Καμμας Παντελης

Διδακτορικη διατριβη

Μαθηματικα μοντελα αναλυσης εισβολων σε δικτυα υπολογιστων με χρηση των χαρακτηριστικων αποδοτικοτητας των δικτυων

Εισηγητής: .....

Επιτροπή

Καθηγητής 1

Καθηγητής 3

Καθηγητής 2

Περιεχόμενα

Εισαγωγή

Τα τελευταία χρόνια το διαδίκτυο αναπτύσσεται και επεκτείνεται με εκθετικούς ρυθμούς τόσο σε επίπεδο πλήθους χρηστών όσο και σε επίπεδο παρεχόμενων υπηρεσιών. Η ευρεία χρήση των κατανεμημένων βάσεων δεδομένων, των κατανεμημένων υπολογιστών και των τηλεπικοινωνιακών εφαρμογών βρίσκει άμεση εφαρμογή και αποτελεί θεμελιώδες στοιχείο στις επικοινωνίες, στην άμυνα, στις τράπεζες, στα χρηματιστήρια, στην υγεία, στην εκπαίδευση και άλλους σημαντικούς τομείς. Το γεγονός αυτό, έχει κάνει επιτακτική την ανάγκη προστασίας των υπολογιστικών και δικτυακών συστημάτων από απειλές που μπορούν να τα καταστήσουν τρωτά σε κακόβουλους χρήστες και ενέργειες. Αλλά για να προστατεύσουμε κάτι θα πρέπει πρώτα να καταλάβουμε και να αναλύσουμε από τι απειλείται. Η διαθεσιμότητα αξιόπιστων μοντέλων σχετικά με τη διάδοση απειλών στα δίκτυα υπολογιστών, μπορεί να αποδειχθεί χρήσιμη με πολλούς τρόπους, όπως το να προβλέψει μελλοντικές απειλές (ένα νέο Code Red worm) ή να αναπτύξει νέες μεθόδους αναχαίτισης. Αυτή η αναζήτηση νέων και καλύτερων μοντέλων αποτελεί ένα σημαντικό τομέα έρευνας στην ακαδημαϊκή και όχι μόνο κοινότητα.

Σκοπός της παρούσης εργασίας είναι η δημιουργία κάποιων επιδημιολογικών μοντέλων. Αναλύουμε για κάθε μοντέλο τις υποθέσεις που έχουν γίνει, όπως επίσης τα δυνατά και αδύνατα σημεία αυτών. Τα επιδημιολογικά μοντέλα που παρουσιάζουμε και αναλύουμε είναι εμπνευσμένα από τα αντίστοιχα βιολογικά, που συναντάμε σήμερα σε τομείς όπως για παράδειγμα ο τομέας της επιδημιολογίας στην ιατρική που ασχολείται με μολυσματικές ασθένειες. Αυτά τα μοντέλα χρησιμοποιούνται για να μοντελοποιηθεί η διάδοση αρκετών απειλών στα δίκτυα υπολογιστών, όπως για παράδειγμα οι ιοί και τα σκουλήκια (viruses and worms). Οι ιοί και τα σκουλήκια είναι δύο διαφορετικά είδη (θα τα παρουσιάσουμε αναλυτικά στο επόμενο κεφάλαιο), αλλά για τις ανάγκες αυτής της εργασίας όπου μας ενδιαφέρουν γενικότερα οι πληθυσμοί των απειλών, θα τα θεωρούμε το ίδιο. Θα πρέπει ακόμα να αναφέρουμε ότι οι ιοί υπολογιστών και τα σκουλήκια (worms) είναι οι μόνες μορφές τεχνητής ζωής που έχουν μετρήσιμη επίδραση-επιρροή στη κοινωνία. Επίσης αναφέρουμε συγκεκριμένα παραδείγματα όπως το Code Red worm [34], τον οποίων η διάδοση έχει χαρακτηριστεί επιτυχώς από αυτά τα μοντέλα. Αναλύουμε τις βασικές στρατηγικές σάρωσης που χρησιμοποιούν σήμερα τα worms προκειμένου να βρουν και να διαδοθούν σε νέα συστήματα. Επίσης παρουσιάζουμε κάποιες βασικές κατηγορίες δικτύων που συναντάμε σήμερα και χαρακτηρίζουν τα δίκτυα υπολογιστών. Η γνώση αυτή που αφορά την τοπολογία των δικτύων είναι ένα απαραίτητο στοιχείο που σχετίζεται άμεσα με τη διάδοση κάποιων απειλών που μελετάμε στη συγκεκριμένη εργασία.

Παρόλο που ένας μεγάλος αριθμός ερευνητών έχει εστιάσει την προσπάθεια του στην επινόηση νέων τεχνικών για την ανίχνευση και την εξάλειψη των απειλών,

δεν δίνει και τόση μεγάλη σημασία στην ανάπτυξη θεωρητικών μοντέλων που είναι ικανά να προβλέψουν το μέγεθος της εξάπλωσης των απειλών σε ευπαθή δίκτυα. Το 2000 οι Wang κ.α. πρότειναν και ανέλυσαν ένα μοντέλο διάδοσης ιών σε μια κατηγορία δικτύων με μορφή δέντρου [32]. Σύμφωνα με αυτό το μοντέλο, οι ιοί παράγουν αντίγραφα του εαυτού τους και εξαπλώνονται στο δίκτυο με σταθερό ρυθμό, χωρίς να είναι αναγκαία η μεσολάβηση των χρηστών. Το 2002 οι Zou κ.α. μελέτησαν την διάδοση του Code Red Worm, του Nimda και του Slammer worm με βάση τα κλασσικά επιδημικά μοντέλα Kermack-Mckendrick [34, 1, 23]. Άλλοι ερευνητές όπως οι Mannan, van Oorschot [19] μελέτησαν ειδικότερα τα IM worms, δηλαδή ιών που διαδίδονται μέσου του δικτύου επικοινωνίας των χρηστών, και συνοψίζοντας τα κύρια χαρακτηριστικά αυτών των δικτύων κατέληξαν σε θεωρητικά συμπεράσματα για την διάδοση απειλών σε αυτά τα δίκτυα. Σε αυτήν την διατριβή αναλύουμε μοντέλα εισβολών απειλών που συνδέουν τα χαρακτηριστικά αποδοτικότητας των δικτύων (χρόνοι εξυπηρέτησης, ρυθμός χρησιμοποίησης), με την ταχύτητα που διαδίδεται η απειλή.

Ειδικότερα, στο πρώτο κεφάλαιο παρουσιάζουμε κάποιες βασικές έννοιες που χρειάζεται να γνωρίζει κάποιος για να μπορέσει να κατανοήσει το περιεχόμενο και τα συμπεράσματα αυτής της εργασίας. Αρχικά παρουσιάζουμε κάποιες βασικές γνώσεις από τον χώρο της πληροφορικής όπως τι είναι το διαδίκτυο (Internet) και κάποιες εφαρμογές του και ποιες είναι οι απειλές που δέχονται, καθώς και τον τρόπο με τον οποίο μεταδίδονται. Έπειτα κάνουμε μια εισαγωγή στα επιδημιολογικά μοντέλα και παρουσιάζουμε τις σχέσεις που έχουν τα βιολογικά μοντέλα με αυτά που εφαρμόζονται στα δίκτυα υπολογιστών [30]. Να τονίσουμε ότι η πρώτη προσπάθεια για να μοντελοποιηθεί μία διάδοση απειλών σε δίκτυα υπολογιστών, έγινε με βάση το κλασσικό βιολογικό μοντέλο Lotka-Volterra [18, 31]. Στο δεύτερο κεφάλαιο προτείνουμε και αναλύουμε ένα μοντέλο που αποτελείται από ένα σύστημα δύο διαφορικών εξισώσεων, για να μοντελοποιήσουμε την ταυτόχρονη εξέλιξη των πληθυσμών των ιών (viruses) και των λογισμικών προστασίας (antiviruses) σε ένα δίκτυο υπολογιστών [12]. Στην συνέχεια θα επεκτείνουμε αυτό το μοντέλο σε ένα που θα περιέχει και ένα τρίτο είδος, τα traps (παγίδες), τα οποία είναι πανίσχυρα λογισμικά προστασίας. Και για τα δύο αυτά μοντέλα κάνουμε θεωρητική ανάλυση για να βρούμε σημεία ισορροπίας του συστήματος, δηλαδή σημεία στα οποία τα δύο είδη θα συνυπάρχουν χωρίς καμιά μεταβολή στους πληθυσμούς τους. Για το δεύτερο μοντέλο παρουσιάζουμε κάποια αριθμητικά αποτελέσματα λόγω της μη- γραμμικότητας του που μας εμποδίζει να βρούμε κάποια αναλύτικη λύση. Στο τρίτο κεφάλαιο μελετάμε το πρόβλημα διάδοσης και αναχαίτησης ιών σε δίκτυα από μία άλλη οπτική γωνία η οποία αποφεύγει την χρήση των μη γραμμικών διαφορικών εξισώσεων, όπως το μοντέλο Lotka-Volterra. Προτείνουμε και αναλύσουμε ένα μαθηματικό μοντέλο για την ταυτόχρονη εξέλιξη και των δύο πληθυσμών ιών και λογισμικών προστασίας, που βασίζεται στην θεωρία ουρών [11]. Αυτό το μοντέλο λαμβάνειν υπ'οψιν τα χαρακτηριστικά των εξυπηρετητών (servers) και της κίνησης (traffic) του δικτύου στο οποίο επικεντρώνεται η μελέτη μας. Ο πρώτος που μελέτησε τα δίκτυα με αυτόν τον τρόπο ήταν ο Jackson, γι' αυτό και τα δίκτυα υπολογιστών με χαρακτηριστικά ουρών ονομάζονται ανοικτά δίκτυα Jackson (Open Jackson Networks) [3, 16]. Στο τέταρτο κεφάλαιο παρουσιάζουμε συστήματα διαφορικών εξισώσεων, τα οποία μοντελοποιούν ταυτόχρονες εξελίξεις των πληθυσμών, μιας ειδικής κατηγορίας ιών και λογισμικών προστασίας, αυτών των DNS [10]. Αυτά τα είδη χρησιμοποιούν έναν γεννήτορα τυχαίων strings για την εύρεση πιθανών πραγματικών διευθύνσεων και στην συνέχεια θέτουν ερωτήματα στους εξυπηρετητές (Domain Name Servers)

με σκοπό την αποκόμιση της πραγματικής ηλεκτρονικής διεύθυνσης. Θεωρούμε ότι η εξάπλωση αυτών των ιών γίνεται σε IPv6 δίκτυα, ένα νέο πρωτόκολο δικτύων που χρησιμοποιείται τα τελευταία χρόνια λόγω του μεγαλύτερου εύρους των ηλεκτρονικών διευθύνσεων που μπορεί να περιλαμβάνει, για να κάνει δυσκολότερη την γρήγορη διάδοση των απειλών. Όλες οι παράμετροι αυτών των εξισώσεων μοντελοποιούνται από τα χαρακτηριστικά του δικτύου, καθώς και με μαθηματική ανάλυση του τρόπου με τον οποίο λειτουργούν οι εξυπηρετητές DNS. Στην συνέχεια επεκτείνουμε τα μοντέλα με την προσθήκη των dummy honeypot servers, οι οποίοι είναι εικονικοί εξυπηρετητές, και είναι μία ιδέα η οποία έχει προταθεί για καλύτερη αντιμετώπιση μιας ταχείας διάδοσης μίας απειλής. Τα συμπεράσματα μας όμως δεν είναι ενθαρυντικά για όσους βασίζονται σε αυτήν την ιδέα για την προστασία ενός δικτύου. Τέλος στο πέμπτο κεφάλαιο προτείνουμε ένα γενικό μοντέλο για την μελέτη της ταυτόχρονης εξέλιξης των πληθυσμών πολλαπλών κυνηγών, καθώς συναγωνίζονται ή ανταγωνίζονται για το ίδιο θήραμα, και ενσωματώνει την επίδραση που έχουν όλες οι αλληλεπιδράσεις μεταξύ των ειδών των κυνηγών, πάνω στον πληθυσμό του θηράματος [9]. Το προτεινόμενο μοντέλο επεκτείνει τα κλασσικά βιολογικά μοντέλα όπως το Lotka-Voltera και το May-Leonard [21] με στοιχεία από την θεωρία παιγνίων και αναλύουμε ειδικότερα το γνωστό παίγνιο μεταξύ γερακιών και περιστεριών (Hawks- Dove game) [26, 17]. Έχουμε μοντελοποιήσει τις αλληλεπιδράσεις μέσω ενός παιγνίου που ορίζεται μεταξύ των ειδών των κυνηγών, που αναπαριστά την κατανάλωση της λείας όπως και το κόστος απόκτησης της καθώς συνεργάζονται αλλά και καθώς μάχονται. Με αριθμητικά αποτελέσματα από την επίλυση των διαφορικών εξισώσεων του συστήματος, δείχνουμε λογικά συμπεράσματα που παράγονται από τις λύσεις, όταν θεωρήσουμε το παίγνιο σαν παράμετρο του συστήματος. Το παίγνιο επιτρέπει τον ορισμό των κερδών στον πίνακα του παιγνίου σαν μεταβλητές, και ειδικότερα σαν συναρτήσεις του πληθυσμού των ειδών των κηνυγών και του θηράματος, και αναπαριστά το γεγονός ότι η συμπεριφορά των πληθυσμών βασίζεται στην ανάγκη για αποθέματα λείας ή της αφθονίας αυτής.

Π. Καμμάς, Ιωάννινα 2010.

Κεφάλαιο 1

Βασικές έννοιες

1.1 Internet και απειλές

1.1α΄ Γενικά για το internet

Το σημερινό Internet αποτελεί εξέλιξη του ARPANET, ενός δικτύου που άρχισε να αναπτύσσεται πειραματικά στα τέλη της δεκαετίας του 60 στις ΗΠΑ. Το 1993, το εργαστήριο CERN στην Ελβετία παρουσιάζει το World Wide Web (WWW) (Παγκόσμιο Ιστό) που αναπτύχθηκε από τον Tim Berners-Lee. Πρόκειται για ένα σύστημα διασύνδεσης πληροφοριών σε μορφή πολυμέσων (multimedia) που βρίσκονται αποθηκευμένες σε χιλιάδες υπολογιστές του Internet σε ολόκληρο τον κόσμο και παρουσίασής τους σε ηλεκτρονικές σελίδες, στις οποίες μπορεί να περιηγηθεί κανείς χρησιμοποιώντας το ποντίκι. Το γραφικό αυτό περιβάλλον έκανε την εξερεύνηση του Internet προσιτή στον απλό χρήστη. Παράλληλα, εμφανίζονται στο Internet διάφορα εμπορικά δίκτυα που ανήκουν σε εταιρίες παροχής υπηρεσιών Internet (Internet Service Providers - ISP) και προσφέρουν πρόσβαση στο διαδίκτυο για όλους [28]. Οποιοσδήποτε διαθέτει υπολογιστή και modem μπορεί να συνδεθεί με το Internet σε τιμές που μειώνονται διαρκώς.

Η ανακάλυψη του WWW σε συνδυασμό με την ευκολία απόκτησης πρόσβασης στο Internet προσέλκυσε έναν μεγάλο αριθμό καινούργιων χρηστών και έφερε την "έκρηξη" που παρακολουθήσαμε τα τελευταία χρόνια. Σήμερα, το μεγαλύτερο μέρος του πληθυσμού της γης ζει σε χώρες που είναι συνδεδεμένες στο Internet. Παρατηρούμε ότι καθημερινά περιοδικά και εφημερίδες εκδίδονται "on-line" και μας παραπέμπουν στις διευθύνσεις τους, επιχειρήσεις και ιδιώτες φτιάχνουν τις δικές τους σελίδες στο WWW, κλπ. Είναι προφανές ότι το Internet δεν αποτελεί πλέον ένα δίκτυο των φοιτητών και των ερευνητών, αλλά ότι επεκτείνεται και επιδρά στις καθημερινές πρακτικές όλων μας. Ήδη μιλάμε για ηλεκτρονικό εμπόριο, τηλεεργασία, τηλεκπαίδευση, τηλεϊατρική, κλπ, μέσα από το Internet.

