ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
Σχολή Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών
Τομέας Επικοινωνιών, Ηλεκτρονικής & Συστημάτων Πληροφορικής

Σύνθετη Ανίχνευση Ανωμαλιών για Διαχείριση Ασφάλειας στο Διαδίκτυο

Internet Security Management: Data Fusion based Anomaly Detection

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Χρήστος Ι. Σιατερλής

Αθήνα, Ιούλιος 2006

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
Σχολή Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών
Τομέας Επικοινωνιών, Ηλεκτρονικής & Συστημάτων Πληροφορικής

Σύνθετη Ανίχνευση Ανωμαλιών για Διαχείριση Ασφάλειας στο Διαδίκτυο

Internet Security Management: Data Fusion based Anomaly Detection

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Χρήστος Ι. Σιατερλής

Συμβουλευτική Επιτροπή:
Καθηγητής Β. Μάγκλαρης
Καθηγητής Ε. Συκάς
Καθηγητής Ν. Μήτρου

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή την .... Ιουλίου 2006.

.....	.....	.....
Β. Μάγκλαρης	Ε. Συκάς	Ν. Μήτρου
Καθηγητής Ε.Μ.Π.	Καθηγητής Ε.Μ.Π.	Καθηγητής Ε.Μ.Π.
.....	.....	.....
Γ. Παπαβασιλόπουλος	Ι. Βενιέρης	Σ. Παπαβασιλείου
Καθηγητής Ε.Μ.Π.	Καθηγητής Ε.Μ.Π.	Επ. Καθηγητής Ε.Μ.Π.

.....
Σ. Κάτσικας
Καθηγητής Παν. Αιγαίου

Αθήνα, Ιούλιος 2006

.....

Χρήστος Ι. Σιατερλής
Διδάκτωρ Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Περίληψη

Το θέμα της παρούσας διατριβής είναι η ανίχνευση ανωμαλιών στη δικτυακή κίνηση και ειδικότερα η ανίχνευση "κατανεμημένων επιθέσεων άρνησης υπηρεσιών" (Distributed Denial of Service attacks -DDoS), ενός από τα μεγαλύτερα προβλήματα που αντιμετωπίζει σήμερα το Διαδίκτυο, ως μέσο για την ενίσχυση της ασφάλειας δικτύων. Στόχος είναι η ανάπτυξη μιας αρχιτεκτονικής για συστήματα ανίχνευσης ανωμαλιών στη δικτυακή κίνηση (Network Anomaly Detection Systems) που θα βασίζεται στις σημερινές πρακτικές διαχείρισης δικτύων και θα είναι επεκτάσιμη ώστε να προσαρμόζεται στην αναμενόμενη εξέλιξη των χαρακτηριστικών της δικτυακής κίνησης και των ανωμαλιών της. Η προτεινόμενη αρχιτεκτονική βασίζεται στην σύνθεση δεδομένων (data fusion) από πολλούς αισθητήρες (sensors) που συλλέγουν δεδομένα για τη δικτυακή κίνηση με χρήση διάφορων τεχνικών παθητικής παρακολούθησης δικτύων (π.χ. μέσω συλλογής πακέτων, SNMP MIB's και τεχνολογίας Netflow). Στην απλούστερη μορφή τους οι αισθητήρες χρησιμοποιούν ως αλγόριθμο ανίχνευσης σταθερές ή προσαρμοζόμενες συναρτήσεις κατωφλίου. Τα αποτελέσματα της ανίχνευσης κάθε στοιχειώδους αισθητήρα συνδυάζονται με έναν αλγόριθμο σύνθεσης δεδομένων όπως είναι η θεωρία των Dempster-Shafer (D-S).

Στα πλαίσια της διατριβής υλοποιήθηκε ένα σύστημα ανίχνευσης κατανεμημένων επιθέσεων άρνησης υπηρεσιών σύμφωνα με την προτεινόμενη αρχιτεκτονική. Κατά την υλοποίηση αναπτύχθηκε λογισμικό παρακολούθησης δικτύου, υπολογισμού μετρικών, ανίχνευσης και σύνθεσης δεδομένων. Παράλληλα αναπτύχθηκε εργαλείο εξομοίωσης επιθέσεων άρνησης υπηρεσιών με ελεγχόμενα χαρακτηριστικά. Με τον τρόπο αυτό διεξήχθησαν πειράματα σε περιβάλλον πραγματικού δικτύου υψηλών ταχυτήτων (στη γραμμή σύνδεσης του Εθνικού Μετσόβιου Πολυτεχνείου και του Εθνικού Δικτύου Έρευνας και Τεχνολογίας ταχύτητας 1Gbps). Η ανάλυση των πειραματικών αποτελεσμάτων ανέδειξε χρήσιμα μετρικά για την ανίχνευση δικτυακών ανωμαλιών και απέδειξε την αποτελεσματικότητα της σύνθεσης δεδομένων. Συγκεκριμένα καταδείχθηκε πως η ανίχνευση ανωμαλιών στους ακραίους δρομολογητές του δικτύου ενός παρόχου (Provider Edge routers) είναι εφικτή και πως η σύνθεση δεδομένων από αισθητήρες που βασίζονται σε διαφορετικά μετρικά, μπορεί να βελτιώσει την απόδοση ενός συστήματος ανίχνευσης.

Abstract

We increasingly consider Internet a standard utility, like electricity or telephone access. Reliability of its offered services becomes then critical and even a short downtime can cost hundreds of dollars. Distributed Denial of Service attacks (DDoS) are the prime cause for such cut-offs, especially because they are planned and executed by wicked individuals. The development of systems that are able to detect such attacks as well as a wider range of network traffic anomalies is an essential part of Internet security management. This thesis proposes an architecture for the development of Network Anomaly Detection Systems (NADS). A NADS has to be compatible with standard operating procedures of network administrators and at the same time flexible enough to adapt to the constant evolution of Internet traffic anomalies.

We propose an architecture whereby several simple metrics (traffic features) are being monitored and used to detect traffic anomalies. To that end, we heavily rely on data fusion concepts and algorithms developed within the "Dempster-Shafer Theory of Evidence" (D-S). For consistency with the terminology used extensively in "Multisensor data fusion" bibliography we adopt the term "sensor" to denote basic detection elements. Each sensor monitors, detects and reports its own perspective (belief) of the observed network traffic features. The sensors can rely on complementary network monitoring technologies like passive capturing, SNMP, Netflow. The beliefs of several sensors are then combined (fused) in order to detect network traffic anomalies.

To test and evaluate the proposed architecture we developed a NADS prototype and performed extensive experiments on high speed production networks (the Greek Research and Academic Network - GRNET and the National Technical University of Athens - NTUA). In order to conduct these experiments in a controlled fashion we developed a DDoS attack emulator. Through our analysis we identified several metrics that can serve as anomaly indicators and demonstrated that data fusion can improve detection performance. Finally we show that detecting traffic anomalies at the border routers of a network provider (Provider Edge routers) is feasible and effective in terms of countermeasures applicability.

Ευχαριστίες

Οι άνθρωποι που με βοήθησαν καθ' όλη τη διάρκεια της εκπόνησης της διδακτορικής μου διατριβής είναι σίγουρα πολλοί και προέρχονται τόσο από το ακαδημαϊκό και εργασιακό, όσο και από το φιλικό και συγγενικό μου περιβάλλον. Κατά την πολυετή αυτή προσπάθεια ήταν πολλές οι στιγμές που η υποστήριξη τους αποδείχθηκε ανεκτίμητη.

Ο επιβλέπων καθηγητής κ. Βασίλης Μάγκλαρης με στήριξε και με καθοδήγησε με σοφία μέχρι την ολοκλήρωση της. Η δική του παρότρυνση του ήταν καταλυτική και για την φοίτηση μου στο University of Southern California (U.S.C.) όπου και ξεκίνησα τη διδακτορική μου διατριβή. Τα πρώτα βήματα της ερευνητικής μου προσπάθειας έγιναν εκεί με την πολύτιμη καθοδήγηση του καθηγητή κ. Χρήστου Παπαδόπουλου. Η πορεία αυτή δεν θα είχε ολοκληρωθεί χωρίς τη συμπαράσταση και τη συνεργασία των επιστημονικών υπευθύνων, συναδέλφων και φίλων μου στο Κέντρο Δικτύων του Ε.Μ.Π. (ΚΕΔ). Εκεί, εκτός από ένα ευχάριστο και δημιουργικό εργασιακό περιβάλλον μου προσφέρθηκε και πρόσβαση σε δικτυακές υποδομές που αποδείχθηκε απαραίτητη για την διεξαγωγή της συγκεκριμένης έρευνας. Μια ακόμη σημαντική πτυχή της συμβολής των συναδέλφων-φίλων στο ΚΕΔ, στο Εργαστήριο Διαχείρισης & Βέλτιστου Σχεδιασμού Δικτύων (NETMODE) και στο Computer Networks and Distributed Systems Laboratory (U.S.C.) είναι ο γόνιμος διάλογος μέσα από τον οποίο προέκυψαν στοχευμένες, ακριβείς και ιδιαίτερα χρήσιμες παρατηρήσεις, οδηγίες, σχόλια και ιδέες. Για όλα αυτά τα μικρά και μεγάλα θέλω να ευχαριστήσω όλους αυτούς τους ανθρώπους, που ο περιορισμένος χώρος δε μου επιτρέπει να αναφέρω αναλυτικά.

Τέλος, θα ήθελα να ευχαριστήσω τη Μαρούσα, τις αδελφές μου Χριστίνα και Μαρία και τους γονείς μου Γιάννη και Μαρία για την αμέριστη συμπαράστασή τους όλα αυτά τα χρόνια. Η αγάπη και η συναισθηματική τους υποστήριξη ήταν απαραίτητη για να συνεχίσω μέχρι την ολοκλήρωση της διατριβής αυτής, παρά τα εμπόδια που εμφανίστηκαν. Για όλα αυτά και για όσα πολλά ακόμα προσφέρουν στη ζωή μου, τους αγαπώ και τους ευχαριστώ.

Περιεχόμενα

Κατάλογος Σχημάτων

Κατάλογος Πινάκων

Κεφάλαιο 1

Εισαγωγή

1.1 Περιγραφή του προβλήματος

Η παρούσα διδακτορική διατριβή πραγματεύεται την ανίχνευση ανωμαλιών δικτυακής κίνησης με έμφαση στην ανίχνευση κατανεμημένων επιθέσεων άρνησης υπηρεσιών (Distributed Denial of Service attacks -DDoS) που αποτελούν ένα από τα μεγαλύτερα προβλήματα που αντιμετωπίζει σήμερα το Διαδίκτυο (Internet) [1]. Οι επιθέσεις αυτές έχουν ως μόνο στόχο την "άρνηση υπηρεσιών," παρεμποδίζουν δηλαδή την πρόσβαση των χρηστών σε ηλεκτρονικές υπηρεσίες και συστήματα. Το φαινόμενο αυτό διαφέρει από τις εισβολές[1] σε υπολογιστικά συστήματα με την παραδοσιακή έννοια, καθώς σε εκείνες το κίνητρο της παραβίασης της ασφάλειας ενός συστήματος είναι ο έλεγχος του ή η απόκτηση κρίσιμων πληροφοριών. Στην περίπτωση των κατανεμημένων επιθέσεων άρνησης υπηρεσιών, οι επιτιθέμενοι εκμεταλλεύονται εγγενείς αδυναμίες των πρωτοκόλλων δικτύωσης TCP/IP, την ελεύθερη αποστολή πακέτων στο δίκτυο και την "δύναμη των πολλών." Αφού λοιπόν αποκτήσουν τον έλεγχο ενός μεγάλου συνήθως πλήθους υπολογιστών, τους χρησιμοποιούν για να εξαπολύσουν έναν καταιγισμό πακέτων (packet flooding), που κατευθυνόμενα προς έναν συγκεκριμένο προορισμό έχουν ως αποτέλεσμα την εξάντληση κά-

[1] στα αγγλικά χρησιμοποιείται ο όρος Intrusion

ποιου δικτυακού ή υπολογιστικού πόρου και κατ' επέκταση την υποβάθμιση της ποιότητας της παρεχόμενης υπηρεσίας. Παράδειγμα εξαντλούμενου πόρου είναι η χωρητικότητα (bandwidth) μίας δικτυακής σύνδεσης. Συνεπώς, ο όρος "επιθέσεις καταιγισμού πακέτων" μπορεί να χρησιμοποιηθεί αντί του "κατανεμημένες επιθέσεις άρνησης υπηρεσιών," αναδεικνύοντας την μέθοδο της επίθεσης και όχι την τοπολογία και τον στόχο των επιτιθέμενων. Η σημασία του προβλήματος των επιθέσεων καταιγισμού πακέτων φαίνεται από τις σημαντικότατες επιπτώσεις τους τόσο σε λειτουργικό όσο και σε οικονομικό επίπεδο. Όταν επενδύονται τεράστια χρηματικά ποσά στην ανάπτυξη ηλεκτρονικών υποδομών και υπηρεσιών δεν μπορεί να κινδυνεύει η εύρυθμη λειτουργία τους από τις ενέργειες ενός κακόβουλου χρήστη. Μάλιστα σύμφωνα με πρόσφατη μελέτη του CSI-FBI το 2004 [2] οι επιθέσεις άρνησης υπηρεσιών αν και αποτελούν μόνο το 17% του συνολικού αριθμού των περιστατικών ασφάλειας που καταγράφησαν, αντιπροσωπεύουν την δεύτερη κυριότερη πηγή οικονομικών ζημιών. Παράλληλα ολοένα και περισσότερες από τις καθημερινές μας δραστηριότητες βασίζονται στην καλή λειτουργία του Διαδικτύου, με αποτέλεσμα να απαιτούμε σταθερότητα και αξιοπιστία ανάλογη με αυτή του τηλεφωνικού δικτύου. Η επένδυση της βιομηχανίας στην επίλυση του προβλήματος είναι μάλιστα τόσο μεγάλη ώστε τα λιγοστά συστήματα που ισχυρίζονται ότι προσφέρουν προστασία από κατανεμημένες επιθέσεις άρνησης υπηρεσιών κοστίζουν πολλές χιλιάδες δολάρια [3–7].

Ένα εύλογο ερώτημα είναι πόσο εύκολα εξαπολύεται μια επίθεση άρνησης υπηρεσιών. Για να απαντήσουμε πρέπει να αναλογιστούμε ότι τα υπολογιστικά συστήματα είναι ολοένα και πιο ευάλωτα σε προβλήματα ασφάλειας λογισμικού, ιούς (viruses), σκουλήκια[2] (worms) και δούρειους ίππους (Trojan horses) [8, 9]. Εκμεταλλευόμενος κανείς τα προβλήματα αυτά μπορεί αρκετά εύκολα με αυτοματοποιημένο τρόπο να λάβει υπό τον έλεγχο του χιλιάδες υπολογιστές, δημιουργώντας έναν "ηλεκτρονικό στρατό"[3] [10, 11]. Η δημοσιο-

[2]αυτο-αναπαραγόμενα προγράμματα παρόμοια με ιούς που μεταδίδονται μέσω δικτύου.

[3]ελεύθερη μετάφραση του όρου botnet που προέρχεται από τις λέξεις ro-bot και net-work.

ποίηση κώδικα για την εκμετάλλευση των προβλημάτων ασφάλειας επιτρέπει ακόμη και σε μη τεχνικά καταρτισμένους χρήστες να προχωρήσουν σε τέτοιου είδους κακόβουλες ενέργειες και πολλαπλασιάζει τις πιθανότητες εκδήλωσης επιθέσεων [12]. Υπάρχουν λοιπόν αναφορές για την δημιουργία ηλεκτρονικών στρατών με χιλιάδες μέλη, έτοιμα να εξαπολύσουν κατανεμημένες επιθέσεις άρνησης υπηρεσιών με μία απλή εντολή [1, 10]. Συνεπώς είναι πιθανό να επαναληφθούν γεγονότα ανάλογα με μαζικές επιθέσεις του 2000 κατά τα οποία τέθηκαν εκτός λειτουργίας ένας μεγάλος αριθμός δικτυακών τόπων όπως το Yahoo, E-bay, Amazon και CNN [13].

Το θέμα της διατριβής είναι η ανίχνευση ανωμαλιών στη δικτυακή κίνηση και ειδικότερα των κατανεμημένων επιθέσεων άρνησης υπηρεσιών, ως μέσο για την ενίσχυση της ασφάλειας δικτύων. Στόχος μας είναι η ανάπτυξη συστημάτων ανίχνευσης ανωμαλιών στη δικτυακή κίνηση (Network Anomaly Detection Systems) σε αντιδιαστολή με τα συστήματα ανίχνευσης επιθέσεων (Intrusion Detection Systems-IDS) που εντοπίζουν συγκεκριμένες προσπάθειες παραβίασης υπολογιστικών συστημάτων. Προσανατολιζόμαστε στην ανίχνευση ανωμαλιών καθώς το Διαδίκτυο αποτελεί ένα συνεχώς εξελισσόμενο, "ζωντανό" σύστημα και οι ανωμαλίες στη δικτυακή κίνηση θα εξελίσσονται παράλληλα. Χαρακτηριστικό παράδειγμα είναι η εκρηκτική εξάπλωση των σκουληκιών (worms), που αν και ως φαινόμενο έχει τις ρίζες του στα πρώτα χρόνια του Διαδικτύου με το περίφημο Morris worm [14], τα τελευταία χρόνια αναδείχθηκε σε μία μεγάλη πληγή στην ασφάλεια των υπολογιστών. Καθώς λοιπόν αναμένουμε την εξέλιξη των φαινόμενων που προσπαθούμε να ανιχνεύσουμε, η προσέγγιση μας επιβάλλεται να είναι ευέλικτη και ανοικτή σε μελλοντικές επεκτάσεις. Αυτή η συνεχής εξέλιξη του Διαδικτύου και η εγγενής πολυπλοκότητα του ως συνδετικού ιστού ετερογενών δικτυακών συσκευών ανά την υφήλιο, αποτελούν τα βασικά αίτια για τη δυσκολία μοντελοποίησης της φυσιολογικής δικτυακής κίνησης. Το γεγονός αυτό δυσχεραίνει την ανάπτυξη αποτελεσματικών συστημάτων ανίχνευσης ανωμαλιών.

Παρόλο που υπάρχουν προληπτικά μέτρα εναντίον του φαινομένου των κατανεμημένων επιθέσεων άρνησης υπηρεσιών [15–17] αλλά και μέθοδοι καταστολής τους όταν αυτές εντοπιστούν [18–22], το πρόβλημα παραμένει ουσιαστικά άλυτο, καθώς απαιτείται σημαντική βελτίωση των συστημάτων ανίχνευσης προκειμένου να γίνουν αποτελεσματικά. Τα συνηθέστερα προβλήματα τους είναι το μεγάλο ποσοστό λανθασμένων εκτιμήσεων για την κανονικότητα της δικτυακής κίνησης[4] και η ανίχνευση επιθέσεων μόνο όταν τα περιθώρια αντιμετώπισης του προβλήματος είναι λιγοστά, όπως θα δούμε αναλυτικά στην ενότητα 2.3. Επίσης πολλά από τα προτεινόμενα συστήματα είναι ιδιαίτερα πολύπλοκα και τεχνικά δύσκολο να λειτουργήσουν σε πραγματικές συνθήκες, γεγονός που καταδεικνύει την ανάγκη μιας στενότερης συνεργασίας μεταξύ της ερευνητικής κοινότητας και των διαχειριστών δικτύου. Σήμερα η συνήθης πρακτική για την αντιμετώπιση κατανεμημένων επιθέσεων άρνησης υπηρεσιών από τους διαχειριστές δικτύων είναι η λήψη αντιμέτρων είτε κατόπιν ειδοποίησης των υπηρεσιών της Τεχνικής Στήριξης (helpdesk) από το θύμα, είτε αν εντοπιστεί σοβαρή δυσλειτουργία μιας δικτυακής συσκευής από τον διαχειριστή [23, 24]. Συνοψίζοντας, μπορούμε να πούμε ότι η έγκαιρη, αξιόπιστη, πρακτικά εφαρμόσιμη και μελλοντικά επεκτάσιμη ανίχνευση κατανεμημένων επιθέσεων άρνησης υπηρεσιών είναι ένας απαραίτητος συνδετικός κρίκος μεταξύ προληπτικών και κατασταλτικών μέτρων, που λείπει σήμερα από τον χώρο της ασφάλειας των δικτύων.

1.2 Προτεινόμενη λύση

Η ανίχνευση κατανεμημένων επιθέσεων άρνησης υπηρεσιών έχει συγκεντρώσει τα τελευταία χρόνια το ενδιαφέρον της ερευνητικής κοινότητας καθώς αποτελεί ένα δυσεπίλυτο πρόβλημα, κρίσιμης σημασίας για την ασφάλεια στο Διαδίκτυο. Το μεγάλο πλήθος των δημοσιεύσεων γύρω από το συγκεκρι-

[4]οι εσφαλμένες εκτιμήσεις είναι δύο τύπων: ανίχνευση μη υπαρκτής ανωμαλίας (false positive) και μη ανίχνευση υπαρκτής ανωμαλίας (false negative).

μένο θέμα είναι ενδεικτικό της δυσκολίας ανάπτυξης ενός αποτελεσματικού συστήματος ανίχνευσης επιθέσεων. Η ύπαρξη λιγοστών αναφορών όπως των Moore et al [25] και Hussain et al [26], για τα χαρακτηριστικά πραγματικών κατανεμημένων επιθέσεων άρνησης υπηρεσιών που συμβαίνουν σε σύγχρονα δίκτυα αποτελεί ένα ακόμη πρόβλημα και αποδεικνύει ότι οι γνώσεις μας για τα φαινόμενα αυτά είναι περιορισμένες. Οι υπάρχουσες προσεγγίσεις για την ανίχνευση κατανεμημένων επιθέσεων άρνησης υπηρεσιών μπορούν να ταξινομηθούν ανάλογα με το σημείο στο οποίο επιχειρούν την ανίχνευση [27], δηλαδή σε δίκτυα πελάτες (Customer Networks - CN), που μπορεί να είναι πηγές ή θύματα μιας επίθεσης, και σε δίκτυα παρόχων (Provider Networks - PN).

Η ανίχνευση στο δίκτυο του θύματος (CN) υπήρξε η πρώτη προσέγγιση. Σύμφωνα με αυτή, η ανίχνευση στην περιοχή του θύματος είναι πιο εύκολη καθώς εκεί τελικά συγκλίνουν τα πακέτα μιας επίθεσης προκαλώντας συμφόρηση (congestion) [19]. Καθώς όμως έχουν ήδη καταναλωθεί πολύτιμοι δικτυακοί πόροι απαιτείται συνήθως συνεργασία με το πάροχο δικτύου (upstream provider) για την εφαρμογή αντιμέτρων [28]. Αυτή η συνεργασία πελάτη - παρόχου είναι δύσκολο να επιτευχθεί για δύο κυρίως λόγους. Πρώτον, ο πάροχος επιβαρύνεται με μεγάλο διαχειριστικό κόστος για την εφαρμογή των απαραίτητων αντιμέτρων σε κάθε αίτημα των πελατών του. Δεύτερον, η συνεργασία βασίζεται στον ανθρώπινο παράγοντα, γεγονός που την καθιστά αργή και δύσκολα κλιμακούμενη σε περίπτωση μεγάλου πλήθους πελατών και συμβάντων.

Αντίθετα, προσπαθώντας να αντιμετωπίσουμε μια επίθεση στα δίκτυα των πηγών της (CN) [29, 30], όπου τα αντίμετρα είναι γρήγορα εφαρμόσιμα και αποτελεσματικά, ερχόμαστε αντιμέτωποι με το δύσκολο πρόβλημα της ανίχνευσης μιας επίθεσης μέσα στη φυσιολογική δικτυακή κίνηση. Μια επίθεση κατά τη γέννηση της συχνά δεν έχει αποκτήσει ικανό μέγεθος ώστε να την διαφοροποιήσει από την φυσιολογική κίνηση. Γενικότερα το πρόβλημα της ανίχνευσης ανωμαλιών είναι ιδιαίτερα δύσκολο λόγω των ευμετάβλητων χα-

ρακτηριστικών της δικτυακής κίνησης και της έλλειψης μοντελοποίησης της. Μάλιστα στα οπτικά δίκτυα υψηλών ταχυτήτων η λεπτομερής παρακολούθηση και ανάλυση της δικτυακής κίνησης είναι τεχνικά δύσκολη και τυχόν ανωμαλίες στη δικτυακή κίνηση μπορούν να περάσουν απαρατήρητες. Γενικά μπορούμε να πούμε ότι υπάρχει μια σχέση αντίστροφης αναλογίας ανάμεσα στην αποτελεσματικότητα της αντιμετώπισης των κατανεμημένων επιθέσεων άρνησης υπηρεσιών και στην ευκολία ανίχνευσης τους όπως θα αναλύσουμε στην ενότητα 2.3.

Η προσέγγιση, που υιοθετεί η παρούσα διατριβή, είναι η ανίχνευση ανωμαλιών στη δικτυακή κίνηση **στους ακραίους δρομολογητές του δικτύου ενός παρόχου** (Provider Edge - PE routers) προκειμένου να είναι δυνατή η γρήγορη και αποτελεσματική εφαρμογή αντιμέτρων. Το σημείο που θα επιλεγεί για να εγκατασταθεί ένα σύστημα ανίχνευσης και αντιμετώπισης πρέπει να διαθέτει ελεύθερο εύρος ζώνης (underutilized link), μεγάλη ισχύ δικτυακού εξοπλισμού (processing power in packets per second) και θέση κοντά στο δίκτυο του πελάτη (PE routers). Ειδικότερα προτείνεται μια αρχιτεκτονική ανίχνευσης ανωμαλιών που βασίζεται στις σημερινές πρακτικές διαχείρισης δικτύων και είναι επεκτάσιμη ώστε να μπορέσει να προσαρμοστεί στην αναμενόμενη εξέλιξη των χαρακτηριστικών της δικτυακής κίνησης και των ανωμαλιών της. Καινοτομία της προσέγγισης μας είναι αφενός η σύνθεση δεδομένων από πολλές πηγές (multisensor data fusion) [31, 32] που βελτιώνει την αποτελεσματικότητα της ανίχνευσης και ο διαχωρισμός της φάσης ανίχνευσης της επίθεσης από την φάση της ταυτοποίησης της. Η ανίχνευση μιας επίθεσης συνίσταται στην εξαγωγή ενός απλού συμπεράσματος ύπαρξης ή μη κάποιας ανωμαλίας (detection phase). Η ταυτοποίηση μιας επίθεσης αποτελεί ξεχωριστή διαδικασία με στόχο να προσδιορίσει τα χαρακτηριστικά μιας επίθεσης, π.χ. ο ορισμός των IP διευθύνσεων των πηγών μιας επίθεσης, ώστε να μπορούν στην συνέχεια να εφαρμοστούν αντίμετρα. Η κατάτμηση αυτή του προβλήματος αποσκοπεί στην ευκολότερη επίλυση του όπως αναλύεται στην

ενότητα 2.5.1.

Πιο αναλυτικά, η προτεινόμενη αρχιτεκτονική ορίζει τη χρήση αισθητήρων (sensors) που συλλέγουν δεδομένα για τη δικτυακή κίνηση με χρήση διάφορων τεχνικών παθητικής παρακολούθησης δικτύων (π.χ. μέσω συλλογής πακέτων-Packet Capturing, SNMP MIB's και τεχνολογίας Netflow[5]). Από τα δεδομένα αυτά απομονώνονται και μετρώνται συγκεκριμένα χαρακτηριστικά του υπό παρακολούθηση δικτύου (μετρικά-metric). Επιπρόσθετα κάθε αισθητήρας αποτελεί και μία στοιχειώδη μονάδα ανίχνευσης, επεξεργάζεται δηλαδή με έναν αλγόριθμο το καταγεγραμμένο μετρικό με σκοπό την ανίχνευση ανωμαλιών στη δικτυακή κίνηση. Στην απλούστερη μορφή τους οι αισθητήρες χρησιμοποιούν ως αλγόριθμο ανίχνευσης σταθερές ή προσαρμοζόμενες συναρτήσεις κατωφλίου. Τα αποτελέσματα της ανίχνευσης κάθε στοιχειώδους αισθητήρα συνδυάζονται με έναν αλγόριθμο σύνθεσης δεδομένων όπως είναι η θεωρία των Dempster-Shafer (D-S) [34], τα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks -ANN) [35, 36] κτλ. Η αρχιτεκτονική είναι αρκετά ανοικτή και επεκτάσιμη καθώς κάθε αισθητήρας λειτουργεί αυτόνομα χωρίς να προϋποθέτει την ύπαρξη ενός μοντέλου λειτουργίας του δικτύου. Στόχος μας είναι η σύνθεση των αποτελεσμάτων των στοιχειωδών αισθητήρων / ανιχνευτών που μπορεί να βασίζονται σε εντελώς διαφορετικά μετρικά αλλά και μεθοδολογίες ανίχνευσης. Συνεπώς μπορούν να συνδυαστούν διαφορετικές μεθοδολογίες όπως η ανίχνευση ανωμαλιών (anomaly detection) και η ανίχνευση κακής χρήσης (misuse detection).

Ο σχεδιασμός της προτεινόμενης αρχιτεκτονικής ανίχνευσης ανωμαλιών έγινε μετά από ενδελεχή διερεύνηση ενός πλήθους ερευνητικών ερωτημάτων. Κρίσιμα ερωτήματα που θέτει η παρούσα διατριβή είναι: Ποια χαρακτηριστικά του δικτύου (μετρικά) πρέπει να παρακολουθήσουμε και με ποιον τρόπο; Ποιοι αλγόριθμοι μπορούν να χρησιμοποιηθούν για ανίχνευση; Πως μπορούμε να ελαχιστοποιήσουμε την παραμετροποίηση των αισθητήρων μας αλλά πα-

[5]τεχνολογία για την καταγραφή και παρακολούθηση της δικτυακής κίνησης [33]

ράλληλα να εκμεταλλευτούμε την προϋπάρχουσα γνώση (expert knowledge); Υπάρχει βελτίωση της απόδοσης ανίχνευσης με την χρήση πολλαπλών μετρικών; Με τι κριτήρια πρέπει να γίνει επιλογή του αλγόριθμου σύνθεσης δεδομένων που χρησιμοποιούμε; Μέσα την διερεύνηση των θεμάτων αυτών μπορούμε να δώσουμε μια σαφή εικόνα για το πρόβλημα της ανίχνευσης ανωμαλιών και μια οριστική θετική απάντηση για την δυνατότητα ανάπτυξης αποτελεσματικών συστημάτων ανίχνευσης ανωμαλιών που θα υποστηρίζουν αυτοματοποιημένες διαδικασίες αντιμετώπισης τους. Τέλος, η προτεινόμενη αρχιτεκτονική ενοποιεί τις δύο προσεγγίσεις ανίχνευσης ανωμαλιών κοντά στην πηγή και κοντά στο θύμα σε ένα σύστημα που μπορεί να εντοπίζει ανωμαλίες σε σημεία όπου δεν έχουν εξαντληθεί ακόμα οι δικτυακοί πόροι και υπάρχουν περιθώρια τοπικής αντιμετώπισης. Με τον τρόπο αυτό έχουμε δυνατότητα γρήγορης και αποτελεσματικής επέμβασης δίχως να απαιτείται συνεργασία μεταξύ διαφορετικών διαχειριστικών τομέων (administrative domains).

1.3 Υλοποίηση της προτεινόμενης λύσης και πειραματικά αποτελέσματα

Στα πλαίσια της παρούσας διατριβής υλοποιήθηκε ένα σύστημα ανίχνευσης κατανεμημένων επιθέσεων άρνησης υπηρεσιών σύμφωνα με την προτεινόμενη αρχιτεκτονική. Η υλοποίηση περιλαμβάνει την ανάπτυξη λογισμικού παρακολούθησης δικτύου, υπολογισμού μετρικών, ανίχνευσης και σύνθεσης δεδομένων. Παράλληλα αναπτύχθηκε εργαλείο εξομοίωσης επιθέσεων άρνησης υπηρεσιών με ελεγχόμενα χαρακτηριστικά. Με τον τρόπο αυτό διεξήχθησαν πειράματα σε περιβάλλον πραγματικού δικτύου υψηλών ταχυτήτων (στη γραμμή σύνδεσης Εθνικού Μετσόβιου Πολυτεχνείου (Ε.Μ.Π.) και του Εθνικού Δικτύου Έρευνας και Τεχνολογίας (Ε.Δ.Ε.Τ./GRNET) ταχύτητας 1Gbps). Η ανάλυση των πειραματικών αποτελεσμάτων ανέδειξε χρήσιμα μετρικά [37] για την ανίχνευση δικτυακών ανωμαλιών και απέδειξε την αποτε-

λεσματικότητα της σύνθεσης δεδομένων [38]. Συγκεκριμένα καταδεικνύουμε πως η σύνθεση δεδομένων από αισθητήρες που βασίζονται σε διαφορετικά μετρικά, μπορεί να βελτιώσει την απόδοση ενός συστήματος ανίχνευσης. Το σύστημα που αναπτύχθηκε επιβεβαιώνει και πρακτικά, ότι η προσέγγιση μας για ανίχνευση ανωμαλιών στους ακραίους δρομολογητές του δικτύου ενός παρόχου (PE routers) είναι εφικτή. Το αποτέλεσμα αυτό είναι ενθαρρυντικό για την δυνατότητα ανάπτυξης υπηρεσιών ανίχνευσης και αντιμετώπισης ανωμαλιών δικτυακής κίνησης από παρόχους δικτύου. Η προτεινόμενη αρχιτεκτονική έχει υιοθετηθεί ως αρχικό σχέδιο (blueprint) από την ερευνητική ομάδα εργασίας για την ασφάλεια δικτύων του πανευρωπαϊκού ακαδημαϊκού δικτύου Geant-2 για την κατασκευή ενός συνόλου εργαλείων για ανίχνευση περιστατικών ασφάλειας [39].

1.4 Παρουσίαση περιεχομένων

Τα κεφάλαια που ακολουθούν παρουσιάζουν αναλυτικά τις διαφορετικές πτυχές του προβλήματος των κατανεμημένων επιθέσεων άρνησης υπηρεσιών και της προτεινόμενης αρχιτεκτονικής ανίχνευσης ανωμαλιών. Αρχικά στο κεφάλαιο 2 γίνεται μια εισαγωγή στο πρόβλημα των κατανεμημένων επιθέσεων άρνησης υπηρεσιών μέσα από μια ιστορική αναδρομή της εξέλιξης τους. Παρουσιάζονται επίσης όλες οι συνιστώσες αντιμετώπισης τους: τα πιθανά σημεία άμυνας, τα προληπτικά και τα κατασταλτικά μέτρα (ανίχνευση και μέθοδοι καταστολής). Στο κεφάλαιο αυτό ορίζεται η προσέγγιση μας για τη ανάπτυξη ενός συστήματος ανίχνευσης ανωμαλιών με δύο φάσεις: την ανίχνευση με σύνθεση πολλαπλών μετρικών και την ταυτοποίηση. Στην συνέχεια στο κεφάλαιο 3 παρουσιάζεται το θεωρητικό υπόβαθρο ανάπτυξης συστημάτων σύνθεσης δεδομένων και αναλύονται τα κριτήρια επιλογής αλγορίθμων σύνθεσης δεδομένων για το πρόβλημα της ανίχνευσης ανωμαλιών δικτυακής κίνησης. Το κεφάλαιο 4 αποτελεί την παρουσίαση της προτεινόμενης αρχιτεκτονικής και καλύπτει τα θέματα της παρακολούθησης δικτύου και επιλογής

μετρικών, αλγόριθμων ανίχνευσης και σύνθεσης. Το κεφάλαιο συμπληρώνεται με την ανάλυση της φάσης ταυτοποίησης επιθέσεων. Το κεφάλαιο 5 περιέχει τις λεπτομέρειες υλοποίησης του πρότυπου συστήματος που αναπτύχθηκε και πειραματικά αποτελέσματα για την αξιολόγηση των επιδόσεων του. Μετά την παρουσίαση του πρότυπου συστήματος γίνεται σύγκριση του με άλλες προσεγγίσεις που έχουν προταθεί στην βιβλιογραφία. Η διατριβή ολοκληρώνεται με το κεφάλαιο 6 στο οποίο συνοψίζονται τα συμπεράσματα και υποδεικνύονται ανοικτά ερευνητικά θέματα και μελλοντικές βελτιώσεις της προτεινόμενης αρχιτεκτονικής.

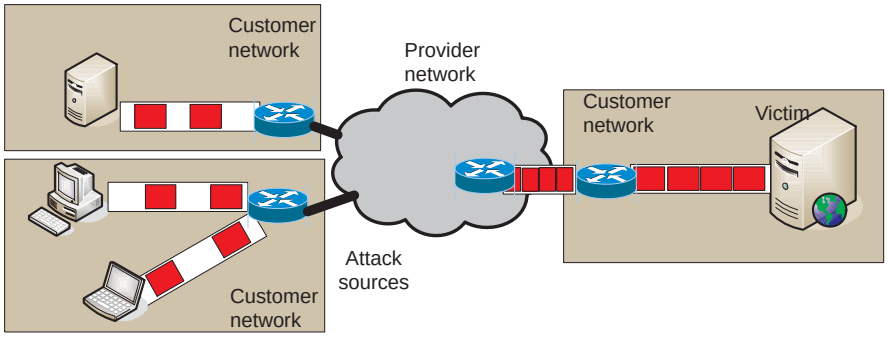
Κεφάλαιο 2

Αντιμετώπιση κατανεμημένων επιθέσεων άρνησης υπηρεσιών

2.1 Εισαγωγή

Η ασφάλεια υπολογιστικών συστημάτων και δικτύων είναι ένας τομέας με έντονη δραστηριότητα. Συνεχώς εμφανίζονται νέοι τρόποι παραβίασης της ασφάλειας ενώ παράλληλα εξελίσσονται και οι μέθοδοι άμυνας απέναντι στις επιθέσεις αυτές. Ένα βασικό στοιχείο της ασφάλειας δικτύων, όπως προσδιορίζεται από σχετική μελέτη της National Security Agency (NSA) των Η.Π.Α. [40,41], είναι η προστασία από επιθέσεις άρνησης υπηρεσιών (**Denial of Service attacks - DoS**). Οι επιθέσεις με στόχο την άρνηση υπηρεσιών έχουν διαφορετικά χαρακτηριστικά από τις κλασσικές επιθέσεις εισβολής (**intrusion**) σε ένα υπολογιστικό σύστημα. Κατά την εισβολή σε ένα σύστημα πλήγεται συνήθως η εμπιστευτικότητα (data confidentiality) ή/και η ακεραιότητα των δεδομένων και επικοινωνιών (data and communications integrity). Στην περίπτωση της άρνησης υπηρεσιών όπως προκύπτει και από τη λεκτική περιγραφή, στόχος είναι η παύση ή μείωση της ποιότητας μιας παρεχόμενης ηλεκτρονικής υπηρεσίας (availability). Οι υπηρεσίες αυτές κυμαίνονται από πολύ απλές, όπως η περίπτωση της απομακρυσμένης πρόσβασης σε ένα

σύστημα μέσω του πρωτοκόλλου Telnet, έως σύνθετες όπως μία υπηρεσία e-banking. Ένα δεύτερο χαρακτηριστικό που σε πολλές περιπτώσεις διαφοροποιεί ποιοτικά τις εισβολές από τις επιθέσεις άρνησης υπηρεσιών είναι η μέθοδος επίτευξης του στόχου. Στην πρώτη περίπτωση έχουμε εκμετάλλευση αδυναμιών λογισμικού ή ρυθμίσεων ενός συστήματος, ενώ στην δεύτερη οι επιτιθέμενοι συνήθως προκαλούν και εκμεταλλεύονται την εξάντληση κάποιου περιορισμένου πόρου σε συνδυασμό με εγγενείς αδυναμίες των πρωτοκόλλων δικτύωσης TCP/IP. Παραδείγματα εξαντλούμενων πόρων είναι η χωρητικότητα μίας γραμμής, η ισχύς ενός επεξεργαστή και η διαθέσιμη μνήμη ενός συστήματος ή μιας διεργασίας (process). Ο δημοφιλέστερος τύπος επιθέσεων άρνησης υπηρεσιών είναι σήμερα οι κατανεμημένες επιθέσεις άρνησης υπηρεσιών (**Distributed Denial of Service attacks - DDoS**). Ένα τυπικό σενάριο μιας επίθεσης DDoS παρουσιάζεται στο σχήμα 2.1 και περιλαμβάνει την αποστολή ενός καταιγισμού πακέτων (packet flooding) από μεγάλο πλήθος υπολογιστών, τα οποία όταν συγκεντρώνονται κοντά στο δίκτυο του θύματος προκαλούν συμφόρηση (congestion). Μέσα από την διαδικασία αυτή επιτυγχάνεται η υποβάθμιση της ποιότητας ή η πλήρης διακοπή των παρεχόμενων υπηρεσιών του θύματος. Συνεπώς, ο όρος "επιθέσεις καταιγισμού πακέτων" (packet flooding attacks) αποδίδει καλύτερα την μέθοδο των επιθέσεων DDoS και στην συνέχεια θα τον χρησιμοποιούμε εναλλακτικά, παρόλο που στη βιβλιογραφία συναντάται σπανιότερα. Καθώς οι κατανεμημένες επιθέσεις άρνησης υπηρεσιών αποσκοπούν στη διακοπή μιας παρεχόμενης υπηρεσίας, ο στόχος δεν είναι πάντα ένα συγκεκριμένο υπολογιστικό σύστημα. Στόχος μπορεί να είναι ένα σύνολο συστημάτων, ο δικτυακός εξοπλισμός που υποστηρίζει την πρόσβαση σε αυτά, ή βασικές υπηρεσίες δικτύου όπως η υπηρεσία ονοματολογίας (Domain Name Service -DNS). Έχοντας κάνει λοιπόν τον διαχωρισμό μεταξύ κατανεμημένων επιθέσεων άρνησης υπηρεσιών και εισβολών, είναι προφανής η διαφορά μεταξύ συστημάτων ανίχνευσης επιθέσεων-εισβολών (Intrusion Detection Systems - **IDS**) και συστημάτων ανίχνευσης κατανεμη-



Σχήμα 2.1: Χαρακτηριστικό σενάριο κατανεμημένης επίθεσης άρνησης υπηρεσίας, όπου το δίκτυο του θύματος αδυνατεί να αντιμετωπίσει την επίθεση χωρίς την επέμβαση του παρόχου.

μένων επιθέσεων άρνησης υπηρεσιών. Μπορούμε να χρησιμοποιούμε λοιπόν τον όρο **Network Anomaly Detection Systems - NADS** για να περιγράψουμε συστήματα που ανιχνεύουν επιθέσεις DDoS και γενικότερα ανωμαλίες στη δικτυακή κίνηση. Χρησιμοποιώντας τους όρους που εμφανίζονται στην βιβλιογραφία για την ταξινόμηση συστημάτων ανίχνευσης επιθέσεων, τα συστήματα ανίχνευσης των ανωμαλιών της δικτυακής κίνησης (NADS) αντλούν στοιχεία από την παρακολούθηση του δικτύου, δηλαδή είναι network-based και όχι host-based. Σε αντίθετη περίπτωση θα εντόπιζαν επιθέσεις DDoS σε επίπεδο τερματικών σταθμών (host), δηλαδή των θυμάτων, με ελάχιστα πλέον περιθώρια αντίδρασης. Τα NADS ακολουθούν κατά κύριο λόγο την προσέγγιση ανίχνευσης ανωμαλιών[1] (anomaly detection) και όχι την προσέγγιση ανίχνευσης κακής χρήσης (misuse detection). Στο πρώτο τμήμα του κεφαλαίου και συγκεκριμένα στη παράγραφο 2.2 παρουσιάζεται η εξέλιξη του φαινομένου των κατανεμημένων επιθέσεων άρνησης υπηρεσιών μέσα από μια σύντομη ιστορική αναδρομή. Στην παράγραφο 2.3 ακολουθεί μία ανάλυση των πιθανών σημείων άμυνας από τις επιθέσεις καταιγισμού πακέτων, που βοηθά την κατανόηση κρίσιμων παραμέτρων του προβλήματος της αντιμετώπισης των

[1]Σύμφωνα με την προσέγγιση ανίχνευσης ανωμαλιών εντοπίζονται αποκλίσεις από κάποιο φυσιολογικό πρότυπο - προφίλ, ενώ στην ανίχνευση κακής χρήσης εντοπίζονται συγκεκριμένα ιδιαίτερα χαρακτηριστικά (signatures) των επιθέσεων.

κατανεμημένων επιθέσεων άρνησης υπηρεσιών. Η παράγραφος 2.4 απαριθμεί τα προληπτικά μέτρα που μπορούμε να λάβουμε, ενώ στην παράγραφο 2.5.1 εστιάζουμε στην ανάλυση του προβλήματος της ανίχνευσης, που αποτελεί και το βασικό θέμα της διατριβής. Το κεφάλαιο ολοκληρώνεται στην παράγραφο 2.5.2 με την παρουσίαση των μεθόδων καταστολής που μπορούν να εφαρμοστούν μετά από την επιτυχημένη ανίχνευση μιας επίθεσης.

2.2 Ιστορική αναδρομή

Το φαινόμενο των κατανεμημένων επιθέσεων άρνησης υπηρεσιών, αν και η εξέλιξη του στην μορφή που το γνωρίζουμε σήμερα ξεκίνησε το 1998, έχει τις ρίζες του αρκετά χρόνια νωρίτερα. Πριν από το 1998 οι επιθέσεις άρνησης υπηρεσιών (Denial of Service attacks - **DoS**) γίνονταν σε μικρή κλίμακα και ήταν άμεσες επιθέσεις ενός επιτιθέμενου σε ένα θύμα (direct, 1-tier attacks). Οι πρώτες επιθέσεις αξιοποιούσαν αδυναμίες λογισμικού για την επίτευξη του στόχου τους και χαρακτηριστικά παραδείγματα είναι οι επιθέσεις TearDrop, Land και Ping of Death [42,43]. Ένας χρήστης μπορεί εύκολα να προστατευτεί από τις επιθέσεις αυτές αντικαθιστώντας το λογισμικό που χρησιμοποιεί με μια μη ευάλωτη έκδοση του. Πέρα από τις περιπτώσεις αυτές, υπάρχουν δείγματα επιθέσεων που χρονολογούνται πριν το 1996 και δεν εκμεταλλεύονται προβλήματα στην υλοποίηση ενός πρωτοκόλλου αλλά εγγενείς αδυναμίες του Διαδικτύου. Οι αναφορές του Computer Emergency Response Team των Η.Π.Α. (CERT) το 1996 για επιθέσεις τύπου "TCP SYN" και "UDP packet storm" είναι ενδεικτικές για την ιστορία των επιθέσεων καταιγισμού πακέτων [44, 45]. Αυτές είναι και οι πρώτες επιθέσεις άρνησης υπηρεσιών με καταιγισμό πακέτων. Στην περίπτωση μαζικής αποστολής πακέτων TCP με SYN flag εξαντλούνται πόροι του λειτουργικού συστήματος ενώ στην δεύτερη περίπτωση αποστολής UDP πακέτων μεγάλου μεγέθους έχουμε εξάντληση της χωρητικότητας (bandwidth) μιας δικτυακής σύνδεσης.

