Gradul Global de Recunoaştere, încât sistemul va raporta traficul ca aparţinând unei staţii cu adresă MAC falsificată.

Un alt aspect important metodei DTF este acela că se poate aplica atât în reţele fără fir, cât şi în reţele cu fir. Modalitatea de amprentare prezentată în cadrul metodei DTF nu foloseşte caracteristici ale comunicaţiei wireless sau wired, având drept urmare faptul că se va putea folosi la fel de bine în oricare din cele două, şi chiar mai mult, în reţele care conţin părţi wireless şi părţi wired.

Metoda DTF nu necesită instalarea unui software pe calculatorul client. Este adevărat că sistemul foloseşte un software de monitorizare al reţelei, dar acesta este instalat pe un anumit calculator din reţea, fără să aibă nevoie de module care să funcţioneze pe clienţii din reţea.

Faptul că o metodă cere prezenţa unor programe instalate pe calculatoarele care vin în reţea, atrage după sine o serie de limitări. După cum s-a menţonat deja, în metoda DTF nu este necesară intervenţia pe calculatoarele client, ceea ce reprezintă un real avantaj.

Problema suprapunerii traficului a fost discutată în capitolul precedent, menţionându-se că există metode care nu se pot aplica decât în cazul în care atât sursa autorizată, cât şi cea neautrizată funcţionează concomitent.

În ceea ce priveşte metoda DTF, această limitare nu există, metoda va detecta intrusul şi în absenţa traficului staţiei autorizate. Amprenta de trafic a atacatorului fiind diferită de cea a staţiei autorizate, sistemul de monitorizare va semnala prezenţa unei surse falsificate.

Dacă cele două surse funcţionează concomitent, metoda DTF continuă să dea rezultate chiar mai rapid, întrucât va sesiza două adrese IP asociate cu acelaşi MAC. Există posibilitatea ca o staţie să fie conectată la două sau mai multe reţele prin interfeţe diferite, dar lucrul acesta s-ar observa prin clase IP difeirte.

Pentru metoda DTF, mobilitatea staţiilor şi trecerea lor dintr-o subreţea în alta nu reprezintă un impediment. Calculele nu se bazează pe date care sunt influenţate de localizarea geografică, chiar dacă lucrul acesta presupune trecerea dintr-o reţea wired într-o reţea wireless sau invers. Calculele şi estimările sunt realizate în metoda DTF la nivel de minut, ceea ce înseamnă că eventualele întârzieri provocate de trecerea într-o reţea wireless, nu vor putea afecta atât de mult amprenta actuală de trafic.

Metoda DTF oferă rezultate favorabile şi în acest caz. Amprentarea unei staţii nu ţine cont de modelul echipamentului, ci de activitatea sa în reţea. De aici rezultă faptul că un atacator, chiar dacă aduce un echipament identic cu cel falsificat, nu poate reproduce cu exactitate mulţimea aplicaţiilor software instalate pe staţia originală, şi nici comportamentul acesteia în reţea.

Aplicabilitatea metodei DTF se referă în mod direct la echipamente de tip Desktop / Laptop. Scopul metodei este acela de a identifica persoane neautorizate, care încearcă să pătrundă în reţea cu un calculator care nu ar avea drepturile necesare.

În finalul capitolului, se poate afirma cu certitudine faptul că metoda DTF aduce o serie de avantaje care sunt demne de remarcat. Mai mult, ţinând cont de Tabelul 4, importanţa metodei DTF derivă nu numai din faptul că are avantaje, ci şi din faptul că are capacitatea să grupeze la un loc o serie de avantaje pe care le întâlnim în mod disparat la alte metode.

4. TOOLBOX SOFTWARE PENTRU STUDIUL METODEI DTF

Capitolul reprezintă o contribuţie a autorului în domeniul simulatoarelor de reţea. În prezent există o serie de simulatoare de reţea disponibile, după cum reiese din: [Mil-12], [Reh-12], [Kum-12], [Bar-12], [Nak-12], [Vuc-11], [Nao-10], [Bor-09], [Bul-09], [Mel-09], [Nis-09], [Car-08], [Gal-08], [Wan-08], [Esc-08], [Gon-07], [Fek-07], [Hos-07], [Drz-07].

Fiecare simulator are anumite particularităţi şi oferă o serie de avantaje, apropiindu-se cât mai mult de funcţionarea reală. Motivul însă pentru care a fost necesară conceperea unui nou simulator, este acela că studiul metodei DTF necesită acces direct la pachetele din reţea şi prelucrarea acestora într-o manieră particulară, specifică metodei DTF [Sas-12b], [Sas-10b].

Toolbox-ul software este alcătuit în esenţă din două aplicaţii software dezvoltate de către autor. Pe de-o parte este vorba despre "Packet Recorder", program destinat captării packetelor din reţea şi pe de altă parte de "Network Detector", program care reprezintă simulatorul de reţea şi care foloseşte datele accumulate de "Packet Recorder".

4.1. Packet Recorder.

Este o aplicaţie software dezvoltată cu scopul de a captura traficul la nivelul plăcilor de reţea. Are la bază driverul WinPcap, ce poate fi descărcat gratuit de pe Internet.

Aplicaţia rulează permanent şi captează pachetele care circulă pe fiecare placă de reţea identificată. Pachetele sunt înregistrate în MySQL, sub forma unor înregistrări cu următoarele câmpuri:

- identificatorul plăcii de reţea;
- momentul captării pachetului;
- adresa IP a sursei;
- adresa MAC a sursei;
- adresa IP a destinaţiei;
- adresa MAC a destinaţiei;
- numărul de pachete.

Fig. 4.1 prezintă interfaţa aplicaţiei software, organizată pe câteva zone distincte. În partea superioară se afişează lista interfeţelor de reţea disponibile pe staţia monitorizată. Pentru fiecare în parte se precizează numele, numărul de pachete captate, numărul de conectări şi starea actuală a monitorizării.

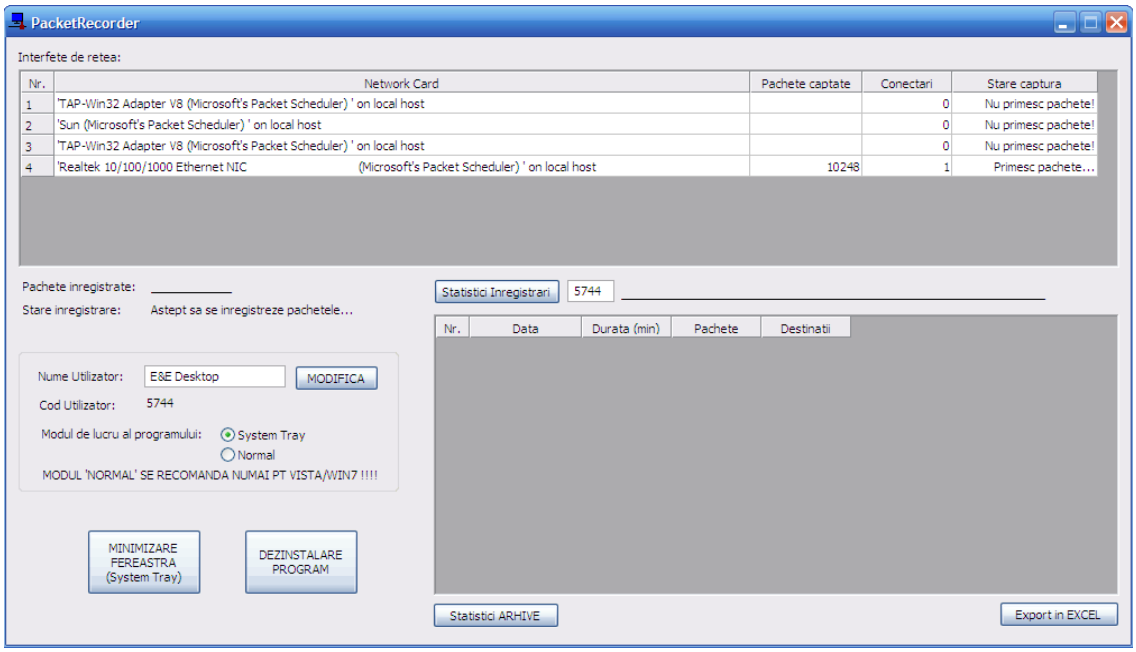


Fig. 4.1 Captarea pachetelor din reţea cu PacketRecorder

Sub tabelul cu interfeţele de reţea (în partea dreaptă), se află un alt tabel care oferă un centralizator zilnic aferent datelor înregistrate, precizându-se durata de funcţionare (în minute), numărul de pachete captate, numărul de destinaţii distincte identificate în trafic, pentru fiecare zi în parte.

Aplicaţia permite generarea unor statistici cu privire la datele înregistrate. Dacă există mai multe arhive cu date salvate în procesul de monitorizare, permite de asemenea o incursiune în arhivele respective, pentru a prezenta succint conţinutul acestora.

Ca setări, aplicaţia software poate funcţiona fie ca fereastră afişată pe ecran, fie ca aplicaţie în System Tray. A doua variantă permite o execuţie de lungă durată, fără a încurca activitatea utilizatorului.

4.2. Network Detector.

Este o aplicaţie software mai complexă, folosită în studiul metodei DTF. Are câteva module importante, care vor fi prezentate detaliat în ceea ce urmează.

4.2.1 Modulul „Network Setup”. Alcătuirea unei pseudo-reţele, formată din înregistrări individuale.

Modulul „Network Setup” este descris grafic de interfaţa prezentată în Fig. 4.2 şi se foloseşte pentru a crea contextul folosit la pasul de simulare. Datele înregistrate de PacketRecorder pe diverse calculatoare, pot fi adunate la un loc sub

forma unei pseudo-reţele. Fiecare arhivă este adăugată în sistem şi prelucrată, astfel încât să ofere condiţii mai bune pentru studiul metodei DTF.

În continuare sunte prezentate principalele funcţionalităţi a modulului software „Network Setup”.

În primul rând se extrag câteva elemente statistice din baza de date, necesare prelucrărilor ulterioare:

- perioada în care au fost înregistrate pachetele (data şi ora pentru primul şi ultimul pachet);
- numărul total de ore aferent perioadei înregistrate (nu ţine cont de pauze);
- numărul total de pachete înregistrate;
- numărul total de adrese IP identificate;
- numărul total de adrese MAC identificate.

O altă prelucrare importantă transformă înregistrările traficului, astfel încât să se obţină înregistrări orientate pe IP destinaţie şi minut. Astfel, înregistrările vor conţine perechi unice de forma:

- numărul minutului;
- IP destinaţie;
- adresa MAC.

Datele sunt astfel mult mai accesibile şi pot fi extrase în rapoarte, cu viteză semnificativ mai mare. În continuare, datele din pseudo-reţea pot fi folosite ca sursă de date pentru simulator.

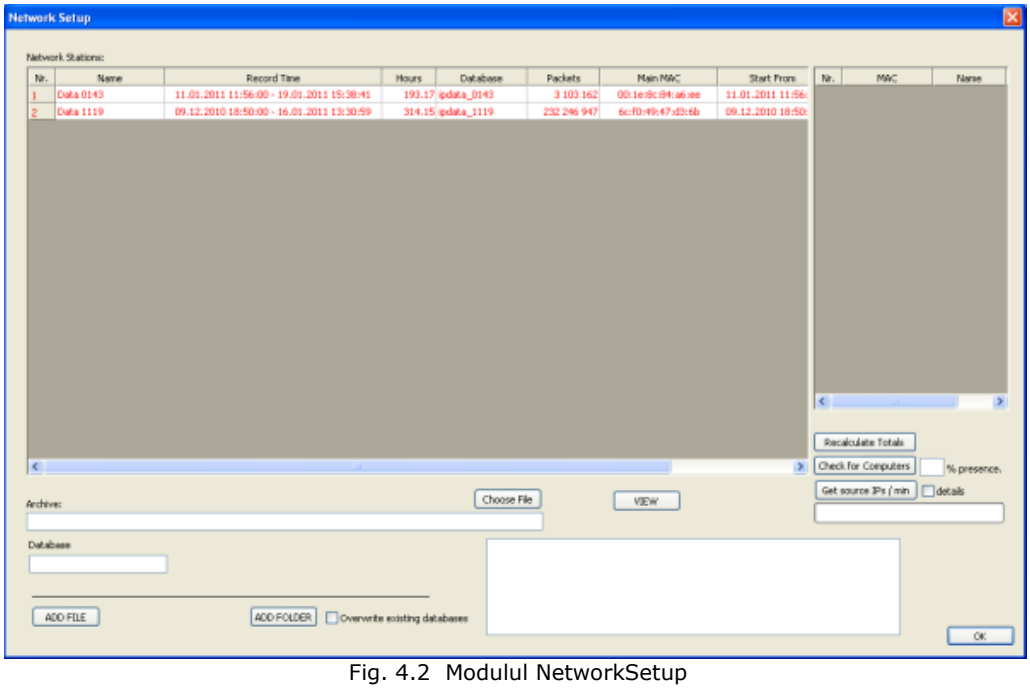


Fig. 4.2 Modulul NetworkSetup

După integrarea datelor în Modulul „Network Setup”, fiecare arhivă apare sub forma unei linii în tabelul din Fig. 4.2. Coloanele comunică informaţii importante despre date, după cum urmează:

- numele atribuit arhivei. Acest nume provine chiar din Paclet Recorder, dar poate fi modificat ulterior oricând se doreşte acest lucru;
- perioada monitorizată, identificată prin data şi ora primului pachet captat, respectiv data şi ora ultimului pachet captat;
- totalul de ore de funcţionare, calculat ca diferenţă dintre momentul final şi cel iniţial;
- numele bazei de date MySQL unde sunt stocate informaţiile;
- numărul total de pachete captate;
- adresa MAC considerată principală pentru staţia respectivă;
- numărul total de adrese IP identificate;
- numărul total de adrese MAC identificate.

Pentru a putea crea un mediu de simulare şi testare adecvat, modulul permite selectarea din tabel a uneia sau mai multor poziţii, care vor constitui surse de trafic în simulator.

Practic, prin selectarea surselor, se crează o reţea virtuală, ce poate fi configurată în mod avantajos, prin stabilirea momentului la care să intre „în acţiune” fiecare sursă.

Una din funcţiile modulului permite utilizatorilor să vizualizeze un caz de suprapunere în trafic a două sau mai multe adrese IP sursă, din aceeaşi clasă IP, care în acelaşi timp trimit date folosind o singură adresă MAC. Cazul acesta apare atunci când un atacator falsifică o adresă MAC şi intră in sistem în timp ce staţia autorizată este în funcţiune. Simulatorul semnalează cazul acesta, însă modulul Network Setup oferă o funcţie prin care să se detecteze suprapunerile.

Pentru a determina aceste situaţii, modulul va verifica fiecare minut din arhiva selectată şi va genera ca răspuns un tabel de tipul Tabelului 5.

Tabelul 5 – Vizualizarea suprapunerii sursei autorizate cu cele falsificate

Mnt	Adresa MAC şi destinaţiile IP aferente fiecărioa
14	**MAC: [00:11:6b:73:05:d8]**: 192.168.1.250 **MAC: [00:18:38:02:3b:01]**: 192.168.1.29 **MAC: [00:1c:25:36:24:af]**: 192.168.1.27 **MAC: [00:1d:6a:95:14:31]**: 192.168.1.2 **MAC: [00:1d:7d:0c:5e:2a]**: 0.0.0.0 **MAC: [00:1e:8c:84:a6:ee]**: 166.238.192.168 192.168.1.169 **MAC: [00:22:6b:43:0b:a3]**: 11.163.192.168 192.168.1.1 38.118.85.50 80.97.209.16 80.97.209.19 **=>** **SUPRAPUNERE pentru clasa [80.97.209#]!!!** 81.196.146.74 **MAC: [00:25:22:2e:4c:9e]**: 192.168.1.4 **MAC: [00:30:6e:d2:d6:a7]**: 192.168.1.100
15	**MAC: [00:11:6b:73:05:d8]**: 192.168.1.250 **MAC: [00:18:38:02:3b:01]**: 0.0.0.0 59.1.192.168 **MAC: [00:1c:25:36:24:af]**: 0.0.0.0 192.168.1.27 **MAC: [00:1d:7d:0c:5e:2a]**: 0.0.0.0 94.42.192.168 **MAC: [00:1e:8c:84:a6:ee]**: 0.0.0.0 166.238.192.168 192.168.1.169

4.2.2 Modulul „Fingerprint Generation”. Extragerea amprentelor de trafic.

Fig. 4.3 prezintă interfaţa grafică a modulului „Fingerprint Generation”, care extrage amprenta de referinţă pentru o adresă MAC.

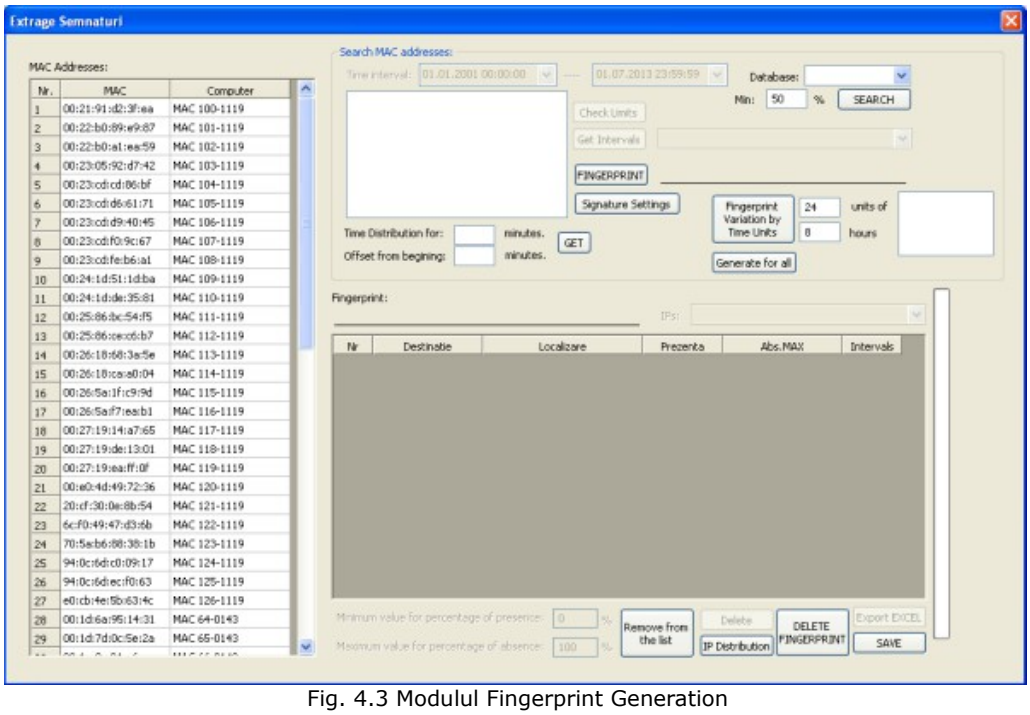


Fig. 4.3 Modulul Fingerprint Generation

Modulul are definite următoarele funcţii:

- definirea setărilor pentru extragerea amprentei de trafic;
- detectarea adreselor MAC dintr-o bază de date MySQL, care au trafic semnificativ;
- generarea amprentei de trafic pentru adresa MAC selectată, folosind setările definite;
- memorarea amprentei de trafic;
- ştergerea unei amprente de trafic;
- trasarea unui grafic pentru distribuţia traficului către o destinaţie IP selectată;
- afişarea listei tuturor adreselor MAC, şi a amprentelor de trafic aferente (dacă sunt înregistrate);
- generarea unui raport de urmărire a variaţiei amprentei de trafic, pe unităţi de timp;

Pentru a descrie modul de lucru al funcţiilor, se va parcurge un exemplu concret, pornind de la căutarea adreselor MAC semnificative, până la generarea amprentei de trafic şi folosirea ei în timp real. Se consideră o bază de date, care în

cele aproximativ 300 ore înregistrate, conţine 232 milioane pachete, repartizate pe 1779547 adrese IP şi 2264 adrese MAC.

