Ανοικτό Πανεπιστήμιο Κύπρου

Σχολή Θετικών και Εφαρμοσμένων Επιστημών

Μεταπτυχιακό Πρόγραμμα Σπουδών
Ασφάλεια υπολογιστών και δικτύων

Πτυχιακή Εργασία

Αντιμετώπιση απειλών με χρήση Dark-Web και Τεχνητής νοημοσύνης

Ισίδωρος Μουλάς

Επιβλέπων Καθηγητής
Δρ. Σταύρος Σιαηλής

Νοέμβριος 2019

Ανοικτό Πανεπιστήμιο Κύπρου

Σχολή Θετικών και Εφαρμοσμένων Επιστημών

Μεταπτυχιακό Πρόγραμμα Σπουδών

Ασφάλεια υπολογιστών και δικτύων

Πτυχιακή Εργασία

Αντιμετώπιση απειλών με χρήση Dark-Web και Τεχνητής νοημοσύνης

Ισίδωρος Μουλάς

Επιβλέπων Καθηγητής
Δρ. Σταύρος Σιαηλής

Η παρούσα πτυχιακή εργασία υποβλήθηκε προς μερική εκπλήρωση των απαιτήσεων για απόκτηση πτυχιακού τίτλου σπουδών στο Μεταπτυχιακό Πρόγραμμα Σπουδών Ασφάλεια υπολογιστών και δικτύων από τη Σχολή Θετικών και Εφαρμοσμένων Επιστημών του Ανοικτού Πανεπιστημίου Κύπρου.

Νοέμβριος 2019

ΛΕΥΚΗ ΣΕΛΙΔΑ

Περίληψη

Η αντιμετώπιση των απειλών από τους κυβερνοεγκληματίες είναι ένα πεδίο εργασίας και έρευνας το οποίο είναι εξαιρετικά περίπλοκο, χαοτικό και απαιτεί εξειδικευμένες γνώσεις. Οι επιθέσεις που καλούμαστε να αντιμετωπίσουμε είναι απρόβλεπτες, πολύπλοκες και τις περισσότερες φορές είναι πολύ δύσκολο να αντιμετωπιστούν αποτελεσματικά.

Στο διάστημα που μεσολαβεί από την δημοσίευση της ευπάθειας έως την διόρθωση αλλά και την τελική αναβάθμιση των εγκατεστημένων εκδόσεων του λογισμικού, μεσολαβεί αρκετός χρόνος ώστε οι κυβερνοεγκληματίες ή άλλων ερευνητών να προσπαθήσει να δοκιμάσει αν η συγκεκριμένη ευπάθεια είναι πραγματική. Σε περίπτωση που η ευπάθεια είναι πραγματική τότε τα συγκεκριμένα συστήματα είναι ευάλωτα και κινδυνεύουν άμεσα.

Από την άλλη πλευρά υπάρχουν οι ευπάθειες που ανακαλύπτονται από κυβερνοεγκληματίες και δημοσιεύονται στο dark-web με σκοπό την πώληση σε ιστοσελίδες αγοραπωλησιών (marketplaces). Υπάρχουν βέβαια και οι περιπτώσεις που οι ευπάθειες δημοσιεύονται στο dark-web με σκοπό την ενημέρωση των υπολοίπων που παρακολουθούν την ιστοσελίδα ή την συζήτηση (forum). Αυτές οι επιθέσεις που βασίζονται σε ευπάθειες που μόλις έχουν δημοσιοποιηθεί ονομάζονται επιθέσεις μηδενικής ημέρας (zero day attacks) και προσβάλουν όλα τα συστήματα τα οποία είναι ευάλωτα στην συγκεκριμένη ευπάθεια (vulnerability).

Ο στόχος της πτυχιακής εργασίας είναι να δημιουργηθεί ένα σύστημα το οποίο προστατεύει ένα τοπικό δίκτυο από zero day attacks χρησιμοποιώντας πληροφορίες οι οποίες αλιεύονται μέσω crawlers από το ελεύθερο διαδίκτυο και το dark-web. Οι πληροφορίες αυτές συγκρίνονται και αξιολογούνται σε πραγματικό χρόνο (live) προκειμένου να αξιοποιούνται άμεσα με μοναδικό στόχο την προστασία του τοπικού δικτύου. Η διαδικασία είναι αυτοματοποιημένη και δεν χρειάζεται αλληλεπίδραση από τον χρήστη. Το σύστημα ελέγχει και καταγράφει αυτόματα τις υπηρεσίες του τοπικού δικτύου και ταυτόχρονα αναζητά για τις ευπάθειες που δημοσιεύονται και αφορούν τις συγκεκριμένες υπηρεσίες.

Το αποτέλεσμα της πτυχιακής είναι ότι οι διαχειριστές του τοπικού δικτύου ενημερώνονται άμεσα για τις νέες ευπάθειες των υπηρεσιών που χρησιμοποιούν πριν την διόρθωση του λογισμικού από τον κατασκευαστή. Στη συνέχεια μπορούν να

πραγματοποιήσουν διορθωτικές ενέργειες και επιπλέον ελέγχους μέχρι ο κατασκευαστής δημοσιοποιήσει μία νέα αναβάθμιση του λογισμικού.

Summary

Cyber threats are a field of work and research that is extremely complex, chaotic and requires specialized knowledge. The cyberattacks we are called upon to deal with are unpredictable, complex, and often difficult to deal with effectively in the short term.

Between the publication of the vulnerability to the software patch and the final upgrade of the installed versions of the software, it takes time for the community of cyber criminals or other researchers to try to determine if the vulnerability is real. If the vulnerability is real, then these systems are vulnerable and are at immediate risk of anyone having the necessary knowledge and time available.

On the other hand, there are vulnerabilities discovered by cyber criminals and published on the dark-web for sale on marketplaces. There are also cases where vulnerabilities are published on the dark-web for the purpose of informing others who are browsing over the site or the forum. These attacks that have just been released are called zero-day attacks and affect all systems that are vulnerable to that vulnerability.

The goal of the thesis is to create a system that protects a local network from zero-day attacks using information that is captured via crawlers from the free internet and the dark-web. This information is compared and evaluated in real time in order to be used immediately with the sole purpose of protecting the local network. The process is automated and requires no user interaction. The system automatically checks and records the services of the local network and at the same time searches for vulnerabilities that are published and related to those services.

The result of the thesis is that the local network administrators are immediately informed of the new vulnerabilities of the services they use before the software is patched-fixed by the manufacturer. They can then take corrective actions and additional checks until the manufacturer releases a new software upgrade.

Ευχαριστίες

Η παρούσα διατριβή αποτελεί μια μεγάλη προσωπική προσπάθεια σε ένα εξαιρετικά δυναμικό και ενδιαφέρον πεδίο έρευνας. Είναι μία προσπάθεια ταυτόχρονης έρευνας και ανάπτυξης του λογισμικού που περιγράφεται στις επόμενες σελίδες. Σημαντικός παράγοντας στην υλοποίηση της παρούσας διατριβής αποτελεί ο επιβλέπων καθηγητής Δρ. Σταύρος Σιαηλής, ο οποίος με τις κατευθυντήριες γραμμές και οδηγίες του κατέστη δυνατό η υλοποίηση του συστήματος και η συγγραφή της παρούσας έρευνας.

Κατάλογος εικόνων

Κατάλογος διαγραμμάτων

Ακρωνύμια

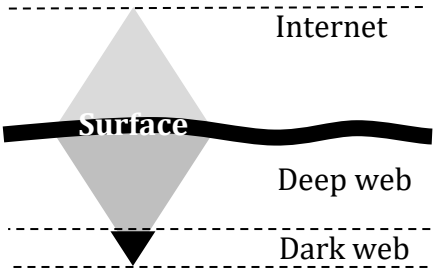
IP ..... Internet Protocol
IDS ..... Intrusion Detection System
IPS ..... Intrusion Prevention System
API ..... Application Programming Interface
SQL ..... Structured Query Language
NVD ..... National Vulnerabilities Database
POC και PoC ..... Proof Of Concept
nmap ..... Network Mapper
TCP ..... Transmission Control Protocol
UDP ..... User Datagram Protocol
ICMP ..... Internet Control Message Protocol
ARP ..... Address Resolution Protocol
MAC ..... Media Access Control
AI ..... Τεχνητή Νοημοσύνη

Περιεχόμενα

Κεφάλαιο 1

Εισαγωγή

Το Dark-web είναι το κρυφό κομμάτι του διαδικτύου που δεν ανιχνεύεται και δεν δεικτοδοτείται (indexed) από μηχανές αναζήτησης όπως το Google. Γενικά, παρουσιάζουμε το διαδίκτυο σαν ένα παγόβουνο (Εικόνα 1). Πάνω από την επιφάνεια του παγόβουνου έχουμε το συνηθισμένο διαδίκτυο που χρησιμοποιούμε καθημερινά όπως ειδησεογραφικές ιστοσελίδες, κοινωνικά δίκτυα, ιστοσελίδες καταστημάτων, εταιρειών κ.α. και είναι διαθέσιμο σε όλους χωρίς περιορισμούς. Κάτω από την επιφάνεια του παγόβουνου έχουμε το Deep Web, το οποίο είναι το μέρος του διαδικτύου στο οποίο οι μηχανές αναζήτησης δεν έχουν πρόσβαση όπως το email και άλλες ιστοσελίδες που απαιτούν πρόσβαση με χρήση λογαριασμού πρόσβασης (όνομα χρήστη και κωδικός πρόσβασης).



Εικόνα 1: The internet iceberg

Στη συνέχεια, πολύ κάτω από την επιφάνεια έχουμε το Dark-web το οποίο στην πραγματικότητα είναι μέρος του Deep Web. Σύμφωνα με ερευνητές μόνο το 4% του

διαδικτύου είναι διαθέσιμο στο ευρύ κοινό και το υπόλοιπο 96% είναι διαθέσιμο μόνο στο deep web (Farook Bin Rafiuddin, Minhas, & Singh Dhubb, 2017).

1.1 Dark-web

Το Dark-web που συχνά αναφέρεται και ως το σκοτεινό μέρος του ίντερνετ, δεν δημιουργήθηκε με αποκλειστικό σκοπό την παρανομία. Το Dark-web είναι ένα υποσύνολο του deep web στο οποίο πωλούνται ναρκωτικά, όπλα, παιδική πορνογραφία, παράνομο λογισμικό κ.α. Πάρα πολλές έρευνες έχουν δείξει ότι πολλές εξτρεμιστικές και τρομοκρατικές οργανώσεις έχουν χρησιμοποιήσει και συνεχίζουν να χρησιμοποιούν το δημόσιο διαδίκτυο για τις παράνομες ενέργειες τους. Εντούτοις πολύ λίγα είναι γνωστά για το τι συμβαίνει στο dark-web σχετικά με τις παράνομες ενέργειες αυτών των οργανώσεων (Zulkarnine, Frank, Mitchell, & Davies, 2016). Βέβαια ένα αποκεντρωμένο και κρυφό σύστημα που βασίζεται στην μυστικότητα και την ανωνυμία, δεν θα μπορούσε να χρησιμοποιηθεί για καλούς σκοπούς. Στο dark-web διευκολύνεται η πώληση ναρκωτικών χρησιμοποιώντας το ηλεκτρονικό κρυπτό νόμισμα του bitcoin. Εκτός από ναρκωτικά, στο Dark-web εκτελούνται αγοραπωλησίες όπλων και πολλών ακόμη παράνομων προϊόντων. Επίσης υπάρχει πρόσβαση σε πληρωμένους δολοφόνους και πλαστά διαβατήρια μέχρι και κατά παραγγελία χάκερς. Όλα τα παραπάνω είναι βέβαια σχετικά αφού και ο δυνητικός πελάτης των παραπάνω προϊόντων και υπηρεσιών μπορεί να πέσει θύμα απάτης.

Από την άλλη πλευρά στις χώρες που διαθέτουν απολυταρχικά καθεστώτα για παράδειγμα, πολλοί χρήστες του ίντερνετ εκμεταλλεύονται το Dark-web για να χρησιμοποιήσουν τα συνήθως απαγορευμένα κοινωνικά δίκτυα ή για να επισκεφτούν κλειδωμένες ενημερωτικές σελίδες. Το Dark-web είναι επίσης ιδιαίτερα δημοφιλής τόπος ανταλλαγής πληροφοριών από κρατικά μυστικά έως νέες ευπάθειες και ιοί ηλεκτρονικών υπολογιστών. Ειδικά οι ευπάθειες και όλα τα σχετικά λογισμικά με ιούς (malware, bots κα) είναι τα αγαπημένα παιχνίδια της σκοτεινής κοινότητας. Γενικότερα σε μία ιστοσελίδα στο dark-web μπορεί να δημοσιευτεί οτιδήποτε και είναι εντελώς αχαρτογράφητο αν οι πληροφορίες που παρουσιάζονται είναι αξιόπιστες.

Οι ιστοσελίδες του dark-web έχουν γίνει ο κύριος χώρος για ηλεκτρονική αγορά κακόβουλων προϊόντων και υπηρεσιών hacking από εγκληματίες στον κυβερνοχώρο. Ένα παράδειγμα που απεικονίζει αυτό το γεγονός είναι μια ευπάθεια που στοχεύει τη Microsoft. Η ευπάθεια για το λειτουργικό σύστημα των Windows ήταν προς πώληση σε μια αγορά dark-web τον Μάρτιο του 2015 (Nunes, και συν., 2016). Η ευπάθεια αποκαλύφθηκε από τη Microsoft ένα μήνα νωρίτερα, χωρίς την δημόσια δημοσίευση εκείνη τη στιγμή. Τέσσερις μήνες μετά τη διαθεσιμότητα της ευπάθειας FireEye1, ανακαλύφθηκε το Dyre Banking trojan το οποίο εκμεταλλευόταν την συγκεκριμένη ευπάθεια, σχεδιασμένο να στοχεύσει χρηματοπιστωτικούς οργανισμούς για να κλαπούν πληροφορίες πιστωτικών καρτών. Πολλά marketplaces ενδεχομένως συνδέονται με το Dyre Banking Trojan, και θα μπορούσε να αναγνωριστεί αυτόματα αν τουλάχιστον ένας από αυτούς είχε ήδη επιβεβαιωθεί ότι προσφέρει την ευπάθεια προς πώληση. Σε πιο περίπλοκα σενάρια, οι κοινότητες αυτές (marketplaces, forums) θα μπορούσαν να αντιστοιχούν σε ομάδες ατόμων που ασχολούνται με παρόμοια προϊόντα ή υπηρεσίες σε πολλαπλά πεδία hacking ταυτόχρονα, όπως κάρτες, phishing και keyloggers (Marin, Almukaynizi, Nunes, & Shakarian, 2018).
Για να έχουμε πρόσβαση στο Dark-web θα πρέπει να χρησιμοποιήσουμε ειδικό λογισμικό.
Το πιο γνωστό λογισμικό είναι το Tor το οποίο μας δίνει την δυνατότητα να έχουμε πρόσβαση σε ιστοσελίδες με επέκταση .onion (αντίστοιχο του .com, .gr, .cy) οι οποίες είναι διαθέσιμες μόνο μέσω του Dark-web. Σε αυτές τις σελίδες στις οποίες δεν υπάρχει κανένας έλεγχος από καμία αρχή, κράτος ή υπηρεσία, είναι διαθέσιμες πληροφορίες για πωλήσεις όπλων, ναρκωτικών, παιδικής πορνογραφίας κ.α.

1.2 Vulnerabilities - Ευπάθειες

Ένα σημαντικό μέρος του Dark-web, το οποίο μας ενδιαφέρει άμεσα, είναι όλες οι πληροφορίες που δημοσιεύονται και αφορούν νέες ευπάθειες (vulnerabilities). Οι ευπάθειες ενός λογισμικού (software) ή υλικού (hardware) μπορεί να εντοπιστούν κατευθείαν από τον κατασκευαστή, από ανεξάρτητους ερευνητές ή από τρίτα πρόσωπα όπως κυβερνοεγκληματίες. Συνήθως οι ευπάθειες δημοσιεύονται από τους ερευνητές σε αξιόπιστες ιστοσελίδες στο συνηθισμένο διαδίκτυο και ταυτόχρονα ενημερώνουν τον κατασκευαστή με ηλεκτρονικό ταχυδρομείο, με στόχο ο κατασκευαστής να προχωρήσει σε

διόρθωση του προϊόντος. Τις περισσότερες φορές η διόρθωση γίνεται με αναβάθμιση λογισμικού (software upgrade, patch, service pack) προκειμένου να αντιμετωπιστεί αποτελεσματικά η ευπάθεια που δημοσιεύτηκε.

Μία ευπάθεια στο λογισμικό είναι ένα παράδειγμα ενός σφάλματος στις προδιαγραφές, την ανάπτυξη ή τη διαμόρφωση του λογισμικού έτσι ώστε να μπορεί να εκτελεστεί χωρίς προβλήματα. Οι αναφορές ευπάθειας για εφαρμογές έχει μελετηθεί εδώ και πολλά χρόνια. Για το λογισμικό, συνολικά, ο αριθμός των ανακοινωθέντων ευπαθειών έχει αλλάξει από 3 ευπάθειες ανά ημέρα το 2000 έως πάνω από 16 ανά ημέρα το 2008 σύμφωνα με την National Vulnerabilities Database (NVD, n.d.). Αυτό το υψηλό ποσοστό ευπαθειών έχει εστιάσει την προσοχή στο πολύ πρακτικά και άμεσα θέματα διαχείρισης της ενημερωμένης έκδοσης κώδικα, στις διαδικασίες γνωστοποίησης ευπάθειας και στη ταχύτητα εγκατάστασης των νέων εκδόσεων της εφαρμογής (McQueen, McQueen, Boyer, & Chaffin, 2009).

Οι επιθέσεις zero day attacks εκμεταλλεύονται μια άγνωστη ευπάθεια του λογισμικού και μπορεί να επηρεάσει σημαντικά όλα τα ευάλωτα συστήματα. Ανάλογα την σπουδαιότητα της ευπάθειας, η ζημιά σε ένα πληροφοριακό σύστημα μπορεί να πάρα πολύ μεγάλη έως ασήμαντη. Συχνά η διάγνωση της ευπάθειας αλλά και η διόρθωση του λογισμικού παίρνουν σημαντικό χρόνο. Για όσο διάστημα δεν έχει δημοσιοποιηθεί ενημερωμένη έκδοση του λογισμικού, το πληροφοριακό σύστημα είναι ευάλωτο στην απειλή. Αφού εντοπιστεί η ευπάθεια οι διαχειριστές και οι προγραμματιστές προσπαθούν να διορθώσουν την ευπάθεια όσο το δυνατόν γρηγορότερα.

Η ευπάθεια ονομάζεται ευπάθεια μηδενικής ημέρας (zero day vulnerability) επειδή η ευπάθεια έχει μόλις ανακαλυφθεί και τα πληροφοριακά συστήματα που πλήττονται δεν έχουν ενημερωθεί ακόμα για να αντιμετωπίσουν την απειλή με ασφάλεια.

Όταν οι κυβερνοεγκληματίες ή οι ερευνητές κατασκευάζουν μικρά λογισμικά (PoC), για να αποδείξουν στον κατασκευαστή ή στον υπόλοιπο κόσμο την ευπάθεια, την ίδια στιγμή ο κατασκευαστής είτε εργάζεται στην διόρθωση της ευπάθειας ή ακόμα δεν γνωρίζει την ύπαρξη της. Σε ακριβώς αυτό το σημείο που η ευπάθεια γίνεται γνωστή στο κοινό μέσω δημοσιεύσεων σε ερευνητικές ιστοσελίδες τότε η ευπάθεια γίνεται μία ευπάθεια μηδενικής ημέρας. Εάν μια ευπάθεια ανακαλύπτεται από τους ερευνητές - εταιρείες λογισμικού ασφάλειας στο διαδίκτυο ή προμηθευτές λογισμικού - η τάση είναι να παραμείνει κρυφό

μέχρι ο κατασκευαστής λογισμικού να έχει μια ενημερωμένη έκδοση λογισμικού η οποία διορθώνει την ευπάθεια.. Ωστόσο, σε ορισμένες περιπτώσεις, οι ερευνητές ασφάλειας ή οι κατασκευαστές του λογισμικού πρέπει να ανακοινώσουν δημοσίως το ελάττωμα, επειδή οι χρήστες θα μπορούσαν να αποφύγουν το πρόβλημα, για παράδειγμα, κάνοντας διαγραφή ενός συγκεκριμένου προγράμματος ή βεβαιώνοντας ότι δεν θα ανοίξει ένα συγκεκριμένο συνημμένο αρχείο ηλεκτρονικού ταχυδρομείου. Ή η ευπάθεια μπορεί να ανακαλυφθεί από ένα χρήστη και να τελειώσει σε ένα blog ή να δημοσιοποιηθεί με άλλο τρόπο. Το χειρότερο σενάριο είναι η ευπάθεια να ανακαλυφθεί από κυβερνοεγκληματίες και να τεθεί προς πώληση σε κάποιο marketplace του dark-web (What are zero-day attacks, n.d.).

Πολλές φορές οι ευπάθειες μηδενικής ημέρας γίνονται αντικείμενο αγοραπωλησίας στο dark-web. Για τον άνθρωπο ή την ομάδα που ανακάλυψε την ευπάθεια (συνήθως κάποιος κυβερνοεγκληματίας) είναι ένα εύρημα το οποίο αποφέρει σημαντικά έσοδα αλλά και σημαντική φήμη, αφού όλα τα πληροφοριακά συστήματα που αφορούν την συγκεκριμένη ευπάθεια είναι ευάλωτα. Από την άλλη πλευρά υπάρχουν πολλοί αγοραστές που θα αγοράσουν το λογισμικό που αποδεικνύει την ευπάθεια προκειμένου να εκμεταλλευτούν την ευπάθεια για δικό τους όφελος ή απλά για να κατανοήσουν τον τρόπο λειτουργίας του συγκεκριμένου ευρήματος. Αυτές οι επιθέσεις που βασίζονται σε άγνωστες ευπάθειες ή οποίες μόλις έγιναν γνωστές μπορούν να περάσουν απαρατήρητες για μεγάλο χρονικό διάστημα. Πολλές φορές αυτός (ένας ή ομάδα) που ανακαλύπτει την ευπάθεια, την χρησιμοποιεί για δικό του όφελος εντελώς αθόρυβα και χωρίς να έχει δημοσιοποιηθεί κανένα στοιχείο.

Οι επιθέσεις μηδενικής ημέρας είναι απρόβλεπτες και εξαιρετικά δύσκολο να προβλεφθούν. Ο σκοπός είναι η αναζήτηση των zero-days exploits που αναρτώνται σε διάφορες ιστοσελίδες στο ελεύθερο διαδίκτυο και στο dark-web με στόχο την λήψη έγκαιρων μέτρων για την ασφάλεια των συστημάτων. Η αναζήτηση στο dark-web, το οποίο είναι ασταθές και απρόβλεπτο αφού όλο το σκεπτικό του dark-web είναι δυναμικό, είναι μία διαδικασία επίπονη η οποία χρειάζεται συνεχώς παρακολούθηση και διορθώσεις των προγραμμάτων που πραγματοποιούν τις αναζητήσεις. Οι ιστοσελίδες που δημοσιεύουν τα στοιχεία έχουν διαφορετικές δομές και περιεχόμενο και αναμένεται οι πληροφορίες να είναι μη δομημένες και σχετικά ασαφής για το περιεχόμενο τους.

1.3 Ερευνητικά Ερωτήματα

Το βασικό ερευνητικό ερώτημα που εγείρεται είναι:

- αν μπορούν να αξιοποιηθούν τα δεδομένα από το dark-web για την προστασία ενός τοπικού δικτύου
- και αν αυτά τα δεδομένα μπορούν να αξιοποιηθούν με αυτοματοποιημένο τρόπο χωρίς την αλληλεπίδραση του χρήστη.

Σε ένα τοπικό δίκτυο, οι απειλές αφορούν εσωτερικούς παράγοντες από τους ίδιους τους χρήστες τSου τοπικού δικτύου, ενώ οι απειλές από εξωτερικούς παράγοντες μπορεί να περιλαμβάνουν παραβιάσεις από την χρήση κακόβουλου λογισμικού (malware) ή επιθέσεις κυβερνοεγκληματιών εκτός του δικτύου. Στατιστικά, το μεγαλύτερο μέρος των παραβιάσεων οφείλεται σε εξωτερικούς παράγοντες αλλά ανάλογα την χρήση και το είδος του τοπικού δικτύου ενδέχεται οι παραβάσεις να είναι μεγαλύτερες από εσωτερικούς παράγοντες. Για την πρόληψη και την αντιμετώπιση των απειλών αυτών, εξετάζονται και συγκρίνονται τα υπάρχοντα εργαλεία ανίχνευσης απειλών δικτύου. Για τον έλεγχο ενός δικτύου, το πρώτο βήμα αποτελεί η χαρτογράφησή του, η οποία μπορεί να πραγματοποιηθεί με διάφορα εργαλεία.

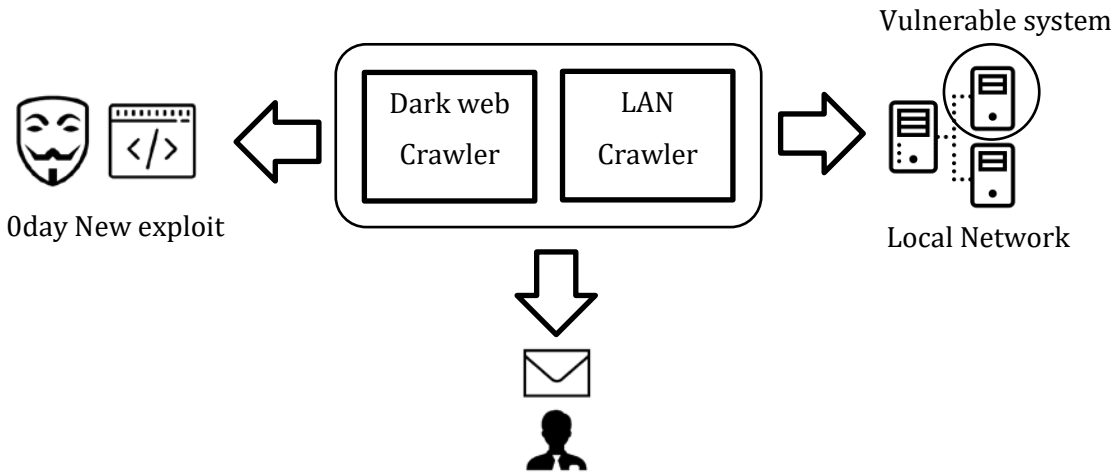
Το σημερινό περιβάλλον απειλών από όλη την κοινότητα που ασχολείται με την ασφάλεια εξελίσσεται συνεχώς και απαιτεί μεγαλύτερη προστασία και έλεγχο με τα πιο προηγμένα εργαλεία ασφάλειας και ανάλυσης. Από την μία πλευρά είναι οι ερευνητές-εταιρείες που αναζητούν ευπάθειες για την προστασία των πελατών όπως η McAfee, Norton, Kaspersky, Microsoft. Αυτές οι εταιρείες ερευνούν για ευπάθειες και αγοράζουν όσο το δυνατόν γρηγορότερα τις πληροφορίες για τις νέες ευπάθειες. Από την άλλη πλευρά οι κυβερνοεγκληματίες και όλων εκείνων που θέλουν να εκμεταλλευτούν τις ευπάθειες είτε πουλώντας τις πληροφορίες στις εταιρείες ή για δικό τους όφελος.

