



Σχολή Θετικών Επιστημών & Τεχνολογίας  
Πληροφορική

Προστασία και Ασφάλεια Συστημάτων Υπολογιστών (ΠΛΗ-35)  
4<sup>η</sup> Γραπτή Εργασία 2021-22

## **Διερεύνηση λειτουργίας botnet**

Καραγκούνης Λεωνίδας Χρήστος ΑΜ: 114163  
Τμήμα: ΗΛΕ-42, ΣΕΠ Μαυρίδης Ιωάννης

## Πίνακας Περιεχομένων

Πίνακας Εικόνων .....	ii
1. Εισαγωγή .....	1
2. Υπόβαθρο.....	2
2.1 Σύστημα Ονοματοδοσίας Τομέων .....	2
2.1.1 Λειτουργία .....	2
2.1.1 Θέματα ασφάλειας – κατάχρησης .....	4
2.2 Δίκτυα υπολογιστών ρομπότ.....	5
2.1.1 Λειτουργία .....	5
2.1.1 Τεχνικές γρήγορης ροής και ροής τομέα .....	6
3. Περιγραφή λύσεων .....	8
3.1 Ερώτημα Α.....	8
3.2 Ερώτημα Β .....	11
3.3 Ερώτημα Γ.....	12
3.4 Ερώτημα Δ .....	13
3.5 Ερώτημα Ε .....	13
4. Συμπεράσματα .....	14
Βιβλιογραφία .....	15
Παράρτημα.....	18
split.py.....	18
GE4.ipynb .....	18

## Πίνακας Εικόνων

Εικόνα 1. Δενδρική δομή DNS (Alakbarov et al., 2017).....	2
Εικόνα 2. Διαδικασία επίλυσης DNS (Dancs, 2020).....	3
Εικόνα 3. Ανάλυση πρόσφατα καταχωρημένων τομέων (Chen, Wang and Kwan, 2019).....	4
Εικόνα 4. Κεντρική (a) και αποκεντρωμένη (b) δομή botnet (Han and Im, 2011, p. 590) .....	5
Εικόνα 5. Λειτουργία DGA botnet (Casino et al., 2021).....	7
Εικόνα 6. Στατιστικά στοιχεία πρωτοκόλλου DNS .....	8
Εικόνα 7. Πακέτα με κωδικό απάντησης rcode:3 .....	9
Εικόνα 8. Καταληκτικά σημεία φίλτρου rcode:3 .....	9
Εικόνα 9. Φίλτρο ονομάτων τομέα κατάληξης .ru .....	10
Εικόνα 10. Καταληκτικά σημεία φίλτρου ονομάτων τομέα κατάληξης .ru .....	10
Εικόνα 11. Εντολή tshark για εξαγωγή ονομάτων τομέα κατάληξης .ru .....	11
Εικόνα 12. Αποτελέσματα αλγορίθμου ταύτισης .....	12

## 1. Εισαγωγή

Το Σύστημα Ονοματοδοσίας Τομέων (Domain Name System – DNS) είναι ένας από τους ακρογωνιαίους λίθους λειτουργίας του σύγχρονου διαδικτύου καταστώντας το πιο προσιτό στον απλό χρήστη αντιστοιχώντας διευθύνσεις πρωτοκόλλου διαδικτύου (IP) στα πιο ευανάγνωστα ονόματα τομέα (domain names). Το DNS για αρκετά χρόνια έχει γίνει αντικείμενο κατάχρησης ως μέρος διαφόρων κακόβουλων επιθέσεων όπως ηλεκτρονικό «ψάρεμα» (phishing), δίκτυα υπολογιστών ρομπότ (botnets), δηλητηρίαση προσωρινής μνήμης DNS (DNS cache poisoning – DNS spoofing) κ.λ.π.

Στην παρούσα εργασία θα ασχοληθούμε με ένα περιστατικό κατάχρησης DNS (DNS abuse) και συγκεκριμένα τη διερεύνηση λειτουργίας ενός μολυσμένου υπολογιστή ρομπότ (bot). Αρχικά αναφέρονται πληροφορίες σχετικά με τη λειτουργία, ασφάλεια και κατάχρηση του DNS και τα δίκτυα υπολογιστών ρομπότ και τις μεθόδους που χρησιμοποιούνται για τη λειτουργία τους. Έπειτα ακολουθεί ανάλυση και απαντήσεις στα ερωτήματα της εργασίας. Τέλος αναγράφονται πιθανά συμπεράσματα που εξήχθησαν κατά την εκπόνηση της εργασίας.

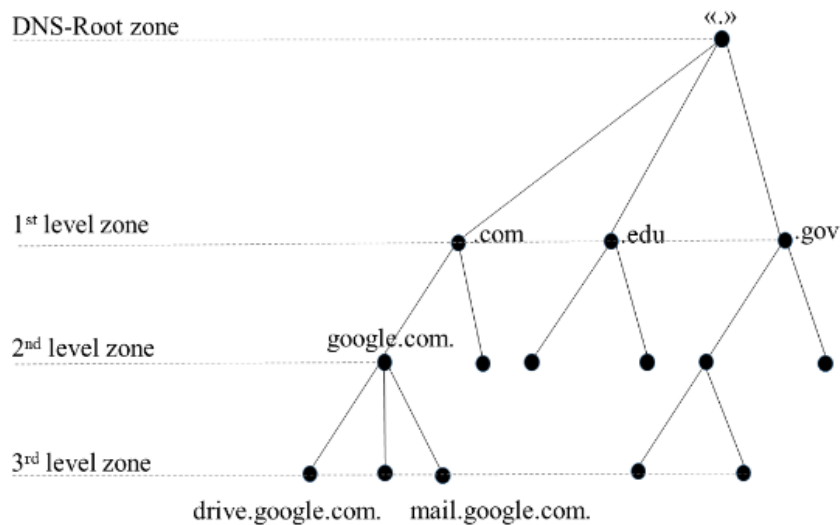
## 2. Υπόβαθρο

### 2.1 Σύστημα Ονοματοδοσίας Τομέων

Το διαδίκτυο, αλλά και κάθε είδος δικτύου, δουλεύει με ανάθεση διευθύνσεων πρωτοκόλλου διαδικτύου<sup>1</sup> (IP) στα καταληκτικά σημεία του (endpoints). Αν θέλουμε να συνδεθούμε σε κάποιο σημείο πρέπει να γνωρίζουμε την διεύθυνσή του. Η απομνημόνευση αριθμητικών διευθύνσεων όμως είναι αρκετά δύσκολη, τόσο εξαιτίας του πλήθους αλλά και του γεγονότος πιθανής μεταβολής τους. Έτσι προς διευκόλυνση των χρηστών χρησιμοποιείται το Σύστημα Ονοματοδοσίας Τομέων (Domain Name System), γνωστό και ως DNS, το οποίο πραγματοποιεί αντιστοίχιση μιας ή περισσοτέρων διευθύνσεων πρωτοκόλλου διαδικτύου με τα πιο ευανάγνωστα και εύκολα στη χρήση ονόματα τομέα (domain names).

#### 2.1.1 Λειτουργία

Το DNS είναι ένα ιεραρχικά δομημένο σύστημα παρόμοιο με τη δομή του συστήματος αρχείων Unix και απεικονίζεται ως ένα ανεστραμμένο δέντρο (Εικόνα 1).