1.1β΄ Παράγοντες που βοηθούν τη διάδοση απειλών στο διαδίκτυο

Σε αυτή την ενότητα θα αναφερθούμε σε εκείνους τους παράγοντες μέσω των οποίων διευκολύνεται η διάδοση κάποιων κοινών απειλών που συναντάμε σήμερα όλοι μας στο διαδίκτυο. Το ποιες είναι αυτές οι απειλές στις οποίες αναφερόμαστε θα τις δούμε παρακάτω όταν θα κάνουμε μία αναλυτική παρουσίαση αυτών. Προκειμένου να καταλάβουμε καλύτερα αυτούς τους παράγοντες και στο πως βοηθούν στη διάδοση κάποιων γνωστών απειλών, θα πρέπει να αναφερθούμε σε κάποια στοιχεία

τα οποία κρίνουμε απαραίτητα. Οι λεγόμενες απειλές, οι οποίες μας ενδιαφέρουν στη συγκεκριμένη εργασία, όπως θα παρατηρήσουμε και στη συνέχεια, δεν είναι τίποτα άλλο παρά κακόβουλο λογισμικό, το οποίο αποτελείται από διάφορα είδη ανάλογα με τον τρόπο που λειτουργεί και διαδίδεται το καθένα από αυτά.

Το κακόβουλο λογισμικό, που θα αναλύσουμε στη συνέχεια, για να διαδοθεί, θα πρέπει να εκμεταλλευτεί αυτά που ονομάζουμε τρωτά σημεία. Τι είναι όμως ένα τρωτό σημείο· Με τον όρο τρωτό εννοούμε ένα αδύναμο σημείο που εκμεταλλεύεται κάποιος που θέλει να βρει ένα τρόπο να εισβάλει χωρίς εξουσιοδότηση σε ένα υπολογιστικό/ δικτυακό σύστημα. Όταν με χρήση του τρωτού σημείου γίνει εισβολή, τότε μιλάμε για περιστατικό παραβίασης της ασφάλειας. Τα τρωτά σημεία οφείλονται σε σχεδιαστικά και κατασκευαστικά λάθη, αλλά και κάποιους άλλους παράγοντες, όπως μερικούς που παρουσιάζουμε παρακάτω:

(i) Ελαττώματα στο λογισμικό ή στο σχεδιασμό πρωτοκόλλων,

(ii) Αδυναμίες στην υλοποίηση του λογισμικού ή του πρωτοκόλλου,

(iii) Αδυναμίες στη διαμόρφωση των συστημάτων και δικτύων,

(iv) Μεγαλύτερες βάσεις δεδομένων χωρίς τις απαιτούμενες γνώσεις,

(v) Αφαιρούμενα μέσα,

(vi) Web browsing,

(vii) Εφαρμογές instant messaging,

(viii) Downloading,

(ix) Ηλεκτρονικό ταχυδρομείο,

(x) Έλλειψη χρήσης προγραμμάτων προστασίας.

Σε αυτό το σημείο θα παρουσιάσουμε αναλυτικά κάποιους από τους παραπάνω παράγοντες που μας ενδοιαφέρουν περισσότερο σε αυτήν την εργασία, για τον λόγο ότι χρήστες κινούνται μέσα σε κάποιο δίκτυο:

Εφαρμογές instant messaging

Πολλοί χρήστες χρησιμοποιούν αυτές τις εφαρμογές προκειμένου να επικοινωνήσουν με φίλους, να στείλουν ή να δεχθούν αρχεία, μηνύματα, καθώς αυτές οι εφαρμογές προσπαθούν να ξεγελάσουν τα προγράμματα που φιλτράρουν τις πληροφορίες που εισέρχονται και εξέρχονται από ένα δίκτυο, ώστε να περάσουν και δεδομένα που μπορεί να περιέχουν κακόβουλο λογισμικό. Ωστόσο οι χρήστες δεν συνειδητοποιούν τους κινδύνους που κρύβουν αυτές οι εφαρμογές και την πιθανή καταστροφή που μπορεί να επιφέρουν. Ποτέ δεν μπορούμε να είμαστε σίγουροι για το ποιος είναι στο άλλο άκρο της γραμμής. Μπορεί πράγματι να είναι κάποιος φίλος μας ή ένας κακόβουλος χρήστης. Οι περισσότερες από αυτές τις εφαρμογές περιέχουν τρωτά σημεία, η εκμετάλλευση των οποίων από γνώστες του είδους θα μπορούσε να δημιουργήσει σοβαρά προβλήματα. Ένα αρχείο που θα πάρουμε μπορεί να είναι μολυσμένο και να οδηγήσει με η σειρά του στη μόλυνση του συστήματος και του δικτύου γενικότερα.

Web browsing

Το διαδίκτυο και γενικότερα η περιήγηση σε αυτό, το γνωστό σε όλους Web surfing, αποτελεί μία μόνιμη πηγή κινδύνου για την ασφάλεια των σταθμών εργασίας αλλά και γενικότερα του δικτύου σε ένα οργανισμό ή σε μία εταιρεία. Οι χρήστες επισκέπτονται διαδικτυακές τοποθεσίες που περιέχουν επισφαλές περιεχόμενο και συχνά κατεβάζουν από αυτές αρχεία τα οποία το πιθανότερο είναι να είναι μολυσμένα με κακόβουλο λογισμικό. Αυτό το εκμεταλλεύονται μετά κάποιοι κακόβουλοι χρήστες προκειμένου να αποκτήσουν πρόσβαση στο σύστημα και να το κάνουν δικό τους. Επίσης συχνό είναι και το φαινόμενο όπου χρήστες επισκέπτονται σελίδες με πορνογραφικό περιεχόμενο με αποτέλεσμα στις περισσότερες των περιπτώσεων να εγκαθίσταται εν αγνοία του χρήστη ένας dialer που κάνει κλήσεις σε απομακρυσμένες περιοχές.

Ηλεκτρονικό ταχυδρομείο

Το ηλεκτρονικό ταχυδρομείο αποτελεί μία από τις πιο διαδεδομένες και χρησιμοποιούμενες υπηρεσίες του διαδικτύου. Είναι το κύριο μέσο επικοινωνίας μεταξύ των χρηστών και αποτελεί ένα από τα βασικά εργαλεία μιας εταιρείας ή ενός οργανισμού. Αποτελεί όμως και τον κύριο εκπρόσωπο μέσω του οποίου διαδίδεται κακόβουλο λογισμικό με τη μορφή επισυναπτόμενων αρχείων. Ο χρήστης ανοίγει τα αρχεία που δέχεται χωρίς να τα ελέγχει με αποτέλεσμα να μολύνεται το σύστημα. Η μη σωστή χρήση του αποτελεί ένα τρωτό σημείο το οποίο μπορεί να οδηγήσει στη παραβίαση της ασφάλειας του δικτύου. Ενδεικτικά μπορούμε να αναφέρουμε σαν μη σωστή χρήση, όπως είπαμε και πιο πάνω, το άνοιγμα κάποιων αρχείων από άτομα που δεν ξέρουμε, τη μη χρήση πρωτοκόλλων ασφάλειας για κρυπτογράφηση και επιβεβαίωση της πηγής που έστειλε το μήνυμα. Συχνά οι χρήστες δίνουν το email που διατηρούν σε μια εταιρεία σε ιστοσελίδες με αμφίβολο περιεχόμενο προκειμένου να καλύψουν κάποιες προσωπικές ανάγκες. Και αυτό αποτελεί ένα τρωτό σημείο διότι μπορεί να χρησιμοποιηθεί για την παραλαβή spam μηνυμάτων προς τον χρήστη.

1.1γ΄ Είδη Απειλών

Όπως έχουμε αναφέρει και παραπάνω, τα είδη απειλών με τα οποία θα ασχοληθούμε σε αυτή την εργασία και που μας ενδιαφέρει ο τρόπος με τον οποίο διαδίδονται και αντιμετωπίζονται, είναι αυτά που καλούμε σήμερα κακόβουλο λογισμικό, μαζί με τις διάφορες παραλλαγές αυτού που συναντούμε. Μιλώντας κανείς για κακόβουλο λογισμικό μπορεί να δώσει διάφορες ερμηνείες. Αρχικά όμως ας θεωρήσουμε ότι μιλάμε για εχθρικό κώδικα. Τι είναι όμως αυτό που ονομάζουμε εχθρικό κώδικα ή κακόβουλο λογισμικό· Οφείλουμε να ομολογήσουμε πως υπάρχουν δεκάδες ορισμοί σήμερα και που ο καθένας θα περιέγραφε αυτό που θέλουμε να ορίσουμε. Στην περίπτωσή μας λοιπόν ας χρησιμοποιήσουμε τον εξής ορισμό:

"Εχθρικό κώδικα ονομάζουμε ένα σύνολο από εντολές που τρέχουν σε ένα σύστημα-υπολογιστή και τον αναγκάζουν να εκτελεί αυτά που θέλει ο εισβολέας, που δεν είναι άλλος από αυτόν που τον δημιούργησε ή απλώς χρησιμοποιεί τον εχθρικό κώδικα."

Γενικά, όταν εχθρικός κώδικας τοποθετείται στον υπολογιστή ενός χρήστη, το γεγονός αυτό και μόνο δίνει στον εισβολέα (ο οποίος μπορεί να είναι ο δημιουργός του εχθρικού κώδικα ή κάποιος που απλά τον χρησιμοποιεί) σημαντικά δικαιώματα όσον αφορά στον έλεγχο του συστήματος αυτού του χρήστη. Ο κώδικας

λοιπόν αυτός μπορεί να δράσει σαν ένας κατάσκοπος που επιτρέπει στον εισβολέα να κάνει πράξη τα σχέδιά του. Αν ένας εισβολέας λοιπόν μπορεί να εγκαταστήσει εχθρικό κώδικα στον υπολογιστή μας ή να μας ξεγελάσει ώστε να κατεβάσουμε ένα πρόγραμμα που έχει μολυνθεί, τότε ο υπολογιστής μας δρα σαν μια μαριονέτα στα χέρια του εισβολέα που εκτελεί τις εντολές του. Την ίδια στιγμή το σύστημά μας παύει να εκτελεί τις δικές μας εντολές.

Δυστυχώς για εμάς και κυρίως για αυτούς που ασχολούνται με την ασφάλεια των υπολογιστικών συστημάτων και δικτύων, σήμερα συναντάμε μία μεγάλη ποικιλία από κακόβουλο λογισμικό, τα είδη του οποίου αναφέρουμε και αναλύουμε αμέσως παρακάτω:

Ιός virus

Η λέξη "ιός" μπορεί να έχει διάφορες σημασίες ανάλογα με το ποιον ρωτά κανείς κάθε φορά. Ένας ορισμός που μπορεί να δοθεί και είναι σχετικός με το αντικείμενό μας είναι ο ακόλουθος:

"ιός είναι ένα αυτο-αντιγραφόμενο κομμάτι κώδικα που μπορεί και προσκολλάται σε άλλα προγράμματα και συνήθως απαιτεί την βοήθεια του χρήστη για να διαδοθεί-εξαπλωθεί".

Ένα από τα κύρια χαρακτηριστικά των ιών είναι η αδυναμία τους να σταθούν σαν αυτόνομα εκτελέσιμα.

Σκουλήκι worm

Τα σκουλήκια των δικτύων υπολογιστών πρέπει να θεωρηθούν ως οντότητες διαφορετικές από τους ιούς, αν θέλουμε να καταλάβουμε τον τρόπο με τον οποίο λειτουργούν, διαδίδονται και καταπολεμούνται. Πιθανή αδυναμία, στο διαχωρισμό τους από τους ιούς, μπορεί να οδηγήσει σε αναποτελεσματικές μεθόδους ανίχνευσης και καταπολέμησής τους. Όπως οι ιοί, έτσι και τα σκουλήκια, τροποποιούν τη φυσιολογική λειτουργία των μηχανημάτων που μολύνουν. Τα σκουλήκια, συνήθως, εγκαθιστούν τους εαυτούς τους στα μολυσμένα μηχανήματα και κατόπιν αρχίζουν την εκτέλεσή τους. Κατά την εκτέλεσή τους χρησιμοποιούν τους πόρους του μολυσμένου μηχανήματος, όπως άλλωστε και οποιοδήποτε φυσιολογικό πρόγραμμα. Ορίσουμε ως σκουλήκι υπολογιστών:

"Ένα πρόγραμμα το οποίο αντιγράφεται και μολύνει ανεξάρτητα και αυτόνομα και το οποίο δύναται να ανακαλύψει νέα συστήματα και να τα μολύνει χρησιμοποιώντας το δίκτυο."

Ένα σημαντικό ζήτημα είναι γιατί τα τελευταία χρόνια χρησιμοποιούνται τόσο πολύ τα worms. Είναι μία σημαντική ερώτηση με πολλές απαντήσεις που κάθε μία έχει τις δικές της ιδιαιτερότητες και τη δικιά της σημασία. Οι κακόβουλοι χρήστες χρησιμοποιούν τα worms επειδή προσφέρουν δυνατότητες τις οποίες δεν μπορούν να πετύχουν εύκολα με άλλες μορφές επίθεσης. Οι βασικοί στόχοι χρησιμοποίησης είναι οι ακόλουθοι:

(i) Μπορούν και κυριαρχούν σε ένα μεγάλο αριθμό συστημάτων,

(ii) Κάνουν τον εντοπισμό πιο δύσκολο,

(iii) Μπορούν και προκαλούν μεγαλύτερη καταστροφή.

Μερικές βασικές διαφορές που παρατηρούνται μεταξύ των ιών και των σκουληκιών είναι οι ακόλουθες:

(i) Τόσο τα σκουλήκια, όσο και οι ιοί, μεταδίδονται σε άλλα μηχανήματα. Ωστόσο, οι ιοί συνήθως μεταδίδονται προσκολλώντας τον εαυτό τους σε άλλα αρχεία (είτε αρχεία δεδομένων, είτε εκτελέσιμα προγράμματα). Η μετάδοσή τους προϋποθέτει και την μετάδοση του μολυσμένου αρχείου. Σε αντίθεση με τους ιούς, τα σκουλήκια έχουν τη δυνατότητα να μεταδίδονται αυτόνομα από σύστημα σε σύστημα μέσω του δικτύου, χωρίς να χρειάζονται τη βοήθεια άλλου λογισμικού.

(ii) Ένα σκουλήκι είναι ένα ενεργό σύστημα διανομής, το οποίο ελέγχει και χρησιμοποιεί το δίκτυο για να φτάσει στο σύστημα στόχο. Αντίθετα, ένας ιός είναι ένα στατικό μέσο, το οποίο δεν μπορεί να ελέγξει και να κάνει χρήση του συστήματος διανομής, δηλαδή του δικτύου.

(iii) Σε αρκετές περιπτώσεις, οι διάφοροι κόμβοι του δικτύου του σκουληκιού (τα μηχανήματα που έχουν μολυνθεί από το σκουλήκι) μπορούν να ανταλλάξουν πληροφορία μεταξύ τους ή με κάποιον κεντρικό κόμβο. Αντίθετα οι ιοί δεν έχουν τη δυνατότητα να επικοινωνήσουν με εξωτερικά συστήματα.

1.1δ΄ Στρατηγικές σάρωσης των worms

Πριν αρχίσει μία επίθεση κάποιο δικτυακό worm, αυτό θα πρέπει να ελέγξει τα συστήματα τα οποία στοχεύει προκειμένου να δει αν έχουν τα τρωτά σημεία που αυτό στοχεύει. Μία καλή στρατηγική σάρωσης μπορεί να επιταχύνει τη διάδοση ενός worm. Ένα worm με μία ιδανική στρατηγική σάρωσης μπορεί να βρει όλα τα πιθανά συστήματα που είναι να μολύνει στο διαδίκτυο στο λιγότερο χρόνο. Με βάση τους διαφορετικούς τρόπους με τους οποίους ένα worm μπορεί να επιλέξει το πεδίο εκείνων των διευθύνσεων που θέλει να στοχεύσει, οι λεγόμενες στρατηγικές σάρωσης μπορούν να κατηγοριοποιηθούν ως εξής:

(i) Επιλεκτική τυχαία σάρωση (selective random scan),

(ii) Σειριακή σάρωση (sequential scan),

(iii) Τοπική σάρωση,

(iv) Σάρωση με βάση μία λίστα,

(v) Σάρωση διαίρει-και-βασίλευε (Divide and Conquer scan)

(vi) Υβριδική σάρωση (Hybrid scan)

(vii) DNS scan.

DNS σάρωση

Ο δημιουργός ενός worm μπορεί να χρησιμοποιήσει τις IP διευθύνσεις που αποκτά από DNS servers προκειμένου να δημιουργήσει τη βάση με τα συστήματα που πρόκειται να σαρώσει. Το πλεονέκτημα αυτής της μεθόδου είναι ότι όλες οι διευθύνσεις που συλλέγονται (αν όχι όλες οι περισσότερες) χρησιμοποιούνται ήδη. Παρόλα αυτά όμως έχει και κάποια προβλήματα. Αρχικά δεν είναι εύκολο να αποκτήσει κανείς τις λίστες που έχουν οι DNS servers, δεύτερον, ο αριθμός των συστημάτων περιορίζεται μόνο σε εκείνα που έχουν ένα public domain name. Από παρατηρήσεις που έγιναν για ένα γνωστό worm, ξέρουμε ότι τα μισά περίπου από

τα θύματα του Code Red δεν είχαν κάποια DNS εγγραφή. Τρίτον, επειδή το worm θα πρέπει να κουβαλά μία αρκετά μεγάλη βάση, η διάδοσή του θα είναι αρκετά αργή.