Η αναζήτηση των ευπαθειών στο dark-web θα μπορούσε να γίνει χρησιμοποιώντας λέξεις-κλειδιά ή φράσεις σε σελίδες ενδιαφέροντος. Οι κοινές μηχανές αναζήτησης όπως το google.com αναζητούν και παρουσιάζουν τα δεδομένα σχετικά με το επίπεδο επιφάνειας του διαδικτύου. Η ίδια αναζήτηση θα πραγματοποιείται στο dark-web συλλέγοντας

κρίσιμα δεδομένα και εξοικονομώντας πολύτιμο χρόνο στην πρόληψη και στην καταπολέμηση των νέων ευπαθειών.

Από την μία πλευρά έχουμε τις νέες ευπάθειες που καθημερινά εμφανίζονται και από την άλλη πλευρά έχουμε να προστατέψουμε ένα δίκτυο συσκευών και υπολογιστών. Ο χρόνος που μεσολαβεί από την εμφάνιση της ευπάθειας μέχρι την διορθωτική κίνηση στο σύστημα το οποίο εντοπίστηκε το πρόβλημα, είναι πάρα πολύ σημαντικός.

Ο στόχος του παρόντος κειμένου είναι η δραματική μείωση του χρόνου στον οποίο γίνονται οι διορθωτικές κινήσεις από τους διαχειριστές (οποιεσδήποτε και αν είναι αυτές), έως την στιγμή που έχει ήδη εντοπιστεί μία ευπάθεια.



Διάγραμμα 1: Αρχιτεκτονική

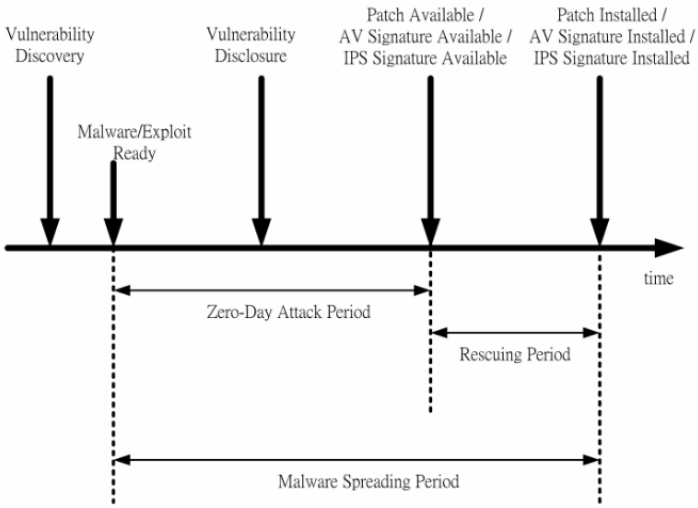
Οι διαχειριστές του δικτύου να μπορούν να ενημερώνονται έγκαιρα προκειμένου να προστατεύσουν το δίκτυο υπολογιστών που διαχειρίζονται.

Η διάρθρωση της διατριβής είναι το Κεφάλαιο 2 παρουσιάζει παρόμοιες εργασίες που έχουν υλοποιηθεί σχετικά με crawlers στο dark-web. Στο κεφάλαιο αυτό καταγράφονται οι τεχνικές και παρόμοιες υλοποιήσεις που έχουν ολοκληρωθεί στο παρελθόν. Στο Κεφάλαιο 3 παρουσιάζεται η μεθοδολογία που ακολουθήθηκε για να υλοποιηθούν τα ερευνητικά ερωτήματα. Επίσης γίνεται λεπτομερής αναφορά για τον τρόπο σχεδιασμού και υλοποίησης όλων των crawlers, scripts που χρησιμοποιήθηκαν. Τέλος, στο Κεφάλαιο 4 αναφέρονται τα συμπεράσματα από την έρευνα και τα μελλοντικά βήματα για την περαιτέρω αξιοποίηση των ευρημάτων.

Κεφάλαιο 2

Παρόμοιες Εργασίες

Οι επιθέσεις μηδενικής ημέρας (zero day attack) είναι πάρα πολύ σημαντικές. Το Διάγραμμα 2 εμφανίζει την διάρκεια ζωής μίας τέτοιας επίθεσης σε σχέση με τον χρόνο. Αφού ανακαλυφθεί μία ευπάθεια, οι κυβερνοεγκληματίες είναι έτοιμοι να την εκμεταλλευτούν την ευπάθεια, να την διαδώσουν αλλά και να την εκμεταλλευτούν. Οι κατασκευαστές λογισμικού πολλές φορές ανακαλύπτουν την ευπάθεια αφού έχουν μολυνθεί αρκετοί χρήστες και συστήματα.



Διάγραμμα 2: Η διάρκεια μίας επίθεσης

Μία έρευνα (Kao, και συν., 2015) προτείνει την δημιουργία honeypots στο τοπικό δίκτυο τα οποία θα λειτουργούν ως ενεργά θύματα του τοπικού δικτύου. Το προτεινόμενο σύστημα καταγράφει την συνηθισμένη κίνηση στο τοπικό δίκτυο και ανιχνεύει οποιαδήποτε ανωμαλία προκύψει, όπως η προσπάθεια σύνδεσης σε κάποια πόρτα (tcp/udp) που δεν

χρησιμοποιείται συχνά. Αν αυτή η προσπάθεια επαναληφθεί και σε άλλο honeypot, τότε το σύστημα θεωρεί ότι πρόκειται για μία επίθεση μηδενικής ημέρας και αυτόματα μπλοκάρει την κίνηση στους πραγματικούς διακομιστές. Το όλο σύστημα στηρίζεται σε καταγραφή στατιστικών δεδομένων (στην έρευνα υπήρχαν στοιχεία 6 χρόνων). Ομοίως με τον Kao σε μία άλλη έρευνα (Sornalakshmi, 2017) προτείνεται η παρακολούθηση των υπηρεσιών ενδιαφέροντος και η δημιουργία honeypots τα οποία θα αποπροσανατολίσουν τον επιτιθέμενο και δυσκολέψουν την αναζήτηση του πραγματικού υπολογιστή ή εφαρμογής που έχει μία ευπάθεια μηδενικής ημέρας. Συμπληρωματικά ο Albert (Sagala, 2015) προτείνει την δημιουργία honeypots και ταυτόχρονα την αυτόματη δημιουργία κανόνων σε ένα IDS/IPS όπως το snort (4.2 Μελλοντική Εργασία) αλλά και οι Utpal-Girish (Upadhyay & Khilari, 2016) προτείνουν παρόμοια αρχιτεκτονική με honeypots.

Ένα Honeypot είναι ένα σύστημα που ενεργεί ως διακομιστής που αποθηκεύει ή χειρίζεται σημαντικές ή σημαντικές πληροφορίες ορίζεται ως σύστημα honeypot. Το σύστημα αφήνεται εσκεμμένα ευάλωτο στους επιτιθέμενους προκειμένου να μπορεί να μελετηθούν και να αναλυθούν οι τεχνικές και τα εργαλεία που χρησιμοποιεί ο εισβολέας. Όλα τα δεδομένα που χρησιμοποιούνται από αυτό το σύστημα είναι ψεύτικα και παραπλανητικά και δεν παρέχει καμία αξία στον επιτιθέμενο. Οι επιτιθέμενοι είναι εξαπατημένοι αποκαλύπτοντας τις γνώσεις τους, τα εργαλεία και κάποια στιγμή δεδομένα και δεν παίρνουν τίποτα σε αντάλλαγμα παρά μόνο ψεύτικα στοιχεία και δεδομένα (Lihet & Pr.Dr. Dadarlat, 2018).

Μία άλλη έρευνα (Bau, Bursztein, Gupta, & Mitchell, 2010) προτείνει την δημιουργία web εφαρμογής η οποία θα αναζητά αυτόματα για γνωστές ευπάθειες σε ιστοσελίδες. Οι ερευνητές επικοινώνησαν με τις ποιο γνωστές εταιρείες που διαθέτουν εμπορικά εργαλεία για αναζήτηση ευπαθειών (vulnerabilities scanners) και σύγκρινε τα αποτελέσματα από τις αναζητήσεις. Το εργαλείο στηρίζεται σε ευπάθειες που έχουν ήδη ενσωματωθεί στους scanners και διαπιστώθηκε ότι ανάμεσα στα εμπορικά εργαλεία υπήρχαν διαφορές με αποτέλεσμα να μην αναγνωρίζουν τις ίδιες ευπάθειες.

Σύμφωνα με την έρευνα (Graham, Maynor, & Security, 2011), οι εταιρείες που διαθέτουν προϊόντα ασφάλειας (antivirus, firewall) αγοράζουν και πουλούν πληροφορίες για τις νέες ευπάθειες και στη συνέχεια ενσωματώνουν τις νέες πληροφορίες στα προϊόντα τους είτε πρόκειται για scanners, IPS, IDS, anti-virus κτλ. Μία άλλη τακτική που ακολουθούν οι

κυβερνοεγκληματίες είναι να αποκωδικοποιούν τις αναβαθμίσεις των προϊόντων ασφάλειας (virus definition updates) προκειμένου να ανακαλύπτουν τις νέες ευπάθειες. Ειδικά σε προϊόντα ανοικτού κώδικα όπως το Snort, οι αναβαθμίσεις των ευπαθειών είναι σε απλό κείμενο και είναι πολύ εύκολο κάποιος να καταλάβει το είδος και την λειτουργία της ευπάθειας αλλά και σε πιο προϊόν ή υπηρεσία αναφέρεται. Το γενικό συμπέρασμα της έρευνας είναι ότι οι ευπάθειες μηδενικής ημέρας μπορούν πολύ εύκολα να αναζητηθούν από τα προϊόντα ασφάλειας αφού οι περισσότεροι κατασκευαστές αναβαθμίζουν καθημερινά τα προϊόντα με τις νέες ευπάθειες.

Η έρευνα (Shiaeles, Kolokotronis, & Bellini, 2019) προτείνει τη δημιουργία ενός συστήματος το οποίο ερευνά το dark-web αλλά και άλλες πηγές για νέες ευπάθειες με αυτοματοποιημένο τρόπο. Οι crawlers της έρευνας καταγράφουν τα στοιχεία σε μία τοπική βάση δεδομένων. Τα στοιχεία καταγράφονται σε κατηγορίες ανάλογα την ευπάθεια και μπορούν να αντληθούν χρήσιμα στατιστικά στοιχεία. Τα στοιχεία που βρέθηκαν από την λειτουργία του συστήματος συγκρίθηκαν με τις τελευταίες αναφορές για ευπάθειες και βρέθηκαν πάρα πολλά κοινά στοιχεία.

Η επιτυχία μια επίθεσης μηδενικής ημέρας εξαρτάται από το χρονικό διάστημα που μεσολαβεί από την δημοσιοποίηση της ευπάθειας και την λήψη μέτρων από την πλευρά του πληροφοριακού συστήματος. Πολλές φορές μία ευπάθεια από την στιγμή που δημοσιοποιηθεί μπορεί να είναι εξαιρετικά χρονοβόρο μέχρι την εφαρμογή της ενημερωμένης έκδοσης σε όλα τα συστήματα που είναι ευάλωτα. Μία ενημερωμένη έκδοση η οποία αντιμετωπίζει μία ευπάθεια (όσο σημαντική και ασήμαντη και αν είναι) απαιτεί τις περισσότερες φορές αλληλεπίδραση του χρήστη που θα πρέπει να αναβαθμίσει χειροκίνητα την συγκεκριμένη ευάλωτη έκδοση του λογισμικού. Μία έρευνα αναφέρει ότι οι χρήστες καθυστερούν μία σημαντική αναβάθμιση του λογισμικού που αφορά το λειτουργικό σύστημα τουλάχιστον 80 ημέρες (Fransesco, Joanna, Aurelien, Michel, & Wendy, May 2017) από την ημέρα που δημοσιεύεται η αναβάθμιση. Πολλοί λίγοι χρήστες προετοιμάζονται για την αναβάθμιση του λογισμικού και ακόμα λιγότεροι αναφέρουν ότι ο λόγος που εφαρμόζουν την συγκεκριμένη αναβάθμιση είναι κάποιο κενό ασφαλείας που έχει εντοπιστεί. Οι εξειδικευμένοι χρήστες προχωρούν σε αναβάθμιση τουλάχιστον 50% γρηγορότερα από τους υπόλοιπους χρήστες.

Μία άλλη έρευνα προτείνει μία αρχιτεκτονική η οποία υιοθετεί μια προσέγγιση δύο φάσεων για τη συλλογή δεδομένων (Koloveas, Chantzios, Tryfonopoulos, & Skiadopoulos, 2019). Αρχικά χρησιμοποιείται ένας μηχανισμός ανίχνευσης με μηχανική μάθηση για να αναζητά ευπάθειες από τις ιστοσελίδες ενδιαφέροντος, ενώ στη δεύτερη φάση χρησιμοποιούνται στατιστικά μοντέλα για την απεικόνιση των αποτελεσμάτων. Η προτεινόμενη αρχιτεκτονική υλοποιείται χρησιμοποιώντας αποκλειστικά εργαλεία ανοιχτού κώδικα και μια προκαταρκτική αξιολόγηση που διενεργήθηκε στα αποτελέσματα από το πλήθος των στοιχείων αποδεικνύει την αποτελεσματικότητά της. Στην έρευνα γίνεται αναφορά για την κατασκευή crawlers για το διαδίκτυο και το dark-web καθώς επίσης για διάφορες άλλες πηγές όπως κοινωνικά δίκτυα. Στη συνέχεια τα αποτελέσματα αξιολογούνται και βαθμολογούνται (rank) σύμφωνα με τους αλγόριθμους της έρευνας και καταγράφονται σε μία βάση δεδομένων.

Μία άλλη έρευνα προτείνει στην κατασκευή ενός αυτοματοποιημένου crawler (Schäfer, et al., 2019) ο οποίος θα αναζητά όλες τις διευθύνσεις .onion του δικτύου dark-web ακολουθώντας όλες τις πιθανές υπερσυνδέσεις που υπάρχουν στην κάθε σελίδα. Στην έρευνα παρουσιάστηκε ότι οι ιστοσελίδες που αναζητήθηκαν από τον crawler στη συνέχεια μετά από μερικές ημέρες αυτές δεν υπήρχαν πλέον και ήταν ανενεργές, το οποίο επιδεικνύει τον βαθμό δυσκολίας χαρτογράφησης του dark-web. Για την ταχύτερη αναζήτηση των υπηρεσιών χρησιμοποιήθηκαν πολλαπλές μηχανές και υπολογιστικά συστήματα.

Το λογισμικό Sixgill (https://www.cybersixgill.com/) παρέχει πρόσβαση στο dark-web και παρουσιάζει δεδομένα που αναζητούνται μέσω crawlers. Μία έρευνα χρησιμοποίησε τα δεδομένα του SixGill (KADOGUCHI, HAYASHI, HASHIMOTO, & OTSUKA, 2019) τα οποία προέρχονται από διάφορες δημοσιεύσεις στο dark-web. Τα δεδομένα επεξεργάστηκαν με το εργαλείο doc2vec και διαπιστώθηκε ότι μία τεχνική μηχανικής μάθησης (machine learning) είναι αποδοτική όσο αφορά τις δημοσιεύσεις στο dark-web. Βέβαια, όπως καταλήγει και η έρευνα, το μεγαλύτερο πρόβλημα για την σωστή εκπαίδευση του μοντέλου είναι ο περιορισμένος αριθμός δεδομένων. Για ένα εκπαιδευτεί ένα μοντέλο μηχανικής μάθησης χρειάζεται μεγάλος όγκος δεδομένων.

Μία ακόμα έρευνα (Baravalle, Lopez, & Lee, 2016) επικεντρώθηκε στην ανάλυση των δεδομένων ενός γνωστού market place (agora). Στην έρευνα έγινε προσπάθεια για την

κατασκευή ενός crawler ο οποίος θα ήταν αυτοματοποιημένος και δεν θα γινόταν αντιληπτός από τα συστήματα ασφάλειας της ιστοσελίδας. Πολλές φορές κατά την διάρκεια της έρευνας αυτό έγινε αντιληπτό με αποτέλεσμα το crawling της σελίδας να γίνεται δυσκολότερο και ακόμα πιο αργό αφού ο διακομιστής διέκοπτε συχνά την σύνδεση. Ένα άλλο σημαντικό πρόβλημα που εντοπίστηκε ήταν η χρήση captcha το οποίο καθυστερούσε την διαδικασία και έπρεπε πολλές φορές κατά την διάρκεια της αναζήτησης ο χρήστης να πληκτρολογεί τον σωστός κωδικό captcha προκειμένου ο crawler να συνεχίσει την εργασία του. Η διαδικασία ήταν σχεδόν αυτοματοποιημένη και από τα συμπεράσματα προκύπτει ότι μπορούν να εξαχθούν χρήσιμα δεδομένα με τις κατάλληλες ενέργειες.

Από τις παραπάνω έρευνες προκύπτει ότι μία αναζήτηση στο dark-web με λέξεις κλειδιά είναι εφικτή αρκεί για κάθε σελίδα του dark-web να υπάρχει ένα μοναδικό script που θα ελέγχει και θα καταγράφει τα δεδομένα της σελίδας (crawling). Όλες οι έρευνες αναζητούν ευπάθειες στο dark-web που αφορούν οποιαδήποτε υπηρεσία, συσκευή λογισμικό με σκοπό να καταγράψουν την ευπάθεια σε μία μεγάλη βάση δεδομένων. Επίσης είναι φανερό ότι το dark-web μπορεί να χρησιμοποιηθεί ως πηγή άντλησης στοιχείων (Shiaeles, Kolokotronis, & Bellini, 2019) και παρέχει σημαντικές πληροφορίες σχετικά με τις ευπάθειες μηδενικής ημέρας. Η παρούσα έρευνα επικεντρώνεται στην αυτοματοποιημένη άντληση στοιχείων από το dark-web σχετικά με τις συγκεκριμένες υπηρεσίες που υπάρχουν στο τοπικό δίκτυο ενδιαφέροντος. Η έρευνα επικεντρώνεται στο τοπικό δίκτυο ενδιαφέροντος και αναζητά ευπάθειες που αφορούν το συγκεκριμένο δίκτυο είτε πρόκειται για ευπάθειες σε συσκευές, υπηρεσίες ή λογισμικό. Η έρευνα έχει σκοπό στην άμεση ενημέρωση των διαχειριστών για τις νέες πιθανές ευπάθειες του τοπικού δικτύου είτε πρόκειται για νέες ευπάθειες που αφορούν υπάρχουσες συσκευές, είτε πρόκειται για παλιές ευπάθειες οι οποίες αφορούν νέες συσκευές (με παλιές εκδόσεις λογισμικών) οι οποίες συνδέονται στο τοπικό δίκτυο.

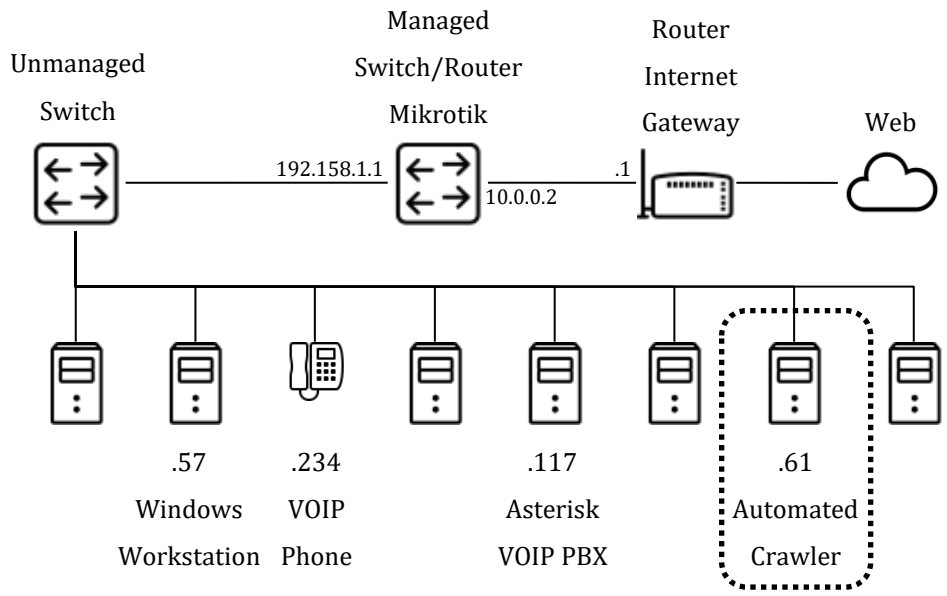
Κεφάλαιο 3

Προτεινόμενη Μεθοδολογία

Σε ένα ιδιωτικό δίκτυο μίας εταιρείας ή οργανισμού (παράδειγμα Διάγραμμα 3) υπάρχουν διάφοροι υπολογιστές και δικτυακές συσκευές συνδεδεμένες όπως δρομολογητές (routers), εξυπηρετητές (servers), θέσεις εργασίας (workstations) ακόμα και τηλεφωνικές συσκευές. Κάθε ένα από αυτά τα συστήματα μπορεί να έχει διάφορες υπηρεσίες διαθέσιμες προς τους υπόλοιπους χρήστες όπως για παράδειγμα είναι ο web server, ssh, sql, pop3 κ.α. Στο δίκτυο μπορεί να συνδεθούν ή αποσυνδεθούν διάφορες συσκευές ή να προστεθούν ή αφαιρεθούν διάφορες υπηρεσίες από τις συσκευές.

Το προτεινόμενο σύστημα αφορά ένα αυτοματοποιημένο σύστημα ελέγχου του τοπικού δικτύου το οποίο θα χαρτογραφεί το δίκτυο και στη συνέχεια θα αναζητά στο dark-web για ευπάθειες που αφορούν τις υπηρεσίες που είναι εκτεθειμένες στο δίκτυο ελέγχου. Οι αυτοματοποιημένες διαδικασίες θα πραγματοποιούνται μέσω crawlers οι οποίοι είναι διαφορετικοί για κάθε εργασία. Οι crawlers είναι προγράμματα-scripts, συνήθως μικρού μεγέθους, τα οποία πραγματοποιούν μία συγκεκριμένη εργασία. Οι crawlers που χρειάζονται είναι για την αυτόματη χαρτογράφηση του τοπικού δικτύου και ένας crawler για κάθε πηγή-ιστοσελίδα από την οποία θα αντληθούν τα δεδομένα. Πολλές ιστοσελίδες προσφέρουν έτοιμα εργαλεία ή API για ευκολία στην αναζήτηση των ευπαθειών στην βάση δεδομένων τους.

Η εύρεση πληροφοριών σχετικά με το Dark-web δεν είναι εύκολη υπόθεση, έτσι θα χρειαστούμε ένα εργαλείο για να επιθεωρήσουμε πολλές αγορές κάθε φορά και αυτό το εργαλείο θα ήταν ένας crawler. Ο crawler είναι ένα πρόγραμμα που επιθεωρεί τα περιεχόμενα ενός δεδομένου ιστοσελίδα και εντοπίζει προκαθορισμένα δεδομένα. Ωστόσο, το πρόγραμμα πρέπει να γνωρίζει τον ιστότοπο που σκανάρει και να βρίσκει δεδομένα με βάση τη δομή του η ιστοσελίδα. Έτσι, ο crawler πρέπει να είναι μοναδικός και ο καθένας ο ιστότοπος θα πρέπει να έχει το δικό του script (Shiaeles, Kolokotronis, & Bellini, 2019).



Διάγραμμα 3: Τοπικό δίκτυο

Το εργαλείο ελέγχου θα μπορεί διενεργήσει ελέγχους:

- για τις ανοικτές θύρες σε περιμετρικές συσκευές του τοπικού δικτύου.
- θα αναγνωρίσει τις συγκεκριμένες εκδόσεις των υπηρεσιών.
- θα εντοπίσει έγκυρες ή μη υποστηριζόμενες εκδόσεις λειτουργικών συστημάτων.
- θα ανακαλύψει τις εφαρμογές και υπηρεσίες του τοπικού δικτύου.

Η αναζήτηση των υπηρεσιών θα ανιχνεύσει:

- τις γνωστές θύρες (TCP/UDP) οι οποίες είναι από το 1 ως και 1023. Οι πόρτες αυτές χρησιμοποιούνται κυρίως από γνωστές διεργασίες του συστήματος ή από εφαρμογές όπως ο web server (TCP 80), το SSH (TCP 22). Κάθε πόρτα χρησιμοποιεί συγκεκριμένο πρωτόκολλο επικοινωνίας και χρησιμοποιείται για την ανταλλαγή πληροφοριών.
- τις γνωστές καταγεγραμμένες πόρτες (Registered Ports) των γνωστών εφαρμογών από 1024 ως και 49151. Οι καταγεγραμμένες πόρτες παρέχονται στις συνηθισμένες διεργασίες και προγράμματα. Οι πόρτες αυτές, είναι καταγεγραμμένες επίσημα από την IANA (IANA.org, n.d.).
- τις δυναμικές πόρτες (Dynamic Ports), από 49152 ως και 65535. Η χρήση των δυναμικών θυρών γίνεται κυρίως από εφαρμογές που δεν είναι γνωστές στο ευρύ κοινό.

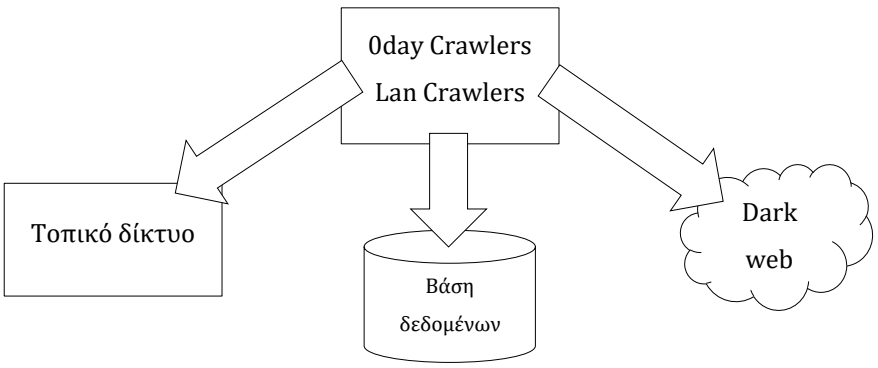
Υπάρχουν πολλές περιπτώσεις στις οποίες οι διαχειριστές αλλάζουν τις συνηθισμένες πόρτες όπως για παράδειγμα την πρόσβαση σε ένα διακομιστή ssh (TCP 22) σε κάποια άλλη πόρτα (οποιαδήποτε πάνω από 1024). Με αυτόν το τρόπο δυσκολεύουν το έργο ενός επίδοξου εισβολέα αφού ο εισβολέας θα πρέπει να ανακαλύψει σε ποια πόρτα από το σύνολο των 64511 που είναι διαθέσιμες βρίσκεται αυτή η υπηρεσία. Η εύρεση της πόρτας γίνεται ακόμα πιο δύσκολη αν το σύστημα προστατεύεται από κάποιο IDS το οποίο ανιχνεύει το scanning του εισβολέα. Σε περίπτωση ανίχνευσης εισβολέα τότε το IDS αυτόματα μπλοκάρει την κίνηση από την διεύθυνση IP για κάποιες ώρες ή ημέρες. Η ανίχνευση είναι μια σχετικά απλή διαδικασία η οποία παρακολουθεί τα TCP/UDP πακέτα που φθάνουν στην κάρτα δικτύου. Αν υπάρχει κάποια συνέχεια σε διάφορες πόρτες τότε αυτό θεωρείται μία απειλή και μπλοκάρεται αυτόματα.