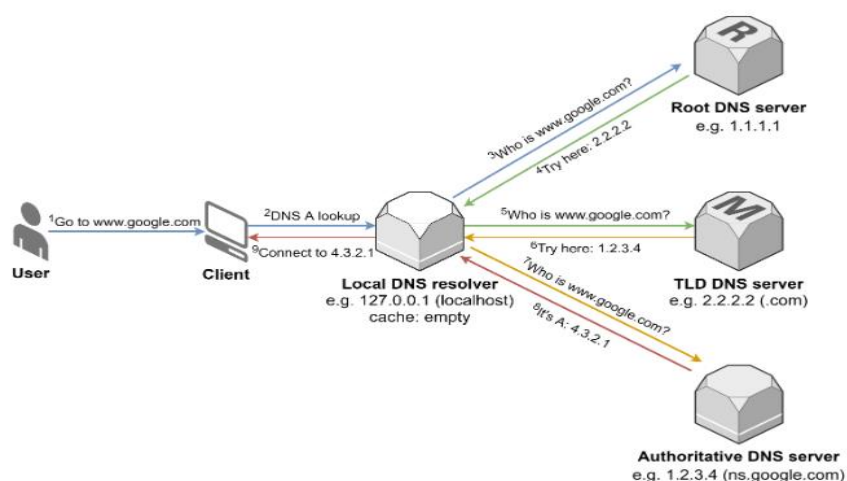


Εικόνα 1. Δενδρική δομή DNS (Alakbarov et al., 2017)

<sup>1</sup> Τοπικές ή παγκοσμίως μοναδικές

Κάθε τομέας, π.χ. “www.example.com.”, αποτελείται από πολλαπλές ετικέτες (labels) ονομάτων τομέων που χωρίζονται με μια τελεία “.”. Οι εξουσιοδοτημένοι εξυπηρετητές ονομάτων (Authoritative Name Servers) είναι η αρμόδια αρχή για μια συγκεκριμένη ζώνη του χώρου ονομάτων τομέα. Οι πληροφορίες οργανώνονται και διατηρούνται από τους διαχειριστές τομέων ενώ αποτελούν μια πλήρη βάση δεδομένων για ένα συγκεκριμένο υποδέντρο του χώρου των τομέων περιέχοντας τις εγγραφές πόρων (RRs) της ζώνης. Στην κορυφή του δέντρου βρίσκεται ο κόμβος ρίζα, ο οποίος παριστάνεται με μια τελεία “.” που προστίθεται στο τέλος του τομέα, μη ορατός στο χρήστη, για τον οποίο είναι εξουσιοδοτημένος ο Εξυπηρετητής Ονομάτων Ρίζας (Root Name Server) . Ακολουθούν οι Τομείς Ανωτάτου Επιπέδου<sup>2</sup> (Top-Level Domains - TLDs), οι Δεύτεροι Τομείς Ανωτάτου Επιπέδου (Second-Top-Level Domains - 2-TLDs) και στη συνέχεια οποιοσδήποτε άλλος αριθμός χαμηλότερων επιπέδων (Dancs, 2020).

Η επίλυση αιτήματος (query) ονόματος τομέα από τον χρήστη διεκπεραιώνεται από τον αναδρομικό επιλύτη DNS (recursive resolver). Αν ο επιλύτης είχε την απάντηση στην κρυφή του μνήμη το αίτημα επιλυεται άμεσα, αλλιώς αποστέλλει επαναληπτικά ερωτήματα που ξεκινούν από τον Εξυπηρετητή Ονομάτων Ρίζας ο οποίος αν δεν γνωρίζει την απάντηση, τον κατευθύνει στον εξουσιοδοτημένο για την επόμενη ετικέτα Τομέα Ανωτάτου Επιπέδου (Εικόνα 2).



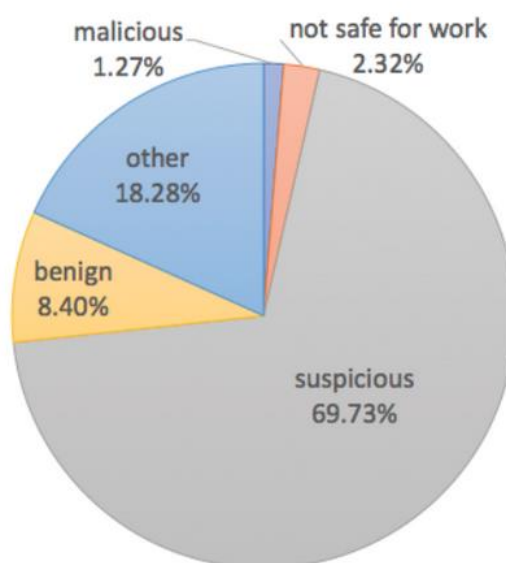
Εικόνα 2. Διαδικασία επίλυσης DNS (Dancs, 2020)

<sup>2</sup> Οι οποίοι χωρίζονται στους γενικούς (gTLDs) και σε αυτούς των κωδικών χωρών (ccTLDs) και είναι στο πλήθος 1591 κατά την στιγμή της συγγραφής της παρούσας εργασίας

Η διαδικασία επαναλαμβάνεται έως ότου ένας εξυπηρετητής ονομάτων γνωρίζει την απάντηση στο ερώτημα και επιστραφεί η ζητούμενη διεύθυνση. Σε περίπτωση που ο επιλύτης δεν κατάφερε να ανακτήσει τις πληροφορίες που απαιτούνται, τότε επιστρέφει μια απάντηση σφάλματος, όπως το μη υπαρκτό όνομα τομέα (NXDomain).

### 2.1.1 Θέματα ασφάλειας – κατάχρησης

Το DNS είναι από τα πιο αναγκαία πρωτόκολλα για την σωστή λειτουργία του διαδικτύου. Παρά την σημαντικότητά του όμως, το πρωτόκολλο παρουσιάζει αδυναμίες ασφάλειας στον σχεδιασμό του με αποτέλεσμα φορείς απειλών (threat actors) να μπορούν εύκολα να το καταχραστούν. Σύμφωνα με δημοσίευμα της Ευρωπαϊκής Ένωσης που αφορά τη μελέτη κατάχρησης του DNS, κατάχρηση αποτελεί κάθε δραστηριότητα που κάνει χρήση ονομάτων τομέα ή το πρωτόκολλο DNS για να διεξάγει επιβλαβείς ή παράνομες δραστηριότητες (Directorate-General for Communications Networks, Content and Technology et al., 2022). Εκτός από τα προβλήματα ασφάλειας του πρωτοκόλλου, στα φαινόμενα κατάχρησης συνεισφέρουν και οι οργανισμοί καταχώρησης (registrars) λόγω των ιδιαίτερα χαλαρών κριτηρίων κατά την εγγραφή νέων τομέων. Σε ανάλυση που πραγματοποιήθηκε, 70% των πρόσφατα καταχωρημένων τομέων που εξετάστηκαν είναι "κακόβουλα", "ύποπτα" ή "μη ασφαλή για εργασία" (Chen, Wang and Kwan, 2019) (Εικόνα 3).



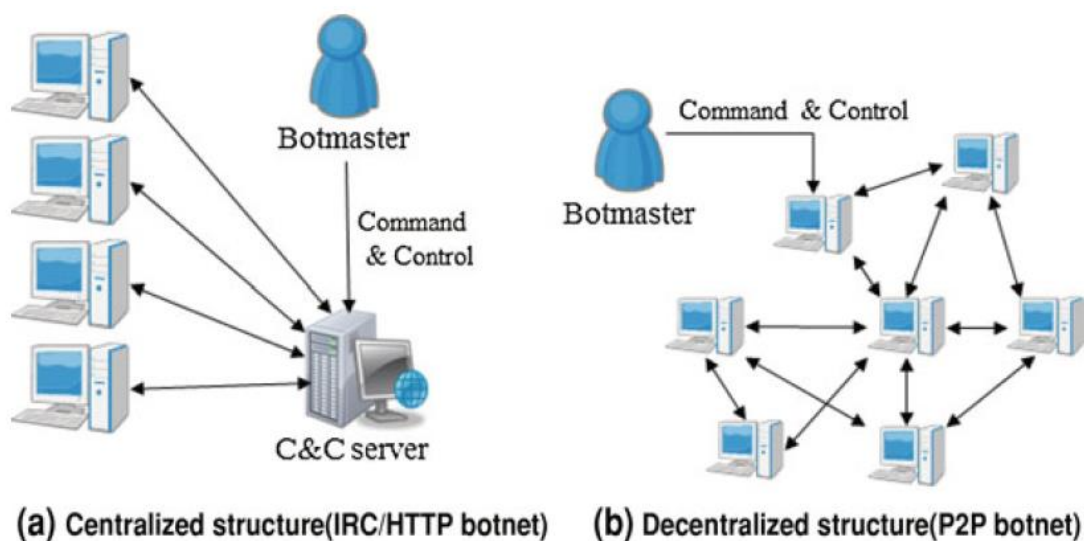
Εικόνα 3. Ανάλυση πρόσφατα καταχωρημένων τομέων (Chen, Wang and Kwan, 2019)

Ενδεικτικά, κακόβουλες ενέργειες κατάχρησης του πρωτοκόλλου DNS περιλαμβάνουν επιθέσεις επανασύνδεσης (rebinding attacks), επιθέσεις σήραγγας (tunneling attacks), δηλητηρίαση της κρυφής μνήμης DNS (DNS cache poisoning) και επιθέσεις DNSSEC. Αντίθετα επιθέσεις με χρήση πρόσφατα καταχωρημένων τομέων περιλαμβάνουν μεταξύ άλλων, συστήματα διανομής κακόβουλου λογισμικού, ηλεκτρονικό «ψάρεμα», ιστοσελίδες απάτης για κλοπή προσωπικών στοιχείων και διαπιστευτηρίων και καταχώρηση τομέων για χρήση ως εξυπηρετητή εντολών και ελέγχου (command and control server) ενός δικτύου υπολογιστών ρομπότ.