Tabelul 6 – Adrese MAC semnificative, identificate în procesul de verificare automată a datelor înregistrate prin Packet Recorder

Adresă MAC	***Denumire automată***	***Adresă MAC***	***Denumire automată***
00:00:ca:f5:7a:92	MAC 135-1119	00:21:91:44:be:ef	MAC 164-1119
00:0a:e6:cb:4c:b0	MAC 136-1119	00:21:91:75:11:81	MAC 165-1119
00:0f:3d:31:ed:ea	MAC 137-1119	00:21:91:d2:3f:ea	MAC 166-1119
00:13:02:7d:67:0a	MAC 138-1119	00:22:b0:89:e9:87	MAC 167-1119
00:13:8f:f1:6d:be	MAC 139-1119	00:22:b0:a1:ea:59	MAC 168-1119
00:14:d1:5c:e8:2f	MAC 140-1119	00:23:05:92:d7:42	MAC 169-1119
00:15:e9:e0:be:a4	MAC 141-1119	00:23:cd:cd:86:bf	MAC 170-1119
00:15:e9:e3:5e:a4	MAC 142-1119	00:23:cd:d6:61:71	MAC 171-1119
00:19:5b:e1:3d:5f	MAC 143-1119	00:23:cd:d9:40:45	MAC 172-1119
00:19:e0:79:e6:3d	MAC 144-1119	00:23:cd:f0:9c:67	MAC 173-1119
00:1a:4d:22:a5:a8	MAC 145-1119	00:23:cd:fe:b6:a1	MAC 174-1119
00:1b:11:fb:c2:59	MAC 146-1119	00:24:1d:51:1d:ba	MAC 175-1119
00:1b:11:fe:e3:51	MAC 147-1119	00:24:1d:de:35:81	MAC 176-1119
00:1c:c0:60:d7:51	MAC 148-1119	00:25:86:bc:54:f5	MAC 177-1119
00:1c:f0:7d:2a:d1	MAC 149-1119	00:25:86:ce:c6:b7	MAC 178-1119
00:1c:f0:7e:a9:e1	MAC 150-1119	00:26:18:68:3a:5e	MAC 179-1119
00:1c:f0:86:d8:df	MAC 151-1119	00:26:18:ca:a0:04	MAC 180-1119
00:1d:72:05:6b:ee	MAC 152-1119	00:26:5a:1f:c9:9d	MAC 181-1119
00:1d:7d:9a:24:f9	MAC 153-1119	00:26:5a:f7:ea:b1	MAC 182-1119
00:1e:33:77:3d:d4	MAC 154-1119	00:27:19:14:a7:65	MAC 183-1119
00:1e:58:0b:8b:9b	MAC 155-1119	00:27:19:de:13:01	MAC 184-1119
00:1e:58:11:d2:0d	MAC 156-1119	00:27:19:ea:ff:0f	MAC 185-1119
00:1e:58:14:45:63	MAC 157-1119	00:e0:4d:49:72:36	MAC 186-1119
00:1e:58:14:46:fb	MAC 158-1119	20:cf:30:0e:8b:54	MAC 187-1119
00:1e:58:18:22:3b	MAC 159-1119	6c:f0:49:47:d3:6b	MAC 188-1119
00:1e:8c:87:1d:fa	MAC 160-1119	70:5a:b6:88:38:1b	MAC 189-1119
00:21:04:1a:4c:f9	MAC 161-1119	94:0c:6d:c0:09:17	MAC 190-1119
00:21:04:1c:b4:8a	MAC 162-1119	94:0c:6d:ec:f0:63	MAC 191-1119
00:21:04:1c:cc:e4	MAC 163-1119	e0:cb:4e:5b:63:4c	MAC 192-1119

În primul rând, trebuie verificat traficul aferent adreselor MAC, pentru a elimina adresele MAC nesemnificative. Astfel, din totalul de 2264 de adrese MAC, se extrag doar cele care au trafic peste o limită stabilită aproximativ la jumătate din numărul total de minute înregistrate. Rămân astfel doar 58 adrese MAC, denumite automat de către sistem, şi menţionate în Tabelul 6.

Denumirea este utilă pentru a identifica mai uşor o adresă. Modul de codificare automată nu prezintă importanţă în studiul de faţă, în principiu, se poate folosi orice metodă de numerotare unică.

În continuare se va discuta generarea amprentei de trafic pentru adresa MAC "00:00:ca:f5:7a:92", codificată automat de sistem prin notaţia „[MAC 135-1119]".

Pasul 1: - *Stabilirea setărilor*

Pentru studiul de faţă s-a folosit următoarele setări:

- amprenta se extrage din 24 ore de trafic efectiv;
- cel puţin 1% procent de prezenţă;
- intervalul total se împarte în 4 subintervale, din care sunt obligatorii minim două;
- după 10 minute de inactivitate, staţia este declarată ca „oprită".

Pasul 2: - *Se caută perioada aferentă celor 24 ore de funcţionare efectivă şi se extrag înregistrările aferente perioadei*

- ➔ primul minut evaluat: 1
- ➔ ultimul minut evaluat: 3959
- ➔ număr total de minute considerate pentru evaluarea curentă: 3959

Pasul 3: - *Se identifică eventualele opriri şi se reface traficul, astfel încât să se elimine intervalele respective*

S-au identificat 3 opriri, aşa după cum reiese din Tabelul 7:

Tabelul 7 – Opriri identificate în perioada evaluată

Număr oprire	***De la minutul***	***Până la minutul***	***Minute de oprire***
Oprire 0	215	778	564
Oprire 1	1533	2278	746
Oprire 2	3163	3679	517
		TOTAL:	**1827**

Se recalculează numărul total de minute considerate pentru evaluarea curentă: 2132.

Pasul 4: - *Se extrag IP-urile destinaţiilor, care au prezenţă cel puţin cât s-a stabilit în setările iniţiale*

Rezultă doar 5 adrese IP. Ele sunt introduse într-o listă, calculându-se inclusiv procentul de prezenţă.

Pasul 5: - *Pentru fiecare IP, se calculează încadrarea în subintervale*

Toate cele 5 IP-uri rămân mai departe şi rezultă amprenta de trafic din Tabelul 8

Tabelul 8 – Amprentă de trafic formată din cinci destinaţii IP cu trafic constant

Nr. crt	***Adresă IP***	***Procent Prezenţă***
1	224.0.0.1	47,05%
2	40.1.0.0	28,71%
3	239.255.255.250	13,60%
4	255.255.255.255	4,55%
5	239.192.152.143	2,39%

Pentru a dovedi că aceste adrese au trafic constant, se prezintă graficele de distribuţie în timp, pentru fiecare în parte, axa absciselor fiind etalonată în minute iar pe ordonată se va marca valoarea „0” dacă în minutul *M* nu există trafic către destinaţia urmărită, sau „1” în caz contrar. Zonele haşurate în culoare albastru deschis reprezintă momentele în care staţia a fost oprită.

Fig. 4.4 prezintă distriubţia în timp a traficului către destinaţia 224.0.0.1. Se observă prezenţa traficului într-o manieră constantă, pe toată durata de funcţionare a staţiei.

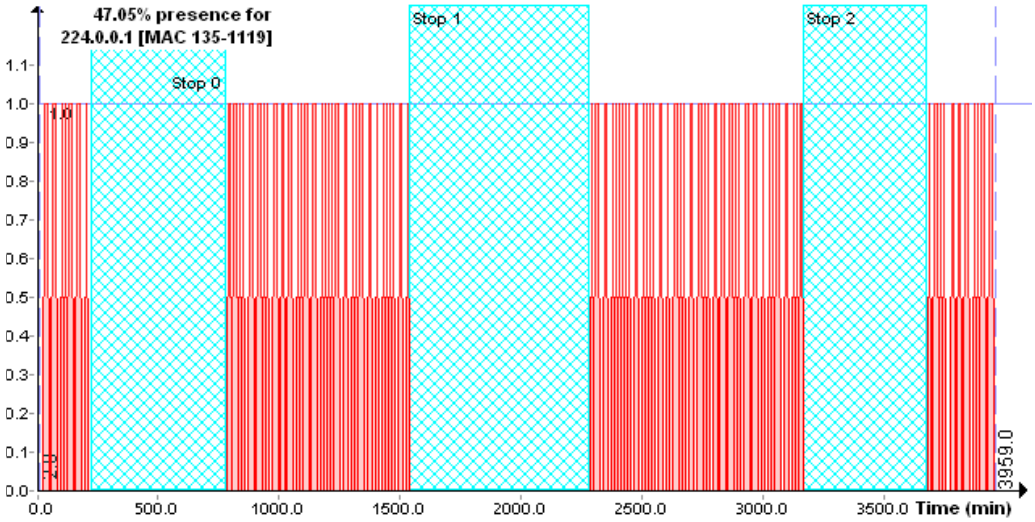


Fig. 4.4 Distribuţia în timp, a traficului către destinaţia 224.0.0.1

În Fig. 4.5, se reprezintă o altă destinaţie IP. Deşi valoarea procentului de prezenţă este mai mică, totuşi se observă şi în cazul acesta o repartizare constantă pe toată durata de funcţionare.

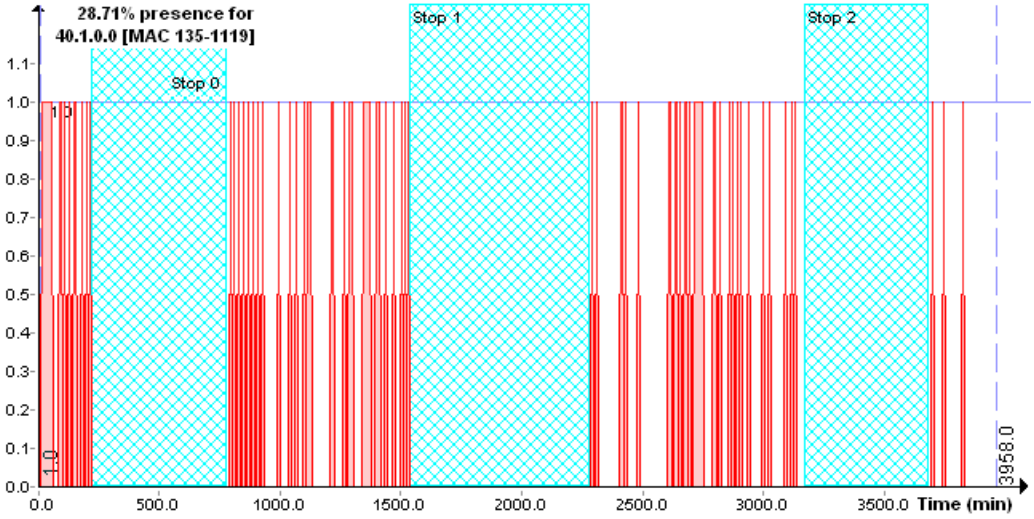


Fig. 4.5 Distribuţia în timp, a traficului către destinaţia 40.1.0.0

Fig. 4.6 prezintă o situaţie similară, dar procentul de prezenţă este şi mai mic.

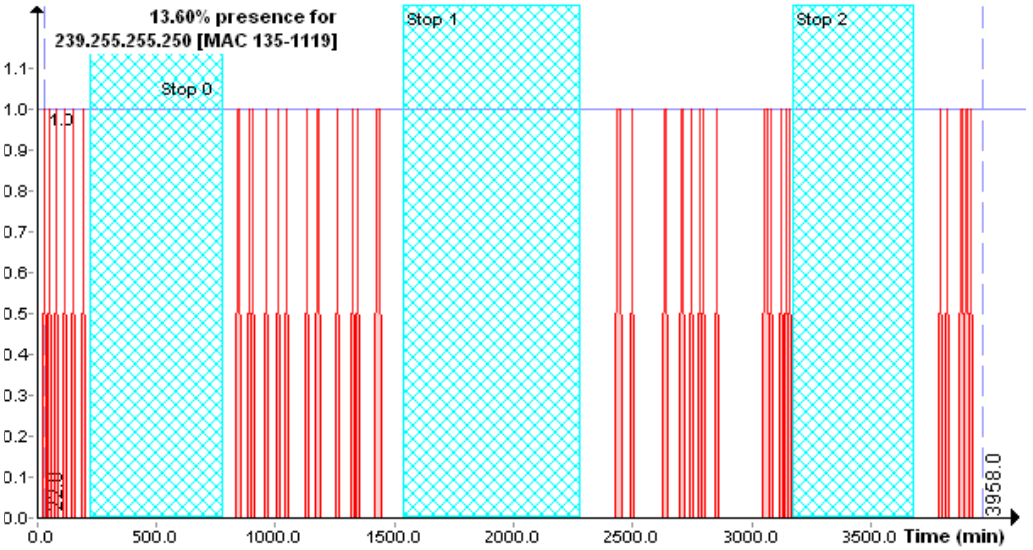


Fig. 4.6 Distribuţia în timp, a traficului către destinaţia 239.255.255.250

În Fig. 4.7, minutele cu trafic către adresa IP urmărită sunt mult mai puţine. Totuşi, distribuţia lor permite să fie încadrată în categoria celor constante.

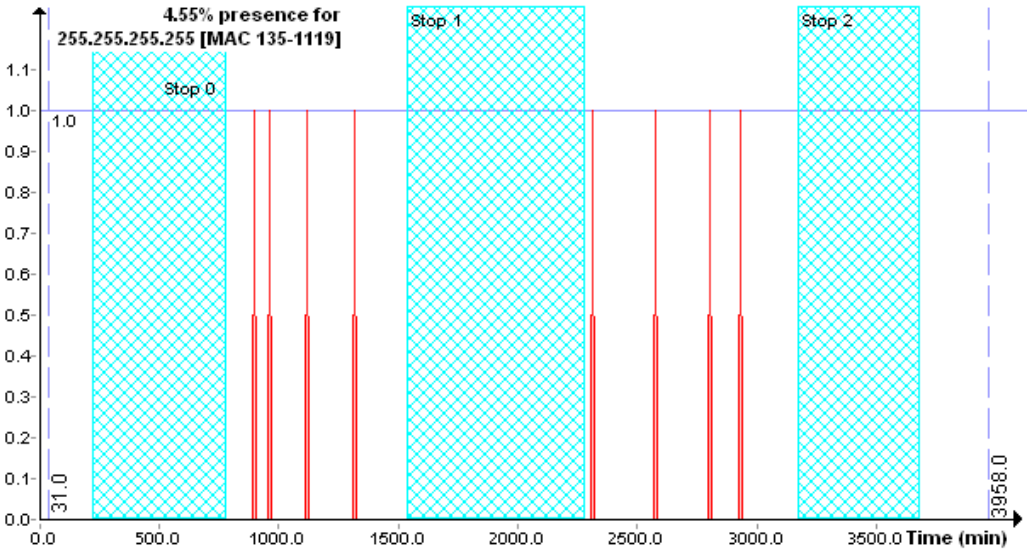


Fig. 4.7 Distribuţia în timp, a traficului către destinaţia 255.255.255.255

Un caz mai aparte este cel reprezentat Fig. 4.8, unde aparent s-ar putea considera un caz de prezenţă punctuală. Totuşi, după cum se va observa în cele ce urmează, adresa IP a destinaţiei urmărite are trafic constant, chiar dacă procentul de prezenţă este foarte mic.

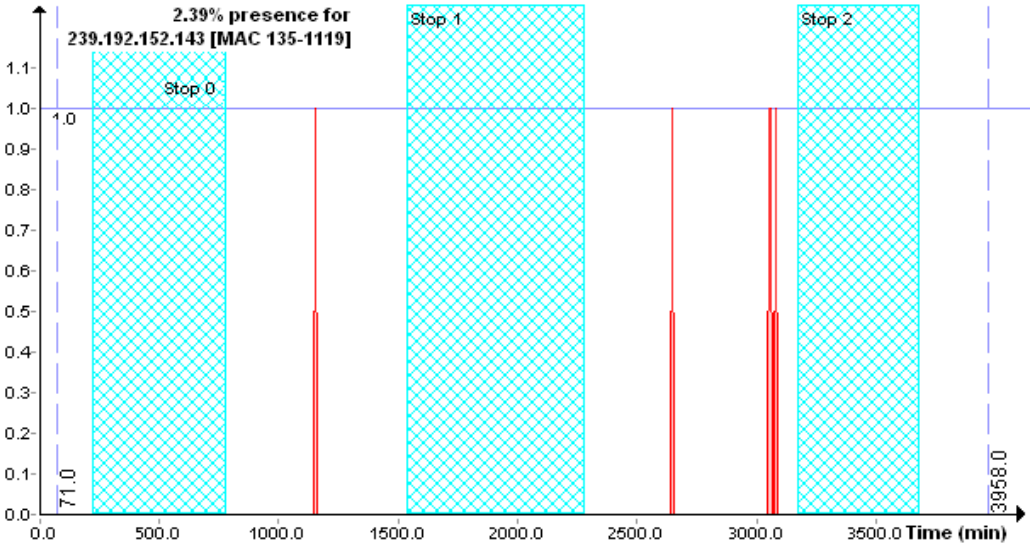


Fig. 4.8 Distribuţia în timp, a traficului către destinaţia 239.192.152.143

În calculul amprentei de trafic s-au folosit datele extrase din trafic pentru o perioadă limitată în timp.