Οι πρώτες επιθέσεις άρνησης υπηρεσιών που δεν ήταν άμεσες αλλά περιε-

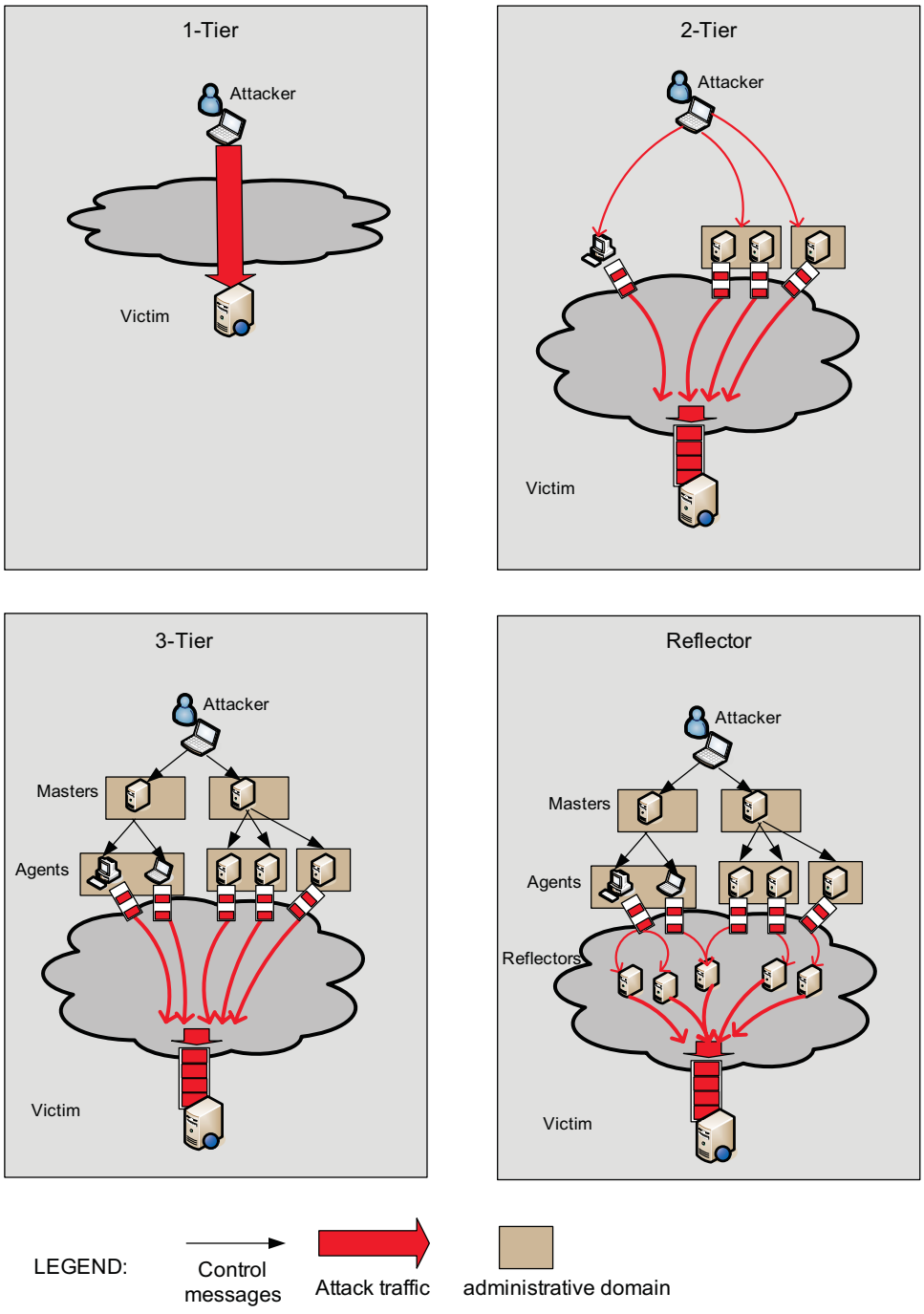
λάμβαναν την χρήση και ενδιάμεσων συστημάτων πέρα του επιτιθέμενου και του θύματος (2-tier attacks) εμφανίστηκαν περίπου το 1998 με τα ονόματα smurf και fraggle attacks [46]. Και οι δύο τύποι επιθέσεων χρησιμοποιούσαν την τεχνική αποστολής παραποιημένων πακέτων (**spoofing**). Σύμφωνα με αυτή ένα σύστημα μπορεί να αποστείλει IP πακέτα στο Internet τοποθετώντας στο πεδίο της διεύθυνσης αποστολέα (source address) μια τυχαία διεύθυνση και όχι την πραγματική του IP διεύθυνση. Το σημαντικό αυτό πρόβλημα της δυνατότητας αποστολής παραποιημένων πακέτων ήταν γνωστό ήδη από το 1989 μέσα από το άρθρο "Security Problems in the TCP/IP Protocol Suite" του S.M. Bellovin [47] και συνεχίζει μέχρι σήμερα να επηρεάζει καθοριστικά την ασφάλεια στο Διαδίκτυο.

Στην εικόνα 2.2 περιγράφονται σχηματικά οι αρχιτεκτονικές εξαπόλυσης επιθέσεων DDoS όπως αυτές εξελίχτηκαν από τα μέσα της δεκαετίας του 90 μέχρι σήμερα.

Τα πρώτα προγράμματα κατανεμημένων επιθέσεων άρνησης υπηρεσιών (Distributed Denial of Service attacks - **DDoS**) υπολογίζεται ότι άρχισαν να κυκλοφορούν σε διάφορες κλειστές ομάδες μέσα στο 1998. Οι επιθέσεις τύπου DDoS αποτελούν μια εξέλιξη των επιθέσεων DoS. Χρησιμοποιούν μια αρχιτεκτονική τριών στρωμάτων (3-tier attacks) όπου ο επιτιθέμενος (attacker or client) δίνει εντολή για επίθεση σε μικρό αριθμό αντιπροσώπων του (handler or master), που με την σειρά τους την προωθούν σε μεγάλο αριθμό παραβιασμένων συστημάτων (agents). Στην συνέχεια οι agents εξαπολύουν μια επίθεση καταιγισμού πακέτων προς το θύμα. Οι handler και οι agents είναι ουσιαστικά λογισμικό που έχει εγκατασταθεί σε παραβιασμένα συστήματα και συνθέτουν ένα δίκτυο ελέγχου μεγάλου πλήθους παραβιασμένων συστημάτων. Με τη χρήση του δικτύου αυτού ο επιτιθέμενος μπορεί να κρύψει τα ίχνη του αλλά και να στείλει εύκολα και γρήγορα εντολή για επίθεση με ένα μόνο μήνυμα που προωθείται αυτόματα μέσα από τους handlers στους agents. Το χρονικό διάστημα 1998 έως 2000 τα προγράμματα επιθέσεων DDoS ωρίμασαν προσφέρο-

ντας κρυπτογραφημένα κανάλια επικοινωνίας μεταξύ των client, handlers και agents αλλά και μεγαλύτερο έλεγχο στον τύπο των αποστελλόμενων πακέτων και στην χρήση της τεχνικής spoofing. Τα δημοφιλέστερα προγράμματα είναι: Trin00, TFN, TFN2k, mstream, Stacheldraht. Τα εργαλεία αυτά μπορούν να εξαπολύσουν διάφορες επιθέσεις καταιγισμού πακέτων, π.χ. **TCP-flood, SYN-flood, UDP-flood και ICMP-flood**. Τα ονόματα τους αντιστοιχούν στον τύπο των αποστελλόμενων πακέτων, δηλαδή πακέτα TCP, TCP με SYN flag, UDP και ICMP. Στο διάστημα αυτό το φαινόμενο των επιθέσεων DDoS γενικεύτηκε και ύστερα από μια σειρά σημαντικών γεγονότων έγινε ευρύτατα γνωστό μέσα από τα Μέσα Μαζικής Ενημέρωσης όλου του πλανήτη. Ενδεικτικά αναφέρουμε κάποια περιστατικά, ξεκινώντας από τις μαζικές επιθέσεις του 2000 εναντίων πολλών εμπορικών ιστοσελίδων όπως του Yahoo, του E-bay και του CNN [13]. Το 2001 μετά από ένα περιστατικό εμπλοκής πολεμικών αεροσκαφών Κίνας και Η.Π.Α. εξαπολύθηκε μεγάλο κύμα επιθέσεων DDoS σε αμερικανικά δίκτυα, συμπεριλαμβανομένου του Λευκού Οίκου [48]. Τα κίνητρα για τις επιθέσεις DDoS όπως αποδεικνύεται από τα ίδια τα γεγονότα ποικίλουν. Το οικονομικό όφελος, που προέρχεται από δραστηριότητες όπως το spam, οδήγησε για παράδειγμα στην εξαπόλυση συνεχόμενων επιθέσεων ενάντια στην εταιρία Osirusoft που πρόσφερε υπηρεσίες blacklisting ενάντια στο spam. Οι ζημιές ήταν τόσο μεγάλες που η εταιρία έκλεισε [49]. Σε άλλη περίπτωση, εξαπολύθηκαν επιθέσεις εναντίον του δικτύου ενημέρωσης "AlJazeera" [50] (2003). Ένα από τα σημαντικότερα γεγονότα (2002) ήταν η εξαπόλυση επίθεσης εναντίον της ίδιας της υποδομής του Διαδικτύου και συγκεκριμένα των κεντρικών εξυπηρετητών ονοματολογίας, δηλαδή των 13 root name servers [51]. Σε αυτή την περίπτωση οι επιτιθέμενοι διερευνούσαν το μέγεθος της δύναμης που είχαν στα χέρια τους. Τα περιστατικά αυτά αποδεικνύουν ότι οι κατανεμημένες επιθέσεις άρνησης υπηρεσιών είναι ένα σημαντικό όπλο ηλεκτρονικού πολέμου.

Ένας άλλος τύπος επιθέσεων που ονομάστηκε **DDoS reflector** attacks

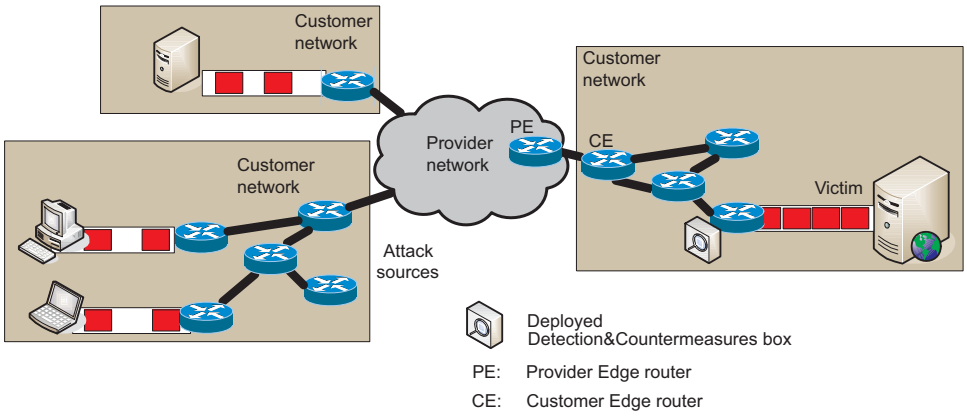


Σχήμα 2.2: Αρχιτεκτονικές οργάνωσης υπολογιστών για εξαπόλυση επιθέσεων DDoS.

[52] εμφανίστηκε το 2001. Οι επιθέσεις αυτές βασίζονται στην αποστολή παραποιημένων πακέτων (spoofing) προς δικτυακούς κόμβους με μεγάλες ταχύτητες διασύνδεσης και μεγάλη υπολογιστική ισχύ. Οι κόμβοι αυτοί προσπαθώντας να απαντήσουν στα παραποιημένα πακέτα [2] δημιουργούν ένα καταιγισμό πακέτων προς το θύμα. Ένα σημαντικό στοιχείο αυτού του τύπου επιθέσεων είναι ότι χρησιμοποιούνται ως πηγές της επίθεσης συστήματα που δεν έχουν παραβιαστεί.

Η τελευταία εξέλιξη στον χώρο των επιθέσεων DDoS είναι η χρήση των worms για την μαζική και αυτοματοποιημένη παραβίαση συστημάτων με παράλληλη εγκατάσταση μηχανισμών εξαπόλυσης επιθέσεων καταιγισμού πακέτων [53, 54]. Ο έλεγχος και η οργάνωση των παραβιασμένων συστημάτων γίνεται πλέον με την αξιοποίηση των πρωτοκόλλων επικοινωνίας του Internet Relay Chat (IRC) και κλιμακώνεται εύκολα σε μεγάλο πλήθος συστημάτων. Τα παραβιασμένα συστήματα αντί να περιμένουν παθητικά εντολές (σε ένα listening port) συνδέονται ενεργά σε προκαθορισμένα IRC κανάλια [3] μέσα από τα οποία λαμβάνουν εντολές. Με τον τρόπο αυτό σχηματίζεται ένα δίκτυο που αναφέρεται ως **botnet** (από τις λέξεις ro-bot και net-work). Σε ελεύθερη μετάφραση ένα botnet αποτελεί έναν "ηλεκτρονικό στρατό" [10, 11], έτοιμο να εξαπολύσει επιθέσεις. Τα botnets δεν χρησιμοποιούνται αποκλειστικά για DDoS επιθέσεις αλλά και για άλλες παράνομες δραστηριότητες όπως η μαζική αποστολή μηνυμάτων ηλεκτρονικού ταχυδρομείου (spam) [55] και η λειτουργία ιστοσελίδων (δικτυακών τόπων) με παράνομο περιεχόμενο. Η σημερινή κατάσταση πήρε αυτές τις διαστάσεις λόγω της πλημμελούς διαχείρισης ασφάλειας των συστημάτων που συνδέονται στο Διαδίκτυο, της κατανεμημένης διαχείρισης συστημάτων σε παγκόσμια κλίμακα, των εγγενών αδυναμιών των πρωτοκόλλων δικτύωσης και της δυσκολίας εντοπισμού των υπευθύνων για περιστατικά ασφάλειας. Είναι δεδομένο ότι τα φαινόμενα των επιθέσεων DDoS, τα botnets, τα worms, το spam κ.α. δεν είναι ανεξάρτητα και η κα-

[2] π.χ. απαντώντας σε ένα TCP SYN πακέτο με περισσότερα TCP SYN-ACK πακέτα.
[3] ξεπερνώντας με τον τρόπο αυτό εγκατεστημένα firewall



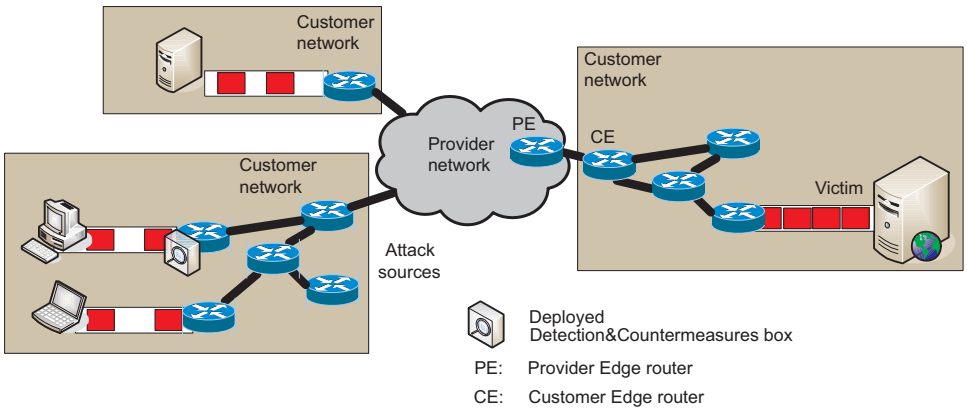
Σχήμα 2.3: Αντιμετώπιση επιθέσεων καταιγισμού πακέτων στο δίκτυο του θύματος

ταπολέμηση καθενός από αυτά δρα θετικά στο γενικότερο επίπεδο ασφάλειας στο Διαδίκτυο. Ένας σημαντικός κλάδος της διαχείρισης ασφάλειας στο Διαδίκτυο είναι λοιπόν η αντιμετώπιση των επιθέσεων DDoS και γενικότερα των ανωμαλιών στη δικτυακή κίνηση. Για την επίτευξη του στόχου αυτού καθίσταται απαραίτητη η ανάπτυξη αποτελεσματικών συστημάτων ανίχνευσης ανωμαλιών στη δικτυακή κίνηση.

2.3 Σημεία ανάπτυξης μηχανισμών άμυνας απέναντι σε επιθέσεις DDoS

Μετά από την σύντομη επισκόπηση της εξέλιξης των επιθέσεων καταιγισμού πακέτων προχωράμε στην παρουσίαση των πιθανών σημείων στα οποία μπορούν να αναπτυχθούν μηχανισμοί άμυνας. Οι πρώτες ιστορικά προσεγγίσεις πρότειναν **ανίχνευση και αντιμετώπιση επιθέσεων DDoS στο δίκτυο του θύματος (Customer Network - CN)** μιας επίθεσης. Η ανίχνευση μιας επίθεσης στην περιοχή του θύματος (σχήμα 2.3) είναι αρκετά απλή καθώς εκεί συγκεντρώνονται τα πακέτα των επιτιθέμενων, προκαλώντας ασυνήθιστα μεγάλη δικτυακή κίνηση και συμφόρηση (congestion) [19]. Μια τέτοια δραματική αλλαγή στο προφίλ της κίνησης εντοπίζεται εύκολα ακόμη και με χρήση των

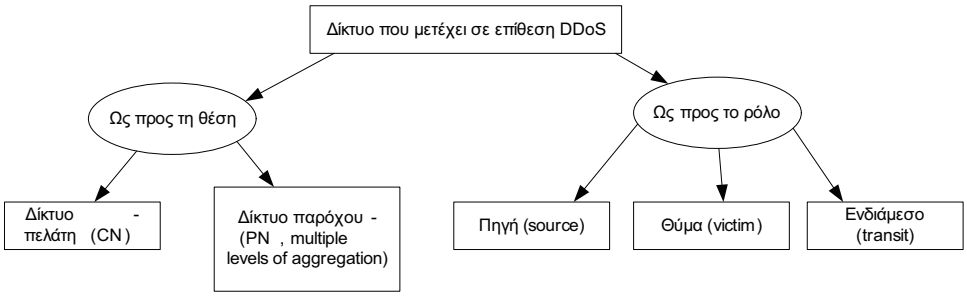
υπαρχόντων τεχνικών παρακολούθησης δικτύων, όπως η μέτρηση της χρησιμοποίησης (utilization) μιας γραμμής. Στην περίπτωση αυτή όμως έχουν ήδη καταναλωθεί πολύτιμοι δικτυακοί πόροι και οι δυνατότητες αντιμετώπισης της επίθεσης είναι περιορισμένες. Προκειμένου να αντιμετωπιστεί αποτελεσματικά η επίθεση πρέπει να ληφθούν κατασταλτικά μέτρα σε κάποιο σημείο όπου δεν έχουν ακόμα εξαντληθεί οι δικτυακοί πόροι. Δυστυχώς είναι σύνηθες το δίκτυο του θύματος να παρακολουθεί και να ανιχνεύει επιθέσεις στον ακραίο δρομολογητή του (**Customer Edge router - CE**) και σε περίπτωση μιας επίθεσης να εξαντλούνται οι δικτυακοί πόροι της γραμμής διασύνδεσης με τον **πάροχο δικτύου (Provider Network - PN)**. Καθώς ο διαχειριστής του δικτύου του θύματος, που ανιχνεύει την επίθεση, δεν μπορεί να αποτρέψει τη συμφόρηση της γραμμής διασύνδεσης, η μοναδική του διέξοδος είναι να ζητήσει τη λήψη κατασταλτικών μέτρων από τον πάροχο στον ακραίο δρομολογητή του παρόχου (**Provider Edge router - PE**). Το πρόβλημα αυτό απεικονίζεται γραφικά στο σχήμα 2.1. Η αντιμετώπιση όμως του προβλήματος με τέτοιου είδους συνεργασία πελάτη - παρόχου είναι αναποτελεσματική για δύο κυρίως λόγους. Πρώτον, υπάρχει μεγάλο διαχειριστικό κόστος από την πλευρά του παρόχου για την εφαρμογή των απαραίτητων αντιμέτρων για κάθε πελάτη σε κάθε ανάλογο συμβάν. Δεύτερον, όσο η συνεργασία δεν είναι αυτοματοποιημένη και βασίζεται στον ανθρώπινο παράγοντα, είναι αργή και δύσκολα κλιμακούμενη σε μεγάλο πλήθος πελατών και συμβάντων. Τα προβλήματα αυτά ίσως να επιλυθούν με την ανάπτυξη αυτοματοποιημένων μηχανισμών συνεργασίας μεταξύ δικτύων [28]. Η κατανεμημένη διαχείριση του Διαδικτύου και η έλλειψη εμπιστοσύνης μεταξύ των διαχειριστών, που έχει ως αποτέλεσμα να μην κοινοποιούνται τα περιστατικά παραβίασης της ασφάλειας και ο τρόπος χειρισμού τους, αποτελούν σε κάθε περίπτωση σημαντικά εμπόδια. Η προσέγγιση λοιπόν της **ανίχνευσης και αντιμετώπισης επιθέσεων στο δίκτυο του θύματος (Victim CN)** απαιτεί σε πολλές περιπτώσεις την επίλυση του προβλήματος της συνεργασίας μεταξύ διαχειριστών δικτύου.



Σχήμα 2.4: Αντιμετώπιση επιθέσεων καταιγισμού πακέτων στο δίκτυο της πηγής

Μια δεύτερη προσέγγιση, που εικονίζεται στο σχήμα 2.4, είναι η **ανίχνευση και αντιμετώπιση μιας επίθεσης DDoS στο δίκτυο πηγή μιας επίθεσης** (DDoS source CN) [29]. Τα πλεονεκτήματα είναι ότι τα αντίμετρα είναι γρήγορα εφαρμόσιμα και αποτελεσματικά, καθώς η πηγή του προβλήματος -συνήθως παραβιασμένα συστήματα- βρίσκονται υπό τον έλεγχο του διαχειριστή που διέγνωσε το πρόβλημα. Επίσης η δικτυακή κίνηση στην πηγή (before aggregation) είναι μικρή σε όγκο και συνεπώς η λεπτομερής παρακολούθηση της είναι εφικτή. Παράλληλα όμως η προσέγγιση αυτή αντιμετωπίζει το δύσκολο πρόβλημα της ανίχνευσης μιας μικρού μεγέθους ανωμαλίας μέσα στην φυσιολογική δικτυακή κίνηση. Λόγω της ευμετάβλητης φύσης της δικτυακής κίνησης και της έλλειψης μοντελοποίησης της το πρόβλημα της ανίχνευσης ανωμαλιών είναι ιδιαίτερα δύσκολο. Ένα ακόμη μειονέκτημα της προσέγγισης αυτής είναι η δυσκολία εγκατάστασης συστημάτων ανίχνευσης επιθέσεων σε κάθε πιθανό δίκτυο πηγή, καθώς η επένδυση σε υλικό (hardware) αλλά και σε χρόνο εργασίας δεν αντισταθμίζεται από κάποιο όφελος πέρα από την προστασία της κοινότητας του Διαδικτύου. Για το λόγο αυτό έχει προταθεί η επιβολή χρηματικών ποινών ως κίνητρο για την λήψη μέτρων από τα δίκτυα πηγές των επιθέσεων DDoS [56].

Ένα δίκτυο ανάλογα με τον ρόλο του κατά την διάρκεια μιας επίθεσης

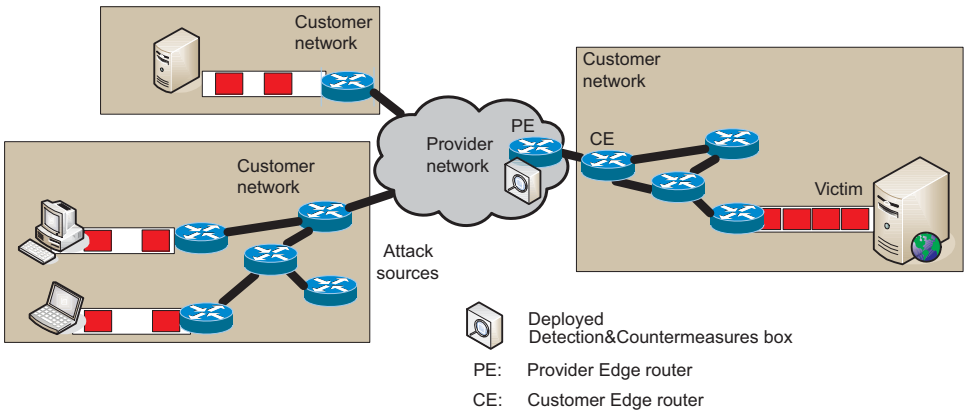


Σχήμα 2.5: Ταξινόμηση δικτύων που μετέχουν σε μια επίθεση καταιγισμού πακέτων.

DDoS μπορεί να χαρακτηριστεί ως πηγή, θύμα ή ενδιάμεσο. Ανάλογα με την θέση του χαρακτηρίζεται ως δίκτυο πελάτη ή παρόχου (σχήμα 2.5). Οι παραπάνω προσεγγίσεις αναφέρονται **σε δίκτυα πελατών** (Customer Networks - CN) καθώς αυτά αποτελούν συνήθως τις πηγές και τα θύματα μιας επίθεσης. Μια άλλη προσέγγιση είναι η ανίχνευση και αντιμετώπιση του φαινομένου σε ενδιάμεσα **δίκτυα παρόχων** (Provider Networks - PN). Ο χαρακτηρισμός ενός δικτύου ως ενδιάμεσου δεν είναι ιδιαίτερα κατατοπιστικός καθώς η διασύνδεση των σημερινών δικτύων είναι πολύπλοκη και δεν υπάρχει μια αυστηρή ιεραρχική δομή στην διασύνδεση τους. Σε γενικές γραμμές με τον όρο ενδιάμεσα αναφερόμαστε σε δίκτυα παρόχων δικτυακών υπηρεσιών διαφόρων επιπέδων (Network Service Providers, Internet Service Providers -ISP). Οι ISPs χαρακτηρίζονται ως Tier 1 όταν συνδέονται με το υπόλοιπο Διαδίκτυο μόνο μέσω peering με άλλους ISPs Tier 1. Δηλαδή δεν εξαρτώνται από άλλους ISPs ανώτερου επιπέδου (no upstream providers). Αντίθετα ISPs επιπέδου 2 λαμβάνουν μαζικό Internet feed από κάποιον IPS Tier 1 [57].

Η ανίχνευση επιθέσεων καταιγισμού πακέτων στα δίκτυα παρόχων (σχήμα 2.6) είναι αρκετά δύσκολη καθώς οι συνδέσεις που παρακολουθούμε είναι γρήγορες και έχουν χαμηλή χρησιμοποίηση [4] με αποτέλεσμα μια επίθεση να μην προκαλεί συμφόρηση. Θετικό όμως είναι το γεγονός ότι οι δρομολογητές που διαθέτουν τα δίκτυα αυτά είναι αρκετά ισχυροί και μπορούν να αντιμετωπίσουν

[4] οι μεγάλοι ISPs έχουν παραδοσιακά overprovisioned δίκτυα κορμού (backbone)



Σχήμα 2.6: Αντιμετώπιση επιθέσεων καταιγισμού πακέτων στο δίκτυο του παρόχου

αποτελεσματικά επιθέσεις εφόσον αυτές εντοπιστούν χωρίς να επηρεαστεί η ομαλή λειτουργία τους.

Γενικά υπάρχει μια σχέση αντίστροφης αναλογίας μεταξύ της αποτελεσματικότητας της αντιμετώπισης των επιθέσεων DDoS και της ευκολίας ανίχνευσης τους, όπως φαίνεται στον πίνακα 2.1. Η προσέγγιση που υιοθετεί η διατριβή είναι **η ανίχνευση ανωμαλιών στους ακραίους δρομολογητές του δικτύου ενός παρόχου (PE)** με σκοπό την προστασία δικτύων-πελατών, τα οποία ανάλογα με την περίπτωση μπορεί να είναι πηγές ή θύματα. Η επιλογή αυτή καθιστά δυνατή τη γρήγορη και αποτελεσματική εφαρμογή αντιμέτρων. Πρέπει όμως να αντιμετωπιστούν οι δυσκολίες που παρουσιάζει η παρακολούθηση του δικτύου και η ανίχνευση επιθέσεων σε υψηλές ταχύτητες.

Πίνακας 2.1: Βαθμός αποτελεσματικότητας των αναγκαίων ενεργειών για την αντιμετώπιση επιθέσεων DDoS ανάλογα με το σημείο που εφαρμόζονται.

Σημείο Άμυνας	Παρακολούθηση	Ανίχνευση	Αντιμετώπιση
Δίκτυο Θύματος	***	***	*
Δίκτυο Παρόχου	**	**	***
Δίκτυο Πηγής	***	*	***

Η επιλογή του σημείου εγκατάστασης ενός συστήματος ανίχνευσης και αντιμετώπισης επιθέσεων DDoS από έναν πάροχο δικτύου πρέπει να γίνει με

βάση τα εξής κριτήρια: πρέπει να διαθέτει ελεύθερο εύρος ζώνης (underutilized link), μεγάλη ισχύ δικτυακού εξοπλισμού (processing power in packets per second) και να βρίσκεται κοντά στο σύνορο παρόχου-πελάτη (PE-CE). Η προσέγγιση αυτή ενδιαφέρει και τους παρόχους δικτύου καθώς τους δίνει την δυνατότητα να προσφέρουν νέες, προηγμένες υπηρεσίες προστασίας των πελατών τους από επιθέσεις άρνησης υπηρεσιών και άλλες ανωμαλίες της δικτυακής κίνησης. Αξιοποιώντας την υποδομή του δικτύου κορμού και κάνοντας μια επένδυση σε συστήματα παρακολούθησης, ανίχνευσης και αντιμετώπισης επιθέσεων ένας πάροχος μπορεί να έχει μεγάλο περιθώριο κέρδους από τις υπηρεσίες αυτές λόγω οικονομίας κλίμακας. Τέτοιες υπηρεσίες προσφέρονται σήμερα από λίγους μεγάλους παρόχους Tier 1 όπως η ATT [58] και η Sprint [59].

Πέρα από τις προαναφερθείσες προσεγγίσεις ανίχνευσης και αντιμετώπισης στον ίδιο διαχειριστικό τομέα (δίκτυο πηγής, δίκτυο θύματος, δίκτυο παρόχου) υπάρχει η δυνατότητα συνεργασίας δύο δικτύων: το ένα ανιχνεύει μια επίθεση και το άλλο την αντιμετωπίζει. Στη γενική περίπτωση όπου τα δύο συνεργαζόμενα δίκτυα δεν έχουν ήδη σχέσεις εμπιστοσύνης, οι μέθοδοι συνεργασίας αποτελούν ανοικτό ερευνητικό θέμα [28]. Πρέπει να αντιμετωπιστούν θέματα πολιτικής, όπως η δυνατότητα κοινοποίησης στοιχείων (προστασία προσωπικών δεδομένων) [5] ή η δικαιοδοσία εγκατάστασης ενός φίλτρου απόρριψης κίνησης. Σε κάθε περίπτωση απαιτείται έγκαιρη και αποτελεσματική ανίχνευση των επιθέσεων και η παρούσα διατριβή διερευνά τρόπους για την ανίχνευση ανωμαλιών σε δίκτυα υψηλών ταχυτήτων χρησιμοποιώντας ενδείξεις πέραν της συμφόρησης (congestion).

2.4 Προληπτικά μέτρα

Οι διαχειριστές δικτύου ανεξάρτητα από τον τύπο του δικτύου που διαχειρίζονται, δηλαδή δίκτυα παρόχων ή πελατών, επιβάλλεται να λάβουν μια

[5] π.χ. οι επικεφαλίδες πακέτων που αποδεικνύουν μια επίθεση (packet traces)

σειρά από προληπτικά μέτρα για την αντιμετώπιση του φαινομένου των κατανεμημένων επιθέσεων άρνησης υπηρεσιών. Τα μέτρα αυτά αποσκοπούν στη προφύλαξη τόσο του δικτύου που βρίσκεται υπό την επίβλεψη τους όσο και του υπόλοιπου Διαδικτύου. Ακολουθεί μια σύντομη απαρίθμηση των μέτρων αυτών:

1. **Αντιμετώπιση της αποστολής παραποιημένων πακέτων.** Η αποστολή IP πακέτων στο Διαδίκτυο που στο πεδίο της διεύθυνσης αποστολέα (source address) έχουν μια τυχαία διεύθυνση και όχι την πραγματική IP διεύθυνση του αποστολέα ονομάζεται spoofing. Αν και η δυνατότητα αποστολής παραποιημένων πακέτων (spoofing) δεν μπορεί να εξαλειφθεί εντελώς, όπως θα εξηγήσουμε παρακάτω, ο έλεγχος της εισερχόμενης και εξερχόμενης δικτυακής κίνησης (Egress/Ingress filtering) μπορεί να μειώσει το μέγεθος του φαινομένου [15]. Σύμφωνα με την μέθοδο ελέγχου εξερχόμενης δικτυακής κίνησης (Egress filtering), τα πακέτα που εξέρχονται από τους συνοριακούς δρομολογητές ενός δικτύου X πρέπει να έχουν διεύθυνση αποστολέα μέσα από το τμήμα διευθύνσεων του X. Όταν χρησιμοποιείται ο έλεγχος της εισερχόμενης δικτυακής κίνησης (Ingress filtering) οι συνοριακοί δρομολογητές του δικτύου X (edge routers) δεν επιτρέπουν την είσοδο πακέτων με διεύθυνση αποστολέα (source address) από το τμήμα των διευθύνσεων του X (assigned address block). Προφανώς η προστασία που προσφέρει στο δίκτυο X η μέθοδος ελέγχου της εισερχόμενης δικτυακής κίνησης (Ingress filtering) είναι περιορισμένη. Καμία όμως από τις δύο πρακτικές δεν εμποδίζει έναν υπολογιστή ενός συγκεκριμένου δικτύου να στείλει πακέτα με διεύθυνση αποστολέα μια τυχαία IP από το δίκτυο του αντί με την πραγματική του IP διεύθυνση. Το κέρδος του επιτιθέμενου από την αποστολή παραποιημένων πακέτων είναι ότι δεν αποκαλύπτεται το σύστημα από το οποίο πράγματι πηγάζει μια επίθεση. Ακόμη και στην περίπτωση χρήσης spoofed addresses από το πραγματικό υποδίκτυο του

επιτιθέμενου, ως πηγές μπορούν παραπλανητικά να εμφανίζονται δεκάδες υπολογιστές. Η καθυστέρηση εντοπισμού του επιτιθέμενου υπολογιστή δυσκολεύει τη λήψη μέτρων αντιμετώπισης καθώς αυτά πρέπει να δράσουν σε ένα ολόκληρο υποδίκτυο προξενώντας παράπλευρες απώλειες (collateral damage). Μια συμπληρωματική μέθοδος με το όνομα Reverse Path Filtering (RPF) προβλέπει τον έλεγχο της εγκυρότητας της διεύθυνσης αποστολέα [16] με δυναμικό τρόπο, αντίθετα από τα στατικά φίλτρα που προαναφέρθηκαν. Η περίπτωση όμως χρήσης πλαστής διεύθυνσης αποστολέα (spoofing) από το ίδιο υποδίκτυο με την πραγματική πηγή του πακέτου δεν αντιμετωπίζεται με τις παραπάνω τεχνικές. Αυτό συμβαίνει γιατί ένα interface εισόδου/εξόδου στο δρομολογητή αντιστοιχεί σε ένα υποδίκτυο υπολογιστών και όχι σε μεμονωμένους υπολογιστές. Στην ίδια κατηγορία προληπτικών μέτρων ανήκουν προτάσεις επέκτασης της τεχνικής RPF σε δίκτυα κορμού όπως αναλύεται από τους Park et al [22]. Οι επεκτάσεις αυτές δεν υλοποιούνται σήμερα καθώς απαιτείται αλλαγή των πρωτοκόλλων δρομολόγησης στο Διαδίκτυο.

2. **Προληπτικός περιορισμός κίνησης ICMP.** Το πρωτόκολλο ICMP χρησιμοποιείται κυρίως για διαχειριστικές λειτουργίες όπως ο εντοπισμός προβλημάτων [60] και συνεπώς ο όγκος της δικτυακής κίνησης που του αναλογεί είναι πρακτικά μικρός. Ο περιορισμός του ρυθμού αποστολής ICMP κίνησης προς αποφυγή επιθέσεων καταιγισμού πακέτων τύπου ICMP είναι μια καλή πρακτική. Το όριο που θα επιβληθεί πρέπει όμως να είναι αρκετά μεγάλο ώστε να μην παρεμποδιστούν φυσιολογικές λειτουργίες όπως η διαδικασία MTU path discovery [61].

3. **Ενίσχυση ασφάλειας δικτυακών συσκευών.** Ο όρος δικτυακές συσκευές είναι γενικός και περιγράφει πέρα από ηλεκτρονικούς υπολογιστές και δρομολογητές κάθε συσκευή που συνδέεται στο δίκτυο. Τα

συστήματα αυτά αποτελούν πιθανό στόχο και κατ' επέκταση πηγή δυνητική επιθέσεων. Η συνεχής διόρθωση προβλημάτων λογισμικού (software upgrades) και η χρήση κρυπτογραφημένων πρωτοκόλλων πρόσβασης (ssh/sftp αντί telnet/ftp) είναι δύο από τα μέτρα που μπορούν να αποτρέψουν να λάβει τον έλεγχο μιας δικτυακής συσκευής ένας επιτιθέμενος. Αναλυτικές οδηγίες για την προστασία δρομολογητών, μεταγωγέων και υπολογιστών με δημοφιλή λειτουργικά συστήματα παρέχονται από την NSA των Η.Π.Α [62]. Όσο πιο αποτελεσματικά προστατευτούν όλες οι δικτυακές συσκευές τόσο πιο δύσκολη θα είναι η δημιουργία ενός πολυάριθμου "ηλεκτρονικού στρατού" [10, 11], έτοιμου να εξαπολύσει επιθέσεις καταιγισμού πακέτων.

4. **Περιορισμός πρόσβασης στη δικτυακή υποδομή.** Ο εξοπλισμός που αποτελεί το δίκτυο κορμού (backbone) ενός δικτύου, όπως είναι οι δρομολογητές και οι μεταγωγείς, μπορεί να προστατευτεί από επιθέσεις καταιγισμού πακέτων με χρήση φίλτρων που επιτρέπουν συνδέσεις μόνο από προκαθορισμένα συστήματα διαχείρισης. Τα φίλτρα αυτά (Infrastructure Access Control Lists - ACL) [17] απορρίπτουν κάθε προσπάθεια σύνδεσης στον εξοπλισμό από μη εξουσιοδοτημένες διευθύνσεις στις κάρτες εισόδου ενός δρομολογητή. Κατά συνέπεια μια επίθεση άρνησης υπηρεσιών στο διαχειριστικό επίπεδο (management plane), που υλοποιείται στην κεντρική μονάδα επεξεργασίας του δρομολογητή, είναι πολύ δύσκολη. Παράλληλα μειώνεται ο κίνδυνος παραβίασης των συσκευών αυτών και η χρήση τους ως πηγές μιας επίθεσης καταιγισμού πακέτων. Το σενάριο αυτό θα ήταν ιδιαίτερα επικίνδυνο λόγω της δυνατότητας των δρομολογητών να χειριστούν ένα μεγάλο αριθμό πακέτων ανά δευτερόλεπτο και συνεπώς να εξαπολύσουν ισχυρότατες επιθέσεις καταιγισμού πακέτων.

2.5 Μηχανισμοί ανίχνευσης και καταστολής

Τα προληπτικά μέτρα που είδαμε στην προηγούμενη ενότητα δεν είναι δυνατόν να σταματήσουν την εκδήλωση επιθέσεων καταιγισμού πακέτων. Για την πλήρη αντιμετώπιση του φαινομένου απαιτείται έγκαιρη ανίχνευση και καταστολή. Οι υπάρχουσες μέθοδοι καταστολής προϋποθέτουν την επιτυχημένη ανίχνευση και ταυτοποίηση μιας επίθεσης DDoS. Με τον όρο ταυτοποίηση εννοούμε τον προσδιορισμό των χαρακτηριστικών μιας επίθεσης ώστε να μπορέσει να διαχωριστεί από την υπόλοιπη "φυσιολογική" κίνηση. Όπως αναλύεται στην ενότητα 2.5.2, για την εφαρμογή των μεθόδων καταστολής απαιτείται η ταυτοποίηση μιας επίθεσης. Η ταυτοποίηση γίνεται με τον προσδιορισμό των δικτυακών ροών (flows) που την αποτελούν. **Σύμφωνα με τη τεχνολογία Netflow [33] ένα flow ορίζεται ως το σύνολο των πακέτων που διέρχονται από μια δικτυακή ζεύξη προς μια συγκεκριμένη κατεύθυνση (είσοδος, έξοδος) και έχουν κοινές τιμές για τα παρακάτω πέντε στοιχεία της IP επικεφαλίδας (header) τους: Πρωτόκολλο, IP προέλευσης, IP προορισμού, Port προέλευσης, Port προορισμού.**

Ο προσδιορισμός των χαρακτηριστικών αυτών πραγματοποιείται συνήθως κατά την ανίχνευση μιας επίθεσης. Μια από τις καινοτομίες της διατριβής είναι η ανάλυση της ανίχνευσης σε τρεις διαδικασίες: την παρακολούθηση δικτύου, την ανίχνευση ανωμαλιών και την φάση ταυτοποίησης της ανωμαλίας. Η ανάλυση αυτή κρίνεται σκόπιμη για λόγους απλοποίησης αλλά και γιατί τα επιμέρους προβλήματα χαρακτηρίζονται από διαφορετικές απαιτήσεις στον τρόπο επίλυσης τους. Συγκεκριμένα η παρακολούθηση του δικτύου και η ανίχνευση ανωμαλιών είναι συνεχείς διαδικασίες που πρέπει να εκτελούνται με μεγάλη συχνότητα για να έχουμε έγκαιρη ειδοποίηση ενώ η ταυτοποίηση μιας ανωμαλίας εκτελείται ασύγχρονα (μόνο μετά τον εντοπισμό μιας ανωμαλίας) και μπορεί να διαρκεί περισσότερο.

Η παρούσα διατριβή εστιάζει στην ανάπτυξη ενός αποτελεσματικού συστήματος ανίχνευσης επιθέσεων καταιγισμού πακέτων. Στην ενότητα 2.5.1.1

παρουσιάζεται το θεωρητικό πλαίσιο αξιολόγησης συστημάτων ανίχνευσης. Στην ενότητα 2.5.1.2 αναλύεται η ανίχνευση επιθέσεων σε επιμέρους διαδικασίες και στην ενότητα 2.5.1.3 αιτιολογούμε την προσέγγιση αυτή. Η ενότητα 2.5.1.4 συνοψίζει τις σημαντικότερες προσεγγίσεων που απαντώνται στη βιβλιογραφία για την ανίχνευση κατανεμημένων επιθέσεων άρνησης υπηρεσιών. Το κεφάλαιο ολοκληρώνεται στην ενότητα 2.5.2 με μια σύντομη αναφορά στις μεθόδους καταστολής που μπορούν να εφαρμοστούν μετά από την επιτυχημένη ανίχνευση μιας επίθεσης.

2.5.1 Ανίχνευση επιθέσεων

2.5.1.1 Αξιολόγηση της απόδοσης συστημάτων ανίχνευσης

Χρησιμοποιώντας ορισμούς από την θεωρία ανίχνευσης (detection theory) θα παρουσιάσουμε τα κριτήρια αξιολόγησης ενός συστήματος ανίχνευσης επιθέσεων. Θεωρούμε ότι έχουμε ένα σύνολο N δειγμάτων όπου κάθε ένα αντιστοιχεί σε μια φυσιολογική κατάσταση K_1 ή μια κατάσταση επίθεσης K_2 (attack). Έστω ότι έχουμε N_1 δείγματα K_1 και N_2 δείγματα K_2. Προφανώς $N = N_1 + N_2$. Ένα σύστημα ανίχνευσης καλείται να αξιολογήσει κάθε δείγμα (τεστ). Αν ένα δείγμα χαρακτηριστεί ως επίθεση K_2 (τεστ θετικό -D_2) τότε παράγεται ένα μήνυμα συναγερμού (alert). Αν το τεστ είναι αρνητικό D_1 δεν παράγεται alert. Στα πλαίσια του τεστ ορίζουμε ως :

- true positives (TP) = αριθμός από alert σε καταστάσεις επίθεσης ($D_2|K_2$).
- false negatives (FN) = αριθμός απουσίας alert σε καταστάσεις επίθεσης($D_1|K_2$).
- false positives (FP) = αριθμός από alert σε φυσιολογικές καταστάσεις($D_2|K_1$).
- true negatives (TN) = αριθμός απουσίας alert σε φυσιολογικές καταστάσεις($D_1|K_1$).

Οι παραπάνω ορισμοί αποτυπώνονται στον πίνακα 2.2.

Πίνακας 2.2: Ορισμοί απόκρισης συστήματος ανίχνευσης.

	Alert (D_2)	Non Alert (D_1)	Total
Attack (K_2)	True Positive	False Negative	N_2
Non Attack (K_1)	False Positive	True Negative	N_1

Καθώς σε κάθε περίπτωση το σύστημα αποκρίνεται με μήνυμα συναγερμού ή όχι (δεν υπάρχει ενδιάμεση κατάσταση) έχουμε:

$$TP + FN = N_2 \Rightarrow \frac{TP}{N_2} + \frac{FN}{N_2} = 1$$

$$FP + TN = N_1 \Rightarrow \frac{FP}{N_1} + \frac{TN}{N_1} = 1$$

Η **ευαισθησία (sensitivity)** ορίζεται ως η πιθανότητα το τεστ να είναι θετικό σε περίπτωση μίας επίθεσης. Η ευαισθησία που παριστάνεται με το σύμβολο S_n δίνεται από την σχέση:

$$S_n \hat{=} P[Τεστ\ θετικο|Επιθεση]$$

και εκτιμάται από την σχέση:

$$S_n \cong \frac{TP}{(TP + FN)} = \frac{TP}{N_2}$$

Υψηλή ευαισθησία σημαίνει πως ένα αρνητικό τεστ μπορεί με μεγάλη εμπιστοσύνη να απορρίψει την ύπαρξη επίθεσης.

Η **ειδικότητα (specificity)** ορίζεται ως η πιθανότητα το τεστ να είναι αρνητικό σε περίπτωση μίας φυσιολογικής κατάστασης. Η ειδικότητα παριστάνεται με το σύμβολο S_p και δίνεται από την σχέση:

$$S_p \hat{=} P[Τεστ\ αρνητικο|Μη\ επιθεση]$$

και εκτιμάται από την σχέση:

$$S_p \cong \frac{TN}{(TN + FP)} = \frac{TN}{N_1}$$

Υψηλή ειδικότητα ενός συστήματος ανίχνευσης σημαίνει ότι ένα θετικό τεστ μπορεί με μεγάλη εμπιστοσύνη να προσδιορίσει την ύπαρξη επίθεσης.

Η αξιολόγηση της απόδοσης των συστημάτων ανίχνευσης γίνεται συνήθως βάσει δύο κριτηρίων: την πιθανότητα true positives (TP ratio - TPR)

$$TPR = S_n \cong \frac{TP}{(TP + FN)} = \frac{TP}{N_2}$$

και την πιθανότητα false positives (FP ratio - FPR)

$$FPR = 1 - S_p \cong \frac{FP}{(TN + FP)} = \frac{FP}{N_1}$$

Με χρήση των κριτηρίων αυτών σχεδιάζονται οι χαρακτηριστικές καμπύλες ROC (Receiver Operating Characteristic curves) που παρουσιάζουν την σχέση (tradeoff) μεταξύ ευαισθησίας και ακρίβειας ενός συστήματος ανίχνευσης βάσει των παραμέτρων λειτουργίας του. Τέλος γνωρίζοντας τα TPR και FPR οι πιθανότητες των true negatives και false negatives προκύπτουν από τις σχέσεις:

$$FNR = 1 - TPR$$

$$TNR = 1 - FPR$$

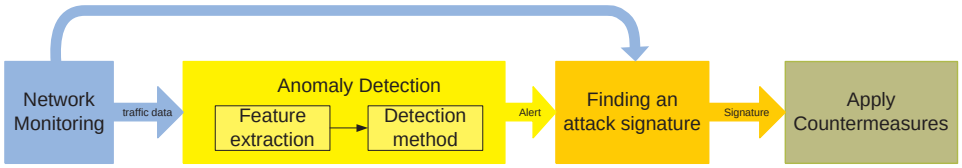
Οι ορισμοί και οι έννοιες αυτές ακολουθούν το παράδειγμα παλαιότερων ανιχνευτών. Για παράδειγμα στο χώρο της ιατρικής υπάρχουν πολλά τεστ για την ανίχνευση ασθενειών που παρουσιάζουν υψηλή ευαισθησία αλλά χαμηλή ειδικότητα [63].

2.5.1.2 Ανάλυση της ανίχνευσης επιθέσεων σε επιμέρους διαδικασίες

Για την αποτελεσματική αντιμετώπιση μιας ανωμαλίας στη δικτυακή κίνηση και εν προκειμένου μιας επίθεσης DDoS είναι απαραίτητο να προηγηθεί η ακριβής ταυτοποίηση της. Δηλαδή για την εφαρμογή μέτρων κατασταλτικού χαρακτήρα απαιτείται η περιγραφή μιας επίθεσης (attack signature) με την μέγιστη δυνατή ακρίβεια, για παράδειγμα ορίζοντας τις δικτυακές ροές (**flows**) που την συνθέτουν. Σύμφωνα με τη τεχνολογία Netflow [33] ένα flow ορίζεται ως το σύνολο των πακέτων που διέρχονται από μια δικτυακή ζεύξη προς μια συγκεκριμένη κατεύθυνση (unidirectional) και έχουν όλα κοινές τιμές στα παρακάτω πέντε στοιχεία της IP επικεφαλίδας (header) τους: Πρωτόκολλο, IP προέλευσης, IP προορισμού, Port προέλευσης, Port προορισμού. Πριν όμως ταυτοποιηθεί μια επίθεση είναι προφανές ότι πρέπει να έχει ανιχνευτεί. Η λογική ακολουθία των βημάτων που συνθέτουν την αντιμετώπιση μιας επίθεσης DDoS παρουσιάζεται στο σχήμα 2.7. Στην βιβλιογραφία οι επιμέρους διαδικασίες της παρακολούθησης δικτύου, της ανίχνευσης και της ταυτοποίησης επιθέσεων δεν διαχωρίζονται επαρκώς και παρουσιάζονται συνολικά ως ανίχνευση επιθέσεων. Όπως θα αναλύσουμε στην συνέχεια η ανίχνευση επιθέσεων είναι μια σύνθετη διαδικασία που υπόκειται σε πολλούς πρακτικούς περιορισμούς. Για το λόγο αυτό προτείνεται στην παρούσα διατριβή ο διαχωρισμός της ανίχνευσης σε τρεις επιμέρους διαδικασίες - βήματα που προηγούνται της λήψης αντίμετρων (countermeasures):

1. Παρακολούθηση δικτύου (Network monitoring)
2. Εξαγωγή χαρακτηριστικών (μετρικών) και εκτέλεση αλγόριθμων ανίχνευσης (Anomaly detection)
3. Ταυτοποίηση ανιχνευθεισών ανωμαλιών (Finding an attack signature)

Στα πλαίσια της πρώτης διαδικασίας (παρακολούθηση του δικτύου) καταγράφονται όλες οι απαραίτητες πληροφορίες που θα διευκολύνουν το έργο της



Σχήμα 2.7: Φάσεις ανίχνευσης και καταστολής επιθέσεων

ανίχνευσης επιθέσεων. Η διαδικασία αυτή κυμαίνεται από απλή συλλογή δεδομένων δικτυακής κίνησης που θα δεχτούν περαιτέρω επεξεργασία μέχρι την μέτρηση συγκεκριμένων χαρακτηριστικών. Στην παρούσα διατριβή ο υπολογισμός μετρικών από τα δεδομένα που συλλέγονται κατά την παρακολούθηση δικτύου ονομάζεται εξαγωγή χαρακτηριστικών - μετρικών (feature extraction). Ο **προσδιορισμός των μετρικών** στα οποία θα βασιστεί η διαδικασία της ανίχνευσης (detection) είναι ιδιαίτερα σημαντικός καθώς κρίνει καθοριστικά την αποτελεσματικότητα του συνολικού συστήματος ανίχνευσης. Ο υπολογισμός τους υπόκειται στους περιορισμούς των χρησιμοποιούμενων τεχνολογιών παρακολούθησης δικτύων και ιδιαίτερα στις δυνατότητες συλλογής και επεξεργασίας στοιχείων δικτυακής κίνησης σε υψηλές ταχύτητες. Η διαδικασία ανίχνευσης ανωμαλιών έχει ως στόχο να απαντήσει στην ερώτηση κατά πόσο το υπό παρακολούθηση δίκτυο βρίσκεται σε φυσιολογική ή μη κατάσταση. Η απάντηση δίνεται με την εφαρμογή κάποιου αλγόριθμου ανίχνευσης, που επεξεργάζεται τα μετρικά που έχουν υπολογιστεί. Σε περίπτωση που ανιχνεύεται μια επίθεση ενεργοποιείται η διαδικασία ταυτοποίησης της. Η διαδικασία αυτή προσδιορίζει τα χαρακτηριστικά της επίθεσης ώστε να μπορέσει να διαχωριστεί από την υπόλοιπη "φυσιολογική" κίνηση. Τα χαρακτηριστικά αυτά πρέπει να είναι όσο το δυνατό πιο ακριβή για να αποφεύγονται παράπλευρες απώλειες (collateral damage), δηλαδή η επιβολή μέτρων καταστολής σε φυσιολογική κίνηση που εσφαλμένα χαρακτηρίστηκε ως επίθεση.