## 2.2 Δίκτυα υπολογιστών ρομπότ

### 2.1.1 Λειτουργία

Ένα δίκτυο υπολογιστών ρομπότ αποτελείται από ένα σύνολο μολυσμένων συσκευών ρομπότ που συνδέονται σε εξυπηρετητές εντολών και ελέγχου με σκοπό την εκτέλεση συντονισμένων παράνομων ενεργειών που κατευθύνονται από τον κάτοχο του δικτύου (botmaster) (Εικόνα 4).



Εικόνα 4. Κεντρική (a) και αποκεντρωμένη (b) δομή botnet (Han and Im, 2011, p. 590)

Συνήθως η μόλυνση επιτυγχάνεται χρησιμοποιώντας τεχνικές κοινωνικής μηχανικής (social engineering) με το ανοιγμα επισυναπτόμενου σε μήνυμα ηλεκτρονικού ταχυδρομείου κακόβουλου λογισμικού ή συνδέσμου ιστοσελίδας. Η χρήση τέτοιων δικτύων ρομπότ ποικίλλει



ανάλογα με τους στόχους του κατόχου αλλά και την οικογένεια του κακόβουλου λογισμικού<sup>3</sup> και περιλαμβάνει μεταξύ άλλων την αυτοματοποίηση επιθέσεων μεγάλης κλίμακας, συμπεριλαμβανομένης της κλοπής δεδομένων, της διάδοσης κακόβουλου λογισμικού, της δημιουργίας μηνυμάτων spam και της δημιουργίας κακόβουλης κίνησης για κατανεμημένες επιθέσεις άρνησης παροχής υπηρεσιών (DDoS).

### 2.1.1 Τεχνικές γρήγορης ροής και ροής τομέα

Είναι ζωτικής σημασίας για το botnet να αποκρύπτει το κανάλι επικοινωνίας του, τις ταυτότητες του εξυπηρετητή εντολών και ελέγχου και του botmaster, ώστε να διατηρείται κάποιος βαθμός μη διασυνδεσιμότητας και να αποφευχθούν τυχόν προσπάθειες ταυτοποίησης και κατάργησής του. Ως εκ τούτου, οι επιτιθέμενοι προσπαθούν να αποκρύψουν τους εξυπηρετητές εντολών και ελέγχου, αλλά χρειάζεται να διατηρηθεί ένας τρόπος ώστε τα bot να εντοπίζουν και να επανασυνδέονται σε αυτούς. Κατά ειρωνικό τρόπο, η ευελιξία και η παγκόσμια διαθεσιμότητα του DNS παρέχει στους botmasters έναν τρόπο να δημιουργήσουν μια ανθεκτική υποδομή για τους εξυπηρετητές εντολών και ελέγχου. Η γρήγορη ροή και η ροή τομέα είναι δύο τεχνικές που χρησιμοποιούνται ευρέως για την επίτευξη αυτού του στόχου.

Η γρήγορη ροή (fast-flux)<sup>4</sup> έχει ως στόχο την απόκρυψη εξυπηρετητών εντολών και ελέγχου χρησιμοποιώντας μια σειρά ταχέως μεταβαλλόμενων διευθύνσεων που συνδέονται με την επίλυση ενός αιτήματος ονόματος τομέα με την χρήση διευθυνσιοδότησης round-robin και χαμηλού TTL για μια εγγραφή πόρου DNS (MITRE, 2020). Αντίθετα στην ροή τομέα (domain flux) αντί της μεταβολής διευθύνσεων έχουμε μεταβολή των ονομάτων τομέα. Ένας αριθμός ονομάτων τομέων ή υποτομέων (subdomains) είναι κωδικοποιημένα στο εκτελέσιμο αρχείο του bot ή συχνότερα πλέον δημιουργούνται από έναν ενσωματωμένο στο bot αλγόριθμο που ονομάζεται αλγόριθμος παραγωγής τομέων (Domain Generation Algorithm - DGA). Χρησιμοποιώντας μια αρχική τιμή<sup>5</sup> (seed) δημιουργεί ένα πλήθος ψευδοτυχαίων ονομάτων τομέων<sup>6</sup> με τα οποία τα bot προσπαθούν να επικοινωνήσουν επαναληπτικά για να βρουν τον εξυπηρετητή εντολών και ελέγχου

---

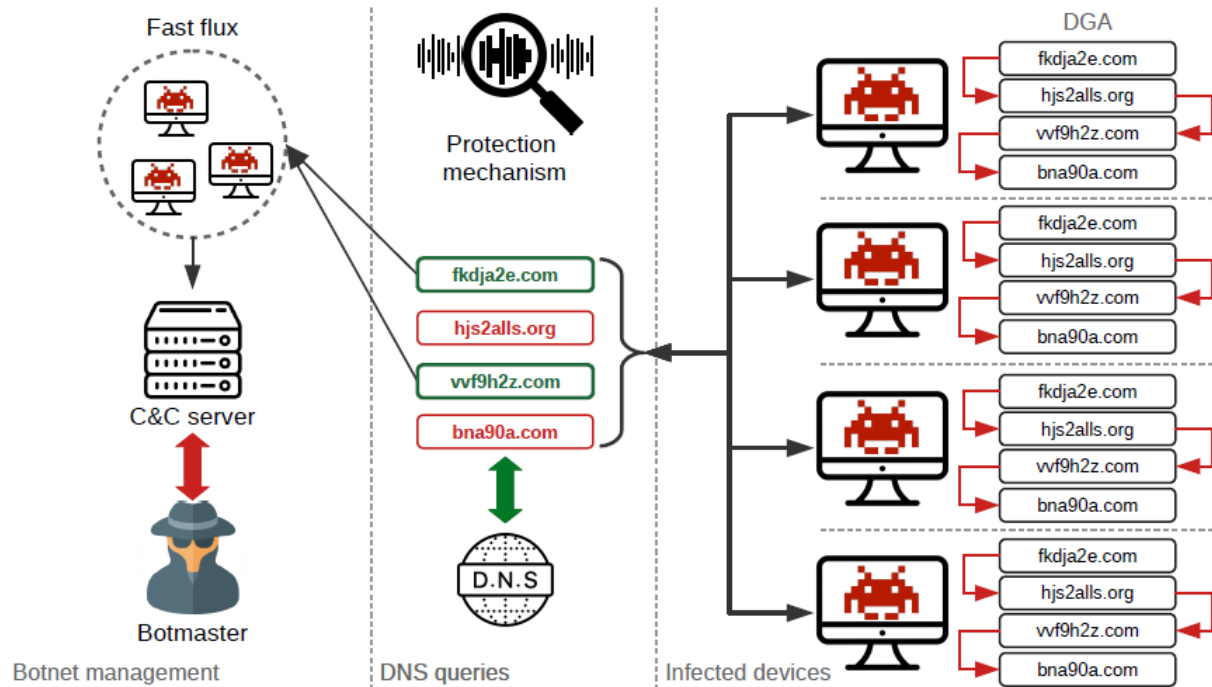
<sup>3</sup> Γνωστά παραδείγματα αποτελούν τα Mirai, Emotet, Zbot/ZeuS, Cutwail, Trickbot, Ramnit, Storm, Dridex/Cridex, ZeroAccess, Mariposa κ.λ.π.