Το διαδίκτυο σήμερα επιτρέπει στους ιούς και τα σκουλήκια, να διαδοθούν σε ένα μεγαλύτερο αριθμό από συστήματα ποιο γρήγορα και πιο εύκολα. Το Morris worm το 1988 ήταν το πρώτο δικτυακό σκουλήκι, και εκμεταλλευόταν τα τρωτά σημεία αρκετών προγραμμάτων και εφαρμογών με σκοπό να διαδοθεί. Μπόρεσε και μόλυνε το 10% των συστημάτων που ήταν συνδεδεμένα στο διαδίκτυο εκείνο το καιρό. Το SQL Slammer worm του 2003, επίσης εκμεταλλευόταν κάποια τρωτά σημεία εφαρμογών με αποτέλεσμα να καταφέρει να μολύνει το 15% των συνδεδεμένων στο δίκτυο υπολογιστών. Παρόλα αυτά, ο αριθμός των συστημάτων που ήταν συνδεδεμένα στο δίκτυο το 2003 ήταν πολύ μεγαλύτερο σε σχέση με αυτά του 1988, οπότε το SQL Slammer worm προξένησε μεγαλύτερη ζημιά. Η ζημία που προκαλείται από τα δικτυακά σκουλήκια, δεν περιορίζεται μόνο σε απώλεια χρημάτων ή στη δυσφήμηση που προκαλείται σε εταιρείες και πανεπιστήμια. Σημαντικές υπηρεσίες, όπως είναι κάποιες υπηρεσίες κοινής ωφελείας (Schneier 2003) ή αυτές που σχετίζονται με τη ασφάλεια των πτήσεων των αεροσκαφών (Airwise News 2003) , έχουν επηρεαστεί από κάποια διαδικτυακά σκουλήκια. Δυστυχώς, οι διαχειριστές των δικτυακών και υπολογιστικών συστημάτων, δεν μπορούν να ελέγξουν αυτές τις απώλειες που δημιουργούνται (χρήματα, κ.α,) διότι δεν παράγονται κάποια προειδοποιητικά σήματα σχετικά με μία επικείμενη επίθεση από κάποιο ιό ή κάποιο σκουλήκι. Μερικά worms, όπως το SQL Slammer μπόρεσε να διαδοθεί στο διαδίκτυο σε 10 λεπτά περίπου. Οι περισσότεροι διαχειριστές δεν είχαν τον απαιτούμενο χρόνο προκειμένου να διασφαλίσουν να συστήματα και τα δίκτυα απέναντι σε μία τέτοια επίθεση.

Η διάδοση ενός worm εξαρτάται από τα χαρακτηριστικά που έχει αυτό το worm (για παράδειγμα ποιο τρωτό σημείο εκμεταλλεύεται και ποια τεχνική προκειμένου να δημιουργήσει και να επιλέξει τις IP διευθύνσεις στις οποίες θα επιτεθεί), καθώς επίσης και από την "ευαισθησία" που παρουσιάζει ένα υπολογιστικό σύστημα στο τρωτό σημείο που εκμεταλλεύεται το worm. Τα περισσότερα σημερινά worms χρησιμοποιούν κάποιο μηχανισμό αυτό-πολλαπλασιασμού με σκοπό να πετύχουν τη μετάβαση από υπολογιστή σε υπολογιστή. Αυτά τα worms για τα οποία μιλάμε, μπορούν είτε να αυτό-ενεργοποιηθούν είτε να κάνουν χρήση κάποιου εξωτερικού σήματος που να τα ενεργοποιεί. Για παράδειγμα, στη περίπτωση του ιού Sobig, ο χρήστης θα έπρεπε να ανοίξει ένα μήνυμα ηλεκτρονικού ταχυδρομείου προκειμένου να μολυνθεί ο υπολογιστής. Στη περίπτωση των αυτό-διαδομένων σκουληκιών, δεν είναι απαραίτητη η δράση από το χρήστη προκειμένου να διαδοθεί το worm και να μολύνει κάποιο υπολογιστή. Με αυτό το τρόπο, η διάδοση αυτών των worms είναι ταχεία. Ένα τέτοιο worm συνήθως χρησιμοποιεί μία τεχνική τυχαίας παραγωγής IP διευθύνσεων ώστε να μετακινηθεί στο επόμενο θύμα του. Αυτού του είδους τα worms, χρησιμοποιούν τα πρώτα ένα ή δύο bytes της IP διεύθυνσης του μολυσμένου συστήματος προκειμένου να παράγουν νέες διευθύνσεις και να μολύνουν νέα συστήματα. Για παράδειγμα, το worm Code Red II είχε 3 στις 8 πιθανότητα να χρησιμοποιήσει τα πρώτα δύο bytes της IP διεύθυνσης του μολυσμένου υπολογιστή, ώστε να παράγει και να επιλέξει την επόμενη IP διεύθυνση για να μολύνει. Στη πραγματικότητα, η δημιουργία IP διευθύνσεων δεν είναι εντελώς τυχαία σαν διαδικασία. Η διαδικασία καθορισμού της ροής των computer worms στο δίκτυο, είναι κρίσιμη διότι από αυτή θα επιλεγούν και τα κατάλληλα μέτρα αντίστασης και αντεπίθεσης. Για να φτάσουμε όμως στο σημείο να αντεπιτεθούμε, θα πρέπει πρώτα να έχουμε ανιχνεύσει κάποια εισβολή. Ας δώσουμε έναν ορισμό

για αυτήν την πράξη:

"Ανίχνευση εισβολών καλείται το πρόβλημα του εντοπισμού πράξεων που έχουν σαν σκοπό να διαβάλλουν την ακεραιότητα, την αξιοπιστία ή την διαθεσιμότητα ενός υπολογιστικού πόρου."

Στα πλαίσια αυτής της εργασίας λοιπόν, θα μελετήσουμε στα επόμενα κεφάλαια κάποια μοντέλα που αποτελούνται στην ουσία από μαθηματικές εξισώσεις, οι οποίες καταφέρνουν και αποκαλύπτουν σημαντικά στοιχεία για τη διάδοση των worms.

1.2 Δίκτυα και Επιδημιολογία

Τα δίκτυα και η επιδημιολογία των άμεσα μεταδοτικών μολυσματικών ασθενειών είναι βασικά συνδεδεμένα. Τα θεμέλια της επιδημιολογίας και τα πρώιμα επιδημιολογικά μοντέλα βασίστηκαν σε πληθυσμούς ευρέως τυχαία ανάμεικτους, αλλά στην πράξη κάθε άτομο έχει ένα ορισμένο αριθμό επαφών στις οποίες μπορεί να περάσει την μόλυνση. Το σύνολο όλων αυτών των επαφών διαμορφώνει ένα "ανάμεικτο δίκτυο". Η γνώση της δομής του δικτύου επιτρέπει στα μοντέλα να υπολογίσουν τη δυναμική της επιδημίας στην αναλογία του πληθυσμού από το επίπεδο της ατομικής συμπεριφοράς των μολύνσεων. Διάφορες μορφές δικτύων δημιουργημένων σε υπολογιστή έχουν μελετηθεί στα πλαίσια της μετάδοσης ασθενειών. Κάθε ένα από αυτά τα ιδανικά δίκτυα μπορούν να προσδιοριστούν από το πώς τα άτομα κατανέμονται στο χώρο (γεωγραφικά και κοινωνικά) και πώς γίνονται οι επαφές, συνεπώς απλοποιώντας και καταστώντας σαφή τις διάφορες σύνθετες διαδικασίες που εμπεριέχονται στη δημιουργία ενός δικτύου σε πραγματικούς πληθυσμούς.

Ο όρος "επιδημικός" (epidemic), έχει οριστεί σαν "ένα ξέσπασμα μιας μεταδοτικής ασθένειας η οποία εξαπλώνεται ταχύτατα και σε ένα μεγάλο εύρος όσον αφορά τη περιοχή μόλυνσης". Με ένα παρόμοιο λοιπόν τρόπο, η επιδημιολογία στην επιστήμη των υπολογιστών μπορεί να οριστεί σαν "ένα ιός υπολογιστή ή ένα σκουλήκι που διαδίδεται γρήγορα και σε μεγάλο βαθμό μολύνοντας υπολογιστικά συστήματα σε μία περιοχή ή σε ένα πληθυσμό ταυτόχρονα" (Symantec 2000). Η επιδημιολογία στην επιστήμη των υπολογιστών, μελετήθηκε αρχικά από τους Kephart, Chess και White, οι οποίοι περιέγραψαν το τρόπο με τον οποίο διαδίδονται τα computer worms/viruses. Βρέθηκε ότι υπάρχουν αρκετές αναλογίες στο τρόπο που διαδίδεται μία επιδημία αν το μελετά κάποιος από τη σκοπιά της βιολογίας ή τη σκοπιά της επιστήμης των υπολογιστών [15].

1.2α΄ Επιδημιολογία στη βιολογία και στους υπολογιστές

Η επιδημιολογία σε γενικές γραμμές προσπαθεί να εξηγήσει αυτό που συμβαίνει στο ανθρώπινο είδος και κυρίως στις ασθένειες που εμφανίζονται. Γενικά μπορούμε να πούμε ότι περιγράφει μία επιστημονική μεθοδολογία στη βιολογία ώστε να μελετήσει τη φύση, την υπερίσχυση και τις αιτίες εκείνες που προκαλούν μια ασθένεια. Αυτή η επιστήμη, παρέχει μία μέθοδο για την κατανόηση αλλά και ανταπόκριση σε μία ασθένεια καθώς αυτή διαδίδεται σε ένα πληθυσμό. Η επιδημιολογία χρησιμοποιεί μαθηματικά μοντέλα για να ποσοτικοποιήσει, χαρακτηρίσει, και να προβλέψει τη διάδοση και επίδραση μιας ασθένειας. Η δημογραφική ανάλυση που συνήθως εφαρμόζεται, χρησιμοποιείται για να καθορίσει τη σχέση μεταξύ της ασθένειας και του πληθυσμού. Ο ρόλος αυτών που ασχολούνται με την επιδημιολογία είναι να καταστρέψουν ή να βλάψουν αυτή τη σχέση ώστε να

προλάβουν τη μόλυνση του πληθυσμού. Ο βασικός στόχος της επιδημιολογίας είναι να εμποδίσει τη διάδοση της ασθένειας και να προλάβει πιθανή μελλοντική επανεμφανισή της.

Η λεγόμενη ψηφιακή επιδημιολογία, εφαρμόζει την βιολογική επιδημιολογία που είδαμε παραπάνω στον κυβερνοχώρο, και γενικότερα σε αυτό που ονομάζουμε σήμερα επιστήμη των υπολογιστών. Οι διαχειριστές δικτύων και συστημάτων, καθώς και οι ερευνητές, αντιλήφθηκαν ότι η ασφάλεια ενός συστήματος εξαρτάται από την ασφάλεια ολόκληρου του πληθυσμού όπου μπορεί να περιλαμβάνει το υποδίκτυο, το πανεπιστήμιο ή το δίκτυο της εταιρείας, ή ακόμη και ολόκληρο το διαδίκτυο. Τεχνικές από την βιολογική επιδημιολογία προσφέρουν μεθόδους προκειμένου να κατανοήσουμε και να αντιμετωπίσουμε τα θέματα ασφάλειας που απειλούν την υγεία αυτού του πληθυσμού [5, 6].

Ομοιότητες μεταξύ βιολογικών και ψηφιακών ασθενειών

Τα worms διαδίδονται μεταξύ των δικτύων υπολογιστών με το να εισβάλουν στα συστήματα και μετά να διαδίδονται σε σε άλλα συστήματα. Αυτή η διαδικασία, μπορεί να παραλληλιστεί με τη διάδοση μιας μολυσματικής ασθένειας στη βιολογία, όπου μια ξένη οντότητα διαδίδεται μέσω του πληθυσμού με το να μολύνει κάποιο άτομο, το οποίο με τη σειρά του μολύνει και άλλα. Η χρήση αυτής της επιστημονικής μεθόδου, εστίασε στις βασικές αρχές της επιδημιολογίας με ενδιαφέρον στη μετάδοση μιας ασθένειας σε ένα ζωντανό πληθυσμό. Η προσαρμογή αυτής της μεθόδου ξεκίνησε με τη χαρτογράφηση της βιολογικής ορολογίας σε ένα ψηφιακό παράδειγμα όπως μπορούμε να δούμε και στον παρακάτω πίνακα 1.1. Οι 5 πρώτοι όροι- πληθυσμός, οργανισμός, ασθένεια, περιβάλλον και μεταφορά, περιγράφουν τα βασικά στοιχεία της διάδοσης μιας ασθένειας. Οι όροι υγεία, αρρώστια και σύμπτωμα περιγράφουν την εκδήλωση μιας ασθένειας και διευκολύνουν την ανίχνευση και αναγνώριση. Η διάγνωση, η πρόγνωση και η θεραπεία, οι 3 τελευταίοι όροι του πίνακα, καθορίζουν μία διαδικασία ανίχνευσης, κατανόησης, πρόβλεψης και ελέγχου μιας ασθένειας. Η επιδημιολογία βασικά ενδιαφέρεται για την υγεία των οντοτήτων, ειδικότερα των ζωντανών οργανισμών. Όπως ακριβώς και το πεδίο της ασφάλειας των συστημάτων και δικτύων που σκοπό έχει να διατηρεί τη φυσιολογική λειτουργία και διαθεσιμότητα των υπολογιστικών συσκευών.

Ομάδες από τέτοιες συσκευές σχηματίζουν ένα δίκτυο ενώ στη βιολογία οι οντότητες σχηματίζουν ένα πληθυσμό. Μια βιολογική ασθένεια, επιδρά εχθρικά στην υγεία των οντοτήτων όπως και το κακόβουλο λογισμικό επηρεάζει τη κανονική συμπεριφορά των υπολογιστικών συστημάτων. Η ψηφιακή ασθένεια, στη περίπτωση ενός network stelth worm, μεταδίδεται με βάση τις υπάρχουσες δικτυακές συνδέσεις ενώ σε μία βιολογική ασθένεια, η μόλυνση μπορεί να μεταφέρεται από τον αέρα, το έδαφος ή από άλλους οργανισμούς. Μια βιολογική ασθένεια και μία οντότητα, μοιράζονται ένα περιβάλλον που χαρακτηρίζεται από τη θερμοκρασία, την υγρασία, τις κοινωνικές αλληλεπιδράσεις, και άλλα παραπλήσια στοιχεία. Η ύπαρξη τρωτών σημείων, οι τα αντιδραστικά μέτρα άμυνας, το ανοιχτό μοντέλο επικοινωνίας όπως είναι το διαίκτυο, και η έλλειψη ποικιλίας συνθέτουν το περιβάλλον ενός κακόβουλου λογισμικού. Η κανονική /φυσιολογική συμπεριφορά μιας οντότητας είναι γνωστή σαν υγεία ενώ σε ένα υπολογιστικό σύστημα αυτό χαρακτηρίζεται σαν κανονική λειτουργία. Ένα τέτοιο σύστημα βρίσκεται σε αντικανονική λειτουργία όταν προκληθεί μια εχθρική απόκλιση από τη κανονική λειτουργία, μια κατάσταση όπου είναι γνωστή σαν ασθένεια στους βιολογικούς οργανισμούς. Ένας ζωντανός οργανισμός εμφανίζει κάποιο σύμπτωμα, παρό-

μοιο με την ανωμαλία σε μία ψηφιακή υπολογιστική συσκευή, όπου και δείχνει μία ασθένεια ή αντικανονική λειτουργία. Η αναγνώριση και καθορισμός των παραγόντων μιας ασθένειας προήλθε σαν αποτέλεσμα της αυτοψίας και μελέτης των δεδομένων. Αυτό είναι ανάλογο με την ανίχνευση και χαρακτηρισμό του κακόβουλου λογισμικού σε ένα υπολογιστικό σύστημα που έχει αντικανονική λειτουργία. Η ανάλυση και πρόβλεψη συνέβαλε στη κατανόηση των επιδράσεων και αποτελεσμάτων του κακόβουλου λογισμικού σε ένα σύστημα, όπου είναι παρόμοια με τη πρόγνωση που γίνεται σε ένα βιολογικό οργανισμό που έχει μολυνθεί από μία ασθένεια. Ο σκοπός της πρόγνωσης είναι τελικά να αναγνωρισθεί η θεραπεία για την ασθένεια προκειμένου ο οργανισμός να επιστρέψει στην υγειή κατάσταση που βρισκόταν. Στο ψηφιακό κόσμο έχουμε τη λεγόμενη αντίδραση προκειμένου να απαλείψουμε το κακόβουλο λογισμικό και το σύστημα να αναρώσει.