2.5.1.3 Αιτιολόγηση της προτεινόμενης αρχιτεκτονικής ανίχνευσης

Η αρχιτεκτονική που προτείνεται από την παρούσα διατριβή διαχωρίζει το πρόβλημα της ανίχνευσης σε τρεις επιμέρους διαδικασίες για λόγους απλοποίησης αλλά και γιατί τα επιμέρους προβλήματα χαρακτηρίζονται από διαφορετικές απαιτήσεις ως προς τον τρόπο επίλυσης τους. Συγκεκριμένα η παρακολούθηση του δικτύου και η ανίχνευση ανωμαλιών είναι συνεχείς διαδικασίες που πρέπει να εκτελούνται με μεγάλη συχνότητα για να έχουμε έγκαιρη ειδοποίηση. Αντίθετα η ταυτοποίηση μιας ανωμαλίας εκτελείται ασύγχρονα, δηλαδή μόνο μετά τον εντοπισμό μιας ανωμαλίας, και μπορεί να διαρκεί μεγαλύτερα χρονικά διαστήματα. Διαχωρίζουμε λοιπόν τους αλγόριθμους που χρησιμοποιούνται από την διαδικασία ανίχνευσης ανωμαλιών και από τη διαδικασία ταυτοποίησης. Η πρώτη έχει ως στόχο να απαντήσει στην ερώτηση κατά πόσο το υπό παρακολούθηση δίκτυο βρίσκεται σε φυσιολογική ή μη κατάσταση. Στο στάδιο αυτό δεν μας απασχολεί ποια είναι η ανωμαλία που τυχόν παρατηρείται αλλά απλά αν είναι υπαρκτή. Συνεπώς δεν χρειάζεται να δώσουμε τελικές απαντήσεις για την ταυτότητα του θύματος ή για τις πηγές της επίθεσης αλλά απλά να διαγνώσουμε μια ανωμαλία στη δικτυακή κίνηση βάσει ενός ή περισσοτέρων μετρικών. Η διαδικασία αυτή πρέπει να είναι υπολογιστικά απλή ώστε ο αλγόριθμος ανίχνευσης να εκτελείται συνεχώς και να μπορεί να διαγνώσει έγκαιρα σύντομες χρονικά επιθέσεις (time constraint). Υποστηρίζοντας λοιπόν συχνότερη δειγματοληψία μετρικών και συχνότερες εκτιμήσεις για την ύπαρξη ανωμαλιών επιτυγχάνεται ταχύτερη και ακριβέστερη ανίχνευση. Για να επιτευχθεί αυτό ο όγκος της πληροφορίας που επεξεργάζεται πρέπει να είναι περιορισμένος (storage constraint).

Ακόμη και αν γνωρίζουμε την ύπαρξη μιας επίθεσης, η εύρεση μιας ακριβούς και εύστοχης περιγραφής της (attack signature) είναι δύσκολη. Η περιγραφή αυτή πρέπει να είναι αρκετά ακριβής ώστε όταν εφαρμοστούν μέθοδοι καταστολής, όπως η εγκατάσταση φίλτρων απόρριψης πακέτων, να μη αδικηθεί φυσιολογική δικτυακή κίνηση. Παράλληλα για να λειτουργήσουν αποτελεσμα-

τικά οι μέθοδοι καταστολής πρέπει να είναι αρκετά εύστοχη και να καλύπτει ένα μεγάλο ποσοστό της κίνησης που προέρχεται από την επίθεση. Η διαδικασία της ταυτοποίησης αναφέρεται στην βιβλιογραφία ως εντοπισμός ενός συνοθηλεύματος κίνησης (traffic aggregate), δηλαδή ενός συνόλου πακέτων με κοινά χαρακτηριστικά [19]. Η διαδικασία αυτή έχει αρκετές διαφορές από την ανίχνευση ανωμαλιών. Πρώτον δεν απαιτείται να εκτελείται συνεχώς αλλά καλείται ασύγχρονα, μονάχα σε περίπτωση εντοπισμού κάποιας ανωμαλίας. Δεύτερον καθώς πρέπει να χαρακτηρίσει επακριβώς την επίθεση είναι προτιμότερο να βασίζεται σε λεπτομερή στοιχεία κίνησης. Τρίτον, μπορούν να γίνουν παραχωρήσεις σχετικά με την ταχύτητα της διαδικασίας ταυτοποίησης. Γενικά ένα σύστημα ανίχνευσης ανωμαλιών της δικτυακής κίνησης θα πρέπει να συλλέγει πληροφορίες κίνησης, να υπολογίζει μετρικά και να ανιχνεύει σε σχεδόν πραγματικό χρόνο (near real time), ώστε να μπορεί να παρακολουθεί τις διακυμάνσεις της κίνησης που επιτρέπει η υψηλή ταχύτητα των σύγχρονων δικτύων. Αντίθετα η διαδικασία της ταυτοποίησης δεν απαιτείται να παράγει τόσο γρήγορα αποτελέσματα, καθώς ένα χρονικό διάστημα αρκετών λεπτών μέχρι την καταστολή μιας επίθεσης είναι σήμερα ανεκτό. Βλέπουμε λοιπόν ότι ο αλγόριθμος ταυτοποίησης μπορεί να σχεδιαστεί χωριστά με μικρότερους περιορισμούς σε πολυπλοκότητα χρόνου και χώρου (time and storage complexity) από ότι οι αλγόριθμοι ανίχνευσης.

2.5.1.4 Βιβλιογραφική ανασκόπηση προσεγγίσεων ανίχνευσης

Έχοντας παρουσιάσει τα βασικά στάδια επεξεργασίας που συνθέτουν την διαδικασία ανίχνευσης μπορούμε να προχωρήσουμε σε μια σύντομη παρουσίαση των βασικότερων προσεγγίσεων της βιβλιογραφίας για την επίλυση του προβλήματος της ανίχνευσης κατανεμημένων επιθέσεων άρνησης υπηρεσιών. Οι προσεγγίσεις αυτές δεν διαχωρίζουν σαφώς τις διαδικασίες ανίχνευσης και ταυτοποίησης. Οι αλγόριθμοι ανίχνευσης στις περισσότερες περιπτώσεις εντοπίζουν και κάποια χαρακτηριστικά που περιγράφουν μια επίθεση. Για το λόγο

αυτό χρησιμοποιούνται συνήθως λεπτομερή μετρικά (π.χ. πλήθος πακέτων ανά διεύθυνση IP) που σε πολλές περιπτώσεις δεν είναι μετρήσιμα σε δίκτυα υψηλών ταχυτήτων, δεδομένων των υπαρχόντων τεχνολογιών παρακολούθησης δικτύων.

1. **Παρακολούθηση δικτύου και μετρικά.** Οι περισσότερες βιβλιογραφικές αναφορές δεν αναλύουν την διαδικασία εξαγωγής μετρικών από τους υπάρχοντες μηχανισμούς παρακολούθησης δικτύου. Τα συχνότερα χρησιμοποιούμενα μετρικά είναι ο ρυθμός αποστολής bits (bit rate), που χρησιμοποιήθηκε από τους Li et al [64], Ramanarran [65], Manikopoulos et al [66], Jiang et al [67], ο ρυθμός αποστολής πακέτων (packet rate), που χρησιμοποιήθηκε από τους Barford et al [68, 69], Akella et al [70], Hussain et al [71] ή ο ρυθμός αιτήσεων σε επίπεδο εφαρμογών (application layer request rates) π.χ. ρυθμός κλήσεων HTTP όπως προτάθηκε από τους Jung et al [72]. Τα μετρικά αυτά συχνά διαχωρίζονται ανάλογα με την κατεύθυνση, π.χ. σε εισερχόμενα/εξερχόμενα πακέτα, ή ανάλογα με το πρωτόκολλο (TCP, UDP, ICMP) όπως αναφέρεται από τους Cabrera et al [73]. Κατά πόσο τα απλά αυτά μετρικά είναι ικανά να ανιχνεύσουν επιτυχώς επιθέσεις καταιγισμού πακέτων είναι αμφίβολο όπως θα δούμε στο κεφάλαιο 4. Στην βιβλιογραφία έχουν προταθεί διάφορα εναλλακτικά μετρικά όπως:

   - το CPU load ενός δρομολογητή από τον M. Behringer [74].
   - το πλήθος SYN, FIN, RST πακέτων από τους Wang et al [75], Blazek et al [76] και Noh et al [77].
   - η κατανομή του μήκους ενός flow από τους Akella et al [70] και Siaterlis et al [37].
   - η κατανομή του μεγέθους IP πακέτων από τους Blazek et al [76].
   - το πλήθος των νέων διευθύνσεων προέλευσης που παρατηρούνται από τους Peng et al [78] και Jung et al [72].

- το πλήθος των διαφορετικών προθεμάτων διεύθυνσης προέλευσης ανά διεύθυνση προορισμού από τους Akella et al [70] και Hussain et al [71].
- το πλήθος των UDP flows ανά διεύθυνση προορισμού από τους Mirkovic et al [29].
- ο αριθμός των flows ανά δευτερόλεπτο που παρατηρεί ένας δρομολογητής από τους Barford et al [68, 69].
- ο λόγος των εξερχόμενων προς εισερχόμενων TCP, UDP ή ICMP πακέτων ανά προορισμό από τους Mirkovic et al [29] και Gil et al [30].
- ο ρυθμός απόρριψης πακέτων στην ουρά ενός δρομολογητή (drop rate) από τους Ioannidis et al [19].

Τα μετρικά που προτείνονται από την παρούσα διατριβή επεκτείνουν τις παραπάνω ιδέες. Η επιλογή των προτεινόμενων μετρικών έγινε μετά από ενδελεχή αξιολόγηση της χρησιμότητας, της πολυπλοκότητας και της ευκολίας μέτρησης τους. Η αξιολόγηση αυτή βασίστηκε σε πλήθος πειραμάτων, δοκιμών και μετρήσεων σε πραγματικά δίκτυα υψηλών ταχυτήτων. Ενδεικτικά αναφέρουμε στο σημείο αυτό, ότι μετρικά που απαιτούν την τήρηση στατιστικών ανά IP διεύθυνση προορισμού ή προέλευσης είναι πολύ δύσκολο να μετρηθούν σε δίκτυα υψηλών ταχυτήτων λόγω πρακτικών περιορισμών σε μνήμη και υπολογιστική ισχύ. Όπως όμως θα δούμε στην ενότητα 4.3 υπάρχουν άλλα πιο αφηρημένα-γενικά μετρικά που είναι εύκολα μετρήσιμα και αρκετά αποτελεσματικά.

2. **Τεχνικές προ-επεξεργασίας (pre processing).** Σε πολλές περιπτώσεις τα μετρικά δεν χρησιμοποιούνται στην πρωτογενή τους μορφή (όπως παράγονται κατά την παρακολούθηση δικτύου) αλλά υφίστανται κάποια προ-επεξεργασία προκειμένου να αποκαλυφθούν κάποια ενδιαφέροντα

χαρακτηριστικά (feature extraction). Στην βιβλιογραφία έχουν προταθεί αρκετές διαφορετικές τεχνικές προ-επεξεργασίας όπως:

- η χρήση αλγόριθμων clustering βασισμένων σε χαρακτηριστικά δικτυακών ροών (flow) από τους Estan et al [79] και βασισμένων σε source IP διευθύνσεις από τους Jung et al [72].
- η ανάλυση με χρήση wavelets από τους Barford et al [68], Ramanarran [65] και Li et al [64].
- η ανάλυση στο πεδίο της συχνότητας (Fourier) από τους Cheng et al [80].
- ο αλγόριθμος 'sample and hold', που εντοπίζει δημοφιλείς προορισμούς πακέτων και εκτιμά το μέγεθος των δικτυακών ροών προς κάθε έναν από αυτούς, από τους Akella et al [70].
- αλγόριθμοι πρόβλεψης της εξέλιξης μιας χρονοσειράς (time series) βάση προηγούμενων δειγμάτων της από τους Jiang et al [67] και Barford et al [68].
- η μέθοδος ανάλυσης σε κύριες συνιστώσες (Principal Component Analysis - PCA) από τους Lakhina et al [81].

3. **Αλγόριθμοι ανίχνευσης.** Οι περισσότερες μέθοδοι ανίχνευσης που έχουν προταθεί χρησιμοποιούν προ-επεξεργασμένα ή μη μετρικά πάνω στα οποία εφαρμόζουν απλές συναρτήσεις κατωφλίου (thresholds). Η ρύθμιση του κατωφλίου μπορεί να γίνεται αυθαίρετα ή κατά το χρόνο εκπαίδευσης (training period) όπως αναφέρεται από τους Mirkovic et al [29], Ramanarran [65] και Gil et al [30]. Σε άλλες περιπτώσεις η τιμή της συνάρτησης κατωφλίου προσαρμόζεται δυναμικά με χρήση ενός κυλιόμενου παραθύρου (sliding window) όπως αναφέρεται από τους Jiang et al [67]. Ορισμένες προσεγγίσεις, όπως των Barford et al [68], λαμβάνουν υπόψη το μέγεθος και την διάρκεια των διακυμάνσεων ενός μετρικού. Στη κατηγορία αυτή ανήκουν υλοποιήσεις με τη χρήση αλγόριθμων

τύπου Cumulative Sum (CUSUM) όπως των Siris et al [82], Wang et al [75], Peng et al [83] κ.α.

4. **Τελική επεξεργασία (post processing).** Σε πολλές περιπτώσεις προκειμένου να μειωθεί ο αριθμός των false positives ή το πλήθος των ανιχνευόμενων συμβάντων στο χρόνο χρησιμοποιείται κάποια τεχνική τελικής επεξεργασίας πριν να παραχθεί ένα μήνυμα συναγερμού (alert). Για παράδειγμα μπορεί να θεωρηθεί ότι η απόκλιση από μία προβλεπόμενη τιμή ή από το "φυσιολογικό" προφίλ δεν είναι αρκετή για την παραγωγή ενός μηνύματος αλλά πρέπει να παρατηρηθούν επαναλαμβανόμενες αποκλίσεις για την παραγωγή του μηνύματος συναγερμού. Ανάλογες προτάσεις έχουν γίνει από τους Akella et al [70] και Jiang et al [67]. Οι τεχνικές αυτές αποτελούν ουσιαστικά ένα δεύτερο στάδιο ανίχνευσης και χρησιμοποιούν μεθόδους που ήδη αναφέραμε.

2.5.2 Μέθοδοι καταστολής

Μετά την ανίχνευση μιας δικτυακής ανωμαλίας και την ταυτοποίηση των χαρακτηριστικών της είναι δυνατή η εφαρμογή διάφορων μεθόδων καταστολής. Η πιο συχνά χρησιμοποιούμενη μέθοδος είναι η εφαρμογή ενός φίλτρου δικτυακής κίνησης (access list) από ένα firewall. Εναλλακτικά υπάρχει η δυνατότητα ελέγχου της ροής της δικτυακής κίνησης - **rate limiting** [18, 19]. Τα μέτρα αυτά εφαρμόζονται συνήθως με ανθρώπινη παρέμβαση σε έναν ή περισσότερους δρομολογητές και firewalls. Σε περιπτώσεις μεγάλων δικτύων η διαδικασία αυτή είναι κουραστική και χρονοβόρα. Για το λόγο αυτό αναπτύχθηκαν τεχνικές που μπορούν να φιλτράρουν την δικτυακή κίνηση δίνοντας μία εντολή σε έναν δρομολογητή η οποία προωθείται σε περισσότερα σημεία μέσω των πρωτοκόλλων δρομολόγησης. Σύμφωνα με την τεχνική **blackhole routing**, πακέτα με συγκεκριμένες IP διευθύνσεις προορισμού δρομολογούνται στο null interface ενός δρομολογητή ενώ με την τεχνική **traffic shunt** [21] προωθούνται σε έναν συγκεκριμένο προορισμό (IP address). Οι τεχνικές

αυτές υλοποιούνται ουσιαστικά ως κανόνες δρομολόγησης που μεταφέρονται μέσω του iBGP [20] ή eBGP [84]. Η τεχνική traffic shunt παρουσιάζει ένα ιδιαίτερο ενδιαφέρον καθώς προωθώντας την δικτυακή κίνηση σε ειδικό εξοπλισμό θα μπορούσαμε να την "καθαρίσουμε", δηλαδή να την απαλλάξουμε από την παρατηρούμενη ανωμαλία. Με τον καθαρισμό της δικτυακής κίνησης η τεχνική traffic shunt συνδυάζει την αντιμετώπιση περιστατικών ασφάλειας και την δυνατότητα ελεύθερης επικοινωνίας στο Διαδίκτυο[6].

Οι τεχνικές που αποσκοπούν να εντοπίσουν τις πραγματικές πηγές μιας ανωμαλίας μπορούν να θεωρηθούν ως ένας άλλος τύπος κατασταλτικών μέτρων. Αυτές οι τεχνικές συναντώνται στην βιβλιογραφία με τον όρο **traceback** και προσπαθούν να εντοπίσουν τις πηγές μιας επίθεσης με παραποιημένα πακέτα (spoofed packets) προκειμένου στην συνέχεια να ενημερωθούν οι υπεύθυνοι διαχειριστές, να αποκοπούν τα υπεύθυνα συστήματα από το δίκτυο, να εφαρμοστούν αποτελεσματικά φίλτρα, να επιδιορθωθούν και να ασφαλιστούν τα παραβιασμένα συστήματα προβλήματα αλλά και να συλλεχθούν οι απαραίτητες πληροφορίες για να εντοπιστούν οι υπαίτιοι της παραβίασης των συστημάτων. Τα ευρήματα αυτά μπορούν να στηρίζουν διώξεις και αιτήματα για αποζημιώσεις. Ενδεικτικά αναφέρουμε τις τεχνικές που βασίζονται στο μαρκάρισμα πακέτων (probabilistic packet marking) [85–87], στην αποστολή ξεχωριστών πακέτων π.χ. με επέκταση του πρωτοκόλλου ICMP [88] ή στην αποθήκευση πληροφοριών για την ανάκτηση των πηγών μέσα στους δρομολογητές [89]. Οι τεχνικές αυτές μέχρι σήμερα δεν έχουν χρησιμοποιηθεί. Τα θύματα μιας επίθεσης είναι δύσκολο να ευαισθητοποιήσουν τους διαχειριστές των πηγών της για τη λήψη μέτρων και την μελλοντική αποφυγή του προβλήματος. Η συνεργασία μεταξύ διαχειριστών συστημάτων και δικτύων στο παγκόσμιο Internet πραγματοποιείται κυρίως μέσα από τις δραστηριότητες οργανωμένων ομάδων αντιμετώπισης περιστατικών ασφάλειας (Computer

[6]Πρέπει να σημειωθεί πως συζητείται γενικότερα, αν είναι δεοντολογικά ορθό να αποκόπτεται ολοκληρωτικά η δικτυακή κίνηση ενός συστήματος για λόγους ασφάλειας.

Security Incident Response Team CSIRT[7]).

[7]Το όνομα CSIRT χρησιμοποιείται ως συνώνυμο του Computer Emergency Response Team (CERT) καθώς το όνομα CERT είναι καταχωρημένο στο Patent and Trademark Office των Η.Π.Α. από το Carnegie Mellon University

Κεφάλαιο 3

Συστήματα σύνθεσης δεδομένων (Data Fusion Systems)

3.1 Εισαγωγή στην σύνθεση δεδομένων

Για τον εντοπισμό μιας κατανεμημένης επίθεσης άρνησης υπηρεσιών και γενικότερα μίας ανωμαλίας κίνησης στο Διαδίκτυο απαιτείται συχνά συνεργασία διαφορετικών συστημάτων και αξιολόγηση πολλών πληροφοριών. Παρόμοιες απαιτήσεις έχουν και άλλα προβλήματα ανίχνευσης όπως για παράδειγμα ο εντοπισμός ενός εχθρικού αεροσκάφους στον εναέριο χώρο κάποιου κράτους. Για την επίλυση σύνθετων προβλημάτων έχουν αναπτυχθεί τα συστήματα σύνθεσης δεδομένων. Η **σύνθεση δεδομένων (data fusion)** αναφέρεται στην διαδικασία συλλογής πληροφοριών από πολλαπλές και πιθανόν ετερογενείς πηγές και στην σύνθεση τους με στόχο την εξαγωγή ενός περισσότερου περιγραφικού, μεγαλύτερης αφαιρετικότητας και σημασίας συμπεράσματος. Η διαδικασία αυτή είναι προφανής για τον ανθρώπινο εγκέφαλο. Όμως όπως θα δούμε παρακάτω η αυτοματοποίηση και η περιγραφή της με αλγόριθμους δεν είναι εύκολη. Συστήματα σύνθεσης δεδομένων έχουν αναπτυχθεί σε πολλούς τομείς της επιστήμης. Τα χαρακτηριστικότερα παραδείγματα είναι στρατιωτικά συστήματα για ανίχνευση και αξιολόγηση κινδύνων

και συστήματα μετεωρολογικών προγνώσεων. Επιχειρώντας να αποδώσουμε συνοπτικά την ουσία της σύνθεσης δεδομένων μπορούμε να πούμε πως **"είναι μια διαδικασία που χρησιμοποιεί ροές δεδομένων από πολλές πηγές και τις συνθέτει σε μία με μεγαλύτερη αφαιρετικότητα και περισσότερη χρήσιμη πληροφορία μέσα από στάδια ανίχνευσης, συσχέτισης και εκτιμήσεων (detection, association, correlation, estimation)."** Τα συστήματα σύνθεσης δεδομένων δεν ακολουθούν όλα μία ενιαία αρχιτεκτονική. Πέρα από τις ιδιαιτερότητες κάθε συστήματος υπάρχουν κοινά στάδια επεξεργασίας της πληροφορίας που συνθέτεται. Αντί για μια μακροσκελή παρουσίαση των διαφορετικών αρχιτεκτονικών θα παρουσιάσουμε τα βασικά στάδια επεξεργασίας συνοπτικά και λογικά συγκεντρωμένα στην ενότητα 3.2. Το γενικό πλαίσιο και η ορολογία που θα χρησιμοποιήσουμε μπορεί να συνοψιστεί ως εξής:

- Ένα σύστημα σύνθεσης δεδομένων έχει ως σκοπό να περιγράψει (χαρακτηρίσει) την κατάσταση στην οποία βρίσκεται το παρατηρούμενο περιβάλλον, που στο εξής θα αποκαλούμε **υπό παρακολούθηση σύστημα**. Υποθέτουμε ότι έχουν οριστεί οι πιθανές καταστάσεις του υπό παρακολούθηση συστήματος.

- Η διαδικασία αυτή πραγματοποιείται με χρήση πολλών **αισθητήρων** που εγκαθίστανται στο υπό παρακολούθηση σύστημα και συλλέγουν πληροφορίες.

- Με την βοήθεια των αισθητήρων ανιχνεύονται και αναγνωρίζονται επιμέρους αντικείμενα, οντότητες, γεγονότα, ενέργειες, σήματα κ.τ.λ., για τα οποία θα χρησιμοποιούμε τον αφηρημένο όρο **"στόχους" (targets)**.

- Ένας μηχανισμός απόφασης εκτιμά την κατάσταση στην οποία βρίσκεται το υπό παρακολούθηση σύστημα βάσει των δεδομένων που παρέχονται από τους αισθητήρες.

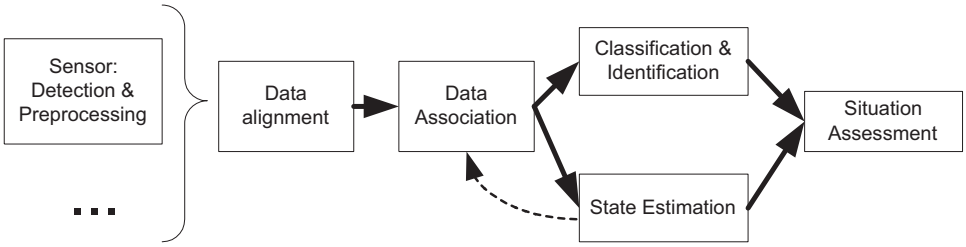
Η αρχική ιδέα της χρήσης αρχιτεκτονικών και αλγορίθμων από το χώρο των συστημάτων σύνθεσης δεδομένων για την ανάπτυξη συστημάτων ανίχνευσης επιθέσεων (Intrusion Detection Systems - IDS) αποδίδεται στον T. Bass [31]. Η παρούσα διατριβή παρουσιάζει συγκεκριμένους τρόπους χρήσης τους για την ανάπτυξη συστημάτων ανίχνευσης ανωμαλιών στη δικτυακή κίνηση (Network Anomaly Detection Systems - NADS) και εστιάζει στην ανίχνευση επιθέσεων καταιγισμού πακέτων. Συνεπώς επικεντρωνόμαστε στο πρόβλημα της ανίχνευσης ανωμαλιών, το δεύτερο βήμα (ενότητα 2.5.1.2) για την αντιμετώπιση του προβλήματος των επιθέσεων DDoS στο Διαδίκτυο. Η προσέγγιση μας προβλέπει την συλλογή πληροφοριών από πολλούς ανεξάρτητους αισθητήρες, που πιθανώς ακολουθούν διαφορετικές προσεγγίσεις ανίχνευσης, και την σύνθεση τους με χρήση ενός σαφούς μαθηματικού μοντέλου.

Η ενότητα 3.2 παρουσιάζει τα βασικά στάδια επεξεργασίας που συναντώνται σε συστήματα σύνθεσης δεδομένων. Στην ενότητα 3.3 παραθέτουμε μια ταξινόμηση των κυριότερων αλγόριθμων σύνθεσης δεδομένων και στην ενότητα 3.4 αναλύουμε τα κριτήρια επιλογής αλγορίθμου σύνθεσης. Η ενότητα 3.5 παρουσιάζει τη "θεωρία των Dempster-Shafer" που χρησιμοποιήσαμε στο πλαίσιο της προτεινόμενης αρχιτεκτονικής. Το κεφάλαιο ολοκληρώνεται στην ενότητα 3.5.3 με έναν σύντομο σχολιασμό της εφαρμογής της "Θεωρίας των Dempster-Shafer" στον χώρο της ανίχνευσης ανωμαλιών στη δικτυακή κίνηση.

3.2 Αρχιτεκτονική συστημάτων σύνθεσης δεδομένων

Τα βασικά στάδια επεξεργασίας που συναντώνται σε πολλά συστήματα σύνθεσης δεδομένων [32, 90] είναι (σχήμα 3.1):

- Συλλογή δεδομένων (Data Collection): **Αισθητήρες** μετρούν διάφορα χαρακτηριστικά, ανιχνεύουν και ανακοινώνουν επιμέρους στόχους που βοηθούν τον προσδιορισμό της κατάστασης του υπό παρακολούθηση συ-



Σχήμα 3.1: Τυπική αρχιτεκτονική συστήματος σύνθεσης δεδομένων.

στήματος.

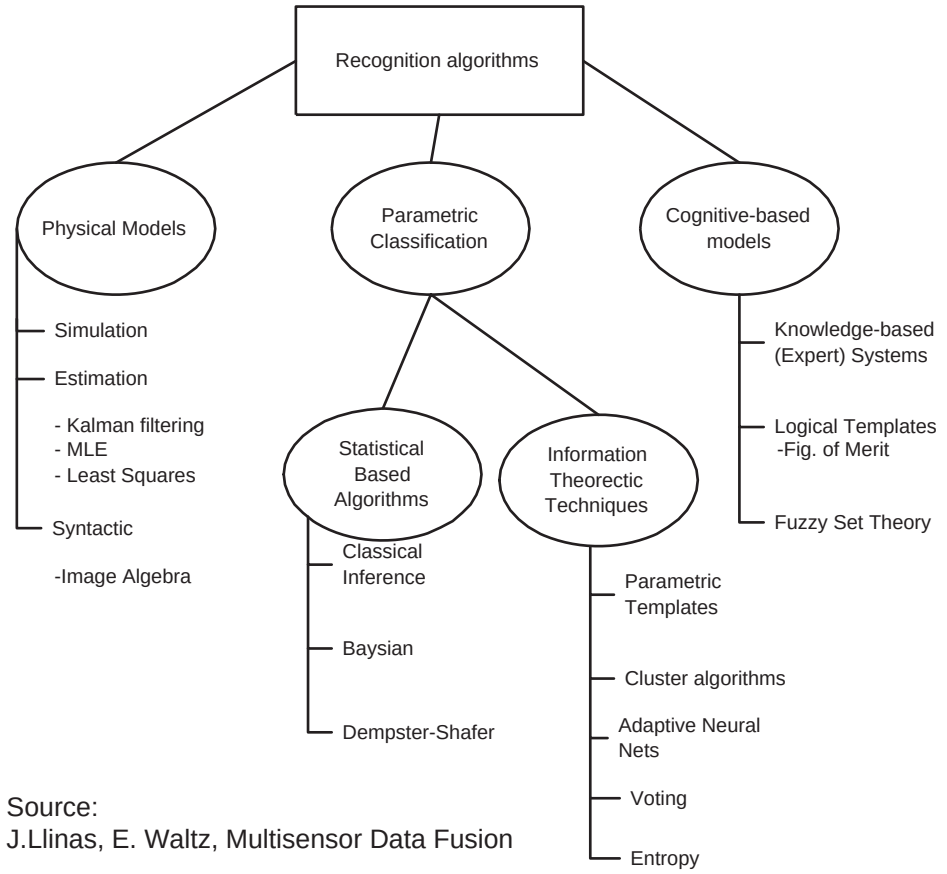
- Εναρμόνιση και συσχέτιση δεδομένων (Data Alignment & Association): Τα δεδομένα που προέρχονται από διαφορετικούς αισθητήρες μπορούν να παρουσιάζουν ανομοιογένεια στο χρόνο, χώρο ή στις μονάδες μέτρησης. Τα δεδομένα αυτά πρέπει να εναρμονιστούν πριν την σύνθεση τους.

- Εκτίμηση κατάστασης (State Estimation): Ένας **αλγόριθμος σύνθεσης δεδομένων** εκτιμά ποια είναι η τρέχουσα κατάσταση του υπό παρακολούθηση συστήματος βασιζόμενος στους στόχους και τις υπόλοιπες πληροφορίες που έχουν συλλεχθεί από τους αισθητήρες.

- Ταξινόμηση χαρακτηριστικών και αναγνώριση (Attribute classification & Identification): Στο στάδιο αυτό αναγνωρίζονται και ταξινομούνται οι διάφοροι στόχοι που έχουν ανιχνευθεί.

- Αξιολόγηση κατάστασης (Situation Assessment): Στο στάδιο αυτό υλοποιείται το υψηλότερο επίπεδο σύνθεσης πληροφοριών όπου βασιζόμενοι στις εκτιμήσεις της κατάστασης και στους στόχους που έχουν αναγνωριστεί (από τα δύο προηγούμενα στάδια επεξεργασίας) μπορούμε να συμπεράνουμε την γενική κατάσταση του υπό παρακολούθηση συστήματος.

Σε πολλές περιπτώσεις οι μαθηματικές μέθοδοι και οι αλγόριθμοι που χρησιμοποιούν τα παραπάνω στάδια επεξεργασίας είναι κοινοί. Αναφέρονται συ-



Σχήμα 3.2: Αλγόριθμοι σύνθεσης δεδομένων

χνά ως αλγόριθμοι σύνθεσης δεδομένων και προέρχονται από διάφορες επιστημονικές περιοχές. Ακόμη και μια σύντομη αναφορά σε αυτούς θα ήταν ιδιαίτερα μακροσκελής. Για το λόγο αυτό παραθέτουμε στην επόμενη ενότητα μια συνοπτική ταξινόμηση συχνά χρησιμοποιούμενων μεθόδων που δίνεται από τους Llinas, Waltz και Hall [90] και φαίνεται στο σχήμα 3.2. Η παρουσίαση μας εστιάζει στη χρήση τους για την ανίχνευση ανωμαλιών στη δικτυακή κίνηση. Μια πιο λεπτομερής ανάλυση των μεθόδων που φαίνονται στο σχήμα 3.2 υπάρχει στο [90].

3.3 Αλγόριθμοι σύνθεσης δεδομένων

3.3.1 Φυσικά μοντέλα (Physical models)

Στη πρώτη κατηγορία ανήκουν αλγόριθμοι που βασίζονται στην ακριβή μοντελοποίηση του υπό παρακολούθηση συστήματος και αντιστοιχίζουν τα μετρούμενα δεδομένα με ένα μοντέλο. Ένας από τους σημαντικότερους αντιπρόσωπους είναι το φίλτρο Kalman που βρίσκει την λύση (εκτίμηση κατάστασης) που ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα μεταξύ της πραγματικής κατάστασης του συστήματος και της εκτίμησης της. Απαιτεί όμως γνώση του πίνακα μετάβασης καταστάσεων (state transition matrix) και προϋποθέτει ότι οι μετρήσεις αλλοιώνονται από λευκό θόρυβο με γνωστό πίνακα συμμεταβλητότητας (covariance matrix). Η χρήση τέτοιων μεθόδων καθίσταται αρκετά δύσκολη με βάση όσα γνωρίζουμε για την σχέση της κατάστασης μιας δικτυακής ζεύξης (link), π.χ. ύπαρξη μιας επίθεσης, με τα χαρακτηριστικά της δικτυακής κίνησης που μεταφέρει.

3.3.2 Παραμετρική ταξινόμηση (Parametric classification)

Στην δεύτερη κατηγορία ανήκουν αλγόριθμοι που αντιστοιχίζουν χαρακτηριστικά (features) στο χώρο ταξινόμησης (classification space). Ένας τέτοιος αλγόριθμος αξιολογεί τα δεδομένα εισόδου (input data) για να συμπεράνει την κατάσταση του υπό παρακολούθηση συστήματος. Τα δεδομένα εισόδου είναι τα χαρακτηριστικά που χρησιμοποιούνται για την ταξινόμηση στον χώρο των καταστάσεων του συστήματος, δηλαδή στο χώρο ταξινόμησης. Ενδεικτικά θα αναφέρουμε ορισμένους δημοφιλείς αλγορίθμους:

- Οι στατιστικές μέθοδοι, όπως η συμπερασματολογία κατά Bayes (Bayesian inference) που παρουσιάζεται στην ενότητα 3.5.1, χρησιμοποιούν την έννοια των πιθανοτήτων για να εκφράσουν την πίστη ότι το σύστημα βρίσκεται σε μια πιθανή κατάσταση. Η Bayesian inference βασίζεται σε προϋπάρχουσα γνώση σχετικά με την πιθανότητα εμφάνισης των δεδο-

μένων εισόδου (χαρακτηριστικών) δεδομένης μιας πιθανής κατάστασης του υπό παρακολούθηση συστήματος. Η θεωρία Dempster-Shafer (ενότητα 3.5.2) αποτελεί επέκταση της Bayesian inference που προσφέρει ευελιξία στην μοντελοποίηση της αβεβαιότητας.

- Τα τεχνητά νευρωνικά δίκτυα (Artificial Neural Networks - ANN) αποτελούν μια ιδιαίτερα δημοφιλή και αποτελεσματική προσέγγιση για τη λύση προβλημάτων ταξινόμησης καθώς δεν απαιτούν μοντελοποίηση του υπό παρακολούθηση συστήματος. Ένα ANN λαμβάνει ως είσοδο ένα σύνολο χαρακτηριστικών και εμφανίζει ως έξοδο το αποτέλεσμα της ταξινόμησης που προκύπτει μετά την επιτυχημένη εκπαίδευση των νευρώνων του. Η εκπαίδευση των νευρώνων μπορεί να γίνει με μεθόδους επιβλεπόμενης (π.χ. υποδεικνύοντας το σωστό αποτέλεσμα ταξινόμησης για πλήθος δεδομένων εισόδου) και μη-επιβλεπόμενης μάθησης (π.χ. αυτο-οργάνωση) [91]. Τα νευρωνικά δίκτυα διαφοροποιούνται ανάλογα με την αρχιτεκτονική τους (αριθμό νευρώνων, κρυμμένα στρώματα), την συνάρτηση ενεργοποίησης κάθε νευρώνα (λογική κατωφλίου, σιγμοειδής συνάρτηση) και άλλα χαρακτηριστικά. Η χρήση νευρωνικών δικτύων έχει προταθεί κυρίως για συστήματα ανίχνευσης επιθέσεων (IDS) που εκμεταλλεύονται συγκεκριμένες αδυναμίες συστημάτων (exploits) [92].

- Η ψηφοφορία (Voting Methods) είναι μια από τις πιο εύκολα κατανοητές και υπολογιστικά απλές προσεγγίσεις. Τα δεδομένα εισόδου αντιστοιχίζονται σε ψήφους και το αποτέλεσμα της σύνθεσης είναι το αποτέλεσμα της ψηφοφορίας π.χ. με πλειοψηφικό κανόνα. Η μέθοδος αυτή είναι ιδιαίτερα χρήσιμη στην περίπτωση όπου δεν γνωρίζουμε εκ των προτέρων πιθανότητες (a priori statistics). Υπάρχουν πολλές παραλλαγές της προσέγγισης της ψηφοφορίας όπως η χρήση βαρών (weighted voting) και ενδιάμεσων αποτελεσμάτων σχηματίζοντας ένα δέντρο αποφάσεων (decision tree).

3.3.3 Γνωστικοί αλγόριθμοι (Cognitive algorithms)

Οι γνωστικοί αλγόριθμοι προσομοιάζουν τον τρόπο με τον οποίο εξάγει συμπεράσματα ο άνθρωπος. Χαρακτηριστικά παραδείγματα είναι:

- Η μέθοδος Logical Templating επιχειρεί την αντιστοίχηση των δεδομένων εισόδου, δηλαδή των παρατηρήσεων μας, σε προκαθορισμένα λογικά πρότυπα. Η μέθοδος αποτελεί ουσιαστικά αναγνώριση προτύπων (pattern recognition) με χρήση της μαθηματικής λογικής. Ένα τυπικό πρότυπο διαθέτει ένα κατώφλι αποδοχής, ένα κατώφλι απόρριψης καθώς και αναγκαίες (necessary conditions) και ικανές συνθήκες (sufficient conditions) για την ισχύ του. Οι συνθήκες αυτές εκφράζονται ως boolean expressions. Εξετάζοντας επαναληπτικά τα δεδομένα εισόδου και τα πρότυπα που έχουν οριστεί επιχειρείται να ικανοποιηθούν οι συνθήκες για την ισχύ τους. Μέσα από την διαδικασία αυτή τα διαφορετικά πρότυπα "συναγωνίζονται" για την επικράτηση κάποιου από αυτά. Είναι προφανές ότι η μέθοδος Logical Templating προϋποθέτει μια σαφή περιγραφή των πιθανών καταστάσεων του υπό παρακολούθηση συστήματος με την μορφή λογικών προτύπων. Η περιγραφή αυτή σε πολλές περιπτώσεις, όπως στην ανίχνευση ανωμαλιών στη δικτυακή κίνηση, είναι δύσκολη.

- Τα έμπειρα συστήματα (Expert Systems) αξιολογούν τα δεδομένα εισόδου με χρήση λογικών κανόνων που αποθηκεύονται σε μια βάση γνώσης (knowledge base). Η βάση γνώσης είναι ένα σύνολο λογικών προτάσεων και δημιουργείται από τους σχεδιαστές κάθε συστήματος που θεωρούνται ειδικοί στην συγκεκριμένη εφαρμογή για την οποία προορίζεται. Η γνώση που καταγράφεται μπορεί να είναι γεγονότα (facts), ακόμη και ολόκληροι αλγόριθμοι. Πέρα από τη βάση γνώσης τα Expert Systems έχουν έναν αλγόριθμο εξαγωγής συμπερασμάτων (inference algorithm) και μια μονάδα ελέγχου (rule interpreter). Σημαντικό πλεονέκτημα

της προσέγγισης αυτής είναι ότι μπορεί να αποτυπωθεί ο συλλογισμός που οδήγησε στην εξαγωγή ενός συμπεράσματος παρουσιάζοντας τους λογικούς κανόνες που οδήγησαν το σύστημα από τα δεδομένα εισόδου στο συμπέρασμα. Τα συστήματα αυτά δεν πραγματοποιούν πάντα εξαντλητική αναζήτηση στην βάση γνώσης για την εξαγωγή ενός συμπεράσματος. Η μαθηματική θεωρία στην οποία συνήθως στηρίζονται, είναι η κλασσική λογική "First Order Logic" [93]. Σημαντικό μειονέκτημα της κλασσικής λογικής είναι ότι δεν μπορεί να αποτυπώσει όλο το φάσμα μεταξύ πίστης και αμφισβήτησης μιας λογικής πρότασης (belief and disbelief) καθώς χρησιμοποιεί μόνο τους χαρακτηρισμούς αληθές και ψευδές. Επίσης η γνώση του συστήματος πρέπει να αυξάνει μονότονα αφού μια αληθής πρόταση δεν μπορεί στην συνέχεια να αναιρεθεί, δηλαδή να θεωρηθεί ψευδής (complete and consistent knowledge). Τα Expert Systems έχουν χρησιμοποιηθεί επιτυχώς σε πλήθος εφαρμογών. Για τη χρήση τους στην ανίχνευση ανωμαλιών στη δικτυακή κίνηση πρέπει να ξεπεραστεί η δυσκολία περιγραφής της διαδικασίας ανίχνευσης με λογικούς κανόνες.

- Τα ασαφή σύνολα και η ασαφής λογική (Fuzzy Logic) αποτελούν επεκτάσεις της κλασσικής λογικής για την εξαγωγή συμπερασμάτων σε συνθήκες αβεβαιότητας. Σύμφωνα με τη θεωρία της ασαφούς λογικής μια πρόταση δεν είναι απλά αληθής ή ψευδής αλλά συσχετίζεται με μια τιμή ανάμεσα στο 0 και 1, όπου 0 αντιπροσωπεύει την απολύτως ψευδή πρόταση και 1 την απόλυτα αληθή (συμμετοχή στο σύνολο αλήθειας). Η θεωρία της ασαφούς λογικής είναι ένα καλά θεμελιωμένο μαθηματικό πεδίο και περιέχει κανόνες ασαφούς συλλογιστικής (σύνθεσης). Η θεωρία των Dempster-Shafer, που υιοθετήσαμε στην αρχιτεκτονική μας, έχει αρκετά κοινά στοιχεία με την ασαφή λογική.

3.4 Κριτήρια επιλογής αλγόριθμου σύνθεσης δεδομένων

Όπως φαίνεται από την παραπάνω ταξινόμηση οι διαφορετικοί αλγόριθμοι σύνθεσης δεδομένων που μπορούμε να χρησιμοποιήσουμε στα πλαίσια της ανίχνευσης ανωμαλιών στη δικτυακή κίνηση είναι πολλοί. Στην ενότητα αυτή αναλύονται τα κριτήρια επιλογής αλγόριθμου σύνθεσης δεδομένων.

Η χρήση φυσικών μοντέλων είναι δύσκολη καθώς δεν έχουμε ένα καλό μοντέλο για την φυσιολογική κατάσταση του δικτύου και της δικτυακής κίνησης. Αν χρησιμοποιήσουμε στατιστικές μεθόδους πρέπει να αποφύγουμε να κάνουμε ισχυρές υποθέσεις για τα μετρούμενα μεγέθη χωρίς αυτές να επιβεβαιώνονται στατιστικά. Η χρήση στατιστικών μεθόδων είναι μια καλή επιλογή καθώς είναι μια ευρέως διαδεδομένη λύση για σχετικά προβλήματα ανίχνευσης, όπως ο εντοπισμός ανεπιθύμητων μηνυμάτων ηλεκτρονικού ταχυδρομείου (spam). Οι στατιστικές μέθοδοι μπορούν να χειριστούν εύκολα μεγάλο όγκο αριθμητικών δεδομένων.

Ο αλγόριθμος σύνθεσης δεδομένων που θα επιλεχθεί πρέπει να επιτρέπει την χρήση πληροφοριών από πολλαπλές ετερογενείς πηγές (αισθητήρες) με διαφορετική ευαισθησία, αξιοπιστία και πιθανότητα false alarm. Για παράδειγμα θα μπορούσαν να χρησιμοποιηθούν signature-based sensors που βασίζονται σε misuse detection με χαμηλά false positives, ταυτόχρονα με anomaly detection heuristics. Ο αλγόριθμος σύνθεσης δεδομένων πρέπει επιπλέον να μας επιτρέπει να αξιοποιήσουμε προϋπάρχουσα "εξειδικευμένη γνώση" (expert knowledge) π.χ. από τους διαχειριστές δικτύου που γνωρίζουν τα ιδιαίτερα χαρακτηριστικά του δικτύου-συστήματος τους. Η διαδικασία της σύνθεσης δεν πρέπει να βασίζεται εξολοκλήρου στην γνώση αυτή ή να απαιτεί τη χρήση ενός πολύπλοκου συνόλου λογικών κανόνων όπως στην περίπτωση των έμπειρων συστημάτων.

Για την υλοποίηση των επιμέρους αισθητήρων (στάδιο συλλογής δεδομέ-

νων) μπορούμε να επιλέξουμε οποιοδήποτε από τους προαναφερθέντες αλγόριθμους και προσεγγίσεις. Όμως για την σύνθεση δεδομένων (στάδιο εκτίμησης κατάστασης) προτιμούμε έναν ευέλικτο και εύχρηστο αλγόριθμο που θα μπορεί να συνθέτει δεδομένα από πολλούς ετερογενείς αισθητήρες. Όπως θα αναλύσουμε παρακάτω, η θεωρία των Dempster και Shafer (D-S) είναι μία πολλά υποσχόμενη μέθοδος για την σύνθεση δεδομένων και στηρίζεται σε ένα καλά θεμελιωμένο μαθηματικό υπόβαθρο από το χώρο των πιθανοτήτων. Η χρήση της δεν απαιτεί ισχυρές υποθέσεις για την πληροφορία που παρέχουν οι αισθητήρες ενώ παράλληλα είναι δυνατή η περιγραφή των ιδιαιτεροτήτων τους σε "υψηλό επίπεδο". Παράδειγμα της προσέγγισης που ακολουθούμε αποτελεί η χρήση ενός απλού νευρωνικού δικτύου στο επίπεδο ενός αισθητήρα (λόγω της ικανότητας μάθησης που διευκολύνει την διαδικασία της ρύθμισης) και η χρήση της θεωρίας D-S ως αλγόριθμο σύνθεσης δεδομένων (λόγω της ικανότητας μοντελοποίησης των αισθητήρων).

3.5 Εισαγωγή στη θεωρία Dempster-Shafer (D-S)

Η ενότητα αυτή παρουσιάζει τις βασικές έννοιες, ορισμούς και μαθηματικές σχέσεις της θεωρίας των Dempster και Shafer (D-S). Καθώς η θεωρία των D-S αποτελεί ουσιαστικά επέκταση της συμπερασματολογίας κατά Bayes η ενότητα 3.5.1 επαναλαμβάνει τα βασικά σημεία της. Οι όροι που χρησιμοποιούνται από τις μαθηματικές αυτές θεωρίες είναι γενικοί. Στα πλαίσια της εφαρμογής τους για την ανίχνευση ανωμαλιών στη δικτυακή κίνηση ορίζουμε :

- Το υπό παρακολούθηση σύστημα είναι ένα δικτυακό στοιχείο, π.χ. ζεύξη (link), δρομολογητής (router) κ.α.
- Στόχος μας είναι να συμπεράνουμε την πραγματική κατάσταση του υπό παρακολούθηση συστήματος, π.χ. κατάσταση φυσιολογική, ανωμαλίας.
- Οι ενδείξεις για την κατάσταση του υπό παρακολούθηση συστήματος

προέρχονται από την μέτρηση των χαρακτηριστικών της δικτυακής κίνησης (traffic features), για παράδειγμα τα μετρικά: ρυθμός αποστολής πακέτων TCP, UDP κτλ.