<sup>4</sup> Η οποία χωρίζεται στις κατηγορίες μεθόδων μονής και διπλής ροής

<sup>5</sup> Συνήθως χρησιμοποιούνται ημερομηνίες αλλά και άλλες δημόσια γνωστές πληροφορίες (Plohmann et al., 2016)

<sup>6</sup> Διαφορετικό για κάθε οικογένεια DGA

χρησιμοποιώντας αιτήματα DNS (Εικόνα 5). Στη συνέχεια ο botmaster θα προχωρήσει στην καταχώριση ενός ή μερικών από αυτών και για σύντομο χρονικό διάστημα καθιστώντας έτσι εξαιρετικά δύσκολο να προσδιοριστούν και να αποκλειστούν τα εν λόγω ονόματα τομέα (Casino et al., 2021).

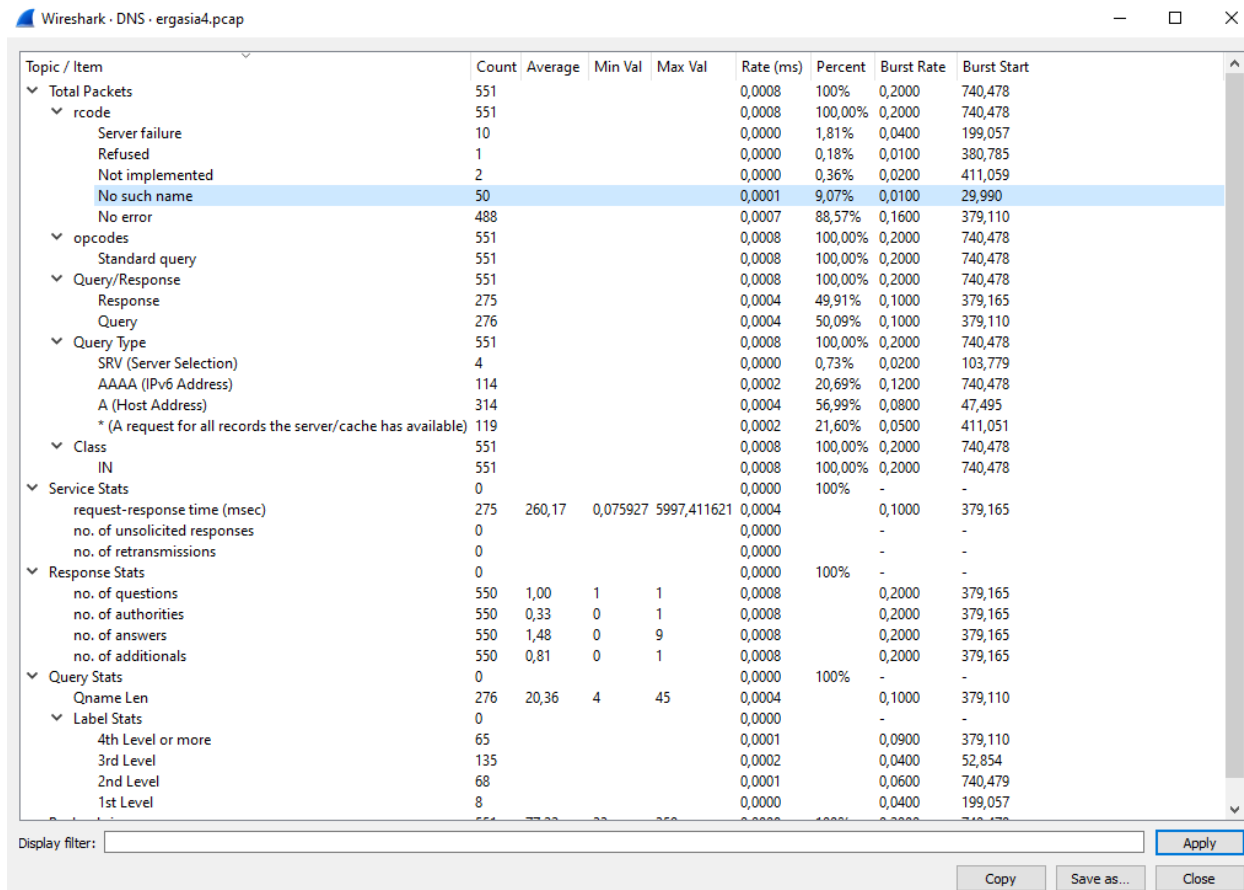


Εικόνα 5. Λειτουργία DGA botnet (Casino et al., 2021)

### 3. Περιγραφή λύσεων

#### 3.1 Ερώτημα Α

Ξεκινάμε την διαδικασία εύρεσης του μολυσμένου υπολογιστή εξετάζοντας στατιστικά στοιχεία για το πρωτόκολλο DNS μέσω του αναδυόμενου μενού Statistics → DNS ( Εικόνα 6).