Se doreşte să se verifice în ce măsură aceste destinaţii considerate cu trafic constant, se regăsesc sau nu se regăsesc ulterior în trafic, sau mai mult, dacă ele îşi păstrează în timp valorile pentru procentele de prezenţă.

În acest sens se va studia modul în care variază amprenta de trafic în timp, prin apelarea unei funcţii de generare a variaţiei semnăturii, prezente în modulul „Fingerprint Generation".

Această variaţie se determină în felul următor: se notează cu *TU* intervalul de timp folosit pentru extragerea amprentei de referinţă, şi se va căuta să se repete determinarea amprentei pentru următoarele intervale *TU*. Fig. 4.9 prezintă un caz în care s-au verificat amprentele pentru 14 unităţi de timp consecutive.

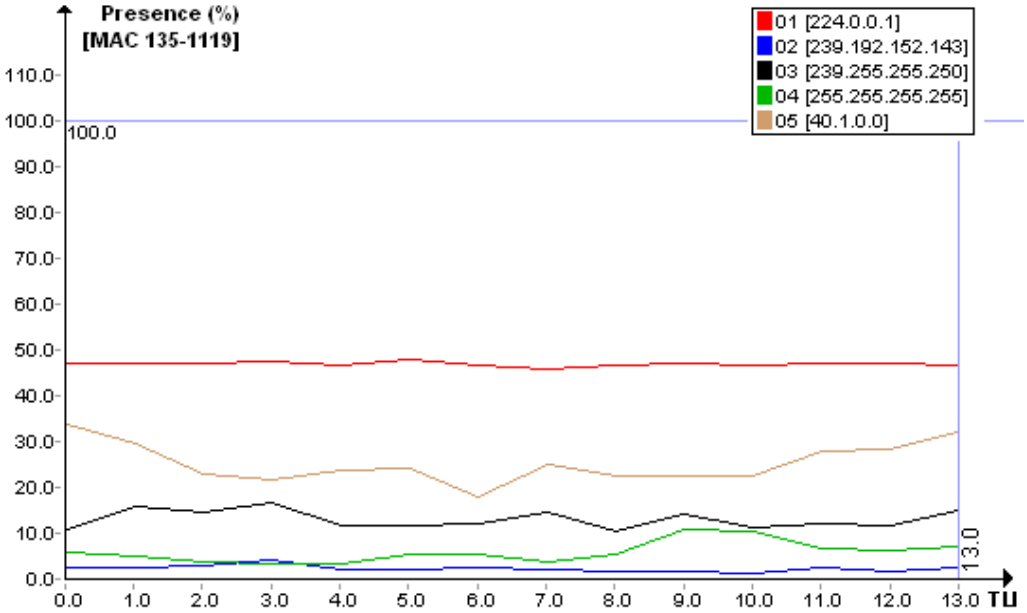


Fig. 4.9 Variaţia amprentei de trafic pentru 14 unităţi de timp

După determinarea amprentelor de trafic pentru fiecare unitate de timp, pe grafic se notează valorile procentelor de prezenţă ale destinaţiilor din amprenta de referinţă, pentru fiecare unitate de timp evaluată. Dacă destinaţiile IP au într-adevăr trafic constant, atunci acest lucru trebuie să fie evidenţiat pe grafic prin două aspecte:

- adresa IP trebuie să se regăsească în amprentele de trafic ale unităţilor de timp succesive;
- valoarea procentului de prezenţă nu trebuie să varieze foarte mult de la o unitate de timp la alta.

Din Fig. 4.9, se observă că cele cinci destinaţii IP se regăsesc în toate cele 14 unităţi de timp, adică se întind pe o durată de 112 ore de funcţionare efectivă, variaţiile fiind mici de la o unitate de timp la alta.

Folosind aceste destinaţii ca amprentă de trafic, se obţine o bună recunoaştere a adresei MAC, după cum se poate observa în Fig. 4.10.

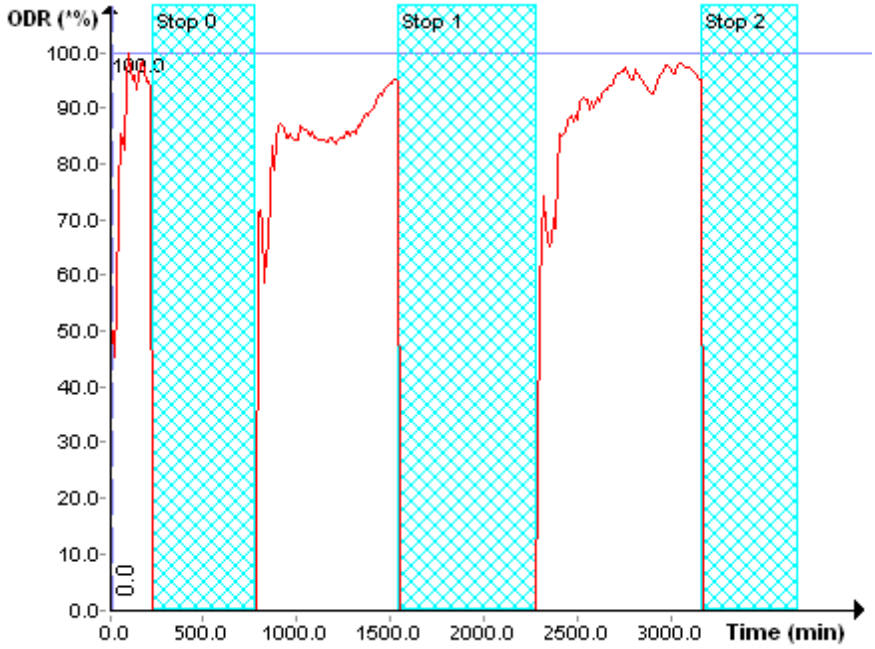


Fig. 4.10 Variaţia Gradului Global de Recunoaştere pentru aproximativ 3500 minute

Fig. 4.10 prezintă variaţia Gradului Global de Recunoaştere al staţiei de-a lungul unei perioade evaluate de aproximativ 3500 minute, folosind amprenta de referinţă stabilită iniţial. Graficul marchează cele trei perioade de oprire ale staţiei, prin zone haşurate în culoarea albastru.

După fiecare oprire a staţiei, în momentul în care staţia începe din nou să transmită pachete în reţea, modulul de monitorizare începe să calculeze la nivel de minut Gradul Global de Recunoaştere. Acest parametru este reprezentat pe grafic şi se observă că ajunge rapid la valori mari, marcând astfel faptul că traficul asociat cu adresa MAC evaluată, provine într-adevăr de la sursa autentică.

4.2.3 Modulul „Fingerprint Variation Report”.

Acest modulul permite observarea variaţiei în timp a amprentei de trafic. Se alege de către utilizator o unitate de timp, notată TU („Time Unit”), şi se stabileşte câte unităţi de timp consecutive vor fi evaluate. Pentru fiecare unitate de timp, se va calcula amprenta de trafic aferentă.

Scopul este acela de a observa dacă există destinaţii IP care apar în amprenta de trafic pe parcursul mai multor unităţi de timp şi care astfel dovedesc existenţa unui trafic constant către aceste destinaţii. Pe baza calculelor descrise anterior, se pot înregistra variaţiile semnăturilor, pentru prelucrări ulterioare, cum ar fi:

- extragerea unor date statistice referitoare la destinaţiile IP;
- desenarea graficelor de prezenţă ale destinaţiilor IP;

- generarea unui raport care prezintă succint rezultatele evaluării destinaţiilor IP;
- generarea unui raport cu privire la adresele MAC întâlnite.

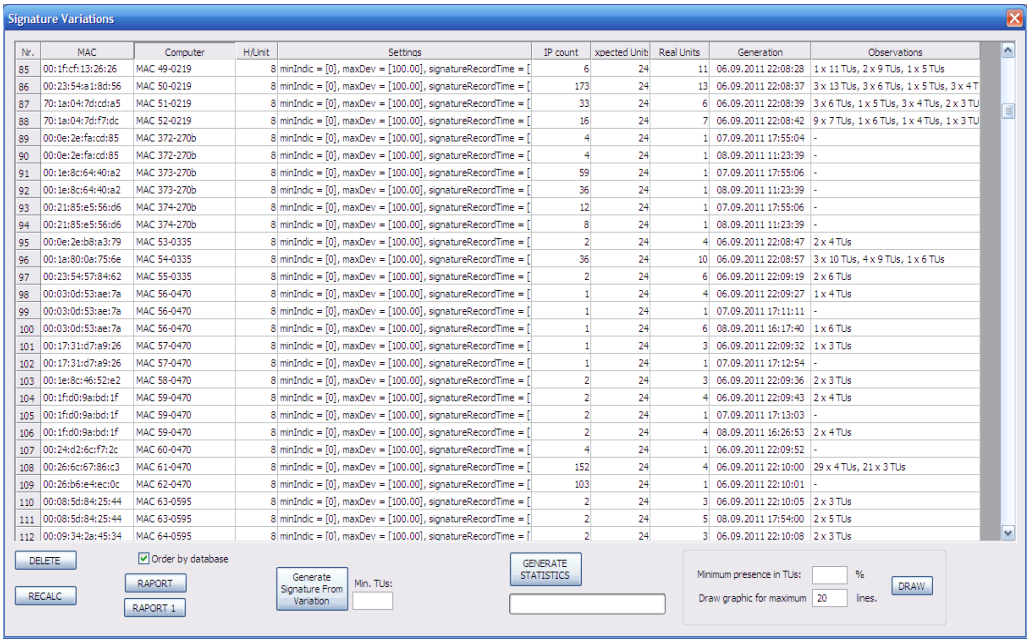


Fig. 4.11 Modulul Fingerprint Variation Report

Fig. 4.11 prezintă interfaţa grafică a modulului „Fingerprint Variation Report”. Interfaţa conţine un tabel cu adresele MAC. Datele statistice referitoare la destinaţiile IP calculează fiecare MAC evaluat, precum şi numărul de destinaţii care se repetă pentru un anumit număr de unităţi de timp. De exemplu, pentru adresa MAC „00:21:97:1a:7b:fe”, cunoscută în sistem sub denumirea „MAC 425-744a”, evaluată pentru 24 de unităţi de timp consecutive, a câte 8 ore fiecare unitate de timp, se pot extrage următoarele date statistice:

4 x 24 TUs, 1 x 15 TUs, 1 x 13 TUs, 1 x 10 TUs, 1 x 9 TUs, 1 x 8 TUs, 3 x 7 TUs, 2 x 6 TUs, 2 x 5 TUs, 6 x 4 TUs, 8 x 3 TUs

ceea ce înseamnă:

- 4 IP-uri găsite în 24 unităţi de timp (4 x 24 TUs)
- 1 IP găsit în 15 unităţi de timp (1 x 15 TUs)
- 1 IP găsit în 13 unităţi de timp (1 x 13 TUs)
- ş.a.m.d.

Pentru exemplul anterior sunt 4 IP-uri care sunt prezente în toate unităţile de timp evaluate. Folosirea lor ca amprentă de trafic va genera o bună recunoaştere a adresei MAC.

4.2.4 Modulul „Network Simulator”.

Permite simularea funcţionării unei reţele de calculatoare, în care se aplică sistemul de monitorizare şi identificare în timp real, pe baza metodei DTF, în vederea validării adreselor MAC întâlnite în trafic. Fig. 4.12 prezintă interfaţa programului pentru simulator.

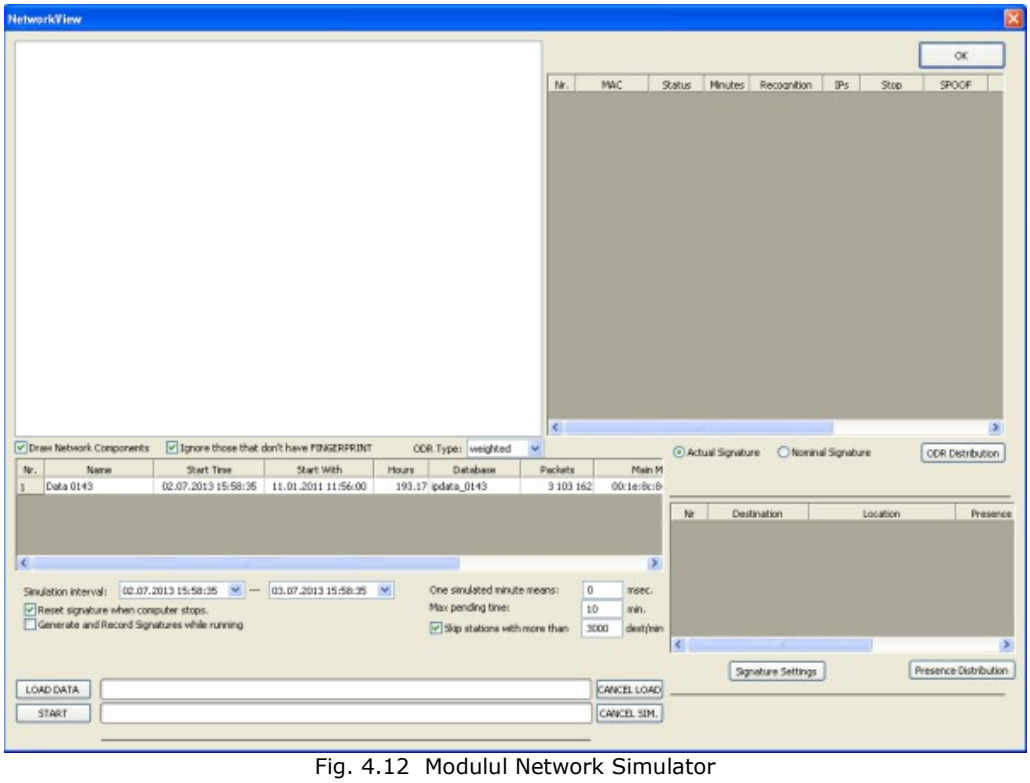


Fig. 4.12 Modulul Network Simulator

Simulatorul are incorporate următoarele funcţii:

- folosirea datelor din arhivele alese în Modulul Network Setup cu modificarea momentelor de timp în care aceste arhive încep injectarea datelor în reţea;
- preîncărcarea datelor pentru mărirea vitezei de simulare;
- reprezentarea grafică a reţelei sub forma unor pictograme care sugerează gradul de recunoaştere al staţiei;
- date actualizate permanent, cu privire la amprentele de trafic şi Gradul Global de Recunoaştere;
- particularizarea simulării prin câteva setări, cum ar fi:
  - activarea / dezactivarea reprezentării grafice;
  - ignorarea staţiilor care nu au memorată o amprentă de trafic;
  - stabilirea duratei în timp a unui minut simulat (în milisecunde);
  - setarea numărului maxim de minute de inactivitate, care determină declararea staţiei ca „oprită”;

  - ignorarea sau acceptarea staţiilor care comunică pe minut cu un număr foarte mare de destinaţii IP (de ordinul miilor);
  - generarea amprentelor de trafic în timpul derulării simulării;
  - resetarea sau păstrarea amprentelor de trafic în urma opririi staţiilor;
  - calcularea Gradului Global de Recunoaştere în manieră standard sau ponderat;
- oprirea forţată a unei staţii prezente în reţea (simulatorul va ignora datele provenite de la staţia respectivă);
- modificarea adresei MAC a unei staţii din reţea, pentru observarea efectelor produse în procesul de identificare în timp real.

Înainte de a porni efectiv simularea, modulul are o faciltate de preîncărcare a traficului, din bazele de date MySQL, în memoria simulatorului. Această preîncărcare facilitează programului o mărire a vitezei de lucru şi evită anumite întârzieri care ar putea fi generate de selecţii multiple din baze de date MySQL de dimensiuni foarte mari.

Trebuie să se menţioneze aici aptul că preîncărcarea este necesară datorită faptului că simulatorul procesează datele mult mai rapid decât s-ar procesa în timp real, pentru a permite utilizatorului să observe fenomenele pe durate mari de timp.

Pentru simulare, fereastra programului pune la dispoziţia utilizatorului trei zone distincte.

Prima zonă este aferentă reprezentării grafice a reţelei, în cadrul ei, adresele MAC întâlnite în trafic sunt desenate sub formă de pictograme, a căror formă şi culoare sugerează gradul de recunoaştere. Se folosesc pictograme diferite pentru staţiile considerate cunoscute faţă de cele necunoscute, pentru staţiile cunoscute, pictogramele fiind colorate diferit, funcţie de gradul de recunoaştere.

A doua zonă este destinată unui centralizator pentru adresele MAC întâlnite în trafic. Pentru fiecare adresă MAC, centralizatorul afişează câteva informaţii statistice. În primul rând se precizează starea curentă, care poate lua una din valorile: „active”, „stoped” şi respeciv „pending”. Starea „pending” se referă la perioada de timp în care staţia nu emite trafic, dar încă nu a expirat durata maximă admisă. După expirarea duratei maxime admise pentru „pending”, starea staţiei devine automat „stopped”.

Tot în centralizator se mai afişează numărul total de minute care au trecut din momentul în care staţia a pornit şi până în momentul prezent al simulării şi numărul total de adrese IP destinaţie identificate, care oferă o imagine despre traficul generat din cadrul unei staţii. Ca obsevaţie, se poate menţiona în contextul acesta faptul că aplicaţiile software de tip „torent”, instalate pe staţiile client, generează un număr foarte mare de destinaţii cu care comunică într-un interval foarte scurt.

Cea mai importantă informaţie afişată în centralizator este Gradul Global de Recunoaştere. O dată cu trecerea fiecărui minut simulat, valoarea este actualizată. Pentru a vizualiza variaţia Gradului Global de Recunoaştere, programul conţine o facilitate de desenare a unui grafic cu toate valorile, din momentul începerii simulării şi până în prezent.

A treia zonă din simulator permite vizualizarea amprentelor de trafic asociate cu adresa MAC selectată în centralizator. Se permite afişarea amprentei de referinţă, sau a amprentei actuale, funcţie de dorinţa utilizatorului, exprimată prin selecţia corespunzătoare a unor butoane radio. În felul acesta, la orice moment al simulării, utilizatorul poate studia transformările care apar în amprentele actuale şi efectul lor asupra Gradului Global de Recunoaştere.

4.3. Concluzii.

Conceperea unui toolbox software pentru studiul metodei DTF reprezintă un suport util întrucât permite pe de-o parte colectarea traficului capturat în reţelele de calculatoare, şi mai apoi folosirea acestora pentru extragerea amprentelor de trafic.