3.5.1 Συμπερασματολογία κατά Bayes (Bayesian inference)

Έστω ότι οι πιθανές καταστάσεις ενός συστήματος είναι $\theta_1, \theta_2, ..., \theta_N \in \Theta$ και ότι οι καταστάσεις αυτές είναι ξένες μεταξύ τους (αμοιβαία αποκλειστικές). Το σύστημα πρέπει να βρίσκεται πάντα σε μία από αυτές τις καταστάσεις. Η πιθανότητα $P(\theta_1)$ είναι μια έκφραση της πίστης ότι το σύστημα βρίσκεται στην κατάσταση θ_1 με απουσία κάποιας άλλης πληροφορίας (*a priori probability*). Έχοντας κάποια επιπλέον γνώση με μορφή μιας ένδειξης[1] E τότε η κατάλληλη έκφραση για την υπόθεση θ_1 είναι η υπό συνθήκη πιθανότητα (conditional probability) $P(\theta_1|E) = \frac{P(\theta_1,E)}{P(E)}$ που καλείται επίσης *a posteriori probability*. Το θεώρημα του Bayes υποδεικνύει:

$$P(\theta_1|E) = \frac{P(E|\theta_1)P(\theta_1)}{\sum_{i=1}^{N} P(E|\theta_i)P(\theta_i)} \tag{3.1}$$

Αν έχουμε πολλές ενδείξεις $E_1, ..., E_M$ τότε μπορούμε να τις συνθέσουμε προκειμένου να εκτιμήσουμε την κατάσταση του συστήματος. Σύμφωνα με τον ορισμό των υπό συνθήκη πιθανοτήτων έχουμε:

$$P(\theta_1|E_1, E_2, ..., E_M) = \frac{P(\theta_1, E_1, E_2, ..., E_M)}{P(E_1, E_2, ..., E_M)} \tag{3.2}$$

Αν θεωρήσουμε τα $E_1, E_2, ...E_M$ πλήρως ανεξάρτητα τότε από την 3.2 προκύπτει:

$$P(\theta_1|E_1, E_2, ..., E_M) = \frac{P(\theta_1)P(E_1, E_2, ..., E_M|\theta_1)}{P(E_1, E_2, ..., E_M)} \Rightarrow$$

$$P(\theta_1|E_1, E_2, ..., E_M) = P(\theta_1) \cdot \frac{P(E_1|\theta_1)}{P(E_1)} \cdot \frac{P(E_2|\theta_1)}{P(E_2)} \cdot ... \cdot \frac{P(E_M|\theta_1)}{P(E_M)} \tag{3.3}$$

[1] οι ενδείξεις στην περίπτωση μας προέρχονται από αισθητήρες (sensors) και εκφράζουν συγκεκριμένες τιμές μιας τυχαίας μεταβλητής.

$$οπου\ P(E_j) = \sum_{i=1}^{N} P(E_j|\theta_i)P(\theta_i),\ j = 1, ..., M$$

Αν πάλι θεωρήσουμε τα $E_1, E_2, ...E_M$ υπό συνθήκη ανεξάρτητα δεδομένης οποιασδήποτε υπόθεσης θ_i, δηλαδή $P(E_1, ..., E_M|\theta_i) = \prod_{j=1}^{M} P(E_j|\theta_i)$, από την 3.2 προκύπτει:

$$P(\theta_1|E_1, E_2, ..., E_M) = \frac{P(\theta_1)P(E_1, E_2, ..., E_M|\theta_1)}{\sum_{i=1}^{N} P(E_1, E_2, ..., E_M|\theta_i)P(\theta_i)} \Rightarrow$$

$$P(\theta_1|E_1, E_2, ..., E_M) = \frac{P(\theta_1) \cdot P(E_1|\theta_1) \cdot P(E_2|\theta_1) \cdot ... \cdot P(E_M|\theta_1)}{\sum_{i=1}^{N} P(E_1|\theta_i) \cdot P(E_2|\theta_i) \cdot ... \cdot P(E_M|\theta_i) \cdot P(\theta_i)} \Rightarrow$$

$$P(\theta_1|E_1, E_2, ..., E_M) = \frac{P(\theta_1) \cdot P(E_1|\theta_1) \cdot P(E_2|\theta_1) \cdot ... \cdot P(E_M|\theta_1)}{\sum_{i=1}^{N} \prod_{j=1}^{M} P(E_j|\theta_i) \cdot P(\theta_i)} \quad (3.4)$$

Με βάση την εξίσωση αυτή και θεωρώντας τα $P(\theta_i)$ και $P(E_j|\theta_i)$ γνωστά, ένας Bayesian classifier συμπεραίνει ότι το σύστημα βρίσκεται στην κατάσταση i όπου:

$$i = \arg\max_i P(\theta_i) \cdot P(E_1|\theta_i) \cdot P(E_2|\theta_i) \cdot ... \cdot P(E_M|\theta_i) \Rightarrow$$

$$i = \arg\max_i P(\theta_i) \prod_{j=1}^{M} P(Ej|\theta_i) \quad (3.5)$$

Τα μειονεκτήματα της προσέγγισης αυτής είναι ότι απαιτεί γνώση της *a priori* κατανομής πιθανότητας των καταστάσεων $P(\theta_1), ..., P(\theta_N)$. Περιοριστικό είναι επίσης το γεγονός ότι κάθε ένδειξη που επιβεβαιώνει μια υπόθεση θ_i ορίζει αυτόματα την απόρριψη της αντίθετης της. Η συμπερασματολογία κατά Bayes έχει χρησιμοποιηθεί για ανίχνευση επιθέσεων τύπου DoS από τους Hershkop et al [94] και τύπου DDoS από τους Noh et al [77].

3.5.2 Θεωρία των Dempster και Shafer

Η θεωρία των Dempster και Shafer ή απλά θεωρία D-S, μπορεί να θεωρηθεί ως επέκταση της συμπερασματολογίας κατά Bayes [93, 95]. Τα θεμέλια της

ορίστηκαν από τον Shafer το 1976 [96]. Η θεωρία μπορεί να ερμηνευθεί με διαφορετικούς τρόπους, π.χ. από μια πιθανοτική ή αξιωματική οπτική γωνία. Μια πολύ καλή και συνοπτική ανασκόπηση των θεωρήσεων αυτών δίνεται από τους Kohla et al [97]. Παρά τις διαφορετικές διατυπώσεις ανάλογα με τον χώρο εφαρμογής της θεωρίας (όπως statistical inference, diagnostics, risk analysis και decision analysis) όλες οι προσεγγίσεις καταλήγουν στους ίδιους μαθηματικούς τύπους. Η δική μας παρουσίαση θυμίζει κυρίως τον χώρο της διαγνωστικής (diagnostics) [98].

Έστω ότι οι πιθανές καταστάσεις ενός συστήματος είναι $\theta_1, \theta_2, ..., \theta_N \in \Theta$ και ότι οι καταστάσεις αυτές είναι ξένες μεταξύ τους (αμοιβαίως αποκλειστικές). Το σύστημα πρέπει να βρίσκεται πάντα σε μία από τις καταστάσεις αυτές. Στα πλαίσια της θεωρίας D-S το σύνολο Θ ονομάζεται *frame of discernment* που μπορεί να αποδοθεί στα ελληνικά ως διαισθητικό πλαίσιο ή πλαίσιο κρίσεως. Θα ονομάζουμε **υποθέσεις** H_i υποσύνολα του Θ, δηλαδή στοιχεία του δυναμοσυνόλου (powerset) 2^Θ.

Η θεωρία μας βοηθά να αποφασίσουμε για τη πραγματική κατάσταση του συστήματος χωρίς να διαθέτουμε ένα συγκεκριμένο μοντέλο για αυτό αλλά απλά βάσει ορισμένων παρατηρήσεων. Οι παρατηρήσεις (μετρήσεις αισθητήρων) μπορούν να θεωρηθούν ως ενδείξεις $E_1, ..., E_M$ προς ορισμένες καταστάσεις του συστήματος (με κάποιο βαθμό αβεβαιότητας).

Για κάθε ένδειξη E_j ορίζεται μια συνάρτηση μάζας (mass function) m_j που αναθέτει πεποιθήσεις $m_j(H_i)$ (beliefs) σε υποθέσεις H_i ή χρησιμοποιώντας τη φράση του Shafer "the measure of belief that is committed exactly to H" [96]. Η συνάρτηση m_j μεταβάλλεται ανάλογα με την έξοδο του αισθητήρα j (ένδειξη E_j). Η συνάρτηση m_j ονομάζεται **basic probability assignment (bpa)**.

$$m_j : 2^\Theta \rightarrow [0,1] \tag{3.6}$$

ή ισοδύναμα

$$m_j(H_i) \in [0,1] \tag{3.7}$$

Η συνάρτηση m_j υπόκειται στους ακόλουθους περιορισμούς:

$$m_j(\emptyset) = 0$$

$$m_j(H) \geq 0, \forall H \subseteq \Theta$$

$$\sum_H m_j(H) = 1, \forall H \subseteq \Theta \tag{3.8}$$

Κάθε υπόθεση H για την οποία $m_j(H) > 0$ ονομάζεται *focal set* και το σύνολο όλων των focal sets αποκαλείται *core*.

Στο σημείο αυτό μπορούμε ήδη να σχολιάσουμε την ευελιξία που μας προσφέρει η θεωρία D-S σε σχέση με την προσέγγιση Bayes, όπου μπορούμε να αναθέσουμε πιθανότητες μονάχα σε απλά στοιχεία του Θ και όχι σε ολόκληρα υποσύνολα του, δηλαδή σύνολα πιθανών καταστάσεων. Το γεγονός αυτό μας δίνει την δυνατότητα να μοντελοποιήσουμε την άγνοια μας. Για παράδειγμα ορισμένες παρατηρήσεις μπορούν να διαχωρίσουν ορισμένες καταστάσεις του συστήματος αλλά παράλληλα αδυνατούν να προσφέρουν οποιαδήποτε ένδειξη για άλλες. Δηλαδή μπορεί να γνωρίζουμε ότι μια ένδειξη E_1 παραπέμπει σε μια υπόθεση $H_1 = \{\theta_1, \theta_2\}$ με μεγάλη πιθανότητα αλλά παράλληλα δεν προσφέρει καμία πληροφορία (πλήρη άγνοια) κατά πόσο το σύστημα είναι στην κατάσταση θ_1 ή θ_2.

Είναι σημαντικό ότι η θεωρία των Dempster-Shafer υπολογίζει τη πιθανότητα ότι οι ενδείξεις υποστηρίζουν μια υπόθεση και όχι την πιθανότητα αυτής καθαυτής της υπόθεσης όπως στην κλασσική θεωρία πιθανοτήτων.

Επιστρέφοντας στο θεωρητικό υπόβαθρο ορίζουμε μια συνάρτηση πεποίθησης (belief function) Bel_j, που περιγράφει τον βαθμό πίστης στην υπόθεση H, ως εξής:

$$Bel_j(H) = \sum_{B|B \subseteq H} m_j(B) \tag{3.9}$$

'Οπου η έκφραση $B|B \subseteq H$ ισοδυναμεί με $\forall B \subseteq H$.

Ο ορισμός αυτός λέει ουσιαστικά ότι ο βαθμός πίστης μας σε μια υπόθεση

H υπολογίζεται αν αθροίσουμε το βαθμό πίστης από κάθε υπόθεση B που περιέχεται στην H, δηλαδή από κάθε $B \subseteq H$. Μια συνάρτηση πεποίθησης έχει τις εξής ιδιότητες:

$$Bel_j(\emptyset) = 0$$

$$Bel_j(\Theta) = 1$$

Παρόμοια ορίζεται η ευλογοφάνεια (Plausibility) μιας υπόθεσης H:

$$Pl_j(H) = \sum_{B|B\cap H\neq\emptyset} m_j(B) \tag{3.10}$$

Η εξίσωση αυτή ορίζει ότι η ευλογοφάνεια μιας υπόθεσης H υπολογίζεται αν αθροίσουμε το βαθμό πίστης από κάθε υπόθεση B που συμπίπτει με την H, δηλαδή $B \cap H \neq \emptyset$. Η ευλογοφάνεια μπορεί να συσχετιστεί με τις αμφιβολίες (doubt) κατά της υπόθεσης H ως εξής:

$$Pl_j(H) = 1 - Doubt_j(H) = 1 - Bel_j(\neg H) \tag{3.11}$$

όπου $\neg H$ είναι το συμπλήρωμα (complement) του H. Η εξίσωση αυτή σημαίνει ότι όσο λιγότερες αμφιβολίες έχουμε για μια υπόθεση H τόσο πιο εύλογη είναι (plausible). Γενικά μπορούμε να χαρακτηρίσουμε το $Bel(_jH)$ ως ένα ποσοτικό μέτρο όλων των στοιχείων που υποστηρίζουν την υπόθεση H (supportive evidence) και το $Pl_j(H)$ ως μέτρο της συμβατότητας των στοιχείων μας με την υπόθεση H, δηλαδή την έλλειψη αμφιβολιών-αντιφάσεων (doubt). Η θεωρία κάνει σαφή διαχωρισμό μεταξύ της έλλειψης υποστήριξης και της ύπαρξης αμφιβολιών για μια υπόθεση (απόρριψη). Αυτό εκφράζεται ως εξής:

- Η έλλειψη υποστήριξης $Bel_j(H) \approx 0$ δεν συνεπάγεται κατ' ανάγκη $Bel_j(\neg H) \approx 1$ ή $Pl_j(H) \approx 0$, δηλαδή δεν προκύπτει αναγκαστικά ύπαρξη αμφιβολιών για την H.

- Αντίθετα, η ύπαρξη αμφιβολιών $Bel_j(\neg H) \approx 1$ συνεπάγεται έλλειψη

υποστήριξης $Bel_j(H) \approx 0$.

Η πραγματική πίστη (true belief) σε μια υπόθεση H βρίσκεται στο διάστημα $[Bel_j(H), Pl(_jH)]$ που ονομάζεται διάστημα εμπιστοσύνης. Ο βαθμός αβεβαιότητας μας εκφράζεται από τη διαφορά $Pl_j(H) - Bel_j(H)$. Η ερμηνεία των διαστημάτων εμπιστοσύνης φαίνεται στον πίνακα 3.1.

$[Bel_j(H), Pl_j(H)]$	Ερμηνεία
$[0,0]$	Η υπόθεση H είναι ψευδής
$[1,1]$	Η υπόθεση H είναι αληθής
$[0,1]$	Πλήρης απουσία πληροφορίας για την αλήθεια της υπόθεσης H
$[p,p]$	Η υπόθεση H είναι αληθής με πιθανότητα $p \in [0,1]$
$[p,1]$	Τάση για πίστη στην υπόθεση H, όπου $p \in [0,1]$
$[0,q]$	Τάση για δυσπιστία στην υπόθεση H, όπου $q \in [0,1]$
$[p,q]$	Αβεβαιότητα, η πιθανότητα της υπόθεσης H βρίσκεται μεταξύ p και q, όπου $0 < p \leq q < 1$

Πίνακας 3.1: Ερμηνεία διαστημάτων εμπιστοσύνης

Ένα πολύ σημαντικό στοιχείο της θεωρίας των Dempster-Shafer αποτελεί ο κανόνας για την σύνθεση ανεξάρτητων ενδείξεων E_1, E_2 σε μια πιο συνοπτική μορφή $m_{12} = m_1 \oplus m_2$.

$$m_{12}(H) = \frac{\sum_{B,C|B\cap C=H} m_1(B)m_2(C)}{1 - \sum_{B,C|B\cap C=\emptyset} m_1(B)m_2(C)} \; \forall B, C \in 2^{\Theta}, H \neq \emptyset \qquad (3.12)$$

Η αλγοριθμική πολυπλοκότητα του **κανόνα σύνθεσης του Dempster** είναι στην χειρότερη περίπτωση εκθετική, αφού απαιτείται να βρεθούν όλα τα ζεύγη των συνόλων B, C όπου $B \cap C = H$. Συγκεκριμένα είναι $O(2^{|\Theta|-|H|} \times 2^{|\Theta|-|H|})$, όπου $|\Theta|$ ο αριθμός των πιθανών καταστάσεων και $|H|$ ο αριθμός των καταστάσεων που ορίζουν την υπόθεση H. Σε περίπτωση που το γεγονός αυτό αποτελέσει πρακτικό πρόβλημα ένας γρήγορος αλγόριθμος υπολογισμού του κανόνα του Dempster δίνεται από τους Wilson et al [99]. Στις περισσότερες εφαρμογές, όπου το πλήθος των καταστάσεων είναι μικρό, η χρήση μιας απλής υλοποίησης, χωρίς προσπάθεια αντιμετώπισης της πολυπλοκότητας, είναι δυνατή.

Ο κανόνας του Dempster είναι αντιμεταθετικός και προσεταιριστικός [95] και επομένως πολλές ενδείξεις $E_1, ..., E_M$, που κωδικοποιούνται με M basic probability assignment (bpa) $m_1, ..., m_M$, μπορούν να συνδυάζονται ανά δύο:

$$m(H) = m_1 \oplus ... \oplus m_M = (...(m_1 \oplus m_2) \oplus m_3)...) \oplus m_{M-1}) \oplus m_M)...)$$

Συνεπώς με επαναληπτική χρήση του απλού κανόνα του Dempster μπορούν να συνδυαστούν περισσότερες ενδείξεις προκειμένου να εκτιμηθεί η πραγματική κατάσταση του συστήματος. Η θεωρία μας παρέχει λοιπόν τον κανόνα συνδυασμού ενδείξεων και το συμπέρασμα εκφράζεται με ένα νέο bpa. Η θεωρία όπως διατυπώθηκε από τον Shafer δεν ορίζει έναν συγκεκριμένο κανόνα απόφασης (συνάρτηση) για την πραγματική κατάσταση του συστήματος. Δηλαδή δεν ορίζεται ένας μηχανισμός απόφασης της πραγματικής κατάστασης θ_i του συστήματος βάσει των τιμών των συναρτήσεων belief και plausibility της συνδυασμένης m. Στα πλαίσια της παρούσας διατριβής χρησιμοποιήσαμε το κανόνα της μέγιστης πίστης [100]: Υποθέτουμε ότι η κατάσταση του συστήματος είναι η θ_k όπου

$$k = \arg \max_{1 \leq l \leq N} [Bel(\theta_l)] \tag{3.13}$$

Μάλιστα καθώς τα θ_l είναι απλές υποθέσεις $Bel(\theta_l) = m(\theta_l)$

$$k = \arg \max_{1 \leq l \leq N} [m(\theta_l)]$$

Συνοψίζοντας, η θεωρία των Dempster-Shafer κάνει σαφή διαχωρισμό μεταξύ αβεβαιότητας και άγνοιας (uncertainty and ignorance) [101] και για το λόγο αυτό είναι χρήσιμη στην "συλλογιστική με αβεβαιότητα με χρήση ελλιπών και πιθανώς αντικρουόμενων πληροφοριών" (reason with uncertainty based on incomplete and possibly contradictory information). Καθώς μάλιστα δεν απαιτεί γνώση με τη μορφή των *a priori* κατανομών πιθανοτήτων των

πιθανών καταστάσεων του συστήματος (όπως η μέθοδος Bayes) είναι ιδιαίτερα χρήσιμη όταν δεν έχουμε ένα μοντέλο για το σύστημα μας. Η χρήση της θεωρίας Dempster-Shafer σε ένα άγνωστο, ασαφές και αβέβαιο περιβάλλον είναι προτιμότερη από τη χρήση του κατηγορικού λογισμού πρώτης τάξης (First Order Logic) που υποθέτει πλήρη και συνεπή γνώση (complete and consistent knowledge) [2] ή τη κλασσική θεωρία των πιθανοτήτων που προϋποθέτει γνώση με την μορφή *a priori* κατανομών πιθανοτήτων.

3.5.3 Χρήση της θεωρίας Dempster-Shafer στην ανίχνευση δικτυακών ανωμαλιών

Η επιλογή της θεωρίας D-S για την σύνθεση δεδομένων στα πλαίσια ενός συστήματος ανίχνευσης δικτυακών ανωμαλιών σχολιάστηκε ήδη στην ενότητα 3.4. Έχοντας πλέον αναλύσει την μαθηματική της διατύπωση είμαστε σε θέση να επεκταθούμε στις δυνατότητες μοντελοποίησης που μας προσφέρει.

Στη παρούσα διατριβή ως **υπό παρακολούθηση σύστημα** θεωρούμε ένα δικτυακό στοιχείο και συγκεκριμένα **μια δικτυακή ζεύξη** (link). Η θεωρία D-S απαιτεί τον ορισμό όλων των καταστάσεων του υπό παρακολούθηση συστήματος. Υπάρχουν πολλές δυνατότητες για την επιλογή των καταστάσεων ανάλογα με την σκοπιμότητα της εφαρμογής μας, δηλαδή τι ακριβώς θέλουμε να ανιχνεύσουμε. Στην απλούστερη περίπτωση μπορούμε να ορίσουμε δυο καταστάσεις: Θ ={φυσιολογική κατάσταση, ύπαρξη ανωμαλίας} ή Θ ={NORMAL, ANOMALY}. Η λογική αυτή μπορεί να επεκταθεί ορίζοντας διαφορετικούς τύπους ανωμαλιών που για παράδειγμα απαιτούν διαφορετικό τρόπο αντιμετώπισης ή είναι διαφορετικής επικινδυνότητας π.χ. Θ ={NORMAL, SYN-flood, UDP-flood, ICMP-flood}. Ένας άλλος πιθανός τρόπος επέκτασης είναι διαχωρίζοντας τις εισερχόμενες από τις εξερχόμενες δικτυακές ανωμαλίες σε μια ζεύξη Θ ={NORMAL, ATTACK SOURCE,

[2] ο κατηγορικός λογισμός πρώτης τάξης είναι "μονότονος", δηλαδή μια αληθής πρόταση δεν μπορεί στην συνέχεια να αναιρεθεί (να θεωρηθεί ψευδής), με αποτέλεσμα τα συμπεράσματα να είναι προσθετικά [93].

ATTACK DESTINATION}. Γενικά βλέπουμε δηλαδή ότι υπάρχει αρκετά μεγάλη ελευθερία στον τρόπο ορισμού των καταστάσεων του συστήματος, αρκεί οι καταστάσεις να είναι πλήρεις και αλληλο-αποκλειόμενες (exhaustive and mutually exclusive). Το πλήθος των καταστάσεων $|\Theta|$ πρέπει να διατηρείται στον μικρότερο δυνατό αριθμό αφού καθορίζει την αλγοριθμική πολυπλοκότητα του κανόνα της σύνθεσης.

Ως αισθητήρες (sensors) ορίζονται τα υποσυστήματα (ενότητα 3.2) που συλλέγουν και αναλύουν πληροφορίες για την κατάσταση του υπό παρακολούθηση συστήματος (δικτυακής ζεύξης). Κάθε **αισθητήρας πέρα από συλλέκτης πληροφοριών, αποτελεί και μία στοιχειώδη μονάδα ανίχνευσης**, δηλαδή εφαρμόζει έναν αλγόριθμο ανίχνευσης σε ένα ή περισσότερα μετρούμενα χαρακτηριστικά του υπό παρακολούθηση δικτυακού στοιχείου. Εμείς θεωρούμε ότι **ένας αισθητήρας βασίζεται σε ένα μετρικό** (αντιστοίχηση ένα προς ένα) χωρίς όμως η σχέση αυτή να είναι γενικά περιοριστική. Σύμφωνα με τη θεωρία D-S απαιτείται όλοι οι αισθητήρες να αναφέρονται πάνω στο ίδιο σύνολο πιθανών καταστάσεων, δηλαδή στο σύνολο Θ (Frame of Discernment), χωρίς βέβαια να είναι υποχρεωτικό να εκφέρουν θετική ή αρνητική γνώμη για όλες τις καταστάσεις.

Οι αισθητήρες βασιζόμενοι στα χαρακτηριστικά που παρακολουθούν, προσφέρουν περιοδικά ενδείξεις για την κατάσταση του υπό παρακολούθηση συστήματος (δικτυακής ζεύξης). Κάθε ένδειξη E_j ενός αισθητήρα j κωδικοποιείται μαθηματικά με τον ορισμό ενός basic probability assignment (bpa) δηλαδή με τον ορισμό μιας συνάρτησης $m_j : 2^{\Theta} \rightarrow [0, 1]$. Ένας αισθητήρας μπορεί να ομαδοποιεί ή να διαχωρίζει τις καταστάσεις ανάλογα με την διακριτική του ικανότητα. Ένας αισθητήρας μπορεί να προσφέρει μια ένδειξη που οδηγεί σε περισσότερες από μια πιθανές καταστάσεις καθώς η αναλογία αισθητήρα προς υπόθεση δεν είναι αναγκαστικά ένα προς ένα. Ιδιαίτερα σημαντική είναι η δυνατότητα έκφρασης "αποθαρρυντικών" στοιχείων για μια κατάσταση (refuting evidence). Δηλαδή ένας αισθητήρας μπορεί να ενθαρρύ-

νει αλλά και να αποθαρρύνει μια απόφαση, δηλαδή ότι το σύστημα βρίσκεται σε μια συγκεκριμένη κατάσταση. Επίσης δίνεται η δυνατότητα να εκφράζει το βαθμό αμφιβολίας στις εκτιμήσεις του αναθέτοντας τιμές στο σύνολο όλων των πιθανών καταστάσεων $m_j(\Theta)$.

Ένα από τα σημεία που έχει θεωρηθεί ως μειονέκτημα της θεωρίας D-S είναι η αδυναμία του κανόνα σύνθεσης του Dempster σε περίπτωση σύνθεσης ισχυρά αντιφατικών ενδείξεων. Το γεγονός αυτό οφείλεται στον παρανομαστή της εξίσωσης 3.12 που μοιράζει ομοιόμορφα μεταξύ των focal sets την τιμή που προσωρινά αποδίδεται στο κενό σύνολο κατά τον συνδυασμό δυο basic probability assignments (κανονικοποίηση). Την περίπτωση αυτή παρουσιάζουμε το παράδειγμα Β που ακολουθεί. Για να αντιμετωπιστεί το φαινόμενο αυτό που σχολιάστηκε πρώτη φορά από τον Zadeh [102], οι Yager, Inagaki, Debois και Zhang έχουν προτείνει διάφορους εναλλακτικούς κανόνες σύνθεσης [95, 103–107]. Ένα ακόμη στοιχείο που μπορεί να θεωρηθεί ως μειονέκτημα είναι το γεγονός ότι η θεωρία μας υποχρεώνει να ορίσουμε όλες τις πιθανές καταστάσεις του συστήματος ώστε να είναι ξένες και πλήρεις. Δηλαδή το σύστημα βρίσκεται πάντα σε μία και μόνο μια από τις καταστάσεις αυτές. Συνεπώς αν θεωρήσουμε ως καταστάσεις διαφορετικά είδη ανωμαλιών τότε υποθέτουμε ότι αυτές δεν θα εμφανίζονται ταυτόχρονα. Το πρόβλημα αυτό μπορεί να αντιμετωπιστεί θεωρώντας την ύπαρξη ταυτόχρονων ανωμαλιών ως ξεχωριστές στοιχειώδεις καταστάσεις, δηλαδή επεκτείνοντας το σύνολο Θ με καταστάσεις όπως SYN-flood-and-UDP-flood, SYN-flood-and-ICMP-flood.

Η χρήση της θεωρίας D-S αποκτά ιδιαίτερη σημασία όταν οι αισθητήρες που χρησιμοποιούνται είναι πολλοί και διαφορετικοί, έχουν δηλαδή διαφορετική γνώση. Η σύνθεση αυτών των διαφορετικών πηγών πληροφορίας μπορεί να γίνει μέσα σε ένα απλό και ενιαίο μαθηματικό πλαίσιο. Συγκεκριμένα τα διαφορετικά basic probability assignments (bpa) που δίνονται από τους αισθητήρες συνθέτονται με χρήση της εξίσωσης 3.12. Για να γίνει κατανοητή η διαδικασία αυτή παραθέτουμε δύο απλοποιημένα παραδείγματα χρήσης της

θεωρίας D-S για την σύνθεση δεδομένων από ανιχνευτές ανωμαλιών στη δικτυακή κίνηση.

3.5.3.1 Παράδειγμα Α

Έστω ότι οι καταστάσεις του υπό παρακολούθηση συστήματος (δικτυακής ζεύξης) είναι : Θ ={NORMAL, SYN-flood, UDP-flood}. Έχουμε δύο αισθητήρες όπου ο πρώτος μετρά το ρυθμό των εισερχόμενων πακέτων τύπου TCP SYN και ο δεύτερος μετρά τον αριθμό των ενεργών flows. Αξιολογώντας τις μετρήσεις αυτές, οι αισθητήρες προσφέρουν ενδείξεις E_1, E_2 που κωδικοποιούνται με τα basic probability assignments m_1 και m_2 αντίστοιχα. Έστω ότι ο αισθητήρας 1 ανακοινώνει:

m_1({SYN-flood}) =0.62

m_1({NORMAL,UDP-flood})=0.23

$m_1(\Theta)$=0.15

και ο αισθητήρας 2 ανακοινώνει :

m_2({SYN-flood,UDP-flood})=0.55

m_2({NORMAL})=0.15

$m_2(\Theta)$=0.30

Συνθέτουμε τις ενδείξεις από τους δύο αισθητήρες εφαρμόζοντας τον κανόνα σύνθεσης του Dempster: $m = m_1 \oplus m_2$. Το πρώτο βήμα υπολογίζει τις τομές των επιμέρους υποθέσεων ανά δύο πολλαπλασιάζοντας τις τιμές των συναρτήσεων $m1$, $m2$:

	m_1({SYN-flood})=0.62	m_1({NORMAL,UDP-flood})=0.23	$m_1(\Theta)$=0.15
m_2({NORMAL})=0.15	$\emptyset$ =0.093	{NORMAL}=0.0345	{NORMAL}=0.0225
m_2({SYN-flood,UDP-flood})=0.55	{SYN-flood} =0.341	{UDP-flood}=0.1265	{SYN-flood,UDP-flood}=0.0825
$m_2(\Theta)$=0.30	{SYN-flood} =0.186	{NORMAL,UDP-flood}=0.069	Θ=0.045

Στο δεύτερο βήμα γίνεται συγχώνευση των τιμών του προηγούμενου πίνακα που αντιστοιχούν στις ίδιες υποθέσεις:

$m_{12}(H) = m_1 \oplus m_2(H) = \frac{\sum_{B,C\|B\cap C=H} m_1(B)m_2(C)}{1-\sum_{B,C\|B\cap C=\emptyset} m_1(B)m_2(C)}$
$m_{12}(\{NORMAL\}) = \frac{0.0345+0.0225}{1-0.093} = \frac{0.057}{1-0.093} = 0.063$
$m_{12}(\{SYN-flood\}) = \frac{0.341+0.186}{1-0.093} = \frac{0.527}{1-0.093} = 0.581$
$m_{12}(\{UDP-flood\}) = \frac{0.1265}{1-0.093} = 0.1395$
$m_{12}(\{SYN-flood, UDP-flood\}) = \frac{0.0825}{1-0.093} = 0.091$
$m_{12}(\{NORMAL, UDP-flood\}) = \frac{0.069}{1-0.093} = 0.076$
$m_{12}(\Theta) = \frac{0.045}{1-0.093} = 0.0496$

3.5.3.2 Παράδειγμα Β

Στο παράδειγμα αυτό θα δούμε την ευαισθησία του κανόνα σύνθεσης του Dempster στην περίπτωση σύνθεσης δεδομένων από δύο αισθητήρες με αντικρουόμενες ενδείξεις. Έστω ότι ο αισθητήρας 1 ανακοινώνει:

$m_1(\{\text{SYN-flood}\}) = 0.9$

$m_1(\Theta) = 0.1$

και ότι ο αισθητήρας 2 ανακοινώνει :

$m_2(\{\text{UDP-flood}\}) = 0.6$

$m_2(\{\text{NORMAL}\}) = 0.4$

Συνθέτουμε τις ενδείξεις από τους δύο αισθητήρες εφαρμόζοντας τον κανόνα σύνθεσης του Dempster: $m = m_1 \oplus m_2$. Το πρώτο βήμα για τον υπολογισμό είναι:

	$m_1(\{\text{SYN-flood}\})=0.9$	$m_1(\Theta)=0.1$
$m_2(\{\text{NORMAL}\})=0.4$	$\emptyset = 0.36$	$\{\text{NORMAL}\}=0.04$
$m_2(\{\text{UDP-flood}\})=0.6$	$\emptyset=0.54$	$\{\text{UDP-flood}\}=0.06$

και το δεύτερο βήμα :

$m_{12}(H) = m_1 \oplus m_2(H) = \frac{\sum_{B,C\|B\cap C=H} m_1(B)m_2(C)}{1-\sum_{B,C\|B\cap C=\emptyset} m_1(B)m_2(C)}$
$m_{12}(\{NORMAL\}) = \frac{0.04}{1-0.90} = 0.4$
$m_{12}(\{UDP-flood\}) = \frac{0.06}{1-0.90} = 0.6$

Παρατηρούμε ότι εκμηδενίζεται η πεποίθηση ύπαρξης επίθεσης SYN. Ο δεύτερος αισθητήρας μέσα από τον ορισμό της m_2 αποκλείει την ύπαρξη επίθεσης SYN καθώς $Pl_2(\{\text{SYN-flood}\})= 0$ και άρα $Doubt_2(\{\text{SYN-flood}\})= 1-0 = 1$. Αντίθετα ο πρώτος αισθητήρας δεν αποκλείει την ύπαρξη επίθεσης UDP αναθέτοντας την τιμή 0.1 στο Θ.

Αν ο δεύτερος αισθητήρας δεν ήταν απόλυτος και ανακοίνωνε :

$m_2(\{\text{UDP-flood}\})$=0.6

$m_2(\{\text{NORMAL}\})$=0.3

$m_2(\Theta)$=0.1

θα είχαμε:

	$m_1(\{\text{SYN-flood}\})$=0.9	$m_1(\Theta)$=0.1
$m_2(\{\text{NORMAL}\})$=0.3	$\emptyset$ =0.27	{NORMAL}=0.03
$m_2(\{\text{UDP-flood}\})$=0.6	$\emptyset$=0.54	{UDP-flood}=0.06
$m_2(\Theta)$=0.1	{SYN-flood} =0.09	Θ=0.01

και το δεύτερο βήμα :

$m_{12}(H) = m_1 \oplus m_2(H) = \frac{\sum_{B,C\|B\cap C=H} m_1(B)m_2(C)}{1-\sum_{B,C\|B\cap C=\emptyset} m_1(B)m_2(C)}$
$m_{12}(\{NORMAL\}) = \frac{0.03}{1-0.81} = 0.16$
$m_{12}(\{UDP-flood\}) = \frac{0.06}{1-0.81} = 0.32$
$m_{12}(\{SYN-flood\}) = \frac{0.09}{1-0.81} = 0.47$
$m_{12}(\Theta) = \frac{0.01}{1-0.81} = 0.05$

Στην περίπτωση αυτή η πιο πιθανή κατάσταση είναι η επίθεση SYN-flood. Το συμπέρασμα είναι ότι οι αισθητήρες δεν πρέπει να είναι απόλυτοι σε ότι

ανακοινώνουν και να συμπεριλαμβάνεται πάντα η πιθανότητα να κάνουν λάθος γιατί σε αντίθετη περίπτωση μπορεί να μας οδηγήσουν σε εσφαλμένες εκτιμήσεις. Προς τη κατεύθυνση αυτή είναι σημαντικό οι αισθητήρες να αναθέτουν μη μηδενικές στο σύνολο των πιθανών καταστάσεων Θ. Στο παράδειγμα B ο δεύτερος αισθητήρας δεν επιτρέπει άλλη διάγνωση πλην των UDP-flood και NORMAL, ενώ ο πρώτος υποδεικνύει με ισχυρή πεποίθηση την ύπαρξη SYN-flood.

Κεφάλαιο 4

Προτεινόμενη αρχιτεκτονική ανίχνευσης ανωμαλιών

4.1 Αρχιτεκτονική

Στο κεφάλαιο αυτό παρουσιάζεται η προτεινόμενη αρχιτεκτονική ανίχνευσης ανωμαλιών στη δικτυακή κίνηση (Network Anomaly Detection System - NADS). Η **προσέγγιση** που έχουμε ακολουθήσει μπορεί να κωδικοποιηθεί στα παρακάτω σημεία:

- Το σύστημα ανίχνευσης ανωμαλιών στη δικτυακή κίνηση λειτουργεί στους ακραίους δρομολογητές του δικτύου ενός παρόχου (Provider Edge routers), όπως αναλύσαμε στην ενότητα 2.3.

- Το σύστημα συλλέγει δεδομένα για την κατάσταση του δικτύου από πολλούς αισθητήρες με στόχο τη βελτίωση της απόδοσης του. Τα δεδομένα ανίχνευσης από πολλούς αισθητήρες συνθέτονται για να εκτιμηθεί η πραγματική κατάσταση του υπό παρακολούθηση δικτυακού στοιχείου. Μέσα από τη διαδικασία αυτή ανιχνεύονται ανωμαλίες στη δικτυακή κίνηση.

- Δίνεται ιδιαίτερη έμφαση στην **πρακτική εφαρμογή** των προτεινόμενων

μεθόδων και την σύνδεση τους με τις υπάρχουσες πρακτικές διαχείρισης δικτύων.

Ακολουθώντας τις παραπάνω βασικές κατευθύνσεις, έγινε προσπάθεια ώστε η σχεδίαση της αρχιτεκτονικής του προτεινόμενου συστήματος να χαρακτηρίζεται (σχεδιαστικές απαιτήσεις) από:

- **επεκτασιμότητα** για να ανταπεξέλθει στην αναμενόμενη εξέλιξη των χαρακτηριστικών της δικτυακής κίνησης και των ανωμαλιών της. Μια ανοικτή αρχιτεκτονική εξασφαλίζει ότι θα είναι δυνατόν να προστεθούν μελλοντικά νέοι αισθητήρες σύμφωνα με νεότερα ερευνητικά αποτελέσματα.

- **αποτελεσματικότητα** στην ανίχνευση ανωμαλιών για να είναι πρακτικά χρήσιμο στους διαχειριστές δικτύων.

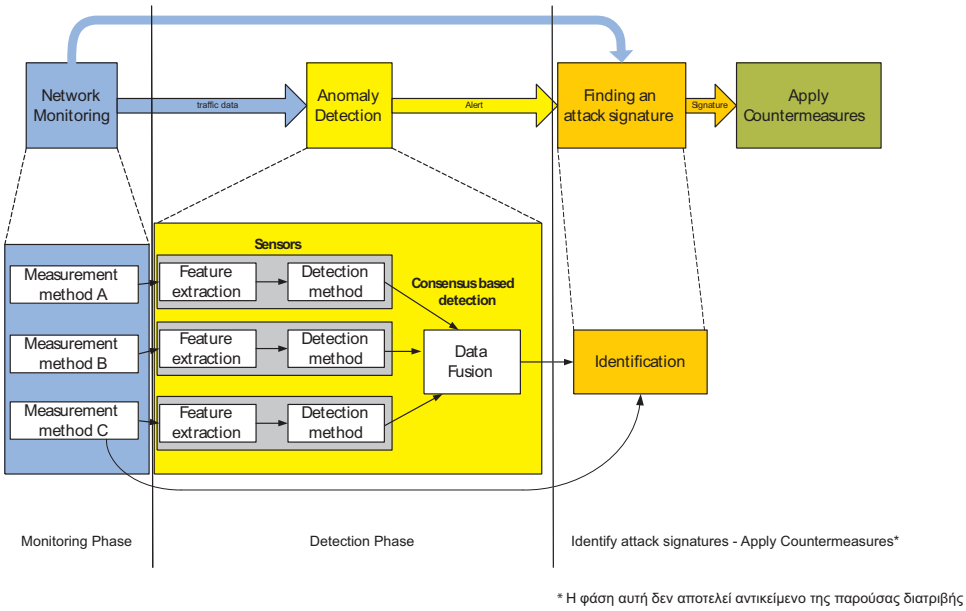
- **αρθρωτή σχεδίαση** (**modularity**)ώστε να επιτευχθεί καταμερισμός των απαιτούμενων σταδίων επεξεργασίας.

- **γρήγορη απόκριση** κατά την ανίχνευση δικτυακών ανωμαλιών (σε πραγματικό χρόνο). Για την ανίχνευση χρονικά σύντομων επιθέσεων όπως pulsing attacks [108], η απόκριση του συστήματος πρέπει να είναι άμεση και να στηρίζεται σε δεδομένα με μεγάλη συχνότητα δειγματοληψίας.

Πριν προχωρήσουμε στην παρουσίαση της προτεινόμενης αρχιτεκτονικής θα συνοψίσουμε τους βασικότερους ορισμούς από τα προηγούμενα κεφάλαια:

- Στόχος του συστήματος που αναπτύσσουμε είναι η ανίχνευση σε κάθε χρονική στιγμή της πραγματικής κατάστασης του υπό παρακολούθηση δικτυακού στοιχείου.

- Οι αισθητήρες (sensors) συλλέγουν και αναλύουν δεδομένα για την κατάσταση του παρακολουθούμενου δικτύου (ενότητα 3.5.3). Κάθε αισθητήρας πέρα από συλλέκτης δεδομένων, αποτελεί και μία στοιχειώδη μο-



Σχήμα 4.1: Βασικά στάδια επεξεργασίας της προτεινόμενης αρχιτεκτονικής.

νάδα ανίχνευσης, δηλαδή επεξεργάζεται με έναν αλγόριθμο ανίχνευσης συγκεκριμένα μετρικά (metrics) του υπό παρακολούθηση δικτύου[1].

- Η σύνθεση δεδομένων είναι μια διαδικασία που χρησιμοποιεί δεδομένα από πολλούς αισθητήρες και τα συνθέτει σε δεδομένα με μεγαλύτερη αφαιρετικότητα και περισσότερη χρήσιμη πληροφορία (ενότητα 3.5).

Σύμφωνα με την προτεινόμενη αρχιτεκτονική προκειμένου ορίζονται τα παρακάτω στάδια συλλογής και επεξεργασίας δεδομένων (σχήμα 4.1):

1. **Παρακολούθηση δικτύου (Network monitoring)**. Αρχικά γίνεται επιλογή των κατάλληλων τεχνολογιών παρακολούθησης δικτύου (Packet Capturing, SNMP, Netflow) και εγκατάσταση του απαραίτητου εξοπλισμού ανάλογα με την υπάρχουσα δικτυακή υποδομή. Στη συνέχεια

[1] στην παρούσα διατριβή θεωρούμε ότι ένας αισθητήρας βασίζεται σε ένα μετρικό αλλά η σχέση αυτή δεν είναι περιοριστική.

ρυθμίζεται ο δικτυακός εξοπλισμός ώστε να τροφοδοτείται το σύστημα ανίχνευσης με δεδομένα.

2. **Εξαγωγή χαρακτηριστικών (μετρικών) και εκτέλεση αλγόριθμων ανίχνευσης (Anomaly Detection).**

   - Οι αισθητήρες του συστήματος συλλέγουν δεδομένα και μετρούν συγκεκριμένα χαρακτηριστικά που υπό παρακολούθηση δικτύου (feature extraction). Τα μετρούμενα χαρακτηριστικά ονομάζονται **μετρικά** (metric). Στη συνέχεια επεξεργάζονται τα καταγεγραμμένα μετρικά με κάποιο αλγόριθμο ανίχνευσης (detection method) σε ένα μετρούμενο χαρακτηριστικό (feature) του υπό παρακολούθηση δικτύου. Ένας αισθητήρας αποτελείται από το υποσύστημα εξαγωγής χαρακτηριστικών (feature extraction) και το υποσύστημα εκτέλεσης αλγόριθμου ανίχνευσης (detection method). Κάθε αισθητήρας λειτουργεί αυτόνομα χωρίς να προϋποθέτει την ύπαρξη ενός μοντέλου λειτουργίας του δικτύου. Το βήμα αυτό αποτελεί το πρώτο στάδιο της ανίχνευσης.

   - Τα δεδομένα των αισθητήρων συγκεντρώνονται και συνθέτονται σε έναν "κόμβο σύνθεσης δεδομένων" (data fusion node). Με τον τρόπο αυτό εξάγεται το συνολικό συμπέρασμα της διαδικασίας ανίχνευσης και γίνεται εκτίμηση της πραγματικής κατάστασης του δικτύου. Το βήμα αυτό αποτελεί το δεύτερο στάδιο της ανίχνευσης.

3. **Ταυτοποίηση ανιχνευθεισών ανωμαλιών (Finding an attack signature) και εφαρμογή μεθόδων καταστολής (Apply Countermeasures)**. Η ταυτοποίηση μιας επίθεσης είναι μια διαδικασία που ενεργοποιείται ασύγχρονα μόνο μετά την ανίχνευση μιας ανωμαλίας στη δικτυακή κίνηση και έχει ως στόχο τον ακριβή προσδιορισμό των χαρακτηριστικών της (attack signature). Η διαδικασία αυτή μπορεί να είναι χρονοβόρα και για το λόγο αυτό δεν υπόκειται σε αυστηρούς χρονικούς

περιορισμούς. Μετά την ταυτοποίηση είναι δυνατή η εφαρμογή διάφορων μεθόδων καταστολής όπως η εφαρμογή ενός φίλτρου δικτυακής κίνησης (access list) από ένα firewall (ενότητα 2.5.2).

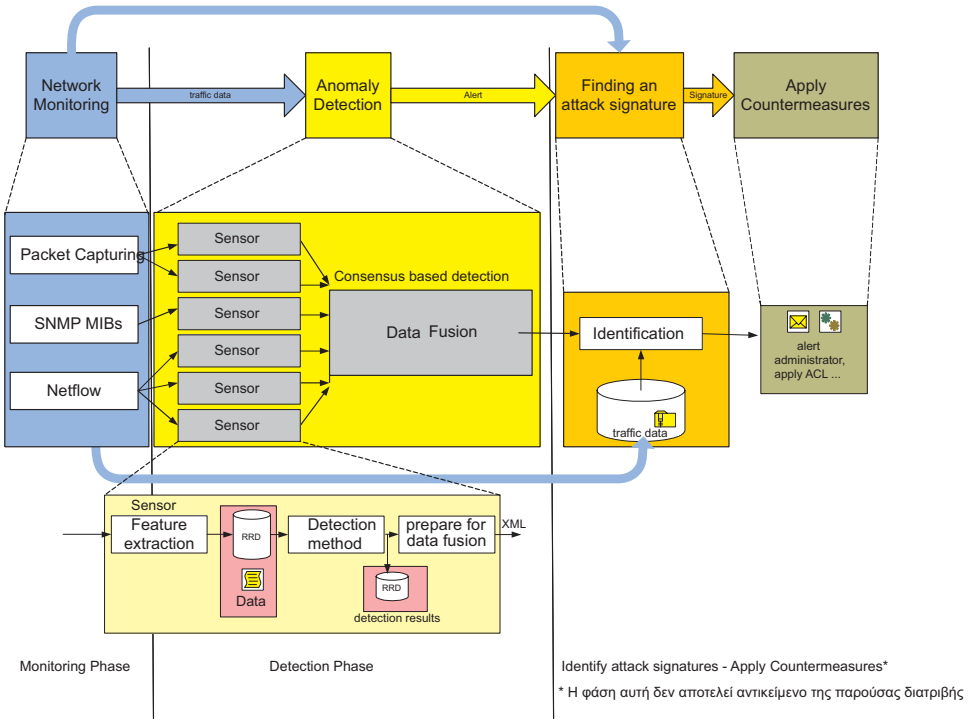
Η προτεινόμενη αρχιτεκτονική (NADS) παρουσιάζει αρκετές ομοιότητες με την ιεραρχική οργάνωση συστημάτων διαχείρισης δικτύων με χρήση του πρωτοκόλλου Simple Network Management Protocol (SNMP) και με το σύστημα ανίχνευσης επιθέσεων SiDS [109]. Στο κατώτερο επίπεδο της ιεραρχίας βρίσκονται οντότητες που συλλέγουν και επεξεργάζονται πληροφορίες με περιορισμένο εύρος π.χ. τοπικό δηλαδή σε ένα συγκεκριμένο link. Οι οντότητες αυτές στο πλαίσιο του SNMP αποκαλούνται agents, στο SiDS controllers και στην προτεινόμενη αρχιτεκτονική NADS sensors. Στο επόμενο επίπεδο υπάρχουν οντότητες που συλλέγουν τις επεξεργασμένες πληροφορίες και τις συνθέτουν, συσχετίζουν κτλ. Οι οντότητες αυτές στο πλαίσιο του SNMP αποκαλούνται managers, στο SiDS analyzers και στην προτεινόμενη αρχιτεκτονική "κόμβοι σύνθεσης δεδομένων" (data fusion nodes).

4.1.1 Επικοινωνία μεταξύ των υποσυστημάτων της αρχιτεκτονικής και η χρήση Round Robin Databases - RRD

Ένα σημαντικό στοιχείο της προτεινόμενης αρχιτεκτονικής είναι οι διεπαφές (interfaces) μέσα από τις οποίες επικοινωνούν τα υποσυστήματα: εξαγωγής χαρακτηριστικών (feature extraction), εκτέλεσης αλγόριθμων ανίχνευσης (detection method) και σύνθεσης δεδομένων (data fusion).