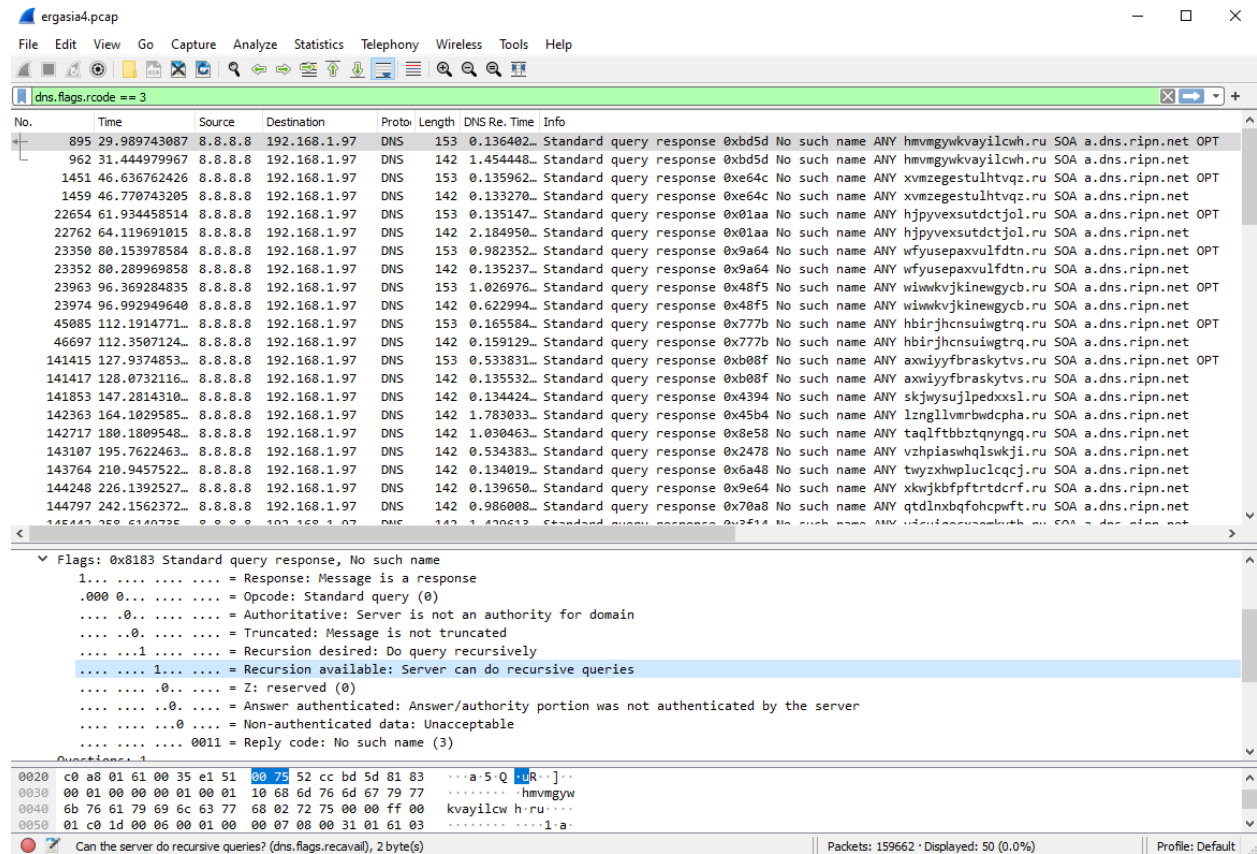
Topic / Item	Count	Average	Min Val	Max Val	Rate (ms)	Percent	Burst Rate	Burst Start
Total Packets	551				0,0008	100%	0,2000	740,478
rcode	551				0,0008	100,00%	0,2000	740,478
Server failure	10				0,0000	1,81%	0,0400	199,057
Refused	1				0,0000	0,18%	0,0100	380,785
Not implemented	2				0,0000	0,36%	0,0200	411,059
No such name	50				0,0001	9,07%	0,0100	29,990
No error	488				0,0007	88,57%	0,1600	379,110
opcodes	551				0,0008	100,00%	0,2000	740,478
Standard query	551				0,0008	100,00%	0,2000	740,478
Query/Response	551				0,0008	100,00%	0,2000	740,478
Response	275				0,0004	49,91%	0,1000	379,165
Query	276				0,0004	50,09%	0,1000	379,110
Query Type	551				0,0008	100,00%	0,2000	740,478
SRV (Server Selection)	4				0,0000	0,73%	0,0200	103,779
AAAA (IPv6 Address)	114				0,0002	20,69%	0,1200	740,478
A (Host Address)	314				0,0004	56,99%	0,0800	47,495
* (A request for all records the server/cache has available)	119				0,0002	21,60%	0,0500	411,051
Class	551				0,0008	100,00%	0,2000	740,478
IN	551				0,0008	100,00%	0,2000	740,478
Service Stats	0				0,0000	100%	-	-
request-response time (msec)	275	260,17	0,075927	5997,411621	0,0004	-	0,1000	379,165
no. of unsolicited responses	0				0,0000	-	-	-
no. of retransmissions	0				0,0000	-	-	-
Response Stats	0				0,0000	100%	-	-
no. of questions	550	1,00	1	1	0,0008	-	0,2000	379,165
no. of authorities	550	0,33	0	1	0,0008	-	0,2000	379,165
no. of answers	550	1,48	0	9	0,0008	-	0,2000	379,165
no. of additionals	550	0,81	0	1	0,0008	-	0,2000	379,165
Query Stats	0				0,0000	100%	-	-
Qname Len	276	20,36	4	45	0,0004	-	0,1000	379,110
Label Stats	0				0,0000	-	-	-
4th Level or more	65				0,0001	-	0,0900	379,110
3rd Level	135				0,0002	-	0,0400	52,854
2nd Level	68				0,0001	-	0,0600	740,479
1st Level	8				0,0000	-	0,0400	199,057

Εικόνα 6. Στατιστικά στοιχεία πρωτοκόλλου DNS

Παρατηρούμε μεγάλο αριθμό (50) αιτημάτων κατηγορίας «No such name» με rcode 3 και μήνυμα επιστροφής που υποδηλώνει πως το συγκεκριμένο όνομα τομέα δεν υπάρχει<sup>7</sup>. Εφαρμόζοντας το φίλτρο `dns.flags.rcode==3` έχουμε τα πακέτα που φαίνονται στην Εικόνα 7. Άμεσα

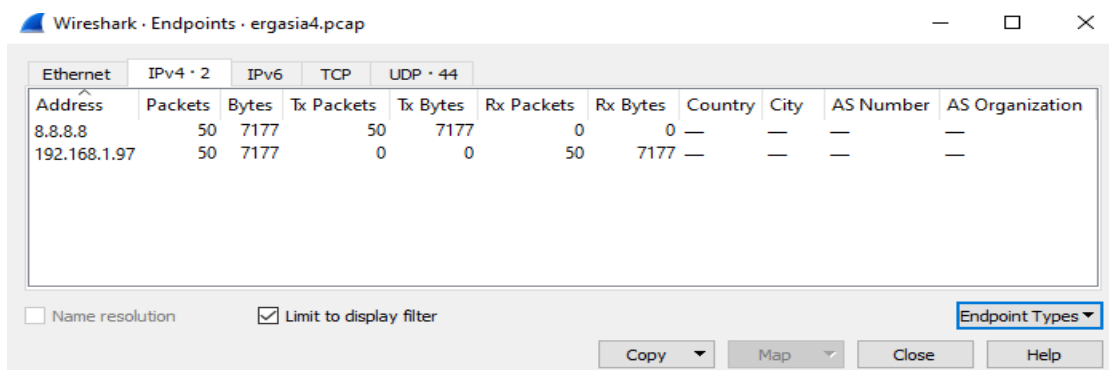
<sup>7</sup> Υποψιαζόμαστε την χρήση αλγορίθμου DGA

παρατηρούμε τα ονόματα τομέα κατάληξης .ru, που φαίνεται πως έχουν παραχθεί με αλγόριθμο DGA, με διεύθυνση παραλήπτη την διεύθυνση τοπικού δικτύου 192.168.1.97.



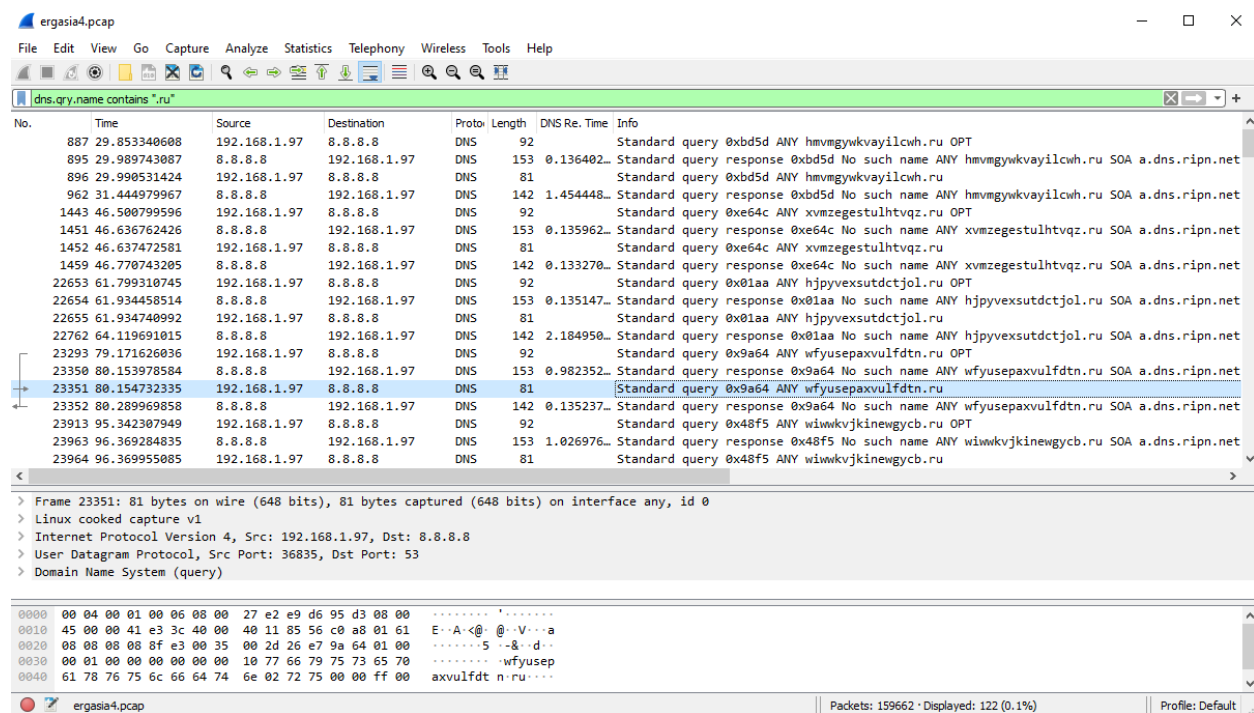
Εικόνα 7. Πακέτα με κωδικό απάντησης rcode:3

Για να επιβεβαιώσουμε ότι όντως η συγκεκριμένη διεύθυνση ανήκει στον ζητούμενο μολυσμένο υπολογιστή, εξετάζουμε στατιστικά στοιχεία για τα καταληκτικά σημεία από το αναδυόμενο μενού Statistics → Endpoints με ενεργή την επιλογή για περιορισμό εμφάνισης μόνο στα δεδομένα του φίλτρου (Εικόνα 8).