Διαφορές μεταξύ βιολογικών και ψηφιακών ασθενειών

Υπάρχουν αρκετές βασικές διαφορές μεταξύ μιας βιολογικής ασθένειας και αυτής που προκαλείται από κακόβουλο λογισμικό. Μία βιολογική ασθένεια τυπικά προσβάλει ένα πληθυσμό σε ένα διάστημα ημερών και εβδομάδων ή ακόμη και δεκάδων ετών. Στη περίπτωση όμως ενός κακόβουλου λογισμικού, η μόλυνση διαδίδεται σε δευτερόλεπτα, ή και ώρες, με λίγα παραδείγματα απαιτούν μήνες. Η περίοδος επώασης μιας βιολογικής ασθένειας είναι πολύ μεγαλύτερη από αυτή μιας ψηφιακής ενώ και το μέσο διάδοσης διαφέρει δραστικά. Τέτοιες διαφορές προσθέτουν επιπρόσθετους περιορισμούς στην λήψη δραστικών και αποτελεσματικών μέτρων για την αντιμετώπιση της μόλυνσης από κάποιο κακόβουλο λογισμικό.

1.2β΄ Εισαγωγή στα επιδημιολογικά μοντέλα

Ένα επιδημιολογικό μοντέλο είναι ένα αρκετά καλό εργαλείο προκειμένου να κατανοήσουμε τη διάδοση μιας μόλυνσης σχετίζοντας τη διαδικασία διάδοσης με τις ιδιότητες που μπορεί να έχει ένας ξενιστής. Ωστόσο, τα επιδημιολογικά μοντέλα δεν είναι τόσο εύκολο να εφαρμοστούν αλλά και να είμαστε σίγουροι για τα αποτελέσματα τους διότι:

(i) Τα αποτελέσματα εξαρτώνται από κάποιους ισχυρισμούς οι οποίοι είναι σπανίως ακριβείς.

(ii) Εξαρτώνται από τις τιμές κάποιων παραμέτρων, όπως είναι για παράδειγμα ο αριθμός του πληθυσμού που έχουμε και των επαφών μεταξύ τους, και οι οποίοι είναι απλά ισχυρισμοί.

(iii) Τα επιδημικά όρια τα οποία προκύπτουν είναι αρκετά ισχυρά και συνήθως παρατηρούνται εύκολα εξαιτίας των τιμών που δίνουμε σε αρκετές παραμέτρους.

Για να κάνουμε την επιδημική διαδικασία δυνατή ή ευκολότερη θα πρέπει να μην είναι πολύπλοκη και να είναι εύκολη και σαφή όσον αφορά τη κατανόηση. Ένας επιδημικός αλγόριθμος ασχολείται με πληθυσμό που μπορεί να αναπαρασταθεί από ένα σύνολο ατομικών οντοτήτων όπου αλληλεπιδρούν μεταξύ τους με βάση κάποιους κανόνες, και οι οποίοι κανόνες έχουν ένα σημαντικότατο ρόλο όσον αφορά τη διάδοση μιας πληροφορίας. Αυτές τώρα οι οντότητες (hosts) θα πρέπει να έχουν κάθε στιγμή μία από τις παρακάτω τρεις καταστάσεις:

(i) Ευάλωτος (susceptible): Ο χρήστης δεν έχει ιδέα σχετικά με τη συγκεκριμένη πληροφορία (virus, worm), αλλά έχει τη δυνατότητα να τη δεχθεί, με άλλα λόγια δηλαδή να μολυνθεί.

(ii) Μολυσματικός (infective): Ο χρήστης γνωρίζει πλέον για αυτή τη πληροφορία, έχει δηλαδή μολυνθεί, και είναι σε θέση να μολύνει και άλλους με το να διαδώσει αυτή τη πληροφορία, σε άλλους ευάλωτους, και οι οποίοι δεν έχουν γίνει ακόμη θύματα.

(iii) Κατάσταση ανάρρωσης (recovered): Ο χρήστης γνωρίζει για τη συγκεκριμένη πληροφορία, και πλέον δεν μπορεί να μολύνει και άλλους, ή να ξαναμολυνθεί αργότερα.

Γενικά τα επιδημιολογικά μοντέλα μπορούμε να τα χωρίσουμε σε δύο μεγάλες κατηγορίες: στα στοχαστικά και στα ντετερμινιστικά. Τα στοχαστικά μοντέλα χρησιμοποιούνται συνήθως για ένα μικρό ή απομονωμένο πληθυσμό, και αυτό διότι επικεντρώνουν τη προσοχή τους σε κάθε χρήστη. Τα μοντέλα αυτά απαιτούν αρκετή εργασία προκειμένου να παραχθεί ένα αποτέλεσμα το οποίο να επιβεβαιώνει τις προβλέψεις που είχαν γίνει. Επίσης αυτά τα μοντέλα είναι δύσκολα στην κατανόηση και έχουν και πολύπλοκα μαθηματικά. Τα ντετερμινιστικά μοντέλα από τη άλλη πλευρά, χρησιμοποιούνται κυρίως σε μεγάλους πληθυσμούς, και προσπαθούν να μας πουν τι γίνεται στο μέσο όρο του πληθυσμού αυτού, με βάση κάποιες αρχικές συνθήκες και καταστάσεις. Αυτά τα μοντέλα τοποθετούν τους χρήστες σε υποκατηγορίες ή καλύτερα σε καταστάσεις. Στην παρούσα εργασία θα ασχολειθούμε με τέτοιου είδους μοντέλα.

Η μετάβαση από τη μία κατάσταση στην άλλη, συμβαίνει με κάποιο ρυθμό, για παράδειγμα ο ρυθμός μόλυνσης είναι ένας αρκετά γνωστός παράγοντας όπου ωθεί τους χρήστες που βρίσκονται σε κατάσταση susceptible να μεταβούν στη κατάσταση infected. Όταν ξεσπά μία επιδημία, ακριβώς επειδή οι χρήστες αλληλεπιδρούν μεταξύ τους, αυτό έχει σαν αποτέλεσμα με το πέρασμα του χρόνου να υπάρχουν αλλαγές στις καταστάσεις που αυτοί βρίσκονται. Και όπως αναφέραμε και πιο πάνω, αυτή η μετάβαση γίνεται με κάποιο ρυθμό. Στην αρχή, κάθε χρήστης, κάθε οντότητα, μπορεί να θεωρηθεί ότι βρίσκεται στη κατάσταση Susceptible (S), με το πέρασμα του χρόνου ο αριθμός αυτός θα μειώνεται και θα αυξάνεται ο αριθμός των άλλων (Infected (I), Exposed (E), Recovered (R)) με κάποιους επιλεγμένους ρυθμούς. Υπάρχουν αρκετοί λόγοι οι οποίοι επηρεάζουν τη διάδοση μιας μόλυνσης:

(i) Ο αριθμός των μολυσμένων κόμβων εκείνη τη στιγμή.

(ii) Ο ρυθμός μόλυνσης.

(iii) Ο αριθμός των ευάλωτων κόμβων.

(iv) Το κατά πόσο ο πληθυσμός παρουσιάζει κάποια τρωτά σημεία.

(v) Τα επίπεδα ανοσίας.

(vi) Ο χρόνος όπου ένα μολυσμένος κόμβος μένει μολυσμένος.

(vii) Ο βαθμός αλληλεπίδρασης- συνδεσιμότητας μεταξύ των κόμβων.

1.2γ΄ Σκοποί και περιορισμοί της επιδημιολογικής μοντελοποίησης

Παρακάτω θα παρουσιάσουμε και θα αναφερθούμε στους σκοπούς που εξυπηρετεί η χρήση επιδημιολογικών μοντέλων, καθώς επίσης και στους περιορισμούς που θέτονται. Κάποιος δεν θα πρέπει να συμπεράνει από τον αριθμό των προθέσεων και των περιορισμών ότι τα πλεονεκτήματα υπερτερούν των μειονεκτημάτων, διότι όπως θα δούμε και παρακάτω ο πρώτος περιορισμός που αναφέρουμε καλύπτει μία αρκετά μεγάλη περιοχή. Θα αναφερθούμε στους περιορισμούς λίγο πιο αναλυτικά, μιας και τα πλεονεκτήματα είναι αρκετά συγκεκριμένα και δεν χρειάζεται να αναλυθούν περαιτέρω.

Βασικοί λόγοι επιδημιολογικής μοντελοποίησης:

(i) Το μοντέλο διαμόρφωσης της διαδικασίας που πάμε να περιγράψουμε, απλοποιεί και επεξηγεί τις διάφορες παραδοχές, μεταβλητές και παραμέτρους που ορίζουμε κάθε φορά.

(ii) Η συμπεριφορά της ακρίβειας των μαθηματικών μοντέλων που χρησιμοποιούμε μπορεί να αναλυθεί χρησιμοποιώντας μαθηματικές μεθόδους και εξομοιώσεις στον υπολογιστή.

(iii) Η μοντελοποίηση επιτρέπει την εξερεύνηση της επίδρασης που έχουν οι διάφοροι ισχυρισμοί που κάνουμε και οι διατυπώσεις.

(iv) Η μοντελοποίηση παρέχει κάποιες βασικές έννοιες όπως είναι τα όρια, οι αριθμοί αναπαραγωγής, κ.α.

(v) Η μοντελοποίηση είναι ένα πειραματικό εργαλείο προκειμένου να τσεκάρουμε κάποιες θεωρίες και να αποτιμήσουμε τις ποσοτικές εικασίες.

(vi) Μοντέλα με κατάλληλη πολυπλοκότητα μπορούν να κατασκευαστούν προκειμένου να απαντηθούν κάποιες συγκεκριμένες ερωτήσεις.

(vii) Η μοντελοποίηση μπορεί να χρησιμοποιηθεί προκειμένου να αξιολογηθούν κάποιες βασικές παράμετροι.

(viii) Τα μοντέλα παρέχουν δομές για οργάνωση, ένωση και διασταύρωση των διαφορετικών κομματιών πληροφορίας.

(ix) Τα μοντέλα μπορούν να χρησιμοποιηθούν προκειμένου να γίνουν συγκρίσεις κάποιων επιδημιών διαφορετικού τύπου, σε διαφορετικές χρονικές στιγμές και σε διαφορετικούς πληθυσμούς.

(x) Τα μοντέλα μπορούν να χρησιμοποιηθούν ώστε να γίνει μία θεωρητική αποτίμηση, σύγκριση ή και βελτίωση διαφόρων προγραμμάτων που σχετίζονται με την ανακάλυψη, την πρόληψη, τη θεραπεία και τον έλεγχο.

(xi) Τα μοντέλα μπορούν να χρησιμοποιηθούν προκειμένου να αποτιμηθεί η ευαισθησία των αποτελεσμάτων που προέρχονται από αλλαγή στις τιμές κάποιων παραμέτρων.

(xii) Η μοντελοποίηση μπορεί να προτείνει τη συλλογή κάποιων σημαντικών δεδομένων, που μπορεί να αγνοούσαμε.

(xiii) Η μοντελοποίηση μπορεί να συμβάλλει στο σχεδιασμό και ανάλυση κάποιων ερευνών σχετικά με την επιδημιολογία (κατά τη φάση του σχεδιασμού, η μοντελοποίηση μπορεί να συμβάλει στον εντοπισμό κάποιων σημαντικών ζητημάτων και ερωτήσεων που θα πρέπει να απαντηθούν ώστε να έχουμε επαρκή στοιχεία για επιτυχή αποτελέσματα).

(xiv) Τα μοντέλα μπορούν να χρησιμοποιηθούν προκειμένου να αναγνωριστούν νέες κατευθύνσεις σε αυτό το τομέα, να γίνουν γενικές προβλέψεις ή ακόμη και να εκτιμηθεί η αβεβαιότητα κάποιων προβλέψεων (γενικά δίδεται η δυνατότητα να γίνουν προβλέψεις σχετικά με τις μελλοντικές επιπτώσεις μιας επιδημίας. Αν και οι άνθρωποι συχνά θεωρούν ότι η πρόβλεψη είναι ο αρχικός ή μοναδικός λόγος της επιδημιολογικής μοντελοποίησης, ο λόγος που αναφέρουμε τώρα είναι πιο σημαντικός).

(xv) Η αξιοπιστία και η ευρωστία των αποτελεσμάτων της μοντελοποίησης μπορεί να καθοριστεί χρησιμοποιώντας ένα εύρος από τιμές όσον αφορά τις παραμέτρους σε πολλά διαφορετικά μοντέλα.

Βασικοί περιορισμοί της επιδημιολογικής μοντελοποίησης.

(i) Ένα επιδημιολογικό μοντέλο δεν είναι κάτι το πραγματικό. Είναι μία υπεραπλούστευση της πραγματικότητας.

(ii) Τα ντετερμινιστικά μοντέλα δεν αντικατοπτρίζουν το ρόλο της τύχης και της πιθανότητας στη διάδοση μιας επιδημίας.

(iii) Τα στοχαστικά μοντέλα εισάγουν την έννοια της τύχης, αλλά είναι συνήθως δυσκολότερα στην ανάλυση σε σχέση με τα ντετερμινιστικά μοντέλα, εξαιτίας των πολύπλοκων μαθηματικών σχέσεων.

Σχετικά τώρα με τους περιορισμούς που ανακύπτουν σε ότι αφορά την επιδημιολογική μοντελοποίηση, έχουμε να αναφέρουμε τα εξής: Επειδή τα μοντέλα διάδοσης είναι απλουστεύσεις, με συνήθως άγνωστες σχέσεις σε σχέση με τις μολύνσεις και τις επιδημίες, δεν μπορεί να είναι ποτέ κανένας σίγουρος σχετικά με τα αποτελέσματα, τις προβλέψεις, τις συγκρίσεις, κ.α. Ακόμη και όταν τα μοντέλα γίνονται πιο πολύπλοκα προκειμένου να πλησιάσουν ακόμη περισσότερο μία επιδημία, εξακολουθούν να είναι μία αφηρημένη έννοια. Αυτός ο οποίος αναπτύσει ένα μοντέλο θα πρέπει να ασκεί συνεχώς την κρίση του ώστε να είναι σε θέση να αποφασίζει ποιοι παράγοντες είναι σχετικοί και ποιοι όχι όταν αναλύει μία μόλυνση ή απαντά κάποιες βασικές ερωτήσεις. Τα ντετερμινιστικά μοντέλα είναι αυτά τα οποία χρησιμοποιούν διαφορικές εξισώσεις προκειμένου να περιγράψουν τις μεταβολές σε σχέση με το χρόνο των μεγεθών κάποιων πληθυσμών. Έχοντας κάποιες αρχικές συνθήκες για ένα καλό-προσφερόμενο ντετερμινιστικό μοντέλο, οι λύσεις σαν συνάρτηση του χρόνου είναι μοναδικές. Στα στοχαστικά μοντέλα, υπάρχουν πιθανότητες σε κάθε βήμα του χρόνου ώστε να πάμε από τη μία κατάσταση στην άλλη. Τα απλά ντετερμινιστικά μοντέλα για επιδημιολογίες έχουν ένα ακριβές όριο το οποίο καθορίζει αν μία επιδημία θα συμβεί ή όχι. Σε αντίθεση, τα στοχαστικά μοντέλα για επιδημίες, εισάγουν κάποιες ποσότητες όπως είναι η πιθανότητα του να συμβεί μία επιδημία ή ο χρόνος που αυτή θα εξαλειφτεί. Οπότε η διαδικασία, οι έννοιες, οι προσεγγίσεις αλλά και οι κατάλληλες ερωτήσεις ώστε να προκύψουν οι απαντήσεις είναι σχετικά διαφορετικές για τα στοχαστικά μοντέλα. Και τα ντετερμινιστικά και τα στοχαστικά μοντέλα, έχουν και άλλους

περιορισμούς πέρα από το γεγονός ότι είναι απλουστεύσεις της πραγματικότητας. Τα ντετερμινιστικά μοντέλα δεν λαμβάνουν υπόψη το ρόλο της τύχης στη διάδοση μιας μόλυνσης. Μερικές φορές οι τιμές κάποιων παραμέτρων στα ντετερμινιστικά μοντέλα θέτονται να είναι ίσες με τη μέση τιμή των παρατηρούμενων τιμών και αγνοούνται κάποιες άλλες πληροφορίες. Ένα σύνολο από αρχικές συνθήκες οδηγεί σε μία μόνο λύση σε ένα ντετερμινιστικό μοντέλο. Οπότε δεν είναι διαθέσιμες κάποιες πληροφορίες σχετικά με την αξιοπιστία των αποτελεσμάτων.