3.1 Ανάλυση του νέου συστήματος

Το σύστημα είναι μία αυτοματοποιημένη υπηρεσία αντιμετώπισης απειλών με χρήση δεδομένων από διάφορες πηγές όπως το διαδίκτυο και το Dark-web. Ο σχεδιασμός περιλαμβάνει την αυτόματη αναζήτηση υπηρεσιών στο τοπικό δίκτυο μέσω crawler. Οι υπηρεσίες που ανιχνεύονται είναι οι ανιχνεύσιμες πόρτες (tcp/udp) των host του τοπικού δικτύου ανεξάρτητα από την χρήση τους. Ταυτόχρονα με την αναζήτηση των υπηρεσιών στο τοπικό δίκτυο γίνεται αναζήτηση για νέες ευπάθειες στις διάφορες πηγές. Σε ένα διάστημα 4 εβδομάδων μπορούν να εντοπιστούν τουλάχιστον 16 ευπάθειες μηδενικής ημέρας, αναζητώντας πληροφορίες στα marketplaces του dark-web (Nunes, et al., 2016).

Σε κάθε νέα ευπάθεια που προκύπτει από την αναζήτηση, γίνεται έξυπνη αναζήτηση με τις υπηρεσίες που έχουν ανιχνευθεί και καταγραφεί στο τοπικό δίκτυο. Αν η ευπάθεια που βρέθηκε αφορά κάποια ή κάποιες από τις υπηρεσίες που ανιχνεύτηκαν στο τοπικό δίκτυο, τότε καταγράφεται το γεγονός στην βάση δεδομένων.

Το τοπικό δίκτυο που πρόκειται να προστατευτεί αποτελείται από άγνωστο αριθμό συσκευών, δρομολογητών, εξυπηρετητών κ.α. Το σύστημα (0day crawlers) ελέγχει το τοπικό δίκτυο για τις διαθέσιμες υπηρεσίες, οποιοσδήποτε και αν είναι αυτές, και στην συνέχεια αναζητεί στο dark-web για νέες ευπάθειες σχετικά με τις υπηρεσίες που βρέθηκαν (Διάγραμμα 4).



Διάγραμμα 4: Σχεδίαση

3.2 Εγκατάσταση λογισμικού

Για την εγκατάσταση του συστήματος θα χρησιμοποιηθεί το λειτουργικό σύστημα Ubuntu server (https://ubuntu.com/). Για την διαχείριση και παρουσίαση των δεδομένων θα χρησιμοποιηθεί το Yii Framework (https://www.yiiframework.com). Το σύστημα μπορεί να εγκατασταθεί σε αυτόνομο μηχάνημα ή σε εικονικό μηχάνημα (virtual machine).

Μετά την εγκατάσταση του λειτουργικού συστήματος Ubuntu server χρειάζεται να εγκατασταθούν τα ακόλουθα πακέτα τα οποία είναι απαραίτητα για την λειτουργία του συστήματος. Η εγκατάσταση των πακέτων γίνεται με την εντολή apt-get install <όνομα πακέτου>.
Αναλυτικές οδηγίες εγκατάστασης του λειτουργικού συστήματος Ubuntu server είναι διαθέσιμες από την επίσημη σελίδα του κατασκευαστή.
https://tutorials.ubuntu.com/tutorial/tutorial-install-ubuntu-server#0.

```
apt-get install nmap apache2 libapache2-mod-php php7.2-cli phpunit php7.2
mbstring mysql-server nbtscan phpmyadmin unzip python python-pip lynx

pip2 install stem
```

Τα παρακάτω πακέτα είναι απαραίτητα για τον crawler του 0day.today (3.5.6 Crawler 0day.today).

```
pip install requests unidecode bs4 lxml
```

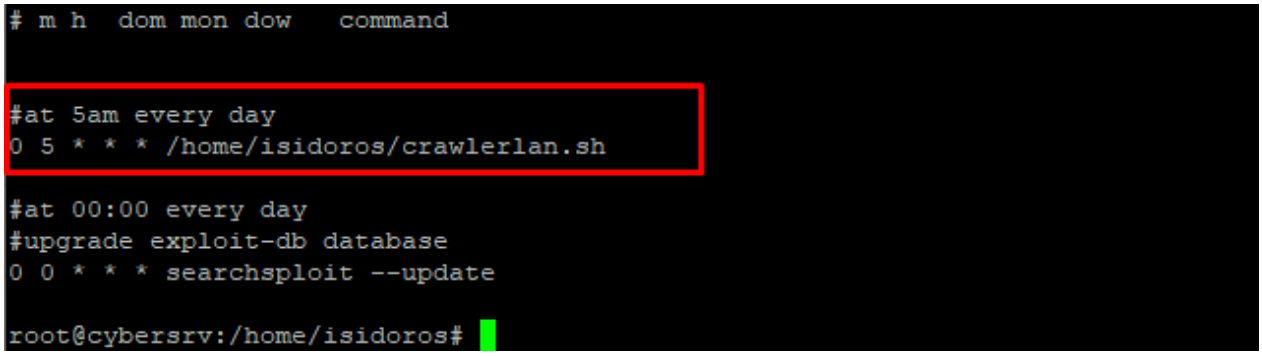
Εγκαθίσταται επίσης το πακέτο exploitdb από το https://github.com/offensive-security/exploitdb το οποίο θα χρησιμοποιηθεί για την αναζήτηση ευπαθείων από την ιστοσελίδα (https://www.exploit-db.com).

```
git clone https://github.com/offensive-security/exploitdb.git /opt/exploitdb

sed 's|path_array+=(.*)|path_array+=("/opt/exploitdb")|g'
/opt/exploitdb/.searchsploit_rc > ~/.searchsploit_rc


sudo ln -sf /opt/exploitdb/searchsploit /usr/local/bin/searchsploit
```

Οι crawlers που παρουσιάζονται παρακάτω εκτελούνται αυτόματα μέσω cronjobs. Τα cronjobs είναι ένα εργαλείο χρονικού προγραμματισμού για την εκτέλεση των crawlers. Ο προγραμματισμός μπορεί να είναι εβδομαδιαίος, μηνιαίος, ωριαίος ή ακόμα και σε συγκεκριμένες ημέρες του μήνα ή της εβδομάδας. Στο σύστημα επιλέχθηκε ο χρονικός προγραμματισμός των λειτουργιών κάθε μέρα στις 5πμ.



Εικόνα 2: Cronjobs

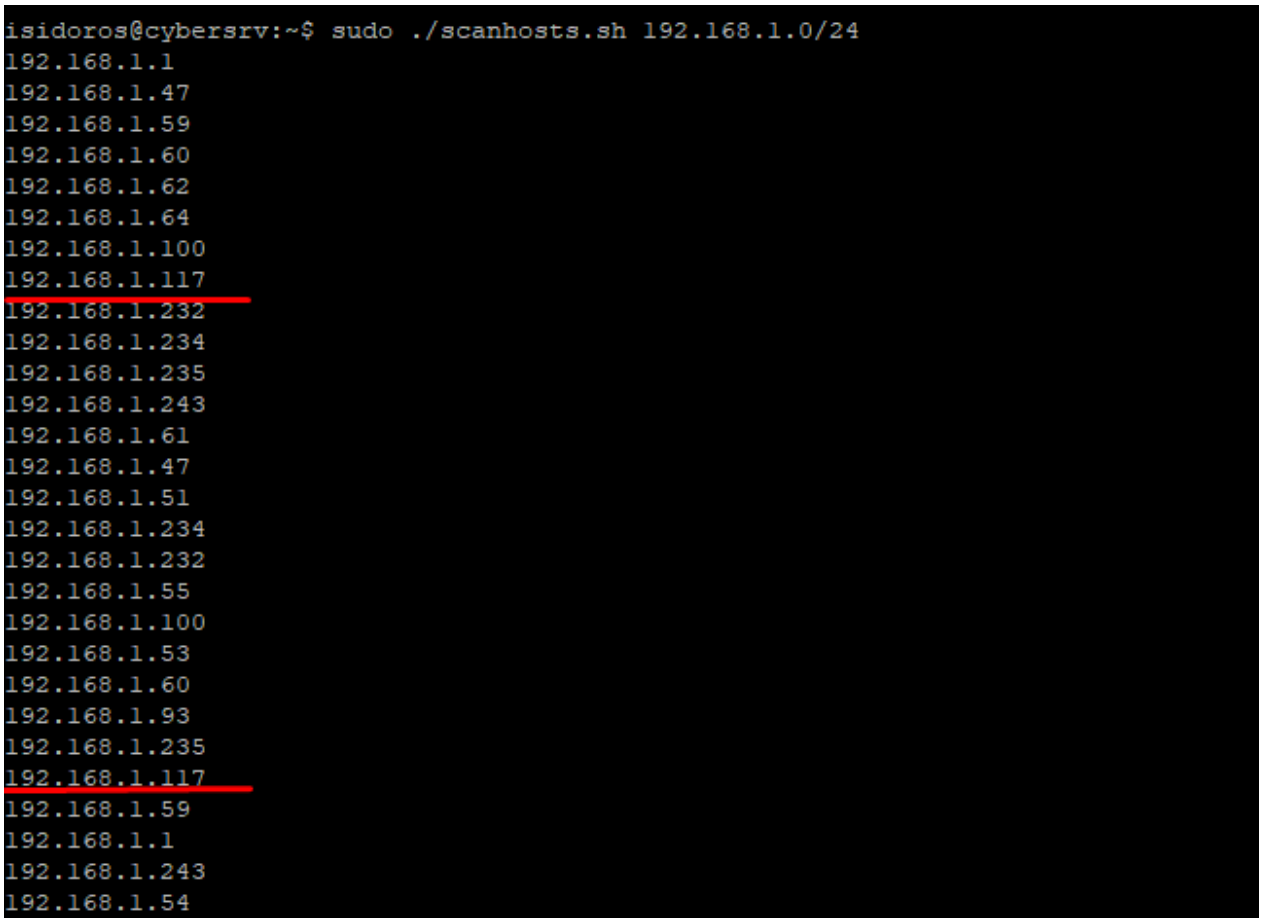
Επίσης, προγραμματίστηκε η αυτόματη ενημέρωση της τοπικής βάσης δεδομένων του εργαλείου searchsploit κάθε μέρα στις 12πμ. Το βήμα αυτό απαραίτητο καθώς με αυτόν το τρόπο θα ενημερώνεται αυτόματα η τοπική βάση δεδομένων του εργαλείου.

3.3 Αναζήτηση hosts στο τοπικό δίκτυο

Για τον έλεγχο του τοπικού δικτύου χρειάζεται μία χαρτογράφηση με το εργαλείο nmap, το οποίο αναβαθμίζεται και επεκτείνεται συνεχώς. Λαμβάνοντας υπόψη τη γενικότερη αποδοχή του εργαλείου επιλέχτηκε ως βασικό εργαλείο για την ανάπτυξη του συστήματος.
Με την εντολή nmap γίνεται αναζήτηση των συσκευών που είναι συνδεδεμένες στο τοπικό δικτύο. Το nmap είναι ένα ελεύθερο και ανοικτού κώδικα εργαλείο για την εξερεύνηση του δικτύου. Λειτουργεί ως σαρωτής και χρησιμοποιείται για να ανακαλύψει υπολογιστές και υπηρεσίες σε ένα δίκτυο υπολογιστών.
Πολλές συσκευές μπορεί να φιλτράρουν με κάποιο τείχος προστασίας ή άλλο λογισμικό ασφάλειας τα ICMP requests, με αποτέλεσμα ένα απλό ping να μην μπορεί να τις εντοπίσει. Συμπληρωματικά με την έρευνα του δικτύου με το nmap θα χρησιμοποιηθεί η εντολή arp που έχει πρόσβαση στο πρωτόκολλο arp. Το πρωτόκολλο arp κάθε φορά που πρέπει να γίνει γνωστή η διεύθυνση MAC που αντιστοιχεί σε κάποια συγκεκριμένη διεύθυνση IP, λαμβάνει χώρα εκπομπή σε όλους τους υπολογιστές (broadcasting) ενός πακέτου δεδομένων που περιέχει τη διεύθυνση IP που θέλουμε να μεταφράσουμε. Ο κάθε ένας από τους υπολογιστές του δικτύου, παραλαμβάνει αυτό το πακέτο, συγκρίνει τη διεύθυνση IP που περιέχει, με τη δική του διεύθυνση IP και εάν οι δύο διευθύνσεις είναι οι ίδιες, αποστέλλει μια απάντηση στον υπολογιστή που υπέβαλλε το ερώτημα. Η απάντηση αυτή περιέχει τη MAC διεύθυνση του υπολογιστή αποστολέα η οποία ταυτοποιείται, απομονώνεται και αποθηκεύεται σε μια ειδική μνήμη ARP cache, έτσι ώστε να μπορεί να χρησιμοποιηθεί στο μέλλον. Η μνήμη αυτή ανανεώνεται σε τακτά χρονικά διαστήματα, διότι τα περιεχόμενα της μπορούν σε κάποια χρονική στιγμή να μεταβληθούν, όπως συμβαίνει για παράδειγμα σε περιπτώσεις κατά τις οποίες αντικαθιστούμε την κάρτα δικτύου του υπολογιστή με κάποια άλλη η οποία έχει τη δική της MAC address.
Θα δημιουργηθεί το παρακάτω script **scanhosts.sh** το οποί συνοψίζει τις δυο λειτουργίες (nmap και arp).

```
#!/bin/bash
nmap $1 -n -sP | grep report | awk '{print $5}'
arp -a -n | awk '{print $2}'| tr -d '()'
```

Το αποτέλεσμα του script εμφανίζεται στην παρακάτω Εικόνα 3.



Εικόνα 3: IP List

Οι διευθύνσεις εμφανίζονται διπλές φορές αφού το script εμφανίζει το συγκεντρωτικό αποτέλεσμα των εντολών nmap και arp. Συνεπώς ενδέχεται να υπάρχουν οι ίδιες διευθύνσεις ως αποτέλεσμα των παραπάνω εντολών (εξαιρετικά πιθανό). Για να αποφύγουμε διπλές εργασίες στη συνέχεια τα αποτελέσματα θα φιλτράρονται με την εντολή sort, uniq και sponge. Το τελικό αποτέλεσμα είναι ένα αρχείο (/home/isidoros/hosts.txt) το οποίο περιέχει μία IP διεύθυνση για κάθε host.

```
cat hosts.txt |sort|uniq|sponge hosts.txt
```

Στη συνέχεια το αρχείο hosts.txt που δημιουργείται θα ελεγχθεί από το monitoring script (3.6 Monitoring) το οποίο θα ελέγξει την κάθε IP διεύθυνση που εντοπίστηκε για υπηρεσίες που είναι εκτεθειμένες στο δίκτυο.

Οι παραπάνω ενέργειες πραγματοποιούνται αυτόματα από τον crawler ο οποίος εκτελεί όλες τις απαραίτητες ενέργειες στο τοπικό δίκτυο για τις ενεργές IP διευθύνσεις. Αυτή η έρευνα πραγματοποιείται ανά τακτά χρονικά διαστήματα (μέσω cronjobs) και τα αποτελέσματα καταγράφονται στην τοπική βάση δεδομένων.

3.4 Αναζήτηση υπηρεσιών στο τοπικό δίκτυο.

Όταν πραγματοποιείται αναζήτηση των υπηρεσιών σε ένα δίκτυο και σε κάθε υπολογιστή αντίστοιχα, είναι εξαιρετικά πιθανό η ανίχνευση που πραγματοποιούμε να εντοπιστεί από το λογισμικό του εξυπηρετητή και στη συνέχεια να μπλοκαριστεί κάθε δικτυακή κίνηση από και προς την IP του συστήματος μας. Μία έρευνα προτείνει την αναζήτηση των πορτών (port scanning) σε μία αρχικά τυχαία αναζήτηση και στη συνέχεια, ανάλογα τα αποτελέσματα, να γίνεται στοχευμένη αναζήτηση στις υπόλοιπες πόρτες με κάποια χρόνο-καθυστέρηση (CHEN & CHENG, 2009).

Για την αναζήτηση των ανοικτών πορτών σε κάποιον υπολογιστή θα χρησιμοποιηθεί το πακέτο nmap. Μία τέτοια αναζήτηση σε ένα μεγάλο δίκτυο Class B ή Class A θα απαιτούσε πάρα πολύ χρόνο και εύρος συνδεσιμότητας (bandwidth) (Shah, Ahmed, Saeed, Junaid, & Khan, 2019). Στο τοπικό δίκτυο (Διάγραμμα 3) έχουμε ένα δίκτυο τύπου Class C 192.168.1.0/24 με μέγιστο αριθμό hosts 254.

Για κάθε IP διεύθυνση που εντοπίζεται αυτόματα από την αναζήτηση στο τοπικό δίκτυο μπορεί να γίνει η αναζήτηση του λειτουργικού συστήματος με την εντολή:

```
nmap -p 22,80,445,65123,56123 -O 192.168.1.60
```

Τα αποτελέσματα της παραπάνω εντολής εμφανίζονται στην Εικόνα 4. Στη συνέχεια η παραπάνω εντολή θα αυτοματοποιηθεί ώστε να γίνεται αυτόματη αναζήτηση των υπηρεσιών της κάθε IP διεύθυνσης. Από την παραπάνω εντολή nmap χρειαζόμαστε τα στοιχεία του λειτουργικού συστήματος.

```
isidoros@cybersrv:~$ sudo  nmap -Pn -p 22,80,445,65123,56123 -O 192.168.1.60
sudo  nmap -Pn -p 22,80,445,65123,56123 -O 192.168.1.60

Starting Nmap 7.60 ( https://nmap.org ) at 2019-11-18 12:26 UTC
Nmap scan report for 192.168.1.60
Host is up (-0.030s latency).

PORT      STATE  SERVICE
22/tcp    open   ssh
80/tcp    closed http
445/tcp   closed microsoft-ds
56123/tcp closed unknown
65123/tcp closed unknown
MAC Address: 00:1A:79:3F:E5:99 (Telecomunication Technologies)
Device type: general purpose
Running: Linux 2.6.X
OS CPE: cpe:/o:linux:linux_kernel:2.6
OS details: Linux 2.6.32 - 2.6.35
Network Distance: 1 hop

OS detection performed. Please report any incorrect results at https://nmap.org/submit/ .
Nmap done: 1 IP address (1 host up) scanned in 3.91 seconds
isidoros@cybersrv:~$
```

Εικόνα 4: OS detection

Για την αυτοματοποίηση και ευκολία της παραπάνω διαδικασίας δημιουργήθηκε το script findos.sh (Εικόνα 5) το οποίο δέχεται ως παράμετρο μία IP διεύθυνση και επιστρέφει την έκδοση του λειτουργικού συστήματος (Εικόνα 6).

```
isidoros@cybersrv:~$ cat findos.sh
#!/bin/bash
#
# Isidoros Moulas
# imoulas@hotmail.com
# 2019
#


datalist=$(sudo nmap -Pn -p 22,80,445,65123,56123 -O $1|grep 'OS details\|MAC Address\|OS CPE')

while read -r line
do
   echo $line
done <<< "$datalist"


isidoros@cybersrv:~$
```

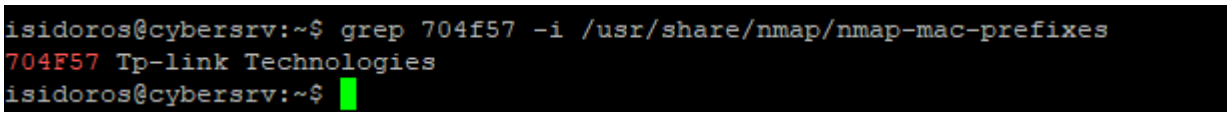
Εικόνα 5: findos.sh

```
isidoros@cybersrv:~$ sudo ./findos.sh 192.168.1.62
OS CPE: cpe:/o:linux:linux_kernel:3
OS details: Linux 3.7 - 3.10
isidoros@cybersrv:~$
```

Εικόνα 6: findos.sh output

Στη συνέχεια μπορούμε να ανακαλύψουμε τον κατασκευαστή της συσκευής με την χρήση του πακέτου nmap-mac-prefixes. Ο εύρεση του κατασκευαστή θα δώσει περισσότερες πληροφορίες για το είδος της συσκευής (υπολογιστής, δρομολογητής, access point κτλ). Η διεύθυνση MAC address (για παράδειγμα 70:4F:57:30:50:56) περιέχει στα τα 3 πρώτα octets (70:4F:57) το μοναδικό κωδικό που χαρακτηρίζει τον κατασκευαστή.
Για παράδειγμα με την παρακάτω εντολή θα πραγματοποιηθεί η αναζήτηση του κατασκευαστή της συγκεκριμένης συσκευής.

```
grep 704f57 -i /usr/share/nmap/nmap-mac-prefixes
```



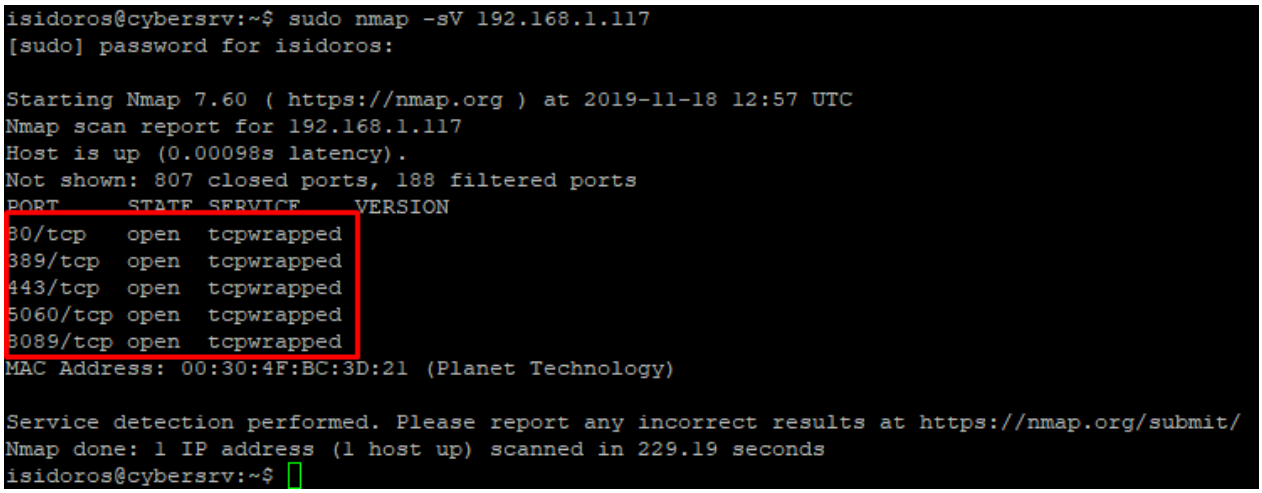
Εικόνα 7: Vendor detection

Η αναζήτηση του κατασκευαστή θα πραγματοποιηθεί από το διαχειριστικό σύστημα (php) το οποίο θα εκτελέσει την παραπάνω εντολή και στη συνέχεια θα αποθηκεύσει τα αποτελέσματα στην βάση δεδομένων.

Στη συνέχεια θα πραγματοποιηθεί αναζήτηση των διαθέσιμων υπηρεσιών/πόρτες για την κάθε IP διεύθυνση με την χρήση της εντολής nmap -sV.

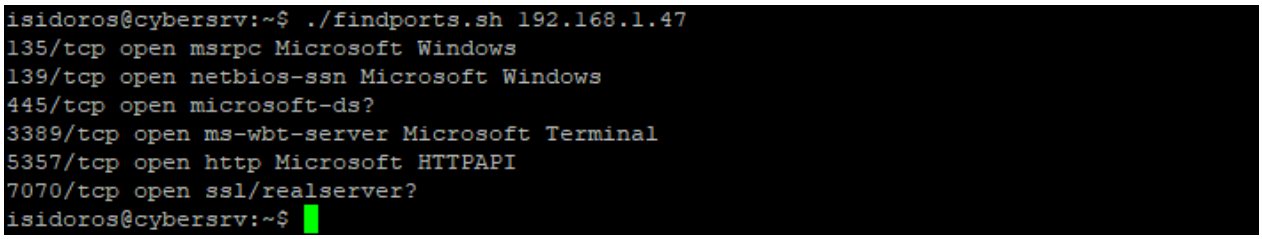
```
nmap -sV 192.168.1.117
```

Τα αποτελέσμα της εντολής εμφανίζονται στην παρακάτω Εικόνα 8.



Εικόνα 8: Open ports nmap

Για την αυτοματοποιημένη αναζήτηση του των υπηρεσιών δημιουργήθηκε το script **findports.sh** (Εικόνα 9). Στο script γίνεται φιλτράρισμα των πορτών σε tcp και udp.



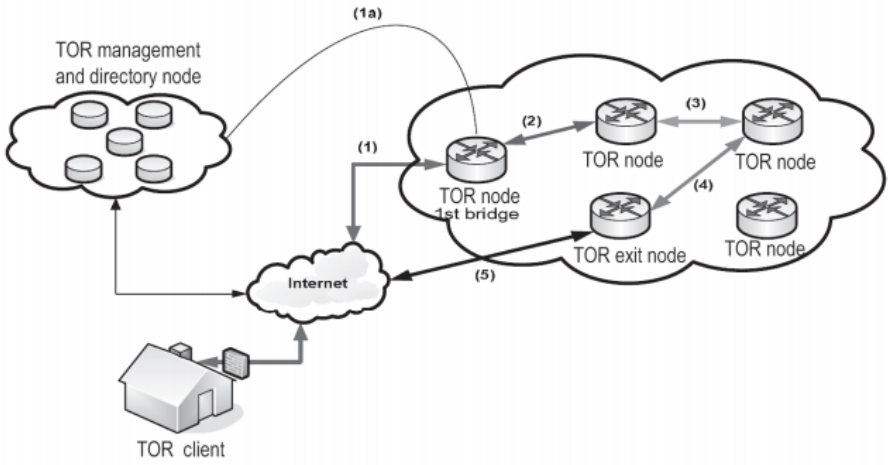
Εικόνα 9: Open ports script

3.5 Crawler του dark-web

Η πρόσβαση στο dark-web απαιτεί ταυτόχρονα πρόσβαση στο διαδίκτυο καθώς το dark-web είναι μέρος του διαδικτύου. Για να μπορέσουμε να έχουμε πρόσβαση στις ιστοσελίδες με επέκταση .onion απαιτείται η χρήση του δικτύου tor. Το Tor είναι δωρεάν λογισμικό ανοικτού κώδικα για την ενεργοποίηση της ανώνυμης επικοινωνίας. Το όνομα προέρχεται από ένα ακρωνύμιο για το αρχικό όνομα του έργου λογισμικού «The Router Onion». Ο Tor κατευθύνει την κυκλοφορία στο Διαδίκτυο μέσω ενός δωρεάν, παγκόσμιου, εθελοντικού δικτύου επικάλυψης που αποτελείται από περισσότερα από επτά χιλιάδες υπολογιστές-διακομιστές για να αποκρύψει την τοποθεσία και τη χρήση του χρήστη από οποιονδήποτε πραγματοποιεί επιτήρηση δικτύου ή ανάλυση κυκλοφορίας.