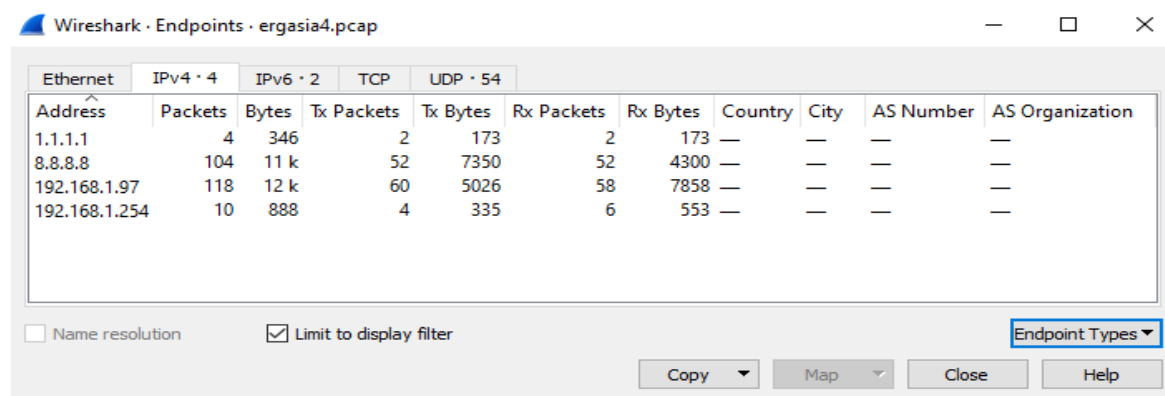
Εικόνα 8. Καταληκτικά σημεία φίλτρου rcode:3

Εκτός της διεύθυνσης 8.8.8.8<sup>8</sup>, η μόνη διεύθυνση είναι η 192.168.1.97. Περιορίζουμε τα αποτελέσματα στα αλγοριθμικά παραγόμενα ονόματα τομέα (AGD) κατάληξης .ru, εφαρμόζοντας το φίλτρο `dns.qry.name contains ".ru"` (Εικόνα 9).



Εικόνα 9. Φίλτρο ονομάτων τομέα κατάληξης .ru

Ακολουθώντας την ίδια διαδικασία για τα καταληκτικά σημεία έχουμε τα δεδομένα όπως φαίνονται στην Εικόνα 10.



Εικόνα 10. Καταληκτικά σημεία φίλτρου ονομάτων τομέα κατάληξης .ru

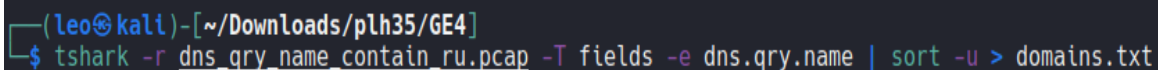
<sup>8</sup> Δημόσιος επιλύτης DNS της Google

Η μόνη διεύθυνση υπολογιστή τοπικού δικτύου είναι η 192.168.1.97<sup>9</sup>. Επομένως χωρίς αμφιβολία μπορούμε να αποφανθούμε πως η ζητούμενη διεύθυνση τοπικού δικτύου μολυσμένου υπολογιστή είναι η 192.168.1.97. Σαν επιπλέον βήμα για τα παρακάτω ερωτήματα κάνουμε εξαγωγή των πακέτων με φίλτρο `dns.qry.name contains ".ru"` και τα αποθηκεύουμε στο αρχείο με όνομα `dns_qry_name_contain_ru.pcap`.

## 3.2 Ερώτημα Β

Είναι προφανές και από τα παραπάνω πως το botnet που έχει συνδεθεί ο μολυσμένος υπολογιστής χρησιμοποιεί αλγόριθμο DGA για την παραγωγή ονομάτων τομέα. Το εργαστήριο έρευνας ασφάλειας 360Netlab διαθέτει μια λίστα με γνωστά ονόματα τομέων DGA καθώς και τις οικογένειες που ανήκουν (360Netlab, 2022). Δεδομένου ότι το μήκος της λίστας είναι πολύ μεγάλο για να ελέγξουμε χειροκίνητα για ταυτίσεις δουλεύουμε ως εξής.

Αρχικά αποθηκεύουμε την λίστα του 360Netlab στο txt αρχείο `netlab360_list.txt`. Εκτελώντας το πρόγραμμα [split.py](#) του Παραρτήματος αποθηκεύουμε μόνο τα ονόματα τομέα και την οικογένεια που ανήκουν στο αρχείο `dga.txt`. Έπειτα κάνοντας χρήση εντολής `tshark` (Εικόνα 11) αποθηκεύουμε τα ονόματα τομέα του αρχείου `dns_qry_name_contain_ru.pcap`, που δημιουργήσαμε στο προηγούμενο ερώτημα, στο αρχείο `domains.txt`.



```
(leo@kali)-[~/Downloads/plh35/GE4]
$ tshark -r dns_qry_name_contain_ru.pcap -T fields -e dns.qry.name | sort -u > domains.txt
```

Εικόνα 11. Εντολή `tshark` για εξαγωγή ονομάτων τομέα κατάληξης `.ru`

Έπειτα μετατρέπουμε τα δύο txt αρχεία σε csv και εκτελούμε τον κώδικα του Jupyter Notebook [GE4.ipynb](#) του Παραρτήματος. Στο τέλος έχουμε τα αποτελέσματα για τις ταυτίσεις των ονομάτων τομέα όπως φαίνεται και στην Εικόνα 12.

---

<sup>9</sup> Η 8.8.8.8 είναι ο δημόσιος επιλύτης DNS της Google, η 1.1.1.1 είναι ο δημόσιος επιλύτης DNS της Cloudflare και η 192.168.1.254 είναι διεύθυνση προεπιλεγμένης πύλης δρομολογητή

```

Results
dns_data.head()
[26]
...
  Family      Domain  Match
0  feodo  anidgwelnidmzueo.ru    1
1  feodo  aopltfxjzspylfh.ru    1
2  feodo  auvqjghelyqwtfsu.ru    1
3  feodo  axwiyyfbraskytvs.ru    1
4  feodo  aygrpumrlmymcwkjh.ru    1

dns_data['Family'].value_counts()
[27]
... feodo    44
     Name: Family, dtype: int64

dns_data['Match'].value_counts()
[28]
... 1    44
     Name: Match, dtype: int64

We can see that 44 out of 44 domains from our data was found in the Netlab 360 Data List and every one of them is classified as feodo dga family

```

Εικόνα 12. Αποτελέσματα αλγορίθμου ταύτισης

Παρατηρούμε πως και τα 44 ονόματα τομέα κατάληξης `.ru` ταυτίστηκαν με αντίστοιχα γνωστά της λίστας 360Netlab και ταυτοποιήθηκαν ως ονόματα τομέων DGA της οικογένειας `feodo` που είναι και το ζητούμενο του ερωτήματος.