Βιολογικό Παράδειγμα	**Ψηφιακό Παράδειγμα**	**Εξήγηση**
Πληθυσμός	Δίκτυο	Το πλήρες σύνολο των οντοτήτων που είναι υπό εξέταση
Οργανισμός	Υπολογιστική Συσκευή	Η οντότητα μέσα στο σύνολο που είναι υπό εξέταση.
Ασθένεια	Κακόβουλο λογισμικό ή κακόβουλος χρήστης	Αντίδραση της οντότητας στην εισβολή ή επιρροή μιας ξένης υπόστασης που επηρέαζει τη κανονική κατάσταση ή συμπεριφορά.
Μέσο Μετάδοσης: αέρας, έδαφος, οργανισμοί, κ.α.	Μέσο Μετάδοσης: δικτυακές συνδέσεις.	Μέθοδος ή μηχανισμός διάδοσης της ασθένειας.
Περιβάλλον: θερμοκρασία, υγρασία, αλληλεπιδράσεις οργανισμών, κ.α.	Περιβάλλον: ύπαρξη τρωτών σημείων, μέτρα άμυνας, κ.α.	Το περιβάλλον όπου η ασθένεια και ο χρήστης συνυπάρχουν.
Ευρωστία	Κανονική λειτουργία	Η κατάσταση κανονικής λειτουργίας και συμπεριφοράς ενός οργανισμού.
Πάθηση/Ασθένεια	Αντικανονική λειτουργία	Εχθρική απόκλιση από τη κανονική λειτουργία ή συμπεριφορά σαν αποτέλεσμα της ασθένειας.
Σύμπτωμα	Ανωμαλία	Σημάδι ή ένδειξη μιας ασθένειας ειδικότερα όταν δείχνει μία εχθρική απόκλιση από τη κανονική λειτουργία ή συμπεριφορά.
Διάγνωση	Ανακάλυψη και χαρακτηρισμός.	Αναγνώριση και καθορισμός της φύσης και αιτίας της ασθένειας μέσα από την αξιολόγηση των χαρακτηριστικών αυτών που νοσούν και διαφόρων άλλων δεδομένων.
Πρόγνωση	Ανάλυση και πρόβλεψη	Πρόβλεψη της πιθανής πορείας μιας ασθένειας και των αποτελεσμάτων αυτής.
Θεραπεία	Αντίδραση	Παροχή θεραπείας σε μία μολυσμένη οντότητα.

Πίνακας 1.1: Πίνακας αντιστοιχήσεων

Κεφάλαιο 2

Ανάλυση μοντέλου διάδοσης ιών τριών πληθυσμών

Σε αυτό το κεφάλαιο προτείνουμε και αναλύουμε ένα μαθηματικό μοντέλο για την ταυτόχρονη εξέλιξη των πληθυσμών των ιών και των λογισμικών προστασίας έναντι ιών σε ένα μολυσμένο δίκτυο υπολογιστών. Στην συνέχεια θα επεκτείνουμε αυτό το μοντέλο σε ένα που θα περιέχει και ένα τρίτο είδος, τις παγίδες (traps), τα οποία είναι πανίσχυρα λογισμικά προστασίας. Η διάκριση αυτών των δύο κατηγοριών λογισμικών προστασίας ωφείλεται στο γεγονός ότι τα "απλά" λογισμικά προστασίας διαδίδονται μέσα στους κόμβους του δικτύου, δηλαδή σε άλλους υπολογιστές, ενώ οι παγίδες παραμένουν ακίνητες στους υπολογιστές που έχουν ήδη εγκατασταθεί. Αρχικά θα μοντελοποιήσουμε ένα σύστημα με δύο πληθυσμούς των ιών και των λογισμικών προστασίας και αναλύοντας το θεωρητικά θα δείξουμε ότι οι προβλέψεις μας συμφωνούν με παρατηρήσεις πραγματικών διαδόσεων ιών. Στην συνέχεια με την προσθήκη των παγίδων, θα καταλήξουμε σε ένα μη- γραμμικό σύστημα τριών διαφορικών εξισώσεων. Αναλύοντας το μοντέλο θεωρητικά χρησιμοποιώντας ανάλυση ιδιοτιμών για πρόβλεψη ευστάθειας, θα δείξουμε με αριθμητικά αποτελέσματα αυτή την φορά ότι οι προβλέψεις μας και πάλι συμφωνούν με πραγματικές παρατηρήσεις.

2.1 Μοντέλο διάδοσης δύο πληθυσμών

Μια πρώτη προσπάθεια για την δημιουργία και ανάλυση ενός μαθηματικού μοντέλου που θα περιγράφει την εξέλιξη των πληθυσμών των ιών και των λογισμικών προστασίας σε ένα δίκτυο υπολογιστών. Αρχικά θα προσπαθήσουμε να προσεγγίσουμε το πρόβλημα με δύο μόνο εξισώσεις, μία που θα περιγράφει την μεταβολή του πλήθους των ιών, και μία των λογισμικών προστασίας. Για την κατασκευή του μοντέλου υποθέτουμε τα παρακάτω:

(i) Το δίκτυο αποτελείται από άπειρο το πλήθος υπολογιστές.

(ii) Τα λογισμικά προστασίας εξουδετερώνουν με κάποια πιθανότητα τους ιούς όταν συνεβρεθούν στον ίδιο κόμβο (υπολογιστή).

(iii) Το πλήθος των ιών αυξάνεται με την πάροδο του χρόνου.

(iv) Τα λογισμικά προστασίας παράγονται μόνο όταν εμφανιστούν ιοί.

(v) Και τα δύο είδη καταστρέφονται με την πάροδο του χρόνου.

2.1α΄ Μια διαισθητική περιγραφή

Αρχικά έχουμε ένα δίκτυο υπολογιστών και κάποιοι υπολογιστές έχουν εγκαταστημένα προγράμματα προστασίας έναντι ιών ενώ σε κάποιους άλλους υπάρχουν ιοί. Στην πάροδο του χρόνου δημιουργούνται νέοι ιοί και αποστέλλονται σε κάποιους υπολογιστές οπότε αυξάνονται στο πλήθος. Αντίθετα τα λογισμικά προστασίας καταστρέφονται στην πάροδο του χρόνου. Αυτό ερμηνεύεται ότι τα ήδη υπάρχοντα λογισμικά προστασίας δεν μπορούν να αντιμετωπίσουν τους νέους ιούς εάν δεν έχουμε κάνει ενημέρωση του λογισμικού (update) και έτσι οι υπολογιστές αυτοί συμπεριφέρονται σαν να μην έχουν εγκατεστημένο πρόγραμμα προστασίας. Όταν όμως εμφανιστούν πολλοί ιοί και γίνει αντιληπτό απο τους χρήστες, τότε αρχίζει η εγκατάσταση καινούριων λογισμικών προστασίας ή update εκδόσεων των ήδη υπαρχόντων και έτσι έχουμε μία αύξηση στο πλήθος αυτών των υπολογιστών. Αλλά και σ΄ αυτήν την περίπτωση μπορεί κάποια λογισμικά προστασίας να μην αντιμετωπίζουν ορισμένους ιούς και γι΄ αυτό έχουμε εισάγει στο μοντέλο μας μία πιθανότητα εξουδετέρωσης. Στο μοντέλο μας θεωρούμε ότι έχουμε εκθετική αύξηση του πλήθους των ιών όταν απουσιάζουν λογισμικά προστασίας ενώ απουσία ιών έχουμε εκθετική μείωση του πλήθους των antivirus λόγω οτι δεν υπάρχει ανάγκη για εγκαταστάσεις προγραμμάτων προστασίας ή update εκδόσεων. Είναι φανερό ότι οι μεταβολές και των δύο πληθυσμών είναι ανάλογες του γινομένου αυτών των πληθυσμών.

2.1β΄ Το μαθηματικό μοντέλο δυο πληθυσμών και η μαθηματική ανάλυση του

Το μαθηματικό μας μοντέλο είναι παρόμοιο με το γνωστό μαθηματικό μοντέλο Lotka-Volterra για δύο ανταγωνιστικά είδη [25]. Έστω $v(t)$ ο πληθυσμός των ιών και $a(t)$ ο πληθυσμός των λογισμικών προστασίας την χρονική στιγμή t, l_1 ο καθαρός ρυθμός αύξησης του πλήθους των ιών, l_2 ο ρυθμός καταστροφής των λογισμικών προστασίας και k η πιθανότητα εξουδετέρωσης των ιών. Και οι δύο ρυθμοί όπως φυσικά και η πιθανότητα είναι θετικές σταθερές. Έχουμε τις παρακάτω διαφορικές εξισώσεις που περιγράφουν την μεταβολή του πλήθους των δύο ειδών:

$$\frac{dv}{dt} = l_1 v - kav, \tag{2.1}$$

$$\frac{da}{dt} = kav - l_2 a \tag{2.2}$$

Πολλές φορές για να μελετήσουμε συστήματα διαφορικών εξισώσεων χρησιμοποιούμε ποιοτική θεωρία. Στόχος της ποιοτικής θεωρίας είναι η κατανόηση της συμπεριφοράς των λύσεων, χωρίς να γνωρίζουμε ακριβώς τις λύσεις. Πρόκειται για την προσέγγιση η οποία είναι γεωμετρικού χαρακτήρα, με στόχο την συλλογή όσο το δυνατόν περισσότερων ποιοτικών πληροφοριών, και όχι ποσοτικών που προκύπτουν όταν γνωρίζουμε ακριβώς τις λύσεις.

Ας δώσουμε πρώτα κάποιους χρήσιμους ορισμούς. Έστω ένα αυτόνομο διανυσματικό πεδίο

$$\dot{x} = F[x], \qquad x \in \Re^n. \tag{2.3}$$

Σημείο ισορροπίας ονομάζεται ένα σημείο $\bar{x} \in \Re^n$ τέτοιο ώστε $F(\bar{x}) = 0$, δηλαδή μία λύση η οποία δεν μεταβάλλεται στον χρόνο. Λέμε ότι ένα σημείο ισορροπίας είναι ευσταθές (stable) αν όλες οι λύσεις με αρχικές συνθήκες κοντά στο $\bar{x}$ παραμένουν κοντά στο $\bar{x}$. Αν οι λύσεις που ξεκινούν με αρχικές συνθήκες κοντά στο $\bar{x}$ συγκλίνουν στο $\bar{x}$ καθώς $t \to \infty$, η λύση λέγεται ασυμπτωτικά ευσταθής. Ένα σημείο ισορροπίας που δεν είναι ευσταθής λέγεται ασταθής [33].

Η έννοια της ευστάθειας είναι δομικό στοιχείο της ποιοτικής θεωρίας των διαφορικών εξισώσεων, τόσο από θεωρητικής πλευράς όσο και από την οπτική γωνία των εφαρμογών. Για παράδειγμα ο έλεγχος της αξιοπιστίας ενός μαθηματικού μοντέλου που χρησιμοποιείται για την μαθηματική περιγραφή μιας φυσικής διαδικασίας, θέτει το ερώτημα του κατά πόσον η συμεριφορά της λύσης μετά από μεγάλο χρονικά διάστημα είναι συμβατή με τις παρατηρήσεις, οι οποίες είναι δυνατόν να καταδεικνύουν κατάληξη του συστήματος σε κάποια κατάσταση ισορροπίας.

Για να προσεγγίσουμε την συμπεριφορά των λύσεων του συστήματος των μηγραμμικών διαφορικών εξισώσεων θα χρησιμοποιήσουμε την τεχνική της γραμμικοποίησης του συστήματος. Έστω ότι έχουμε το σύστημα

$$\begin{aligned} \dot{x} &= f(x,y) \\ \dot{y} &= g(x,y) \end{aligned}$$

και υποθέτουμε ότι το σημείο $(\bar{x},\bar{y})$ είναι σημείο ισορροπίας. Έστω $u = x - \bar{x}, v = y - \bar{y}$ είναι μια μικρή απόκλιση γύρω από το σημείο ισορροπίας. Για να δούμε αν η απόκλιση θα μεγαλώνει ή θα φθείνει, παραγωγίζουμε τις εξισώσεις για τα u και v. Για την πρώτη εξίσωση παίρνουμε:

$$\begin{aligned} \dot{u} &= \dot{x} \Rightarrow \text{ (αφού } \bar{x} \text{ είναι σταθερά)} \\ \dot{u} &= f(u+\bar{x}, v+\bar{y}) = \\ & f(\bar{x},\bar{y}) + u\frac{df}{dx} + v\frac{df}{dy} + O(u^2, v^2, uv) \Rightarrow \qquad (2.4) \\ & \text{(Taylor series expansion)} \\ \dot{u} &= u\frac{df}{dx} + v\frac{df}{dy} + O(u^2, v^2, uv) \qquad (2.5) \\ & \text{(αφού } f(\bar{x},\bar{y}) = 0\text{)} \end{aligned}$$

Για να απλοποιήσουμε τους συμβολισμούς, έχουμε γράψει $df/dx, df/dy$ αλλά πρέπει να θυμόμαστε ότι αυτές οι μερικές παράγωγοι είναι υπολογισμένες πάνω στο σημείο ισορροπίας $(\bar{x},\bar{y})$, οπότε είναι σταθερές και όχι συναρτήσεις. Επίσης στο δεξιό μέλος έχουμε έναν όρο $O(u^2, v^2, uv)$ που δηλώνει τους τετραγωνικούς όρους του u και v. Παρόμοια υπολογίζουμε:

$$\dot{v} = u\frac{dg}{dx} + v\frac{dg}{dy} + O(u^2, v^2, uv)$$

Οπότε η απόκλιση εξελίσσεται σύμφωνα με την σχέση:

$$\begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} = \partial F_{(x,y)} \begin{pmatrix} u \\ v \end{pmatrix} + \text{ τετραγωνικούς όρους.} \qquad (2.6)$$

Ο πίνακας

$$\partial F = \begin{pmatrix} \dfrac{df}{dx} & \dfrac{df}{dy} \\ \dfrac{dg}{dx} & \dfrac{dg}{dy} \end{pmatrix}_{(\bar{x},\bar{y})}$$

ονομάζεται ιακωβιανός πίνακας πάνω στο σημείο ισορροπίας $(\bar{x}, \bar{y})$. Παραλείποντας τους τετραγωνικούς όρους στην παραπάνω σχέση παίρνουμε την γραμμικοποίηση του συστήματος [29].

Για να πάρουμε συμπεράσματα για την ευστάθεια των σημείων ισορροπίας μελετάμε τα πρόσημα των ιδιοτιμών του ιακωβιανού πίνακα πάνω στα αντίστοιχα σημεία. Αν όλα τα πρόσημα των ιδιοτιμών έχουν αρνητικά πραγματικά μέρη, τότε το σημείο ισορροπίας του αρχικού μη-γραμμικού συστήματος είναι ασυμπτωτικά ευσταθές [33]. Αν τουλάχιστον μία ιδιοτιμή έχει θετικό πραγματικό μέρος, τότε το σημείο ισορροπίας είναι ασταθής. Αν τουλάχιστον μία ιδιοτιμή έχει μηδενικό πραγματικό μέρος, τότε δεν μπορούμε να εξάγουμε συμπέρασμα για την ευστάθεια του σημείου.

Ας δούμε τώρα και ένα άλλο εργαλείο για τον καθορισμό της ευστάθειας ενός σημείου ισορροπίας. Έστω το διανυσματικό πεδίο (2.3) και $\bar{x}$ ένα σημείο ισορροπίας. Έστω $V : U \to R$ μία C^1 συνάρτηση ορισμένη σε μια περιοχή U του $\bar{x}$, τέτοια ώστε $V(\bar{x}) = 0$ και $V(x) > 0$ αν $x \neq \bar{x}$, $\dot{V}(x) \leq 0$ στο $U - \{\bar{x}\}$. Τότε το $\bar{x}$ είναι ευσταθές. Ακόμα αν $\dot{V}(x) < 0$ στο $U - \{\bar{x}\}$ τότε το $\bar{x}$ είναι ασυμπτωτικά ευσταθές. Η συνάρτηση V ονομάζεται συνάρτηση Liapunov. Η εικόνα φάσεων είναι το σύνολο όλων των τροχιών που διέρχονται από τα σημεία ισορροπίας, όπου παίζουν τον ρόλο της αρχικής συνθήκης.