3.5.1 Πρόσβαση στο dark-web

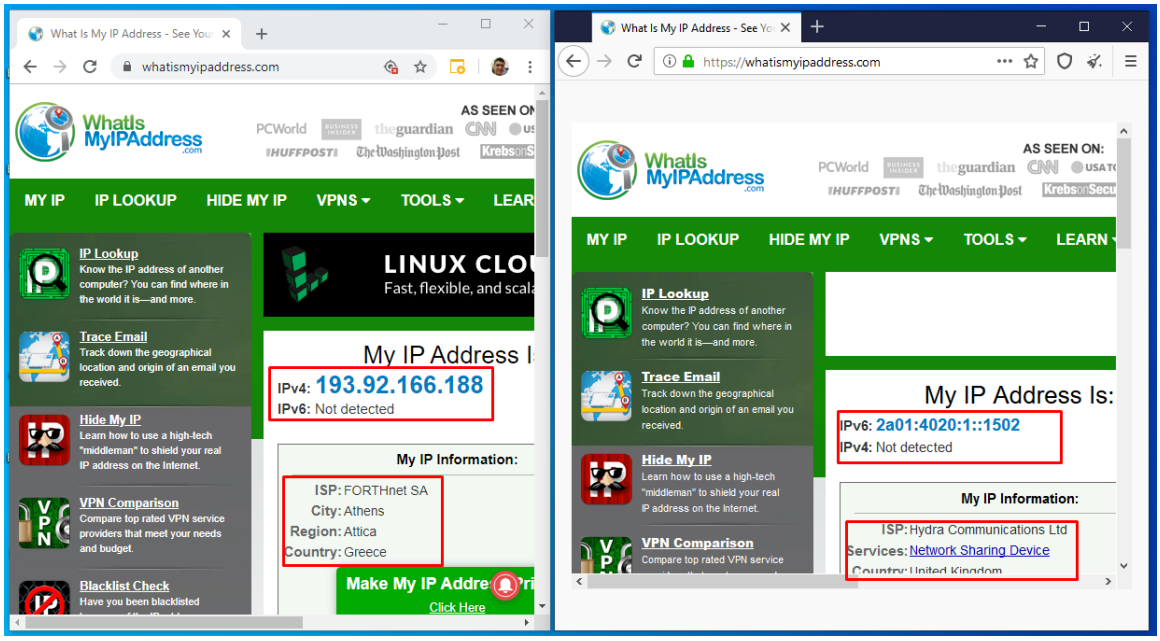
Όταν προσπαθούμε να συνδεθούμε με το δίκτυο tor, με κάποιον από τους παρακάτω τρόπους, η εφαρμογή tor που χρησιμοποιούμαι προσπαθεί να εντοπίσει έναν κόμβο TOR. Αυτός ο κόμβος, ο οποίος λέγεται και γέφυρα (bridge) είναι ένα κόμβος που δέχεται συνδέσεις και είναι ένας κόμβος ο οποίος υπάρχει στην λίστα των πέντε TOR κόμβων που κάνουν την διαχείριση του δικτύου από την ομάδα των διαχειριστών του δικτύου tor. Η εφαρμογή ζητάει (request) έναν από τους 5 κόμβους για έναν κόμβο-γέφυρα. Αφού η εφαρμογή συνδεθεί με τον κόμβο-γέφυρα τότε επιλέγει τυχαία από την λίστα τρείς ακόμα κόμβους TOR (Εικόνα 10: Tor network). Συνεπώς κάθε κόμβος του δικτύου tor γνωρίζει μόνο δύο κόμβους του δικτύου (Haraty & Zantout, Aug. 2014). Η επικοινωνία μεταξύ των κόμβων είναι κρυπτογραφημένη όσο αφορά τις πληροφορίες που ανταλλάσσονται σχετικά με τις διευθύνσεις και το είδος των κόμβων που συνδέεται ο κάθε κόμβος.



Εικόνα 10: Tor network

Για να μπορέσουμε να έχουμε πρόσβαση στις ιστοσελίδες με επέκταση .onion απαιτείται η χρήση κάποιου λογισμικού όπως ο torBrowser (https://www.torproject.org/). Με την χρήση του torBrowser έχουμε πρόσβαση σε όλα τα .onion web sites. Ουσιαστικά ο tor browser είναι ένας firefox web browser με κάποιες απαραίτητες επεκτάσεις ώστε να δρομολογείται η κίνηση που γίνεται μέσα από το περιβάλλον του browser στο dark-web. Πρακτικά αυτό σημαίνει ότι η μόνο η κίνηση η οποία γίνεται μέσα στον tor browser, δηλαδή οι ιστοσελίδες .onion, έχουν πρόσβαση στο dark-web ενώ η υπόλοιπη κίνηση του υπολογιστή δρομολογείται στο διαδίκτυο (Εικόνα 11).

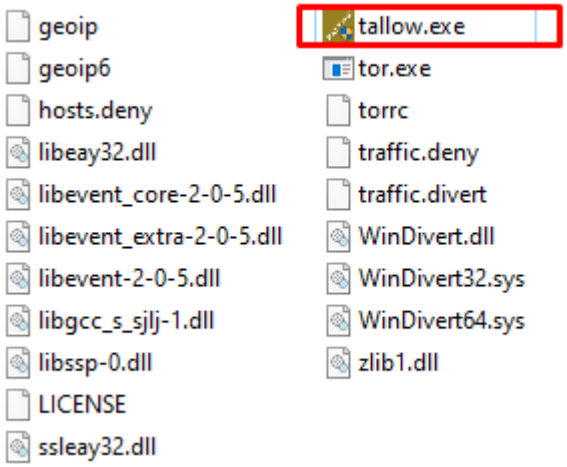
Chrome browser Tor browser



Εικόνα 11: Tor browser

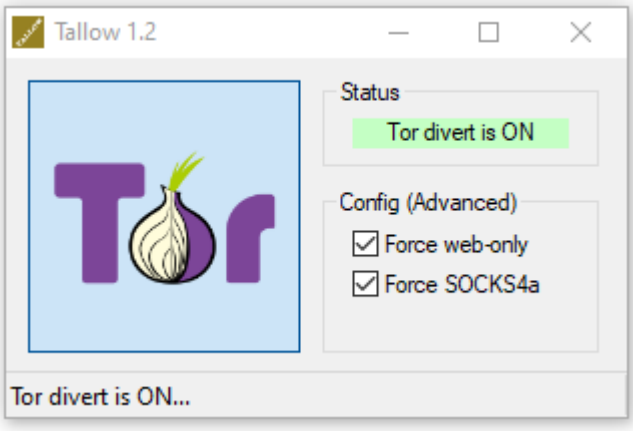
Μία εναλλακτική επιλογή για την πρόσβαση στο dark-web είναι το **tallow για Windows** (https://reqrypt.org/tallow.html). Το tallow δρομολογεί ολόκληρη την κίνηση του υπολογιστή στο dark-web χωρίς να είναι απαραίτητη η χρήση του tor browser. Συγκεκριμένα, όταν ενεργοποιηθεί το tallow, όλη η κίνηση του υπολογιστή με λειτουργικό σύστημα windows δρομολογείται μέσα από το dark-web με αποτέλεσμα και προγράμματα που έχουμε κατασκευάσει να έχουν πρόσβαση στο dark-web απ' ευθείας μέσα από το λειτουργικό σύστημα Windows.

Εκτελώντας το πρόγραμμα tallow.exe (Εικόνα 12: Tallow) το οποίο θα φορτώσει τις απαραίτητες βιβλιοθήκες για την πρόσβαση στο dark-web θα έχουμε την δυνατότητα να συνδεθούμε στο δίκτυο tor. Αφού τελειώσει η φόρτωση των βιβλιοθηκών (υπάρχει σχετικό status bar) μπορούμε να πατήσουμε το εικονίδιο-κουμπί του tor για να ενεργοποιηθεί η σύνδεση με το δίκτυο tor.



Εικόνα 12: Tallow folder

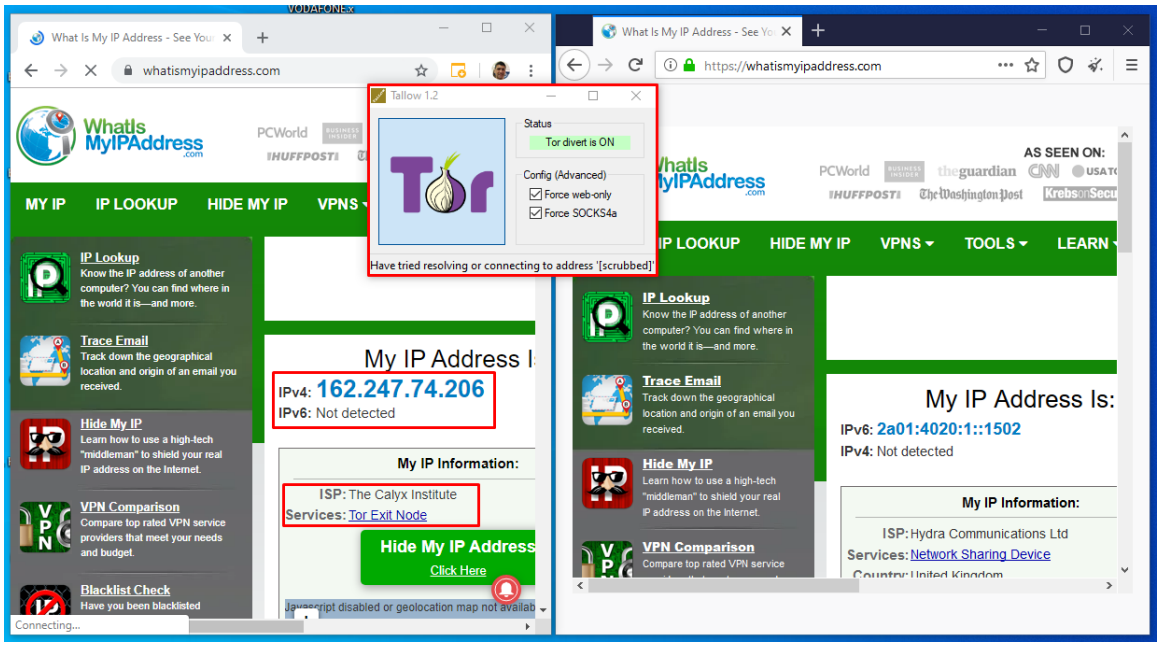
Στη συνέχεια εκτελούμε το αρχείο tallow.exe (Εικόνα 13) και πατάμε το κουμπί Tor το οποίο θα ενεργοποιήσει το σύνδεση του υπολογιστή με το δίκτυο tor.



Εικόνα 13: Tallow launch

Στη συνέχεια οποιαδήποτε πρόσβαση στο διαδίκτυο από τον συγκεκριμένο ηλεκτρονικό υπολογιστή θα προωθείται στο dark-web χωρίς την χρήση του tor browser (Εικόνα 14).
Στην Εικόνα 14 βλέπουμε ότι μετά την ενεργοποίηση του tallow έχουμε πρόσβαση στο δίκτυο tor από τον browser chrome (στο αριστερό πλαίσιο) αλλά και από τον tor browser (στο δεξί πλαίσιο). Ουσιαστικά έτσι όπως φαίνεται στο στιγμιότυπο της οθόνης η συνολική κίνηση του υπολογιστή γίνεται από την διεύθυνση 162.247.74.206 η οποία είναι ένα tor exit node και η κίνηση που γίνεται από τον tor browser γίνεται από την ipv6 διεύθυνση 2a01:4020:1::1502.

Chrome browser Tor browser



Εικόνα 14: Tor via tallow

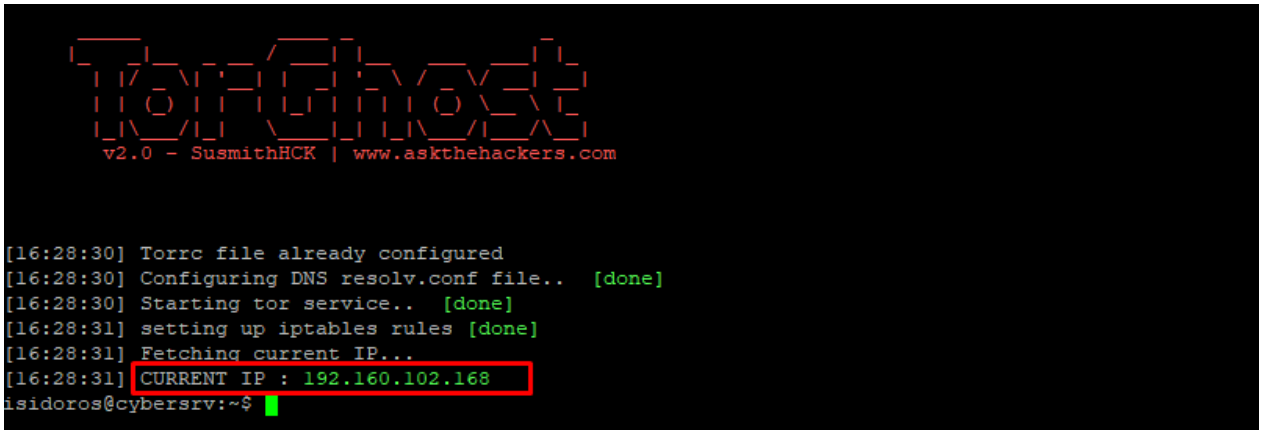
Ο στόχος του crawler που κατασκευάστηκε είναι να μπορεί να αναζητά ευπάθειες μέσα στο dark-web χωρίς την αλληλεπίδραση του χρήστη σε ένα αυτοματοποιημένο περιβάλλον λειτουργίας. Για την κατασκευή επιλέχθηκε το λειτουργικό σύστημα ubuntu server, το οποίο δεν διαθέτει γραφικό περιβάλλον GUI όπως τα Windows. Για την πρόσβαση στο dark-web μπορεί να χρησιμοποιηθεί τα λογισμικά privoxy και torshost. Ο privoxy είναι ένας HTTP proxy ο οποίος ο οποίος με τις κατάλληλες ρυθμίσεις μπορεί να συνδεθεί στο δίκτυο tor (Zulkarnine, Frank, Mitchell, & Davies, 2016). Για την παρούσα έρευνα θα χρησιμοποιηθεί το πακέτο torghost το οποίο έχει την ίδια λειτουργία με τον privoxy για ubuntu αλλά και του tallow για windows χωρίς να είναι απαραίτητο ο χρήστης να πατήσει κάποιο κουμπί για να ενεργοποιηθεί το δίκτυο tor.

Από την στιγμή που θα ενεργοποιηθεί το torghost τότε όλη η κίνηση του λειτουργικού συστήματος θα δρομολογείται μέσω του dark-web χωρίς την ανάγκη παρέμβασης του χρήστη όπως γίνεται στο tallow (χρειάζεται το πάτημα ενός κουμπιού) ή στον tor browser (χρειάζεται να πληκτρολογηθεί το .onion url). Το σύστημα θα μπορεί να πλοηγείται αθόρυβα, χωρίς την αλληλεπίδραση του χρήστη, στο dark-web και να αναζητά για ευπάθειες που έχουν ορισθεί μέσω εξειδικευμένων crawlers στις ιστοσελίδες ελέγχου.

Όταν το torghost είναι ενεργοποιημένο τότε ολόκληρη η κίνηση του λειτουργικού συστήματος δρομολογείται μέσω του δικτύου tor. Αμέσως μετά την απενεργοποίηση του torghost η κίνηση επιστρέφει στην αρχική κατάσταση και δρομολογείται από το συνήθες διαδίκτυο. Με αυτόν τον τρόπο ο crawler που κατασκευάστηκε, συνδεόταν στο dark-web όταν χρειαζόταν και ταυτόχρονα είχε πρόσβαση στο ανοικτό διαδίκτυο ανάλογα την κατάσταση που βρισκόταν.

Το πρόγραμμα torGhost (https://github.com/susmithHCK/torghost) είναι λογισμικό ανοικτού κώδικα και χρησιμοποιείται για πρόσβαση στο δίκτυο dark. Για την εγκατάσταση του torGhost θα χρειαστεί η εγκατάσταση του πακέτου unzip και της python 2.7 (apt install unzip python) για να μπορέσουν να αποσυμπιεστούν τα αρχεία του torghost αλλά και να μπορεί να τρέξει η εφαρμογή torGhost. Για την σωστή εκτέλεση θα πρέπει να εγκατασταθούν και τα πακέτα python-pip (apt install python-pip) και στη συνέχεια το πακέτο stem (pip2 install stem).

Εκτελώντας την εντολή `sudo torghost start` δημιουργείται ένα νέο κανάλι με το dark-web σε μία ανώνυμη διεύθυνση του δικτύου (Εικόνα 15: TorGhost).



Εικόνα 15: TorGhost

Με την παρακάτω εντολή μπορούμε να δούμε ποια είναι η δημόσια διεύθυνση του server πριν και μετά την εκτέλεση του torghost.

```
curl -s http://whatismijnip.nl |cut -d " " -f 5
```

Για επαλήθευση της σύνδεσης θα εκτελέσουμε μία αναζήτηση με την χρήση της εντολής dig για την διεύθυνση http://mvfjfugdwgc5uwho.onion η οποία είναι μία από τις ιστοσελίδες

που θα χειριστούμε στο σύστημα. Παρατηρούμε ότι το dns service απαντάει με μία local διεύθυνση class A στην οποία βρίσκεται το .onion website (Εικόνα 16: Dig onion).

```
isidoros@cybersrv:~$ dig http://mvfjfugdwgc5uwho.onion

; <<>> DiG 9.11.3-1ubuntu1.10-Ubuntu <<>> http://mvfjfugdwgc5uwho.onion
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 33394
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 0, ADDITIONAL: 0

;; QUESTION SECTION:
;http://mvfjfugdwgc5uwho.onion. IN      A

;; ANSWER SECTION:
http://mvfjfugdwgc5uwho.onion. 60 IN    A       10.49.17.22

;; Query time: 0 msec
;; SERVER: 127.0.0.1#53(127.0.0.1)
;; WHEN: Mon Nov 18 16:32:38 UTC 2019
;; MSG SIZE  rcvd: 63

isidoros@cybersrv:~$
```

Εικόνα 16: Dig onion

Παρατηρούμε επίσης ότι η δρομολόγηση του συστήματος δεν έχει αλλάξει και παραμένει όπως αρχικά είχε σχεδιαστεί προκειμένου το σύστημα να έχει πρόσβαση στο διαδίκτυο (Εικόνα 17).

```
sudo route -nv
```

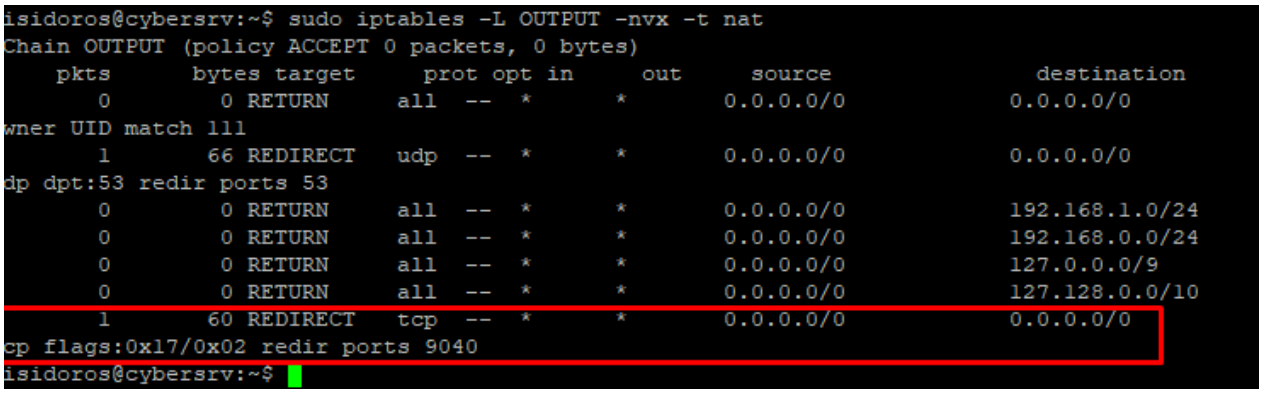
```
root@cybersrv:/home/isidoros# route -nv
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
0.0.0.0         192.168.1.1     0.0.0.0         UG    100    0        0 enp0s3
192.168.1.0     0.0.0.0         255.255.255.0   U     0      0        0 enp0s3
192.168.1.1     0.0.0.0         255.255.255.255 UH    100    0        0 enp0s3
root@cybersrv:/home/isidoros#
```

Εικόνα 17: Δρομολόγηση torghost

Βέβαια το πρόγραμμα torGhost έχει προσθέσει ειδικά services για να έχουμε πρόσβαση στο dark-web και γίνεται η αντίστοιχη ανακατεύθυνση των πακέτων στην πόρτα TCP 9040 του server (Εικόνα 17).

Εκτελούμε την εντολή:

```
sudo iptables -L OUTPUT -nvx -t nat
```



Εικόνα 18: iptables torghost

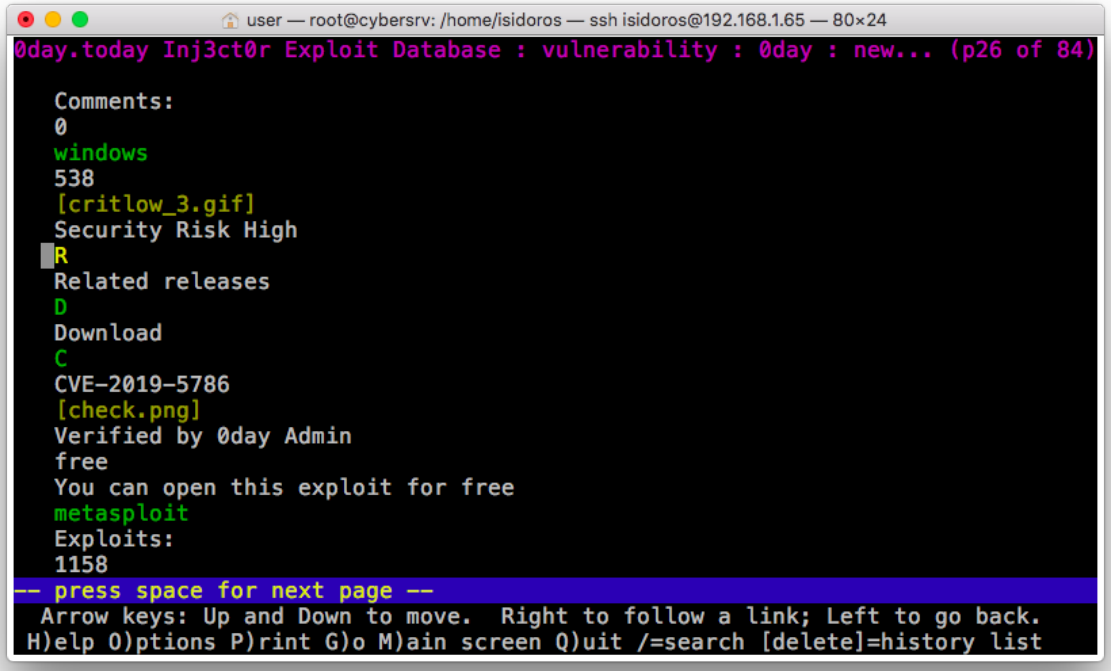
Η ανακατεύθυνση (redirect) της πόρτας UDP 53 και TCP 9040 γίνεται προς την τοπική υπηρεσία του tor.

Καθώς το λειτοργικό σύστημα που χρησιμοποιούμε δεν έχει γραφικό περιβάλλον θα εγκαταστήσουμε έναν υποτυπώδη browser στο bash terminal του ubuntu. Για την εγκατάσταση θα χρειαστούμε το πακέτο links2 (`apt install lynx`) ο οποίος μπορεί να μας δώσει βασικές λειτουργίες πλοήγησης (browsing) χωρίς γραφικά. Με τον lynx θα επιβεβαιώσουμε ότι η σύνδεση με το δίκτυο tor είναι επιτυχημένη και ότι έχουμε πρόσβαση στις ιστοσελίδες .onion.

Εκτελούμε την εντολή:

```
lynx http://mvfjfugdwgc5usho.onion
```

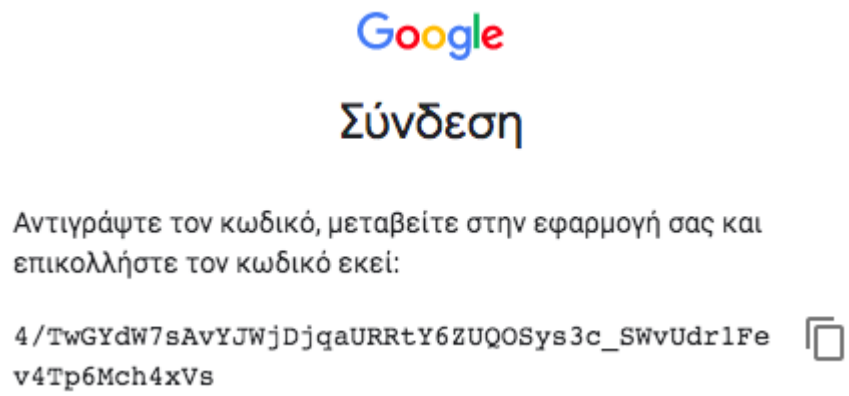
Η παραπάνω εντολή εμφανίζει την ιστοσελίδα http://mvfjfugdwgc5usho.onion χωρίς γραφικά (μόνο κείμενο) στην οθόνη του terminal (Εικόνα 19).



Εικόνα 19: lynx onion

3.5.2 Crawler 0day in the wild

Στο αρχείο "0day in the wild" το οποίο βρίσκεται στη διεύθυνση https://docs.google.com/spreadsheets/d/1lkNJ0uQwbeC1ZTRrxdtuPLCIl7mlUreoKfSIgajnSyY υπάρχουν καταγεγραμμένα όλα τα 0day exploits σε ένα αρχείο spreadsheet του google drive. Για να ανακτήσουμε το αρχείο με αυτοματοποιημένο τρόπο θα πρέπει να εγκαταστήσουμε το αρχείο gdrive (https://github.com/gdrive-org/gdrive). Στη συνέχεια θα δώσουμε δικαιώματα εκτέλεσης στο αρχείο (πχ chmod +x **gdrive-linux-x64**) και θα εκτελέσουμε το πρόγραμμα με την εντολή (./**gdrive-linux-x64**). Το πρόγραμμα θα μας εμφανίσει μία διεύθυνση που θα πρέπει να ακολουθήσουμε στον browser προκειμένου να λάβουμε το κλειδί (Εικόνα 20) για να συνεχίσει η εγκατάσταση (θα χρειαστεί σύνδεση με λογαριασμό google/gmail).