Simulatorul oferă un cadru de testare adecvat, folosind ca surse date reale, care sunt prelucrate în aceeaşi manieră cu sistemul real de monitorizare şi control. Fiecare aspect poate fi corect observat şi interpretat. Ţinând cont de faptul că mediul de testare este unul simulat, se pot aplica metode de testare care în mod normal nu se puteau aplica pe un sistem real, datorită posibilităţii de apariţie a unor efecte nedorite cu consecinţe dăunătoare.

Rezultatele obţinute cu ajutorul simulatorului atestă avantajele pe care metoda DTF le oferă în domeniul detecţiei adreselor MAC falsificate, pătrunderile neautorizate fiind detectate şi semnalate conform amprentelor de referinţă generate.

5. REZULTATE EXPERIMENTALE

Capitolul urmăreşte să demonstreze utilitatea şi validitatea metodei DTF, prin intermediul unui experiment, care a dus la obţinerea unor rezultate relevante. Experimentul s-a realizat prin intermediul unor teste, care au avut ca scop demonstrarea faptului că se poate stabili o amprentă a staţiei pe baza traficului de reţea generat de către aceasta [Sas-12a].

5.1. Descrierea testelor.

Pentru a putea verifica validitatea şi utilitatea metodei DTF, sunt necesare teste aplicate unui trafic real. Primele teste au avut în vedere un număr mic de calculatoare (2-3). Pentru fiecare IP s-a calculat procentul de prezenţă şi procentul de absenţă maximă. IP-urile intră în componenţa amprentei dacă cei doi parametrii (PP şi PA) se încadrează în limitele admise, limite stabilite manual (LP = 4%, LA = 2%, iar condiţiile sunt PP >= LP iar PA <= LA). Folosirea combinaţiei dintre procentul de prezenţă şi cel de absenţă maximă, îşi are limitările ei, după cum s-a discutat în capitolul 3.

Al doilea test s-a realizat pe un număr mult mai mare de calculatoare, folosind o aplicaţie software care înregistra traficul la nivelul plăcilor de reţea. Durata testului s-a întins pe câteva luni, adunând în final 99 baze de date MySQL, cu înregistrări încadrate conform Tabelului 9:

Tabelul 9 – Înregistrări ale programului Packet Recorder pe aproximativ 100 calculatoare

Numărul orelor de funcţionare	***Total baze de date MySQL***
între 0 – 50 ore	20
între 50 – 100 ore	30
între 100 – 150 ore	12
între 150 – 200 ore	19
între 200 – 250 ore	7
între 300 – 400 ore	7
între 400 – 1000 ore	3
peste 2000 ore	1

Testul a căutat în principal să verifice dacă se pot identifica în timp destinaţii IP care apar în amprentele de trafic în mod constant. Pentru a lămuri acest aspect, testul a urmărit extragerea de amprente la intervale succesive de timp, şi a verificat dacă există destinaţii IP care să apară în mai multe, sau chiar în toate intervalele. Concret, s-au aplicat următoarele reguli:

- amprenta de trafic se extrage pentru o unitate de timp (TU = Time Unit), stabilită la 8 ore de funcţionare efectivă (1 TU = 8 ore);
- testul este repetat pentru 24 TU, succesive;
- restricţii pentru IP-urile care intră în amprentă:

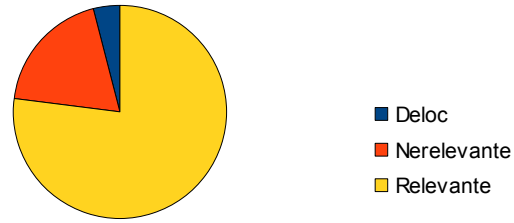
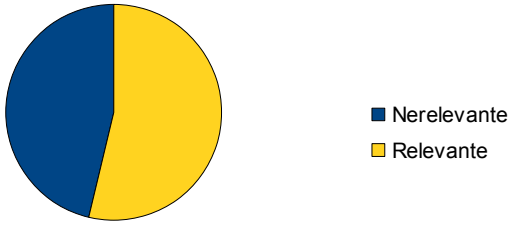
- minim 1% prezenţă per TU;
- dacă se împarte intervalul minutelor unui TU în 4 intervale egale, IP-ul trebuie să se regăsească în cel puţin 2 intervale;
- staţia este declarată ca oprită după 10 minute de inactivitate;

- testul s-a efectuat pentru toate adresele MAC semnificative, descoperite în bazele de date. Pentru ca o adresă MAC să fie considerată semnificativă, trebuie să aibă un număr de pachete mai mare decât cel puţin jumătate din interval.

În concluzie, testul nu şi-a propus să calculeze Gradul Global de Recunoaştere şi nici variaţia lui în timp. S-a urmărit doar generarea amprentelor pentru unităţi de timp succesive, ca ulterior să verifice dacă se observă sau nu adrese IP care se repetă pe mai multe sau chiar pe toate intervalele.

Rezultatele testului pot fi sumarizate ca în Tabelul 10

Tabelul 10 – Observaţii generale asupra rezultatelor experimentale

Din punct de vedere al calculatoarelor pentru care s-au colectat informaţii, s-au identificat: • 4 fără informaţii • 19 cu informaţii nerelevante • 77 cu informaţii relevante	 Relevanţa datelor colectate per calculatoare evaluate
Din punct de vedere al adreselor MAC, s-au identificat 845 cu trafic suficient de mare pentru a merita să fie studiate. Din acestea, 454 conţin informaţii relevante.	 Relevanţa datelor colectate per adrese MAC găsite

Pentru cele 454 adrese MAC, s-au extras o serie de statistici şi grafice, pentru a evidenţia că într-adevăr, se pot identifica adrese IP cu care calculatoarele comunică în mod constant, pe durate mari de timp. Chiar şi în prezenţa unor aplicaţii software care generează trafic cu zeci de mii de adrese IP în intervale scurte (de ordinul orelor), metoda DTF dă rezultate foarte bune.

5.2. Rezultate privind variaţia amprentelor de trafic.

În continuare, se prezentă în detaliu câteva cazuri de studiu cu graficele aferente obţinute. Pe grafice, axa OX reprezintă numărul unităţii de timp, cu convenţia că prima unitate de timp se numerotează cu 0. Aşa cum s-a afirmat deja, 1 TU = 8 ore. Cele 8 ore sunt efectiv ore de funcţionare, pauzele fiind eliminate. Axa OY reprezintă procentul de prezenţă, aferent fiecărei unităţi de timp. Liniile graficului reprezintă diverse destinaţii IP, care intră în componenţa amprentelor de trafic. Pentru a nu se încărca graficul cu foarte multe linii, au fost păstrate maxim 20 de IP-uri, cele care au trafic întins pe cea mai mare durată de timp.

Graficele pot fi clasificate astfel:

- destinaţii IP prezente pe durate mari de timp, cu variaţii reduse ale procentelor de prezenţă;
- destinaţii IP prezente pe durate mari de timp, dar cu variaţii mari ale procentelor de prezenţă.

Fig. 5.1 prezintă graficele este aferente unei adrese MAC întâlnite în trafic pe 18 TU, adică timp de 144 ore de funcţionare. Amprentele de trafic extrase pentru fiecare din cele 18 TU, conţin aceleaşi 5 adrese IP: 105.1.0.0, 192.168.105.127, 255.255.255.255, 94.178.105.127, 94.178.105.3, identificate în fiecare unitate de timp şi cu procente de prezenţă ce variază relativ puţin. Dacă stabilim amprenta de trafic astfel încât să conţină cele 5 destinaţii, recunoaşterea adresei MAC va fi foarte bună.

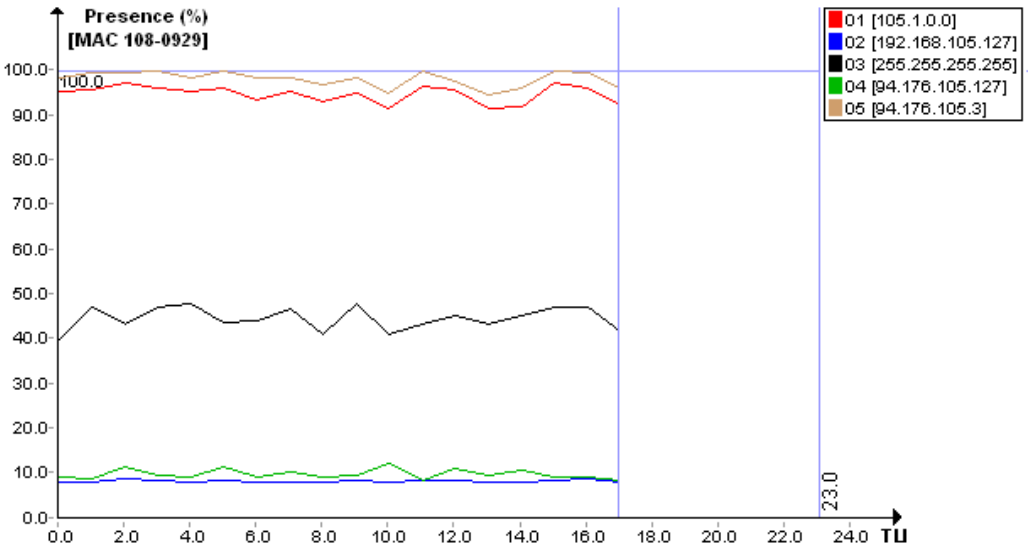


Fig. 5.1 Exemplu de variaţie a amprentei de trafic

Fig. 5.2 prezintă variaţia semnăturii pentru o altă staţie. Adresa IP 13.35.219.16 este prezentă permanent în trafic, pe durata de 13TU, adică 104 ore, cu un procent de prezenţă care variaza de la un TU la altul, între 91% - 98%. Alte câteva destinaţii au prezenţe între 20% - 30%.

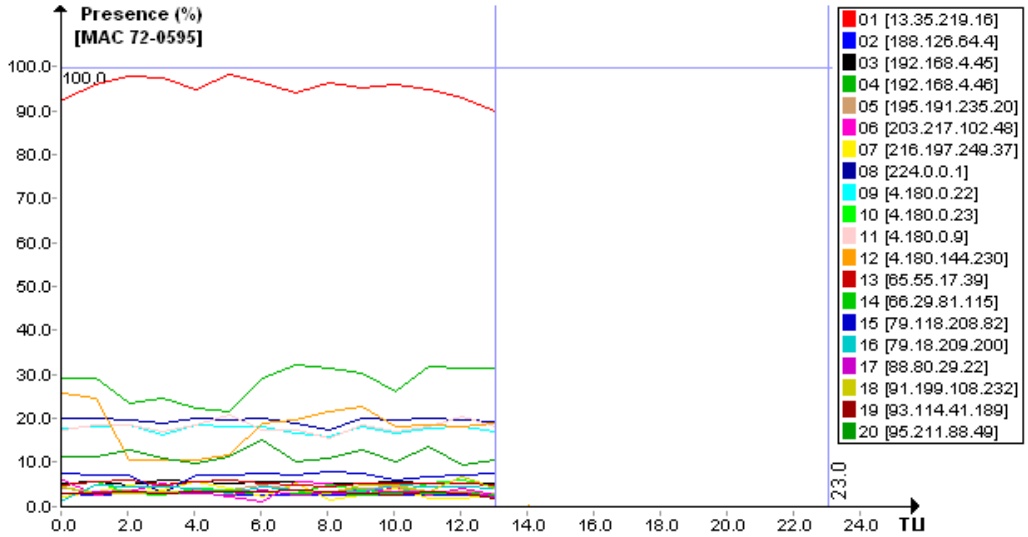


Fig. 5.2 Exemplu de variaţie a amprentei de trafic pentru 13TU

Fig. 5.3 conţine cinci destinaţii, care toate se regăsesc pe un interval de 13TU, adică 104 ore. Unele dintre ele variază foarte puţin, iar altele mai mult, dar toate se încadrează în categoria destinaţiilor cu trafic considerat constant.

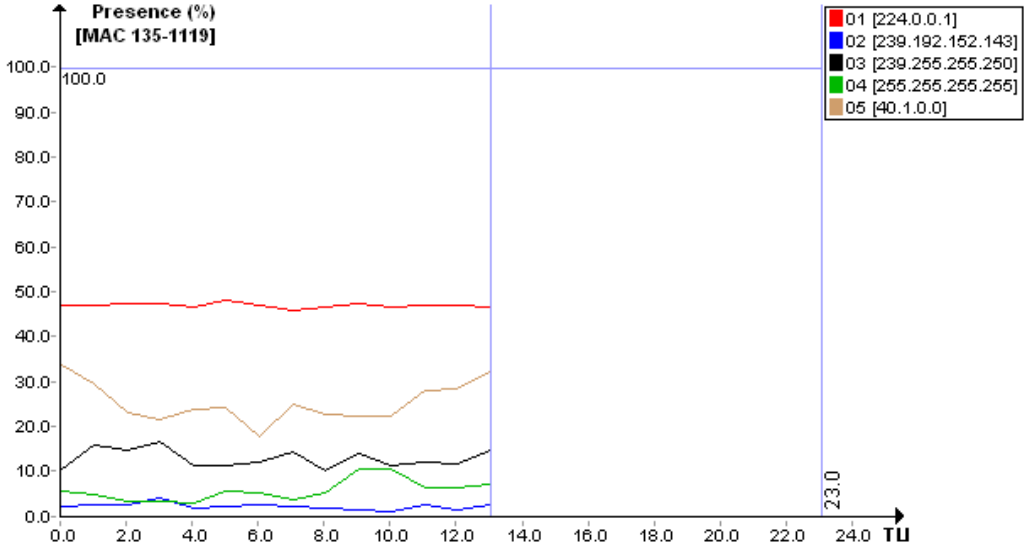


Fig. 5.3 Exemplu de variaţie a amprentei de trafic pentru 13TU, cu 5 destinaţii IP

Fig. 5.4 descrie un caz cu şase destinaţii IP. Cinci dintre acestea sunt cu trafic constant, cuprins între 5% - 20%. O destinaţie are o variaţie foarte puternică, începând cu valori de prezenţă situate aproximativ la 48% şi coborând abrupt câteva ore la valori cuprinse între 5% - 20%.

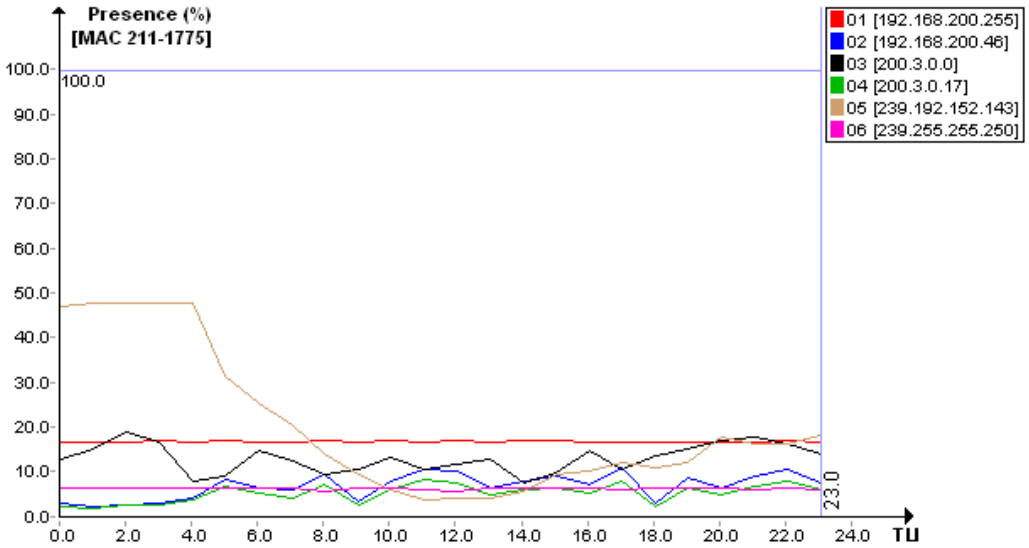


Fig. 5.4 Exemplu de variaţie a amprentei de trafic pentru 23TU, cu valori mici ale procentelor de prezenţă

În Fig. 5.5 se observă patru destinaţii IP cu valori relativ mici de prezenţă, dar totuşi cu un grad ridicat de constanţă. O altă adresă are o prezenţă în jurul valorii de 30%, iar alta variază foarte puternic, între 20% - 83%.

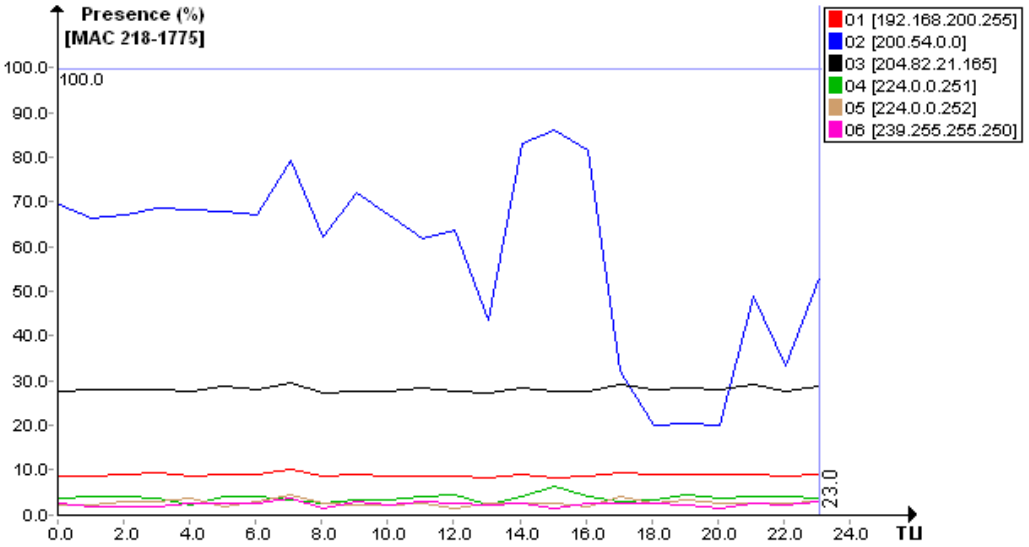


Fig. 5.5 Exemplu de variaţie a amprentei de trafic pentru 23TU, cu valori mici şi medii ale procentelor de prezenţă

Fig. 5.6 prezintă un caz cu număr mare de destinaţii cu trafic constant la valori ridicate ale procentului de prezenţă. Din cele opt adrese IP, şapte au procente de prezenţă cuprinse în intervalul 85% - 100%. Numai una singura are prezenţă în

jurul valorii de 10%, dar se poate observa că este constantă, ceea ce reprezintă un mare avantaj.