Στο εσωτερικό κάθε αισθητήρα τα μετρικά οδηγούνται από το υποσύστημα εξαγωγής χαρακτηριστικών (feature extraction) στο υποσύστημα εκτέλεσης αλγόριθμου ανίχνευσης (detection method) με χρήση ενός ανοικτού de-facto standard: τις Round Robin Databases[2] (RRD) (σχήμα 4.2). Η χρήση των RRD εξασφαλίζει ένα σταθερό Application Programming Interface (API) για την αποθήκευση, ανάκτηση και επεξεργασία μετρικών αλλά και ένα ενιαίο

[2]όπως υλοποιούνται από το RRDtool [110].



Σχήμα 4.2: Προτεινόμενη αρχιτεκτονική ανίχνευσης ανωμαλιών

σχήμα απεικόνισης πληροφοριών (data abstraction layer, data representation). Μια Round Robin Database (RRD) είναι μια κυκλική δομή δεδομένων που αποθηκεύει αριθμητικά δεδομένα σε διαδοχικές θέσεις. Η δομή αυτή είναι ειδικά σχεδιασμένη για την αποθήκευση χρονοσειρών (timeseries) με σταθερή περίοδο δειγματοληψίας. Διαδοχικές τιμές αποθηκεύονται σε διαδοχικές θέσεις και λόγω της κυκλικής φύσης της δομής μετά την εξάντληση των ελεύθερων θέσεων οι παλαιότερες τιμές αντικαθίστανται από νεότερες. Η επιλογή των RRD αντί μιας σχεσιακής βάσης δεδομένων (Relational Database Management System -RDBMS) έγινε κυρίως για τους παρακάτω λόγους:

- Η αποθήκευση των δεδομένων γίνεται σε ένα απλό αρχείο.
- Η ανάκτηση και αποθήκευση δεδομένων σε μια RRD είναι γρήγορη.
- Ο όγκος των αποθηκευόμενων δεδομένων παραμένει σταθερός, δηλαδή

δεν αυξάνεται με την πάροδο του χρόνου κάνοντας την διαχείριση μιας RRD ιδιαίτερα απλή.

- Υποστηρίζεται η σύνοψη δεδομένων (consolidation) ώστε να αποθηκεύονται δεδομένα για το παρελθόν με μικρότερη ακρίβεια σε περιορισμένο χώρο. Εύκολη ρύθμιση της επιθυμητής χρονικής ακρίβειας των αποθηκευόμενων δεδομένων.

- Δίνεται η δυνατότητα εύκολου ορισμού του όγκου των αποθηκευόμενων δεδομένων και της χρονικής ακρίβειας τους (granularity).

- Οι RRD υποστηρίζονται από πολλά λειτουργικά συστήματα και η μεταφορά δεδομένων μεταξύ διαφορετικών συστημάτων είναι απλή.

- Οι RRD υποστηρίζουν εύκολη απεικόνιση των αποθηκευόμενων δεδομένων μέσα από γραφικές παραστάσεις.

- Υποστηρίζεται αυτόματη επίλυση χρονικών καθυστερήσεων, αποκλίσεων στην δειγματοληψία.

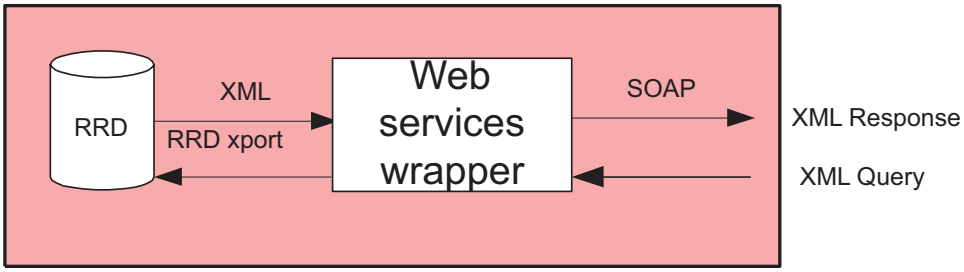
- Δυνατότητα συγχρονισμού δεδομένων και ταυτόχρονη ανάκτηση δεδομένων από πολλά RRD (rrdxport).

- Δυνατότητα περιγραφής αριθμητικών δεδομένων με Extensible Markup Language (XML).

- Δυνατότητα μεταφοράς δεδομένων σύμφωνα με το μοντέλο client-server μέσω TCP/IP.

- Υποστήριξη αλγόριθμου πρόβλεψης χρονοσειρών (Holt-Winters forecasting).

Οι αισθητήρες επικοινωνούν με τον κόμβο σύνθεσης δεδομένων μέσω Extensible Markup Language (XML). Όπως αναφέραμε το εργαλείο RRDtool



Σχήμα 4.3: Συμβατότητα με Service oriented Architectures (SoA).

είναι συμβατό με XML. Τα αποτελέσματα της ανίχνευσης, καθώς αποθηκεύονται σε RRDs, μπορούν να εξάγονται απευθείας σε XML (σχήμα 4.3) και να μεταφέρονται μέσω πρωτοκόλλων όπως το Simple Object Access Protocol (SOAP). Τα στοιχεία αυτό είναι ιδιαίτερα σημαντικό δεδομένης της δυναμικής των Service oriented Architectures (SoA) και των Web Services. Η προτεινόμενη αρχιτεκτονική είναι συμβατή με τις νέες αυτές τεχνολογικές τάσεις. Ας σημειωθεί πως η ερευνητική ομάδα εργασίας για τις μετρήσεις στο πανευρωπαϊκό ακαδημαϊκό δίκτυο Geant-2 [111] προτείνει μια αρχιτεκτονική μετρήσεων χαρακτηριστικών του δικτύου με SoA λογική με χρήση RRD.

Οι Round Robin Databases στην μορφή που υλοποιούνται σήμερα από το εργαλείο RRDtool [110] παρέχουν τη δυνατότητα αποθήκευσης μονάχα αριθμητικών δεδομένων - χρονοσειρών (timeseries). Παρόλο που σήμερα δεν υπάρχει η λειτουργικότητα για την αποθήκευση περισσότερων τύπων δεδομένων όπως σε μια σχεσιακή βάση δεδομένων, η δυνατότητα αυτή μπορεί να προστεθεί με χρήση νέων τύπων δεδομένων (data sources) και νέων consolidation functions. Η προτεινόμενη αρχιτεκτονική δίνοντας έμφαση στην πρακτική εφαρμογή χρησιμοποιεί τις Round Robin Databases.

4.2 Παρακολούθηση δικτύου

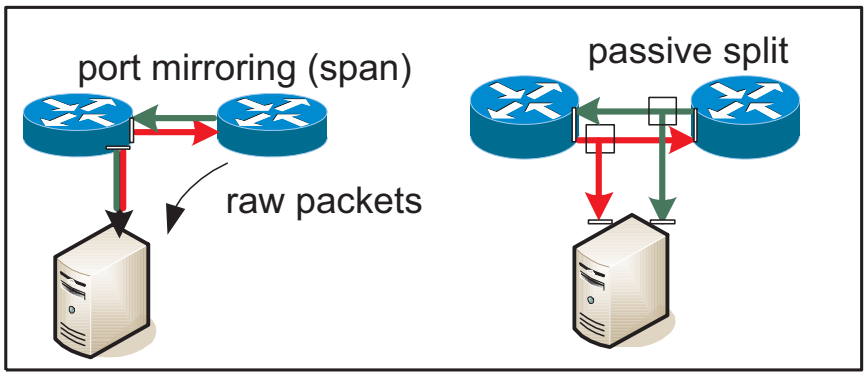
Για τη παρακολούθηση δικτύων ακολουθούνται δύο διαφορετικές προσεγγίσεις. Η παθητική παρακολούθηση δικτύου (passive monitoring) μετρά με-

γέθη και χαρακτηριστικά της υπάρχουσας κίνησης ενός δικτύου ενώ η ενεργή παρακολούθηση (active monitoring) δημιουργεί κίνηση στο δίκτυο εν είδη δοκιμής και μετρά μεγέθη αποκλειστικά για την κίνηση αυτή. Παραδείγματα ενεργής παρακολούθησης είναι η μέτρηση Round-Trip delay Time (RTT) και η εκτίμηση απώλειας πακέτων (packet loss) [112] με ping probes. Η προτεινόμενη αρχιτεκτονική εστιάζει στην **παθητική παρακολούθηση δικτύου** που είναι και η πλέον διαδεδομένη. Οι τεχνικές που χρησιμοποιούνται σήμερα από τους διαχειριστές δικτύων στα πλαίσια της παθητικής παρακολούθησης δικτύου είναι τρεις: η συλλογή πακέτων (Packet Capturing), η χρήση της τεχνολογίας Netflow και του πρωτοκόλλου Simple Network Management Protocol (SNMP). Κάθε μία έχει τις αδυναμίες και τις ιδιαιτερότητες της. Για να μπορέσουμε να διατυπώσουμε μία ρεαλιστική πρόταση για ένα σύστημα ανίχνευσης δικτυακών ανωμαλιών (NADS), που θα μπορεί να χρησιμοποιηθεί εύκολα από τους διαχειριστές δικτύου, πρέπει πρώτα να αναλύσουμε τις πρακτικές δυσκολίες και τους περιορισμούς των διαφορετικών τεχνικών παθητικής παρακολούθησης δικτύου. Η χρήση τους άλλωστε είναι απαραίτητη για να τροφοδοτήσουν ένα σύστημα ανίχνευσης με δεδομένα.

4.2.1 Παρακολούθηση δικτυακής κίνησης με συλλογή πακέτων (Packet Capturing)

Σύμφωνα με την τεχνική παρακολούθησης δικτύου με συλλογή πακέτων (Packet Capturing) ένα σύστημα παρακολούθησης συλλέγει και επεξεργάζεται κάθε πακέτο που μεταδίδεται από μια δικτυακή ζεύξη (link). Το σύστημα παρακολούθησης μπορεί λοιπόν να μετρήσει κάθε μέγεθος που χαρακτηρίζει την κίνηση που μεταφέρεται μέσα από την δικτυακή ζεύξη. Η τεχνική Packet Capturing μπορεί να μας δώσει τις ακριβέστερες και πληρέστερες πληροφορίες για το παρακολουθούμενο link. Παρουσιάζει όμως δυο σημαντικά μειονεκτήματα.

Το πρώτο μειονέκτημα σχετίζεται με την προσπάθεια που απαιτείται να



Σχήμα 4.4: Παθητική παρακολούθηση δικτύου με συλλογή πακέτων (Packet Capturing).

καταβάλει ένας διαχειριστής δικτύου για την εγκατάσταση ενός συστήματος παρακολούθησης Packet Capturing σε ένα εκτεταμένο δίκτυο. Ο διαχειριστής για την εγκατάσταση ενός συστήματος συλλογής πακέτων έχει τις εξής επιλογές (σχήμα 4.4):

- χρήση παθητικών συσκευών (optical splitters/couplers) που διαχωρίζουν το σήμα που διέρχεται από μία οπτική ίνα σε οπτικό επίπεδο και το μεταδίδουν προς δύο κατευθύνσεις. Χρησιμοποιούνται δηλαδή μονάχα παθητικά στοιχεία και δεν γίνεται καμία επεξεργασία σε ηλεκτρικό επίπεδο. Παρουσιάζει όμως το μεγάλο μειονέκτημα ότι με κάθε σύνδεση splitter-συσκευής συλλογής πακέτων μπορεί να παρακολουθηθεί μονάχα η μια φορά της κίνησης που μεταδίδεται από την παρακολουθούμενη δικτυακή ζεύξη (εισερχόμενη ή εξερχόμενη κίνηση). Καθώς δεν γίνεται επεξεργασία σε ηλεκτρικό επίπεδο δεν μπορεί να συνδυαστεί η πληροφορία των δύο κατευθύνσεων σε μία ίνα που θα συνδεθεί στην υποδοχή λήψης (receive) του συστήματος παρακολούθησης.

- χρήση ενεργών συσκευών όπως ενός hub που αντιγράφει κάθε πακέτο που μεταδίδεται σε όλες τις πόρτες του. Με τον τρόπο αυτό μπορούμε να παρακολουθήσουμε την κίνηση που διέρχεται από πολλές δικτυακές ζεύξεις. Για λόγους ασφάλειας αλλά και απόδοσης στην ταχύτητα με-

τάδοσης προτιμάται η χρήση ενός μεταγωγέα (switch). Ένα switch που υποστηρίζει port mirroring μπορεί να μεταδίδει στη πόρτα όπου είναι συνδεδεμένο το σύστημα παρακολούθησης κάθε πακέτο που μεταδίδεται προς ή από μια άλλη πόρτα. Στην περίπτωση αυτή το σύστημα παρακολούθησης λαμβάνει την εισερχόμενη και την εξερχόμενη κίνηση μιας πόρτας.

Δεύτερο μειονέκτημα της παθητικής παρακολούθησης δικτύου με συλλογή πακέτων είναι ότι απαιτεί ένα σύστημα παρακολούθησης με αρκετή υπολογιστική ισχύ. Το σύστημα πρέπει να είναι αρκετά απλό ώστε να μπορεί να ανταπεξέλθει σε υψηλές ταχύτητες μετάδοσης πακέτων στο δίκτυο. Οι μεγάλες ταχύτητες των δικτυακών συνδέσεων π.χ. > 1 Gbps περιορίζουν πρακτικά τα χαρακτηριστικά της δικτυακής κίνησης που μπορούν να μετρηθούν σε πραγματικό χρόνο. Θεωρητικά πάντως υπάρχει δυνατότητα να γίνονται μετρήσεις με μεγάλη χρονική ακρίβεια αφού το σύστημα συλλογής λαμβάνει τα πακέτα σχεδόν παράλληλα με την μετάδοση τους από την παρακολουθούμενη δικτυακή ζεύξη. Για τον περιορισμό που προβλήματος της επεξεργασίας πακέτων σε υψηλές ταχύτητες έχει προταθεί η χρήση εξειδικευμένου υλικού όπως εκείνου που χρησιμοποιήθηκε στο ευρωπαϊκό ερευνητικό πρόγραμμα "Scaleable Monitoring Platform for the Internet" (SCAMPI) [113] και δικτυακών επεξεργαστών (network processors) [114]. Η νέα αυτή τάση έρχεται να αντικαταστήσει τη χρήση ενός απλού υπολογιστή με εξειδικευμένο λογισμικό επεξεργασίας πακέτων.

Συνοψίζοντας, η τεχνική παρακολούθησης δικτύων με συλλογή πακέτων παρουσιάζει προβλήματα κλιμάκωσης καθώς απαιτεί αρκετή προσπάθεια από τον εκάστοτε διαχειριστή και ένα ισχυρό υπολογιστικό σύστημα για κάθε δικτυακή ζεύξη που παρακολουθείται.

4.2.2 Παρακολούθηση δικτυακής κίνησης μέσω Netflow

Η τεχνική παρακολούθησης της δικτυακής κίνησης με πληροφορίες για δικτυακές ροές που συλλέγει ένας δρομολογητής ξεκίνησε το 1996 όταν η Cisco ανακοίνωσε τη τεχνολογία Netflow [115]. Αρχικά χρησιμοποιήθηκε ως ένας μηχανισμός μεταγωγής πακέτων (switching mechanism) αλλά στην συνέχεια εγκαταλείφθηκε και σήμερα χρησιμοποιείται αποκλειστικά για την καταγραφή και παρακολούθηση της δικτυακής κίνησης. Τυπικές χρήσεις της τεχνολογίας Netflow είναι για traffic engineering, capacity planning, χρέωση (billing) και για την ανίχνευση ανωμαλιών. Η επιτυχία της τεχνολογίας Netflow αποτυπώνεται από την υιοθέτηση της από άλλους κατασκευαστές όπως το J-flow της Juniper [116] αλλά την χρήση της ως βάση για την δημιουργία των προδιαγραφών IP Flow Information Export (IPFIX) από το Internet Engineering Task Force (IETF) [117, 118]. Η αρχή λειτουργίας των συστημάτων τύπου Netflow βασίζεται στην ομαδοποίηση των μεταδιδόμενων πακέτων σε **ροές (flows) που σχηματίζουν κλάσεις ισοδυναμίας (equivalence classes) βάσει των ακόλουθων πέντε χαρακτηριστικών της επικεφαλίδας ενός πακέτου (5-tuple of packet header): Πρωτόκολλο, IP προέλευσης, IP προορισμού, Port προέλευσης, Port προορισμού.**

Η βασική λειτουργία ενός συστήματος καταγραφής Netflow [33] μπορεί να περιγραφεί ως εξής:

1. Από την επικεφαλίδα κάθε πακέτου συλλέγεται η πληροφορία που το αντιστοιχίζει μοναδικά με ένα flow.

2. Γίνεται αναζήτηση της εγγραφής του flow στον πίνακα όπου αυτά καταγράφονται. Ο πίνακας ονομάζεται flow-cache και η εγγραφή κάθε flow ονομάζεται flow record. Αν η αναζήτηση αποτύχει δημιουργείται μια νέα εγγραφή.

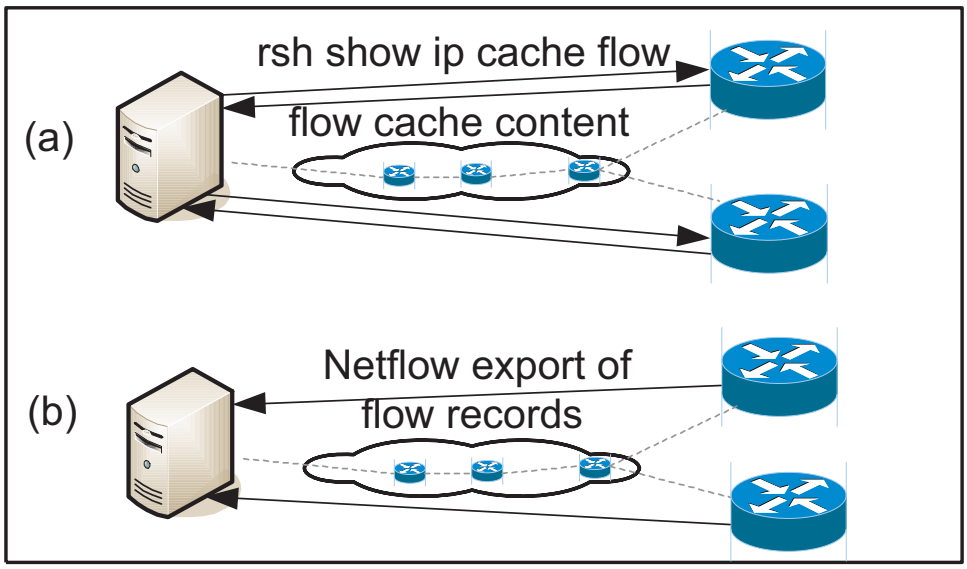
3. Η εγγραφή ενημερώνεται για κάθε πακέτο. Για παράδειγμα αυξάνονται οι μετρητές πλήθους πακέτων και μεταφερόμενων bytes.

4. Σε συγκεκριμένες χρονικές στιγμές flow-records διαγράφονται και αποστέλλονται σε εξωτερικά συστήματα που τα συλλέγουν (collectors) για περεταίρω επεξεργασία (flow export). Η διαδικασία αυτή ενεργοποιείται:

   - όταν ένα flow τερματίζεται (στην περίπτωση TCP flow με την μετάδοση πακέτου με flag RST ή FIN),
   - όταν η flow-cache γεμίζει δημιουργώντας ένα flow learning failure,
   - όταν ένα flow είναι ανενεργό για κάποιο χρονικό διάστημα που ορίζεται από τον inactivity timer,
   - όταν η διάρκεια ζωής ενός flow περάσει την τιμή του active timer.

Η πληροφορία που δίνεται για κάθε ροή (flow) εξαρτάται από την έκδοση του Netflow που χρησιμοποιείται π.χ. 5, 7, 8, 9. Έτσι αν και σε ορισμένες περιπτώσεις κάποια στοιχεία μπορεί να μην είναι διαθέσιμα ενδεικτικά αναφέρουμε ότι παρέχονται οι εξής συγκεντρωτικές πληροφορίες ανά ροή ή ακόμη για σύνολα ροών με κοινά χαρακτηριστικά (aggregation schemes): πλήθος πακέτων (packets in flow), πλήθος bytes (bytes in flow), χρονική στιγμή λήψης πρώτου πακέτου (flow start), χρονική στιγμή λήψης τελευταίου πακέτου (flow end), διεύθυνση προέλευσης/προορισμού (destination/source address), πόρτα εισόδου και εξόδου στον δρομολογητή (source/destination interface), επόμενος δρομολογητής (next hop), Αυτόνομο Σύστημα προέλευσης/προορισμού (source/destination AS) και το λογικό OR μεταξύ των TCP flags των πακέτων (inclusive-OR of TCP flags during flow).

Γενικά η παρακολούθηση της κίνησης με χρήση Netflow έχει το πλεονέκτημα ότι είναι εύκολη στην χρήση από τους διαχειριστές του δικτύου (γίνεται με απλή ρύθμιση των δρομολογητών). Επιπλέον δεν απαιτεί ιδιαίτερους πόρους καθώς υλοποιείται μέσα στους δρομολογητές και αρκεί ένας απλός υπολογιστής για την συλλογή στοιχείων από πολλούς δρομολογητές (σχήμα 4.5). Με τον τρόπο αυτό συλλέγονται ταυτόχρονα στοιχεία για πολλές δικτυακές ζεύξεις. Το μειονέκτημα είναι ότι η πληροφορία που παρέχεται, ως πιο συνοπτική



Σχήμα 4.5: Παθητική παρακολούθηση δικτύου μέσω Netflow.

από τα πλήρη πακέτα, θέτει κάποιους περιορισμούς στο πλήθος των μετρικών που μπορούν να μετρηθούν. Χαρακτηριστικό παράδειγμα είναι ότι δεν είναι εύκολο να μετρηθούν το πλήθος των TCP πακέτων που περιέχουν το SYN flag γιατί η πληροφορία που παρέχεται από ένα flow-record είναι μονάχα αν υπήρξε έστω και ένα πακέτο του flow με το flag αυτό και όχι το πλήθος τους. Σε γενικές γραμμές πάντως, η τεχνολογία Netflow περιέχει πληροφορία αρκετά αναλυτική και συνοψισμένη. Επίσης ιδιαίτερη προσοχή πρέπει να επιδεικνύεται όταν χρησιμοποιείται δειγματοληψία για την δημιουργία των flow-records (sampled Netflow). Η δειγματοληψία είναι υποχρεωτική για ορισμένους δρομολογητές υψηλών ταχυτήτων. Το αποτέλεσμα της χρήσης δειγματοληψίας (sampled Netflow) π.χ. 1/100 ή 1/1000 πακέτα είναι οι μετρήσεις να μην είναι πλέον ακριβείς αλλά να πρέπει να γίνονται κάποιες εκτιμήσεις βάσει του δειγματοληπτούμενου μεγέθους. Η ακρίβεια των στατιστικών εκτιμήσεων σε περίπτωση δειγματοληψίας αποτελεί ανοικτό ερευνητικό θέμα [119,120]. Στη περίπτωση μας η χρήση μετρικών που προέκυψαν από sampled Netflow για την ανίχνευση ανωμαλιών απέδωσε ικανοποιητικά αποτελέσματα αναδεικνύοντας την ύπαρξη ανωμαλιών ακόμη και όταν δεν αποδίδει καλές εκτιμήσεις για τις

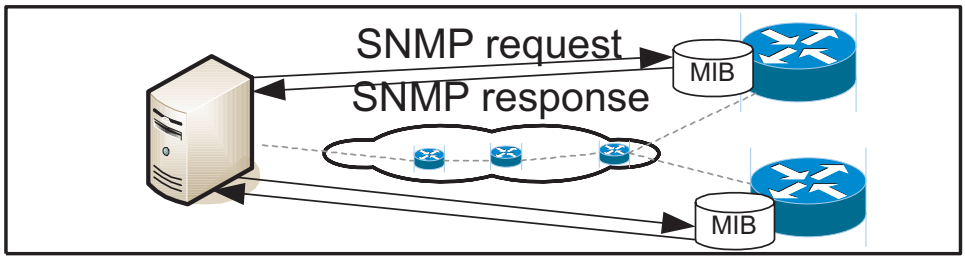
πραγματικές τιμές των μετρικών [121, 122].

Ένα ακόμη κρίσιμο χαρακτηριστικό της παρακολούθησης μέσω του Netflow είναι ότι τα δεδομένα έρχονται με κάποια χρονική καθυστέρηση ανάλογα με τις ρυθμίσεις του δρομολογητή για την διαγραφή των flows (flow export). Συχνά αναφέρεται ως παρακολούθηση σε σχεδόν πραγματικό χρόνο (near real time) καθώς η λήψη των στοιχείων ενός flow από ένα εξωτερικό σύστημα συλλογής γίνεται όταν το flow έχει πλέον διαγραφεί [33].

Μια πρακτική λεπτομέρεια που επηρεάζει σημαντικά τα δεδομένα που συλλέγονται αφορά τον τρόπο με τον οποίο συλλέγεται η πληροφορία της flow-cache. Συγκεκριμένα στα πλαίσια της παρούσας διατριβής παρατηρήθηκε ότι η περιοδική συλλογή των περιεχομένων της flow-cache (π.χ. με χρήση μιας εντολής rsh show ip cache flow - μέθοδος (a) στο σχήμα 4.5) μπορεί να μας δώσει αρκετά διαφορετική εικόνα για την κίνηση από την συλλογή των flow-records που διαγράφονται από τον δρομολογητή (μέθοδος (b) στο σχήμα 4.5). Η διαφορά έγκειται στο πεπερασμένο μέγεθος της flow-cache, που μπορεί να περιέχει εγγραφές που δημιουργούνται και διαγράφονται λόγω έλλειψης χώρου σε χρόνο μικρότερο από την περίοδο δειγματοληψίας. Έτσι είναι ασφαλέστερη η δεύτερη προσέγγιση συλλογής όλων των exported-flows.

4.2.3 Παρακολούθηση δικτυακών συσκευών μέσω SNMP

Ο όρος παρακολούθηση δικτύου μέσω Simple Network Management Protocol (SNMP) περιγράφει το τρόπο (με χρήση του πρωτοκόλλου SNMP) με τον οποίο προσπελαύνονται δεδομένα που μετρούνται από δικτυακές συσκευές, κυρίως δρομολογητές, και αποθηκεύονται σε Management Information Bases (MIBs). Τα δεδομένα αυτά γίνονται διαθέσιμα μέσω των κλήσεων snmp-get, snmp-walk του πρωτοκόλλου SNMP (σχήμα 4.6). Ένα απλό σύστημα χωρίς ιδιαίτερες απαιτήσεις σε λογισμικό ή υλικό μπορεί να συλλέγει στοιχεία ταυτόχρονα από πολλές δικτυακές συσκευές. Για το λόγο αυτό η παρακολούθηση δικτύου με SNMP είναι ιδιαίτερα δημοφιλής ανάμεσα στους διαχειριστές



Σχήμα 4.6: Παθητική παρακολούθηση δικτύου μέσω SNMP.

δικτύου.

Τα δεδομένα που παρέχονται από μια δικτυακή συσκευή εξαρτώνται από τον τύπο της και από το λογισμικό της π.χ. διαφέρουν ανάλογα με το version του Cisco IOS. Τα πιο χρήσιμα δεδομένα που μπορούμε να προσπελάσουμε με χρήση της τεχνικής αυτής αφορούν την λειτουργία των δικτυακών συσκευών όπως η χρησιμοποίηση της cpu , το μέγεθος διαφόρων ουρών (queue sizes), το πλήθος των πακέτων που απορρίπτονται (IP discards) και τον αριθμό των flow learning failures στο σύστημα παρακολούθησης Netflow. Άλλωστε τα μεγέθη αυτά δεν μπορούν να μετρηθούν με εξωτερικές συσκευές. Τα δεδομένα αυτά ενημερώνονται συνήθως με συχνότητα αρκετών δευτερολέπτων π.χ. 30sec. Τα δεδομένα που συνήθως αντλούνται μέσω SNMP είναι οι μετρητές πακέτων και bytes σε επίπεδο 3 (IP packets, bytes). Προς το παρόν δεν παρέχονται μετρητές στις τυποποιημένες MIBs πολλών δρομολογητών (π.χ. MIB-II [123]) για εισερχόμενα, εξερχόμενα πακέτα σε επίπεδο 4 δηλαδή πλήθος TCP, UDP, ICMP πακέτων.

Το συμπέρασμα από την ανάλυση των τεχνικών παθητικής παρακολούθησης δικτύου που προηγήθηκε είναι ότι η χρήση Netflow και SNMP προτιμάται από τους διαχειριστές δικτύου ενώ η παρακολούθηση δικτύου με συλλογή πακέτων (Packet Capturing) είναι ιδιαίτερα πολύτιμη για την επιστημονική έρευνα. Εξετάζοντας όλη την διαθέσιμη πληροφορία (π.χ. packet headers, payload) οι ερευνητές μπορούν να εντοπίσουν και να υποδείξουν χρήσιμα μετρικά για την ανίχνευση ανωμαλιών στη δικτυακή κίνηση και άλλες εφαρμογές.

Τα μετρικά αυτά μπορούν να συμπεριληφθούν μελλοντικά σε SNMP MIBs, όπως η Netflow MIB [124], ή επεκτάσεις της τεχνολογίας Netflow ώστε να χρησιμοποιούνται ευκολότερα από τους διαχειριστές δικτύου. Μια σύνοψη των χαρακτηριστικών των τεχνικών παθητικής παρακολούθησης δικτύων που αναλύσαμε φαίνεται στον πίνακα 4.1.

Πίνακας 4.1: Σύνοψη χαρακτηριστικών των τεχνικών παθητικής παρακολούθησης δικτύων

	Time relevance	Ease of use	Measurement scope
Packet Capturing	*** (real time packet capturing)	* (one packet capture device per link)	*** (full packet header + payload)
Netflow	** (near real time, flow export depends on activity-inactivity timers)	*** (one collector per multiple routers-links)	** (per flow summary, Layer 4 information)
SNMP MIBs	** (constant refresh rate)	*** (one SNMP manager per multiple routers-links)	* (available information depends on MIBs supported by a router)

4.3 Εξαγωγή χαρακτηριστικών και επιλογή μετρικών

Μετά την ανάλυση των τεχνικών παρακολούθησης δικτύου η παρούσα ενότητα παρουσιάζει τα χαρακτηριστικά της δικτυακής κίνησης (**μετρικά**) που προτείνονται για την ανίχνευση ανωμαλιών στη δικτυακή κίνηση. Ανάλογα με την υποδομή ενός δικτύου και τους μηχανισμούς παρακολούθησης που υποστηρίζει, τα μετρικά που μπορούν να υπολογιστούν ποικίλουν. Τα προτεινόμενα μετρικά ξεχωρίζουν λόγω της χρησιμότητας, της μικρής πολυπλοκότητας και της εύκολης μέτρησης τους. Όλα τα μετρικά που προτείνονται μπορούν να υπολογιστούν με χρήση των τριών τεχνικών παθητικής παρακολούθησης δικτύου που αναλύσαμε.

Ορισμένα απλά μετρικά που έχουν προταθεί στην βιβλιογραφία, συχνά δεν καταφέρνουν να ανιχνεύσουν ανωμαλίες στη δικτυακή κίνηση που εμφανίζονται σε πραγματικά δίκτυα υψηλών ταχυτήτων. Ο ρυθμός αποστολής πακέτων (packets per second - PPS) και bytes (bytes per second - BPS) αποτελούν τα απλούστερα μετρικά που παραδοσιακά καταγράφονται από τα συστήματα παρακολούθησης δικτύου (Network Management Systems - NMS). Ο ευκολότερος τρόπος μέτρησης τους είναι με χρήση των αντίστοιχων SNMP counters, οι οποίοι μετρούν πακέτα και bytes σε IP επίπεδο (Layer 3). Η μέτρηση αυτή δεν διαχωρίζει μεταξύ των πρωτοκόλλων του επιπέδου 4 π.χ. ICMP, TCP, UDP. Συνεπώς για την μέτρηση των μετρικών BPS και PPS ανά πρωτόκολλο, που θα τα συμβολίζουμε ως $IBPS$, $IPPS$ για το ICMP, $TBPS$, $TPPS$ για το TCP και $UBPS$, $UPPS$ για το UDP χρησιμοποιούνται οι μέθοδοι Packet Capturing και Netflow. Οι ανωμαλίες στη δικτυακή κίνηση δεν ανιχνεύονται από τα μετρικά αυτά όταν οι υπό παρακολούθηση δικτυακές ζεύξεις έχουν μεγάλη χρησιμοποίηση. Στις περιπτώσεις αυτές περιορίζεται η φυσιολογική κίνηση λόγω συμφόρησης (TCP congestion control) χωρίς να παρατηρείται ουσιαστική αύξηση της συνολικής διερχόμενης κίνησης.

Ένα άλλο απλό μέγεθος που χρησιμοποιήθηκε αρχικά για την ανίχνευση δικτυακών ανωμαλιών είναι ο αριθμός των δικτυακών ροών (flows) σε ένα δρομολογητή που καταγράφει η τεχνολογία Netflow (ενότητα 4.2.2). Πιο συγκεκριμένα πρόκειται για τον αριθμό των flow records (number of active flows - AF) που βρίσκονται στην flow-cache του δρομολογητή. Η μέγιστη τιμή του ισούται με το μέγεθος της flow cache. Στην περίπτωση όμως που εισέρχονται σε ένα δρομολογητή flows με τέτοιο ρυθμό ώστε να γεμίζει η flow cache τότε το μέγεθος αυτό παρουσιάζει έντονες διακυμάνσεις κοντά στην μέγιστη τιμή. Για το λόγο αυτό συχνά δεν μπορεί να αποτυπώσει δικτυακές ανωμαλίες.

Πολλά μετρικά που μπορούν να εξαχθούν από τα packet headers (με την μέθοδο Packet Capturing) ή από τα flow records (με την χρήση του Netflow)

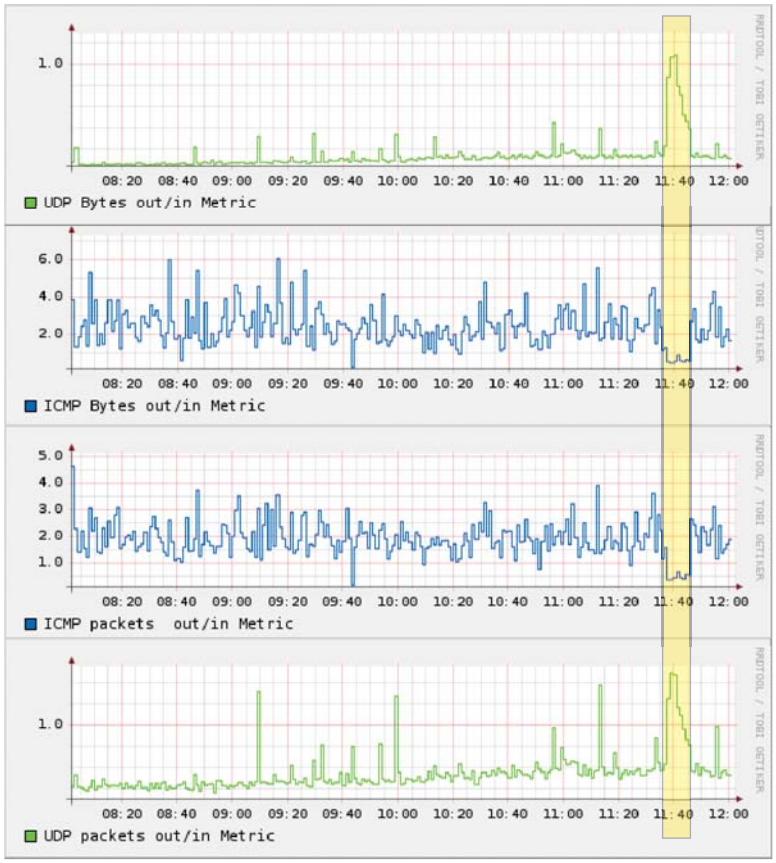
είναι δύσκολο να υπολογιστούν σε πραγματικό χρόνο. Χαρακτηριστικά αναφέρουμε ότι μετρικά που απαιτούν την τήρηση στατιστικών ανά IP διεύθυνση προορισμού ή προέλευσης είναι δύσκολο να μετρηθούν σε δίκτυα υψηλών ταχυτήτων λόγω πρακτικών περιορισμών σε μνήμη και υπολογιστική ισχύ. Μετρικά που δεν είναι απλοί μετρητές αλλά μετρητές ανά IP (π.χ. πλήθος πακέτων ανά IP [125]), ανά port ή κάποιο άλλο μέγεθος με μεγάλο εύρος τιμών (32, 16 bit) είναι δύσκολο να μετρηθούν γιατί απαιτούνται πολλές, χρονοβόρες προσπελάσεις στην μνήμη.

Μέσα από την παρούσα έρευνα εντοπίστηκαν άλλα μετρικά πέρα από τα $IBPS, IPPS, TBPS, TPPS, UBPS, UPPS$ και AF που μπορούν να είναι ιδιαίτερα αποτελεσματικά και ταυτόχρονα εύκολα μετρήσιμα και αποτελεσματικά. Προκειμένου να εντοπιστούν μετρικά που βοηθούν την ανίχνευση ανωμαλιών και να εξεταστούν οι τεχνικές μέτρησης τους πραγματοποιήθηκε ένα μεγάλο πλήθος πειραμάτων στο δίκτυο του Εθνικού Μετσόβιου Πολυτεχνείου (Ε.Μ.Π.) και του Εθνικού Δικτύου Έρευνας και Τεχνολογίας (GRNET).

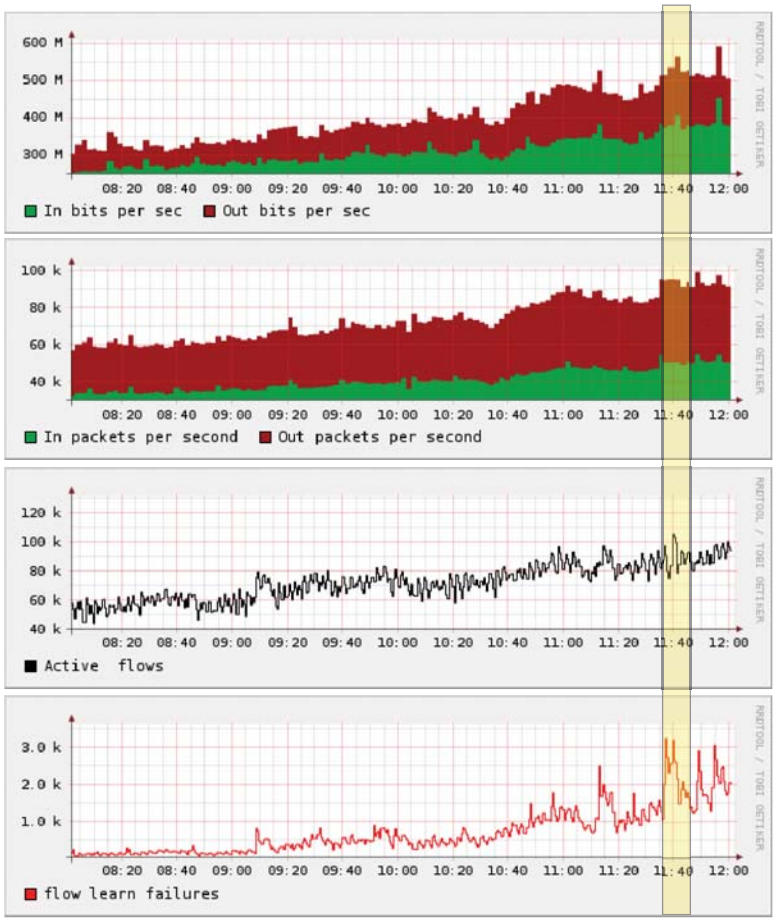
Τα πειράματα συμπεριελάμβαναν επιθέσεις καταιγισμού πακέτων διαφόρων τύπων προς και από το δίκτυο του Ε.Μ.Π. Οι επιθέσεις αυτές ήταν ελεγχόμενης έντασης προκειμένου να μη προκληθούν προβλήματα στα δίκτυα παραγωγής του Ε.Μ.Π. και GRNET. Μέσα από τα πειράματα αυτά δοκιμάστηκαν διάφορες τεχνικές spoofing και port randomization. Ο υπολογισμός των μετρικών έγινε για την κίνηση που διερχόταν μέσα από τη γραμμή διασύνδεσης (CE-PE) του πελάτη (Ε.Μ.Π.) με τον πάροχο (GRNET). Για την παθητική παρακολούθηση της γραμμής διασύνδεσης, ταχύτητας 1Gbps, χρησιμοποιήθηκαν οι τρεις τεχνικές παθητικής παρακολούθησης δικτύων: Packet Capturing, Netflow, SNMP. Χαρακτηριστικά γραφήματα που αποτυπώνουν τις μετρήσεις διαφόρων μετρικών κατά την διεξαγωγή επιθέσεων τύπου UDP και SYN flooding attack φαίνονται στα σχήματα 4.7-4.12.

Τα μετρικά που ξεχώρισαν μέσα από την έρευνα είναι:

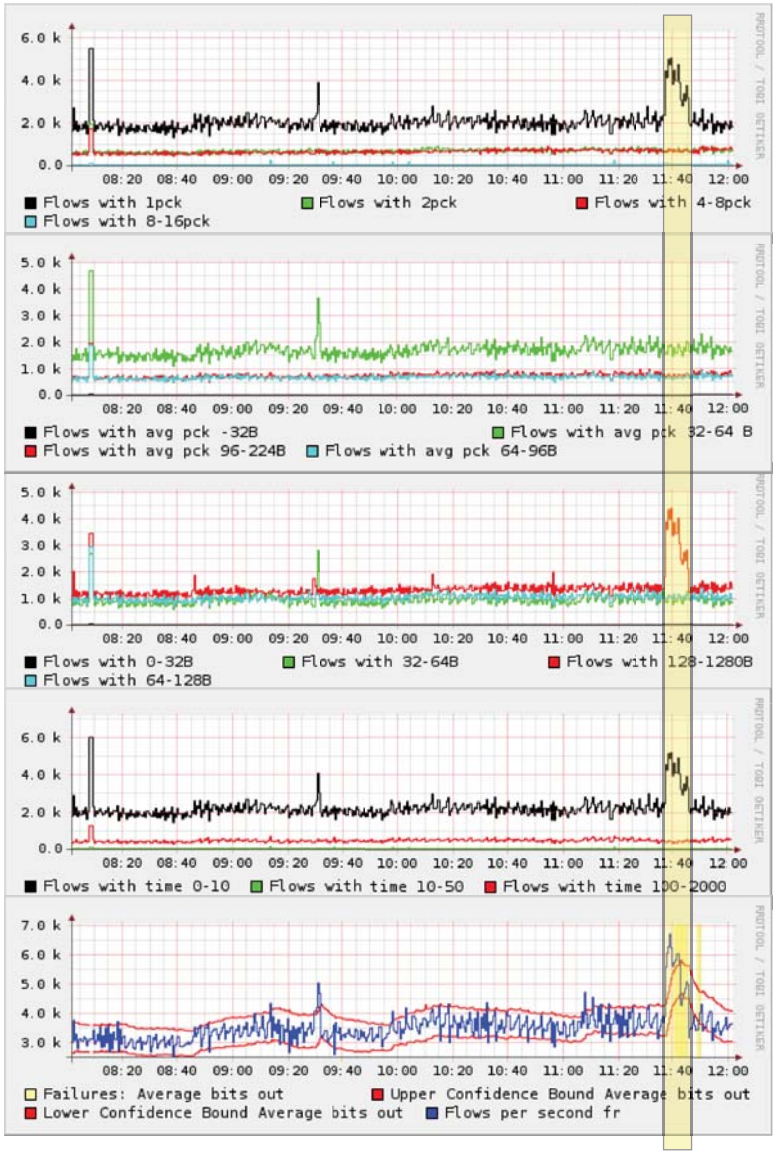
- **Ο λόγος των TCP πακέτων τύπου SYN προς τα πακέτα τύπου**



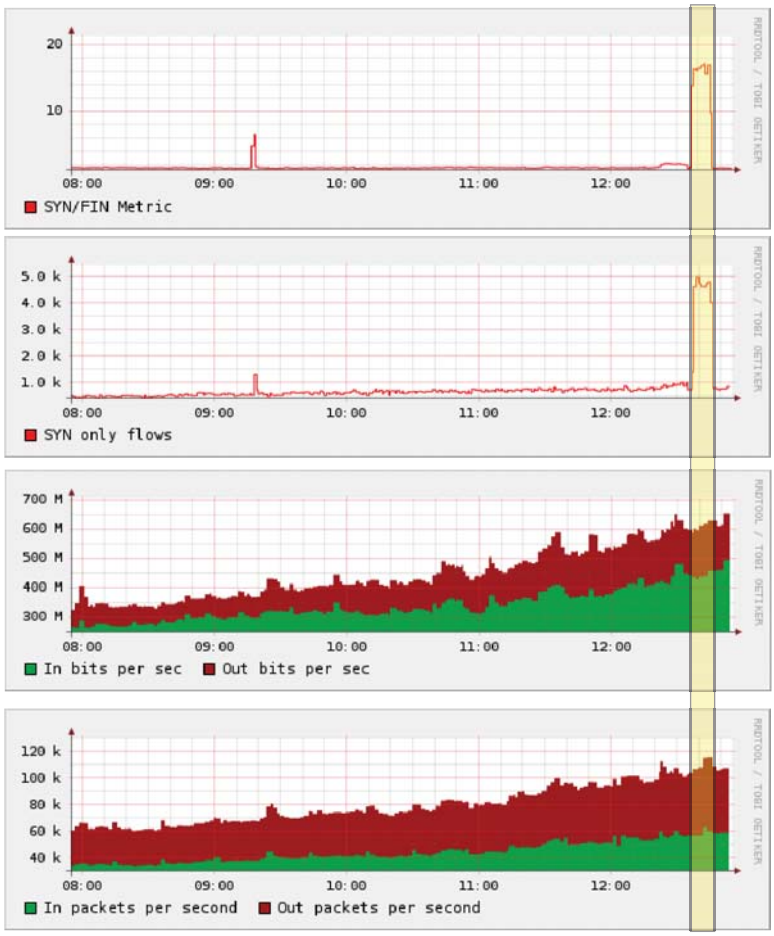
Σχήμα 4.7: Δείγματα μετρικών που περιέχουν επίθεση UDP με εμφανείς ενδείξεις ανωμαλίας (χωρίς data fusion).



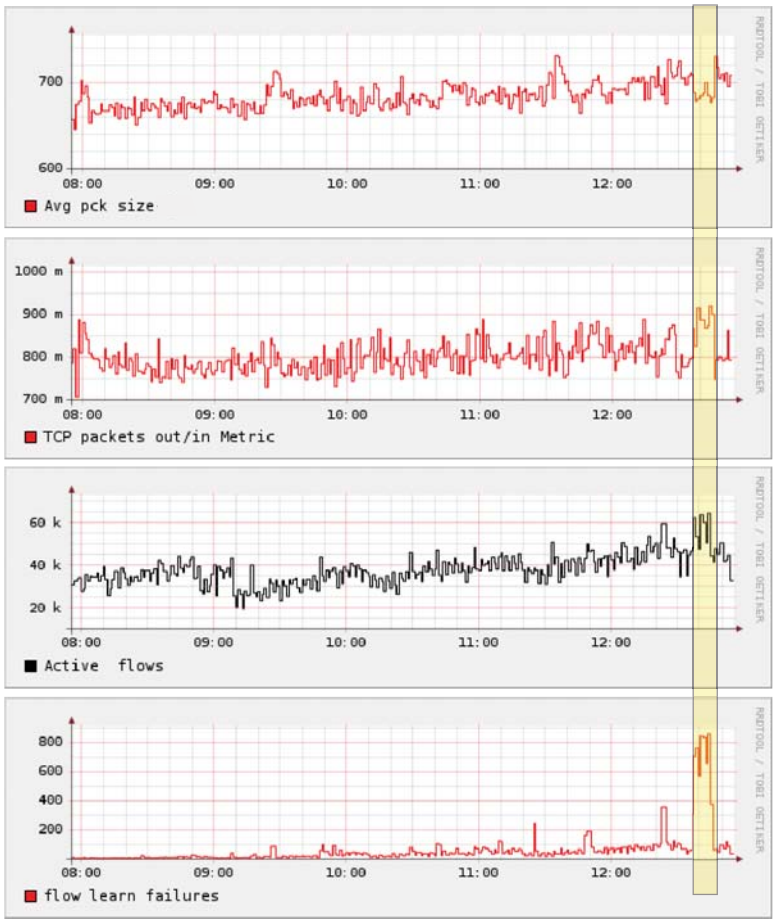
Σχήμα 4.8: Δείγματα μετρικών που περιέχουν επίθεση UDP χωρίς προφανείς ενδείξεις ανωμαλίας (χωρίς data fusion).