### 3.3 Ερώτημα Γ

Αρχικά από το άρθρο των Sood and Zeadally (2016) αντιλαμβανόμαστε ότι πρόκειται για Binary-Based DGA αλφαβητικής διάταξης. Κάνοντας μια σύντομη αναζήτηση στο διαδίκτυο είτε με τα ονόματα τομέα είτε αναζητώντας στοιχεία για μόλυνση από κακόβουλο λογισμικό της οικογένειας `feodo`, συμπεράνουμε πως υπάρχουν διάφοροι τρόποι για την αρχική μόλυνση του υπολογιστή. Κοινός παράγοντας σε όλους είναι τεχνικές κοινωνικής μηχανικής έτσι ώστε να πειστεί το θύμα να ανοίξει κάποιο αρχείο ή σύνδεσμο ιστοσελίδας, τα οποία είναι επισυναπτόμενα σε μήνυμα ηλεκτρονικού ταχυδρομείου. Ο κακόβουλος ιστότοπος είτε μεταφέρει ένα εργαλείο εκμετάλλευσης είτε φορτώνει ένα πακέτο εκμετάλλευσης περιηγητή (BEP) που εκμεταλλεύεται ευπάθειες όπως υπερχείλιση προσωρινής μνήμης (buffer overflow) και αλλοιώσεις μνήμης (memory corruption) στο σύστημα του στόχου για να εγκαταστήσει τον εαυτό του. Από την άλλη πλευρά ένα κακόβουλο λογισμικό, όπως το Troj/MDrop-DZF (Sophos, 2012), ή ένα κακόβουλο

έγγραφο του Office που περιέχει μια συγκεκαλυμμένη μακροεντολή (macro) που εκτελεί μια δέσμη ενεργειών PowerShell, αποσυμπιέζεται απευθείας στη μνήμη και εισάγει κακόβουλα αρχεία στο σύστημα τα οποία είτε εκτελούνται κατευθείαν είτε μέσω νέων διεργασιών που δημιουργούνται (process injection) (Roccia, 2017). Τέλος ο αλγόριθμος DGA αντλεί την αρχική τιμή, ξεκινά τη δημιουργία ονομάτων τομέα και υποβάλλει αιτήματα DNS για να αναζητήσει τον καταχωρημένο τομέα μεταξύ αυτών. Μόλις επιλυθεί το αίτημα και συνδεθεί με τη διεύθυνση IP του καταχωρημένου τομέα, το κακόβουλο λογισμικό αρχίζει να επικοινωνεί με τον εξυπηρετητή εντολών και ελέγχου του δικτύου υπολογιστών ρομπότ.

### 3.4 Ερώτημα Δ

Τα κακόβουλα λογισμικά της οικογένειας feodo έχουν ως στόχο την υποκλοπή ευαίσθητων προσωπικών δεδομένων και διαπιστευτηρίων<sup>10</sup> με καταγραφή των πλήκτρων που πληκτρολογούνται στον μολυσμένο υπολογιστή, υποκλοπή των πεδίων ονόματος χρήστη και κωδικού πρόσβασης που πληκτρολογούνται σε φόρμες σύνδεσης ή ακόμα και παρουσίαση στα θύματα ψεύτικων σελίδων σύνδεσης σε ηλεκτρονικές τραπεζικές υπηρεσίες έτσι ώστε να αποκτήσουν πρόσβαση στον τραπεζικό λογαριασμό τους. Καθώς έχουμε επιβεβαιώσει ότι ο υπολογιστής είναι μολυσμένος και με δεδομένο ότι έχει επιτευχθεί η σύνδεση με τον εξυπηρετητή εντολών και ελέγχου του δικτύου υπολογιστών ρομπότ μέσω αιτημάτων DNS, τότε η εξαγωγή δεδομένων τέτοιων δεδομένων και όχι μόνο είναι εφικτή.

### 3.5 Ερώτημα Ε

Κάνοντας μια αναζήτηση στο διαδίκτυο σχετικά με τα ονόματα τομέων του pcap αρχείου βρίσκουμε το άρθρο του Jenkins (2012) της εταιρίας Spamhaus. Στο εν λόγω άρθρο παρουσιάζονται πειστήρια για την καταχώρηση ονομάτων τομέων για κακόβουλη χρήση από τον ρωσικό οργανισμό καταχώρησης NAUNET. Παρουσιάζεται λίστα ονομάτων τομέων για χρήση από δίκτυα ρομπότ της οικογένειας feodo και παρατηρούμε πως αρκετά από αυτά, συγκεκριμένα 12 από τα 71, συμπίπτουν με αυτά που έχουμε βρεί. Καθώς δεν μπορούμε να αποφανθούμε για τον οργανισμό καταχώρησης που είναι υπεύθυνος από τα δεδομένα του αρχείου pcap, θα μπορούσαμε να κάνουμε την εικασία πως η ζητούμενη κακοδιαχείριση θα μπορούσε να είχε γίνει

---

<sup>10</sup> Κυρίως ηλεκτρονικής τράπεζας (e-banking)



από τον οργανισμό καταχώρησης NAUNET. Από τα 44 ονόματα τομέα του pcap αρχείου μόλις τα 12 (27%) εμφανίζονται στην λίστα με τα 71 ονόματα τομέα feodo του άρθρου. Επομένως η εικασία μας είναι βασισμένη περισσότερο σε διαίσθηση παρά σε αποδείξεις. Δυστυχώς όμως δεν υπάρχουν περισσότερα δεδομένα ώστε να παρέχουμε μια πιο τεκμηριωμένη απάντηση στο ζητούμενο.

## 4. Συμπεράσματα

Στην παρούσα εργασία ασχοληθήκαμε με την διερεύνηση ενός μολυσμένου υπολογιστή ρομπότ και τον τρόπο με τον οποίο το πρωτόκολλο DNS χρησιμοποιήθηκε καταχρηστικά για τη λειτουργικότητά του. Είδαμε πώς χρησιμοποιήθηκε ένας αλγόριθμος DGA για τη δημιουργία ενός αριθμού ονομάτων τομέα που χρησιμοποιήθηκαν για τη σύνδεση με τους εξυπηρετητές εντολών και ελέγχου του botnet. Από αυτή την επίθεση μπορούμε να συμπεράνουμε ότι η κατάχρηση του DNS έχει μεγάλες δυνατότητες ως φορέας επίθεσης επειδή είναι εύκολη στην εκτέλεση, χαμηλού κινδύνου και παρέχει ένα εύκολα αναπτύξιμο σύστημα διανομής κακόβουλου λογισμικού. Αν θέλουμε να αμυνθούμε από αυτήν, πρέπει να βελτιώσουμε την υποδομή DNS και να αποτρέψουμε τους επιτιθέμενους από το να καταχωρούν τους εξυπηρετητές εντολών και ελέγχου με αυστηρότερες πολιτικές καταχωρητών, ή τουλάχιστον να αναπτύξουμε αποτελεσματικές μεθόδους για τον εντοπισμό τέτοιων κακόβουλων ονομάτων τομέα.

## Βιβλιογραφία

360Netlab (2022). *Netlab OpenData Project*. [online] Netlab OpenData Project. Available at: <https://data.netlab.360.com/dga/> [Accessed 7 May 2022].

Alakbarov, R., Rashidov, M., Mustafayev, T. and Yagubov, M. (2017). Internationalized Top Level Domain Names: Their Registration and Problems. *Problems of Information Society*, 08(1), pp.44–51. doi:10.25045/jpis.v08.i1.06.

Amazon (2019). *What is DNS? – Introduction to DNS - AWS*. [online] Amazon Web Services, Inc. Available at: <https://aws.amazon.com/route53/what-is-dns/> [Accessed 9 May 2022].

Casino, F., Lykousas, N., Homoliak, I., Patsakis, C. and Hernandez-Castro, J. (2021). Intercepting Hail Hydra: Real-time detection of Algorithmically Generated Domains. *Journal of Network and Computer Applications*, 190, p.103135. doi:10.1016/j.jnca.2021.103135.

Chen, Z., Wang, J.J. and Kwan, K. (2019). *Newly Registered Domains: Malicious Abuse by Bad Actors*. [online] Unit42. Available at: <https://unit42.paloaltonetworks.com/newly-registered-domains-malicious-abuse-by-bad-actors/> [Accessed 8 May 2022].

Dancs, D. (2020). *Detecting Active Abuse of Leaked Personal Data*. [BSc Thesis] Available at: [https://www.researchgate.net/publication/342003872\\_Detecting\\_active\\_abuse\\_of\\_leaked\\_personal\\_data](https://www.researchgate.net/publication/342003872_Detecting_active_abuse_of_leaked_personal_data) [Accessed 9 May 2022].