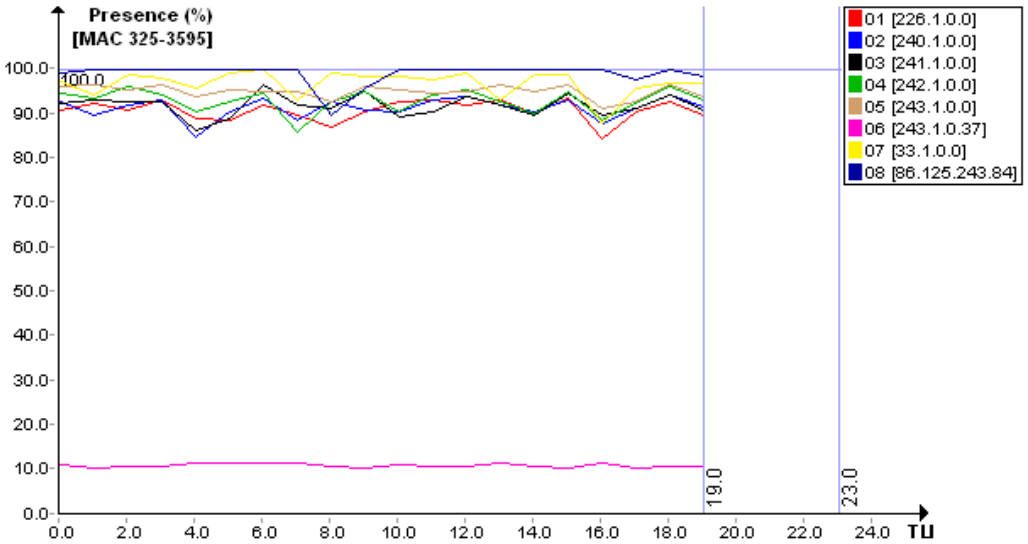


Fig. 5.6 Exemplu de variaţie a amprentei de trafic cu valori ridicate ale procentelor de prezenţă

Fig. 5.7 prezintă un caz cu trei categorii de destinaţii IP. Prima categorie constă în destinaţii IP cu procent ridicat de prezenţă, care variaza în intervalul 80% - 95%. A doua categorie conţine destinaţii IP cu prezenţă redusă, sub 10%, dar constantă. A treia categorie conţine două adrese IP care se opresc în jurul valorii de mijloc a intervalului şi nu mai apar ulterior.

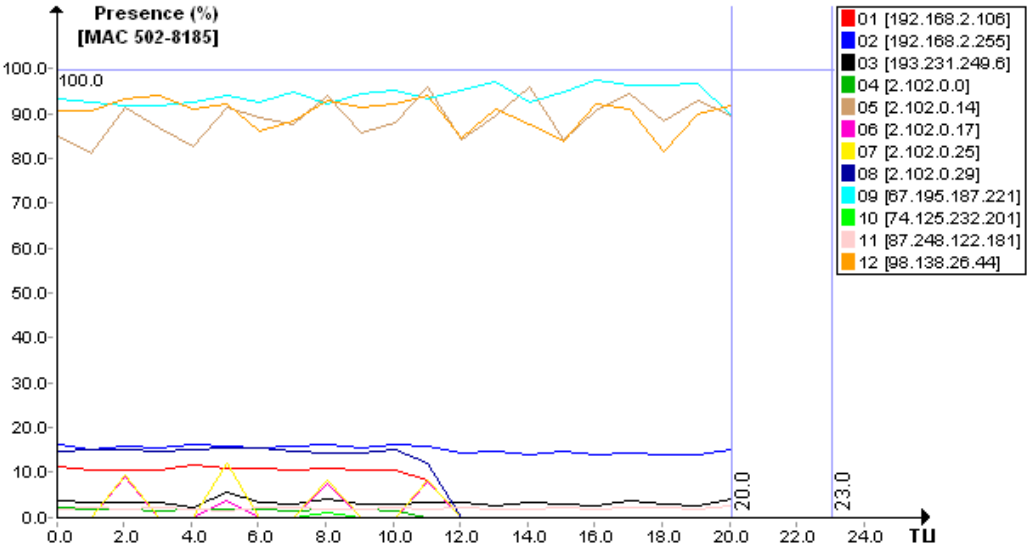


Fig. 5.7 Exemplu de variaţie a amprentei de trafic pentru 21TU

Următorul tip de variaţie constă în destinaţii pentru care procentul de prezenţă variază mult de la o unitate de timp la alta, dar totuşi destinaţiile se regăsesc de-a lungul timpului. Adresele pot fi incluse în amprentele de trafic, cu menţiunea că necesită o stabilire atentă a procentului de prezenţă luat în considerare. Iată câteva situaţii:

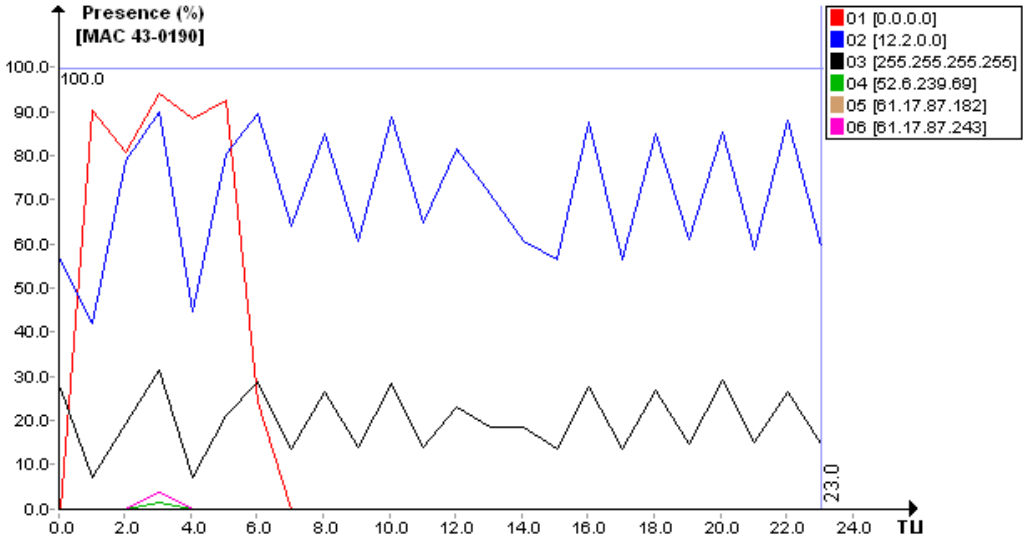


Fig. 5.8 Exemplu de variaţie a amprentei de trafic cu 2 destinaţii IP care prezintă trafic constant

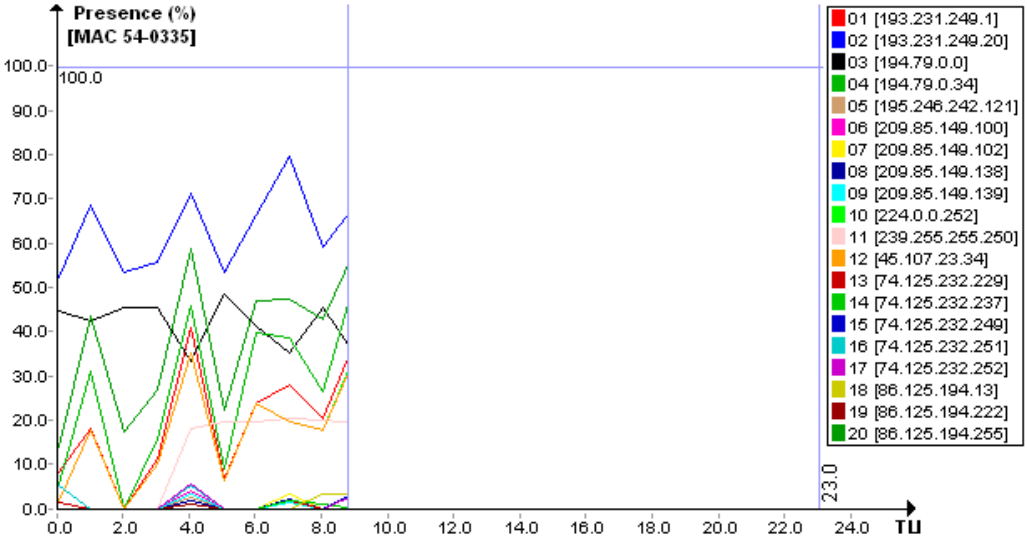


Fig. 5.9 Exemplu de variaţie a amprentei de trafic pentru 9TU, cu variaţii mari de la o unitate de timp la alta

Fig. 5.8 prezintă un caz cu câteva destinaţii IP. Două dintre acestea sunt prezente pe toată durata de 23TU (184 ore), variază puternic, dar se regăsesc pe întreg intervalul evaluat. O altă destinaţie prezintă valori ridicate ale procentului de prezenţă, dar numai pentru 5TU (40 ore), după care se opreşte.

Fig. 5.9 prezintă un număr mai mare de adrese, urmărite pe o durată de 9TU (72 ore). Variaţia lor este mare de la o unitate de timp la alta.

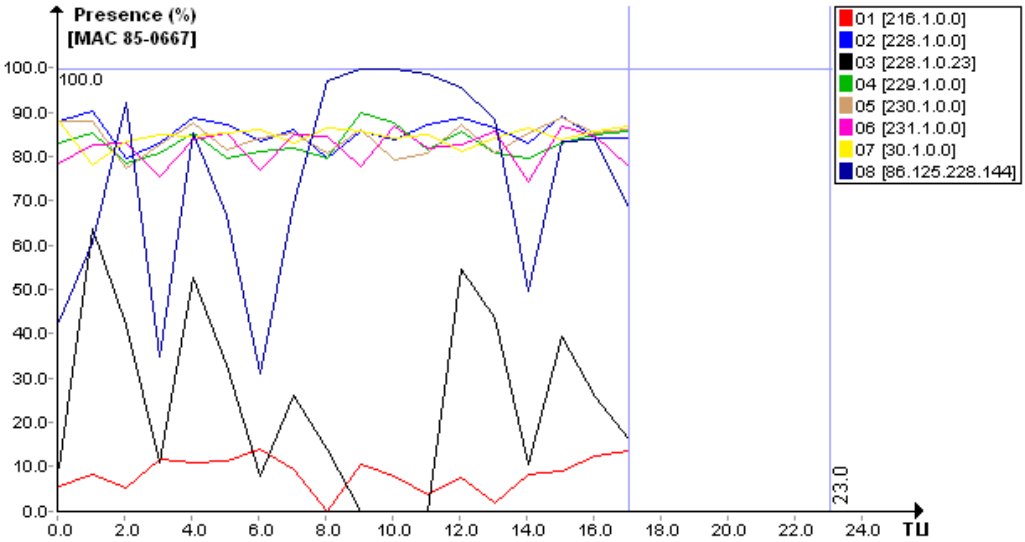


Fig. 5.10 Exemplu de variaţii puternice ale procentelor de prezenţă din cadrul amprentei de trafic, de la o unitate de timp la alta

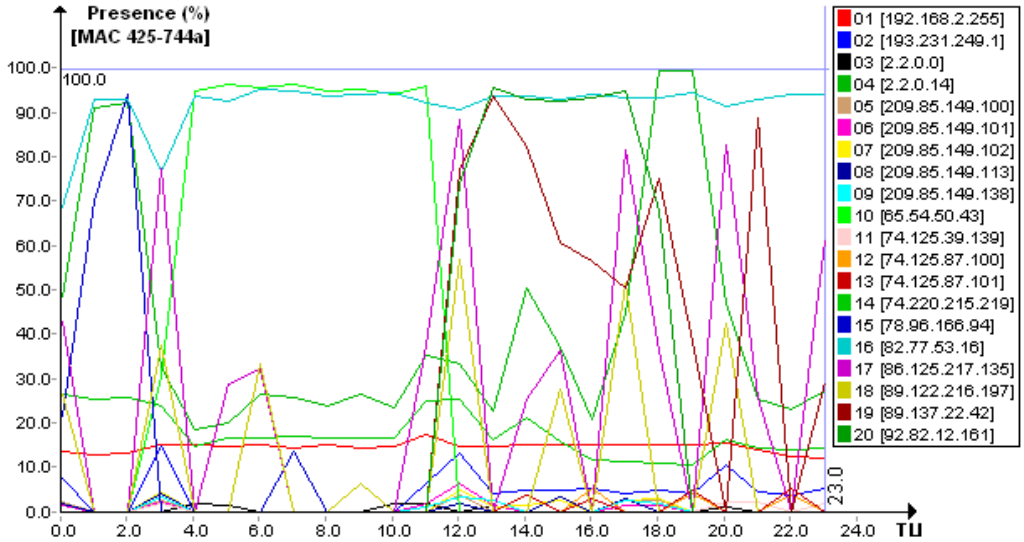


Fig. 5.11 Adresă MAC cu număr mare de destinaţii care prezintă trafic pe o durată de 23TU, dar cu variaţii mari ale procentelor de prezenţă

În Fig. 5.10 se identifică două tipuri combinate de adrese IP. Patru destinaţii IP au prezenţă cu variaţie mică şi la valori mari, cuprinse între 80% - 90%, alte două destinaţii prezintând însă variaţii mari.

Fig. 5.11 descrie un număr mult mai mare de destinaţii cu variaţii puternice. Totuşi, ceea ce interesează, este faptul că există adrese IP care se regăsesc în traficul unei staţii pe o durată mare de timp. În cazul de faţă, evaluarea s-a făcut pentru un interval de 23TU (184 ore).

Fig. 5.12 a urmărit un interval de 17TU (134 ore). Şi în cazul acestei staţii, un număr mare de adrese IP se regăsesc în traficul staţiei pe o durată mare de timp.

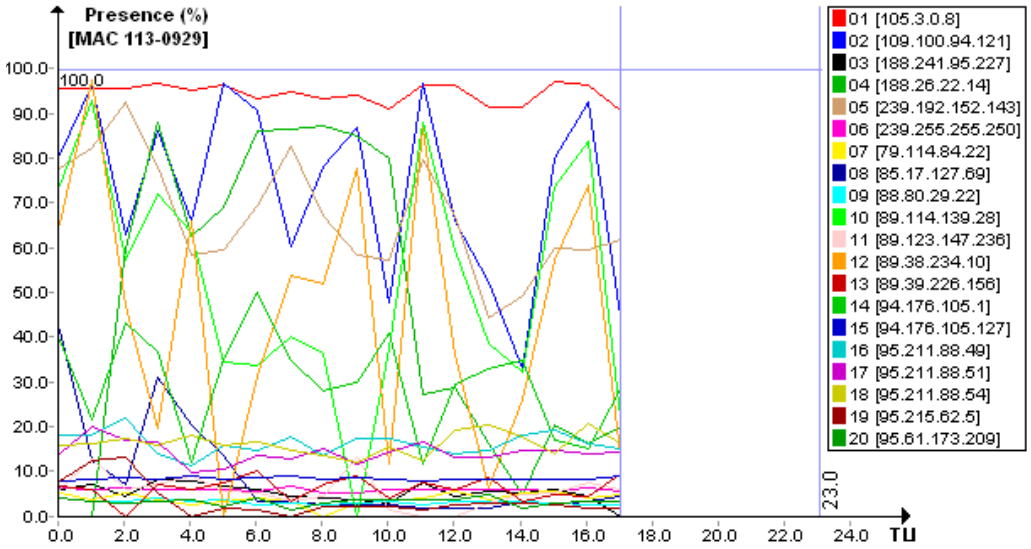


Fig. 5.12 Adresă MAC cu număr mare de destinaţii care prezintă trafic pe o durată de 17TU, cu variaţii mari ale procentelor de prezenţă pentru anumite destinaţii şi constant pentru altele

Fig. 5.13 prezintă patru destinaţii cu variaţie puternică. Una dintre ele chiar dispare pentru câteva TU, fapt care o invalidează de la posibilitatea includerii ei în amprenta de referinţă.

În Fig. 5.14 se observă foarte multe destinaţii cu prezenţă constantă, situată între 5% - 15%. Pe lângă acestea, alte două destinaţii IP variază puternic.

Şi în Fig. 5.15 se consideră un interval format din 23TU (184 ore), în care adresele IP cu variaţie mică se împletesc cu altele care prezintă variaţii puternice. Se observă de exemplu că una dintre destinaţii ajunge la o prezenţă de aproape 100%, dar pentru perioade scurte de timp.