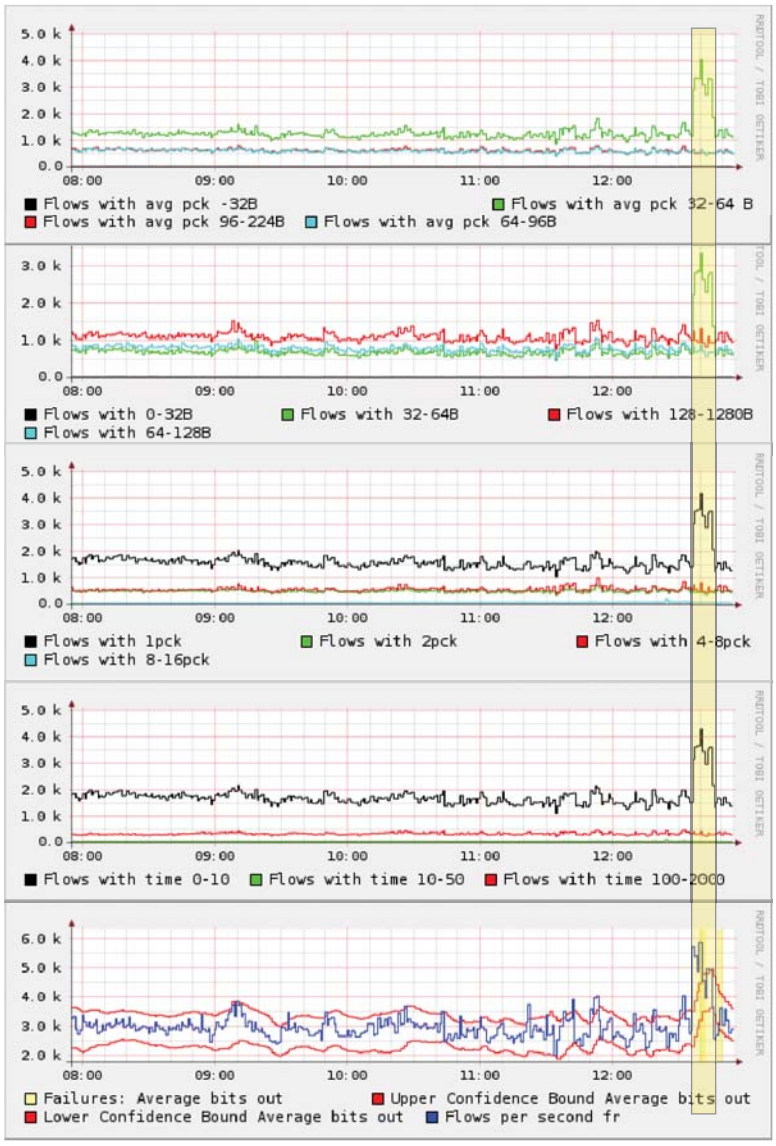
Σχήμα 4.9: Δείγματα μετρικών που περιέχουν επίθεση UDP με εμφανείς ενδείξεις ανωμαλίας (χωρίς data fusion).



Σχήμα 4.10: Δείγματα μετρικών που περιέχουν επίθεση SYN με εμφανείς ενδείξεις ανωμαλίας στα δύο πρώτα μετρικά (χωρίς data fusion).



Σχήμα 4.11: Δείγματα μετρικών που περιέχουν επίθεση SYN με ενδείξεις ανωμαλίας στο τελευταίο μετρικό (χωρίς data fusion).



Σχήμα 4.12: Δείγματα μετρικών που περιέχουν επίθεση SYN με εμφανείς ενδείξεις ανωμαλίας (χωρίς data fusion).

FIN (SYN-FIN Ratio - SFR). Συγκεκριμένα για την ανίχνευση μιας εξερχόμενης επίθεσης [3] υπολογίζουμε το λόγο:

$$SFR = \frac{\epsilon\xi\epsilon\rho\chi o\mu\epsilon\nu\alpha\, SYN\, packets/sec}{\epsilon\iota\sigma\epsilon\rho\chi o\mu\epsilon\nu\alpha\, FIN\, packets/sec}$$

Η αποτελεσματικότητα του μετρικού αυτού βασίζεται στην εγγενή συμμετρία του πρωτοκόλλου TCP. Το παρεμφερές αλλά λιγότερο αποτελεσματικό μετρικό $\frac{\epsilon\xi\epsilon\rho\chi o\mu\epsilon\nu\alpha\, TCP\, packets/sec}{\epsilon\iota\sigma\epsilon\rho\chi o\mu\epsilon\nu\alpha\, TCP\, packets/sec}$ ανά διεύθυνση IP έχει ήδη χρησιμοποιηθεί για τον σχεδιασμό συστημάτων ανίχνευσης επιθέσεων τύπου DDoS στα συστήματα MULTOPS [30] και D-WARD [29]. Ένα ακόμη μετρικό που είναι παρόμοιο με το προτεινόμενο χρησιμοποιήθηκε στο [75] και δεν ξεχώριζε τα πακέτα σε εισερχόμενα/εξερχόμενα, προσφέροντας έτσι τη δυνατότητα σε έναν επιτιθέμενο να ξεγελάσει το σύστημα ανίχνευσης με την αποστολή πακέτων τύπου FIN παράλληλα με SYN πακέτα. Το προτεινόμενο μετρικό SFR δεν παρουσιάζει το μειονέκτημα αυτό. Σε φυσιολογικές καταστάσεις δικτυακής κίνησης και με περίοδο δειγματοληψίας μεγαλύτερες από ένα δευτερόλεπτο το SFR λαμβάνει τιμές κοντά στην μονάδα χωρίς μεγάλες διακυμάνσεις. Συνεπώς επιθέσεις τύπου SYN μπορούν εύκολα να εντοπιστούν με το το μετρικό SFR όπως φαίνεται και από το σχήμα 4.10. Προφανώς το μετρικό SFR μπορεί να αναγνωρίσει μόνο επιθέσεις τύπου SYN αλλά είναι ιδιαίτερα αξιόπιστο και ακριβές. Για τη μέτρηση του SFR απαιτείται χρήση παρακολούθησης δικτύου με συλλογή πακέτων γιατί χρειάζεται εξέταση των flags κάθε TCP πακέτου. Μια παραλλαγή του μετρικού αυτού που βασίζεται σε Netflow είναι το πλήθος των flows μονάχα με SYN flag (SYN only flows -SOF).

- Ο **λόγος των εξερχόμενων προς τα εισερχόμενα bytes** αποτελεί καλή ένδειξη για επιθέσεις τύπου ICMP και UDP. Οι επιθέσεις αυτές

[3] από τον δρομολογητή του παρόχου PE προς το δίκτυο πελάτη CE

αποσκοπούν συνήθως στην εξάντληση του διαθέσιμου bandwidth μιας δικτυακής ζεύξης. Καθώς τα συγκεκριμένα πρωτόκολλα αποτελούν συνήθως μικρό ποσοστό της συνολικής κίνησης, μεγάλες διακυμάνσεις στον αριθμό των μεταδιδόμενων ICMP ή UDP bytes/sec είναι καλές ενδείξεις δικτυακών ανωμαλιών. Το προτεινόμενο μετρικό είναι ο λόγος (**ICMP Ratio -** IR**, UDP Ratio -** UR) :

$$UR = \frac{\varepsilon\xi\varepsilon\rho\chi o\mu\varepsilon\nu\alpha\, UDP\, bit/sec}{\varepsilon\iota\sigma\varepsilon\rho\chi o\mu\varepsilon\nu\alpha\, UDP\, bit/sec}\ ,\ IR = \frac{\varepsilon\xi\varepsilon\rho\chi o\mu\varepsilon\nu\alpha\, ICMP\, bit/sec}{\varepsilon\iota\sigma\varepsilon\rho\chi o\mu\varepsilon\nu\alpha\, ICMP\, bit/sec}$$

Αν και στο πρωτοτόκολο UDP δεν υπάρχει η συμμετρία που παρουσιάζεται στο TCP, οι τιμές του UR παραμένουν σχετικά σταθερές και εξαρτώνται από την χρήση διαφόρων εξυπηρετητών DNS, NFS και streaming στις εκάστοτε πλευρές της παρακολουθούμενης ζεύξης. Στο DWard project [29] οι Mirkovic et al χρησιμοποιούν παρόμοια μετρικά αλλά μετρούν πακέτα αντί για bytes. Η αποτελεσματικότητα των μετρικών αυτών είναι παραπλήσια αλλά οι λόγοι bytes παρουσιάζουν συνήθως μεγαλύτερες αυξομειώσεις στην περίπτωση επιθέσεων UDP. Ένα σημαντικό ερευνητικό αποτέλεσμα είναι πως το μετρικό IR, δηλαδή ο λόγος ICMP $\frac{\varepsilon\xi\varepsilon\rho\chi o\mu\varepsilon\nu\alpha\ bit/sec}{\varepsilon\iota\sigma\varepsilon\rho\chi o\mu\varepsilon\nu\alpha\ bit/sec}$ δεν αποτελεί μονάχα ένδειξη για επιθέσεις ICMP αλλά μπορεί να βοηθήσει και στην ανίχνευση επιθέσεων UDP. Στις περισσότερες περιπτώσεις επιθέσεων UDP δημιουργείται ένα αντίστροφο ρεύμα ICMP ως απάντηση στην λήψη UDP πακέτων σε κλειστό port όπως φαίνεται στο σχήμα 4.7. Συνεπώς το μετρικό IR είναι ευαίσθητο σε δύο δικτυακές ανωμαλίες: στα UDP και ICMP floods.

- Ο **αριθμός των flows με μικρή χρονική διάρκεια** (Flows with duration $< 100ms$ - $FT10$) είναι ένα μετρικό ευαίσθητο στις επιθέσεις που χρησιμοποιούν τυχαία source και destination ports παράγοντας έτσι flows με ελάχιστα πακέτα, με μικρή χρονική διάρκεια.

- Ο **ρυθμός των παραγόμενων flows** (Flow Rate - FR) που αποστέλλο-

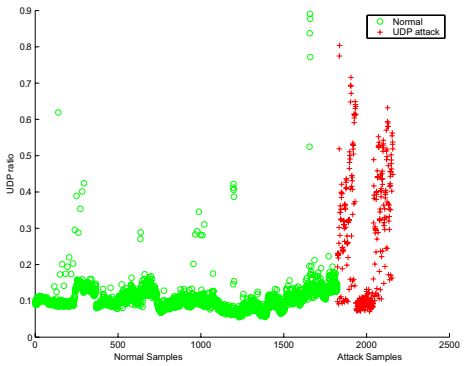
νται από έναν δρομολογητή που υποστηρίζει Netflow προς ένα σύστημα παρακολούθησης αποτελεί καλή ένδειξη για την ύπαρξη μιας δικτυακής ανωμαλίας και συγκεκριμένα επιθέσεων καταιγισμού πακέτων με χρήση της τεχνικής spoofing [126]. Το μετρικό FR διαφέρει από τον αριθμό των ενεργών flows στην flow cache ενός δρομολογητή AF. Το FR μετριέται σε flows per second και εκφράζει τα flow records που εξάγονται από τον δρομολογητή προς το σύστημα παρακολούθησης ανά δευτερόλεπτο, μετρούμε δηλαδή τον ρυθμό παραγωγής flows. Ο αριθμός των ενεργών flows AF μετριέται σε πλήθος flows και απεικονίζει και το ποσοστό πληρότητας της flow-cache στον δρομολογητή. Σε δίκτυα υψηλών ταχυτήτων και όταν δεν γίνεται χρήση sampled Netflow είναι συχνό φαινόμενο η flow cache να γεμίζει γρήγορα και η τιμή των ενεργών flows να κυμαίνεται κοντά στην μέγιστη δυνατή τιμή της. Στις περιπτώσεις αυτές δεν παρατηρείται μεγάλη αύξηση του AF σχήμα (4.8) όπως είναι θεωρητικά αναμενόμενο ενώ αντίθετα αυξάνεται απότομα η τιμή του **αριθμού των flow learning failures** -FLF. Πιο αναλυτικά, όταν τα flows που εισέρχονται σε ένα δρομολογητή δεν προλαβαίνουν να διαγραφούν ομαλά (graceful export) και η flow cache γεμίζει, ο δρομολογητής δεν βρίσκει κενές θέσεις για την καταγραφή νέων flow και το γεγονός αυτό σημειώνεται σαν ένα flow learning failure. Τα παλιότερα flow records σβήνονται άμεσα και συνεπώς ακόμη και οι φυσιολογικές ροές δεν παραμένουν στην cache αλλά δημιουργούνται και σβήνονται συνεχώς [33].

Εκτός από τα παραπάνω μετρικά χρήσιμες ενδείξεις μπορούν να προσφέρουν: ο **αριθμός των flows με 1 ή 2 πακέτα (Flows with 1 or 2 Packets** - $FP1$ **or** $FP2$), ο **αριθμός των flows με μικρό μέγεθος** μεταφερόμενης πληροφορίας (Flows with 32-64 bytes - $FB32$) και ο **αριθμός των flows με μεγάλο μέγεθος** μεταφερόμενης πληροφορίας (Flows with >600 bytes - $FB600$). Υψηλές τιμές των μετρικών $FP1$, $FP2$ και $FB32$ αποτελούν ένδειξη

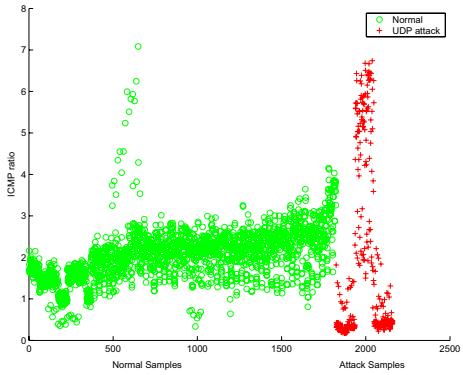
για μη επιτυχείς συνδέσεις και συνεπώς για δικτυακές ανωμαλίες. Ειδικότερα η τυχαία επιλογή source και destination ports κατά την διάρκεια επιθέσεων DDoS, port scans και κατά των εξάπλωση worms αυξάνει τον αριθμό των παραγόμενων flows που αντικατοπτρίζεται στις υψηλές τιμές αυτών των μετρικών. Το μετρικό $FB600$ είναι ευαίσθητο σε επιθέσεις τύπου UDP, ICMP flood που αποσκοπούν την εξάντληση του διαθέσιμου bandwidth και χρησιμοποιούν συνήθως πακέτα μεγάλου μεγέθους. Παραπλήσιο μετρικό είναι το μέσο μέγεθος πακέτων σε ένα flow αλλά στα πλαίσια της παρούσας μελέτης δεν αποδείχτηκε πιο αποτελεσματικό από το $FB600$.

Τα μετρικά που βασίζονται στον αριθμό των flows χωρίς να διαχωρίζουν τα διαφορετικά πρωτόκολλα δεν αποτελούν ενδείξεις για συγκεκριμένες επιθέσεις DDoS όπως SYN, UDP, ICMP floods αλλά αναδεικνύουν γενικά την ύπαρξη κάποιας ανωμαλίας όπως η χρήση των τεχνικών spoofing, port randomization, scanning κτλ. Γενικά, τα μετρικά προσφέρουν ενδείξεις και όχι σίγουρες αποδείξεις για την ύπαρξη μιας ανωμαλίας. Στα σχήματα 4.13(a)-4.13(c) βλέπουμε τις τιμές που λαμβάνουν ορισμένα μετρικά όταν μια δικτυακή ζεύξη βρίσκεται σε φυσιολογική κατάσταση ή υφίσταται μία επίθεση UDP. Τα δείγματα των μετρήσεων χωρίστηκαν αρχικά στις δύο αυτές καταστάσεις και στη συνέχεια αποτυπώθηκαν στα σχήματα δείγματα πολλαπλών χρονοσειρών (πειραμάτων) χωρίς χρονολογική σειρά. Στο αριστερό μέρος των σχημάτων βλέπουμε τις φυσιολογικές καταστάσεις αποτυπωμένες με κύκλους ενώ στο δεξί εμφανίζονται οι καταστάσεις επιθέσεων UDP με σταυρούς. Όπως φαίνεται, η φυσιολογική κατάσταση και μια κατάσταση ανωμαλίας π.χ. επίθεση UDP δεν είναι γραμμικά διαχωρίσιμες στο χώρο των πολλών μετρικών. Καθώς υπάρχει μια αρκετά μεγάλη ζώνη αβεβαιότητας η τιμή ενός μετρικού δεν αρκεί για να αποφανθούμε κατηγορηματικά για την κατάσταση του συστήματος δηλαδή την κατάταξη σε φυσιολογική ή μη κατάσταση. Παρατηρούμε περιπτώσεις ιδιαίτερα αυξημένων τιμών στη φυσιολογική κατάσταση ενώ αντίστοιχα καταγράφηκαν περιπτώσεις καταστάσεων ανωμαλίας που με απλή παρατήρηση

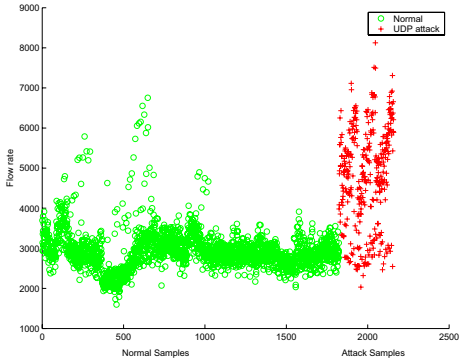
θα κατατάσσονταν ως φυσιολογικές. Επομένως προκύπτει η ανάγκη σύνθεσης πολλαπλών μετρικών για την ασφαλέστερη ανίχνευση ανωμαλιών.



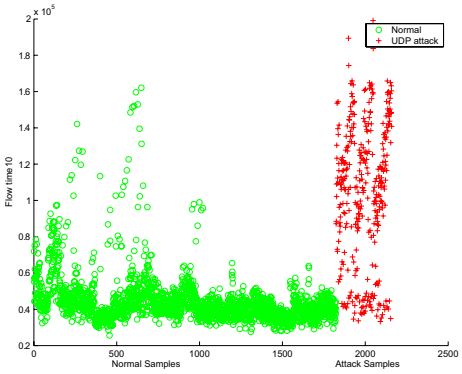
(a) Μετρικό UR



(b) Μετρικό IR



(c) Μετρικό FR



(d) Μετρικό FT10

Σχήμα 4.13: Αποτύπωση φυσιολογικών και καταστάσεων επίθεσης στο χώρο των μετρικών UR, IR, FR, FT10. (πολλαπλά πειράματα)

4.4 Ανίχνευση ανωμαλιών

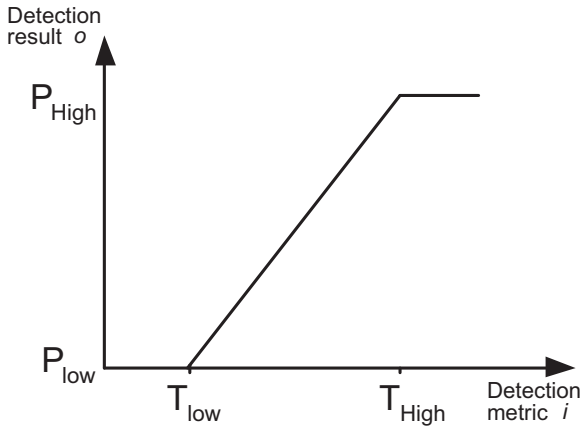
Ένας αισθητήρας μπορεί να ανιχνεύσει μια ανωμαλία στη δικτυακή κίνηση παρατηρώντας τα προτεινόμενα μετρικά. Στην απλούστερη περίπτωση ένας αισθητήρας διακρίνει δύο πιθανές καταστάσεις: την ύπαρξη ή μη ύπαρξη κά-

ποιας ανωμαλίας. Η είσοδος του υποσυστήματος ανίχνευσης κάθε αισθητήρα (detection method) είναι ένα διάνυσμα i με τις τιμές ενός ή περισσότερων μετρικών. Έξοδος o είναι το αποτέλεσμα της διαδικασίας ανίχνευσης. Στην απλή περίπτωση δύο πιθανών καταστάσεων το αποτέλεσμα της ανίχνευσης αποτυπώνεται είτε στις διακριτές τιμές 0 και 1 (binary, $o \in \{0, 1\}$) είτε στο συνεχές διάστημα $o \in [0, 1]$ εκφράζοντας την πιθανότητα ύπαρξης μιας από τις δύο καταστάσεις. Στη γενική περίπτωση ένας αισθητήρας είναι ένας ταξινομητής (classifier) μεταξύ περισσοτέρων των δυο καταστάσεων. Στην συνέχεια παρουσιάζονται σύντομα δημοφιλείς προσεγγίσεις για την υλοποίηση ενός ανιχνευτή.

1. Στην απλούστερη περίπτωση γίνεται χρήση σταθερών **συναρτήσεων κατωφλίου**. Παραδείγματα συναρτήσεων είναι οι βηματικές, γραμμικές και σιγμοειδής. Μια απλή περίπτωση είναι η χρήση μιας τμηματικά γραμμικής συνάρτησης. Για την παραμετροποίηση της από έναν διαχειριστή δικτύου απαιτείται ο ορισμός τεσσάρων τιμών: T_{Low}, T_{High}, P_{Low}, P_{High}. Αν το μετρικό i ξεπεράσει το ανώτατο όριο T_{High} υπάρχει ισχυρή ένδειξη ύπαρξης μιας ανωμαλίας που αντιστοιχεί στην μέγιστη πιθανότητα P_{High}. Αν το μετρικό i βρίσκεται κάτω από το κατώτατο όριο T_{Low} δεν υπάρχει ένδειξη για την ύπαρξη ανωμαλίας και αντιστοιχίζεται η ελάχιστη πιθανότητα P_{Low}. Στο ενδιάμεσο διάστημα υπάρχει μια γραμμική αύξηση της πιθανότητας ύπαρξης ανωμαλίας στο διάστημα $[P_{Low}, P_{High}] \in [0, 1]$. Ένα τέτοιο σύστημα περιγράφεται από την σχέση (σχήμα 4.14):

$$o = \begin{cases} P_{Low} & \text{if } i < T_{Low}; \\ P_{High} & \text{if } i > T_{High}; \\ \frac{(i - T_{Low}) * (P_{High} - P_{Low})}{T_{High} - T_{Low}} + P_{Low} & \text{otherwise.} \end{cases}$$

Υποθέσαμε λοιπόν ότι οι φυσιολογικές τιμές ενός μετρικού είναι γνω-



Σχήμα 4.14: Σταθερή συνάρτηση κατωφλίου ως ανιχνευτής.

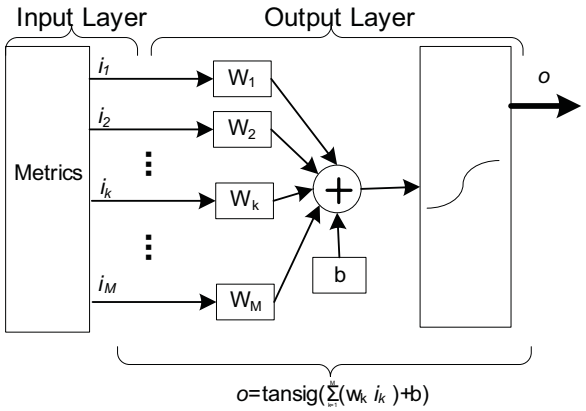
στές[4] και σύμφωνα με αυτές ρυθμίζεται η συνάρτηση κατωφλίου. Συνεπώς η ύπαρξη ανωμαλίας συνδέεται με την λήψη τιμών πέρα από προκαθορισμένα όρια. Σε άλλες περιπτώσεις, ως ύπαρξη ανωμαλίας θεωρείται η απότομη μεταβολή ενός μετρικού. Δηλαδή θεωρείται γνωστή η μέγιστη φυσιολογική μεταβολή της τιμής ενός μετρικού. Η μεταβολή του μετρικού μετριέται σε ένα κυλιόμενο παράθυρο (sliding window). Η προσέγγιση αυτή ονομάζεται συχνά προσαρμοζόμενη συνάρτηση κατωφλίου και υιοθετείται συχνά γιατί η ρύθμιση μιας σταθερής συνάρτησης κατωφλίου, δηλαδή του ορισμού φυσιολογικών ή μη ορίων των μετρούμενων χαρακτηριστικών (μετρικών), απαιτεί αρκετή προσπάθεια από τον διαχειριστή. Παρόλα αυτά η προσέγγιση αυτή παρουσιάζει το πρόβλημα ότι ο ανιχνευτής εστιάζει στο κυλιόμενο παράθυρο με αποτέλεσμα να μην ανιχνεύει ανωμαλίες με μεγάλη χρονική διάρκεια - μικρό ρυθμό μεταβολής.

Στο σημείο αυτό αξίζει να σημειωθεί η επίδραση αλλαγής του σχήματος και της θέσης της συνάρτησης κατωφλίου στην απόδοση του ανιχνευτή. Θεωρώντας λοιπόν έναν τυπικό ανιχνευτή, μετακινώντας την καμπύλη

[4]έχουν οριστεί αυθαίρετα ή καθοριστεί μετά από παρατήρηση, εκπαίδευση και στατιστική επεξεργασία.

πάνω και αριστερά ο ανιχνευτής γίνεται πιο ευαίσθητος αυξάνοντας πιθανώς τα false positives. Αντίθετα μια κίνηση κάτω και δεξιά κάνει τον ανιχνευτή λιγότερο ευαίσθητο αυξάνοντας πιθανώς τα false-negatives. Συμπιέζοντας την καμπύλη ως προς τον άξονα x κινούμαστε προς δυαδική ανίχνευση (binary detection), ενώ επιμηκύνοντάς την ο ανιχνευτής παρουσιάζει μεγαλύτερη αβεβαιότητα αλλά και διάστημα απόκρισης. Τα στοιχεία αυτά είναι ενδεικτικά της δυσκολίας παραμετροποίησης ορισμένων ανιχνευτών με απλή παρατήρηση. Για το λόγο αυτό υπάρχει συχνά η ανάγκη χρήσης προσαρμοζόμενων συναρτήσεων κατωφλίου και σταθερών συναρτήσεων κατωφλίου που έχουν ρυθμιστεί με τεχνικές επιβλεπόμενης μάθησης.

2. Η χρήση ενός **τεχνητού νευρωνικού δικτύου** (Artificial Neural Network - ANN) προσφέρει τη δυνατότητα υλοποίησης ενός πιο σύνθετου ανιχνευτή. Τα ANN μπορούν να χρησιμοποιηθούν ως ανιχνευτές (ταξινομητές) περισσότερων των δύο καταστάσεων. Ένα ANN μπορεί να λαμβάνει ως είσοδο $\boldsymbol{i} = [i_1, i_2, ..., i_M]^T$ ένα ή περισσότερα (M) μετρικά και να παράγει ως έξοδο του ανιχνευτή ένα διάνυσμα $\boldsymbol{o} = [o_1, o_2, ..., o_N]$ στο χώρο των πιθανών καταστάσεων. Είναι επίσης δυνατό να εισάγονται περισσότερες της μίας τιμές για κάθε μετρικό (ιστορική εξέλιξη) όπως στην περίπτωση του ανιχνευτή κυλιόμενου παραθύρου. Η χρήση ενός νευρωνικού δικτύου επιβλεπόμενης μάθησης αντιμετωπίζει αποτελεσματικά το πρόβλημα παραμετροποίησης του ανιχνευτή, δηλαδή της εύρεσης των φυσιολογικών ορίων των μετρικών. Σύμφωνα με την προσέγγιση επιβλεπόμενης offline μάθησης ο διαχειριστής μαθαίνει στο νευρωνικό δίκτυο (ανιχνευτή) τι αποτελεί ανωμαλία και τι όχι, υποδεικνύοντας την σωστή κατάσταση του δικτύου (έξοδο $\boldsymbol{o}$) για διαφορετικά χαρακτηριστικά δείγματα δεδομένων εισόδου $\boldsymbol{i} = [i_1, i_2, ..., i_M]^T$ (μετρικών). Η εκπαίδευση του ANN είναι αρκετά εύκολη για το διαχειριστή αφού το ANN προσαρμόζεται αυτόματα. Καθώς δεν ακολουθείται η λογική της



Σχήμα 4.15: Στοιχειώδες νευρωνικό δίκτυο Perceptron ως ανιχνευτής.

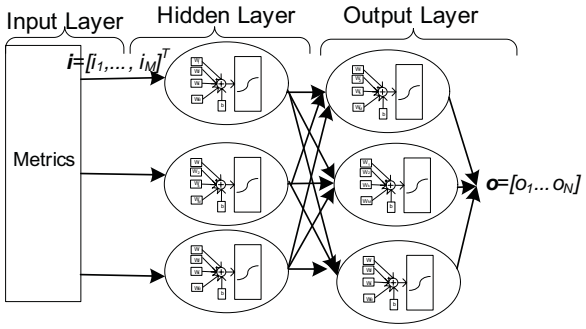
online μάθησης, δεν εμφανίζεται το πρόβλημα της μη επιθυμητής αυτόματης προσαρμογής του συστήματος που είναι εφάμιλλο με αυτό που σχολιάστηκε για τις προσαρμοζόμενες συναρτήσεις κατωφλίου.

Παράδειγμα ενός απλού νευρωνικού δικτύου που μπορεί να χρησιμοποιηθεί ως ανιχνευτής δύο γραμμικά διαχωρίσιμων καταστάσεων είναι το Perceptron [91]. Στο Perceptron κάθε νευρώνας υπολογίζει ένα γραμμικό συνδυασμό των εισόδων του, στον οποίο εφαρμόζει μια συνάρτηση κατωφλίου (ενεργοποίησης). Για τον υπολογισμό του γραμμικού συνδυασμού χρησιμοποιείται ένας πίνακας βαρών $\boldsymbol{IW} = [w_1, w_2, ..., w_M]$ (weights) και το bias b. Οι παράμετροι αυτοί που καθορίζονται κατά την περίοδο εκπαίδευσης. Στο σχήμα 4.15 απεικονίζεται ένα Perceptron που αποτελείται από ένα νευρώνα με σιγμοειδή συνάρτηση ενεργοποίησης και μονοδιάστατη έξοδο o που περιγράφεται από την εξίσωση:

$$o = tansig(\sum_{k=1}^{M}(w_k\ i_k) + b) = tansig(\boldsymbol{IW}\ \boldsymbol{i} + b)$$

όπου $tansig(n) \approx tanh(n)$ (hyperbolic tangent sigmoid).

Στη γενική περίπτωση ταξινόμησης περισσότερων καταστάσεων, όπως για τη ταξινόμηση ιών (viruses) ανάλογα με διάφορα χαρακτηριστικά



Σχήμα 4.16: Ένα Multi-Layer Perceptron ως ιεραρχική δομή από πολλά Perceptrons.

τους [127], πρέπει να χρησιμοποιηθεί ένα Multi-Layer Perceptron (MLP). Ένα Multi-Layer Perceptton είναι μια ιεραρχική δομή από πολλά Perceptrons (σχήμα 4.16). Ένα απλό ευθύ (feed forward) MLP αποτελείται από νευρώνες που οργανώνονται σε τρία στρώματα: M στο στρώμα εισόδου, H νευρώνες στο κρυμμένο στρώμα και N νευρώνες στο στρώμα εξόδου. Σύμφωνα με το "θεώρημα προσέγγισης" του Kolmogorov [91,128,129] χρησιμοποιώντας τουλάχιστον $H = 2M + 1$ νευρώνες στο κρυμμένο στρώμα το νευρωνικό δίκτυο μπορεί να προσεγγίσει κάθε συνεχή συνάρτηση $F : [0,1]^M \rightarrow \Re^N$, δηλαδή μπορεί να αναγνωρίζει N μη-γραμμικά διαχωρίσιμες καταστάσεις. Μια σημαντική απαίτηση είναι η χρήση μη γραμμικών συναρτήσεων ενεργοποίησης όπως μια σιγμοειδής συνάρτηση. Προκειμένου το MLP να μπορεί να ταξινομεί σωστά πρέπει να εκπαιδευτεί με έναν αλγόριθμο εκπαίδευσης όπως ο *Backpropagation*. Μέσα από την επαναλαμβανόμενη εισαγωγή ενός συνόλου εισόδων $\boldsymbol{i}$ και των επιθυμητών εξόδων $\boldsymbol{o}$ στο νευρωνικό δίκτυο, ο αλγόριθμος *Backpropagation* υπολογίζει τις τιμές των βαρών και biases $(\boldsymbol{IW}, \boldsymbol{b})$ των επιμέρους νευρώνων ώστε να ελαχιστοποιηθεί το μέσο τετραγωνικό σφάλμα μεταξύ των πραγματικών και επιθυμητών εξόδων.

4.5 Σύνθεση δεδομένων ανίχνευσης από πολλούς αισθητήρες

Όπως αναφέρθηκε στο κεφάλαιο 3 υπάρχουν πολλές διαφορετικές προσεγγίσεις για την σύνθεση δεδομένων που προέρχονται από τους επιμέρους αισθητήρες ανίχνευσης ανωμαλιών στη δικτυακή κίνηση. Η θεωρία D-S παρουσιάζει ιδιαίτερο ενδιαφέρον λόγω της μεγάλης της ευελιξίας στην μοντελοποίηση των δεδομένων που προέρχονται από τους επιμέρους αισθητήρες ενός Network Anomaly Detection System (NADS). Η ενότητα αυτή παρουσιάζει τον τρόπο με τον οποίο μπορούν να εκφραστούν τα αποτελέσματα ανίχνευσης των αισθητήρων προκειμένου στην συνέχεια να συνθεθούν με χρήση της θεωρίας D-S. Οι πιθανές καταστάσεις του υπό παρακολούθηση δικτυακού στοιχείου θεωρούνται γνωστές και σταθερές. Για λόγους απλότητας θεωρούμε ως υπό παρακολούθηση δικτυακό στοιχείο **μια δικτυακή ζεύξη** και ως πιθανές καταστάσεις (frame of discernment) την φυσιολογική κατάσταση και τους τρεις πιο συχνούς τύπους κατανεμημένων επιθέσεων άρνησης υπηρεσιών [25]: Θ ={NORMAL, SYN-flood, UDP-flood, ICMP-flood}.

Κάθε αισθητήρας ανάλογα με το μετρικό στο οποίο βασίζεται έχει την δυνατότητα να ανιχνεύει ένα συγκεκριμένο σύνολο από αυτές τις δικτυακές ανωμαλίες. Στην απλούστερη περίπτωση ένας αισθητήρας λειτουργεί ως απλός ανιχνευτής, δηλαδή διαχωρίζει μεταξύ δυο συνόλων καταστάσεων. Τότε ο αισθητήρας μπορεί να εκφράσει το αποτέλεσμα της ανίχνευσης με τον ορισμό μιας συνάρτησης m (basic probability assignment) και συγκεκριμένα αναθέτει μια τιμή (m-value) σε τρία σύνολα:

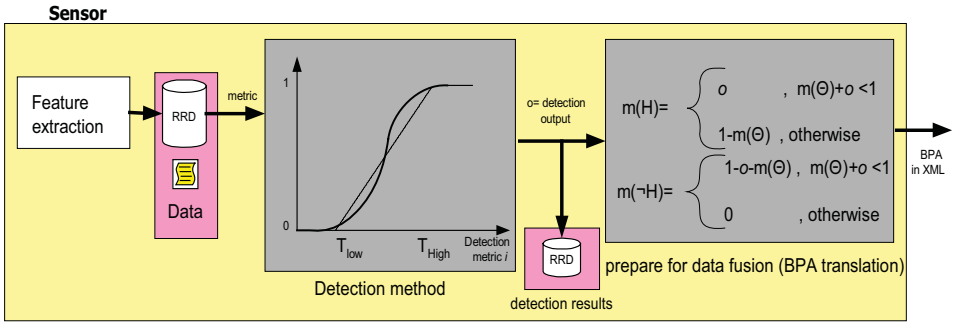
- στο σύνολο H των καταστάσεων όπου ο αισθητήρας μπορεί να αναγνωρίσει και στις οποίες είναι ευαίσθητος : $m(H)$
- στο σύνολο $\neg H$, για να αποτυπωθούν οι ενδείξεις που αμφισβητούν ή αντικρούουν την υπόθεση H: $m(\neg H)$.
- στο σύνολο Θ για να εκφραστεί η άγνοια (ignorance) του αισθητήρα και η πιθανότητα να σφάλει: $m(\Theta)$.

Σύμφωνα με την εξίσωση (3.8) προκύπτει ότι $m(H) + m(\neg H) + m(\Theta) = 1$. Ορίζοντας ένα bpa για κάθε αισθητήρα με αυτό τον τρόπο, μπορούμε να εισάγουμε "εξειδικευμένη γνώση" στο σύστημα (expert knowledge) και συγκεκριμένα γνώση που αφορά τις δυνατότητες ανίχνευσης κάθε αισθητήρα. Ο ειδικός (human expert), που σχεδιάζει έναν αισθητήρα, μπορεί να καθορίσει εξ αρχής προς ποιες κατευθύνσεις βοηθά ο αισθητήρας στην διάγνωση της πραγματικής κατάστασης του υπό παρακολούθηση συστήματος. Για παράδειγμα οι αισθητήρες που βασίζονται στα μετρικά UDP Ratio (UR), SYN-FIN Ratio (SFR), Flow Rate (FR), Flows with duration $< 100ms$ ($FT10$) που είδαμε στην ενότητα 4.3 είναι ευαίσθητοι στις καταστάσεις:

- $H_{UR} = \{UDP - flood\}$, καθώς το UR ανιχνεύει μονάχα UDP-floods.
- $H_{SFR} = \{SYN - flood\}$, καθώς το SFR ανιχνεύει μονάχα SYN-floods.
- $H_{FR} = \{SYN - flood, UDP - flood, ICMP - flood\}$, καθώς το FR ανιχνεύει SYN, UDP, ICMP-floods αλλά δεν μπορεί να τα διαχωρίσει.
- $H_{FT10} = \{SYN - flood, UDP - flood, ICMP - flood\}$, καθώς το $FT10$ ανιχνεύει SYN, UDP, ICMP-floods αλλά δεν μπορεί να τα διαχωρίσει.

Προβλέπεται επίσης η εισαγωγή γνώσης για την αξιοπιστία κάθε αισθητήρα. Κάθε αισθητήρας έχει ορισμένο ποσοστό άγνοιας και πιθανότητας σφάλματος στις εκτιμήσεις του, που μπορεί να είναι ανάλογο με τα false positives που παράγει. Για απλότητα θεωρείται σταθερό και αποτυπώνεται με την ανάθεση της τιμής $m(\Theta) \in [0, 1]$.

Πρέπει να σημειωθεί πως τα παραπάνω αποτελούν γενικές κατευθύνσεις και πως δεν υπάρχει γενικός κανόνας σύμφωνα με τον οποίο ένας αισθητήρας πρέπει να εκφράζει την εκτίμηση του για την κατάσταση του συστήματος. Ο ακριβής τρόπος με τον οποίο μεταφράζεται το αποτέλεσμα της ανίχνευσης σε ένα basic probability assignment (bpa) (σχήμα 4.17) εξαρτάται από τον αλγόριθμο ανίχνευσης που υλοποιεί ο αισθητήρας. Ενδεικτικά ένας απλός αισθητήρας που ανιχνεύει ένα σύνολο καταστάσεων H ορίζει ένα bpa σύμφωνα



Σχήμα 4.17: Ενδεικτικός τρόπος ορισμού ενός basic probability assignment (bpa) από το αποτέλεσμα ανίχνευσης ενός αισθητήρα

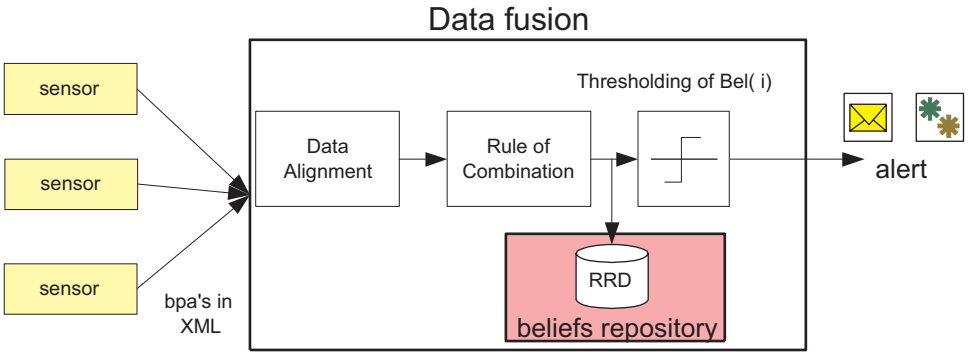
με τις ακόλουθες σχέσεις:

$$m(H) = \begin{cases} o & \text{, if } m(\Theta) + o < 1 \\ 1 - m(\Theta) & \text{, otherwise} \end{cases}$$

$$m(\neg H) = \begin{cases} 1 - o - m(\Theta) & \text{, if } m(\Theta) + o < 1 \\ 0 & \text{, otherwise} \end{cases}$$

όπου ο είναι το αποτέλεσμα του αλγόριθμου ανίχνευσης.

Μετά την παρουσίαση του τρόπου έκφρασης των αποτελεσμάτων ανίχνευσης των επιμέρους αισθητήρων σε bpa, θα σχολιάσουμε τη διαδικασία της σύνθεσης τους (σχήμα 4.18). Αρχικά θα αναφερθούμε στον τρόπο μεταφοράς των bpa στο σημείο σύνθεσης. Οι αισθητήρες περιοδικά μετρούν, ανιχνεύουν και στην συνέχεια μεταφράζουν τα αποτελέσματα ανίχνευσης σε bpa που πρέπει να συγκεντρωθούν στον κόμβο σύνθεσης δεδομένων. Η μετάδοση των δεδομένων αυτών προτείνεται να γίνει με την λογική των Web Services όπου τα δεδομένα κωδικοποιούνται σε XML. Η προσέγγιση αυτή είναι συμβατή και μπορεί να θεωρηθεί ως επέκταση του πρωτοκόλλου IDMEF [130] για την επικοινωνία συστημάτων IDS. Ως παράδειγμα αναφέρουμε ότι ο ορισμός ενός *bpa* (*m*-function definition) μπορεί να γραφεί:



Σχήμα 4.18: Σύνθεση δεδομένων από πολλούς αισθητήρες ανίχνευσης

```
<bpa>
<Timestamp> XXX </Timestamp>
<sensorid> YYYY </sensorid>
<mvalue>
          <hypothesis set>
             <state> θ1 </state>
             ...
             <state> θN </state>
          </hypothesis set>
<value> ZZZ </value>
</mvalue>
<mvalue> ...</mvalue>
</bpa>
```

Κατά την συγκέντρωση των bpa από τους διάφορους αισθητήρες πρέπει να εκτελεστεί ένα σημαντικό στάδιο επεξεργασίας: η εναρμόνιση δεδομένων (Data Alignment). Συγκεκριμένα πρέπει να εναρμονιστούν τα δεδομένα από τους αισθητήρες στο πεδίο του χρόνου (time alignment) καθώς οι αισθητήρες μπορεί να μην είναι συγχρονισμένοι ή να έχουν ακόμη και διαφορετικές περιόδους δειγματοληψίας. Κατά την διαδικασία αυτή τηρείται μια βάση γνώσης (belief pool) που ανανεώνεται ανάλογα με τις περιοδικές εκτιμήσεις των

αισθητήρων. Στη συνέχεια η σύνθεση δεδομένων γίνεται με χρήση του κανόνα σύνθεσης του Dempster (εξίσωση 3.12- Rule of Combination) και της προσεταιριστικής του ιδιότητας για τα δεδομένα όλων των αισθητήρων:

$$m = (...(m_1 \oplus m_2) \oplus m_3)... \oplus m_M)$$

Κατόπιν υπολογίζονται τα belief intervals για κάθε μέλος του συνόλου Θ (frame of discernment). Τέλος η πραγματική κατάσταση του υπό παρακολούθηση συστήματος εκτιμάται με το κανόνα της μέγιστης πίστης:

$$k = \arg \max_{1 \leq l \leq N} [m(\theta_l)]$$

Επιπρόσθετα μπορούμε να ανιχνεύσουμε ανωμαλίες χρησιμοποιώντας μια σταθερή συνάρτηση κατωφλίου στην τιμή Bel (belief function, ενότητα 3.5.2) των πιθανών καταστάσεων. Για παράδειγμα ένας κανόνας παραγωγής μηνύματος συναγερμού (alert) για την ύπαρξη επίθεσης UDP προκύπτει αν θέσουμε τιμή κατωφλίου T=0.7 στην εξίσωση:

$$UDP - alert = \begin{cases} true, & \text{if Bel}(\{\text{UDP}\}) > \text{T}; \\ false, & \text{otherwise.} \end{cases} \tag{4.1}$$

4.6 Ταυτοποίηση επιθέσεων και μέθοδοι καταστολής

Μετά την ανίχνευση μιας ανωμαλίας στη δικτυακή κίνηση και προκειμένου να ληφθούν μέτρα καταστολής της, η ανωμαλία πρέπει να ταυτοποιηθεί. Με τον όρο ταυτοποίηση εννοούμε τον προσδιορισμό των χαρακτηριστικών μιας ανωμαλίας ώστε να μπορεί να διαχωριστεί από την υπόλοιπη "φυσιολογική" κίνηση. Τα χαρακτηριστικά αυτά πρέπει να είναι όσο το δυνατό πιο ακριβή για να αποφεύγονται παράπλευρες απώλειες (collateral damage), δηλαδή η επιβολή μέτρων καταστολής σε φυσιολογική κίνηση που εσφαλμένα χαρακτηρίστηκε ως ανωμαλία. Πρακτικά η ταυτοποίηση γίνεται με τον ορισμό ενός φίλτρου

δικτυακής κίνησης (Access Control List - ACL). Η διαδικασία ταυτοποίησης μιας ανωμαλίας εκτελείται ασύγχρονα (μόνο μετά την ανίχνευση μιας ανωμαλίας) και μπορεί να διαρκεί αρκετά μεγάλο χρονικό διάστημα, ικανό ώστε να επεξεργαστούν όλα τα αποθηκευμένα δεδομένα με αλγόριθμους που θα υπόκεινται σε μικρότερους περιορισμούς σε πολυπλοκότητα χρόνου και χώρου (time and storage complexity) από ότι οι αλγόριθμοι ανίχνευσης ανωμαλιών. Χαρακτηριστικό παράδειγμα τέτοιων αλγόριθμων είναι αλγόριθμοι clustering για την εύρεση ενός ακριβούς, εύστοχου και συγκεκριμένου συνοθηλεύματος κίνησης (traffic aggregate), δηλαδή ενός συνόλου πακέτων με κοινά χαρακτηριστικά [19]. Ένα δεύτερο παράδειγμα είναι η ανάλυση κάθε ροής TCP στο χώρο της συχνότητας και ο χαρακτηρισμός της ως φυσιολογική ή μη ανάλογα με την περιοδικότητα που εμφανίζει [80]. Στην βιβλιογραφία το πρόβλημα της ταυτοποίησης αντιμετωπίζεται ενιαία με το πρόβλημα της ανίχνευσης χρησιμοποιώντας μετρικά ανά διεύθυνση προέλευσης/προορισμού (per destination metrics). Όπως έχει ήδη αναφερθεί τα μετρικά αυτά παρουσιάζουν δυσκολίες στην μέτρηση τους σε δίκτυα υψηλών ταχυτήτων και για το λόγο αυτό προτείνεται ο διαχωρισμός των προβλημάτων της ανίχνευσης και ταυτοποίησης (ενότητα 2.5.1.3). Οι αλγόριθμοι για την ταυτοποίηση ανωμαλιών αποτελούν ένα ξεχωριστό ερευνητικό θέμα το οποίο η παρούσα εργασία δεν ερευνά. Το τελικό βήμα για την αντιμετώπιση μιας δικτυακής ανωμαλίας, έχοντας τα χαρακτηριστικά της με την μορφή μιας ACL, είναι η εφαρμογή μεθόδων καταστολής όπως φίλτρα δικτυακής κίνησης (access list), rate limiting και traffic shunt (ενότητα 2.5.2). Το βήμα αυτό μπορεί εύκολα να αυτοματοποιηθεί με χρήση scripts για να παρακαμφθεί ο ανθρώπινος παράγοντας, επιτυγχάνοντας γρηγορότερη απόκριση.

Κεφάλαιο 5

Υλοποίηση συστήματος ανίχνευσης και πειραματικά αποτελέσματα

5.1 Πρωτότυπη υλοποίηση

Η προτεινόμενη αρχιτεκτονική για την δημιουργία ενός συστήματος ανίχνευσης ανωμαλιών στη δικτυακή κίνηση (Network Anomaly Detection System - NADS) στηρίζεται στην ύπαρξη αισθητήρων που συλλέγουν και αναλύουν δεδομένα (feature extraction) που προέρχονται από την υπάρχουσα υποδομή παρακολούθησης του δικτύου. Οι αισθητήρες του συστήματος επεξεργάζονται τα μετρούμενα χαρακτηριστικά (feature) της δικτυακής κίνησης με αλγόριθμους ανίχνευσης (detection algorithm). Τα μετρούμενα χαρακτηριστικά ονομάζονται **μετρικά** (metrics). Οι αισθητήρες είναι αυτόνομοι και μπορούν να βασίζονται σε εντελώς διαφορετικά μετρικά αλλά και μεθοδολογίες ανίχνευσης π.χ. misuse ή anomaly detection. Τα δεδομένα των αισθητήρων συνθέτονται (data fusion) και εξάγεται το συνολικό συμπέρασμα της διαδικασίας ανίχνευσης.

Κατά την εκπόνηση της παρούσας διατριβής υλοποιήθηκε ένα πρωτότυπο σύστημα που ανιχνεύει επιθέσεις καταιγισμού πακέτων (DDoS) σύμφωνα με την προτεινόμενη αρχιτεκτονική. Το σύστημα τροφοδοτείται με μετρήσεις από μια δικτυακή ζεύξη με χρήση τεχνικών παθητικής παρακολούθησης δικτύου. Υλοποιήθηκαν εργαλεία με διακριτές στοιχειώδεις λειτουργίες που βασίζονται σε λογισμικό ανοικτού κώδικα (open source software). Ιδιαίτερη προσοχή δόθηκε ώστε τα εργαλεία αυτά να είναι ανεξάρτητα και να μπορούν να τροποποιηθούν ή να αντικατασταθούν εύκολα από άλλους ερευνητές που θα μελετήσουν διαφορετικές προσεγγίσεις ανίχνευσης και θα χρησιμοποιήσουν άλλες γλώσσες προγραμματισμού. Το σύστημα υλοποιήθηκε ώστε να μπορεί να συλλέγει και να επεξεργάζεται στοιχεία σε μικρή κλίμακα χρόνου (time scale) προκειμένου να ανιχνεύονται ανωμαλίες πολύ μικρής διάρκειας παρόλο που σε πρακτικό επίπεδο η δυνατότητα αυτή δεν είναι αξιοποιήσιμη μέχρι να αναπτυχθούν αυτοματοποιημένες μέθοδοι καταστολής.