Directorate-General for Communications Networks, Content and Technology, Paulovics, I., Duda, A. and Korczynski, M. (2022). *Study on Domain Name System (DNS) abuse*. [online] Publications Office of the European Union. LU: Publications Office of the European Union. Available at: <https://op.europa.eu/en/publication-detail/-/publication/7d16c267-7f1f-11ec-8c40-01aa75ed71a1> [Accessed 5 May 2022].

Dooley, M. and Rooney, T. (2017). *DNS Security Management*. Wiley-IEEE Press, pp.1–83, 198–206.

Han, K.-S. and Im, E.G. (2011). A Survey on P2P Botnet Detection. *Lecture Notes in Electrical Engineering*, 120, pp.589–593. doi:10.1007/978-94-007-2911-7\_56.

Jenkins, Q. (2012). *Russian registrar NAUNET knowingly harbours Cybercriminals*. [online] www.spamhaus.org. Available at: <https://www.spamhaus.org/news/article/680/russian-registrar-naUNET-knowingly-harbours-cybercriminals> [Accessed 13 May 2022].

Kaspersky (2016). *Kaspersky Threats — Dapato*. [online] threats.kaspersky.com. Available at: <https://threats.kaspersky.com/en/threat/Trojan-Dropper.Win32.Dapato/> [Accessed 9 May 2022].

MITRE (2020). *Dynamic Resolution: Fast Flux DNS, Sub-technique T1568.001 - Enterprise / MITRE ATT&CK®*. [online] attack.mitre.org. Available at: <https://attack.mitre.org/techniques/T1568/001/> [Accessed 10 May 2022].

Nadji, Y., Antonakakis, M., Perdisci, R., Dagon, D. and Lee, W. (2013). Beheading hydras. In: *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security - CCS '13*. doi:10.1145/2508859.2516749.

Nadji, Y., Perdisci, R. and Antonakakis, M. (2017). Still Beheading Hydras: Botnet Takedowns Then and Now. *IEEE Transactions on Dependable and Secure Computing*, 14(5), pp.535–549. doi:10.1109/tdsc.2015.2496176.

Patsakis, C., Casino, F. and Katos, V. (2020). Encrypted and covert DNS queries for botnets: Challenges and countermeasures. *Computers & Security*, 88, p.101614. doi:10.1016/j.cose.2019.101614.

Plohmman, D., Fkie, F., Yakdan, K., Klatt, M., Bader, J. and Gerhards-Padilla, E. (2016). *A Comprehensive Measurement Study of Domain Generating Malware A Comprehensive Measurement Study of Domain Generating Malware*. [online] pp.263–278. Available at: [https://www.usenix.org/system/files/conference/usenixsecurity16/sec16\\_paper\\_plohmman.pdf](https://www.usenix.org/system/files/conference/usenixsecurity16/sec16_paper_plohmman.pdf) [Accessed 7 May 2022].

Prytuluk, M. (2022). *Common DNS return codes for any DNS service (and Umbrella)*. [online] Cisco Umbrella. Available at: <https://support.umbrella.com/hc/en-us/articles/232254248-Common-DNS-return-codes-for-any-DNS-service-and-Umbrella-> [Accessed 4 May 2022].

Roccia, T. (2017). *Emotet Trojan Acts as Loader, Spreads Automatically*. [online] McAfee Blog. Available at: <https://www.mcafee.com/blogs/other-blogs/mcafee-labs/emotet-trojan-acts-as-loader-spreads-automatically/> [Accessed 10 May 2022].

Sood, A.K. and Zeadally, S. (2016). A Taxonomy of Domain-Generation Algorithms. *IEEE Security & Privacy*, 14(4), pp.46–53. doi:10.1109/msp.2016.76.

Sophos (2012). *Detailed Analysis - Troj/MDrop-DZF - Viruses and Spyware - Advanced Network Threat Protection / ATP from Targeted Malware Attacks and Persistent Threats / sophos.com - Threat Center*. [online] www.sophos.com. Available at: <https://www.sophos.com/en-us/threat-center/threat-analyses/viruses-and-spyware/Troj~MDrop-DZF/detailed-analysis> [Accessed 7 May 2022].

Spooren, J., Preuveneers, D., Desmet, L., Janssen, P. and Joosen, W. (2019). Detection of algorithmically generated domain names used by botnets. In: *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*. doi:10.1145/3297280.3297467.

Vormayr, G., Zseby, T. and Fabini, J. (2017). Botnet Communication Patterns. *IEEE Communications Surveys & Tutorials*, [online] 19(4), pp.2768–2796. doi:10.1109/comst.2017.2749442.

## Παράρτημα

### split.py

```
#!/usr/bin/python3
# Script to split the family and domain name from the Netlab360 list -- Discard
the time data
with open('netlab360_list.txt', 'r') as handle:
    lines = handle.readlines()
with open('dga.txt', 'w+') as whandle:
    for i in range(18, len(lines)):
        whandle.write(lines[i].split()[0]+'\\t'+lines[i].split()[1]+'\\n')
```

### GE4.ipynb

```
#Imports
import pandas as pd
import numpy as np
from tqdm.notebook import tqdm

#Loading Data
dga_data = pd.read_csv('dga.csv', sep=';')
dns_data = pd.read_csv('domains.csv', sep=';')
```

#### *#Netlab 360 Data List*

```
dga_data.head()

   Family      Domain
0  nymaim  vcrjdothtrh.net
1  nymaim  kvkriojtu1.net
2  nymaim  astbynvq.info
3  nymaim  nppahcms.org
4  nymaim  gugbnrvck.com
```

#### *#Assignment Data*

```
dns_data.head()

   Family      Domain
0     NaN  anidgwelnidmzueo.ru
1     NaN  aopltfxjzsppylfh.ru
2     NaN  auvqjghelyqwtfsu.ru
3     NaN  axwiyyfbraskytvs.ru
4     NaN  aygrpumrlmymcwkh.ru
```

#### *#Adding Match column and changing Family data to empty strings*

```
dns_data['Match'] = 0
```

```

dns_data['Family'] = ''
dns_data.head()

   Family      Domain  Match
0      anidgwelnidmzueo.ru    0
1      aopltfxjzspylfh.ru    0
2      auvqjghelyqwtfsu.ru    0
3      axwiyyfbraskyts.ru    0
4      aygrpumrlmymcwk.ru    0

dns_dict = dns_data.to_dict(orient='records')11
dga_dict = dga_data.to_dict(orient='records')
print("Starting Search Script for matching Domain Names...")
sum = 0
for row1 in tqdm(dns_dict):
    for row2 in dga_dict:
        if row1['Domain'] == row2['Domain']:
            row1['Family'] = row2['Family']
            row1['Match'] = 1
            break
print("Script Ended")
dns_data = pd.DataFrame.from_dict(dns_dict)
for i in range(len(dns_data)):
    if dns_data['Match'][i] == 1:
        sum += 1
print("Matching Results Found from Netlab 360 Data List:  %d out of %d" %(sum
, len(dns_data)))

```

Starting Search Script for matching Domain Names...

Script Ended

Matching Results Found from Netlab 360 Data List: 44 out of 44

## Results

```
dns_data.head()
```

	Family	Domain	Match
0	feodo	anidgwelnidmzueo.ru	1
1	feodo	aopltfxjzspylfh.ru	1
2	feodo	auvqjghelyqwtfsu.ru	1
3	feodo	axwiyyfbraskyts.ru	1
4	feodo	aygrpumrlmymcwk.ru	1

```
dns_data['Family'].value_counts()
```

---

<sup>11</sup> Έγινε μετατροπή των dataframe σε dictionary έτσι ώστε να επιτευχθεί καλύτερη χρονική πολυπλοκότητα του αλγορίθμου

```
feodo      44  
Name: Family, dtype: int64  
  
dns_data['Match'].value_counts()
```

```
1      44  
Name: Match, dtype: int64
```

We can see that 44 out of 44 domains from our data was found in the [Netlab 360 Data List](#) and every one of them is classified as **feodo** dga family