Τώρα είμαστε έτοιμοι να αναλύσουμε το δικό μας μοντέλο. Το σύστημα (2.1, 2,2) έχει δύο σημεία ισορροπίας, τα $(\bar{v}, \bar{a}) = (0, 0)$ και $(l_2/k, l_1/k)$. Θα γραμμικοποιήσουμε το σύστημα γύρω από τα σημεία ισορροπίας και θα προσπαθήσουμε να βγάλουμε συμπέρασματα για την ευστάθεια τους. Στο σημείο ισορροπίας $(0, 0)$ έχουμε τις εξής ιδιοτιμές από τον ιακωβιανό πίνακα l_1 και $-l_2$, και από την θεωρία των δυναμικών συστημάτων συμπεράινουμε ότι αυτό το σημείο είναι ασταθές. Αυτό είναι αναμενόμενο αφού οι ιοί αναπαράγονται με σταθερό ρυθμό και δεν μπορεί ποτέ ο πληθυσμός τους να φτάσει την τιμή 0 όταν είναι μόνα τους χωρίς αντιμετώπιση. Το ίδιο ισχύει και για τα λογισμικά προστασίας.

Θα εξετάσουμε τώρα το σημείο $(\bar{v}, \bar{a}) = (l_2/k, l_1/k)$. Οι ιδιοτιμές του ιακωβιανού πίνακα πάνω σε αυτό το σημείο είναι οι $l = \pm i\sqrt{l_1 l_2}$ και δεν μπορούμε να συμπεράνουμε ευστάθεια. Είμαστε στην κρίσιμη περίπτωση όπου έχουμε ουδέτερη ευστάθεια, $\operatorname{Re} l = 0$. Θεωρούμε την συνάρτηση

$$H(v, a) = l_2 \ln v + l_1 \ln a - k(v + a)$$

Το σημείο $(\bar{v}, \bar{a}) = (l_2/k, l_1/k)$ είναι τοπικό ελάχιστο για την συνάρτηση H. Οπότε η συνάρτηση $\hat{H}(v, a) = H(v, a) - H(\bar{v}, \bar{a})$ με $\hat{H}(\bar{v}, \bar{a}) = 0$, $\hat{H}(v, a) > 0$ σε μια περιοχή γύρω από το σημείο ισορροπίας και $\dot{\hat{H}}(v, a) = 0$, είναι συνάρτηση Liapunov για το σύστημα μας. Οπότε συμπεραίνουμε ότι το σημείο $(\bar{v}, \bar{a}) = (l_2/k, l_1/k)$ είναι ευσταθές. Για να δούμε την μορφή των λύσεων κάνουμε πρώτα αδιαστατοποίηση του συστήματος. Σύμφωνα με τους μετασχηματισμούς $r = tl_1$, $y = ka/l_1, x = kv/l_2$, $p = l_2/l_1$ το σύστημα γίνεται

$$\frac{dx}{dr} = x(1 - y)$$

$$\frac{dy}{dr} = py(x - 1)$$

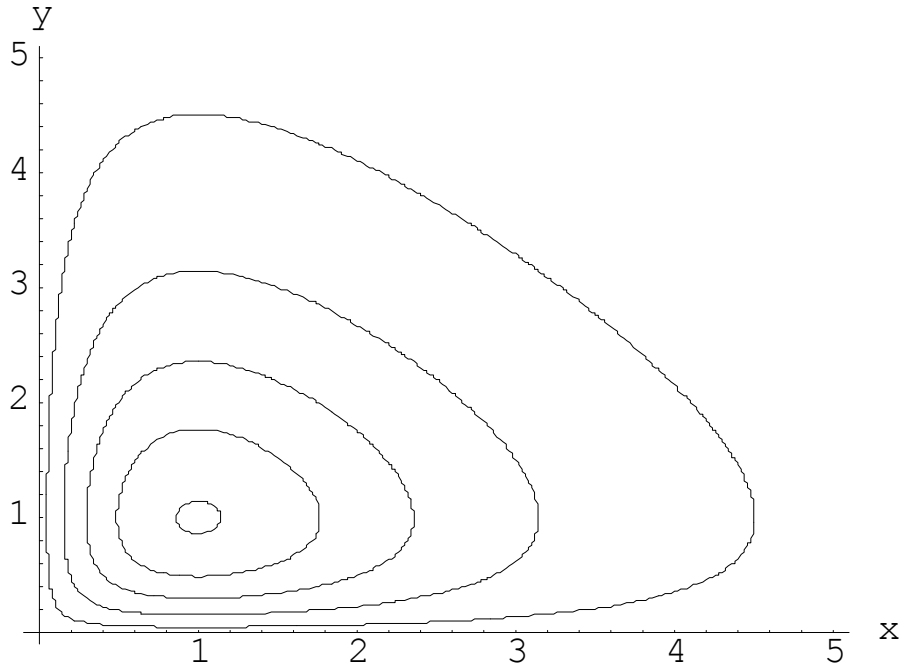
Το παραπάνω σύστημα έχει μόνο μία παράμετρο την p. Για να δούμε την εικόνα φάσεων ολοκληρώνουμε την εξίσωση

$$\frac{dy}{dx} = \frac{py(x - 1)}{x(y - 1)}$$

και παίρνουμε

$$px + y - \ln x^p y = C \ .$$

Αυτή η ποσότητα είναι σταθερή και έχει ελάχιστο στο σημείο ισορροπίας $(1, 1)$ την τιμή $C = 1 + p$.



Σχῆμα 2.1: Διάγραμμα φάσεων για τις διάφορες τιμές του C, όπου βλέπουμε ότι είναι περιοδικές. Αυτές οι τροχιές από μέσα προς τα έξω είναι λύσεις της παραπάνω εξίσωσης για τιμές της σταθεράς $C = 2.01, 2.2, 2.5, 3, 4$.

2.2 Μοντέλο διάδοσης τριών πληθυσμών

Όπως και στην περίπτωση των δύο εξισώσεων, θα παράγουμε και θα αναλύσουμε ένα μαθηματικό μοντέλο που θα περιγράφει την εξέλιξη των πληθυσμών των λογισμικών προστασίας (antivirus) και των ιών σε ένα δίκτυο υπολογιστών. Όμως σε αυτήν την περίπτωση έχουμε και ένα άλλο πλήθος λογισμικών παροστασίας που τα ονομάζουμε παγίδες, γι' αυτό το λόγο χρειαζόμαστε και την τρίτη διαφορική εξίσωση. Για την παραγωγή του μοντέλου κάνουμε τις παρακάτω υποθέσεις:

(i) Το δίκτυο αποτελείται από άπειρο το πλήθος υπολογιστές.

(ii) Τα λογισμικά προστασίας εξουδετερώνουν με κάποια πιθανότητα τους ιούς όταν συνεβρεθούν στον ίδιο κόμβο (υπολογιστή).

(iii) Τα traps εξουδετερώνουν πάντα τους ιούς όταν συνεβρεθούν στον ίδιο κόμβο.

(iv) Το πλήθος των ιών αυξάνεται στην πάροδο του χρόνου.

(v) Τα λογισμικά προστασίας και οι παγίδες παράγονται μόνο όταν εμφανιστεί κάποιος ιός αλλά οι παγίδες με πολύ αργότερο ρυθμό από τα λογισμικά προστασίας.

(vi) Όλα τα είδη καταστρέφονται στην πάροδο του χρόνου αλλά οι παγίδες με πολύ αργότερο ρυθμό από τα άλλα δύο.

2.2α΄ Μια διαισθητική περιγραφή

Σχεδόν επικρατεί η ίδια εικόνα με αυτήν στην περίπτωση των δύο εξισώσεων, με μόνη διαφορά την εμφάνιση των παγίδων. Οι παγίδες είναι κάποιοι υπολογιστές με πολύ βαριά λογισμικά προστασίας που εξουδετερώνουν οποιαδήποτε απειλή το οποίο βρεθεί σε αυτούς τους υπολογιστές. Αυτά τα λογισμικά κάνουν συνεχώς ενημερώσεις (update) γι΄ αυτό αντιμετωπίζουν πάντα κάθε απειλή εκτός μιας μικρής περίπτωσης όπου για κάποια αιτία ο administrator δεν ενημέρωσε το σύστημα και γι΄ αυτό το λόγο έχουν πολύ μικρότερη φθορά στον χρόνο σε σχέση με τα απλά λογισμικά προστασίας. Αυτή είναι και μία άλλη διαφορά των παγίδων με τα απλά λογισμικά προστασίας, ότι δηλαδή η αύξηση των υπολογιστών που λειτουργούν σαν παγίδες δεν επέρχεται σαν αλληλεπίδραση αυτών των υπολογιστών με τους ιούς όπως επέρχεται η αύξηση των υπολογιστών με απλά λογισμικά προστασίας, αλλά μόνο με απόφαση των administrators όταν εμφανίζονται πολλοί ιοί στο δίκτυο.

2.2β΄ Το μαθηματικό μοντέλο τριών πληθυσμών και η μαθηματική ανάλυση του

Έστω $v(t)$, $a(t)$, και $tr(t)$ το πλήθος των virus, των antivirus και των traps την χρονική στιγμή t, m_1 ο καθαρός ρυθμός αύξησης του πλήθους των ιών, m_2 ο ρυθμός καταστροφής των λογισμικών προστασίας, k η πιθανότητα εξουδετέρωσης των ιών από τα λογισμικά προστασίας, m_3 ο ρυθμός αύξησης των παγίδων και m_4 ο ρυθμός φθοράς των παγίδων. Όλοι οι ρυθμοί όπως φυσικά και η πιθανότητα είναι θετικές σταθερές. Έχουμε τις παρακάτω διαφορικές εξισώσεις που περιγράφουν την μεταβολή του πλήθους αυτών των ειδών:

$$\begin{aligned} \frac{dv}{dt} &= m_1 v - kav - trv \\ \frac{da}{dt} &= kav - m_2 a \\ \frac{dtr}{dt} &= m_3 v - m_4 tr \end{aligned}$$

Για την μελέτη αυτού του μοντέλου θα κάνουμε πρώτα αδιαστατοποίηση σύμφωνα με τους ακόλουθους μετασχηματισμούς:

$$r = tm_1, x = \frac{k}{m_1} v, y = \frac{k}{m_1} a, z = \frac{tr}{m_1}, g_1 = \frac{m_2}{m_1}, g_2 = \frac{m_3}{km_1}, g_3 = \frac{m_4}{m_1}$$

όπου $g_1, g_2, g_3 > 0$. Συνεπώς οι διαφορικές μας εξισώσεις γίνονται:

$$\frac{dx}{dr} = x - xy - xz$$

(2.7)
$$\frac{dy}{dr} = yx - g_1 y$$
$$\frac{dz}{dr} = g_2 x - g_3 z$$

Το παραπάνω σύστημα έχει τα παρακάτω τρία σημεία ισορροπίας:

$$(\bar{x}, \bar{y}, \bar{z}) = (0,0,0), (g_3/g_2, 0, 1) \text{ και } (g_1, 1 - g_1 g_2/g_3, g_1 g_2/g_3)\ .$$

Κάνοντας γραμμικοποίηση του παραπάνω συστήματος αποκτούμε τον εξής ιακωβιανό πίνακα:

$$\begin{aligned} f_1(x,y,z) &= x - xy - xz \\ f_2(x,y,z) &= xy - g_1 y \\ f_3(x,y,z) &= g_2 x - g_3 z \end{aligned}$$

όπου

$$\partial F = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial z} \end{pmatrix} = \begin{pmatrix} 1-y-z & -x & -x \\ y & x - g_1 & 0 \\ g_2 & 0 & -g_3 \end{pmatrix}$$

Στο σημείο ισορροπίας $(\bar{x}, \bar{y}, \bar{z}) = (0,0,0)$ έχουμε το παρακάτω πίνακα του γραμμικοποιημένου συστήματος

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -g_1 & 0 \\ g_2 & 0 & -g_3 \end{pmatrix}$$

ο οποίος έχει ιδιοτιμές τις $1, -g_1, -g_3$ και άρα συμπεραίνουμε ότι αυτό το σημείο ισορροπίας είναι ασταθές.
Για το δεύτερο σημείο ισορροπίας $(g_3/g_2, 0, 1)$ έχουμε τις εξής ιδιοτιμές από τον αντίστοιχο πίνακα του γραμμικοποιημένου συστήματος:

$$\{\frac{1}{2}(-\sqrt{g_3 - 4}\sqrt{g_3} - g_3), \frac{1}{2}(\sqrt{g_3 - 4}\sqrt{g_3} - g_3), -g_1 + \frac{g_3}{g_2}\}$$

Επειδή $m_4 < m_1$ διότι έχουμε πολύ αργή φθορά των παγίδων σε σχέση με την αύξηση των ιών έχουμε: $g_3 < 1$. Άρα οι δύο πρώτες ιδιοτιμές έχουν πάντα αρνητικά πραγματικά μέρη. Για την τρίτη ιδιοτιμή αν ισχύει $g_3 < g_1 g_2$, τότε παίρνει αρνητική τιμή και το σημείο ισορροπίας είναι ευσταθές. Αν ισχύει ότι $g_3 > g_1 g_2$, τότε η τρίτη ιδιοτιμή είναι θετική οπότε το σημείο ισορροπίας είναι ασταθές. Αν $g_3 = g_1 g_2$ τότε η ιδιοτιμή μηδενίζεται και έχουμε ουδέτερη ευστάθεια, δηλαδή δεν μπορούμε να εξάγουμε συμπέρασμα για την ευστάθεια του δεύτερου σημείου ισορροπίας.
Για το τρίτο σημείο ισορροπίας $(\bar{x}, \bar{y}, \bar{z}) = (g_1, 1 - g_1 g_2/g_3, g_1 g_2/g_3)$ έχουμε τις εξής περιπτώσεις:

(i) Αν $g_3 < g_1 g_2$, τότε η δεύτερη συντεταγμένη γίνεται αρνητική και το τρίτο σημείο ισορροπίας δεν ανήκει στο χώρο που μελετάμε. Δηλαδή όταν έχουμε ευστάθεια του δευτέρου σημείου ισορροπίας τότε δεν υπάρχει τρίτο σημείο ισορροπίας.

(ii) Αν $g_3 = g_1 g_2$, δηλαδή όταν έχουμε ουδέτερη ευστάθεια στο δεύτερο σημείο ισορροπίας, τότε το τρίτο σημείο ισορροπίας γίνεται $(\bar{x}, \bar{y}, \bar{z}) = (g_1, 0, 1)$, όπου ταυτίζεται με το δεύτερο σημείο ισορροπίας. Ακόμα από τον πίνακα γραμμικοποίησης παίρνουμε τις εξής ιδιοτίμες:

$$\{0, \frac{1}{2}(-g_3 - \sqrt{-4g_1g_2 + g_3^2}), \frac{1}{2}(-g_3 + \sqrt{-4g_1g_2 + g_3^2})\}$$

όπου καταλήγουμε στο ίδιο συμπέρασμα για την ουδέτερη ευστάθεια.