Για την υλοποίηση αισθητήρων-ανιχνευτών αναπτύχθηκαν τα παρακάτω εργαλεία:

- Τα εργαλεία netrrd και snortstat [131] βασίζονται στην παθητική συλλογή πακέτων. Το εργαλείο netrrd είναι βασισμένο στο ngrep [132] ενώ το snortstat είναι ένα preprocessor plugin για το δημοφιλές IDS σύστημα Snort [133]. Και τα δύο είναι γραμμένα στη γλώσσα C και στηρίζονται στην βιβλιοθήκη libpcap [134] για την συλλογή πακέτων. Τα εργαλεία εξετάζουν όλα τα πακέτα που συλλέγονται και με την χρήση απλών μετρητών υπολογίζουν τα μετρικά: SFR, UR, IR, IBPS, IPPS, TBPS, TPPS, UBPS, UPPS που αναλύθηκαν στην ενότητα 4.3. Οι τιμές αυτές καταγράφονται περιοδικά σε μια Round Robin Database [110]. Τα εργαλεία αυτά, λόγω της απλότητας τους, είναι ικανά να λειτουργήσουν σε υψηλές ταχύτητες δικτύου χωρίς να απαιτείται ειδικός εξοπλισμός. Συγκεκριμένα, όταν τα εργαλεία χρησιμοποιήθηκαν για την εξέταση δικτυακής κίνησης της τάξης των 250Mbps σε ένα σύ-

στημα με μια απλή κάρτα δικτύου 1GE χωρίς ειδικούς drivers [1] και με μεγάλη συχνότητα δειγματοληψίας π.χ. f=1/30sec, η απώλεια πακέτων κυμάνθηκε σε επίπεδο $< 0,1\%$ (packet drops). Βεβαίως σε μεγαλύτερες ταχύτητες ο ρυθμός απώλειας πακέτων γίνεται ιδιαίτερα αισθητός και απαιτείται ειδικός εξοπλισμός [113] ή η χρήση δικτυακών επεξεργαστών (network processors) [114]. Υπάρχουν πάντως αναφορές ότι με χρήση εξειδικευμένου λογισμικού επεξεργασίας πακέτων ακόμη και με έναν κοινό υπολογιστή μπορούμε να επεξεργαστούμε πακέτα σε ταχύτητα 1 Gbps [135].

- Το εργαλείο flow-rrd-receiver [136] βασίζεται στα 'OSU Flow-tools' [137] και αποτελεί έναν ειδικό Netflow collector που συλλέγει τα flow-records ενός δρομολογητή και υπολογίζει τα ακόλουθα μετρικά: FP1, FP2, FB32, FT10, FR, SOF, UR, IR, IBPS, IPPS, TBPS, TPPS, UBPS, UPPS που αναλύθηκαν στην ενότητα 4.3. Οι τιμές των μετρικών καταγράφονται περιοδικά σε μια Round Robin Database.

- Το εργαλείο snmpcollector είναι ένα απλό Perl script που εκτελεί SNMP requests σε έναν δρομολογητή, ανακτώντας τιμές από τους πίνακες μιας ή περισσότερων Management Information Base (MIB) και αποθηκεύει τις τιμές τους σε μια Round Robin Database. Χρήσιμα μετρικά που μπορούμε να μετρήσουμε είναι το CPU utilization, per interface packet drops και flow learning failures (FLF).

Όπως επισημάναμε όλα τα εργαλεία καταγράφουν τις μετρήσεις τους σε Round Robin Databases (RRD). Τα εργαλεία δεν χρειάζεται να είναι συγχρονισμένα και να έχουν τις ίδιες περιόδους δειγματοληψίας. Ο συγχρονισμός μεταξύ των μετρήσεων μπορεί να υλοποιηθεί εύκολα χρησιμοποιώντας το εργαλείο rrd-xport. Είναι επίσης σημαντικό ότι κρατείται ιστορικό των μετρήσεων

[1] συγκεκριμένα σε ένα σύστημα Pentium 4 2.4GHz με λειτουργικό σύστημα FreeBSD 5.x

με καθορισμένη χρονική ακρίβεια που ορίζεται κατά την αρχικοποίηση μιας RRD.

Χρησιμοποιώντας τη γλώσσα Perl και τις δυνατότητες του εργαλείου RRDtool υλοποιήθηκαν ανιχνευτές σταθερών και προσαρμοζόμενων συναρτήσεων κατωφλίου που επεξεργάζονται τα καταγεγραμμένα μετρικά. Μέρος της αξιολόγησης του συστήματος πραγματοποιήθηκε στο υπολογιστικό περιβάλλον Matlab όπου χρησιμοποιήθηκαν και ανιχνευτές επιβλεπόμενης μάθησης. Η σύνθεση δεδομένων, δηλαδή η εξαγωγή του τελικού αποτελέσματος της ανίχνευσης βάσει των αποτελεσμάτων των επιμέρους ανιχνευτών, υλοποιείται από ένα εργαλείο εξαγωγής συμπερασμάτων σύμφωνα με την θεωρία Dempster-Shafer [131]. Το εργαλείο υλοποιήθηκε σε C και δέχεται τα basic probability assignments από τους αισθητήρες - ανιχνευτές και εφαρμόζει διαδοχικά τον κανόνα της σύνθεσης (εξίσωση 3.12). Το αποτέλεσμα της σύνθεσης καταγράφεται σε μια RRD με τη μορφή διαδοχικών χρονικά δειγμάτων $Bel(\theta_l)$, $Pl(\theta_l)$ $(1 \leq l \leq N)$.

5.1.1 Ανάπτυξη εργαλείου δημιουργίας ελεγχόμενων επιθέσεων

Στα πλαίσια των πειραμάτων μας μελετήσαμε τη δυνατότητα δημιουργίας ρεαλιστικών επιθέσεων καταιγισμού πακέτων με τα συνηθέστερα χρησιμοποιούμενα προγράμματα DDoS: Stacheldraht [138], Trin00, TFN, mstream. Τα προγράμματα αυτά χρησιμοποιούνται συνήθως για κακόβουλες επιθέσεις τύπου DDoS. Προκειμένου να μη δημιουργήσουμε ανεξέλεγκτες επιθέσεις σε δίκτυα παραγωγής (Ε.Μ.Π., GRNET) τροποποιήσαμε το πρόγραμμα TFN2K (λόγω ευκολίας επέκτασης του κώδικα C και συμβατότητας με περιβάλλον FreeBSD [139]) ώστε να υποστηρίζει:

- ελεγχόμενο ρυθμό αποστολής πακέτων,
- ελεγχόμενο ρυθμό αποστολής bytes,
- ελεγχόμενη αποστολή παραποιημένων πακέτων (**spoofing**),

- επιλογή χρησιμοποιούμενων Layer 4 ports και
- επιλογές σχετικές με τα checksum των UDP πακέτων.

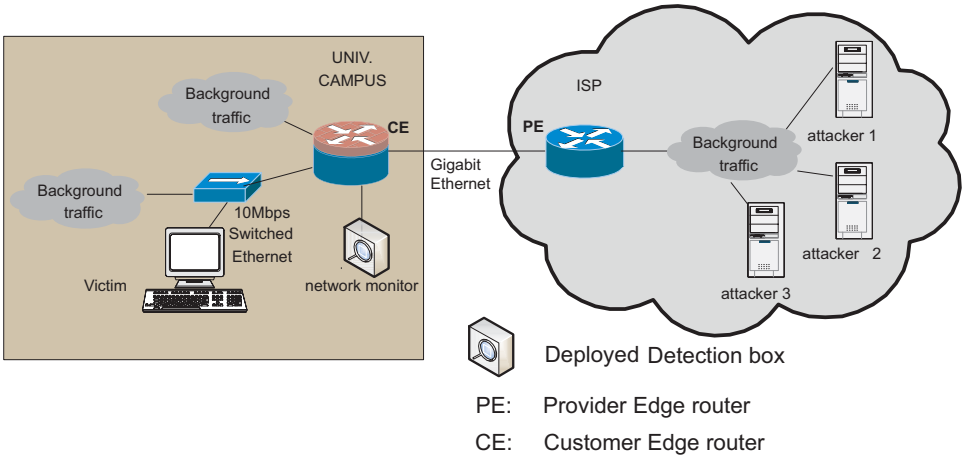
Έτσι το εργαλείο που υλοποιήθηκε δίνει την δυνατότητα εξομοίωσης (emulation) επιθέσεων DDoS με ελεγχόμενα χαρακτηριστικά.

5.2 Πειραματικά αποτελέσματα

5.2.1 Τοπολογία πειραμάτων

Κατά την εκπόνηση της παρούσας διατριβής εκτελέστηκαν πλήθος πειραμάτων στο δίκτυο του Εθνικού Μετσόβιου Πολυτεχνείου (Ε.Μ.Π.) και του Εθνικού Δικτύου Έρευνας και Τεχνολογίας (GRNET) προκειμένου να εντοπιστούν μετρικά που βοηθούν την ανίχνευση ανωμαλιών και να εξεταστούν οι τεχνικές μέτρησης τους, όπως περιγράφηκαν στο κεφάλαιο 4. Στο κεφάλαιο αυτό εστιάζουμε σε ένα υποσύνολο πειραμάτων που έγιναν για την αξιολόγηση της επίδοσης του πρωτότυπου συστήματος ανίχνευσης επιθέσεων καταιγισμού πακέτων που υλοποιήθηκε. Για τα πειράματα αυτά χρησιμοποιήθηκαν οι πρωτότυπες υλοποιήσεις αισθητήρων που αναφέρθηκαν στην προηγούμενη ενότητα. Η αξιολόγηση των αποτελεσμάτων έγινε με τη βοήθεια του προγραμματιστικού περιβάλλοντος Matlab.

Στα πειράματα αυτά (σχήμα 5.1) τρεις υπολογιστές, κατανεμημένοι στο δίκτυο του GRNET, εξαπέλυσαν ένα μεγάλο αλλά ελεγχόμενο πλήθος επιθέσεων DDoS προς ένα θύμα στο εσωτερικό δίκτυο του Ε.Μ.Π. Το θύμα είχε ταχύτητα σύνδεσης 10Mbps ενώ οι τρεις επιτιθέμενοι υπολογιστές είχαν σύνδεση Fast Ethernet στα 100Mbps. Με την επιλογή αυτή προσομοιάζουμε την συγκέντρωση κίνησης από περισσότερους επιτιθέμενους. Η γραμμή διασύνδεσης του πελάτη (Ε.Μ.Π., CE) με τον πάροχο (GRNET, PE) είναι ταχύτητας 1Gbps. Κατά την διάρκεια διεξαγωγής των πειραμάτων πραγματοποιούσαμε συνεχείς παθητικές μετρήσεις στη γραμμή διασύνδεσης πελάτη-πάροχου (CE-PE). Στις μετρήσεις μας χρησιμοποιήσαμε Netflow (not sampled) και Packet



Σχήμα 5.1: Τοπολογία πειραμάτων.

Capturing από το δρομολογητή του Ε.Μ.Π. Οι συγκεκριμένες μετρήσεις θα ήταν ίσως ευκολότερο να πραγματοποιηθούν με τη χρήση του SNMP αλλά αυτό δεν ήταν δυνατό γιατί τα μετρικά μας δεν υποστηρίζονταν από τις MIB των δρομολογητών.

Η χρησιμοποίηση της γραμμής διασύνδεσης ήταν χαμηλή με μέση τιμή δικτυακής κίνησης 250Mbps και με απότομες αυξήσεις που έφταναν τα 450Mbps. Η περίπτωση αυτή όπου υπάρχει διαθέσιμο εύρος ζώνης (bandwidth) στη γραμμή διασύνδεσης πελάτη-πάροχου είναι ιδανική για την ανίχνευση επιθέσεων DDoS καθώς υπάρχει δυνατότητα τοπικής αντιμετώπισης του προβλήματος. Η συγκεκριμένη περίπτωση αποτελεί επίσης χαρακτηριστικό παράδειγμα της παρακολούθησης ενός δικτύου υψηλών ταχυτήτων. Από τη γραμμή διασύνδεσης μεταφέρεται δικτυακή κίνηση που αντιστοιχεί σε ένα πλούσιο μείγμα εφαρμογών όπως τυπικές δικτυακές υπηρεσίες Web, DNS αλλά και προγράμματα peer-to-peer file sharing, online παιχνίδια και streaming audio (πίνακας 5.1). Το γεγονός αυτό είναι ιδιαίτερα σημαντικό καθώς τα προτεινόμενα μετρικά και το σύστημα ανίχνευσης που υλοποιήθηκε, αξιολογήθηκε σε **πραγματικές συνθήκες** και όχι σε ένα απλοποιημένο εργαστηριακό περιβάλλον ή με τη χρήση προσομοιωτών. Λόγω των ιδιαιτεροτήτων της πραγματικής δικτυακής κίνησης είναι πιθανό κάποιες θεωρητικά επιτυχημένες προσεγγίσεις

Πίνακας 5.1: Μερική ανάλυση δείγματος κίνησης (διάρκειας 30 λεπτών) της παρακολουθούμενης δικτυακής ζεύξης

Πρωτόκολλο	Packet rate (pps)	Bit rate (Mbps)
tcp	37184.8(92.55%)	204.47(94.32%)
ftp	1155.3 (2.87%)	8.19 (3.78%)
smtp	168.6 (0.42%)	0.94 (0.43%)
http/s	4011.6 (9.98%)	21.19 (9.78%)
nntp	362.9 (0.90%)	1.76 (0.81%)
p2p	7536,8 (18,75%)	41.53 (19,16%)
other	23911.9 (59.54%)	130.60(60.25%)
udp	2854.1 (7.10%)	12.23 (5.64%)
dns	180.9 (0.45%)	0.19 (0.09%)
realaud	1312.1 (3.27%)	9.93 (4.58%)
other	1,361.10 (3.38%)	2.11 (0.97%)
icmp	111.2 (0.28%)	0.075 (0.03%)
Avg: 216.80Mbps	Stddev:6.53M	Peak: 237.13Mbps

Πίνακας 5.2: Χαρακτηριστικά των επιθέσεων.

Τύπος Επίθεσης	Πλήθος επιθέσεων	Bit rate (Mbps)	Packet rate (pps)
UDP	49	1-6	1000-6000
ICMP	21	1-6	1000-6000
SYN	18	1-6	1000-6000

να καθίστανται μη-εφαρμόσιμες στην πράξη λόγω του υψηλού ρυθμού από false positives που παράγουν, όντας ευαίσθητες σε ορισμένες παρατηρούμενες μορφές δικτυακής κίνησης (traffic patterns).

Στο πίνακα 5.2 φαίνονται τα χαρακτηριστικά των επιθέσεων που θα αναλύσουμε και συμπεριλαμβάνουν 88 διαφορετικές επιθέσεις τύπου SYN, UDP και ICMP. Στα πειράματα αυτά χρησιμοποιήθηκε η τεχνική αποστολής παραποιημένων IP πακέτων (spoofing) και η διεύθυνση αποστολέα επιλέγονταν τυχαία από το πραγματικό υποδίκτυο του επιτιθέμενου προκειμένου να ξεπερνά προληπτικά μέτρα όπως Egress filtering και Reverse Path Filtering (RPF). Κατά την διάρκεια των πειραμάτων έγινε καταγραφή μετρικών (συλλογή δεδομένων) με περίοδο δειγματοληψίας 30sec και συλλέχθηκαν 3863 δείγματα που αντιστοιχούν σε χρονική περίοδο 32 ωρών. Το μέγεθος των επιθέσεων είναι αρκετά μικρό σε σχέση με την συνολική κίνηση που μεταφέρεται από την

γραμμή διασύνδεσης Ε.Μ.Π. - GRNET. Η επιλογή αυτή είναι σκόπιμη, καθώς αν ανιχνεύονται ανωμαλίες μικρού μεγέθους το σύστημα μπορεί να εντοπίσει εύκολα ανωμαλίες μεγαλύτερου μεγέθους.

5.2.2 Αποτελέσματα ανίχνευσης επιθέσεων

Η διαδικασία που ακολουθήσαμε για την αξιολόγηση της ικανότητας ανίχνευσης του πρωτότυπου συστήματος ανίχνευσης ανωμαλιών στη δικτυακή κίνηση (NADS) είναι:

1. Για κάθε επίθεση καταιγισμού πακέτων (πείραμα) καταγράψαμε τις χρονικές στιγμές εκκίνησης και λήξης.

2. Συλλέξαμε τα δεδομένα ανίχνευσης (δείγματα) από τους αισθητήρες του συστήματος για χρονικά διαστήματα που περιείχαν τις επιθέσεις που εκτελέσαμε. Τα μετρικά που χρησιμοποιήθηκαν (ενότητα 4.3) είναι: UDP Ratio (UR), ICMP Ratio (IR), SYN-FIN Ratio (SFR), Flow Rate (FR), Flows with duration $< 100ms$ ($FT10$). Θεωρήσαμε ότι τα δείγματα που συλλέξαμε δεν περιέχουν επιθέσεις πέρα από αυτές που εξομοιώσαμε εμείς.

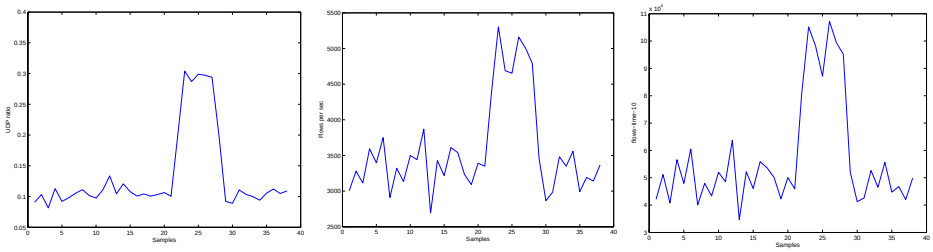
3. Βάσει των χρόνων εκκίνησης και λήξης των επιθέσεων αντιστοιχήσαμε τα δείγματα σε φυσιολογικές καταστάσεις και καταστάσεις επίθεσης SYN, UDP, ICMP. Με τον τρόπο αυτό δημιουργήθηκε ο ορθός πίνακας ταξινόμησης (S_{opt}) των δειγμάτων.

4. Αποθηκεύουμε το αποτέλεσμα της αντιστοίχησης των δειγμάτων σε φυσιολογικές καταστάσεις και καταστάσεις επίθεσης SYN, UDP και ICMP από το σύστημα ανίχνευσης στον πίνακα ταξινόμησης (S_{det}). Η ευαισθησία του συστήματος ανίχνευσης ελέγχεται από τις συναρτήσεις κατωφλίου στο επίπεδο των αισθητήρων-ανιχνευτών αλλά και στο τελικό στάδιο της ταξινόμησης ενός δείγματος μετά τη χρήση της θεωρίας Dempster-Shafer για τη σύνθεση δεδομένων.

5. Συγκρίναμε τα αποτελέσματα του συστήματος ανίχνευσης (S_{det}) με τα ορθά αποτελέσματα (S_{opt}) και μετρήσαμε τον αριθμό των false positives και των true positives.

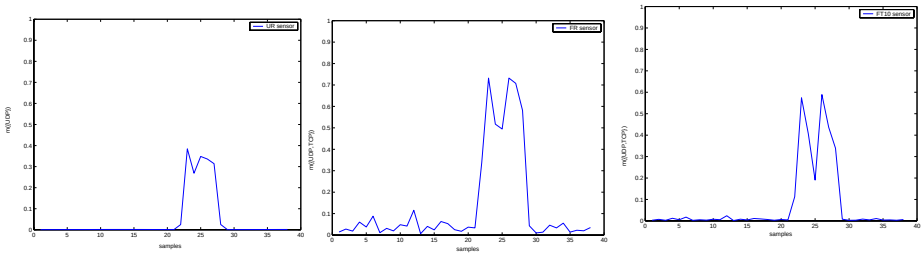
Θα παρουσιάσουμε ενδεικτικά τα στάδια επεξεργασίας κατά την ανίχνευση μιας επίθεσης τύπου UDP ρυθμισμένης έντασης 2Mbps στη γραμμή διασύνδεσης PE-CE ταχύτητας 1Gbps. Στο πείραμα αυτό βλέπουμε ότι ακόμη και αν ένας αισθητήρας ανιχνεύσει μερικώς την ύπαρξη μιας επίθεσης, η συνδυασμένη γνώση από όλους τους αισθητήρες μας οδηγεί στην πεποίθηση ότι πραγματοποιείται μια επίθεση. Στην πραγματικότητα η συγκεκριμένη επίθεση δεν επιτυγχάνει την άρνηση υπηρεσιών αλλά αναδεικνύει την ικανότητα ανίχνευσης του συστήματος. Όπως φαίνεται και στα σχήματα 5.2(a), 5.3(b) το μετρικό UR δεν υποδεικνύει καθαρά την ύπαρξη μιας επίθεσης (0.35 Prob) καθώς το bandwidth της επίθεσης είναι ένα μικρό ποσοστό της συνολικής UDP κίνησης. Παρόλα αυτά με την χρήση των μετρικών FR και $FT10$ μπορούμε να ανιχνεύσουμε επιτυχώς την επίθεση (σχήματα 5.2(b), 5.2(c)). Τα μετρικά αυτά μεταφράζονται σε bpa's που απεικονίζονται στα σχήματα 5.3(a), 5.3(c) και 5.3(b). Η σύνθεση των δεδομένων των τριών στοιχειωδών ανιχνευτών με χρήση του κανόνα του Dempster μας δίνει το συνολικό συμπέρασμα για την κατάσταση της παρακολουθούμενης δικτυακής ζεύξης. Παρατηρώντας το διάστημα εμπιστοσύνης $Bel(\{UDP\})$,$Pl(\{UDP\})$ για την ύπαρξη επίθεσης UDP βλέπουμε ότι η επίθεση ανιχνεύεται επιτυχώς (σχήμα 5.4).

Η απόδοση του συστήματος ανίχνευσης για όλα τα πειράματα αποτυπώνεται στο σχήμα 5.5, με τη Receiver Operating Characteristic (ROC) curve. Το σύστημα εμφανίζει υψηλά ποσοστά από true positives (80%) με false positives κάτω του 3%. Για την αξιολόγηση εσκεμμένα δεν χρησιμοποιούμε το ποσοστό ανίχνευσης (detection ratio) ως κριτήριο, καθώς εμπεριέχει την ανίχνευση της φυσιολογικής κατάστασης ($NORMAL$) που είναι και η συνηθέστερη. Για το λόγο αυτό, το detection ratio συχνά κυμαίνεται άνω του 90% σε οποιαδήποτε περίπτωση και μας δίνει μια παραπλανητική εικόνα για



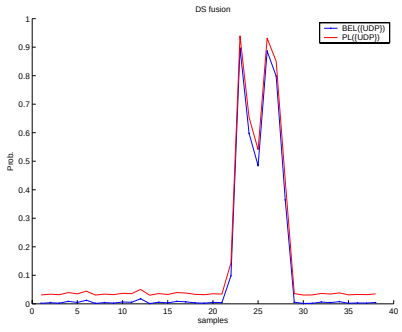
(a) μετρικό UR (b) μετρικό FR (c) μετρικό FT10

Σχήμα 5.2: Μετρικά κατά την διάρκεια μιας επίθεσης UDP.

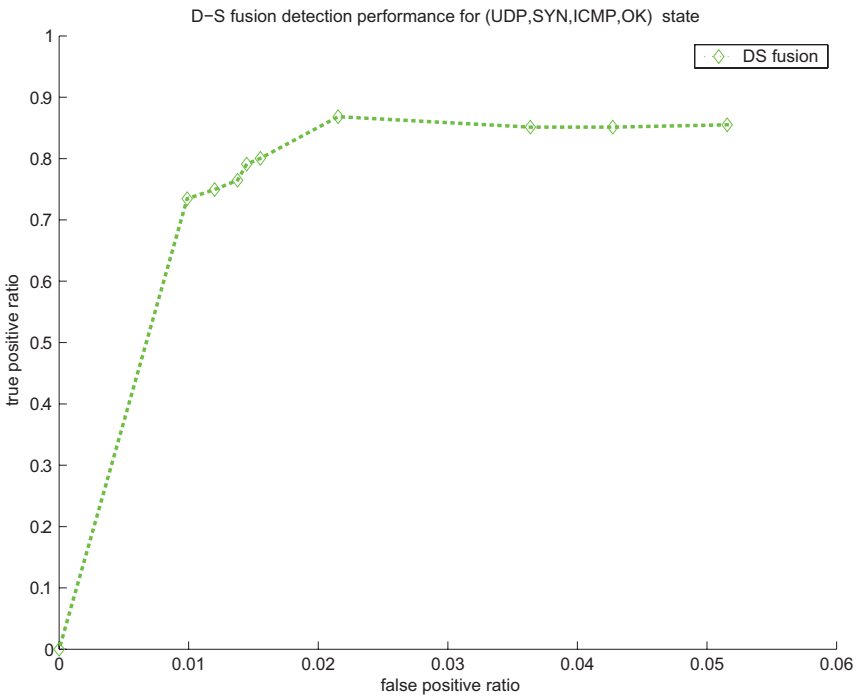


(a) UR -μερική ανίχνευση (b) FR -καλή ανίχνευση (c) FT10 -μερική ανίχνευση

Σχήμα 5.3: Τα basic probability assignments που αντιστοιχούν στα μετρικά 5.2(a), 5.2(b),5.2(c)



Σχήμα 5.4: Διάστημα εμπιστοσύνης για την κατάσταση επίθεσης UDP μετά τη σύνθεση δεδομένων από τα bpa's των σχημάτων 5.3(a), 5.3(b), 5.3(c).



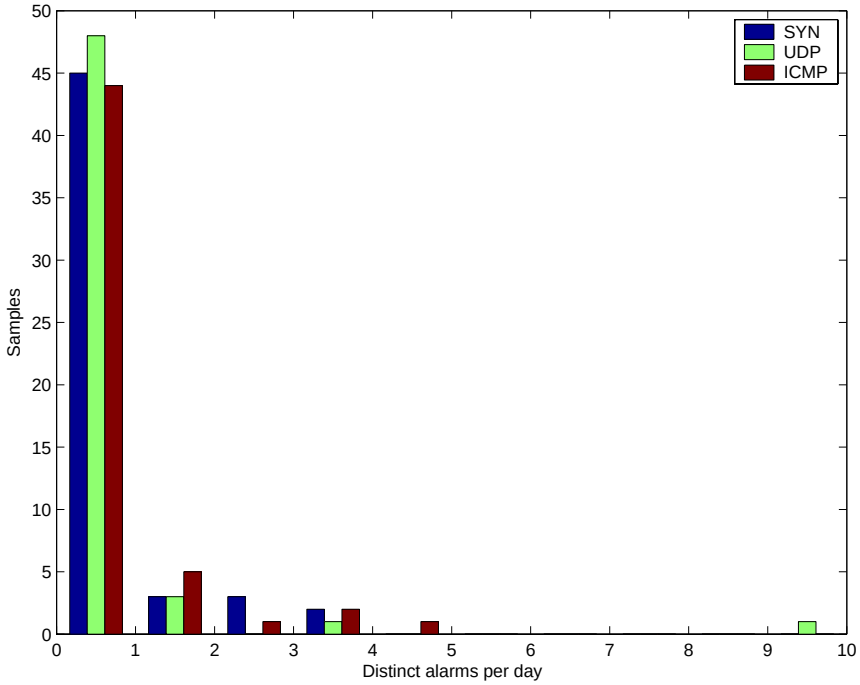
Σχήμα 5.5: Η απόδοση του συστήματος ανίχνευσης με χρήση της θεωρίας D-S (ανίχνευση όλων των καταστάσεων επίθεσης).

την ικανότητα ανίχνευσης.

Πέρα όμως από την αξιολόγηση ενός συστήματος ανίχνευσης με τα true positive και false positive ratios σημαντικός είναι επίσης ο ρυθμός των παραγόμενων alerts (alerts = true positives + false positives) στην μονάδα του χρόνου. Ο αριθμός των true positives και false positives έχει εξεταστεί στην βιβλιογραφία ως απλό ποσοστό. Στην περίπτωση όμως της ανίχνευσης ανωμαλιών στη δικτυακή κίνηση και εφόσον ακόμη δεν χρησιμοποιούμε αυτοματοποιημένες μεθόδους καταστολής, κάθε παραγόμενο alert πρέπει να αξιολογηθεί από έναν διαχειριστή δικτύου. Συνεπώς μέχρι να αναπτυχθούν αυτοματοποιημένα συστήματα καταστολής, ένα σύστημα ανίχνευσης για να είναι πρακτικά χρήσιμο πρέπει να παράγει ένα μικρό αριθμό alert ανά μέρα.

Για να μπορέσουμε να εκτιμήσουμε τον όγκο των παραγόμενων μηνυμάτων συναγερμού στη μονάδα του χρόνου αναλύσαμε την λειτουργία του πρωτότυπου συστήματος ανίχνευσης για 53 συνεχείς ημέρες χωρίς να εξομοιώσουμε επιθέσεις (emulated attacks). Το σύστημα επεκτάθηκε με έναν μηχανισμό καταγραφής λεπτομερών στοιχείων δικτυακής κίνησης (Packet Capturing) για το χρονικό διάστημα ύπαρξης συναγερμού. Τα στοιχεία αυτά αξιολογήθηκαν εμπειρικά ώστε να εκτιμηθεί ο αριθμός των false positives σε πραγματικές συνθήκες για μεγάλα χρονικά διαστήματα[2]. Για το χρονικό διάστημα αυτό οι ρυθμίσεις των αισθητήρων παρέμειναν σταθερές και στο κόμβο σύνθεσης θέσαμε ως κατώφλι για την παραγωγή μηνύματος συναγερμού $T = 0.7$ (εξίσωση 4.1). Τα αποτελέσματα ήταν ιδιαίτερα ενθαρρυντικά καθώς ο μέσος αριθμός από διαφορετικά μηνύματα συναγερμού ανά ημέρα ήταν πολύ μικρός. Η κατανομή του μεγέθους αυτού απεικονίζεται στο σχήμα 5.6. Η ανάλυση των λεπτομερών στοιχείων δικτυακής κίνησης έδειξε ότι σε όλες τις περιπτώσεις υπήρχε μια πραγματική ανωμαλία της δικτυακής κίνησης που ενεργοποίησε το σύστημα ανίχνευσης. Πέρα από προφανείς περιπτώσεις μαζικής αποστολής πακέτων από εφαρμογές file-sharing applications, όλες οι περιπτώσεις

[2]Σημειώνουμε ότι την περίοδο εκείνη (2003-2004) το Ε.Μ.Π. δεν ήταν συχνός στόχος επιθέσεων (1-2 το μήνα)



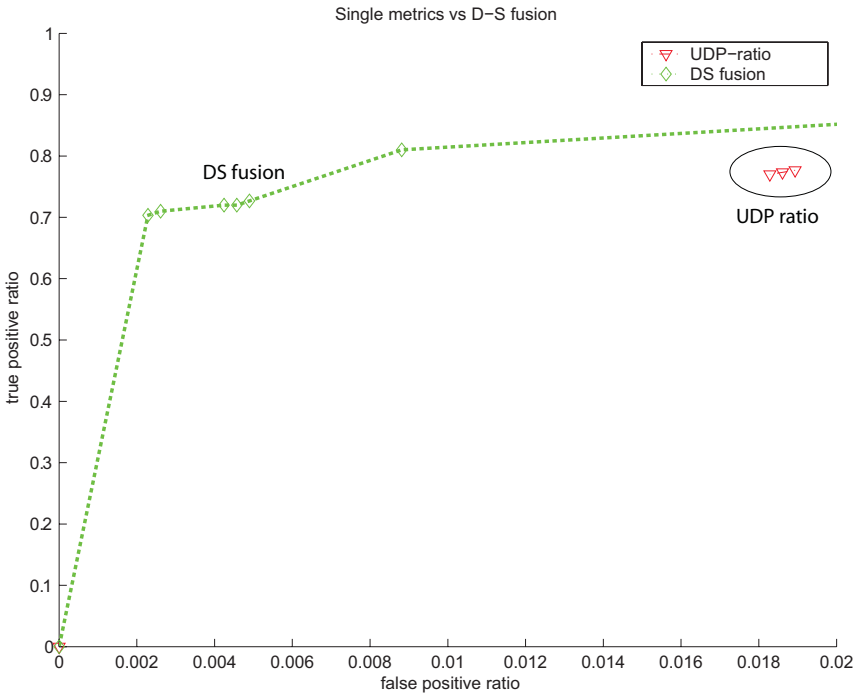
Σχήμα 5.6: Η κατανομή του μέσου αριθμού διαφορετικών μηνυμάτων συναγερμού ανά ημέρα κατά την διάρκεια συνεχούς λειτουργίας του πρωτότυπου συστήματος NADS.

συναγερμού ήταν ιδιαίτερα ύποπτες. Αν και δεν μπορούμε να ισχυριστούμε με σιγουριά ότι ήταν όλες επιθέσεις DoS θα μπορούσαν σίγουρα να φιλτραριστούν ή να ελεγχθεί η ροή τους (rate limiting - ενότητα 2.5.2) χωρίς ανησυχία για τις επιπτώσεις. Το γεγονός αυτό είναι ενθαρρυντικό για τη δυνατότητα αυτόματης εφαρμογής μεθόδων καταστολής.

5.3 Σύγκριση με άλλες προσεγγίσεις

5.3.1 Σύγκριση με απλούς ανιχνευτές κατωφλίου

Η σύνθεση των δεδομένων ανίχνευσης από πολλούς αλληλοσυμπληρούμενους αισθητήρες-ανιχνευτές μπορεί θεωρητικά να βελτιώσει την απόδοση ενός συστήματος ανίχνευσης. Το γεγονός αυτό επιβεβαιώθηκε μέσα από τα πειράματα μας για την αξιολόγηση του πρωτότυπου συστήματος NADS. Συ-



Σχήμα 5.7: Σύγκριση απόδοσης απλού ανιχνευτή ενός μετρικού με σύστημα σύνθεσης δεδομένων (D-S) για επιθέσεις UDP.

γκρίναμε την απόδοση ανίχνευσης του πρωτότυπου συστήματος ανίχνευσης (σύνθεση με χρήση θεωρίας D-S) με έναν απλό ανιχνευτή ενός μετρικού και συγκεκριμένα μια σιγμοειδή συνάρτηση κατωφλίου με είσοδο ένα μετρικό. Η σύγκριση αυτή είναι εξ αρχής σε ένα βαθμό άνιση καθώς το πρωτότυπο σύστημα διακρίνει και ενδεχομένως συγχέει 3 διαφορετικούς τύπους επιθέσεων ενώ οι απλοί ανιχνευτές ενός μετρικού που θα χρησιμοποιήσουμε μπορούν να ανιχνεύσουν ένα μονάχα τύπο επίθεσης. Δηλαδή στην πρώτη περίπτωση έχουμε ταξινόμηση ενός δείγματος σε 4 πιθανές καταστάσεις {NORMAL, SYN, UDP, ICMP} ενώ στην δεύτερη ταξινόμηση σε 2 πιθανές καταστάσεις π.χ. {NORMAL, UDP}.

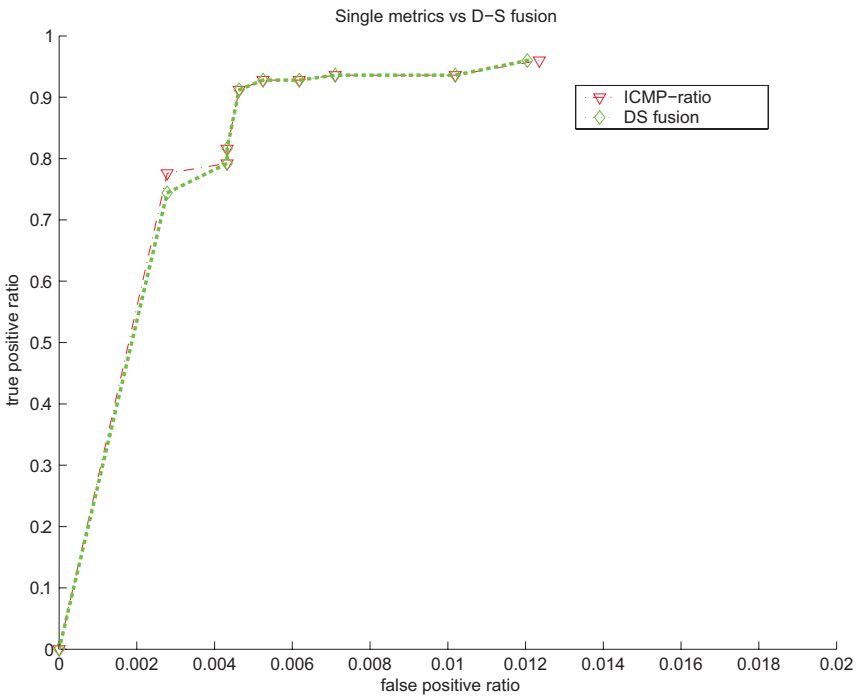
Για την ανίχνευση επιθέσεων UDP χρησιμοποιήθηκε ένας ανιχνευτής κατωφλίου βασισμένος στο μετρικό UR. Η απόδοση του ανιχνευτή στην ταξινόμηση των δειγμάτων ως επιθέσεις UDP και μη φαίνεται στο σχήμα 5.7.

Παρατηρούμε μια σημαντική υπεροχή του πρωτότυπου συστήματος ανίχνευσης με σύνθεση δεδομένων καθώς η ανίχνευση επιθέσεων με μικρό μέγεθος διευκολύνεται από τα flow based μετρικά FR και $FT10$. Ο απλός ανιχνευτής ενός μετρικού μεταβαίνει απότομα από την κατάσταση πλήρους αδυναμίας ανίχνευσης σε δυνατότητα ανίχνευσης 75% αλλά και με false positive ratio 2%. Αντίθετα το σύστημα σύνθεσης δεδομένων D-S παρουσιάζει σταδιακή αύξηση ευαισθησίας με σημαντικά χαμηλότερα false positive ratios.

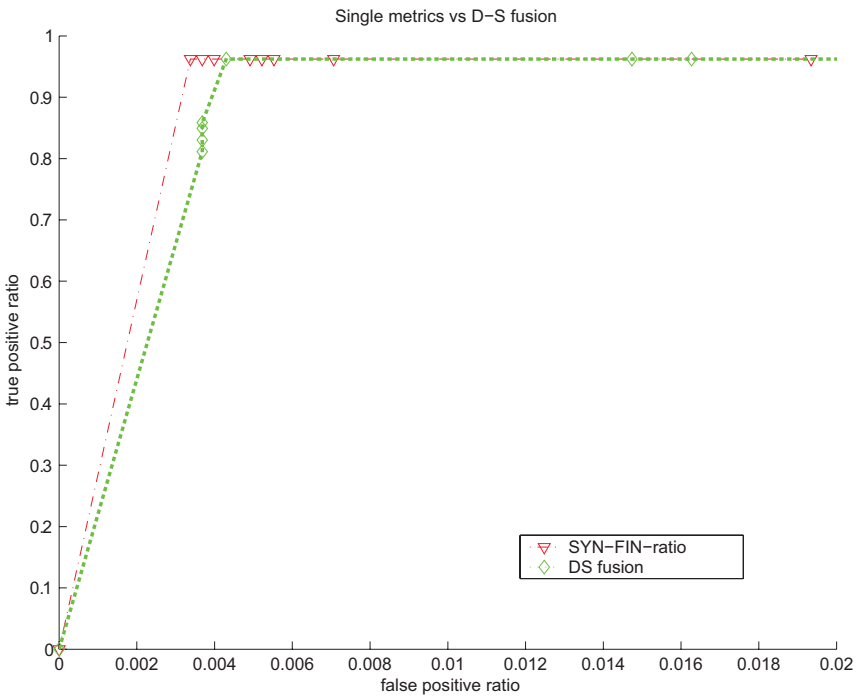
Για την ανίχνευση επιθέσεων ICMP και SYN χρησιμοποιήθηκε ένας ανιχνευτής κατωφλίου βασισμένος στο μετρικό IR και SFR αντίστοιχα. Η απόδοση των δύο προσεγγίσεων είναι παραπλήσια αφού τα μετρικά IR, SFR είναι πολύ αποτελεσματικά και επαρκούν για την ανίχνευση όλων των επιθέσεων (σχήματα 5.8, 5.9). Στις περιπτώσεις όπου ένα μετρικό είναι αρκετό για την ανίχνευση ενός τύπου επίθεσης, το πρωτότυπο σύστημα ανίχνευσης με σύνθεση δεδομένων παρουσιάζει ισοδύναμη δυνατότητα ανίχνευσης με έναν ανιχνευτή κατωφλίου. Η δυνατότητα βελτίωσης της απόδοσης ανίχνευσης ενός συστήματος σύνθεσης δεδομένων είναι δύσκολο να ποσοτικοποιηθεί γιατί εξαρτάται από τα πειραματικά δεδομένα που χρησιμοποιούμε. Υπάρχουν πάντως αποδείξεις για σαφή βελτίωση της απόδοσης ανίχνευσης στις περιπτώσεις όπου η ανίχνευση ανωμαλιών είναι δύσκολη και δεν μπορεί στηριχτεί σε ένα μετρικό.

5.3.2 Σύγκριση με χρήση νευρωνικού δικτύου επιβλεπόμενης μάθησης για σύνθεση δεδομένων

Η σύγκριση της απόδοσης του πρωτότυπου συστήματος ανίχνευσης με ένα σύστημα που βασίζεται σε ένα νευρωνικό δίκτυο (Artificial Neural Network -ANN) για τη σύνθεση δεδομένων έχει ιδιαίτερο ενδιαφέρον. Τα δύο συστήματα κατατάσσουν τα δείγματα εισόδου στις τρεις καταστάσεις επίθεσης και την φυσιολογική: {NORMAL, SYN, UDP, ICMP}. Το πρωτότυπο σύστημα που υλοποιήσαμε βασίζεται στη θεωρία D-S για την σύνθεση δεδομένων από



Σχήμα 5.8: Σύγκριση απόδοσης απλού ανιχνευτή ενός μετρικού με σύστημα σύνθεσης δεδομένων (D-S) για επιθέσεις ICMP.



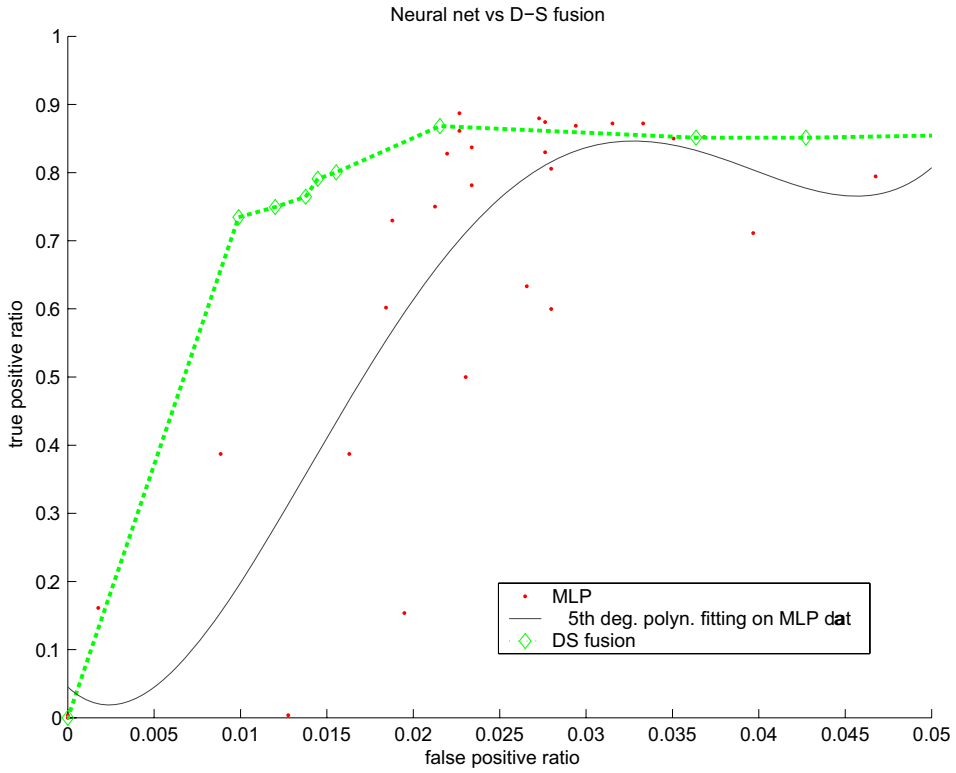
Σχήμα 5.9: Σύγκριση απόδοσης απλού ανιχνευτή ενός μετρικού με σύστημα σύνθεσης δεδομένων (D-S) για επιθέσεις SYN.

πολλούς αισθητήρες ενώ το δεύτερο χρησιμοποιεί ένα απλό νευρωνικό δίκτυο: ένα Multi-Layer Perceptron - MLP (ενότητα 4.4). Το MLP που χρησιμοποιήσαμε είχε τρία στρώματα: το στρώμα εισόδου, ένα κρυμμένο στρώμα και το στρώμα εξόδου. Τα μετρικά που χρησιμοποιήθηκαν (ενότητα 4.3) είναι: UR, IR, SFR, FR, $FT10$. Το στρώμα εισόδου είχε πέντε γραμμικούς νευρώνες και το διάνυσμα εισόδου $\boldsymbol{i}$ ($|\boldsymbol{i}| = M = 5$) είναι :

$$\boldsymbol{i}(t) = [UR(t)\ IR(t)\ SFR(t)\ FR(t)\ FT10(t)]$$

Ο αλγόριθμος εκπαίδευσης που χρησιμοποιήθηκε για την εκπαίδευση του ANN ήταν ο Levenberg-Marquardt [140] λόγω της δυνατότητας του για γρήγορη εκπαίδευση. Κατά την διάρκεια της εκπαίδευσης του νευρωνικού δικτύου, η οποία είχε μέγιστη διάρκεια 400 επαναλήψεις, χρησιμοποιήσαμε ως training set ένα μικρό μέρος των δειγμάτων από τα πειράματα μας (7%).

Τα αποτελέσματα της δυνατότητας ανίχνευσης του ANN φαίνονται στο σχήμα 5.10 όπου απεικονίζεται το αντίστοιχο ROC curve. Καθώς τα σημεία στο επίπεδο (true positive, false positive) είναι διασκορπισμένα, χρησιμοποιήσαμε για λόγους παρουσίασης data fitting με ένα πολυώνυμο πέμπτου βαθμού. Τα αποτελέσματα δείχνουν ότι η επίδοση του συστήματος ανίχνευσης με σύνθεση δεδομένων D-S και του συστήματος με χρήση MLP είναι εφάμιλλη. Η προσέγγιση D-S παράγει λίγο λιγότερα false positives αλλά το σημαντικότερο στοιχείο είναι πως η προσέγγιση αυτή υπερτερεί σε απλότητα. Ένα ANN πρέπει να εκπαιδεύεται ώστε να ρυθμίζονται τα βάρη στους συνδέσμους των νευρώνων καθώς η ευφυΐα του αποθηκεύεται σε αυτά. Επίσης η εκπαίδευση ενός ANN είναι ευαίσθητη στις αρχικές τυχαίες συνθήκες των βαρών του.



Σχήμα 5.10: Σύγκριση απόδοσης συστήματος ανίχνευσης με θεωρία D-S και με ANN

Κεφάλαιο 6

Συμπεράσματα - Ανοικτά θέματα

6.1 Συμπεράσματα

Μέσα από την ανάλυση του προβλήματος των κατανεμημένων επιθέσεων άρνησης υπηρεσιών και την έρευνα για την ανάπτυξη της προτεινόμενης αρχιτεκτονικής συστημάτων ανίχνευσης ανωμαλιών στη δικτυακή κίνηση (NADS) μπορούν να εξαχθούν σημαντικά συμπεράσματα, τα κυριότερα εκ των οποίων θα συνοψίσουμε στην ενότητα αυτή.

Το πρώτο σημείο αφορά την προσέγγιση που ακολουθήσαμε για την ανίχνευση και αντιμετώπιση των ανωμαλιών στη δικτυακή κίνηση και κυρίως των επιθέσεων καταιγισμού πακέτων (DDoS). Ο συγκεκριμένος τύπος ανωμαλίας της δικτυακής κίνησης παρουσιάζει το εξής ιδιαίτερο χαρακτηριστικό: τα μέτρα για την αντιμετώπιση του πρέπει να εφαρμοστούν σε σημεία όπου δεν έχουν εξαντληθεί οι δικτυακοί πόροι που αποτελούν στόχο της επίθεσης. Στα σημεία αυτά η διάγνωση του προβλήματος είναι συνήθως δυσκολότερη από ότι στην περιοχή του θύματος της επίθεσης (ενότητα 2.3). Η παρούσα διατριβή αποδεικνύει μέσα από πειράματα[1] ότι η λύση που προτείνει, δηλαδή η **ανίχνευση στους ακραίους δρομολογητές του δικτύου ενός παρόχου (Provider Edge routers)**, είναι εφικτή και αποτελεσματική. Στα σημεία

[1] όπου ανιχνεύονται επιθέσεις στη δικτυακή ζεύξη 1Gbps μεταξύ GRNET και Ε.Μ.Π.