Εικόνα 20: google key

Σε περίπτωση που χρειαστεί να αλλάξουμε το κλειδί (token) αυτό βρίσκεται στο αρχείο /home/isidoros/**.gdrive**

Στη συνέχεια θα ελέγχουμε το αρχείο "0day in the wild" για την ημερομηνία που πραγματοποιήθηκε η τελευταία αλλαγή (Εικόνα 21) προκειμένου να μην γίνεται άσκοπη επεξεργασία σε περίπτωση που δεν έχουν πραγματοποιηθεί αλλαγές/προσθήκες στο αρχείο.

```
isidoros@cybersrv:~$ ./gdrive-linux-x64 info 1lkNJ0uQwbeC1ZTRrxdtuPLCIl7mlUreoKfSIgajnSyY
Id: 1lkNJ0uQwbeC1ZTRrxdtuPLCIl7mlUreoKfSIgajnSyY
Name: 0day "In the Wild"
Path: 0day "In the Wild"
Mime: application/vnd.google-apps.spreadsheet
Created: 2019-05-14 19:17:57
Modified: 2019-05-16 17:43:48
Shared: True
ViewUrl: https://docs.google.com/spreadsheets/d/1lkNJ0uQwbeC1ZTRrxdtuPLCIl7mlUreoKfSIgajnSyY/edit?usp=d
rivesdk
isidoros@cybersrv:~$
```

Εικόνα 21: gdrive info

Θα κατεβάσουμε το αρχείο σε μορφή zip.

```
isidoros@cybersrv:~$ ./gdrive-linux-x64 export --mime application/zip 1lkNJ0uQwbeC1ZTRrxdtuPLCIl7mlUreo
KfSIgajnSyY --force
Exported '0day "In the Wild".zip' with mime type: 'application/zip'
```

Εικόνα 22: gdrive export

Το αρχείο που κατέβηκε είναι σε μορφή zip και μπορεί να αποσυμπιεστεί με unzip ώστε να έχουμε πρόσβαση στα περιεχόμενα του αρχείου (sheets).

```
isidoros@cybersrv:~$ unzip -o 0day\ \"In\ the\ Wild\".zip
Archive:  0day "In the Wild".zip
  inflating: Introduction.html
  inflating: All.html
  inflating: 2019.html
  inflating: 2018.html
  inflating: 2017.html
  inflating: 2016.html
  inflating: 2015.html
  inflating: 2014.html
  inflating: resources/sheet.css
```

Εικόνα 23: unzip 0day file

Στα φύλλα εργασίας All.html, 2019.html (κτλ) βρίσκονται τα exploits ομαδοποιημένα ανά έτος.

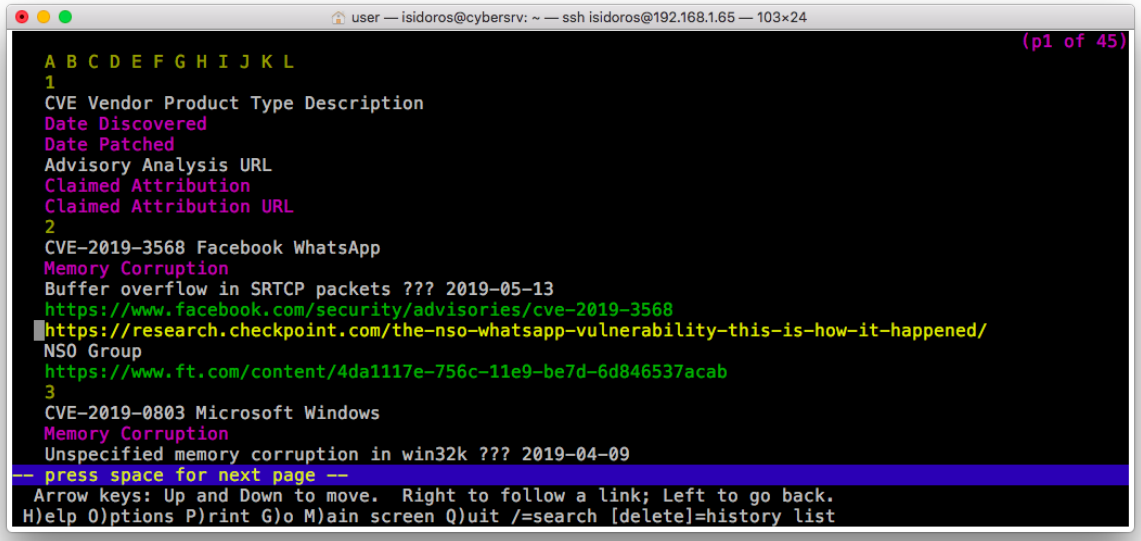
18	CVE-2018-5002	Adobe	Flash	Memory Corruption	Out-of-bounds read/write in AVM li8 opcode	???	2018-06-07	https://helpx.
19	CVE-2018-4990	Adobe	Reader	Memory Corruption	Out-of-bounds free in JPEG2000 CMAP	???	2018-05-14	https://helpx.
20	CVE-2018-8120	Microsoft	Windows	Memory Corruption	NULL pointer dereference in NtUserSetImeInfoE	???	2018-05-08	https://portal.
21	CVE-2018-8174	Microsoft	VBScript	Memory Corruption	Use-after-free in VBScriptClass::Release	???	2018-05-08	https://portal.
22	CVE-2018-4878	Adobe	Flash	Memory Corruption	Use-after-free in MediaPlayer DRM Listener	???	2018-02-06	https://helpx.

Siri roduction All 2019 2018 2017 2016 2015 2014

Εικόνα 24: 0day sheets file

Στη συνέχεια με την εντολή lynx All.html μπορούμε να δούμε το αρχείο σε μορφή html από το bash terminal.

```
lynx All.html
```



Εικόνα 25: lynx All.html

Το αρχείο **All.html** θα περάσει από το script για το σωστό parsing του αρχείου. Το script θα ελέγχει τα exploits και θα τα συγκρίνει με την τοπική βάση δεδομένων. Παρακάτω ακολουθεί ένα δείγμα του αρχείου σε html μορφή.

```
more All.html
```

```
isidoros@cybersrv:~$ more All.html
<meta http-equiv="Content-Type" content="text/html; charset=utf-8"><link type="text/css" rel="styleshee
t" href="resources/sheet.css" >
<style type="text/css">.ritz .waffle a { color: inherit; }.ritz .waffle .s0{background-color:#ffffff;te
xt-align:left;font-weight:bold;color:#000000;font-family:'Arial';font-size:10pt;vertical-align:bottom;w
hite-space:nowrap;direction:ltr;padding:2px 3px 2px 3px;}.ritz .waffle .s2{background-color:#ffffff;tex
t-align:right;color:#000000;font-family:'Arial';font-size:10pt;vertical-align:bottom;white-space:nowrap
;direction:ltr;padding:2px 3px 2px 3px;}.ritz .waffle .s4{background-color:#ffffff;text-align:left;colo
r:#2a2a2a;font-family:'Arial';font-size:10pt;vertical-align:bottom;white-space:nowrap;direction:ltr;pad
ding:2px 3px 2px 3px;}.ritz .waffle .s3{background-color:#ffffff;text-align:left;text-decoration:underl
ine;-webkit-text-decoration-skip:none;text-decoration-skip-ink:none;color:#1155cc;font-family:'Arial';f
```

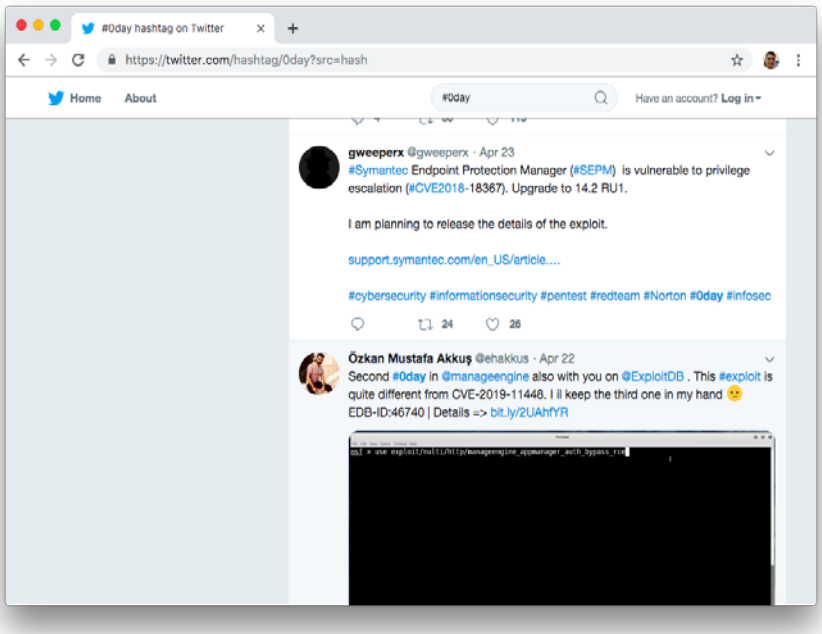
Εικόνα 26: more All.html

Από το αρχείο αναζητούμε λέξεις κλειδιά σχετικά με ευπάθειες που αφορούν το τοπικό δίκτυο μέσα στον HTML κώδικα της σελίδας που ανακτήθηκε από το google drive.

3.5.3 Crawler Twitter #0day

Ο crawler για το twitter θα αναζητά exploits που δημοσιεύονται με το hashtag #0day και είναι διαθέσιμα στην σελίδα https://twitter.com/hashtag/0day?src=hash (Εικόνα 27).
Για την αναζήτηση των tweets με το συγκεκριμένο hashtag θα χρησιμοποιηθεί το πακέτο https://github.com/piroor/tweet.sh το οποίο διαθέτει έναν crawler ο οποίος χρησιμοποιεί το νέο API του twitter. Για την λειτουργία του script πρέπει να δημιουργηθεί developer account στο twitter προκειμένου να λάβουμε τα απαραίτητα tokens και keys για την λειτουργία του script.
Στη συνέχεια με απλές εντολές τύπου SQL μπορούμε να κάνουμε αναζήτηση στα tweets με διάφορους τρόπους όπως εμφανίζονται στο παρακάτω παράδειγμα στο οποίο γίνεται αναζήτηση για ευπάθειες σχετικά με το κέλυφος Bash ή γενικά για το Shell Script.

```
./tweet.sh search -q "queries" -c 10
./tweet.sh search -q "Bash OR Shell Script"
```



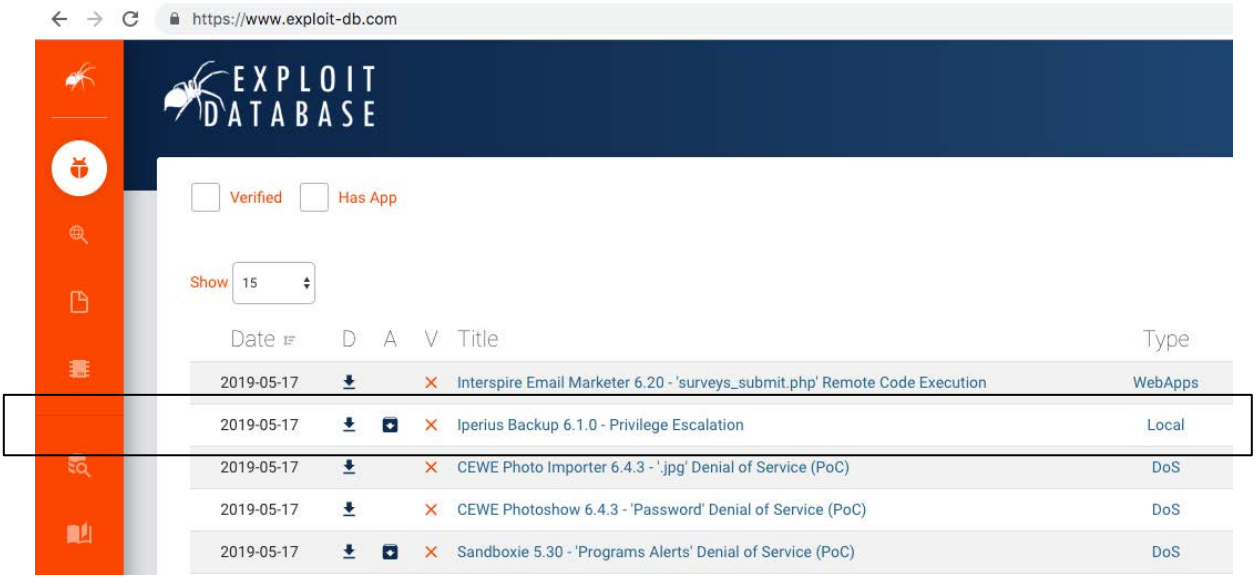
Εικόνα 27: twitter $0day

Το script θα αναζητά αυτόματα σε τακτά χρονικά διαστήματα και θα προσπαθεί να εντοπίσει νέες ευπάθειες οι οποίες στη συνέχεια θα καταγράφονται στη βάση δεδομένων.

Πολλές από τις εγγραφές που θα ανακτηθούν δεν θα είναι χρήσιμες αφού θα αφορούν νέα και ειδήσεις σχετικά με τα 0day exploits και πολύ λίγα από αυτά τα feeds θα αναφέρουν πραγματικές ευπάθειες που μπορούν να καταγραφούν. Σημαντικό σημείο σε αυτόν τον crawler είναι η χρήση λέξεων κλειδιών αφού τα feeds είναι γραμμένα σε φυσική (Αγγλική) γλώσσα.

3.5.4 Crawler exploit-db.com

Η ιστοσελίδα exploit-db.com περιέχει exploits που έχουν καταγραφεί σε μία ενιαία βάση δεδομένων. Η βάση δεδομένων ενημερώνεται καθημερινά και θεωρείται μία από τις πιο αξιόπιστες ιστοσελίδες του δημόσιο διαδικτύου (Εικόνα 28).

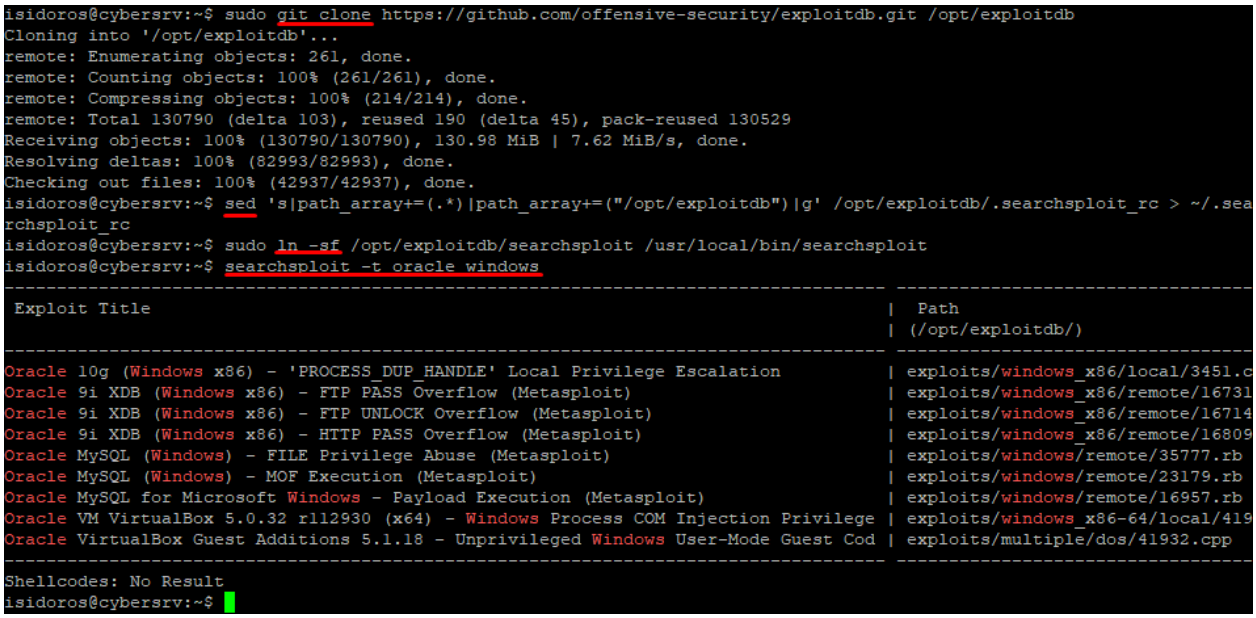


Εικόνα 28: Exploit-db

Ο crawler του exploit-db μπορεί να διαβάζει την ιστοσελίδα και επιστρέφει όλες τις ευπάθειες που βρέθηκαν για την συγκεκριμένη αναζήτηση. Για την ταχύτερη αξιοποίηση των στοιχείων της ιστοσελίδας exploit-db.com να χρησιμοποιηθεί ο crawler ο οποίος έχει κατασκευαστεί από την Offensive Security (https://www.offensive-security.com/) και

διαβάζει τα στοιχεία του exploit-db τα οποία στη συνέχεια καταγράφονται σε αρχείο απλού κειμένου.

Στη συνέχεια μέσω του script searchsploit (https://github.com/offensive-security/exploitdb) μπορεί να γίνει εύκολη και γρήγορη αναζήτηση στην τοπική βάση δεδομένων (Εικόνα 29).



Εικόνα 29: searchsploit

Στην παραπάνω εικόνα γίνεται αναζήτηση για ευπάθειες που αφορούν τις λέξεις κλειδιά oracle και windows.

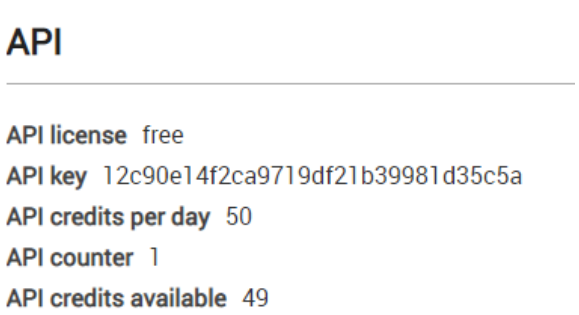
Η βάση δεδομένων του exploit-db ενημερώνεται καθημερινά μέσω του cronjobs.

```
#upgrade exploit-db database
0 0 * * * searchsploit --update
```

3.5.5 Crawler VULDB

Η ιστοσελίδα vuldb.com είναι μία βάση δεδομένων στην οποία επιμελούνται και τεκμηριώνονται όλες οι ευπάθειες ασφαλείας που έχουν δημοσιευθεί για διάφορα προϊόντα λογισμικού. Είναι μία από τις σημαντικότερες πηγές πληροφόρησης του διαδικτύου και ενημερώνεται σε τακτά χρονικά διαστήματα.

Η κοινότητα του vuldb έχει κατασκευάσει crawler για την αυτόματη ανάκτηση των πληροφοριών σε μορφή json. (https://github.com/vuldb/vuldb-api-php-examples). Για την χρήση του crawler χρειάζεται δημιουργία λογαριασμού στο vuldb προκειμένου να έχουμε πρόσβαση στο API. Μετά την εγγραφή στην ιστοσελίδα δημιουργείται το API key το οποίο είναι απαραίτητο για την χρήση του API. Η δωρεάν χρήση της υπηρεσίας έχει όριο 50 κλήσεις την ημέρα. Το σύστημα θα ψάχνει για νέες ευπάθειες κάθε ώρα (σύνολο 24 ανά ημέρα), το οποίο είναι κάτω από το όριο της δωρεάν χρήσης.



Εικόνα 30: Δείγμα api key από VulDB

```
root@cybersrv:/home/isidoros/vuldb-api-php-examples# php index.php 12c90e14f2ca9719df21b39981d35c5a
API RESULT DATA
Array
(
    [0] => Array
        (
            [entry] => Array
                (
                    [id] => 145446
                    [title] => McAfee Total Protection up to 16.0.R22 Microsoft Windows Client privilege escalation
                    [timestamp] => Array
                        (
                            [create] => 1573648916
                        )

                )
```

Εικόνα 31: Vuldb api call

Το script που εκτελέστηκε κάλεσε το API του vuldb και επιστρέφει τις τελευταίες ευπάθειες όπως αυτές δημοσιεύονται στην ιστοσελίδα σε μορφή json.

Created	Base	Temp	Vulnerability	0day	Today	Exp	Rem	CVE
11/13/2019	5.5	5.5	McAfee Total Protection Microsoft Windows Client privilege escalation	$10k-$25k	$10k-$25k			CVE-2019-3648
11/13/2019	5.5	5.5	Lenovo ThinkPad BIOS Tamper Detection privilege escalation	$2k-$5k	$2k-$5k			CVE-2019-6188
11/13/2019	5.5	5.5	Lenovo ThinkPad SMI Callback Code Execution	$2k-$5k	$2k-$5k			CVE-2019-6172

Εικόνα 32: Στιγμιότυπο από την ιστοσελίδα vuldb.com

Στη συνέχεια το json φιλτράρεται και αξιολογείται από τον crawler. Σε περίπτωση που υπάρχουν στοιχεία τα οποία μπορούν να αξιοποιηθούν καταγράφονται στην βάση δεδομένων.

3.5.6 Crawler 0day.today

Η ιστοσελίδα 0day.today περιέχει exploits που έχουν καταγραφεί σε μία ενιαία βάση δεδομένων. Από πάρα πολλούς ερευνητές ασφαλείας θεωρείται η πλέον αξιόπιστη σχετικά με της ευπάθειες μηδενικής ημέρας με βασικό κριτήριο ότι η ιστοσελίδα διαθέτει συχνά ευπάθειες προς αγορά με κρυπτό νομίσματα δημιουργώντας ένα πάρα πολύ γνωστό marketplace του dark-web.

Η σελίδα έχει πρόσβαση και από το ανοικτό διαδίκτυο στη διεύθυνση https://0day.today/ αλλά και από το dark-web μέσω της διεύθυνσης http://mvfjfugdwgc5uwho.onion/. Οι δύο διευθύνσεις παρουσιάζουν τα ίδια στοιχεία και δεν έχει εντοπιστεί διαφορά στην ταχύτητα ενημέρωσης των δύο ιστοσελίδων.

Εικόνα 33: Στιγμιότυπο από την ιστοσελίδα 0day.today

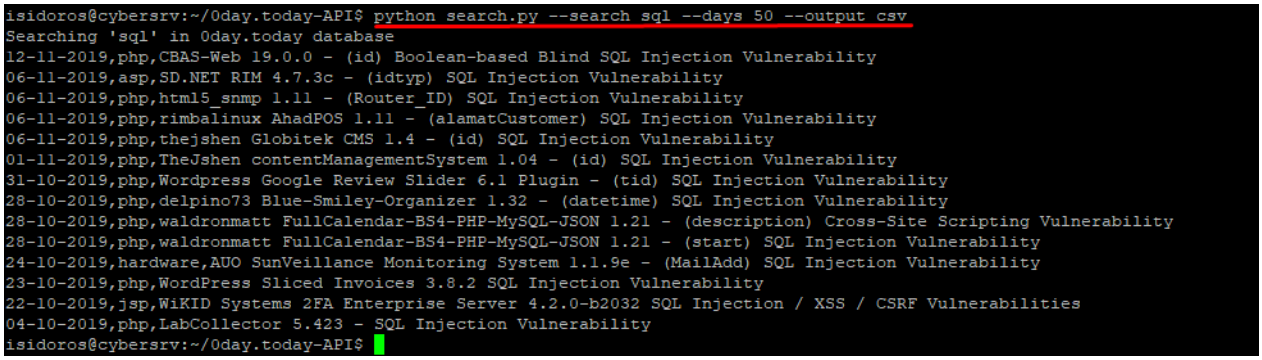
Ο crawler θα πρέπει να διαβάζει την ιστοσελίδα μία φορά την ημέρα (κάθε μέρα για την προηγούμενη) και να καταγράφει τα δεδομένα στη βάση δεδομένων. Η αναζήτηση των στοιχείων μπορεί να γίνει με ανάγνωση του HTML κώδικα που περιέχει η σελίδα. Για την αναζήτηση των ευπαθειών στην ιστοσελίδα θα χρησιμοποιηθεί το πακέτο https://github.com/MrSentex/0day.today-API το οποίο διαθέτει μία επικοινωνία με την ιστοσελίδα παρακάμπτοντας την ασφάλεια της σελίδας για bots.

Το παρακάτω παράδειγμα αναζητά για ευπάθειες σχετικά με το ssh (Εικόνα 34).

```
isidoros@cybersrv:~/0day.today-API$ python python_test.py
Searching 'ssh' in 0day.today database
====== Exploit =======
Date: 08-03-2019
Description: OpenSSH SCP Client - Write Arbitrary Files Exploit
Platform: multiple
Price: free
Author: Harry Sintonen
URL: https://0day.today//exploit/32328
======================
====== Exploit =======
Date: 20-01-2019
Description: OpenSSH 7.6p1 SCP Client - Multiple Vulnerabilities (SSHtranger Things) Exploit
Platform: multiple
Price: free
Author: Mark E. Haase
URL: https://0day.today//exploit/32009
======================
====== Exploit =======
Date: 04-12-2018
Description: OpenSSH < 7.7 - User Enumeration Exploit (2)
Platform: linux
Price: free
Author: Leap Security
URL: https://0day.today//exploit/31730
```

Εικόνα 34: crawler 0day.today

Για τις ανάγκες της αυτόματης αναζήτησης ευπαθειών δημιουργήθηκε το script search.py το οποίο στηρίζεται στο αρχικό script (python_test.py). Το νέο script έχει τη δυνατότητα αναζήτησης ευπαθειών με λέξεις κλειδιά αλλά και τη δυνατότητα να φιλτραριστούν τα αποτελέσματα ημερολογιακά καθώς και να αλλάξει η έξοδος των αποτελεσμάτων σε αρχείο csv. Η παρακάτω εικόνα (Εικόνα 35) εμφανίζει τις ευπάθειες των τελευταίων 50 ημερών σχετικά με την λέξη κλειδί sql σε μορφή csv.