(iii) Αν $g_3 > g_1 g_2$, δηλαδή όταν το δεύτερο σημείο ισορροπίας είναι ασταθές, τότε έχουμε τις παρακάτω ιδιοτιμές από τον πίνακα του γραμμικοποιημένου συστήματος:

Έστω

$$\begin{aligned} A(g_1, g_2, g_3) = &\quad -g_3^3(18g_1^2g_2 + 9g_1(-2 + g_2)g_3 + 2g_3^3) = a \\ B(g_1, g_2, g_3) = &\quad g_3(3g_1^2g_2 - 3g_1(1 + g_2)g_3 + g_3^3) = b \end{aligned}$$

Τότε οι ιδιοτιμές είναι οι εξής:

$$\{-\frac{1}{6g_3}(2g_3^2 + \frac{2^{4/3}b}{(a + \sqrt{a^2 - 4b^3})^{1/3}} + 2^{2/3}(a + \sqrt{a^2 - 4b^3})^{1/3},$$

$$\frac{1}{12g_3}(-4g_3^2 + \frac{2^{4/3}(1 + i\sqrt{3})b}{(a + \sqrt{a^2 - 4b^3})^{1/3}} + 2^{2/3}(1 - i\sqrt{3})(a + \sqrt{a^2 - 4b^3})^{1/3},$$

$$\frac{1}{12g_3}(-4g_3^2 + \frac{2^{4/3}(1 - i\sqrt{3})b}{(a + \sqrt{a^2 - 4b^3})^{1/3}} + 2^{2/3}(1 + i\sqrt{3})(a + \sqrt{a^2 - 4b^3})^{1/3}\}$$

Σε αυτήν την περίπτωση έχουμε:

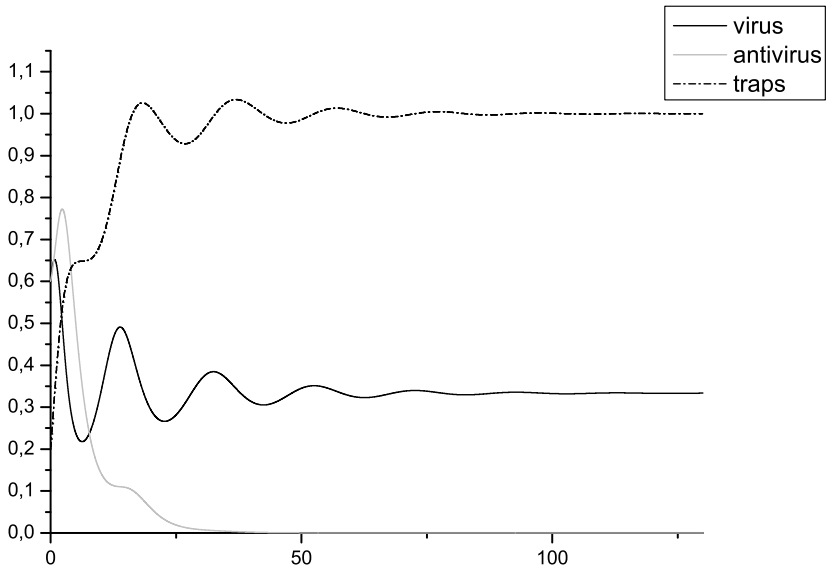
$$g_3 > g_1 g_2 \Rightarrow \frac{m_4}{m_1} > \frac{m_2}{m_1}\frac{m_3}{m_1}\frac{1}{k} \Rightarrow m_4 > \frac{m_2 m_3}{m_1 k}$$

Επειδή ο ρυθμός αύξησης των παγίδων είναι πολύ μικρός όπως και ο ρυθμός φθοράς τους, έχουμε ότι περίπου είναι ίσοι: $m_3 \simeq m_4$. Οπότε παίρνουμε ότι $m_1 k > m_2$. Επειδή $k < 1$, συμπεραίνουμε ότι $m_1 \gg m_2$, δηλαδή ο ρυθμός αύξησης των ιών είναι πολύ μεγαλύτερος από τον ρυθμό φθοράς των λογισμικών προστασίας και φυσικά όλων των υπολοίπων ρυθμών, όπου αυτή είναι και η πιο ρεαλιστική περίπτωση. Στην πραγματικότητα αυτό που παρατηρείται συνεχώς είναι η αύξηση των ιών όπου γίνεται με ραγδαίο ρυθμό. Δηλαδή έχουμε: $g_1 < 1$ και από προηγούμενη ανάλυση $g_3 < 1$. Βλέπουμε ότι οι ιδιοτιμές είναι παραστάσεις των παραμέτρων g_1, g_2, g_3 και είναι πολύ δύσκολο να βρούμε περιοχές τιμών όπου θα έχουν σταθερό πρόσημο για να εξάγουμε συμπέρασμα για την ευστάθεια του σημείου ισορροπίας. Για αυτόν το λόγο θα μελετήσουμε την ευστάθεια για ορισμένες περιπτώσεις τιμών αυτών των παραμέτρων. Πάντως σε κάθε περίπτωση αξίζει να σημειωθεί ότι είναι δυνατόν να υπάρχει μόνο ένα ευσταθές σημείο ισορροπίας.

2.3 Αριθμητικά αποτελέσματα

Σε αυτήν την ενότητα θα παρουσιάσουμε κάποια αριθμητικά αποτελέσματα από την επίλυση των διαφορικών εξισώσεων του συστήματος (2.7). Για την αριθμητική επίλυση χρησιμοποιήσαμε την μέθοδο Runge - Kutta τέταρτης τάξης.

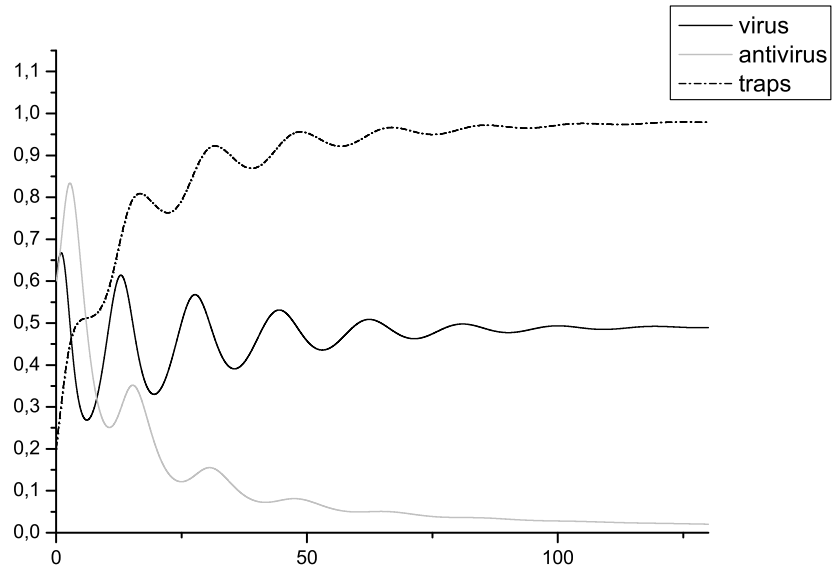
Αρχικά θα μελετήσουμε περιπτώσεις όπου ισχύει $g_3 < g_1 g_2$. Σε αυτήν την περίπτωση από την θεωρητική ανάλυση του μοντέλου έχουμε ότι στο χώρο λύσεων του συστήματος υπάρχει ένα ευσταθές σημείο ισορροπίας, το $(g_3/g_2, 0, 1)$. Στο διάγραμμα του σχήματος 2.2 αναπαριστάμε τις λύσεις των εξισώσεων για τιμές παραμέτρων: $g_1 = 0.5$, $g_2 = 0.3$, $g_3 = 0.1$. Σε όλες τις περιπτώσεις έχουμε αρχικές συνθήκες $x(0) = 0.6, \quad y(0) = 0.6, \quad z(0) = 0.2$. Με αυτές τις τιμές το σύστημα μας έχει ευσταθές σημείο ισορροπίας το $(0.333, 0, 1)$. Όπως βλέπουμε στο διάγραμμα των λύσεων από την αριθμητική επίλυση, οι λύσεις αρχικά κάνουν κάποιες ταλαντώσεις και μετά από κάποιο διάστημα παραμένουν πάνω στο σημείο ισορροπίας όπως είχαμε προβλέψει. Στο σχήμα 2.3 παρουσιάζουμε διάγραμμα



Σχῆμα 2.2:

όπου από τις τιμές των παραμέτρων δεν μπορούμε να προβλέψουμε ευστάθεια. Έχουμε δηλαδή $g_3 = g_1 g_2$ και πιο συγκεκριμένα $g_1 = 0.5$, $g_2 = 0.2$, $g_3 = 0.1$. Βλέπουμε ότι οι λύσεις συμπεριφέρονται σαν την περίπτωση όπου έχουμε ευστάθεια και παραμένουν πάνω στο σημείο ισορροπίας αλλά οι τιμές των antivirus αργούν να φτάσουν σε μηδενική τιμή όπως στην προηγούμενη περίπτωση.

Τώρα θα παρουσιάσουμε διαγράμματα από λύσεις συστημάτων όπου ισχύει $g_3 > g_1 g_2$. Έστω ότι έχουμε τις ακόλουθες τιμές για τις παραμέτρους του συστήματος $g_1 = 0.5$, $g_2 = 0.2$, $g_3 = 0.3$. Τότε από την ανάλυση μας παίρνουμε τις ιδιοτιμές: $\{-0.0305892 + 0.646364i, \; -0.0305892 - 0.646364i, \; -0.238833\}$. Βλέπουμε ότι όλες οι ιδιοτιμές έχουν αρνητικά πραγματικά μέρη άρα το σημείο

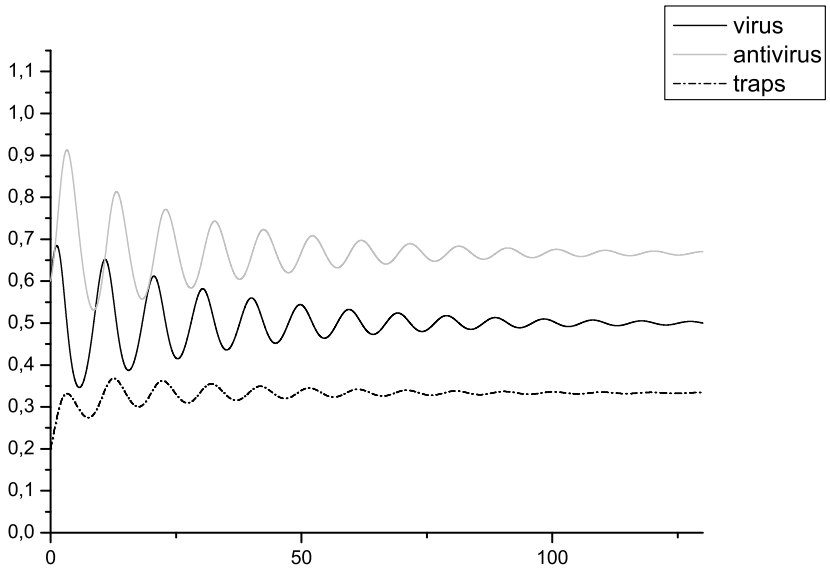


Σχῆμα 2.3:

ισορροπίας $\{0.5,\ 0.666,\ 0.333\}$ είναι ευσταθές. Η παρακάτω αριθμητική λύση (σχήμα 2.4) επιβεβαιώνει τον ισχυρισμό μας.

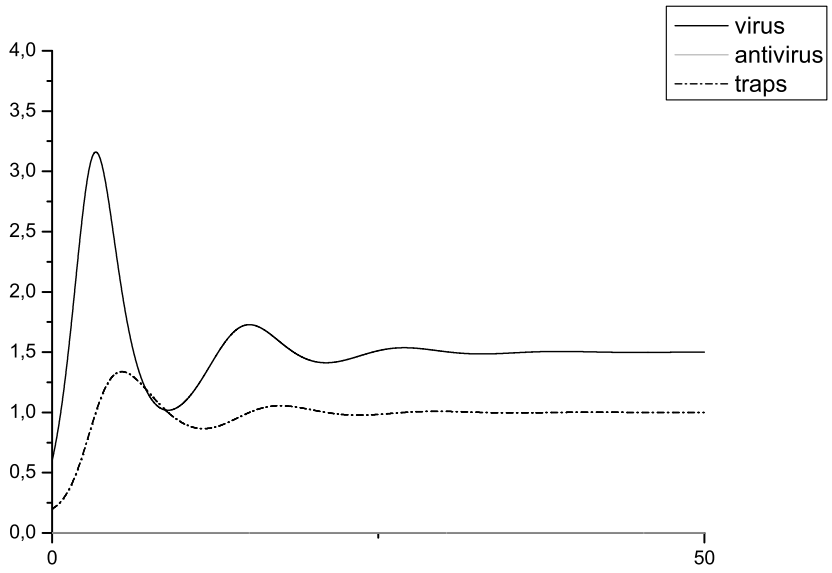
Μια ιδιαίτερη περίπτωση είναι όταν ισχύει $g_3 > g_1 g_2$, αλλά έχουμε μηδενική αρχική τιμή για το πλήθος των antivirus. Τα αριθμητικά αποτελέσματα για αυτήν την περίπτωση τα βλέπουμε στο σχήμα 2.5. Παρατηρούμε ότι οι λύσεις μας συμπεριφέρονται με όμοιο τρόπο όπως όταν έχουμε $g_3 < g_1 g_2$, δηλαδή με σημείο ισορροπίας το $(g_3/g_2, 0, 1)$. Σε αυτήν την περίπτωση που μελετάμε έχουμε $g_1 = 0.5$, $g_2 = 0.2$, οπότε το σημείο ισορροπίας είναι το $(1.5, 0, 1)$.

Η εισαγωγή των παγίδων στο μοντέλο μας βασίζεται στην ιδέα ότι μπορούμε να προστατεύσουμε ένα δίκτυο, τοποθετώντας λιγότερα στο πλήθος λογισμικά προστασίας, αλλά σε ορισμένους μόνο κεντρικούς υπολογιστές, ώστε να μειώσουμε το κόστος από την αγορά λογισμικών προστασίας για κάθε υπολογιστή, όταν έχουμε ένα πολύ μεγάλο δίκτυο. Γι' αυτό και αυτά τα λογισμικά προστασίας πρέπει να είναι πολύ ισχυρά και να έχουν όλες τις απαραίτητες ενημερώσεις αφού ο σκοπός τους δεν είναι η προστασία ενός μόνο υπολογιστή, αλλά σε πολλές περιπτώσεις μπορεί να είναι και το δίκτυο μίας επιχείρησης. Από την θεωρητική μας ανάλυση αποδεικνύουμε ότι μπορεί να υπάρξει μία κατάσταση ισορροπίας του συστήματος με μόνο αυτούς τους δύο μη- μηδενικούς πληθυσμούς, όπου για κάποιες τιμές των παραμέτρων είναι ευσταθής. Στο τελευταίο διάγραμμα φαίνεται ξεκάθαρα ότι ο πληθυσμός των ιών παρουσιάζει απότομη άυξηση στην αρχή της διάδοσης τους, λόγω του πολύ μικρού πλήθους των παγιδών. Στη συνέχεια, με την εμφάνιση πολλων ιών στους υπολογιστές του δικτύου και μετά απο λίγο χρονικό διάστημα που απαιτείται για να ενεργήσουν οι administrators και να εγκαταστήσουν και άλλες λίγες παγίδες στο δίκτυο, ο πληθυσμός των ιών πέφτει κατακόρυφα, σαν αποτέλεσμα της βέβαιης αναχαίτησης των ιών απ' τα πανίσχυρα



Σχῆμα 2.4:

λογισμικά προστασίας. Έπειτα απο μια μικρή ταλάντωση, οι δυο πληθυσμοί θα έρθουν γρήγορα σε κατάσταση ισορροπίας, με τον πληθυσμό των ιών λίγο πιο πάνω απ'την αρχική του τιμή. Σε μια τέτοια περίπτωση μπορούμε να πούμε ότι έχουμε προστατεύσει το δίκτυο αφού δεν πρόκειται ποτέ να παρατηρηθεί εκτός της πρώτης φοράς, έξαρση του πληθυσμού των ιών. Άρα εχουμε αποδείξει τελικά ότι με την χρήση λιγοστών ισχυρών λογισμικών προστασίας μπορούμε να προστατεύσουμε ένα δίκτυο απο επιθέσεις ιών και να έχουμε μείωση του κόστους αγοράς πολλών λογισμικών που είναι και το ζητούμενο.



Σχῆμα 2.5:

Κεφάλαιο 3

Μελέτη εισβολών και αναχαιτίσεων ιών σε δίκτυα υπολογιστών βασισμένη στη Θεωρία Ουρών

3.1 Εισαγωγή

Σε αυτό το κεφάλαιο παρουσιάζουμε ένα μοντέλο διάδοσης και απαλοιφής ιών το οποίο λαμβάνει υπ' όψιν τα χαρακτηριστικά των servers και της κίνησης (traffic) σε ένα δίκτυο υπολογιστών. Αυτό το μοντέλο διαχωρίζει τους κόμβους του δικτύου σε περιμετρικούς και μη-περιμετρικούς κόμβους. Η εσωτερική και εξωτερική κίνηση (incoming/ outgoing traffic) περιορίζεται στην "περίμετρο" του δικτύου, όπου σαν περίμετρο ορίζουμε τους κόμβους του δικτύου που συνδέονται απ' ευθείας με το διαδίκτυο. Οι μη-περιμετρικοί κόμβοι είναι π.χ. υπολογιστές που δεν έχουν απ' ευθείας σύνδεση με το διαδίκτυο, αλλά είναι ένα είδος απομονωμένων κόμβων που τους παρέχεται πρόσβαση στο διαδίκτυο μέσω των περιμετρικών κόμβων. Όλοι οι κόμβοι του δικτύου θεωρούμε ότι λειτουργούν βάση του μοντέλου ουρών M/M/1. Οπότε όλο το μοντέλο του δικτύου συμπεριφέρεται σαν ένα ανοικτό δίκτυο ουρών M/M/1. (open network M/M/1 queue).

Θα μελετήσουμε ξέσπασμα εισβολών ιών (burst intrution) στους περιμετρικούς κόμβους του δικτύου και πως εξελίσσεται αυτή η εισβολή, όταν παράλληλα με αυτή την εισβολή έχουμε και διάδοση λογισμικών προστασίας (antivirus) που κινούνται στο δίκτυο με σκοπό να εμποδίσουν την διάδοση των ιών. Θα προτείνουμε μια ρεαλιστική απεικόνιση της αλληλεπίδρασης των δύο παραγόντων (agents) που θα μας οδηγήσει σε μια κατανομή μορφής γινομένου(product form distribution) σε κατάσταση ισσοροπίας, (steady state) για τον αριθμό των agents στα πακέτα εργασιών, σε κάθε κόμβο του δικτύου, όπως την κατανομή για τους κόμβους ενός Jackson open network.