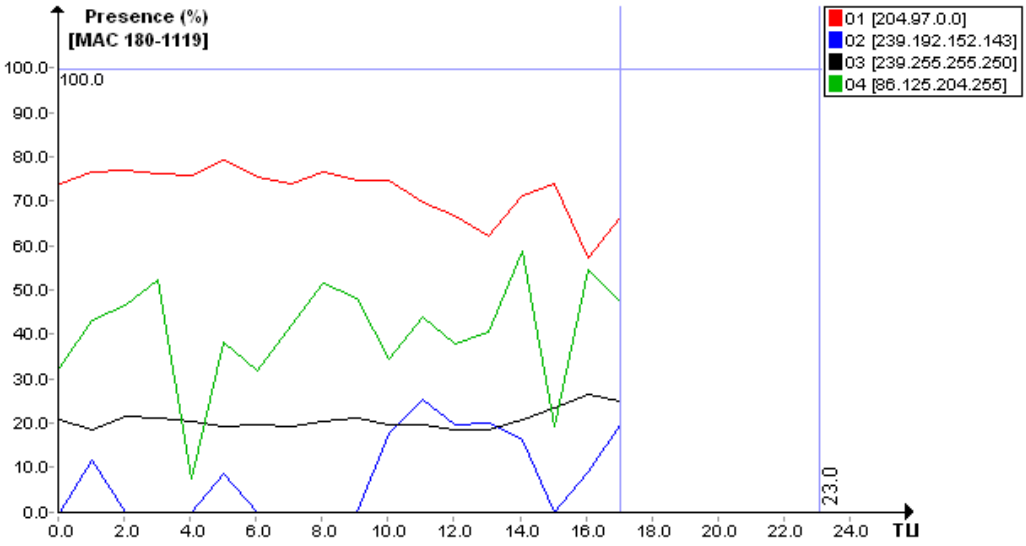


Fig. 5.13 Adresă MAC cu număr mic de destinaţii care prezintă trafic pe o durată de 17TU, dar cu variaţii mari ale procentelor de prezenţă

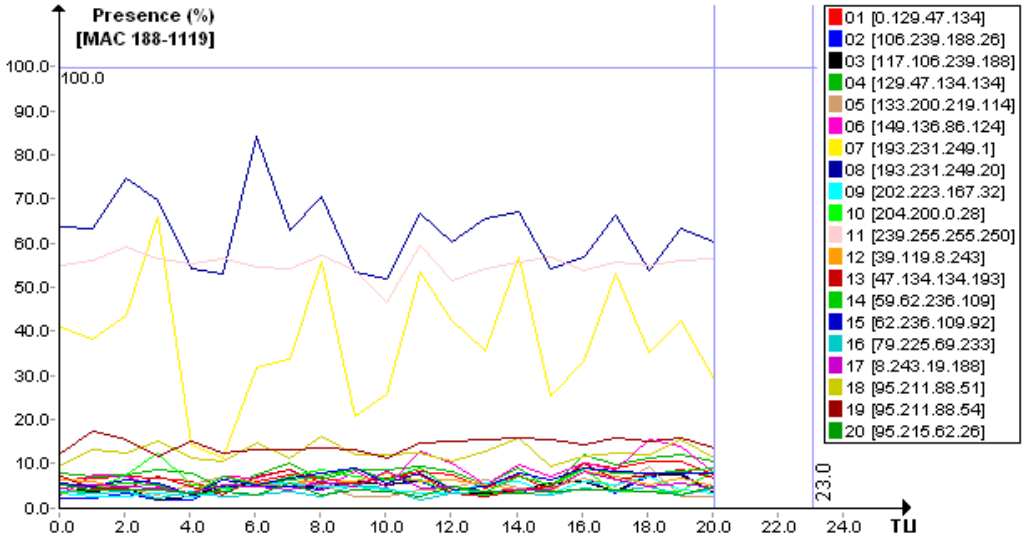


Fig. 5.14 Adresă MAC cu număr mare de destinaţii care prezintă trafic pe o durată de 21TU, cu variaţii mari ale procentelor de prezenţă pentru anumite destinaţii şi constant pentru altele

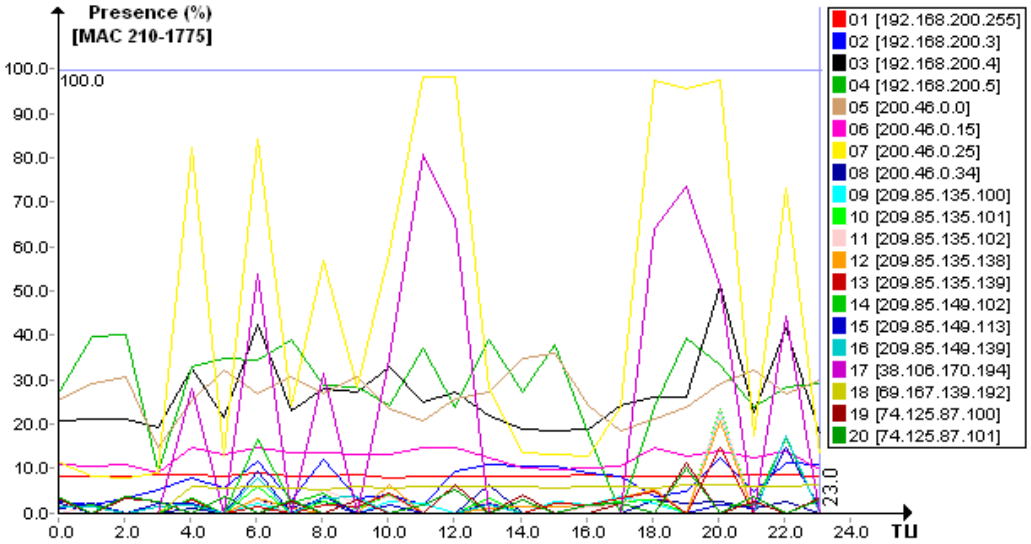


Fig. 5.15 Adresă MAC pentru care destinaţiile IP cu variaţii mari ale procentelor de prezenţă, se împletesc cu destinaţii IP cu variaţii mici

Fig. 5.16 prezintă o structură aparte, cu două destinaţii IP care prezintă un trafic constant pe toată durata de 23TU (184 ore) şi alte trei destinaţii IP care variază puternic.

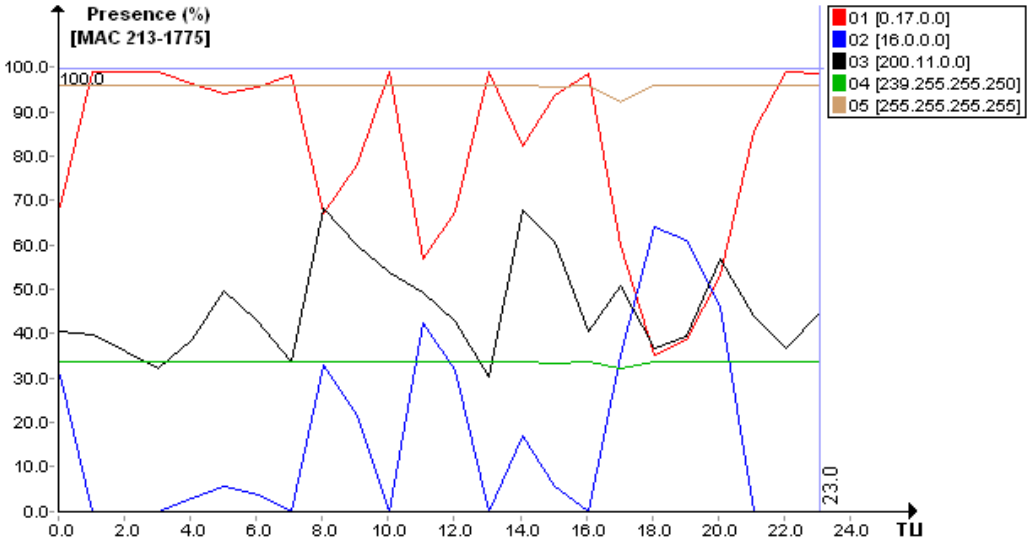


Fig. 5.16 Structură aparte ce conţine două destinaţii IP cu variaţii foarte mici şi altele cu variaţii foarte mari

Sunt situaţii, ca cea descrisă în Fig. 5.17, unde adresele cu variaţii mici sunt aproape absente. Cele mai multe variază puternic, dar se observă că persistă în trafic pe durate mari de timp.

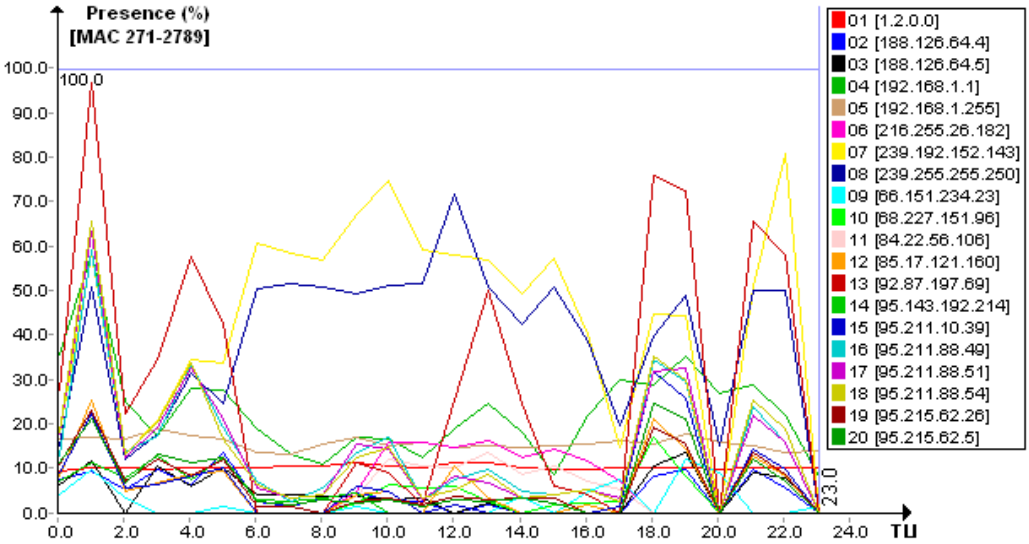


Fig. 5.17 Variaţia amprentei de trafic pentru o adresă MAC ce conţine număr mare de destinaţii IP

În Fig. 5.18, situaţia este asemănătoare cu cea din Fig. 5.17, în sensul că variaţiile procentelor de prezenţă sunt mari de la o unitate de timp la alta.

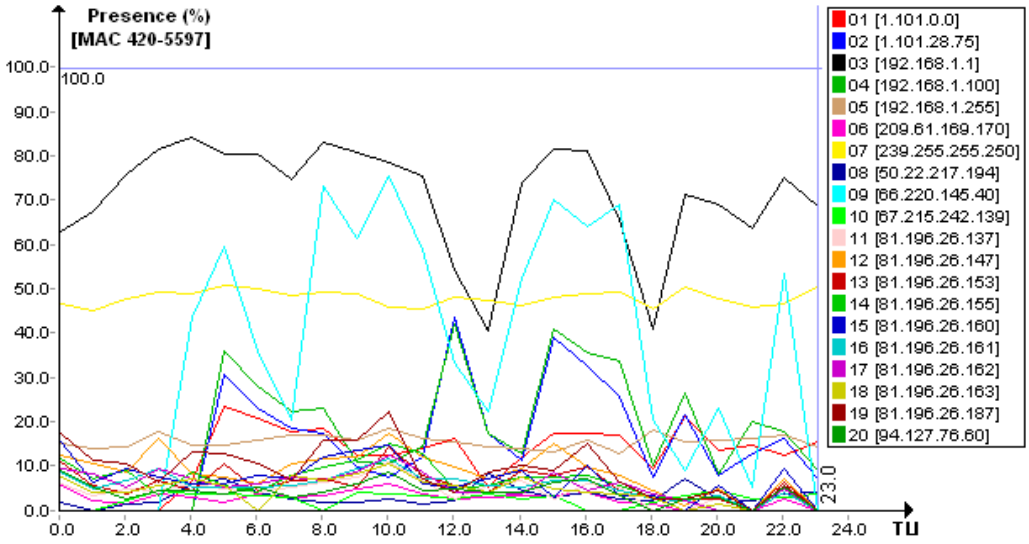


Fig. 5.18 Exemplu de variaţie a amprentei de trafic ce conţine suficiente informaţii pentru identificarea adresei MAC

Fig. 5.19 prezintă un caz cu o serie de destinaţii IP cu variaţii nu foarte mari, situaţie suficient de bună pentru ca identificarea adresei MAC să se realizeze în mod corespunzător.

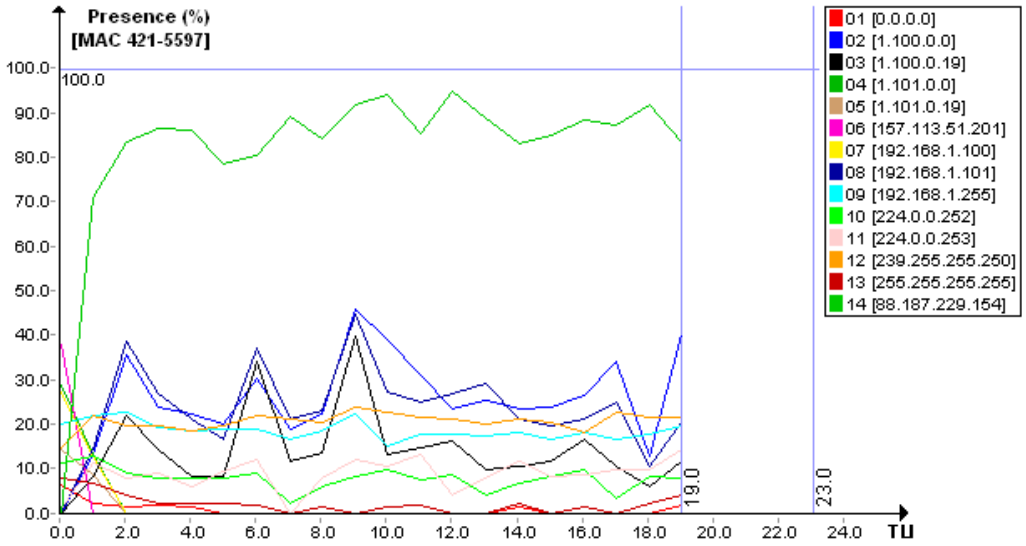


Fig. 5.19 Alt exemplu de variaţie a amprentei de trafic

Fig. 5.20 aduce în atenţie din nou o situaţie combinată, în care adresele IP cu variaţii mici se regăsesc împreună cu adrese IP cu variaţii mai mari. Variaţiile acestea nu încurcă procesul de recunoaştere a adresei MAC, trebuie doar ca amprenta de referinţă să fie configurată corespunzător.

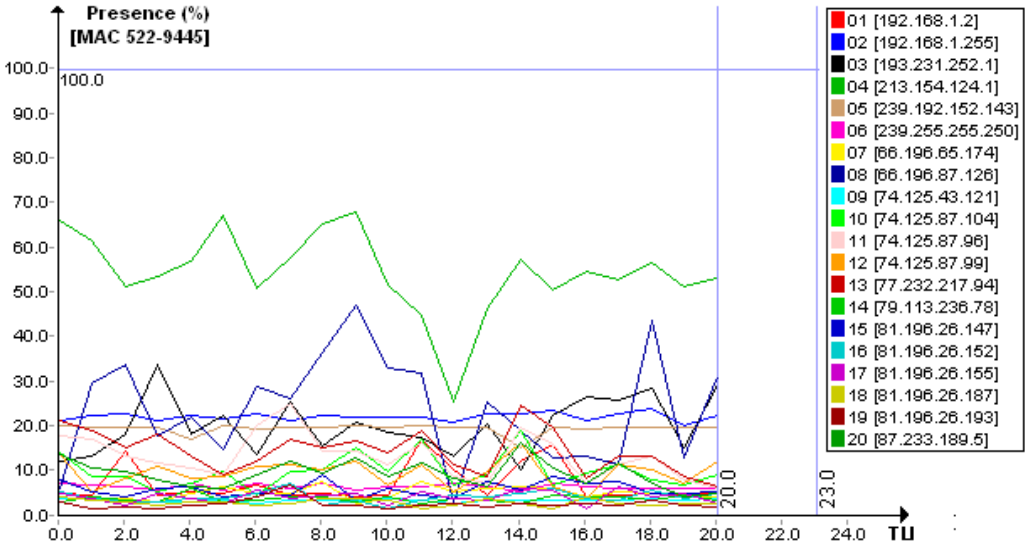


Fig. 5.20 Exemplu de variaţie a amprentei de trafic, cu procente mici de prezenţă

5.3. Concluzii.

Rezultatele obţinute relevă faptul că în timp, există destinaţii IP care se păstrează în traficul unei adrese MAC, şi pot defini o amprentă de trafic. Ţinând cont de durata experimentului şi de numărul calculatoarelor, faptul că s-au identificat un număr semnificativ de cazuri în care traficul spre anumite destinaţii este prezent în mod constant, dovedeşte utilitatea metodei.

Procentul de prezenţă al unei destinaţii IP poate varia în timp mai mult sau mai puţin, dar cel mai important lucru este faptul că adresa poate fi întâlnită în multe unităţi de timp succesive. Variaţii mari pentru procentul de prezenţă complică procesul de stabilire a valorii de referinţă pentru întregul interval, dar cu o configurare atentă, chiar şi aceste situaţii pot fi transformate în componente ale unor amprente de referinţă.

Destinaţiile IP care prezintă variaţii mici ale procentelor de prezenţă, sunt de dorit în componenţa amprentelor de referinţă şi conferă stabilitate procesului de validare în timp real. După cum s-a putut observa în paragrafele precedente, cazuri de genul acesta sunt frecvente. Astfel, dacă se aleg ca amprentă de trafic, destinaţiile IP care se repetă de-a lungul unităţilor de timp, validarea adreselor MAC va fi posibilă cu rate ridicate ale Gradului Global de Recunoaştere.

6. CONCLUZII, CONTRIBUŢII ADUSE ŞI DEZVOLTĂRI PENTRU VIITOR

În cadrul acestui capitol se prezintă concluziile finale şi contribuţiile cele mai importante aduse de către autor, atât ca metode cât şi ca unelte software de testare. Sunt descrise posibile direcţii de dezvoltare pentru viitor, care au ca scop creşterea fiabilităţii aspectelor prezentate.

6.1. Concluzii finale.

Lucrarea de faţă este dedicată cercetărilor din domeniul securizării reţelelor de calculatoare împotriva pătrunderilor neautorizate prin falsificarea adreselor MAC. Problemele abordate sunt de actualitate şi au o importanţă vitală, dat fiind faptul că frecvenţa şi gravitatea atacurilor cibernetice a crescut mult în ultimul timp.

Obiectivul principal al lucrării îl constituie identificarea cât mai rapidă şi cât mai exactă a identităţii unei adrese MAC întâlnite în trafic. Pentru atingerea obiectivului propus, au fost necesare următoarele direcţii de cercetare:

- o analiză critică a stadiului actual în domeniul pătrunderilor neautorizate în reţelele de calculare, prin falsificarea adreselor MAC;
- dezvoltarea unei metode noi de detecţie a falsificării adreselor MAC, care să valideze identitatea unei staţii pe baza comparaţiei dintre amprenta de referinţă şi cea actuală;
- modelarea, implementarea şi analizarea fiabilităţii sistemului de detecţie propus.

Abordarea sistemică a detecţiei adreselor MAC falsificate conferă o formalizare a problematicii studiate, prin care se stabilesc principalele componente ale sistemului şi modul în care acestea sunt interconectate. Reţeaua de calculatoare este văzută sub forma unor staţii individuale care generează trafic de date. Accesul unei staţii în reţea este permis numai pe baza autorizării adresei MAC. Reţeaua este permanent monitorizată şi orice încercare de pătrundere a unei staţii neautorizate este semnalată rapid.

Analiza critică a stadiului actual în domeniul pătrunderilor neautorizate prin falsificarea adreselor MAC, a dus la obţinerea unei imagini de ansamblu cu privire la modul în care este tratată în practică problema falsificării adreselor MAC.

Metodele clasice de detecţie încearcă să valideze identitatea unei adrese MAC, fie prin analizarea unor parametrii din cadrul pachetelor de date monitorizate în reţea, fie prin amprentarea echipamentului care emite pachetelele de date. Fiecare metodă a fost sintetizată cu atenţie, punând în evidenţă avantajele şi dezavantajele pe care le oferă.

Analiza comparativă a metodelor prezentate a luat în calcul răspunsul pe care aceste metode îl au în diverse situaţii specifice, cum ar fi:

- detectarea unui intrus care vine în reţeaua locală, chiar în locul staţiei autorizate;
- limitarea metodelor la reţele wired sau wireless;
- necesitatea instalării unui software pe staţia client;
- aplicabilitate în absenţa suprapunerii traficului provenit de la sursa autorizată, cu cel provenit de la sursa neautorizată;
- aplicabilitate în pofida mobilităţii staţiei client şi trecerii acesteia dintr-o subreţea în alta;
- aplicabilitate în pofida utilizării de către intrus a unui echipament identic cu cel autorizat;
- aplicabilitate pentru echipamente de tip Desktop/Laptop.