αυτά, όπου υπάρχει συνήθως διαθέσιμο bandwidth και υπολογιστική ισχύς στους δρομολογητές του παρόχου[2], τα μέτρα καταστολής (ενότητα 2.5.2) είναι αποτελεσματικά. Ακολουθώντας την προσέγγιση αυτή, ξεπερνάμε το πρόβλημα συνεργασίας μεταξύ διαχειριστικών ομάδων που θα αντιμετωπίζαμε στην περίπτωση όπου η ανίχνευση μιας επίθεσης γινόταν από το δίκτυο του θύματος και η καταστολή της από το δίκτυο ενός παρόχου ή των πηγών της επίθεσης. Οι πάροχοι δικτύου είναι πιθανό να εγκαταστήσουν και να λειτουργήσουν συστήματα ανίχνευσης ανωμαλιών στη δικτυακή κίνηση (NADS) καθώς έχουν άμεσο ενδιαφέρον, τεχνογνωσία, οικονομική δύναμη και υποδομή σε υλικό αλλά και ανθρώπινο δυναμικό. Άλλωστε μια τέτοια επένδυση είναι πιο πιθανό να γίνει από δίκτυα παρόχους και όχι από τα δίκτυα πελάτες λόγω οικονομίας κλίμακας. Η ανίχνευση και αντιμετώπιση ανωμαλιών στη δικτυακή κίνηση (ιδιαίτερα των κατανεμημένων επιθέσεων άρνησης υπηρεσιών) μπορεί να αποτελέσει μια κερδοφόρα υπηρεσία ενός πάροχου δικτύου προς τους πελάτες του.

Το δεύτερο σημαντικό στοιχείο που πρέπει να σημειωθεί είναι ότι οι ανωμαλίες στη δικτυακή κίνηση αποτελούν συνεχώς εξελισσόμενα φαινόμενα. Αναμένουμε λοιπόν την εξέλιξη των χαρακτηριστικών τους όπως εξελίσσονται άλλωστε και τα χαρακτηριστικά της φυσιολογικής κίνησης στο Διαδίκτυο. Για το λόγο αυτό η προτεινόμενη αρχιτεκτονική επιβάλλεται να είναι ευέλικτη και ανοικτή σε μελλοντικές επεκτάσεις. Η αρχιτεκτονική για την ανάπτυξη συστημάτων ανίχνευσης που προτείναμε είναι:

- **ακριβής στην ανίχνευση ανωμαλιών** καθώς χρησιμοποιείται σύνθεση δεδομένων και **αποδοτική στην αποθήκευση** των δεδομένων από πολλούς αισθητήρες. Συγκεκριμένα η χρήση Round Robin Databases (RRD) διευκολύνει και καθιστά ευέλικτη και αποδοτική την αποθήκευση αριθμητικών δεδομένων - χρονοσειρών (timeseries) για μεγάλα χρονικά διαστήματα.

[2]οι μεγάλοι ISPs έχουν παραδοσιακά overprovisioned δίκτυα κορμού (backbone).

- **αρθρωτή** (modular) καθώς προβλέπεται εύκολος σχεδιασμός, υλοποίηση και χρήση νέων αισθητήρων. Τα βασικά στάδια επεξεργασίας ενός αισθητήρα υλοποιούνται σε αυτόνομα τμήματα λογισμικού που μπορούν να επαναχρησιμοποιούνται και να εναλλάσσονται καθώς η είσοδος και η έξοδος τους είναι RRD. Με τον τρόπο αυτό διευκολύνεται η χρήση νέων μετρικών και αλγόριθμων ανίχνευσης. Ένα ακόμη παράδειγμα της αρθρωτής αρχιτεκτονικής είναι ο διαχωρισμός της φάσης ανίχνευσης μιας επίθεσης από την φάση της ταυτοποίησης της που αποσκοπεί στην ευκολότερη επίλυση των επιμέρους προβλημάτων (ενότητα 2.5.1). Η ανίχνευση μιας επίθεσης συνίσταται στην εξαγωγή ενός απλού συμπεράσματος ύπαρξης ή μη κάποιας ανωμαλίας (detection). Η ταυτοποίηση μιας επίθεσης αποτελεί ξεχωριστή διαδικασία με στόχο να προσδιορίσει τα χαρακτηριστικά μιας επίθεσης, π.χ. ο προσδιορισμός των IP διευθύνσεων των πηγών μιας επίθεσης, ώστε να μπορούν στην συνέχεια να εφαρμοστούν αντίμετρα.

- **επεκτάσιμη** καθώς προβλέπει την προσθήκη και εξέλιξη των τεχνικών παρακολούθησης δικτύου, των μετρικών, των αλγόριθμων ανίχνευσης, αλγόριθμων σύνθεσης και αλγόριθμων ταυτοποίησης των χαρακτηριστικών μιας ανωμαλίας. Συνεπώς η προτεινόμενη αρχιτεκτονική μπορεί να επεκταθεί με κατάλληλη προσαρμογή των αισθητήρων και του αλγόριθμου σύνθεσης ώστε να ανιχνεύει πολλά είδη ανωμαλιών στη δικτυακή κίνηση όπως reconnaissance attempts (port-scans, ping sweeps) και εξάπλωση worms. Χαρακτηριστικό παράδειγμα επέκτασης είναι η χρήση δικτυακών τηλεσκοπίων [136] ως αισθητήρες που μας τροφοδοτούν με δεδομένα για την ανίχνευση worms και scans. Ένα δικτυακό τηλεσκόπιο (network telescope) παρακολουθεί δικτυακή κίνηση από ή προς ένα τμήμα IP διευθύνσεων που έχουν ανατεθεί σε κάποιο δίκτυο από ένα Registration Authority και δρομολογούνται κανονικά στο Διαδίκτυο αλλά δεν έχουν αποδοθεί προς χρήση από το δίκτυο αυτό, παραμένουν

δηλαδή αχρησιμοποίητες. Η κίνηση που καταγράφεται από το τηλεσκόπιο είναι εξ ορισμού "ανεπιθύμητη" και η παρακολούθηση της αποκαλύπτει διάφορα περιστατικά όπως η εξάπλωση worms, reconnaissance attempts και απομακρυσμένες επιθέσεις DDoS με spoofed πακέτα λόγω του φαινομένου backscatter [25, 141–143].

Με την ανάπτυξη του πρωτότυπου συστήματος ανίχνευσης επιθέσεων DDoS αποδείξαμε πως η προτεινόμενη αρχιτεκτονική είναι **υλοποιήσιμη** με χρήση υπαρχουσών τεχνολογιών. Επίσης προτείναμε **αποτελεσματικά μετρικά** που μπορούν να μετρηθούν σε δίκτυα υψηλών ταχυτήτων με υπάρχουσες τεχνικές παρακολούθησης δικτύων. Μάλιστα τα μετρικά που βασίζονται στην συμμετρική φύση της δικτυακής κίνησης, π.χ. η ισορροπία στα εξερχόμενα και εισερχόμενα πακέτα ICMP ή TCP με SYN και FIN flags, περιμένουμε να παραμείνουν και μελλοντικά αποτελεσματικά στην ανίχνευση ανωμαλιών καθώς οι επιτιθέμενοι είναι δύσκολο να ελέγξουν τη συμμετρία αυτή.

Κεντρική επιλογή της παρούσας διατριβής είναι **η σύνθεση δεδομένων** από πολλούς ανιχνευτές βασισμένους σε διαφορετικές πηγές δεδομένων και διαφορετικούς αλγόριθμους ανίχνευσης (multisensor data fusion) [31,32]. Τα πλεονεκτήματα της σύνθεσης δεδομένων με χρήση της θεωρίας Dempster Shafer (D-S) συνοψίζονται στα εξής σημεία:

- Η σύνθεση δεδομένων είναι δυνατή χωρίς να απαιτείται λεπτομερής μοντελοποίηση του παρατηρούμενου συστήματος, που στην έρευνα αυτή είναι μια δικτυακή ζεύξη. Για παράδειγμα δεν χρειάζεται να κάνουμε υποθέσεις για την συχνότητα εμφάνισης επιθέσεων (a priori statistics). Η θεωρία D-S μας επιτρέπει να εκφράζουμε την πεποίθηση ότι μια ένδειξη υποδεικνύει συγκεκριμένες καταστάσεις του συστήματος.

- Η χρήση της θεωρίας D-S επιτρέπει σε κάθε αισθητήρα να συνεισφέρει γνώση για την κατάσταση του συστήματος ανάλογα με τη διακριτική του ικανότητα. Για παράδειγμα η αυξημένη χρησιμοποίηση της CPU

ενός δρομολογητή μπορεί να αποτελέσει ένδειξη μιας επίθεσης DDoS αλλά δεν μας επιτρέπει από μόνη της να διακρίνουμε τον τύπο της. Με τον τρόπο αυτό μοντελοποιούμε τις ικανότητες ανίχνευσης κάθε αισθητήρα και εισάγουμε γνώση στο σύστημα για καλύτερα αποτελέσματα ανίχνευσης.

- Η θεωρία D-S μας προσφέρει τη δυνατότητα αποτύπωσης της άγνοιας και αβεβαιότητας στις εκτιμήσεις μας. Μια τυπική προσέγγιση με στόχο την βελτίωση της απόδοσης ενός συστήματος ανίχνευσης είναι η ανάθεση ως βαθμού αβεβαιότητας ενός αισθητήρα-ανιχνευτή της ακρίβειας των παρελθόντων εκτιμήσεων (predicted accuracy of a sensor), όπως το false alarm ratio του.

- Μας προσφέρεται η δυνατότητα ενημέρωσης των πεποιθήσεων μας με νέα δεδομένα (ακόμη και αντικρουόμενα). Για παράδειγμα μπορούμε να ενεργοποιούμε ασύγχρονα αλγορίθμους ανίχνευσης με μεγάλη υπολογιστική πολυπλοκότητα αν δεν έχουμε ένα ξεκάθαρο αποτέλεσμα.

- Η σύνθεση δεδομένων με τον κανόνα του Dempster απαιτεί έναν εύκολα υλοποιήσιμο υπολογισμό χωρίς περιόδους μάθησης[3] (training). Το αποτέλεσμα της σύνθεσης μπορεί επίσης να αναπαράγεται από πολλούς κόμβους σύνθεσης (κατανεμημένα) στην περίπτωση που όλοι μοιράζονται κοινά δεδομένα. Κάθε κόμβος σύνθεσης δεν απαιτεί ιδιαίτερη ευφυΐα και μπορεί να λαμβάνει υπόψη στοιχεία πολιτικής π.χ. να συμπεριλαμβάνει ή όχι τις πεποιθήσεις ενός αισθητήρα ανάλογα με την αξιοπιστία του ή να μειώνει την επιρροή ενός αισθητήρα στο συνολικό αποτέλεσμα με μια διαδικασία που ονομάζεται υποβιβασμός μαρτυρίας [93].

- Λαμβάνοντας υπόψη δεδομένα από πολλές πηγές (ετερογενείς, με διαφορετικές προσεγγίσεις) μπορούμε να βελτιώνουμε τα αποτελέσματα της

[3]Πιθανώς ορισμένοι αισθητήρες να απαιτούν περιόδους μάθησης αλλά η σύνθεση των δεδομένων γίνεται χωρίς εκπαίδευση.

ανίχνευσης. Το στοιχείο αυτό προκύπτει από τα πειράματα που παρουσιάστηκαν αλλά και από αναφορές σε άλλα επιστημονικά πεδία όπως η ανίχνευση συμβάντων σε συγκοινωνιακό δίκτυο (traffic incident detection) [144] ή η διάγνωση της κατάστασης μιας συσκευής (equipment condition monitoring) [98].

Δεν πρέπει πάντως να ξεχνούμε πως η αποτελεσματικότητα κάθε συστήματος εξαρτάται από την ακρίβεια των πηγών από τις οποίες αντλεί γνώση. Συνεπώς νέες αποτελεσματικές προσεγγίσεις (μετρικά, αλγόριθμοι ανίχνευσης) που εντάσσονται σε ένα σύστημα ανίχνευσης προφανώς βελτιώνουν την απόδοση του. Η προτεινόμενη αρχιτεκτονική είναι σε θέση να ενσωματώνει νέες προσεγγίσεις εφόσον τα αποτελέσματα τους μπορούν να εκφραστούν με τη μορφή ενός basic probability assignment (bpa) ώστε να μπορέσουν να συμπεριληφθούν στο τελικό συμπέρασμα.

6.2 Ανοικτά θέματα

Μέσα από την ανάπτυξη της προτεινόμενης αρχιτεκτονικής για την ανίχνευση ανωμαλιών στη δικτυακή κίνηση αποκαλύφθηκαν ένα πλήθος ανοικτών θεμάτων που άπτονται της διαχείρισης ασφάλειας στο Διαδίκτυο.

- Το πρώτο πεδίο μελλοντικής έρευνας αφορά την παρακολούθηση δικτύων. Οι υπάρχουσες πηγές πληροφοριών (SNMP MIBs, Netflow και Packet Capturing) υπόκεινται σε διαφορετικούς περιορισμούς και κάθε μια έχει τα δικά της μειονεκτήματα. Πρόκειται λοιπόν για αλληλοσυμπληρούμενες προσεγγίσεις και απαιτείται περαιτέρω έρευνα στο θέμα της παρακολούθησης δικτύων υψηλών ταχυτήτων ώστε να καταφέρουμε να βελτιώσουμε τις μεθόδους αυτές. Επίσης ως πιθανές πηγές δεδομένων θα πρέπει να μελετηθούν μέθοδοι παρακολούθησης δικτύου με ενεργές μετρήσεις (active measurements) που δεν εξετάστηκαν από την παρούσα διατριβή.

- Η επιλογή μετρικών παραμένει ένα ανοικτό και κρίσιμο θέμα που καθορίζει την απόδοση κάθε συστήματος ανίχνευσης. Νέα μετρικά μπορούν να προκύψουν μέσα από πειραματισμό και έρευνα ώστε να έχουμε περισσότερες και ακριβέστερες ενδείξεις για διάφορους τύπους ανωμαλιών δικτυακής κίνησης. Χαρακτηριστικό παράδειγμα νέου πεδίου έρευνας είναι η χρήση network telescopes [25, 136, 141–143] για την ανίχνευση worms και scans.

- Η έρευνα των αλγόριθμων ανίχνευσης που μπορούν να χρησιμοποιηθούν για την ανίχνευση ανωμαλιών δεν έχει εξαντληθεί. Αναμένεται βελτίωση των μεθόδων που χρησιμοποιούνται και αναφέρουμε χαρακτηριστικά την χρήση multivariate CUSUM test [145] αντί για τον απλό αλγόριθμο CUSUM καθώς και entropy based methods [90]. Ένα σημείο εκκίνησης είναι η διερεύνηση μεθόδων που ανιχνεύουν μεταβολές στον μέσο όρο, στην μεταβλητότητα, στη φασματική δομή κ.α. [146].

- Όπως παρουσιάσαμε στην ενότητα 3.3 οι πιθανοί αλγόριθμοι σύνθεσης δεδομένων είναι πολλοί. Αν και προτείναμε την χρήση της θεωρίας D-S η επιλογή αυτή δεν πρέπει να θεωρηθεί τελική και βέλτιστη. Περαιτέρω έρευνα απαιτείται ώστε να εκτιμηθούν και άλλες προσεγγίσεις π.χ. αυτοοργανούμενα συστήματα από τον χώρο των νευρωνικών δικτύων. Επίσης ακόμη και αν παραμείνουμε στην επιλογή της θεωρίας D-S για τις δυνατότητες μοντελοποίησης που μας προσφέρει θα μπορούσαν να μελετηθούν: η χρήση εναλλακτικών κανόνων σύνθεσης όπως των Yager, Inagaki ή Zhang [95], η χρήση εξαιρέσεων για αντιμετώπιση των περιπτώσεων όπου οι αισθητήρες εκφράζουν αντικρουόμενες απόψεις και η δυνατότητα δυναμικής προσαρμογής των πιθανών καταστάσεων του συστήματος.

- Η απλή μαθηματική περιγραφή των πεποιθήσεων σύμφωνα με την θεωρία Dempster-Shafer μπορεί να αποτελέσει βάση για την δημιουργία ενός

πρωτόκολλου κατανεμημένης ανίχνευσης πολλών αισθητήρων, IDS κ.λ.π. Ως παράδειγμα αναφέρουμε την επέκταση του "Intrusion Detection Exchange Protocol" [147] της IETF ώστε πέρα από την απλή μετάδοση μηνυμάτων συναγερμού (alerts) και των σχετικών αποδεικτικών στοιχείων (logs) να γίνεται σύνθεση των εκτιμήσεων επιμέρους συστημάτων με στόχο την κατανεμημένη ανίχνευση (distributed detection). Άλλωστε η αναπαράσταση της γνώσης με basic probability assignments μπορεί να βοηθήσει στην σύνθεση δεδομένων από οποιαδήποτε πηγή, π.χ. alerts από NMS. Οι πηγές μπορούν να βασίζεται σε διαφορετικές μεθόδους ανίχνευσης όπως η ανίχνευση ανωμαλιών (anomaly detection) και η ανίχνευση κακής χρήσης (misuse detection). Συνεπώς η θεωρία D-S μπορεί να αποτελέσει το μαθηματικό υπόβαθρο ενός πρωτόκολλου κατανεμημένης ανίχνευσης.

- Η ανίχνευση των ανωμαλιών στη δικτυακή κίνηση όπως έχουμε τονίσει αποτελεί μόνο το πρώτο βήμα για την αντιμετώπιση τους. Έχοντας διαχωρίσει την αντιμετώπιση τους σε διακριτά προβλήματα απομένει η διερεύνηση και η βελτίωση των υπαρχόντων τεχνικών για την ταυτοποίηση και την καταστολή τους. Σύμφωνα με την δική μας προσέγγιση, οι αλγόριθμοι για τον προσδιορισμό των χαρακτηριστικών μιας ανωμαλίας στη δικτυακή κίνηση (ταυτοποίηση) δεν υπόκεινται σε αυστηρούς χρονικούς περιορισμούς που περιορίζουν την πολυπλοκότητα των πιθανών αλγόριθμων. Συνεπώς οι αλγόριθμοι ταυτοποίησης, π.χ. η χρήση αλγόριθμων clustering, αποτελούν ανοικτό θέμα. Αφού αναπτυχθούν αλγόριθμοι που με ακρίβεια και χωρίς λάθη θα μπορούν να ανιχνεύουν και να ταυτοποιούν ανωμαλίες στη δικτυακή κίνηση δεν είναι παρά θέμα χρόνου ώστε οι διαχειριστές δικτύου να υιοθετήσουν την αυτόματη λήψη μεθόδων καταστολής όπως για παράδειγμα την τεχνική traffic shunt (ενότητα 2.5.2). Ανοικτό θέμα στο χώρο αυτό είναι επίσης οι μέθοδοι καθαρισμού της δικτυακής κίνησης από μια ανωμαλία.

Τέλος η χρήση τεχνικών multisensor data fusion μπορεί να εφαρμοστεί και σε άλλες ερευνητικές περιοχές[4], όπου είναι δυνατή η σύνθεση πολλαπλών πηγών πληροφοριών. Μάλιστα η ανάπτυξη λογισμικού, που θα υλοποιεί σύνθεση δεδομένων από πολλές πηγές με διαφανή τρόπο, μπορεί να βοηθήσει την εξάπλωση των αρχιτεκτονικών multisensor data fusion σε ευρύτερα επιστημονικά πεδία.

[4]μια συλλογή από βιβλιογραφικές αναφορές για τη χρήση της θεωρίας D-S ανά ερευνητικό πεδίο διατίθεται στο παράρτημα A του [95].

Παράρτημα Α

Απόδοση όρων στα ελληνικά

administrative domain	διαχειριστικός τομέας
attacker	επιτιθέμενος
autocorrelation	αυτοσυσχέτιση
autocovariance	αυτομεταβλητότητα
backbone	δίκτυο κορμού
basic probability assignment	βασική συνάρτηση ανάθεσης πιθανότητας
Bayesian inference	συμπερασματολογία κατά Bayes
belief interval	διάστημα εμπιστοσύνης
compromised system	παραβιασμένο σύστημα
conditional probability	υπό συνθήκη πιθανότητα
correlation	συσχέτιση
correlation coefficient	συντελεστής συσχέτισης
countermeasures	μέθοδοι καταστολής
covariance	συμμεταβλητότητα
cross-correlation	συναρτήσεις συσχέτισης
data fusion	σύνθεση δεδομένων

Dempster's - Shafer's Theory of Evidence	Θεωρία D-S
Distributed Denial of Service attack (DDoS)	κατανεμημένη επίθεση άρνησης υπηρεσιών
edge network	ακραίο δίκτυο
Egress filtering	έλεγχος εξερχόμενης δικτυακής κίνησης
expert systems	Έμπειρα Συστήματα
first order logic	κατηγορικός λογισμός πρώτης τάξης
flow	δικτυακή ροή
focal sets	εστιακές υποθέσεις
Ingress filtering	έλεγχος εισερχόμενης δικτυακής κίνησης
intrusion	εισβολή
joint probability	από κοινού πιθανότητα
membership function	συνάρτηση συμμετοχής
metric	μετρικό
modular	αρθρωτή
mutually exclusive sets	σύνολα αμοιβαίως αποκλειόμενα
Network Anomaly Detection System (NADS)	σύστημα ανίχνευσης ανωμαλιών στη δικτυακή κίνηση
Packet Capturing	συλλογή πακέτων
packet flooding	καταιγισμός πακέτων
packet header	επικεφαλίδα πακέτου
plausibility	ευλογοφάνεια
powerset	δυναμοσύνολο
process	διεργασία
sensitivity	ευαισθησία
specificity	ειδικότητα

spoofed packets	παραποιημένα πακέτα
time series	χρονοσειρά
traffic aggregate	συνοθήλευμα κίνησης
transit networks	ενδιάμεσα δίκτυα
trojan horse	δούρειος ίππος
upstream network	πάροχος δικτύου
utilization	χρησιμοποίηση
virus	ιός
worms	σκουλήκια

Βιβλιογραφία

[1] National Infrastructure Security Co-ordination Centre (NISCC), "Quarterly reviews 2002-2004," 2004. http://www.niscc.gov.uk/niscc/quarterlyRev-en.html.

[2] Computer Security Institute (CSI), Federal Bureau of Investigation (FBI), "2004 computer crime and security survey," 2004. http://www.gocsi.com/press/20040609.jhtml.

[3] M. Andress, "Denial of service: Fighting back," Feb. 2002. Network World Global Test Alliance. http://www.nwfusion.com/reviews/2002/0902rev.html.

[4] Arbor Networks, "The peakflow platform." http://www.arbornetworks.com.

[5] Cs3. Inc, "MANAnet DDoS white papers." http://www.cs3-inc.com/mananet.html.

[6] Mazu Networks, "White papers." http://www.mazunetworks.com/solutions/white_papers/.

[7] Cisco, "Cisco Guard DDoS mitigation appliances." http://www.cisco.com/en/US/products/ps5888/index.html.

[8] CERT/CC, "Overview incident and vulnerability trends," May 2003. http://www.cert.org/present/cert-overview-trends/.

[9] CERT/CC, "Statistics 1988-2004," 2004. http://www.cert.org/stats/cert_stats.html.

[10] Wikipedia, "Botnet definition." http://en.wikipedia.org/wiki/Botnet.

[11] J. T. Johnson, "Hackers step up DDoS assaults with use of 'zombie armies'," Nov 2004. Network World Data Center Newsletter, http://www.nwfusion.com/newsletters/datacenter/2004/1129datacenter1.html.

[12] Wikipedia, "Script kiddie definition." http://en.wikipedia.org/wiki/Script_kiddie.

[13] I. Sager and S. Hamm, "First yahoo! then ebay. the net's vulnerability threatens e-commerce–and you," Feb. 2000. http://www.businessweek.com/2000/00_08/b3669001.htm.

[14] J. Reynolds, "RFC1135 the helminthiasis of the Internet," Dec 1989.

[15] P. Ferguson and D. Senie, "RFC2827 network ingress filtering: Defeating denial of service attacks which employ IP source address spoofing," May 2000.

[16] Cisco, "Unicast reverse path forwarding enhancements for the ISP-ISP edge." ftp://ftp-eng.cisco.com/cons/isp/security/URPF-ISP.pdf.

[17] Cisco, "Protecting your core: Infrastructure protection access control lists." http://www.ripe.net/ripe/meetings/ripe-49/presentations/ripe49-routing-security-discussion.pdf.

[18] Cisco, "Using CAR during DoS attacks." http://www.cisco.com/warp/public/63/car_rate_limit_icmp.html.

[19] J. Ioannidis, R. Mahajan, S. Floyd, S. Shenker, S. M. Bellovin, and V. Paxson, "Controlling high bandwidth aggregates in the network," *Computer Communications Review*, vol. 32, pp. 62–73, Jul 2002.

[20] Cisco, "Remote triggered black hole filtering." ftp://ftp-eng.cisco.com/cons/isp/security/.

[21] Y. Afek, R. Brooks, and N. Fischbach, "MPLS-based synchronous traffic shunt," 2003. Presentation at 28th NANOG Meeting.

[22] K. Park and H. Lee, "On the effectiveness of route-based packet filtering for distributed DoS attack prevention in power-law Internets," in *Proceedings of ACM SIGCOMM 2001*, July 09 2001.

[23] A. D. ZDNET UK, "With ISPs like these, who needs enemies?," December 2004. http://comment.zdnet.co.uk/andrewdonoghue/0,39027004,39175983,00.htm.

[24] National Institute of Standards and Technology, "Computer security incident handling guide," January 2004. SP 800-61.

[25] D. Moore, G. M. Voelker, and S. Savage, "Inferring Internet Denial-of-Service activity," in *10th USENIX Security Symposium*, pp. 9–22, 2001.

[26] A. Hussain, C. Papadopoulos, and J. Heidemann, "A framework for classifying denial of service attacks," in *Proceedings of ACM SIGCOMM 2003*, Feb 2003.

[27] J. Mirkovic and P. Reiher, "A taxonomy of DDoS attack and DDoS defense mechanisms," *ACM SIGCOMM Computer Communications Review*, vol. 34, Apr 2004.

[28] G. Koutepas, F. Stamatelopoulos, and B. Maglaris, "Distributed management architecture for cooperative detection and reaction to DDoS attacks," *Journal of Network and Systems Management*, vol. 12, no. 1, 2004.

[29] J. Mirkovic, G. Prier, and P. Reiher, "Attacking DDoS at the source," in *Proceedings of ICNP 2002*, (Paris, France), pp. 312–321, November 2002.

[30] T. M. Gil and M. Poletto, "MULTOPS: A data-structure for bandwidth attack detection," in *Proceedings of the 10th USENIX Security Symposium* (USENIX, ed.), (Washington, DC, USA), USENIX, Aug 2001.

[31] T. Bass, "Intrusion detection systems and multisensor data fusion," *Communications of the ACM*, vol. 43, pp. 99–105, Apr. 2000.

[32] D. Hall, *Mathematical Techniques in Multisensor Data Fusion.* Norwood, Massachusetts: Artech House, 1992.

[33] Cisco, "Netflow." `http://www.cisco.com/go/netflow`.

[34] C. Siaterlis and B. Maglaris, "Towards multisensor data fusion for DoS detection," in *Proceedings of ACM SAC'04*, (Nicosia, Cyprus), 2004.

[35] C. Siaterlis and B. Maglaris, "Detecting DDoS attacks using a multilayer perceptron classifier," in *Poster session of 9th IFIP/IEEE International Symposium on Integrated Network Management (IM 2005)*, 2005.

[36] C. Siaterlis and B. Maglaris, "Detecting incoming and outgoing DDoS attacks at the edge using a single set of network characteristics," in *Proceedings of the 10th IEEE Symposium On Computers and Communications (ISCC'2005)*, (Spain), 2005.

[37] C. Siaterlis and B. Maglaris, "Detecting DDoS attacks with passive measurement based heuristics," in *Proceedings of the 9th IEEE Symposium On Computers and Communications (ISCC'2004)*, (Alexandria, Egypt), 2004.

[38] C. Siaterlis and V. Maglaris, "One step ahead to multisensor data fusion for ddos detection," *Journal of Computer Security*, vol. 13, no. 5, pp. 779–806, 2005.

[39] Geant-2, "Joint research activity on network security." http://www.geant2.net/server/show/nav.755.

[40] National Computer Security Center, National Security Agency, "Trusted network interpretation of the tcsec (tni)," 1987. NCSC-TG-005 (Red Book).

[41] National Computer Security Center, National Security Agency, "Trusted computer system evaluation criteria," 1987. DOD 5100.28-STD (Orange Book).

[42] CERT/CC, "Cert advisory ca-1997-28 IP Denial-of-Service attacks," Dec 1997. http://www.cert.org/advisories/CA-1997-28.html.

[43] CERT/CC, "Cert advisory ca-1996-26 Denial-of-Service attack via ping," Dec 1996. http://www.cert.org/advisories/CA-1996-26.html.

[44] CERT/CC, "Cert advisory ca-1996-01 UDP port Denial-of-Service attack," Feb 1996. http://www.cert.org/advisories/CA-1996-01.html.

[45] CERT/CC, "Cert advisory ca-1996-21 TCP SYN flooding and IP spoofing attacks," Sep 1996. http://www.cert.org/advisories/CA-1996-21.html.

[46] CERT/CC, "Cert advisory ca-1998-01 smurf IP denial-of-service attacks," Jan 1998. http://www.cert.org/advisories/CA-1998-01.html.

[47] S. M. Bellovin, "Security problems in the TCP/IP protocol suite," *SIGCOMM Comput. Commun. Rev.*, vol. 19, no. 2, pp. 32–48, 1989.

[48] M. A. Vatis, "Cyber attacks during the war on terrorism: A predictive analysis," tech. rep., Institute For Security Technology Studies, Sep 2001. http://www.ists.dartmouth.edu/analysis/cyber_a1.pdf.

[49] Broadbandreports.com, "Osirusoft MIA? Spammers cripple popular blacklist." http://www.broadbandreports.com/shownews/31856.

[50] ITworld.com, "Al-Jazeera hobbled by DDoS attack." http://www.itworld.com/Sec/3834/030327aljazeera/.

[51] ISC/UMD/Cogent, "Operational report: Events of 21-oct-2002," 2002. http://d.root-servers.org/october21.txt.

[52] V. Paxson, "An analysis of using reflectors for distributed denial-of-service attacks," *SIGCOMM Computer Communications Review*, vol. 31, no. 3, pp. 38–47, 2001.

[53] CERT/CC, "Cert advisory ca-2003-20 w32/blaster worm," Aug 2003. http://www.cert.org/advisories/CA-2003-20.html.

[54] U.S. Department of Justice, "Press release: Man pleads guilty to infecting thousands of computers using worm program then launching them in denial of service attacks," Dec 2005. http://www.usdoj.gov/criminal/cybercrime/clarkPlea.htm.

[55] B. McWilliams, "Cloaking device made for spammers." http://www.wired.com/news/business/0,1367,60747,00.html.

[56] R. Cooper, "Internet penalties plan," Sep 2003. http://ntbugtraq.ntadvice.com/default.aspx?sid=1&pid=47&aid=78.

[57] RIPE NCC, "Internet Service Providers (ISP) - the tier hierarchy." http://www.ripe.net/projects/ris/docs/bgpcheat.html#20.

[58] ATT, "Real-time mitigation of denial of service attacks now available with ATT Internet protect," Jun 2004. http://www.att.com/news/2004/06/01-13096.

[59] Sprint, "Sprint IP defender," 2004. http://www.sprint.com/business/products/products/ipDefender.jsp.

[60] J. Postel, "RFC792 Internet Control Message Protocol. DARPA internet program protocol specification," Sep 1981.

[61] J. Mogul and S. Deering, "RFC1191 path MTU discovery," Nov 1990.

[62] National Security Agency, "Security configuration guides," Sep 2005. http://www.nsa.gov/snac/.

[63] R. Beasley, S. Aldington, and G. Robinson, "From medical student to junior doctor: how to approach the interpretation of investigations," Dec 2005.

[64] L. Li and G. Lee, "DDoS attack detection and wavelets," in *12th International Conference On Computer Communications And Networks (ICCCN)*, (Dallas, Texas USA), 2003.

[65] A. Ramanarran, "Wades: A tool for distributed denial of service attack detection," 2002. TAMU Computer Engineering Group Technical Publications, TAMU-ECE-2002. `http://dropzone.tamu.edu/techpubs/`.

[66] C. Manikopoulos and S. Papavassiliou, "Network intrusion and fault detection: a statistical anomaly approach," *IEEE Communications Magazine*, vol. 40, pp. 76– 82, Oct 2002. Issue: 10, ISSN: 0163-6804.

[67] J. Jiang and S. Papavassiliou, "Detecting network attacks in the Internet via statistical network traffic normality prediction," in *Journal of Network and Systems Management*, vol. 12, pp. 51–72, March 2004. Special Issue: Security and Management.

[68] P. Barford, J. Kline, D. Plonka, and A. Ron, "A signal analysis of network traffic anomalies," in *Internet Measurement Workshop*, 2002.

[69] P. Barford and D. Plonka, "Characteristics of network traffic flow anomalies," in *Proceedings of the First ACM SIGCOMM Internet Measurement Workshop*, (New York), pp. 69–74, ACM Press, Nov. 1–2 2001.

[70] A. Akella, A. Bharambe, M. Reiter, and S. Seshan, "Detecting ddos attacks on isp networks," in *ACM SIGMOD/PODS Workshop on Management and Processing of Data Streams [MPDS]*, 2003.

[71] A. Hussain, J. Heidemann, and C. Papadopoulos, "Identification of repeated attack scenarios using network traffic forensics," Tech. Rep. ISI-TR-2003-577, USC/Information Sciences Institute, August 2003.

[72] J. Jung, B. Krishnamurthy, and M. Rabinovich, "Flash crowds and denial of service attacks: Characterization and implications for CDNs and web sites," 2002.

[73] J. B. D. Cabrera, L. Lewis, X. Qin, W. Lee, R. K. Prasanth, B. Ravichandran, and R. K. Mehra, "Proactive detection of distributed denial of service attacks using MIB traffic variables - A feasibility study," in *Proceedings of International Symposium on Integrated Network Management*, Feb. 02 2001.

[74] M. Behringer, "Tracing DoS attacks," Jun 2002. Hi Tech 2002 Workshop, Limmerick, IE.

[75] H. Wang, D. Zhang, and K. G. Shin, "Detecting SYN flooding attacks," in *Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Society (INFOCOM-02)*, vol. 3 of *Proceedings IEEE INFOCOM 2002*, (Piscataway, NJ, USA), pp. 1530–1539, IEEE Computer Society, June 23–27 2002.

[76] R. Blazek, H. Kim, B. Rozovskii, and A. Tartakovsky, "A novel approach to detection of denial of service attacks via adaptive sequential and batch-sequential change-point detection methods," in *IEEE Workshop on Information Assurance and Security*, pp. 220–226, Jun 2001.

[77] S. Noh, C. Lee, K. Choi, and G. Jung, "Detecting distributed denial of service DDoS attacks through inductive learning," *Lecture Notes in Computer Science*, vol. 2690, pp. 286–295, 2003.

[78] T. Peng, C. Leckie, and R. Kotagiri, "Proactively detecting DDoS attack using source IP address monitoring," in *Networking 2004, Athens,Greece*, May 2004.

[79] C. Estan, S. Savage, and G. Varghese, "Automatically inferring patterns of resource consumption in network traffic," in *Proceedings of the ACM SIGCOMM Conference*, (Karlsruhe, Germany), August 2003.

[80] C.-M. Cheng, H. Kung, and K.-S. Tan, "Use of spectral analysis in defense against DoS attacks," in *Proceedings of IEEE GLOBECOM*, 2002. Division of Engineering and Applied Science Harvard University.

[81] A. Lakhina, M. Crovella, and C. Diot, "Diagnosing network-wide traffic anomalies," in *SIGCOMM '04: Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications*, (New York, NY, USA), pp. 219–230, ACM Press, 2004.

[82] V. A. Siris and F. Papagalou, "Application of anomaly detection algorithms for detecting SYN flooding attacks," in *IEEE Globecom 2004 (Security and Network Management Symposium), Dallas, USA*, Nov 2004.

[83] T. Peng, C. Leckie, and K. Ramamohanarao, "Detecting distributed denial of service attacks by sharing distributed beliefs," in *8th Australasian Conference on Information Security and Privacy*, (Wollongong, Australia), Jul 2003.

[84] C. Morrow (UUnet), T. Battles (AT&T), and D. McPherson (Arbor), "Customer-triggered real time blackhole," 2004. Presentation at 30th NANOG Meeting.

[85] S. Savage, D. Wetherall, A. Karlin, and T. Anderson, "Practical network support for IP traceback," in *Proceedings of the 2000 ACM SIGCOMM Conference*, Aug 2000.

[86] D. X. Song and A. Perrig, "Advanced and authenticated marking schemes for IP traceback," in *Proceedings of the Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM-01)*, (Los Alamitos, CA), pp. 878–886, IEEE Computer Society, Apr. 22–26 2001.

[87] D. Dean, M. Franklin, and A. Stubblefield, "An algebraic approach to IP traceback," *ACM Transactions on Information and System Security*, vol. 5, pp. 119–137, May 2002.

[88] S. Bellovin, M. Leech, and T. Taylor, "The ICMP traceback message," Oct 2001.

[89] A. C. Snoeren, C. Partridge, L. A. Sanchez, C. E. Jones, F. Tchakountio, S. T. Kent, and W. T. Strayer, "Hash-Based IP traceback," in *Proceedings of the ACM SIGCOMM 2001 Conference*, vol. 31, 4 of *Computer Communication Review*, (New York), pp. 3–14, ACM Press, Aug. 27–31 2001.

[90] J.Llinas and E. Waltz, *Multisensor Data Fusion*. Norwood, Massachusetts: Artech House, 1990.

[91] Σ. Γ. Τζαφέστας, *Υπολογιστική Νοημοσύνη*. Εθνικό Μετσόβιο Πολυτεχνείο, 2002.

[92] P. Kabiri and A. A. Ghorbani, "Research on intrusion detection and response: A survey," *International Journal of Network Security*, vol. 1, Sep 2005.

[93] Σ. Γ. Τζαφέστας, *Εισαγωγή στην Τεχνητή Νοημοσύνη και τα Έμπειρα Συστήματα*. Εθνικό Μετσόβιο Πολυτεχνείο, 1996. Τεύχος Α', Β' Έκδοση.

[94] S. Mohiuddin, S. Hershkop, R. Bhan, and S. J. Stolfo, "Defending against a large scale DoS attack," *Proceedings of the 3rd Annual IEEE Information Assurance Workshop*, June 2002.

[95] K. Sentz, "Combination of evidence in Dempster-Shafer theory," tech. rep., SAND 2002-0835, Systems Science and Industrial Engineering Department Thomas J. Watson School of Engineering and Applied Science, Binghamton University P.O. Box 6000 Binghamton, NY 13902-6000, Apr 2002.

[96] G. Shafer, *A Mathematical Theory of Evidence*. Princeton: Princeton University Press, 1976.

[97] J. Kohlas and P. Monney, "Theory of evidence - a survey of its mathematical foundations, applications and computational analysis,"

ZOR- Mathematical Methods of Operations Research, vol. 39, pp. 35–68, 1994.

[98] K. Tomsovic and B. Baer, "Fuzzy information approaches to equipment condition monitoring and diagnosis," *Electric Power Applications of Fuzzy Systems, IEEE Press*, pp. 59–84, 1998.

[99] N. Wilson and S. Moral, "Fast markov chain algorithms for calculating Dempster-Shafer belief," in *European Conference on Artificial Intelligence*, pp. 672–678, 1996.

[100] G. L. Rogova, P. Losiewicz, and J. Y. Choi, "Connectionist approach to multiattribute decision making under uncertainty," tech. rep., Department of the Air Force, Air Force Research Laboratory (AFMS), Center for Multisource Information Fusion, State University of New York at Buffalo. 421 Bell Hall Buffalo, NY 14260, Oct 1998. REPORT NO. CMIF-5A-98.

[101] R. Haenni, "Ignoring ignorance is ignorant," Mar 2003. Philosophy, Probability, and Modeling research group, University of Konstanz.

[102] Zadeh, "Review of books: A mathematical theory of evidence," 1984.

[103] D. Dubois and H. Prade, "A set-theoretic view on belief functions: Logical operations and approximations by fuzzy sets," *International Journal of General Systems*, vol. 12, pp. 193–226, 1986.

[104] D. Dubois and H. Prade, "On the combination of evidence in various mathematical frameworks," *Reliability Data Collection and Analysis*, pp. 213–241, 1992.

[105] L. Zhang, *Representation, independence, and combination of evidence in the Dempster-Shafer theory*. New York, NY, USA: John Wiley & Sons, Inc., 1994.

[106] R. R. Yager, "On the Dempster-Shafer framework and new combination rules," *Information Sciences*, vol. 41, no. 2, pp. 93–137, 1987.

[107] T. Inagaki, "Interdependence between safety-control policy and multiple-sensor schemes via Dempster-Shafer theory," in *IEEE Transactions on Reliability*, vol. 40, pp. 182–188, 1991.

[108] X. Luo and R. K. C. Chang, "On a new class of pulsing denial-of-service attacks and the defense.," in *Proceedings of the Network and Distributed System Security Symposium, NDSS 2005, San Diego, California, USA*, The Internet Society, 2005.

[109] P. Astithas, G. Koutepas, A. Moralis, and B. Maglaris, "SIDS - a system for enterprise-wide intrusion detection," in *International System Security Engineering Association Conference '01*, Feb 2001.

[110] T. Oetiker, "About RRDtool." `http://people.ee.ethz.ch/~oetiker/webtools/rrdtool`.

[111] Geant-2, "Joint research activity on performance measurement and monitoring." `http://www.geant2.net/server/show/nav.754`.

[112] A. Habib, M. Hefeeda, and B. Bhargava, "Detecting service violations and DoS attacks," in *Network and Distributed System Security Symposium Conference Proceedings*, Internet Society, 2003.

[113] J. Coppens, E. Markatos, J. Novotny, M. Polychronakis, V. Smotlacha, and S. Ubik, "SCAMPI - a scaleable monitoring platform for the Internet," in *2nd International Workshop on Inter-Domain Performance and Simulation (IPS 2004)*, (Budapest, Hungary), 03 2004.

[114] I. Charitakis, D. Pnevmatikatos, E. Markatos, and K. Anagnostakis, "S2I: a tool for automatic rule match compilation for the ixp network processor," in *Proceedings of the 7th International Workshop on Software and Compilers for Embedded Systems (SCOPES 2003)*, (Vienna), September 2003.

[115] Cisco, "Netflow." `http://www.cisco.com/warp/public/146/pressroom/1996/apr96/292.html`.

[116] J. Networks, "J-flow." `http://www.juniper.net/products/modules/monitoring_pic.html`.

[117] S. Leinen, "IP flow information export (IPFIX) working group," Mar 2006. `http://www.ietf.org/html.charters/ipfix-charter.html`.

[118] S. Leinen, "RFC3955 evaluation of candidate protocols for IP flow information export (IPFIX)," Oct 2004.

[119] C. Lund, M. Thorup, and N. Duffield, "Properties and prediction of flow statistics," in *Internet Measurement Workshop 2002*, Aug. 09 2002.

[120] G. C. Polyzos, H. werner Braun, and K. C. Claffy, "Application of sampling methodologies to network traffic characterization," in *ACM SIGCOMM, 1993*, pp. 194 – 203, 1993.

[121] N. Duffield, C. Lund, and M. Thorup, "Estimating flow distributions from sampled flow statistics," *IEEE/ACM Trans. Netw.*, vol. 13, no. 5, pp. 933–946, 2005.

[122] N. Duffield, C. Lund, and M. Thorup, "Estimating flow distributions from sampled flow statistics," in *SIGCOMM '03: Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, (New York, NY, USA), pp. 325–336, ACM Press, 2003.

[123] K. McCloghrie and F. Kastenholz, "RFC1573 evolution of the interfaces group of MIB-II," Jan 1994.

[124] Cisco, "Netflow MIB and top talkers." `http://www.cisco.com/univercd/cc/td/doc/product/software/ios122s/122snwft/release/122s25/nflowtt.pdf`.

[125] C. Papadopoulos, R. Lindell, J. Mehringer, A. Hussain, and R. Govidan, "COSSACK: coordinated suppression of simultaneous attacks," *discex*, vol. 02, p. 94, 2003.

[126] C. Kotsokalis, D. Kalogeras, and B. Maglaris, "Router-based detection of DoS and DDoS attacks," in *8th Workshop of the HP OpenView University Association*, June 2001.

[127] A. Doumas, K. Mavroudakis, D. Gritzalis, and S. K. Katsikas, "Design of a neural network for recognition and classification of computer viruses.," *Computers & Security*, vol. 14, no. 5, pp. 435–448, 1995.

[128] R. J. P. D. Figueiredo, "Implications and applications of Kolmogorov's superposition theorem," in *IEEE Transactions on Autom. Control*, pp. 1227–1230, 1980.

[129] A. N. Kolmogorov, "On the representation of continuous functions of many variables by superpositions of continuous functions of one variable and addition," 1957. (In Russian).

[130] D. A. Curry and H. Debar, "Intrusion detection message exchange format data model and extensible markup language (XML) document type definition." Internet Draft draft-ietf-idwg-requirements-10.txt, Nov. 2002. Work-in-progress.

[131] Π. Ρόρης, "Ανάπτυξη και εφαρμογή αλγορίθμων σύντηξης δεδομένων για την ανίχνευση επιθέσεων σε δίκτυα υπολογιστών," Ιούλιος 2003. Διπλωματική εργασία. Εθνικό Μετσόβιο Πολυτεχνείο. Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών ΗΥ, Εργαστήριο Διαχείρισης & Βελτίστου Σχεδιασμού Δικτύων.

[132] J. Ritter, "ngrep - network grep." `http://ngrep.sourceforge.net/`.

[133] B. Caswell and M. Roesch., "Snort: The open source network intrusion detection system." `http://www.snort.org`.

[134] Open-source community, "The libpcap project." `http://sourceforge.net/projects/libpcap/`.

[135] L. Deri, "Passively monitoring networks at gigabit speeds using commodity hardware and open source software," in *Passive and Active Measurement Workshop 2003*, NLANR/MNA, April 2003.

[136] Ε. Μεγαλιού, "Ανάπτυξη και εφαρμογή ενός network telescope για ανίχνευση δικτυακών ανωμαλιών," Οκτώβριος 2005. Διπλωματική εργασία. Εθνικό Μετσόβιο Πολυτεχνείο. Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών ΗΥ, Εργαστήριο Διαχείρισης & Βελτίστου Σχεδιασμού Δικτύων.

[137] S. Romig, M. Fullmer, and R. Luman, "The OSU Flow-tools package and CISCO NetFlow logs," in *Proceedings of 14th Systems Administration Conference (LISA 2000)*, pp. 291–303, 2000.

[138] D. Dittrich, "The Stacheldraht distributed denial of service attack tool." `http://staff.washington.edu/dittrich/misc/stacheldraht.analysis`.

[139] J. Barlow and W. Thrower, "TFN2K - an analysis," March 2000. `http://packetstormsecurity.com/distributed/TFN2k_Analysis-1.3.txt`.

[140] M. T. Hagan and M. Menhaj, "Training feedforward networks with the Marquardt algorithm," in *IEEE Transactions on Neural Networks*, vol. 5-6, pp. 989–993, 1994.

[141] R. Pang, V. Yegneswaran, P. Barford, V. Paxson, and L. Peterson, "Characteristics of Internet background radiation," in *IMC '04: Proceedings of the 4th ACM SIGCOMM conference on Internet measurement*, (New York, NY, USA), pp. 27–40, ACM Press, 2004.

[142] E. Cooke, M. Bailey, Z. M. Mao, D. Watson, F. Jahanian, and D. McPherson, "Toward understanding distributed blackhole placement," in *WORM '04: Proceedings of the 2004 ACM workshop on Rapid malcode*, (New York, NY, USA), pp. 54–64, ACM Press, 2004.

[143] V. Yegneswaran, P. Barford, and D. Plonka, "On the design and use of Internet sinks for network abuse monitoring," in *RAID* (E. Jonsson, A. Valdes, and M. Almgren, eds.), vol. 3224 of *Lecture Notes in Computer Science*, pp. 146–165, Springer, 2004.

[144] S. C. Byun, D. B. Choi, and B. H. Ahn, "Traffic incident detection using evidential reasoning based data fusion.," in *Proceeding of the 6th World Congress on Intelligent Transport Systems*, (Toronto, Canada), 1999.

[145] P. Qiu and D. Hawkins, "A nonparametric multivariate cumulative sum procedure for detecting shifts in all directions," *Journal of the Royal Statistical Society: Series D (The Statistician)*, vol. 52, p. 151, Jul 2003. http://www.blackwell-synergy.com/doi/abs/10.1111/1467-9884.00348.

[146] S. Rodionov, "A brief overview of the regime shift detection methods," Jun 2005. http://www.beringclimate.noaa.gov/regimes/Regime_shift_methods_list.htm.

[147] B. S. Feinstein, G. A. Matthews, and J. C. C. White, "The intrusion detection exchange protocol (IDXP)." Internet Draft draft-ietf-idwg-beep-idxp-07.txt, Oct. 2002. Work-in-progress. http://www.ietf.org/internet-drafts/draft-ietf-idwg-beep-idxp-07.txt.