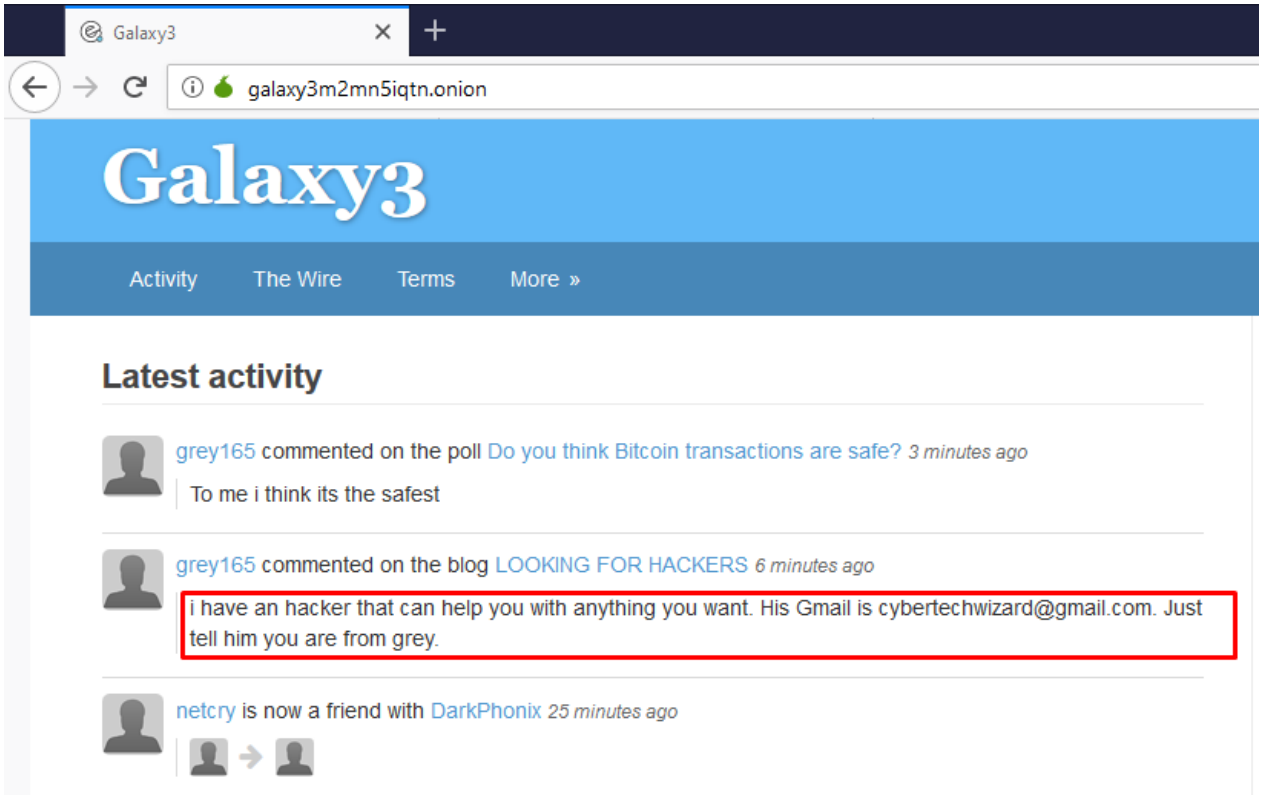


Εικόνα 35: Αποτελέσματα αναζήτησης 0day

Η παραπάνω λειτουργία προστέθηκε στο αρχικό project και βρίσκεται στην διεύθυνση του github (https://github.com/imoulas/0day.today-API)

3.5.7 Crawler Galaxy3

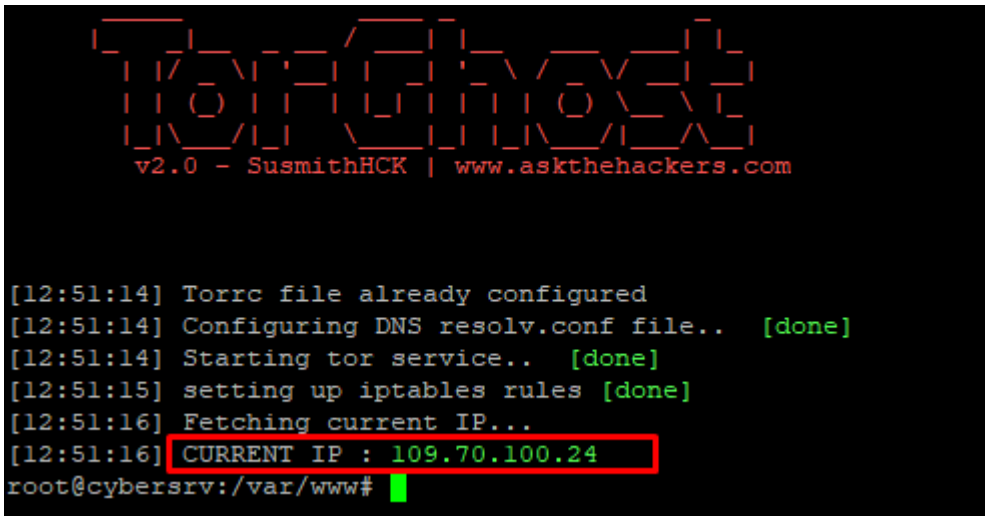
Το Galaxy3 είναι ένας ιστότοπος ο οποίος είναι προσβάσιμος μόνο μέσω του tor network στο dark-web. Δεν υπάρχει πρόσβαση από το ελεύθερο διαδίκτυο όπως συμβαίνει με άλλες ιστοσελίδες ενδιαφέροντος όπως το 0day.today. Η διεύθυνση του Galaxy3 είναι η http://galaxy3m2mn5iqtn.onion και πρόκειται για μία ιστοσελίδα στην οποία οι χρήστες ανεβάζουν διάφορες παράνομες δραστηριότητες προς πώληση, ενημέρωση ή αναζήτηση πληροφοριών (Εικόνα 36).



Εικόνα 36: Galaxy3

Για παράδειγμα στην παραπάνω Εικόνα 36: Galaxy3 εμφανίζεται κάποιος χρήστης ο οποίος έχει επαφή ή είναι ο ίδιος με κάποιον κυβερνοεγκληματία και παρουσιάζεται το email του. Αντίστοιχα αν εμφανιστεί στη σελίδα κάποιο κείμενο το οποίο θα περιέχει κάποιες από τις λέξεις κλειδιά που δημιουργούνται κατά την διάρκεια τις αναζήτησης των hosts στο τοπικό δίκτυο, τότε θα καταγράφονται οι πληροφορίες και θα αξιολογούνται.

Για να δημιουργηθεί ένας αυτόματος crawler για την συγκεκριμένη ιστοσελίδα προϋποθέτει ότι υπάρχει σύνδεση στο dark-web μέσω του torghost. Το torghost θα δημιουργήσει ένα νέο κανάλι στο λειτουργικό σύστημα το οποίο θα δρομολογεί όλη την κίνηση στο dark-web. Για την σύνδεση στο δίκτυο tor εκτελούμε την εντολή torghost start (3.5.1 Πρόσβαση στο dark-web). Στη συνέχεια ο crawler θα μπορεί να έχει πρόσβαση στον ιστότοπο galaxy3 του dark-web.



Εικόνα 37: torghost start

Ο crawler που θα δημιουργηθεί θα είναι αυτοματοποιημένος χωρίς την αλληλεπίδραση του χρήστη, θα διαβάζει το περιεχόμενο της σελίδας το οποίο είναι HTML και στη συνέχεια θα εξάγει τα χρήσιμα στοιχεία βασισμένα σε λέξεις κλειδιά. Τα στοιχεία που θα εξάγει θα αξιολογούνται και αν κρίνονται ότι μπορούν να χρησιμοποιηθούν και περιέχουν χρήσιμες πληροφορίες θα καταγράφονται στην βάση δεδομένων για περαιτέρω επεξεργασία από τις υπόλοιπες υπηρεσίες του συστήματος όπως είναι η ενημέρωση των διαχειριστών.

Ο crawler κατασκευάστηκε σε γλώσσα προγραμματισμού php χρησιμοποιώντας το πακέτο curl (https://www.php.net/manual/en/book.curl.php) για να διαβάζει τον πηγαίο κώδικα της ιστοσελίδα (source code) ο οποίος είναι σε γλώσσα προγραμματισμού HTML. Με την χρήση του curl διαβάζουμε ολόκληρο τον πηγαίο κώδικα της ιστοσελίδας σε μία μεταβλητή της php που στη συνέχεια θα επεξεργαστούμε. Για την ίδια λειτουργία μπορεί να χρησιμοποιηθεί και η γλώσσα προγραμματισμού python με την βιβλιοθήκη selenium (Shiaeles, Kolokotronis, & Bellini, 2019). Για την αποθήκευση των στοιχείων στη βάση δεδομένων και την παρουσίαση τους σε μία ιστοσελίδα με όλες τις γνωστές ευκολίες όπως αναζήτηση, εξαγωγή (search, export) χρησιμοποιήθηκε το yii framework αλλά μπορεί να χρησιμοποιηθεί οποιοδήποτε άλλο παρόμοιο framework όπως το Django.

Το πρόγραμμα διαβάζει την ιστοσελίδα με curl και θέτει τον user agent σε κάποιο γνωστό browser ώστε η ιστοσελίδα να ανιχνεύει ότι το request γίνεται από κάποιον χρήστη και όχι

από κάποιο bot. Πολλές ιστοσελίδες μπλοκάρουν την πρόσβαση σε headless browsers που στις περισσότερες περιπτώσεις πρόκειται για bots όπως ο crawler που κατασκευάζουμε.

```
$ch = curl_init();

// set url
curl_setopt($ch, CURLOPT_URL, "http://galaxy3m2mn5iqtn.onion/");

//return the transfer as a string
curl_setopt($ch, CURLOPT_RETURNTRANSFER, 1);
curl_setopt($ch, CURLOPT_USERAGENT,'Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.8.1.13) Gecko/20080311

// $output contains the output string
$output = curl_exec($ch);
```

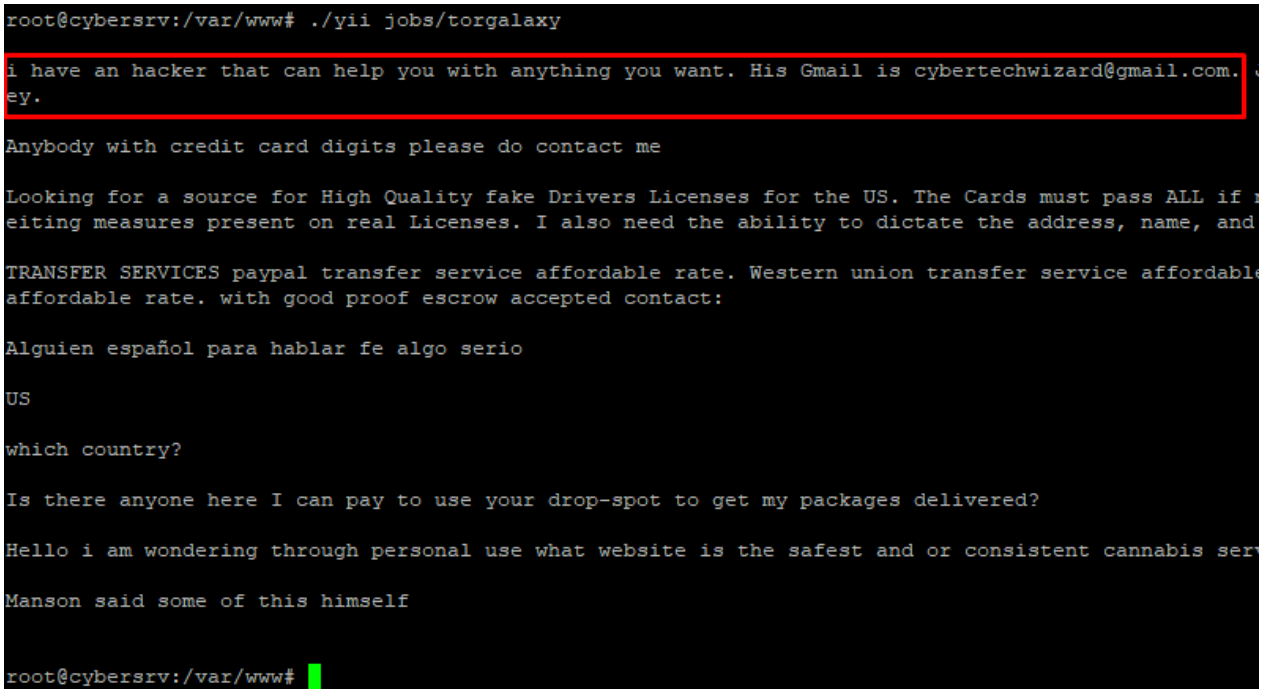
Εικόνα 38: Galaxy3 curl

Στη συνέχεια ο crawler πραγματοποιεί αναζήτηση μέσα στον πηγαίο κώδικα για τα συγκεκριμένα div που περιέχουν το συγκεκριμένο class (Εικόνα 39: Galaxy3 div) μέσα στην html σελίδα. Το σημείο κλειδί της συγκεκριμένης αναζήτησης είναι η φράση elgg-river-message η οποία βρίσκεται στη δήλωση του div και πρόκειται για την CSS κλάση (class) που χρησιμοποιείται για την μορφοποίηση των γραφικών στον browser του χρήστη (χρώματα, γραμματοσειρές, στοίχιση κ.α.). Στο συγκεκριμένο div το οποίο αρχίζει με <div class … και τελειώνει με εμπεριέχεται το κείμενο του κάθε post της ιστοσελίδας Galaxy3.

```
▼ <div class="elgg-river-message">
    i have an hacker that can help you with anything you want. His Gmail is
    cybertechwizard@gmail.com. Just tell him you are from grey.
  </div>
```

Εικόνα 39: Galaxy3 div

Το παραπάνω div επαναλαμβάνεται πολλές φορές στην ιστοσελίδα. Η διαδικασία επαναλαμβάνεται για όλα τα divs με το συγκεκριμένο class έως το σημείο που θα τελειώσει ο πηγαίος κώδικας της HML σελίδας. Κατά την διάρκεια της ανίχνευσης, ο crawler εμφανίζει τα αποτελέσματα στην οθόνη (Εικόνα 40: Galaxy3 output).

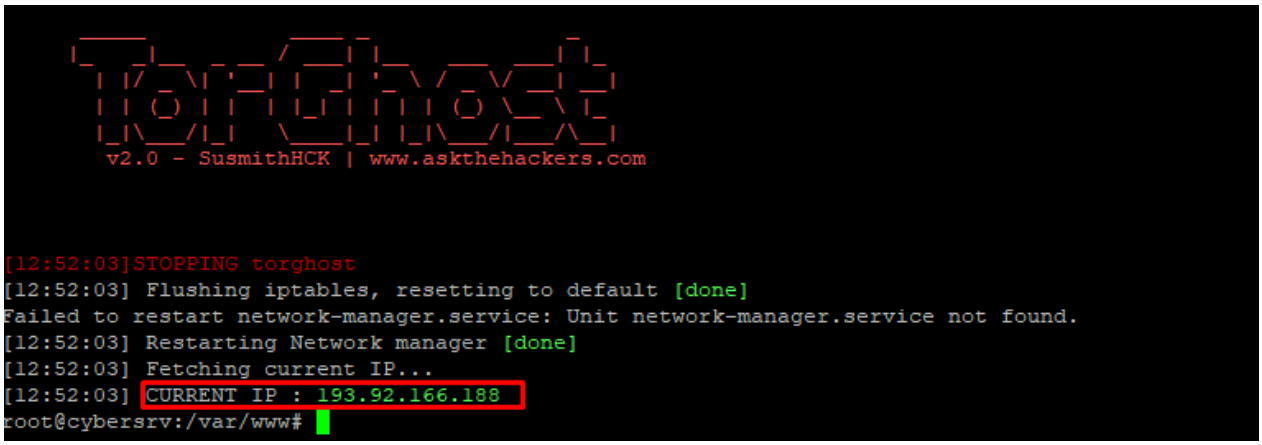


Εικόνα 40: Galaxy3 output

Στην παραπάνω εικόνα βλέπουμε τα αποτελέσματα από την εκτέλεση του crawler. Τα αποτελέσματα έχουν καταγραφεί στην βάση δεδομένων και η διαδικασία επαναλαμβάνεται σε τακτά χρονικά διαστήματα κατά την διάρκεια της ημέρας.

Σε περίπτωση που η δομή της σελίδας αλλάξει τότε θα πρέπει να γίνουν διορθωτικές κινήσεις στον crawler για να υποστηρίζει την νέα δομή HTML της σελίδας.

Μετά το τέλος της διαδικασίας το script κλείνει το δίκτυο tor με την εντολή torghost stop και η δρομολόγηση επιστρέφει στην αρχική κατάσταση μέσω του συνηθισμένου διαδικτύου (Εικόνα 41: torghost stop).



Εικόνα 41: torghost stop

Τα παραπάνω βήματα αυτοματοποιούνται με την δημιουργία του script crawlertor.sh το οποίο εκτελεί όλα τα βήματα σειριακά (Εικόνα 42).

```
#!/bin/bash
#
# Isidoros Moulas
# imoulas@hotmail.com
# 2019
#

torghost start
sleep 5
/var/www/yii jobs/torgalaxy
torghost stop
```

Εικόνα 42: crawlertor.sh

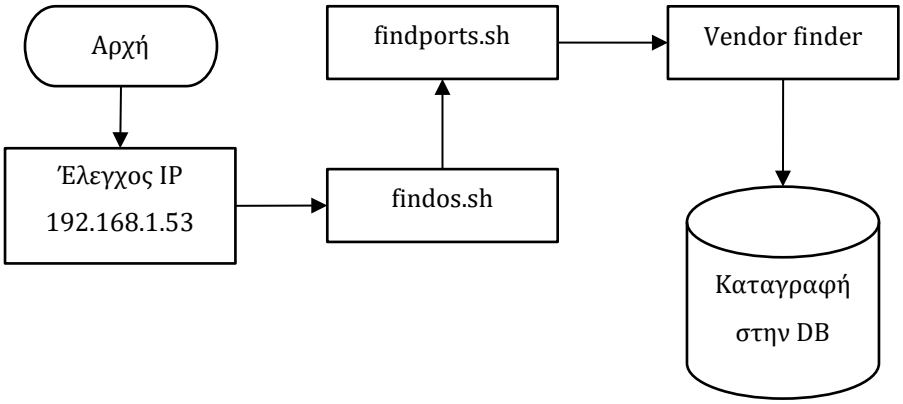
3.6 Monitoring

Ένα σημαντικό μέρος του συστήματος είναι η παρακολούθηση (monitoring) του τοπικού δικτύου. Η διαδικασία της παρακολούθησης πραγματοποιείται αφού ολοκληρωθεί η διαδικασία αναζήτησης των συσκευών στο τοπικό δίκτυο. Το τοπικό δίκτυο είναι ένα ελεγχόμενο δίκτυο στο οποίο υπάρχουν συνδεδεμένες διάφορες δικτυακές συσκευές. Στο δίκτυο αυτό δύναται να προστεθούν συσκευές με νέες υπηρεσίες και στις υπάρχουσες συσκευές να προστεθούν νέες υπηρεσίες, για παράδειγμα σε έναν διακομιστή να προστεθεί μία νέα υπηρεσία όπως web server. Η βασική εργασία του monitoring είναι να ελέγχει την τρέχουσα κατάσταση του δικτύου και να την συγκρίνει με την προηγούμενη κατάσταση που ήδη έχει καταγραφεί στην τοπική βάση δεδομένων.

Το monitoring script έχει γραφεί σε γλώσσα προγραμματισμού php και έχει χρησιμοποιηθεί το Yii Framework (https://www.yiiframework.com/). Το script διαβάζει το αρχείο hosts.txt το οποίο περιέχει τις IP διευθύνσεις που βρέθηκαν στο τοπικό δίκτυο. Για κάθε διεύθυνση εκτελούνται διαδοχικά τα script findos.sh και findports.sh (3.4 Αναζήτηση υπηρεσιών στο τοπικό δίκτυο.) τα οποία επιστρέφουν την έκδοση του λειτουργικού συστήματος και τις ανοικτές-εκτεθειμένες πόρτες της κάθε συσκευής. Στη συνέχεια γίνεται αναζήτηση του κατασκευαστή της συσκευής από το πακέτο nmap mac prefixes (Διάγραμμα 5).

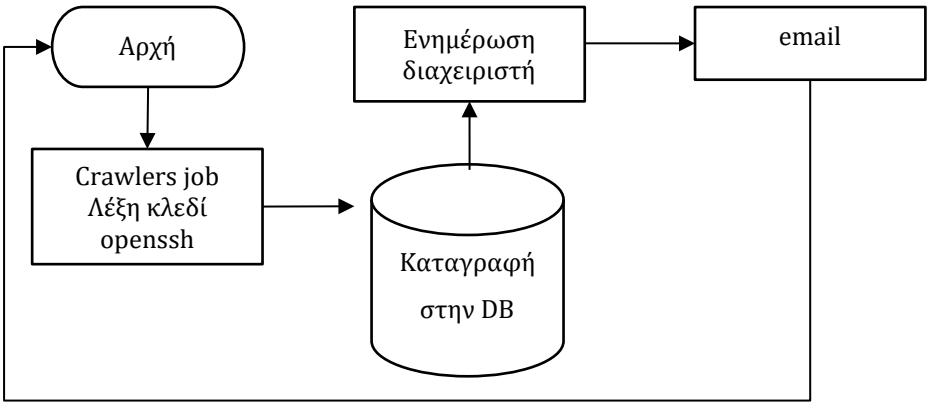
Αφού ολοκληρωθεί η εκτέλεση των παραπάνω scripts γίνεται αναζήτηση των καταγεγραμμένων αποτελεσμάτων που υπάρχουν στη βάση δεδομένων με τα νέα αποτελέσματα. Αν υπάρχουν διαφορές τότε καταγράφονται τα νέα στοιχεία και ενημερώνεται ο σχετικός πίνακας στην βάση δεδομένων (MySQL).

Ο ενημερωμένος πίνακας (με όνομα host) στην βάση δεδομένων (MySQL) θα αποτελέσει στη συνέχεια το φίλτρο για την αναζήτηση ευπαθειών στο διαδίκτυο και στο Dark-Web.



Διάγραμμα 5: Ροή εργασίας

Στη συνέχεια για κάθε εγγραφή που βρίσκεται στον πίνακα host και για κάθε υπηρεσία που εντοπίστηκε γίνεται η αναζήτηση για τις πιθανές ευπάθειες που αφορούν την συγκεκριμένη υπηρεσία (Διάγραμμα 6). Για παράδειγμα αν ο παραπάνω host με διεύθυνση IP 192.168.1.53 έχει εκτεθειμένη την υπηρεσία SSH (port 22/TCP), τότε το script θα αναζητήσει τις πιθανές ευπάθειες με λέξεις κλειδιά όπως ssh. Αν έχει εντοπιστεί η έκδοση του ssh, για παράδειγμα openssh 5.7, τότε η αναζήτηση θα αναζητήσει ευπάθειες και με την έκδοση του λογισμικού.

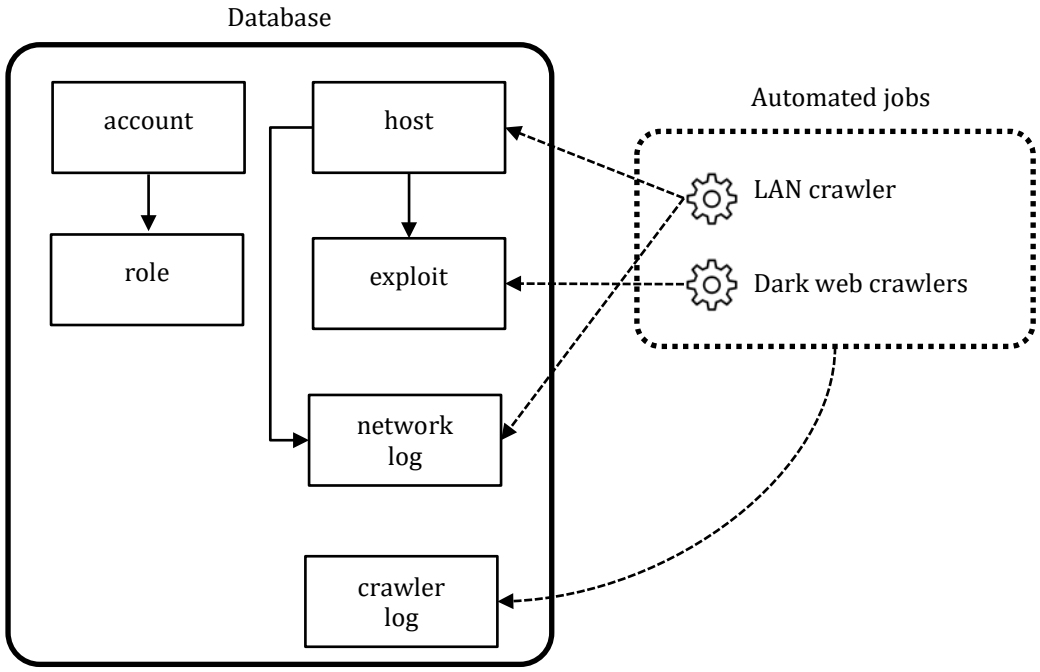


Διάγραμμα 6: Crawlers flow

Αν η υπηρεσία (πχ openssh) έχει εντοπιστεί σε περισσότερους hosts τότε η αναζήτηση πραγματοποιείται μία φορά για όλες τις συσκευές και γίνεται μαζική ενημέρωση της ενημέρωση της βάσης δεδομένων και στη συνέχεια των διαχειριστών με ηλεκτρονικό ταχυδρομείο.

3.7 Παρουσίαση δεδομένων

Όλα τα δεδομένα που αναζητούν οι crawlers καταγράφονται σε μία σχεσιακή βάση δεδομένων (MySQL). Η βάση δεδομένων (Διάγραμμα 7) περιέχει στοιχεία για όλες τις ενέργειες που πραγματοποιούνται από το σύστημα κατά την διάρκεια των αυτοματοποιημένων λειτουργιών του.



Διάγραμμα 7: Βάση δεδομένων

Στη βάση δεδομένων υπάρχουν πίνακες για την διαχείριση των χρηστών (account) και ρόλων (role) τα οποία είναι απαραίτητα για τον έλεγχο πρόσβασης στα στοιχεία της εφαρμογής και της βάσης δεδομένων.

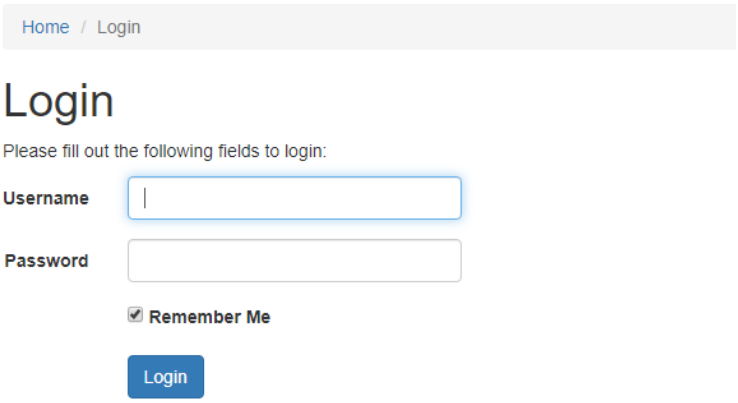
Ο πίνακας host περιέχει όλες τις πληροφορίες για τις συσκευές που εντοπίστηκαν στο τοπικό δίκτυο, ανεξάρτητα από τον τύπο και το είδος της συσκευής. Στον πίνακα καταγράφονται αναλυτικά όλα τα στοιχεία όπως διεύθυνση IP, mac address, λειτουργικό σύστημα, πόρτες που ανιχνεύτηκαν καθώς επίσης και η ημερομηνία/ώρα που πραγματοποιήθηκε η τελευταία ανίχνευση από τους crawlers.

Ο πίνακας networklog περιέχει ένα αρχείο καταγραφής συμβάντων με ότι εντοπίζεται στο τοπικό δίκτυο. Για παράδειγμα καταγράφονται πληροφορίες ότι ο host με διεύθυνση IP εντοπίστηκε στο δίκτυο ή ότι σε κάποιον host εντοπίστηκε μία νέα υπηρεσία που δεν είχε εντοπιστεί τις προηγούμενες ημέρες.

Στον πίνακα exploit καταγράφονται τα πιθανά exploit που έχουν εντοπίσει οι crawlers για το darkweb σε σχέση με τους hosts που έχουν εντοπιστεί στη βάση δεδομένων. Στον πίνακα καταγράφεται επίσης η ημερομηνία του εντοπισμού και το είδος του exploit (όπου αυτό είναι διαθέσιμο).