Analiza comparativă a dus la concluzia că cele mai multe dintre metodele folosite în prezent pentru validarea adreselor MAC întâlnite în trafic, au valabilitate limitată. Situaţiile specifice menţionate anterior sunt des întâlnite în trafic, dar din păcate lipsesc abordări care să asigure validarea adreselor MAC în toate aceste situaţii specifice.

În cadrul tezei, s-a dezvoltat o metodă nouă de detecţie a adreselor MAC falsificate, metodă care acoperă toate cazurile menţionate anterior. Metoda a fost numită „Destination Traffic Fingerprint” (DTF) şi foloseşte pentru validare o amprentă de referinţă, alcătuită din adresele IP ale destinaţiilor cu care staţia comunică în mod constant.

Amprentele sunt reprezentate matematic sub forma unor mulţimi de perechi, unde fiecare pereche este compusă dintr-o destinaţie IP şi procentul de prezenţă pe care îl are în timp traficul către destinaţia respectivă.

Domeniul de aplicabilitate al metodei s-a demonstrat a fi suficient de larg, pentru ca algoritmul să poată fi utilizat în practică. Pentru ca metoda să fie utilizabilă pentru o anumită staţie din reţea, este necesar ca să se poată genera o amprentă de referinţă pentru staţia respectivă. După cum s-a dovedit în capitolele precedente, modul de utilizare al calculatoarelor, permite în cele mai multe cazuri extragerea unei amprente de referinţă. Sunt rare cazurile în care nu se poate extrage o amprentă de referinţă, şi aceste cazuri de obicei se întâlnesc în reţele cu acces public, unde o staţie este folosită de către un număr foarte mare de utilizatori.

În plus, s-a demonstrat în teza de faţă faptul că domeniul de aplicabilitate este extins şi datorită utilizării pe scară largă a unor tehnologii şi servicii, care favorizează aplicarea metodei DTF, cum ar fi:

- utilizarea aplicaţiilor software de tip ERP;
- comunicarea între angajaţi prin intermediul unor servere de e-mail proprii;
- utilizarea reţelelor VPN;
- virtualizare şi cloud.

Determinarea traficului constant reprezintă partea cea mai importantă din întreg procesul de validare. Din totalitatea destinaţiilor IP cu care o staţie comunică în reţea, trebuie filtrate adresele care sunt asociate cu un trafic temporar sau punctual, astfel încât să rămână în final doar cele aferente traficului constant. Evaluarea traficului s-a realizat la nivel de minut, verificându-se nu „cantitatea” de trafic ci prezenţa sau absenţa traficului.

În acest context, traficul constant s-a definit ca fiind cel pentru care minutele de prezenţă sunt relativ constante pe o durată mare de timp. Traficul temporar a fost definit ca fiind traficul pentru care minutele de prezenţă sunt

regăsite doar în anumite subintervale de timp din intervalul evaluat iar traficul punctual s-a definit ca fiind traficul pentru care minutele de prezenţă se regăsesc izolat pe axa timpului.

Pentru a caracteriza traficul constant, au fost propuşi următorii parametrii:

- Procentul de Prezenţă;
- Procentul de Absenţă Maximă;
- Criteriul de prezenţă pe subintervale egale;
- Puterea amprentei de referinţă.

Procentul de Prezenţă s-a stabilit ca fiind raportul dintre numărul de minute în care există trafic către o destinaţie anume, şi numărul total de minute aferente perioadei de evaluare. Acest parametru este foarte important în procesul de stabilire a traficului constant, dar nu este o măsură suficientă datorită faptului că nu oferă informaţii despre variaţia sa de-a lungul perioadei de timp evaluate.

S-a demonstrat în lucrarea de faţă faptul că există situaţii în care valoarea procentului de prezenţă este constantă în timp sau variază foarte puţin. Scopul metodei DTF este să determine cu precizie aceste situaţii şi să includă în amprenta de referinţă destinaţiile IP aferente unui astfel de trafic.

Totuşi, valoarea procentului de prezenţă nu este suficientă pentru stabilirea unui trafic ca fiind „constant". O valoare ridicată a procentului de prezenţă pe anumite subintervale, ar putea „ascunde" absenţa traficului în restul timpului şi ar putea conduce la o interpretare greşită, însoţită de inserarea în amprenta de referinţă a unor destinaţii care nu prezintă trafic constant. Din acest motiv, în caracterizarea traficului constant trebuie să fie obligatoriu să se ţină cont de următorii parametrii.

Procentul de Absenţă Maximă este următorul parametru care poate fi utilizat în scopul determinării traficului constant. Procentul acesta urmăreşte să determine cât reprezintă cea mai mare „pauză" de trafic, din totalul timpului evaluat şi este definit ca fiind raportul dintre numărul maxim de minute în care nu a avut loc trafic către o destinaţie, şi numărul total de minute evaluate.

Parametrul acesta ajută în determinarea traficului constant, în sensul că scoate în evidenţă anumite cazuri de trafic temporar sau punctual. Totuşi, problema principală a acestui indicator este aceea că e puternic influenţat de perioada de timp evaluată. Din acest motiv, folosirea Procentului de Absenţă Maximă în vederea caracterizării traficului constant, nu este o măsură concludentă.

Criteriul de prezenţă pe subintervale egale este mult mai relevant decât Procentul de Absenţă Maximă şi poate identifica adresele IP cu trafic temporar sau punctual. Criteriul acesta împarte intervalul de timp evaluat într-un număr de subintervale egale şi verifică traficul pe fiecare din aceste subintervale. Traficul constant este declarat cel regăsit în toate sau aproape toate subintervalele.

Puterea amprentei de referinţă este o valoare numerică şi reprezintă o măsură asupra calităţii amprentei de referinţă. Cu cât valoarea aceasta este mai mare, cu atât calitatea amprentei este mai bună. Evaluarea se poate face prin intermediul unor factori, cum ar fi: stabilitatea, credibilitatea şi viteza de validare. Două amprente de referinţă stabilite pentru aceeaşi adresă MAC, vor fi comparate pe baza puterii calculate, iar sitemul va păstra pe cea care prezintă valoarea

maximă. Calculul puterii amprentei se realizează prin însumarea procentelor de prezenţă ale destinaţiilor IP din componenţa amprentei.

Pentru validarea identităţii adreselor MAC s-a introdus Gradul Global de Recunoaştere, calculat pe baza comparaţiei dintre amprenta de referinţă şi amprenta actuală. Rezultatul este dat sub forma unui procent, care comunică o măsură a recunoaşterii adresei MAC. Calculul se poate realiza în manieră standard sau ponderat.

Calculul standard determină un grad individual de recunoaştere pentru fiecare adresă IP care intră în componenţa amprentei de referinţă, şi apoi stabileşte Gradul Global de Recunoaştere ca fiind media aritmetică a gradelor individuale. Gradele individuale sunt stabilite ca raport între procentul de prezenţă actual cu cel din amprenta de referinţă.

Calculul ponderat ţine cont de valoarea procentului de prezenţă a fiecărei destinaţii ce intră în componenţa amprentei de referinţă. Gradele individuale intră în calculul Gradului Global cu o anumită pondere, calculată prin intermediul raportului dintre procentul de prezenţă şi valoarea medie a puterii amprentei de referinţă.

S-a demonstrat în lucrarea de faţă, faptul că varianta ponderată de calcul a Gradului Global de Recunoaştere este mult superioară calculului standard, întrucât are capacitatea de a oferi rezultate bune chiar şi în prezenţa variaţiilor temporare de trafic. Absenţa temporară a traficului către o destinaţie cu prezenţă redusă, poate duce la interpretări greşite în cazul calculului standard.

Modelarea matematică a funcţionării metodei DTF, atât pentru determinarea amprentei de referinţă, cât şi pentru validarea adreselor MAC în timpul funcţionării. permite o abordare sistemică a problematicii studiate, prezentând formalismele fiecărui subsistem şi modul în care aceste subsisteme sunt legate între ele pentru obţinerea funcţionării de ansamblu. Pentru fiecare bloc din schemă s-a formalizat matematic mărimea sau mărimile de intrare, şi s-a determinat modul de calcul al ieşirii.

Modelarea Fuzzy permite identificarea traficului constant pe baza relaţiilor şi regulilor definite în cadrul unui sistem de tip Mandami, sistem definit cu patru variabile de intrare şi o ieşire. Fiecare din cele patru variabile de intrare reprezintă o subdiviziune a intervalului de timp evaluat, şi are asociat procentul de prezenţă calculat pe subintervalul respectiv. Valorile funcţiilor de apartenenţă, definite: „Continuu”, „Prezenţă-Ridicată”, „Prezenţă-Medie”, „Prezenţă-Mică” şi respectiv „Absent”, încadrează valoarea procentului de prezenţă asociat subintervalului. Regulile definite vor ţine cont de aceste valori pentru cele patru subintervale, şi vor genera la ieşire răspunsul sistemului, sub forma unei valori care poate fi: „Constant” , „Aproape Constant”, „Variabil”, „Puternic Variabil” şi respectiv „Ocazional”.

Toolbox-ul software conceput în vederea dezvoltării metodei DTF, oferă un mediu de simulare foarte util, ce permite observarea tuturor fazelor de lucru, de la generarea amprentei de referinţă şi până la identificarea adreselor MAC. Rezultatele obţinute cu ajutorul simulatorului atestă avantajele metodei DTF în domeniul detecţiei adreselor MAC falsificate.

În final, trebuie remarcat faptul că teza de faţă aduce contribuţii semnificative în domeniul detecţiei adreselor MAC falsificate, prin conceperea

metodei „Destination Traffic Fingerprint”, o abordare nouă, originală, ce oferă o serie de avantaje, cum ar fi:

- costuri reduse pentru monitorizare şi control;
- extragerea amprentei de referinţă se realizează fără să fie necesară o intervenţie directă asupra staţiei evaluate, ci doar prin monitorizarea traficului într-un punct al reţelei
- extragerea amprentei de referinţă se poate realiza în timpul normal de funcţionare al staţiei, fără să fie necesară întreruperea activităţii utilizatorilor;
- detectează atât atacuri provenite din exteriorul reţelei, cât şi cele provenite din interior, chiar în locul staţiei autorizate;
- nu are limitări cu privire la tehnologia de implementare a reţelei (wired sau wireless);
- semnalieză intrusul indiferent dacă acesta încearcă să pătrundă în reţea atunci când staţia autorizată este oprită sau când staţia autorizată este pornită;
- se poate aplica şi pentru utilizatorii mobili, care trec dintr-o subreţea în alta;
- stabileşte identitatea adresei MAC, indiferent dacă intrusul foloseşte un echipament identic sau diferit de cel autentic;

În concluzie, folosirea metodei DTF pentru stabilirea şi validarea identităţii adreselor MAC întâlnite în trafic, reprezintă o abordare atractivă şi se recomandă în contextul în care atacurile cibernetice sunt frecvente şi duc la consecinţe negative uneori deosebit de grave.

Semnalarea din timp a utilizatorilor neautorizaţi care încearcă să pătrundă în reţea, oferă posibilitatea de reacţie rapidă, în vederea localizării intrusului şi luarea măsurilor necesare pentru fiecare situaţie particulară în parte.

6.2. Contribuţii originale.

Pornind de la obiectivele declarate, în continuare se prezintă principalele contribuţii originale:

- realizarea unei analize sistemice a procesului de detecţie a adreselor MAC falsificate;
- elaborarea unui studiu critic al metodelor actuale de detecţie a falsificării adreselor MAC şi evaluarea acestor metode prin prisma răspunsului pe care îl oferă în diverse situaţii specifice;
- elaborarea unei noi metode de detecţie a adreselor MAC falsificate, pe baza unei amprente de trafic alcătuite din destinaţiile IP cu care o staţie comunică în mod constant, metodă denumită Destination Traffic Fingerprint (DTF);
- definirea traficului constant, temporar şi punctual;

- definirea matematică a procentului de prezenţă al unei destinaţii IP;
- definirea matematică a procentului de absenţă maximă a unei destinaţii IP;
- definirea criteriului de prezenţă pe subintervale egale;
- definirea matematică a puterii amprentei de referinţă şi a mediei puterii amprentei de referinţă;
- definirea calităţii amprentei de referinţă, prin stabilitate, credibilitate şi viteză de validare;
- definirea matematică a relaţiilor ce caracterizează amprenta de referinţă şi amprenta actuală;
- definirea matematică a gradului de recunoaştere a unei destinaţii, în variantă standard şi variantă ponderată;
- definirea Gradului Global de Recunoaştere al unei adrese MAC, în variantă standard şi variantă ponderată;
- stabilirea modului de validare a unei adrese MAC, pe baza Gradului Global de Recunoaştere;
- formalizarea matematică a spaţiului de adrese IP v.4;
- formalizarea matematică a unui pachet de date din reţea;
- dezvoltarea unui model sistemic destinat determinării amprentei de referinţă, cu definirea formalismelor subsistemelor componente;
- dezvoltarea unui model sistemic destinat validării în trafic a unei adrese MAC, cu definirea formalismelor subsistemelor componente;
- dezvoltarea unui model Fuzzy pentru determinarea traficului constant;
- analiză comparativă a generării amprentei de referinţă, în manieră standard respectiv modelare Fuzzy;
- analiza unor servicii/tehnologii care favorizează utilizarea metodei DTF;
- implementare în Matlab a celor două variante (standard şi ponderată) de determinare a Gradului Global de Recunoaştere;
- realizarea unei aplicaţii software numită „Packet Recorder”, care are capacitatea de a capta pachetele din reţea şi de a le transforma într-un tipar necesar aplicării metodei DTF;
- dezvoltarea unui simulator de reţea, numit „Network Detector”, care include următoarele funcţii importante:
    - alcătuirea unei reţele virtuale cu date provenite din arhivele realizate de Packet Recorder;
    - parametrizarea generării amprentelor de referinţă;
    - generarea şi memorarea amprentelor de referinţă;
    - generarea unor rapoarte de urmărire a variaţiei amprentelor de-a lungul mai multor unităţi de timp;
    - generarea unor grafice pentru observarea distribuţiei traficului către destinaţiile IP;
    - simularea unei reţele în care este activă detecţia DTF, cu următoarele facilităţi:
        - selectarea şi configurarea datelor care intră în simulator;
        - calculul Gradului Global de Recunoaştere şi vizualizarea grafică a gradului de recunoaştere pentru staţiile ce funţionează în reţea;
        - informaţii actualizate despre amprentele actuale;
        - determinarea amprentelor în timpul derulării simulării;
        - stabilirea modului de calcul al Gradului Global de Recunoaştere, chiar şi în timpul simulării;

  - oprirea forţată a unei staţii (simulatorul ignoră traficul staţiei respective);
  - falsificarea adresei MAC, astfel încât o staţie să îşi poată schimba adresa MAC şi în felul acesta să se observe identificarea realizată prin intermediul metodei DTF;
- realizarea unui test experimental pentru validarea strategiei propuse.

6.3. Direcţii de cercetare generate de studiile efectuate.

Rezultatele obţinute în cadrul acestei teze pot fi dezvoltate prin continuarea cercetării în domeniul abordat, în câteva direcţii, dintre care se pot aminti:

- adaptarea automată a amprentelor de referinţă, în cazul unor schimbări apărute în traficul constant generat de staţia în cauză;
- optimizarea calculului Gradului Global de Recunoaştere, prin introducerea altor modele, separat de varianta standard şi varianta ponderată;
- reducerea timpului necesar validării identităţii unei adrese MAC;
- completarea detecţiei adreselor MAC falsificate, cu algoritmi de localizare a intrusului;
- completarea detecţiei adreselor MAC falsificate, cu elemente de eliminare automată a intrusului care a pătruns în reţea:

Bibliografie.

[Ahm-13] A.M.Y. Ahmed, Q. Depei, "An optimization of security and trust management in distributed systems", IACC 2013, ISBN: 978-1-4673-4527-9

[Ara-10] C. Arackaparambil, S. Bratus, A. Shubina, D. Kotz, „On The Reliability of Wireless Fingerprinting using Clock Skews”, ACM WiSec 2010, ISBN 978-1-60558-923-7

[Avi-04] Algirdas Avizienis, Jean-Claude Laprie, Brian Randell, Carl Landwehr, „Basic concepts and Taxonomy of Dependable and Secure Computing”, IEEE Trasactions on Dependable and Secure Computing, 2004

[Ban-08] Richa Bansal, Siddharth Tiwari, Divya Bansal, “Non-Cryptographic Methods of MAC Spoof Detection in Wireless LAN”, ICON 2008, 978-1-4244-3805-1/08

[Bar-12] I.B. Barla, D.A. Schupke, G. Carle, "Virtual Network Simulator Architecture", UKSim 2012, ISBN: 978-1-4673-1366-7

[Bea-09] A. Beach, M. Gartrell, R. Han, „Solutions to Security and Privacy Issues in Mobile Social Networking”, CSE 2009, ISBN 978-1-4244-5334-4

[Bor-09] K.C. Borries, G. Judd, D.D. Stancil, P. Steenkiste, "FPGA-Based Channel Simulator for a Wireless Network Emulator", VTC 2009, ISBN: 978-1-4244-2517-4

[Bra-08] S. Bratus, C. Cornelius, D. Kotz, D. Peebles, „Active Behavioral Fingerprinting of Wireless Devices”, ACM, 2008, 978-1-59593-814-5

[Bri-08] V. Brik, S. Banerjee, M. Gruteser, S. Oh, „Wireless Device Identification with Radiometric Signatures”, ACM MobiCom 2008, ISBN 978-1-60558-096-8

[Bul-09] D. Bultmann, M. Muhleisen, K. Klagges, M. Schinnenburg, "OpenWNS - open Wireless Network Simulator", EW 2009, ISBN: 978-1-4244-5935-3

[Bzo-11] P. Bzoch, J. Safarik, "Security and reliability of distributed file systems", IDAACS 2011, ISBN: 978-1-4577-1426-9

[Car-08] L. Carter, J. Dyal, S. Doshi, R. Bagrodia, "A hardware-in-the-loop (HWIL) network simulator for analysis, and evaluation of large-scale military wireless communication systems", MILCOM 2008, ISBN: 978-1-4244-2676-8

[Cha-09] G. Chandrasekaran, J. Francisco, V. Ganapathy, M. Gruteser, W. Trappe, „Detecting Identity Spoofs, in IEEE 802.11e Wireless Networks”, IEEE GLOBECOM Proceedings, 2009, 978-1-4244-4148-8

[Chu-12] L. Chunli, L. Donghui, „Computer Network Security Issues and Countermeasures”, ISRA 2012, ISBN 978-1-4673-2205-8

[Cor-11] E. Corchado, Á. Herrero, "Neural Visualization of Network Traffic Data for Intrusion Detection", Applied Soft Computing, Volume 11, Issue 2, 2011

[Del-10] P. De Lutiis, „Managing Home Networks Security Challenges Security Issues and Countermeasures”, ICIN 2010, ISBN 978-1-4244-7443-1

[Dik-10] T.K. Dikaliotis, A.G. Dimakis, T. Ho, "Security in distributed storage systems by communicating a logarithmic number of bits", ISIT 2010, ISBN: 978-1-4244-7890-3

[Dis-11] S. Distefano, A. Puliafito, "Achieving Distributed System Information Security", CIS 2011, ISBN: 978-1-4577-2008-6

[Drz-07] L. Drzewiecki, M. Antoniak-Lewandowska, "Flow Simulator - a flow-based network simulator", EUROCON 2007, ISBN: 978-1-4244-0813-9

[Edm-09] M. Edman, B. Yener, „Active Attacks Modulation-based Radiometric Identification. Technical Report 09-02”, 2009

[Elk-94] C. Elkan, H.R. Berenji, B. Chandrasekaran, C.J.S de Silva, „The Paradoxical Success of Fuzzy Logic”, IEEE Expert, Volume 9, Issue 4, 1994, ISSN: 0885-9000

[Esc-08] R. Escola, C. Puzat, A. Chaffiol, B. Yvert, I.E. Magnin, R. Guillernaud, "SIMONE: A Realistic Neural Network Simulator to Reproduce MEA-Based Recordings", IEEE Transactions on Neural Systems and Rehabilitation Engineering, Volume 16, Issue 2, ISSN: 1534-4320

[Esf-10] A. Esfandi, „Challenges and problems in the ERP implementation and its application”, ICCSIT 2010, ISBN 978-1-4244-5537-9

[Esw-13] T. Eswari, V. Vanitha, "A novel rule based intrusion detection framework for Wireless Sensor Networks", ICICES 2013, ISBN 978-1-4673-5786-9

[Fek-07] S.P. Fekete, A. Kroller, S. Fischer, D. Pfisterer, "Shawn: The fast, highly customizable sensor network simulatorShawn: The fast, highly customizable sensor network simulator", INSS 2007, ISBN: 1-4244-1231-5

[Gal-08] E. Galli, G. Cavarreta, S. Tucci, "HLA-OMNET++: An HLA Compliant Network Simulator", DS-RT 2008, ISBN: 978-0-7695-3425-1

[Gao-10] K. Gao, C. Corbett, R. Benyah, „A Passive Approach to Wireless Device Fingerprinting”, IEEE/IFIP International Conference on Dependable Systems and Networks, 2010, ISBN 978-1-4244-7500-1

[Gao-11] H. Gao, J. Hu, T. Huang, J. Wang, Y. Chen, „Security Issues in Online Social Networks”, IEEE Internet Computing, Volume 15, Issue 4, ISSN 1089-7801

[Gau-13] N. Gaurha, D. Mishra, P. Trivedi, "Data Security in Distributed System Using Fully Homomorphic Encryption and Linear", CSNT 2013, ISBN: 978-1-4673-5603-9

[Goe-09] Shikha Goel, Sudesh Kumar, “An Improved Method of Detecting Spoofed Attack in Wireless LAN”, NetCom - First International Conference on Networks and Communications, 978-0-7695-3924-9/09, 2009

[Gon-07] M.E. Gonzalez, "A Generalized Packet Traffic Simulator for 4G Network Dimensioning Tools", VTC 2007, ISBN: 1-4244-0266-2

[Gup-10] M. Gupta, S. Malhotra, „Comparison of Security Issues in Wireless Networks”, ICSTE 2010, ISBN 978-1-4244-8667-0

[Hos-07] H. Hossain, M. Ahmed, A. Al-Nayeem, T.Z. Islam, M.M. Akbar, "Gpnocsim - A General Purpose Simulator for Network-On-Chip", ICICT 2007, ISBN: 984-32-3394-8

[STD-802.11] 802.11-1997 – IEEE Standard for Information Technology – Telecommunications and Information Exchange Between Systems-Local and Metropolitan Area Networks-Specific Requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications

[Ish-01] H. Ishihashi, N.Yamai, K. Abe, T. Matsura, „A Protection Method against Unauthorized Access and Address Spoofing for Open Network Access Systems”, IEEE, 2001

[Jan-08] S. Jana, and S.K. Kasera, “On Fast and Accurate Detection of Unauthorized Wireless Access Points Using Clock Skews”, *MobiCom '08, September 14-19, 2008, San Francisco, California, USA*, ACM 978-1-60558-096-8

[Jan-10] S. Jana, S.K. Kasera, „On Fast and Accurate Detection Of Unauthorized Wireless Access Points Using Clock-Skews”, Mobile Computing, IEEE Transactions on (Volume 9, Issue 3), 2010, ISSN 1536-1233

[Kal-12] D. Kalbande, C. Shah, A. Nigam, P. Kothawade, „Integrating ERP to accelerate business process agility: A case study and critical research review in Indian pharmaceutical industry", ICCICT 2012, ISBN 978-1-4577-2077-2

[Kuf-13] L. Kufel, "Security Event Monitoring in a Distributed Systems Environment", IEEE Security and Privacy, Volume 11, Issue 1

[Kum-12] A. Kumar, S.K. Kaushik, R. Sharma, P. Raj, "Simulators for Wireless Networks: A Comparative Study", ICCS 2012, ISBN: 978-1-4673-2647-6

[Lan-12] F. Lanze, A. Panchenko, B. Braatz, A. Zinnen, „Clock Skew Based Remote Device Fingerprinting Demystified", GLOBECOM 2012, ISBN: 978-1-4673-0920-2

[Lee-90] C.C. Lee „Fuzzy Logic in Control Systems: Fuzzy Logic Controler.I", IEEE Transactions on Systems, Man and Cybernetics, Volume 20, Issue 2, 1990, ISSN: 0018-9472

[Liq-06] Q. Li, W. Trappe, „Relationship-based Detection of Spoofing-related Anomalous Traffic in Ad Hoc Networks", IEEE Secon 2006 Proceedings, 1-4244-0626-9

[Loh-08] D.C.C Loh, C. Y. Cho, C. P. Tan, R. S. Lee, „Identifying Unique Devices through Wireless Fingerprinting, ACM, WiSec 2008, 978-1-59593-814-5

[Mam-77] E.H. Mamdani, „Application of Fuzzy Logic to Approximate Reasoning Using Linguistic Synthesis", IEEE Transactions on Computers, Volume C-26, Issue 12, 1977, ISSN: 0018-9340

[Mar-08] A. Martinez, U. Zurutuza, R. Uribeetxeberria, M. Fernandez, J. Iizarraga, A. Serna, I. Velez, „Beacon Frame Spoofing Attack Detection in IEEE 802.11 Networks", ARES 2008, ISBN 978-0-7695-3102-1

[MatW] MathWorks, Fuzzy Logic Toolbox http://www.mathworks.com/products/fuzzy-logic/

[Mel-09] F. Melakessou, T. Engel, "Network Traffic Simulator 2.0: Simulating the internet traffic", OSSC 2009, ISBN: 978-1-4244-4452-6

[Mie-10] A. Miede, N. Nedyalkov, C. Gottron, A. Konig, N. Repp, R. Steinmetz, "A Generic Metamodel for IT Security Attack Modeling for Distributed Systems", ARES 2010, ISBN: 978-1-4244-5879-0

[Mil-12] V. Miletic, B. Mikac, M. Dzanko, "Modelling optical network components: A network simulator-based approach", BIHTEL 2012, ISBN: 978-1-4673-4875-1

[Moh-10] H.S. Mohan, A.R. Reddy, „An Effective Defense against Distributed Denial of Service in Grid", ICIIC 2010, ISBN 978-1-4244-7963-4

[Nag-10] V. Nagarajan, V. Arasan, D. Huang, „Using Power Hopping to Counter MAC Spoof Attacks in WLAN", IEEE CCNC 2010 Proceedings, 978-1-4244-5176-0

[Nak-12] K. Nakajima, T. Hieda, I. Taniguchi, H. Tomiyama, H. Takada, "A Fast Network-on-Chip Simulator with QEMU and SystemC", ICNC 2012, ISBN: 978-1-4673-4624-5

[Nao-10] L. Nagowah, M.H.Domun, M. Umar IbnWaliyullah, "The network simulator of tomorrow - JNS", ICCSIT 2010, ISBN: 978-1-4244-5537-9

[Nat-12] S. Natarajan, T. Wolf, „Security Issues in Network Virtualization for the Future Internet", ICNC 2012, ISBN 978-1-4673-0008-7

[NetP] Hewlet Packard Co., Netperf: A Network Performance Benchmark", http://www.netperf.org

[Nis-05] R. Nishi, H. Morioka, K. Sakurai, „Trends and Issues for Security of Home-Nework Based on Power Line Communication", AINA 2005, ISBN 0-7695-2249-1

[Nis-09] T. Nishioka, Y. Sakumoto, H. Osaki, M. Imase, "Design and Implementation of Flow-Level Simulator for a Network with Heterogeneous Flows", SAINT-09, ISBN: 978-1-4244-4776-3

[Oht-11] M. Ohta, Y. Kanda, K. Fukuda, T. Sugawara, „Analysis of Spoofed IP Traffic Using Time-to-Live and Identification Fields in IP Headers”, WANA 2011, ISBN 978-1-61284-829-7

[Pin-13] A.K. Pinnaka, D. Tharashasank, V.S.K. Reddy, "Cost performance analysis of intrusion detection system in mobile wireless ad-hoc network", IACC 2013, ISBN 978-1-4673-4527-9

[Pop-11] F. Pop, A. Arcalianu, C. Dobre, V. Cristea, "Enhanced security for monitoring services in large scale distributed systems", ICCP 2011, ISBN: 978-1-4577-1479-5

[Pua-11] Somnuk Puangpronpitag and Atthapol Suwannasa, "A Design of Egress NAC using an Authentication Visa Checking mechanism to Protect against MAC Address Spoofing Attacks”, ECTI 2011, ISBN: 978-1-4577-0425-3

[Raz-13] A. Razzaq, A. Hur, H. Ahmad, M. Masood, "Cyber security: Threats, reasons, challenges, methodologies and state of the art solutions for industrial applications”, ISADS 2013, 978-1-4673-5069-3

[Reh-12] A. Rehman, S.M. Bilal, M. Othman, "A performance comparison of open source network simulators for wireless networks", ICCSCE 2012, ISBN: 978-1-4673-3142-5

[Rit-08] N. Rittammanart, W. Wongyued, M.N. Dailey, „ERP application development frameworks: Case study and evaluation”, ECTI-CON 2008, ISBN 978-1-4244-2101-5

[Sas-09] E.C. Sasu, O. Proştean, V. Groza, „MSHOSIM – A Tool for Mobile Station Handover Simulation”, SACI 2009, ISBN: 978-1-4244-4478-6

[Sas-10a] E.C. Sasu, O. Proştean, „Using Constant Traffic to Specific IP Destinations for Detecting Spoofed MAC Addresses in Local Area Networks”, ICCC-CONTI 2010, ISBN: 978-1-4244-7431-8

[Sas-10b] E.C. Sasu, O. Proştean, Voicu Groza, „Proving the Efficiency of DTF Method in a Local Area Network”, ICCC-CONTI 2010, ISBN: 978-1-4244-7431-8

[Sas-12a] E.C. Sasu, O. Proştean, „Testing DTF Method for Applicability in a Real Environment”, SACI 2012, ISBN: 978-1-4673-1011-6

[Sas-12b] E.C. Sasu, O. Proştean, „Network Simulation for MAC Spoofing Detection, using DTF Method”, SACI 2012, ISBN: 978-1-4673-1011-6

[Sas-12c] E.C. Sasu, O. Proştean, „Using Fuzzy Logic to Determine If the Traffic Is Constant or Not in DTF Method”, SOFA 2012, ISBN: 978-3-642-33940-0

[Sas-12d] E.C. Sasu, „Mathematical Model and Formalization for DTF Method”, SOFA 2012, ISBN: 978-3-642-33940-0

[Sar-13] V. Saravanan, A. Neeraja, „Security Issues in Computer Networks and Stenography”, ISCO 2013, ISBN 978-1-4673-4359-6

[Sha-12] S.K. Sharma, P. Pandey, S. Trwari, M.S. Sisodia, "An improved network intrusion detection technique based on k-means clustering via Naïve bayes classification", ICAESM 2012, ISBN 978-1-4673-0213-5

[She-08] Y. Sheng, K. Tan, G. Chen, D. Kotz, A. Campbell, „Detecting 802.11 MAC Layer Spoofing Using Received Signal Strength”, 978-1-4244-2026-1, IEEE 2008

[Som-10] R. Sommer, V. Paxon, "Outside the Closed World: On Using Machine Learning for Network Intrusion Detection", IEEE Symposium on Security and Privacy, 2010, ISBN 978-1-4244-6894-2

[Tan-03] Andrew Tanenbaum, „Computer Networks, Fourth Edition”, Prentice Hall, 2003, ISBN 0-13-066102-3

[Vuc-11] M. Vucicevic, A. Miljkovic, I. Papp, D. Samardzija, "Event driven wireless network simulator", TELFOR 2011, ISBN: 978-1-4577-1499-3

[Wan-07] H. Wang, C. Jin, K. Shin, „Defense Against Spoofed IP Traffic Using Hop-Count Filtering”, IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 15, NO. 1, FEBRUARY 2007

[Wan-08] S.Y. Wang, C.C. Lin, "NCTUns 5.0: A Network Simulator for IEEE 802.11(p) and 1609 Wireless Vehicular Network Researches", VTC 2008, ISBN: 978-1-4244-1721-6

[Wax-11] M.C. Waxman, "Cyber-Attacks and the Use of Force: Back to the Future of Article 2(4)", ale Journal of International Law, Vol. 36, 2011

[Woo-13] S. Woo, J. On, M. Lee, “Behaviour Ontology: A Framework to Detect Attack Patterns for Security”, WAINA 2013, 978-1-4673-6239-9

[Wri-03] Joshua Wright, „Detecting Wireless LAN MAC Address Spoofing”, GCIH, CCNA, 2003

[Xia-10] G. Xiaoqing, G. Hebin, C. Luyi, "Network intrusion detection method based on Agent and SVM", ICIME 2010, ISBN 978-1-4244-5263-7

[Xin-11] H. Xinyi, X. Yang, A. Chonka, Z. Jianying, R.H. Deng, "A Generic Framework for Three-Factor Authentication: Preserving Security and Privacy in Distributed Systems", IEEE Transactions on Parallel and Distributed Systems, Volume 22, Issue 8, ISSN: 1045-9219

[Yan-11] Q. Yanzhen, Y. Weiwen, "Improving the security of the distributed enterprise data warehouse system", ISI 2011, ISBN: 978-1-4577-0082-8

[Zad-62] L.A. Zadeh, „General System Theory”, IRE Transactions on Education, Volume 5, Issue 2, 1962, ISSN 0893-7141

[Zad-94] L.A. Zadeh, „The role of Fuzzylogic in modeling, identification and control”, Modeling, Identification and Control, Volume 15, Issue 3, 1994

[Zad-96] L.A. Zadeh, „Fuzzy Logic = computing with words”, IEEE Transactions on Fuzzy Systems, Volume 4, Issue 2, 1996, ISSN 1063-6706

[Zad-05] L.A. Zadeh, „The concept of a generalized constraint - a bridge from natural languages to mathematics”, IEEE International Workshop on Intelligent Signal Processing, 2005, ISBN 0-7803-9030-X

[Zad-07] L.A. Zadeh, „A New Frontier in Computation-Computation with Information Described in Natural Language”, IEEE International Workshop on Intelligent Signal Processing, 2007, ISBN 978-1-4244-0829-0

[Zha-09] X. Zhaoyong, „The Research of ERP Development Based on Network Era”, IFITA 2009, ISBN 978-0-7695-3600-2

[Zha-12] X. Zhao, "Research on the network intrusion detection method introduced with the view of quantum optimization", FNCES 2012, ISBN 978-1-4673-5033-4

Anexe.

A1. Rezultate obținute pe parcursul stagiului doctoral

Lucrări publicate ca prim autor în cadrul unor conferințe ISI

- Emanuel Ciprian Sasu, Octavian Proştean, „Mathematical Model and Formalization for DTF Method", 5th International Workshop on Soft Computing Applications, August 22-24, 2012, Szeged – Hungary, pg 391-402, ISBN 978-3-642-33940-0
- Emanuel Ciprian Sasu, Octavian Proştean, „Using Fuzzy Logic to Determine If the Traffic Is Constant or Not in DTF Method", 5th International Workshop on Soft Computing Applications, August 22-24, 2012, Szeged – Hungary, pg 117-128, ISBN 978-3-642-33940-0
- Emanuel Ciprian Sasu, Octavian Proştean, Voicu Groza, „MSHOSIM – A Tool for Mobile Station Handover Simulation", 5th International Symposium on Applied Computational Intelligence and Informatics, May 28-29, 2009, Timisoara – Romania, pg 363-367, ISBN 978-1-4244-4478-6

Lucrări publicate ca prim autor în cadrul unor conferințe BDI

- Emanuel Ciprian Sasu, Octavian Proştean, „Network Simulation for MAC Spoofing Detection, using DTF Method", IEEE 7th International Symposium on Applied Computational Intelligence and Informatics, May 24-26, 2012, Timisoara – Romania, pg 291-296, ISBN 978-1-4673-1011-6
- Emanuel Ciprian Sasu, Octavian Proştean, „Testing DTF Method for Applicability in a Real Environment", IEEE 7th International Symposium on Applied Computational Intelligence and Informatics, May 24-26, 2012, Timisoara – Romania, pg 285-289, ISBN 978-1-4673-1011-6
- Emanuel Ciprian Sasu, Octavian Proştean, Voicu Groza, „Proving the Efficiency of DTF Method in a Local Area Network", IEEE International Joint Conferences on Computational Cybernetics and Technical Informatics, May 27-29, 2010,Timisoara – Romania, pg 683-687, ISBN 978-1-4244-7431-8
- Emanuel Ciprian Sasu, Octavian Proştean, „Using Constant Traffic to Specific IP Destinations for Detecting Spoofed MAC Addresses in Local Area Networks", IEEE International Joint Conferences on Computational Cybernetics and Technical Informatics, May 27-29, 2010,Timisoara – Romania, pg 677-681, ISBN 978-1-4244-7431-8