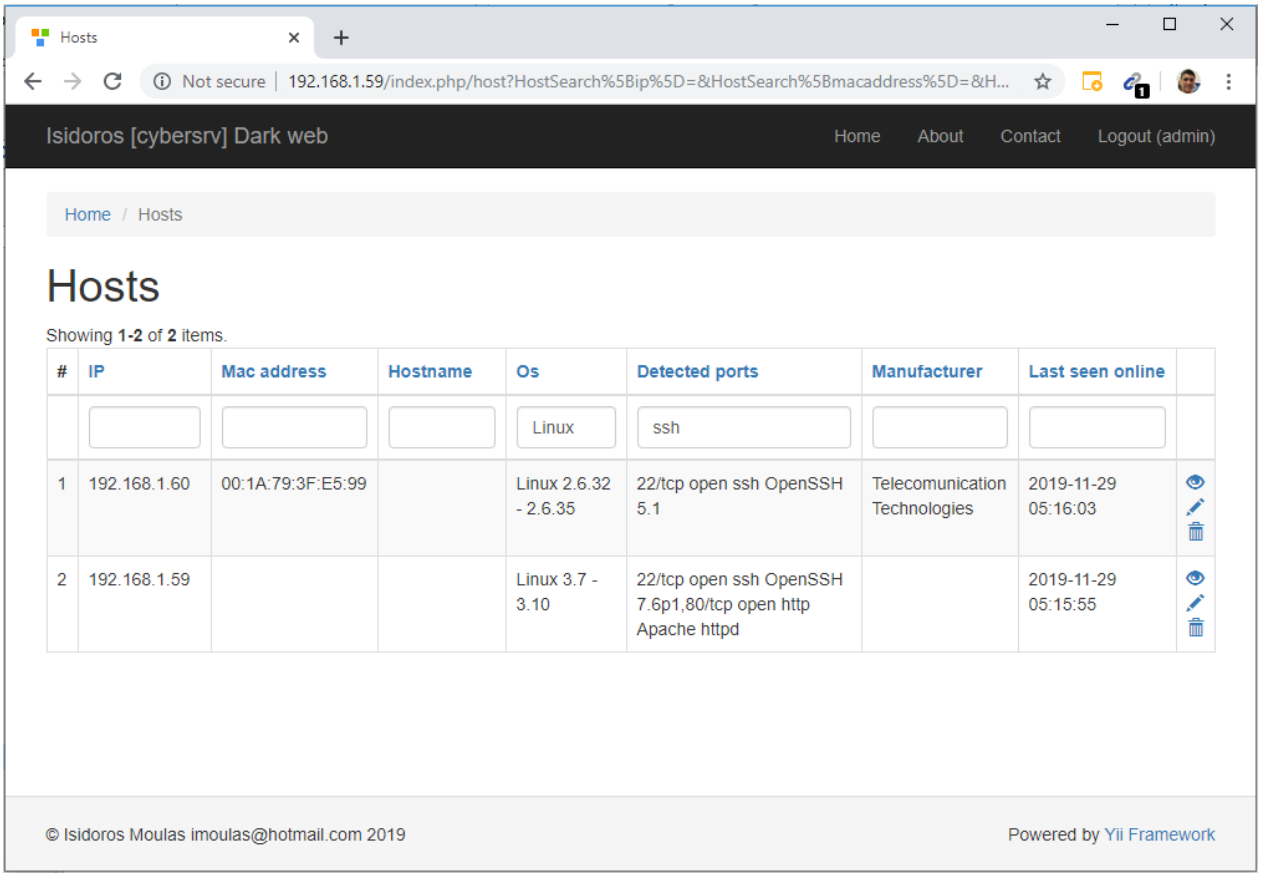
Ο πίνακας crawlerlog περιέχει ένα αρχείο καταγραφή συμβάντων σχετικά με τους crawlers του συστήματος. Ο πίνακας περιέχει το όνομα του crawler, την ημερομηνία που ξεκίνησε η αναζήτηση αλλά και το τέλος της εργασίας αναζήτησης.

Για την παρουσίαση των δεδομένων με εύκολο και εύχρηστο τρόπο χρησιμοποιήθηκε το Yii Framework. Στη συνέχεια κατασκευάστηκε μία νέα εφαρμογή (web app) στον τοπικό web για την παρουσίαση των στοιχείων. Αρχικά, δημιουργήθηκε ένα σύστημα πρόσβασης στην ιστοσελίδα με username και password (Εικόνα 43) το οποίο είναι απαραίτητο για την πρόσβαση στην εφαρμογή.



Εικόνα 43: webapp login

Μετά την επιτυχημένη πρόσβαση στην εφαρμογή μπορούμε να δούμε τις ανιχνευμένες υπηρεσίες του τοπικού δικτύου. Η εφαρμογή μας δίνει την δυνατότητα για αναζήτηση/φιλτράρισμα των στοιχείων της βάσης δεδομένων (Εικόνα 44).



Εικόνα 44: webapp hosts

Στην παραπάνω εικόνα εμφανίζονται τα στοιχεία από την βάση δεδομένων. Στο στιγμιότυπο έχει εφαρμοστεί φιλτράρισμα στο λειτουργικό σύστημα (Operating system Linux) και στις ανιχνευμένες υπηρεσίες του τοπικού δικτύου (ssh). Η εφαρμογή έχει την δυνατότητα για πολλαπλά φίλτρα σε οποιαδήποτε στήλη-στοιχείο του πίνακα.

Κάθε εγγραφή του πίνακα παρέχει περαιτέρω ανάλυση στα στοιχεία της εγγραφής και μπορούμε να δούμε αναλυτικά στοιχεία (Εικόνα 45).

ID	6
IP address	192.168.1.60
Mac address	00:1A:79:3F:E5:99
Hostname	
Os	Linux 2.6.32 - 2.6.35
Detected ports	22/tcp open ssh OpenSSH 5.1
Manufacturer	Telecomunication Technologies
Last seen online	2019-11-29 05:16:03

Εικόνα 45: webapp details

Στην παραπάνω εικόνα εμφανίζονται αναλυτικά στοιχεία για τον host με IP διεύθυνση 192.168.1.60.

Κεφάλαιο 4

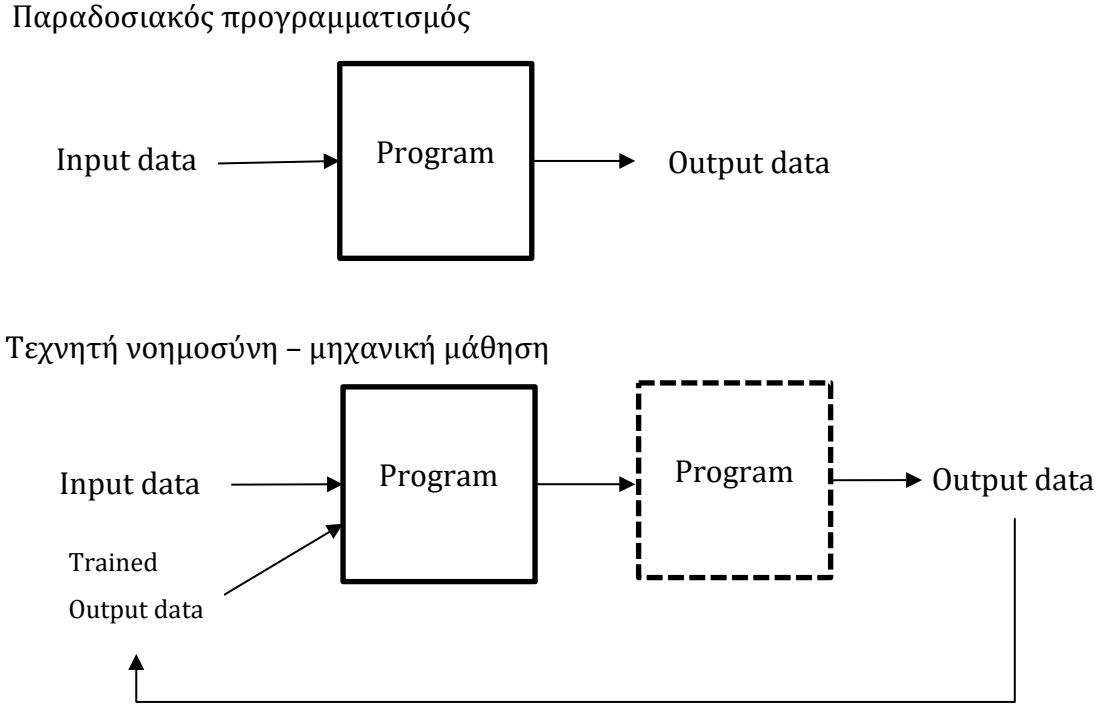
Τεχνητή νοημοσύνη

Ο όρος τεχνητή νοημοσύνη αναφέρεται στον κλάδο της πληροφορικής ο οποίος ασχολείται με τη σχεδίαση και την υλοποίηση υπολογιστικών συστημάτων που μιμούνται στοιχεία της ανθρώπινης συμπεριφοράς τα οποία υπονοούν έστω και στοιχειώδη ευφυΐα: μάθηση, προσαρμοστικότητα, εξαγωγή συμπερασμάτων, κατανόηση από τα συμπεράσματα, επίλυση προβλημάτων (Wikipedia). Η τεχνητή νοημοσύνη (AI) δίνει τη δυνατότητα στις μηχανές να μάθουν από την εμπειρία, να προσαρμόζονται σε νέες εισροές και να εκτελούν ανθρώπινες εργασίες. Τα περισσότερα γνωστά παραδείγματα αφορούν σκακιστικούς υπολογιστές μέχρι τα αυτοκίνητα με αυτόνομη οδήγηση, τα οποία βασίζονται σε μεγάλο βαθμό στη βαθιά εκμάθηση (deep learning) και την επεξεργασία της φυσικής γλώσσας. Χρησιμοποιώντας αυτές τις τεχνολογίες, οι υπολογιστές μπορούν να εκπαιδεύονται για να ολοκληρώσουν συγκεκριμένες εργασίες, επεξεργάζοντας μεγάλους όγκους δεδομένων και αναγνωρίζοντας πρότυπα στα δεδομένα από τα προ εγκατεστημένα μοντέλα εκμάθησης της μηχανής. Η εξόρυξη δεδομένων είναι ένας γνωστός τομέας στην εξαγωγή χρήσιμων πληροφοριών από διάφορες πηγές με την χρήση τεχνητής νοημοσύνης.

Οι έξυπνες ανιχνευτές (crawlers) πρέπει να αναπτυχθούν και να χρησιμοποιηθούν για να ξεπεράσουν το συνεχώς αυξανόμενο όγκο πληροφοριών του διαδικτύου και του dark web. Η μηχανική μάθηση (machine learning) χρησιμοποιείται ευρέως για την ανάπτυξη λογισμικού το οποίο οδηγεί στην ανάπτυξη λογισμικού επόμενης γενιάς. Η εξόρυξη των δεδομένων είναι μία εντατική και επίπονη δραστηριότητα. Οι αλγόριθμοι εξόρυξης δεδομένων διαδραματίζουν σημαντικό ρόλο εισάγοντας τεχνητή νοημοσύνη στους ανιχνευτές (crawlers).

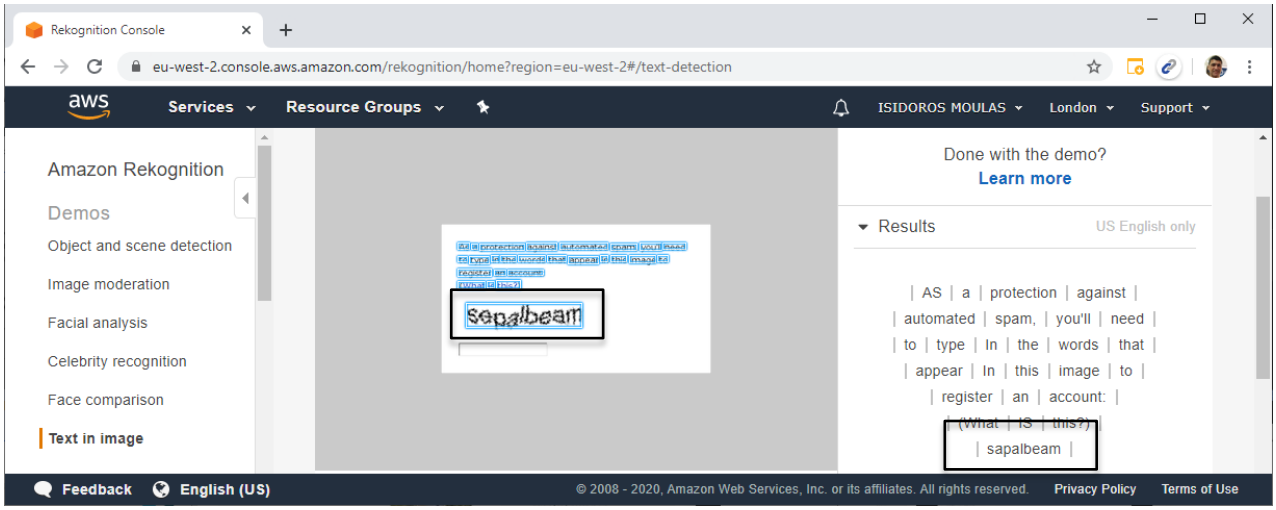
Για να ενσωματώσουμε τεχνητή νοημοσύνη και μηχανική μάθηση στους crawlers θα πρέπει πρώτα να τους εκπαιδεύσουμε με πληροφορίες. Οι πληροφορίες αυτές θα απαντούν στο ερώτημα για το αν μία πληροφορία που ανιχνεύτηκε πρέπει να αξιολογηθεί και να καταχωρηθεί στο σύστημα μας. Κατά την διάρκεια της εκπαίδευσης τα δεδομένα που ανιχνεύονται αξιολογούνται από τον εκπαιδευτή – ειδικό τεχνικό προσωπικό ασφάλειας υπολογιστικών συστημάτων. Σε αυτό το στάδιο καταγράφονται τα δεδομένα πρότυπα/μοντέλα τα οποία το σύστημα θα χρησιμοποιήσει προκειμένου να είναι σε θέση να προβλέψει τις νέες απαιτήσεις της εκάστοτε αναζήτησης.

Ο παραδοσιακός προγραμματισμός έχει μία είσοδο δεδομένων στην οποία γίνεται επεξεργασία των δεδομένων και παράγεται ένα αποτέλεσμα. Στην τεχνητή νοημοσύνη κατά την διάρκεια της μηχανικής μάθησης συνεχίζουμε να έχουμε μία είσοδο δεδομένων αλλά γνωρίζουμε περίπου ποια είναι τα αναμενόμενα αποτελέσματα με σκοπό να κατασκευαστεί το κατάλληλο πρόγραμμα από τον υπολογιστή προκειμένου να επιτευχθεί ο σκοπός (Εικόνα 46).



Εικόνα 46 Τεχνητή νοημοσύνη

Ένας άλλος τομέας που η τεχνητή νοημοσύνη βοηθά στην αναζήτηση πληροφοριών στο dark web είναι η παράκαμψη της ασφάλειας των ιστοσελίδων που αφορούν τα bots και τους crawlers. Πολλές φορές η αναζήτηση των δεδομένων σε μία ιστοσελίδα δεν είναι εφικτή αν προηγουμένως ο χρήστης δεν κάνει κάποια ενέργεια όπως να πατήσει κάποιο κουμπί ή υπερσύνδεση σε κάποιο αναδυόμενο παράθυρο (popup window) ή ακόμα να πρέπει να συμπληρώσει ένα captcha προκειμένου να συνδεθεί στη σελίδα. Με την χρήση της τεχνητής νοημοσύνης είναι εφικτό (υπό προϋποθέσεις) να πραγματοποιηθούν ενέργειες για λογαριασμό του χρήστη, οι οποίες συνήθως απαιτούν την παρέμβαση του.
Για παράδειγμα (Εικόνα 47) στην περίπτωση της προστασίας με κωδικό captcha μπορεί να χρησιμοποιηθεί μία υπηρεσία αναγνώρισης εικόνων (face recognition) οι οποία θα επιστρέψει την σωστή απάντηση από την αλλοιωμένη εικόνα captcha.



Εικόνα 47 Image recognition

Σε περίπτωση που η εικόνα είναι αρκετά αλλοιωμένη ενδέχεται η αναγνώριση να αποτύχει. Στο επόμενο βήμα μετά την αποτυχημένη πρόσβαση, η ιστοσελίδα θα εμφανίσει μία διαφορετική εικόνα captcha στην οποία το αυτόματο σύστημα τεχνητής νοημοσύνης θα ανιχνεύσει και θα αποστείλει για αναγνώριση. Η διαδικασία επαναλαμβάνεται για όσες φορές χρειαστεί.
Μία άλλη περίπτωση χρήσης της τεχνητής νοημοσύνης στο σύστημα μας είναι η παράκαμψη του συστήματος ασφαλείας μίας ιστοσελίδας. Η εφαρμογή (Εικόνα 48) μετακινεί αυτόματα το δείκτη του ποντικιού από οποιαδήποτε θέση βρίσκεται στην

επιφάνεια εργασίας του υπολογιστή στη σωστή θέση πάνω στο παράθυρο που εμφανίζεται η ιστοσελίδα. Στη συνέχεια, αυτόματα μετά από 2 δευτερόλεπτα πραγματοποιεί αριστερό κλικ στην υπερσύνδεση που βρίσκεται στο σημείο αυτό ακριβώς κάτω από τον κέρσορα του ποντικιού.

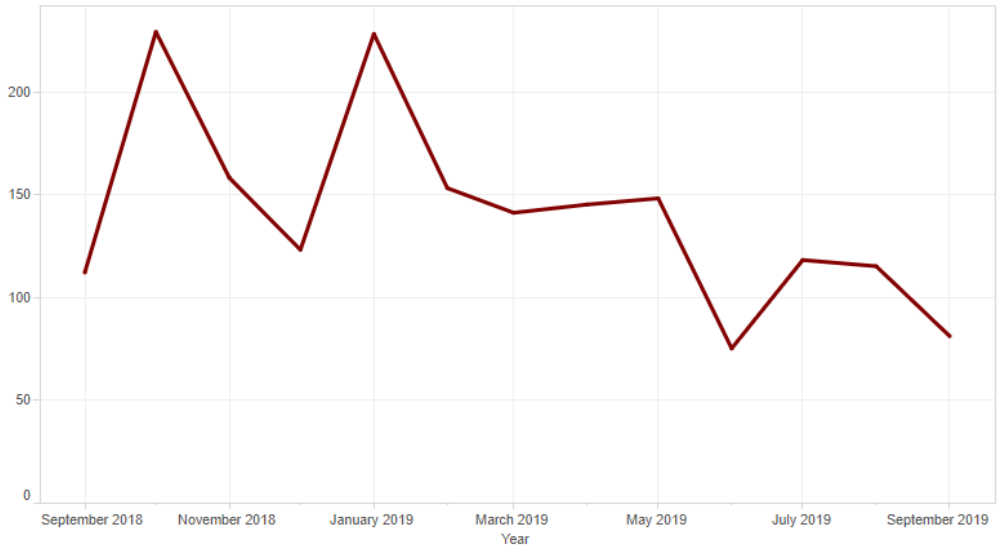
Εικόνα 48 AI cursor move

Η τεχνητή νοημοσύνη σε συνδυασμό με την μηχανική μάθηση έχει τεράστιο πεδίο εφαρμογής στο dark web αφού η κοινότητα συνεχώς εξελίσσεται και μεταβάλλεται. Με την χρήση της τεχνολογίας θα μπορούν να αξιοποιηθούν οι περισσότερες πληροφορίες από το διαδίκτυο και το dark web προκειμένου η αναζήτηση των ευπαθειών μηδενικής ημέρας να είναι όσο το δυνατόν πιο άμεση, γρήγορη και έγκυρη.

Κεφάλαιο 5

Συμπεράσματα – Μελλοντική Εργασία

Η σελίδα exploit-db.com έχει καταγεγραμμένο έναν μεγάλο αριθμό ευπαθειών και ενημερώνεται καθημερινά για όλες τις νέες ευπάθειες που ανακαλύπτονται από ερευνητές. Η παρακάτω Εικόνα 49 παρουσιάζει τις ευπάθειες που έχουν καταγραφεί στην ιστοσελίδα τους τελευταίους 13 μήνες (από Σεπτέμβριο 2018 έως και Σεπτέμβριο 2019). Από το διάγραμμα βγαίνει το συμπέρασμα ότι εντοπίζονται περισσότερες ευπάθειες τους μήνες Οκτώβριο έως και Ιανουάριο ενώ οι υπόλοιποι μήνες βρίσκονται σε χαμηλότερα επίπεδα.



Εικόνα 49: Στατιστικά exploit-db.com

4.1 Συμπεράσματα

Στην έρευνα που διενεργήθηκε στο τοπικό δίκτυο (Εικόνα 3) υπήρχαν συσκευές οι οποίες ήταν συνεχώς συνδεδεμένες στο δίκτυο αλλά και συσκευές που συνδέθηκαν κατά την διάρκεια της έρευνας. Αρκετές από τις συσκευές ήταν εκτεθειμένες σε διάφορες ευπάθειες σε όλη την διάρκεια της έρευνας. Οι συσκευές που ήταν συνδεδεμένες στο τοπικό δίκτυο μπορούν να χαρακτηριστούν στις παρακάτω 3 κατηγορίες.

a) Προσωρινές συσκευές

   Συσκευές που εμφανιζόντουσαν στο δίκτυο για μερικές μέρες όπως συσκευές τηλέφωνα/tablet Android/iOS και φορητοί υπολογιστές.

b) Νέες συσκευές

   Συσκευές που προστέθηκαν στο δίκτυο μετά την πρώτη ημέρα λειτουργίας του συστήματος, όπως μία νέα θέση εργασίας, ένας δικτυακός εκτυπωτής, ένας δρομολογητής.

c) Υπάρχουσες συσκευές

   Συσκευές οι οποίες υπήρχαν στο δίκτυο πριν την έναρξη λειτουργίας του συστήματος.

Ειδικά για τις συσκευές που ήταν συνεχώς συνδεδεμένες στο δίκτυο είτε είναι νέες συσκευές (κατά τη διάρκεια της έρευνας), είτε παλιές που ήδη έχουν εντοπιστεί από το σύστημα (πριν την έρευνα), γίνεται αναζήτηση των υπηρεσιών που διαθέτει η κάθε συσκευή ανεξάρτητα πότε η συσκευή εντοπίστηκε στο δίκτυο. Υπάρχουν περιπτώσεις που σε κάποια συσκευή προστίθεται ή καταργείται μία υπηρεσία. Για παράδειγμα σε κάποιον εξυπηρετητή προστέθηκε υπηρεσία DNS (UDP/53). Σε αυτές τις συσκευές γίνεται πρόσθετος έλεγχος για τις πιθανές ευπάθειες της νέας υπηρεσίας προκειμένου να διασφαλιστεί αν έχει εγκατασταθεί κάποια παλιά έκδοση του λογισμικού.

Στον παρακάτω πίνακα εμφανίζονται συσκευές και υπηρεσίες τους που υπήρχαν στο τοπικό δίκτυο. Ο πίνακας είναι συνοπτικός και παρουσιάζει τα πιο χαρακτηριστικά παραδείγματα της έρευνας.

Συσκευή	Λειτουργικό σύστημα	Υπηρεσία
Planet VOIP PBX	DD-WRT v24 or v30 (Linux 3.10)	21/tcp ftp, 22/tcp ssh, 23/tcp telnet, 25/tcp smtp, 53/tcp domain, 110/tcp pop3, 143/tcp imap, 199/tcp smux, 256/tcp fw1-secureremote, 554/tcp rtsp, 993/tcp imaps, 995/tcp pop3s, 1025/tcp NFS-or-IIS, 1720/tcp h323q931, 1723/tcp pptp, 5900/tcp vnc, 8080/tcp http-proxy, 8888/tcp sun-answerbook
Grandstream VOIP phone	Linux 2.6.13 - 2.6.32	23/tcp telnet Grandstream VoIP, 80/tcp http BusyBox httpd
Windows Desktop	Windows 10	135/tcp msrpc Microsoft Windows, 139/tcp netbios-ssn Microsoft Windows, 445/tcp microsoft-ds?, 3389/tcp ms-wbt-server Microsoft Terminal, 5357/tcp http Microsoft HTTPAPI, 7070/tcp ssl/realserver?
WiFi AP	Tp-Link	23/tcp telnet BusyBox telnetd, 80/tcp http TP-LINK TD-W8968, 1900/tcp upnp Portable SDK
Linux server	Ubuntu Linux 3.7 - 3.10	22/tcp ssh OpenSSH 7.6p1, 80/tcp http Apache httpd
Router	Mikrotik	22/tcp ssh MikroTik RouterOS, 53/tcp domain MikroTik RouterOS, 80/tcp http MikroTik router, 443/tcp ssl/https?, 1723/tcp pptp MikroTik (Firmware:, 2000/tcp bandwidth-test MikroTik bandwidth-test

Το σύστημα αναζητούσε σε τακτά χρονικά διαστήματα μέσω των crawlers για νέες ευπάθειες που αφορούν τις συγκεκριμένες συσκευές που προηγουμένως είχαν εντοπιστεί στο τοπικό δίκτυο. Οι ευπάθειες που εντοπίζονταν κάθε φορά καταγράφονται σε βάση δεδομένων για την ενημέρωση των διαχειριστών. Όταν μία νέα συσκευή συνδεθεί στο δίκτυο τότε γίνεται αναζήτηση για τις ευπάθειες που έχουν εντοπιστεί για την συγκεκριμένη συσκευή.

Για παράδειγμα, στο δίκτυο προστέθηκε-συνδέθηκε ο δρομολογητής Mikrotik. Αμέσως μετά την σύνδεση του το σύστημα εντόπισε τον δρομολογητή και τις υπηρεσίες που είχε ο δρομολογητής εκτεθειμένες. Στη συνέχεια έγινε αναζήτηση των ευπαθειών μέσω των crawlers για τον εντοπισμό πιθανών ευπαθειών. Η αναζήτηση επέστρεψε τα αποτελέσματα των crawlers (Εικόνα 50 και Εικόνα 51) τα οποία έδειξαν ότι εντοπίστηκαν διάφορες ευπάθειες σχετικά με την συγκεκριμένη συσκευή.

```
Searching 'mikrotik' in 0day.today database
31-10-2019,hardware,MikroTik RouterOS 6.45.6 - DNS Cache Poisoning Exploit
21-02-2019,hardware,MikroTik RouterOS < 6.43.12 (stable) / < 6.42.12 (long-term) - Firewall and NAT Bypass
09-08-2018,windows,Mikrotik WinBox 6.42 - Credential Disclosure Exploit
16-03-2018,hardware,MikroTik RouterOS < 6.41.3/6.42rc27 - SMB Buffer Overflow Exploit
13-03-2018,hardware,MikroTik RouterOS < 6.38.4 (x86) - Chimay Red Stack Clash Remote Code Execution Exploit
13-03-2018,hardware,MikroTik RouterOS < 6.38.4 (MIPSBE) - Chimay Red Stack Clash Remote Code Execution Exploit
24-01-2018,hardware,MikroTik RouterOS < 6.38.5 Remote Command Execution Exploit
16-12-2018,hardware,Mikrotik RouterOS Telnet Arbitrary Root File Creation Vulnerability
13-04-2018,linux,MikroTik 6.41.4 - FTP daemon Denial of Service PoC
28-03-2017,hardware,MikroTik RouterBoard 6.38.5 - Denial of Service Exploit
04-03-2017,hardware,MikroTik Router Denial Of Service | ARP Table OverFlow Exploit
```

Εικόνα 50: Mikrotik 0day

```
isidoros@cybersrv:~$ searchsploit -t mikrotik
---------------------------------------------------------------------------- ---------------------------------
 Exploit Title                                                              |  Path
                                                                            | (/opt/exploitdb/)
---------------------------------------------------------------------------- ---------------------------------
MikroTik 6.40.5 ICMP - Denial of Service                                    | exploits/hardware/dos/43317.c
MikroTik 6.41.4 - FTP daemon Denial of Service PoC                          | exploits/linux/webapps/44450.txt
MikroTik Router - ARP Table OverFlow Denial Of Service                      | exploits/hardware/dos/41601.c
MikroTik RouterBoard 6.38.5 - Denial of Service                             | exploits/hardware/dos/41752.pl
MikroTik RouterOS - sshd (ROSSSH) Remote Heap Corruption                    | exploits/hardware/remote/28056.tx
MikroTik RouterOS 3.0 - SNMP SET Denial of Service                          | exploits/hardware/dos/31102.c
MikroTik RouterOS 3.13 - SNMP write (Set request)                           | exploits/hardware/remote/6366.c
MikroTik RouterOS 6.45.6 - DNS Cache Poisoning                              | exploits/hardware/remote/47566.cp
MikroTik RouterOS < 6.38.4 (MIPSBE) - 'Chimay Red' Stack Clash Remote Code Ex | exploits/hardware/remote/44283.py
MikroTik RouterOS < 6.38.4 (x86) - 'Chimay Red' Stack Clash Remote Code Execu | exploits/hardware/remote/44284.py
MikroTik RouterOS < 6.41.3/6.42rc27 - SMB Buffer Overflow                   | exploits/hardware/remote/44290.py
MikroTik RouterOS < 6.43.12 (stable) / < 6.42.12 (long-term) - Firewall and N | exploits/hardware/remote/46444.tx
Mikrotik Router - Denial of Service                                         | exploits/hardware/dos/18817.py
Mikrotik Syslog Server for Windows 1.15 - Denial of Service (Metasploit)    | exploits/windows/dos/24968.rb
Mikrotik WinBox 6.42 - Credential Disclosure (Metasploit)                   | exploits/windows/remote/45170.py
Mikrotik WinBox 6.42 - Credential Disclosure (golang)                       | exploits/hardware/webapps/45209.g
Web Interface for DNSmasq / Mikrotik - SQL Injection                        | exploits/php/webapps/39817.php
---------------------------------------------------------------------------- ---------------------------------
Shellcodes: No Result
```

Εικόνα 51: Mikrotik exploit-db

Η πληροφορίες καταγράφονται στην τοπική βάση δεδομένων και στη συνέχεια αυτόματα από το σύστημα ενημερώνονται οι διαχειριστές για την ευπάθεια που εντοπίστηκε στη συγκεκριμένη συσκευή. Η ενημέρωση των διαχειριστών γίνεται με αυτοματοποιημένο email με το παρακάτω πρότυπο (template).

From: system

To: **{adminemail}**

Subject: [Warning] Vulnerability detected **{ip}**

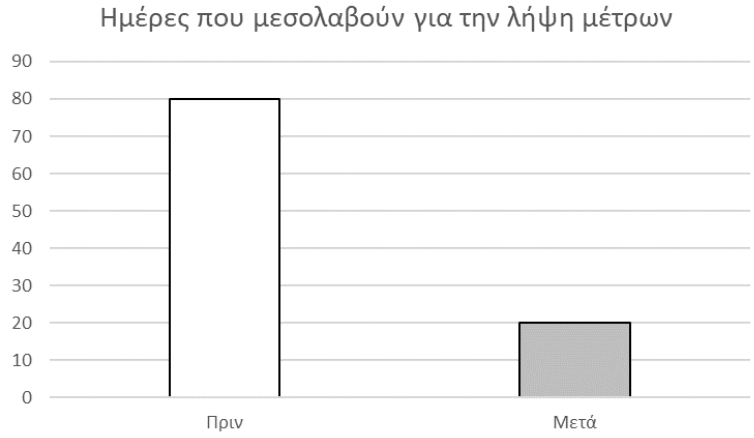
The device **{manufacturer}** with **{ip}** IP address is detected vulnerable to the following exploits:

{exploits_list}

Οι διαχειριστές οι οποίοι θα λάβουν από το σύστημα την αυτοματοποιημένη ενημέρωση με ηλεκτρονικό ταχυδρομείο και θα πρέπει να πραγματοποιήσουν διορθωτικές κινήσεις

προκειμένου να αντιμετωπιστούν οι απειλές, είτε αναβαθμίζοντας το λογισμικό (software upgrade) ή ρυθμίζοντας αυστηρότερους κανόνες στο τείχος προστασίας (firewall) του δικτύου αλλά και της συσκευής (εφόσον διαθέτει την δυνατότητα).

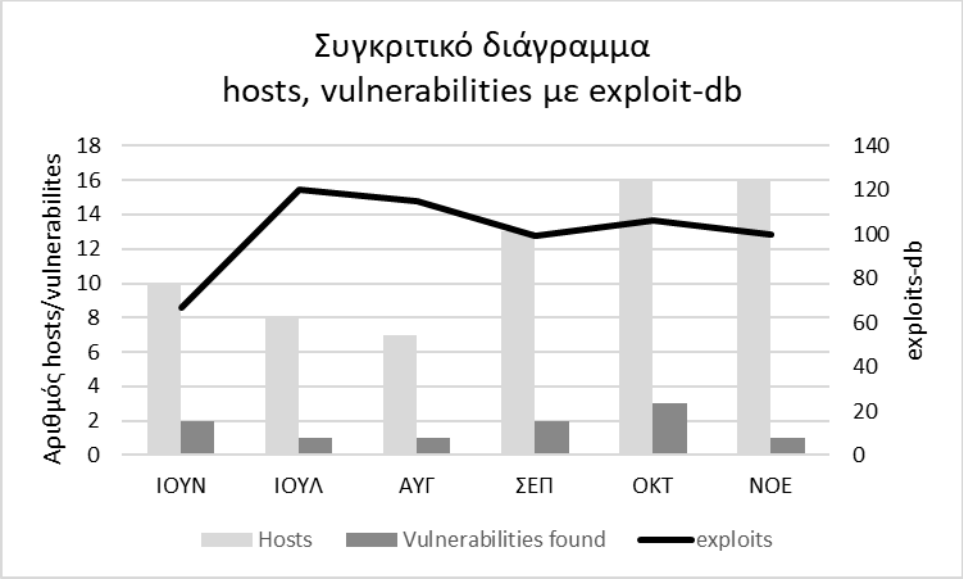
Ο χρόνος απόκρισης των διαχειριστών μειώθηκε σημαντικά και οι διαχειριστές είχαν την δυνατότητα να εφαρμόσουν πολιτικές ασφαλείας εγκαίρως πριν προκύψει μία επίθεση.



Διάγραμμα 8: Χρόνος που μεσολαβεί για την λήψη μέτρων

Ο καταγεγραμμένος χρόνος μειώθηκε σε 20 ημέρες από 80 ημέρες σύμφωνα με την έρευνα (Fransesco, Joanna, Aurelien, Michel, & Wendy, May 2017) με αποτέλεσμα το Διάγραμμα 2 το οποίο παρουσιάζει την διάρκεια μια επίθεσης μηδενικής ημέρας, να συρρικνωθεί σημαντικά αφού όλοι οι ενδιάμεσοι χρόνοι μειώνονται από την έγκαιρη ενημέρωση των διαχειριστών.

Στο παρακάτω Διάγραμμα 9 εμφανίζονται το σύνολο των vulnerabilities που καταγράφτηκαν από την ιστοσελίδα exploit-db.com ανά μήνα (με μαύρη γραμμή). Εμφανίζονται επίσης ο μέσος όρος των hosts που ήταν συνδεδεμένοι στο τοπικό δίκτυο ανά μήνα καθώς επίσης και οι ευπάθειες που βρέθηκαν από το σύστημα για τους συγκεκριμένους hosts.



Διάγραμμα 9: Συγκριτικό exploits/hosts/vulnerabilities

Παρατηρούμε ότι ο συνολικός αριθμός vulnerabilities που καταγράφεται μηνιαία είναι πάρα πολύ μεγάλος (exploit-db). Τα καταγεγραμμένα exploits αφορούν ευπάθειες για το σύνολο των διαθέσιμων συσκευών και ευπαθειών. Προφανώς σε ένα τοπικό δίκτυο αναμένουμε να έχουμε σημαντικά πολύ λιγότερες ευπάθειες αφού οι συσκευές και οι εφαρμογές είναι πολύ λιγότερες αλλά και δεν υπάρχουν σημαντικές αλλαγές από μέρα σε μέρα. Από το σύνολο των καταγεγραμμένων vulnerabilities ένας μικρός αριθμός αφορά το τοπικό δίκτυο και τις συσκευές που είναι συνδεδεμένες σε αυτό. Αυτό διευκολύνει σημαντικά την διαχείριση ενός συστήματος αφού γίνεται στοχευμένη προσέγγιση στις ευπάθειες σε συνδυασμό με τους hosts του δικτύου.

Το συμπέρασμα της έρευνας είναι ότι το dark-web μπορεί να χρησιμοποιηθεί ως βάση δεδομένων αρκεί να κατασκευαστούν ειδικοί crawlers για όλες τις σελίδες ενδιαφέροντος. Οι σελίδες ενδιαφέροντος συνεχώς αλλάζουν ή προστίθεται νέες και αυτό πρακτικά σημαίνει ότι η συντήρηση των crawlers θα πρέπει να είναι συνεχής προκειμένου οι crawlers για το dark-web να λειτουργούν αδιάκοπτα. Στη συνέχεια μπορούμε να αντλήσουμε στοιχεία για ευπάθειες που αφορούν το τοπικό δίκτυο χωρίς την αλληλεπίδραση του χρήστη. Τα δεδομένα που αντλούνται μπορούν να αξιοποιηθούν για την προστασία του τοπικού δικτύου όπως η άμεση ενημέρωση των διαχειριστών. Η έρευνα έδειξε ότι οι διαχειριστές του συστήματος ενημερώνονται σε πολύ σύντομο χρονικό

διάστημα (σχεδόν άμεσα) και αμέσως μόλις εντοπιστεί ένας σύστημα με ευπάθειες στο τοπικό δίκτυο.

4.2 Μελλοντική Εργασία

Η παρούσα εργασία μπορεί να βελτιωθεί σε δύο τομείς. Ο πρώτος τομέας είναι οι crawlers στους οποίους η αναζήτηση γίνεται με λέξεις κλειδιά σε συγκεκριμένες ιστοσελίδες ενδιαφέροντος. Η κατασκευή επιμέρους crawlers για κάθε ιστοσελίδα ενδιαφέροντος απαιτεί πολλές ώρες εργασίας και έρευνας.

Σε μία έρευνα προτείνεται η κατασκευή ενός crawler ο οποίος ακολουθεί όλα τα link που εντοπίζει στο dark-web δημιουργώντας ένα δαιδαλώδες δέντρο υπερσυνδέσεων. Ο crawler αυτός θα λειτουργεί συνέχεια μέχρι κάποια στιγμή να τερματίσει το οποίο είναι εξαιρετικά απίθανο αφού το δέντρο επικαλύπτεται από τα κλαδιά των υπερσυνδέσεων και πολλά κλαδιά θα ενώνονται αφού θα υπάρχουν υπερσυνδέσεις μεταξύ των ιστοσελίδων. Βέβαια έρευνες έχουν δείξει ότι το 87% των ιστότοπων στο dark-web δεν έχει σύνδεση με άλλες ιστοσελίδες και θα μπορούσε να χαρακτηριστεί ότι κάθε ιστότοπος είναι αυτόνομος και αποξενωμένος από το υπόλοιπο δίκτυο. Κατά την διάρκεια της πλοήγησης από σελίδα σε σελίδα θα ψάχνει για λέξεις κλειδιά και θα καταγράφει τις ιστοσελίδες που έχει εντοπίσει τα περισσότερα αποτελέσματα ώστε τις συγκεκριμένες ιστοσελίδες να τις ελέγχει συχνότερα. Όλα τα ευρήματα θα αποθηκεύονται σε μία μεγάλη βάση δεδομένων για την περαιτέρω επεξεργασία των πληροφοριών (Schäfer, και συν., 2019).

Οι κατασκευή των crawlers είναι μία τεχνική η οποία είναι ευρέως γνωστή και πάρα πολλοί ερευνητές ασφαλείας έχουν προσπαθήσει να κατασκευάσουν crawlers με επιτυχία τις περισσότερες φορές. Το πρόβλημα με τους crawlers είναι ότι στηρίζονται στο κείμενο HTML της σελίδας το οποίο μπορεί να αλλάξει ανά πάσα χρονική στιγμή. Πολλοί προγραμματιστές προκειμένου να προστατεύουν τις σελίδες από crawlers/bots δημιουργούν την σελίδα με ελάχιστα διαφορετικό HTML κάθε φορά ώστε στον browser να εμφανίζεται η σελίδα χωρίς αλλαγές αλλά στην πραγματικότητα ο πηγαίος κώδικας της σελίδας είναι διαφορετικός. Επίσης οι crawlers έρχονται αντιμέτωποι με τα συστήματα ασφάλειας της σελίδας είτε αυτό πρόκειται για ένα απλό login με username/password, είτε

με συστήματα antibot/anticrawler κ.α. Αυτό πρακτικά σημαίνει ότι ο crawler (που στην ουσία είναι ένα bot με ένα headless browser) να μην μπορεί να διαβάσει την σελίδα και να πρέπει να γίνουν διορθωτικές κινήσεις προκειμένου να συνεχίσει να λειτουργεί. Η διαδικασία αυτή της διόρθωσης του πηγαίου κώδικα είναι εξαιρετικά χρονοβόρα και πολλές φορές κάποια συστήματα ασφαλείας είναι πολύ ανθεκτικά σε τέτοιες τεχνικές.

Πολλές ιστοσελίδες που παρέχουν πληροφορίες για τα vulnerabilities διαθέτουν επικοινωνία μέσω API το οποίο είναι πολύ καλύτερο από την χρήση crawler, αφού με αυτόν το τρόπο η ιστοσελίδα επιτρέπει σε κάποιο crawler/bot να κάνει αυτόματη αναζήτηση στην σελίδας. Συνεπώς το πρόβλημα που εγείρεται είναι ότι οι σελίδες του dark-web που δεν έχουν API και διαθέτουν συστήματα ασφαλείας για τους crawlers είναι επί του παρόντος πολύ δύσκολο να διαβαστούν και να επεξεργαστούν οι πληροφορίες που υπάρχουν εκεί. Για να ξεπεραστεί αυτό το βήμα μπορεί να κατασκευαστεί ένας crawler που μιμείται τις κινήσεις του χρήστη (κούνημα ποντικιού) και την καθυστέρηση στην πληκτρολόγηση των στοιχείων πρόσβασης. Προφανώς αν αλλάξει κάτι στην ιστοσελίδα τότε και αυτός ο crawler θα πρέπει να διορθωθεί με τα νέα δεδομένα.

Ένας άλλος έξυπνος crawler θα μπορούσε να κατασκευαστεί ο οποίος να χρησιμοποιεί τεχνητή νοημοσύνη και να καταλαβαίνει πότε μία ευπάθεια που εντοπίστηκε έχει νόημα και πότε πρέπει να αγνοηθεί. Βέβαια τέτοιου είδους συστήματα απαιτούν την εκπαίδευση του μοντέλου από τον κατασκευαστή και την συνεχή βελτιστοποίηση του αλγόριθμου τουλάχιστον για ένα σημαντικό αριθμό ευρημάτων.

Αφού γίνει η ανακάλυψη των ευπαθειών με όλους τους παραπάνω τρόπους και υπάρχουν σημαντικές ενδείξεις ότι πρόκειται για βάσιμες απειλές τότε θα μπορούν να γίνονται αυτόματες ρυθμίσεις σε συστήματα IDS όπως το Snort.

Τα συστήματα IDS/IPS εντοπίζουν πιθανές διεισδύσεις στα δίκτυα και αποτρέπονται οι πιθανές επιθέσεις, καθώς επίσης και οι επιθέσεις σε επίπεδο εφαρμογής, όπως από είναι τα worms, τα trojans, τα προγράμματα κατασκόπους (spyware), καταγραφείς πληκτρολόγησης (keyloggers) κ.α. Σε περίπτωση εντοπισμού κάποιου συμβάτος, ενημερώνεται ο διαχειριστής για το κάθε συμβάν. Η αρχιτεκτονική και λειτουργία των συστημάτων αυτών βασίζεται σε ενσωματωμένα scripts τα οποία ελέγχουν την κίνηση για κάποιο ίχνος/υπογραφή ή για κάποιο μοτίβο και βασίζεται σε ήδη γνωστές απειλές. Η

διαφορά τους από τα τείχη προστασίας είναι ότι τα συστήματα IPS ελέγχουν την πρόσβαση σε επίπεδο εφαρμογή.

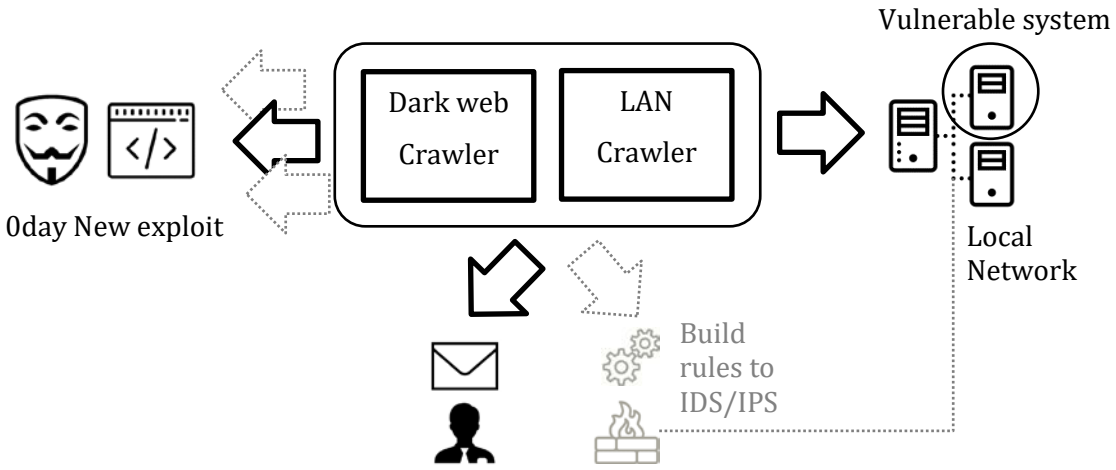
Για παράδειγμα στο σύστημα snort θα πρέπει να ενημερωθεί το αρχείο /etc/snort/rules για τις νέες ευπάθειες όπως το παρακάτω παράδειγμα.

```
alert tcp $EXTERNAL_NET any -> $HOME_NET $HTTP_PORTS
(msg:"COMMUNITY SQL-INJECTION OpenBB board.php";
flow:to_server,established; uricontent:"/board.php";
pcre:"/board.php\x3F\w+\x3D[0-9]+\s/Ui"; classtype:web-
application-attack; reference:bugtraq,7404; sid:100000108;
rev:1;)
```

Ο παραπάνω κανόνας αφορά την εφαρμογή OpenBB και μία ευπάθεια τύπου SQL injection στο αρχείο board.php της εφαρμογής.

Το snort ως ένα σύστημα IDS/IPS έχει αναλυθεί από τους ερευνητές πάρα πολλές φορές. Οι ερευνητές έχουν καταλήξει στο συμπέρασμα ότι το snort μπορεί να προστατεύσει ένα δίκτυο αποτελεσματικά εφόσον έχει τους κατάλληλους κανόνες και υπογραφές (Gaddam & Dr. Nandhini, 2017). Στην έρευνα του Albert (Sagala, 2015) προτείνεται η δημιουργία honeypots και στη συνέχεια να γίνεται αυτόματη δημιουργία κανόνων στο snort σε περίπτωση που εντοπιστεί κυβερνοεπίθεση στο honeypot. Το αποτέλεσμα της έρευνας ήταν ότι όλες οι επιθέσεις που πραγματοποιήθηκαν στο honeypot είχαν σαν αποτέλεσμα σε δημιουργία κανόνων στο snort που με τη σειρά τους προστάτεψαν το δίκτυο από τον επιτιθέμενο. Στη συνέχεια σε μία παρόμοια έρευνα (Upadhyay & Khilari, 2016) δημιουργούνται αυτόματοι κανόνες στο snort σχετικά με επιθέσεις τύπου SQL Injections που πραγματοποιούνται σε κάποιο honeypot.

Η μελλοντική εργασία (Διάγραμμα 10) της παρούσας έρευνας θα επικεντρώνονταν στην περαιτέρω ανάπτυξη των crawlers με περισσότερες πηγές και στην αυτοματοποιημένη δημιουργία κανόνων σε κάποιο IDS/IPS όπως το snort.



Διάγραμμα 10: Μελλοντική εργασία

Βιβλιογραφία

Baravalle, A., Lopez, S., & Lee, S. (2016). Mining the Dark Web Drugs and fake ids . *IEEE 16th International Conference on Data Mining Workshops*, (σσ. 350-356). London, United Kingdom.

Bau, J., Bursztein, E., Gupta, D., & Mitchell, J. (2010). State of the Art: Automated Black-Box Web Application Vulnerability Testing. *IEEE Symposium on Security and Privacy* (σσ. 332-345). Stanford University.

CHEN, J.-j., & CHENG, X.-j. (2009). A Novel Fast Port Scan Method Using Partheno-Genetic Algorithm. *2009 2nd IEEE International Conference on Computer Science and Information Technology*, (σσ. 219-222). Beijing.

Farook Bin Rafiuddin, M., Minhas, H., & Singh Dhubb, P. (2017). A Dark Web Story In-Depth Research and Study Conducted on the Dark Web based on Forensic Computing and Security in Malaysia. *IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI-2017)*, (σσ. 3049-3055). Malaysia.

Fransesco, V., Joanna, M., Aurelien, T., Michel, B.-L., & Wendy, M. (May 2017). *High Costs and Small Benefits: A Field Study of How Users Experience Operating System Upgrades.* Denver, United States. pp. 4242-4253 ff10.1145/3025453.3025509. hal-01493415.

Gaddam, R., & Dr. Nandhini, M. (2017). An Analysis of Various Snort Based Techniques to Detect and Prevent Intrusions in Networks. *International Conference on Inventive Communication and Computational Technologies*, (σσ. 10-15).

Graham, R., Maynor, D., & Security, E. (2011). *A Simpler Way of Finding 0day*. Ανάκτηση από blackhat.com: https://www.blackhat.com/presentations/bh-usa-07/Maynor_and_Graham/Whitepaper/bh-usa-07-maynor_and_graham-WP.pdf

Haraty, R., & Zantout, B. (Aug. 2014). The TOR Data Communication System. *ournal of Communications and Networks*, 415-420.

IANA.org. (χ.χ.). Ανάκτηση από iana.org: https://www.iana.org/assignments/service-names-port-numbers/service-names-port-numbers.xhtml

KADOGUCHI, M., HAYASHI, S., HASHIMOTO, M., & OTSUKA, A. (2019). Exploring the Dark Web for Cyber Threat Intelligence using Machine Leaning. *2019 IEEE International Conference on Intelligence and Security Informatics (ISI)*, (σσ. 200-202). Shenzhen, China.

Kao, C.-N., Chang, Y.-C., Huang, N.-F., Salim, S., Liao, I.-J., Liu, R.-T., & Hung, H.-W. (2015). A Predictive Zero-day Network Defense using. *IEEE Conference on Communications and Network Security (CNS)*, (σσ. 695-696). Florence.

Koloveas, P., Chantzios, T., Tryfonopoulos, C., & Skiadopoulos, S. (2019). A crawler architecture for harvesting the clear, social, and dark web. *IEEE World Congress on Services (SERVICES)*, (σσ. 3-8).

Lihet, M., & Pr.Dr. Dadarlat, V. (2018). Honeypot in the cloud. *2018 17th RoEduNet Conference: Networking in Education and Research (RoEduNet)*, (σσ. 1-6). Cluj-Napoca.

Marin, E., Almukaynizi, M., Nunes, E., & Shakarian, P. (2018). Community Finding of Malware and Exploit Vendors on Darkweb Marketplaces. *1st International Conference on Data Intelligence and Security*, (σσ. 81-84). Tempe, Arizona.

McQueen, M., McQueen, T., Boyer, W., & Chaffin, M. (2009). Empirical Estimates and Observations of 0Day Vulnerabilities . *Proceedings of the 42nd Hawaii International Conference on System Sciences*, (σσ. 1-12).

Nunes, E., Diab, A., Gunn, A., Marin, E., Mishra, V., Paliath, V., . . . Shakarian, P. (2016). Darknet and Deepnet Mining for Proactive Cybersecurity Threat Intelligence. *2016 IEEE Conference on Intelligence and Security Informatics (ISI)*, (σσ. 7-12). Tucson, AZ.

NVD. (χ.χ.). Ανάκτηση από National Vulnerability Database: https://nvd.nist.gov/

Sagala, A. (2015). Automatic SNORT IDS Rule Generation Based on Honeypot Log. *7th International Conference on Information Technology and Electrical Engineering*, (σσ. 576-580). Chiang Mai, Thailand.

Schäfer, M., Fuchs, M., Strohmeier, M., Engel, M., Liechti, M., & Lenders, V. (2019). BlackWidow: Monitoring the Dark Web for Cyber Security Information. *11th International Conference on Cyber Conflict: Silent Battle*, (σσ. 1-21). Tallinn.

Shah, M., Ahmed, S., Saeed, K., Junaid, M., & Khan, H. (2019). Penetration Testing Active Reconnaissance Phase - Optimized Port Scanning With Nmap Tool. *International Conference on Computing, Mathematics and Engineering Technologies.*

Shiaeles, S., Kolokotronis, N., & Bellini, E. (2019). IoT Vulnerability Data Crawling and Analysis. *IEEE World Congress on Services (SERVICES)*, (σσ. 78-83).

Sornalakshmi, K. (2017). Detection of DoS attack and Zero Day Threat with SIEM. *International Conference on Intelligent Computing and Control Systems*, (σσ. 1-7).

Upadhyay, U., & Khilari, G. (2016). SQL Injection Avoidance for Protected Database with ASCII using SNORT and Honeypot. *International Conference on Advanced Communication Control and Computing Technologies*, (σσ. 596-599).

What are zero-day attacks. (χ.χ.). Ανάκτηση από BullGuard: https://www.bullguard.com/bullguard-security-center/pc-security/computer-threats/what-are-zero-day-attacks.aspx

Zulkarnine, A., Frank, R., Mitchell, J., & Davies, G. (2016). Surfacing Collaborated Networks in Dark Web to Find Illicit and Criminal Content. *International CyberCrime Research Center (ICCRC)*, (σσ. 109-114). Burnaby, Canada .