Ειδικότερα σε αυτό το κεφάλαιο θα μελετήσουμε το πρόβλημα διάδοσης και αναχαίτησης ιών σε δίκτυα από μία άλλη οπτική γωνία η οποία αποφεύγει την χρήση των μη γραμμικών διαφορικών εξισώσεων, όπως το μοντέλο Lotka-Volterra. Αντιθέτως θα προτείνουμε και θα αναλύσουμε ένα μαθηματικό μοντέλο για την

εξέλιξη και των δύο πληθυσμών ιών και λογισμικών προστασίας, που βασίζεται στην θεωρία ουρών. Μοντελοποιούμε το δίκτυο σαν ένα δίκτυο υπερσυνδεδεμένων κόμβων, οι οποίοι έχουν συνδέσεις μεταξύ τους αλλά και με τον εξωτερικό κόσμο [3, 16]. Με τον όρο εξωτερικό κόσμο ενοούμε υπολογιστές οι οποίοι δεν ανήκουν στο τοπικό δίκτυο που μελετάμε, ή ακόμα και το διαδίκτυο (internet). Κάθε κόμβος ο οποίος εξυπηρετεί πακέτα εργασιών που κινούνται στο δίκτυο, μοντελοποιείται σαν μία M/M/1 ουρά. Αυτά τα πακέτα εργασιών μπορεί να περιέχουν ιό ή και λογισμικό προστασίας. Η ιδέα πίσω από αυτό το μοντέλο είναι ότι και τα λογισμικά προστασίας είναι σαν τυπικά πακέτα και χρειάζονται να εκτελεστούν πρώτα από τους χρήστες για να αρχίσει η διάδοση τους. Αυτό το μοντέλο συνδέει τα χαρακτηριστικά του δικτύου (χρόνοι εξυπηρέτησης, ρυθμός χρησιμοποίησης) με την ταχύτητα με την οποία διαδίδεται ο ιός. Ένα επιπλέον στοιχείο αυτού του μοντέλου είναι η αντίδραση μεταξύ του ιού και του λογισμικού προστασίας. Ο κανόνας μας είναι απλός, όταν ένα λογισμικό προστασίας συναντήσει έναν ιό, τότε το λογισμικό προστασίας "σκοτώνει" τον ιό και μετά "σκοτώνεται" και το ίδιο, έχουμε δηλαδή μια αλληλοεξουδετέρωση για να περιορίζεται η φόρτωση του δικτύου. Μας ενδοιαφέρει η εξέλιξη και των δύο πληθυσμών μέσα στο δίκτυο καθώς ακολουθούν αυτόν τον κανόνα της αντίδρασης. Θα δείξουμε ότι η κατανομή του πλήθους και των δύο ειδών στους κόμβους του δικτύου μπορεί να γραφεί σε μορφή γινομένου όπως την λύση για τα ανοικτά δίκτυα Jackson M/M/1 ουρών.

Το θεώρημα Jackson είναι πολύ σημαντικό στην ανάπτυξη της θεωρίας δικτύων ουρών [16].

Θεώρημα Jackson

Έστω ένα ανοικτό δίκτυο ουρών με τα εξής χαρακτηριστικά:

(i) M ο αριθμός των ουρών του δικτύου (όσοι και οι κόμβοι)

(ii) m_i ρυθμός εξυπηρέτησης της i ουράς.

(iii) l_i ρυθμός αφίξεων στην i ουρά.

(iv) p_i χρησιμοποίηση της i ουράς που ισούται με l_i/m_i.

(v) $n_i(t)$ πλήθος εργασιών στην ουρά i την χρονική στιγμή t.

(vi) $\underline{n}(t) = (n_1(t), n_2(t), ...n_M(t))$ η κατάσταση του συστήματος την χρονική στιγμή t.

(vii) $P(k_1, k_2, ..., k_M; t) = Pr[\underline{n}(t) = (k_1, k_2, ..., k_M)]$

(viii) $P(k_1, k_2, ..., k_M) = lim_{t\to\infty} P(k_1, k_2, ..., k_M; t)$

Οι αφίξεις των εργασιών ακολουθούν κατανομή Poisson και οι χρόνοι εξυπηρέτησης των ουρών ακολουθούν εκθετική κατανομή. Τότε

$$P(k_1, k_2, ..., k_M) = \prod_{i=1}^{M} (1 - p_i) p_i^k i$$

3.2 Μοντέλο διάδοσης εισβολών και απαλοιφής

Μοντελοποιούμε ένα δίκτυο υπολογιστών σαν ένα ανοικτό δίκτυο Jackson Opan Jackson Network με κόμβους (υπολογιστές) οι οποίοι συνδέονται όλοι μεταξύ τους και επιπλέον όλοι οι υπολογιστές έχουν την δυνατότητα να συνδέονται με εξωτερικά δίκτυα, δηλαδή με δίκτυα εκτός από αυτό που μελετάμε, ακόμα και με το διαδίκτυο (outside world)[1]. Κάθε κόμβος του δικτύου μοντελοποιείται σαν μία M/M/1 ουρά απείρου μεγέθους, δηλαδή δεν υπάρχει πιθανότητα να απορριφθεί κανένα πακέτο που έχει αποφασιστέι να σταλθεί σε αυτόν τον κόμβο. Ο χρόνος εξυπηρέτησης της ουράς ακολουθεί εκθετική κατανομή, ενώ ο αριθμός των αφίξεων των πακέτων ακολουθεί κατανομή Poisson. Υποθέτουμε ότι ο χρόνος εξυπηρέτησης των πακέτων σε κάθε ουρά, είναι ανεξάρτητος από το χρόνο εξυπηρέτησης των άλλων ουρών. Ακόμα υποθέτουμε ότι ο χρόνος μεταφοράς ενός πακέτου σε μία ουρά είναι ίδιος για όλες τις ουρές του δικτύου, και είναι περίπου ίσος με την αντίστροφη ταχύτητα μετάδοσης του συνδέσμου που οδηγεί στη ουρά. Αυτό είναι γενικά αποδεκτό για όλα τα πακέτα που μεταφέρονται μέσα σε ένα δίκτυο, και ακόμα μπορούμε να υποθέσουμε αν το μέγεθος του πακέτου είναι πολύ μικρό, ότι ο χρόνος μεταφοράς είναι σταθερός. Οι χρόνοι εξυπηρέτησης για ένα πακέτο καθώς ταξιδεύει στους κόμβους του δικτύου προς τον κόμβο προορισμού είναι ανεξάρτητοι μεταξύ τους (αρχή ανεξαρτησίας του Kleinrock).

Αρχή Ανεξαρτησίας του Kleinrock

(i) Οι χρόνοι αφίξεων των πακέτων στις ουρές του δικτύου είναι ανεξάρτητοι.

(ii) Οι χρόνοι εξυπηρέτησης για ένα συγκεκριμένο πακέτο στις ουρές του δικτύου είναι ανεξάρτητοι.

(iii) Οι χρόνοι εξυπηρέτησης και οι χρόνοι αφίξεων είναι ανεξάρτητοι.

Κάθε φορά που ένα πακέτο εξυπηρετηθεί σε μία ουρά, διαλέγει τον επόμενο κόμβο που θα μεταπηδήσει, ή εξέρχεται του δικτύου με μια συγκεκριμένη πιθανότητα (Markov Routing). Επίσης το μοντέλο επιτρέπει ντετερμινιστική εξέλιξη, όπου η επιλογή για τον επόμενο κόμβο είναι προκαθορισμένη. Το δίκτυο είναι ανοικτό για αφίξεις πακέτων από άλλα δίκτυα σε όλους τους κόμβους, και για κάθε κόμβο η εισερχόμενη κίνηση μοντελοποιείται με κατανομή Poisson. Οι παράμετροι του μοντέλου είναι οι παρακάτω:

(i) N ο αριθμός των ουρών του δικτύου (όσοι και οι κόμβοι)

(ii) l_i η παράμετρος της κατανομής Poisson που χρησιμοποιείται για να μοντελοποιεί τον ρυθμό αφίξεων των πακέτων στην κάθε ουρά, είτε περιέχουν ιό είτε λογισμικό προστασίας διότι όταν φτάνει ένα πακέτο στην ουρά, δεν γνωρίζουμε αν δεν εκτελεστεί τι περιέχει και γι' αυτό το λόγο χρησιμοποιούμε μία παράμετρο και για τα δύο είδη.

(iii) mi η παράμετρος της εκθετικής κατανομής για τον χρόνο εξυπηρέτησης της i ουράς.

(iv) ϱ_i ο ρυθμός χρησιμοποίησης utilization της ουράς που ισούται με l_i/m_i.

(v) (a_i, v_i, d_i) ο αριθμός των λογισμικό προστασίας (antivirus), των ιών (virus) και των αντιδράσεων στην i ουρά σε συνθήκες μόνιμης κατάστασης.

(vi) $Pr[a]$ η πιθανότητα ένα πακέτο να περιέχει λογισμικό προστασίας.

(vii) $Pr[v]$ η πιθανότητα ένα πακέτο να περιέχει ιό.

(viii) l_{ij} ο ρυθμός με τον οποίο ένα πακέτο αφήνει την i ουρά για να μεταπηδήσει στην ουρά j.

(ix) q_{ij} η πιθανότητα ένα πακέτο να αφήσει την i ουρά για να πάει στην ουρά j.

(x) ορίζουμε σαν διάνυσμα κατάστασης το διάνυσμα που περιέχει N τριάδες, μία για κάθε κόμβο του δικτύου, που περιγράφει τον αριθμό των λογισμικών προστασίας, των ιών και των αντιδράσεων που συμβαίνουν μεταξύ των ειδών των δύο πλυθυσμών.

$$n(t) = ((a_1(t), v_1(t), d_1(t)), \ldots, (a_N(t), v_N(t), d_N(t)))$$

είναι η κατάσταση του συστήματος την χρονική στιγμή t. Ειδικότερα, για έναν συγκεκριμένο κόμβο i, η τριάδα $(a_i(t), v_i(t), d_i(t))$ δηλώνει ότι την χρονίκή στιγμή t, ο κόμβος i περιέχει $a_i(t)$ πλήθος λογισμικών προστασίας, $v_i(t)$ πλήθος ιών και έχουν συμβεί $d_i(t)$ αντιδράσεις μεταξύ των ιών και των λογισμικών προστασίας. Παρατηρούμε ότι το άθροισμα $a_i(t)+v_i(t)+2d_i(t)$ είναι ίσο με τον συνολικό αριθμό των ιών και των λογισμικών προστασίας που έχουν εξυπηρετηθεί από τον κόμβο i.

(xi) Η συνάρτηση πυκνότητας πιθανότητας για τα πιθανά διανύσματα κατάστασης του συστήματος ορίζεται ως εξής:

$$\begin{aligned} P_{((a_1,v_1,d_1),(a_2,v_2,d_2),\ldots,(a_N,v_N,d_N):t)} &= Pr[n(t) \\ &= ((a_1,v_1,d_1),(a_2,v_2,d_2),\ldots,(a_N,v_N,d_N))] \end{aligned}$$

(xii) Η κατανομή σταθερής κατάστασης (steady state distribution) για τα διανύσματα κατάστασης ορίζεται ως εξής:

$$P_{((a_1,v_1,d_1),(a_2,v_2,d_2),\ldots,(a_N,v_N,d_N))} = \lim_{t\to\infty} P((a_1,v_1,d_1),(a_2,v_2,d_2),\ldots,(a_N,v_N,d_N):t)$$

3.3 Κατανομή σταθερής κατάστασης

Σε αυτό το σημείο παρουσιάζουμε την κατανομή μορφής γινομένου για τα πλήθη δύο ειδών στους κόμβους του δικτύου. Η κατανομή αυτή γενικεύει την κατανομή του Jackson για δύο πληθυσμούς. Για την απόδειξη χρησιμοποιήσαμε και μια τρίτη μεταβλητή, η οποία υπολογίζει τον αριθμό των αντιδράσεων σε κάθε κόμβο, ώστε να γνωρίζουμε τον ακριβή αριθμό των πακέτων που έχουν εξυπηρετηθεί κάθε χρονική στιγμή.

Θεώρημα
Έστω ένα δίκτυο υπολογιστών που μοντελοποιείται με τον τρόπο που παρουσιάσαμε στην προηγούμενη ενότητα. Η συνάρτηση κατανομής για το διάνυσμα του συστήματος σε σταθερή κατάσταση είναι:

$$P((a_1,v_1,d_1),(a_2,v_2,d_2),....,(a_N,v_N,d_N)) = \prod_{i=1}^{N}(1-p_i)p_i^{a_i+v_i+2d_i} \tag{3.1}$$

Απόδειξη. Η απόδειξη του θεωρήματος ακολουθεί την γενική ιδέα του θεωρήματος Jackson για έναν πληθυσμό σε ανοικτά δίκτυα ουρών. Αρχικά καταγράφουμε όλα τα πιθανά γεγονότα που μπορεί να συμβούν σε ένα απειροελάχειστο χρονικό διάστημα dt.

(i) Ένα πακέτο καταφτάνει σε μία ουρά του δικτύου.

(ii) Ένα πακέτο αφήνει μία ουρά και εξέρχεται του δικτύου.

(iii) Ένα πακέτο αφήνει μία ουρά και εισέρχεται σε άλλη ουρά του δικτύου.

(iv) Μία αντίδραση λαμβάνειν χώρα σε μία ουρά του δικτύου.

(v) Τίποτα από τα παραπάνω.

Συμπεριλαμβάνοντας όλα τα παραπάνω γεγονότα για τον υπολογισμό της πιθανότητας της αλλαγής του διανύσματος κατάστασης του συστήματος παίρνουμε:

$$
\begin{aligned}
&P((a_1,v_1,d_1),(a_2,v_2,d_2),....,(a_N,v_N,d_N):t+dt) = \\
&\sum_{j=1}^{N} P((a_1,v_1,d_1),...(a_j-1,v_j,d_j),...,(a_N,v_N,d_N):t) l_{0j} Pr(a) dt \\
&+\sum_{j=1}^{N} P((a_1,v_1,d_1),...(a_j,v_j-1,d_j),...,(a_N,v_N,d_N):t) l_{0j} Pr(v) dt \\
&+\sum_{i=1}^{N} P((a_1,v_1,d_1),...(a_i+1,v_i,d_i),...,(a_N,v_N,d_N):t) m_i q_{i0} Pr(a) dt \\
&+\sum_{i=1}^{N} P((a_1,v_1,d_1),...(a_i,v_i+1,d_i),...,(a_N,v_N,d_N):t) m_i q_{i0} Pr(v) dt \\
&+\sum_{i=1}^{N}\sum_{j=1}^{N} P((a_1,v_1,d_1),...(a_i{+}1,v_i,d_i),..,(a_j{-}1,v_j,d_j)..,(a_N,v_N,d_N):t) m_i q_{ij} Pr(a) dt \\
&+\sum_{i=1}^{N}\sum_{j=1}^{N} P((a_1,v_1,d_1),...(a_i,v_i{+}1,d_i),..,(a_j,v_j{-}1,d_j)..,(a_N,v_N,d_N):t) m_i q_{ij} Pr(v) dt \\
&+\sum_{i=1}^{N} P((a_1,v_1,d_1),...(a_i+1,v_i+1,d_i-1),...,(a_N,v_N,d_N):t) m_i^2 Pr(a) Pr(v) dt \\
&+P((a_1,v_1,d_1),(a_2,v_2,d_2),....,(a_N,v_N,d_N):t)(1-dt\sum_{j=1}^{N}(l_{0j}+m_j+mj^2 Pr(a) Pr(v)))
\end{aligned}
\tag{3.2}
$$

Ο πρώτος και ο δεύτερος όρος της παραπάνω εξίσωσης υπολογίζουν την πιθανότητα ένα λογισμικό προστασίας ή ένας ιός να είσελθει σε μία ουρά του δικτύου. Ο τρίτος και ο τέταρτος όρος υπολογίζουν την πιθανότητα ένα λογισμικό προστασίας ή ένας ιός να εγκαταλείψουν το δίκτυο από κάποια ουρά του δικτύου. Ο πέμπτος και ο έκτος όρος αντίστοιχα υπολογίζουν την πιθανότητα ένα λογισμικό προστασίας ή ένας ιό να μετακινηθούν από μία ουρά του δικτύου σε μία άλλη εντός του

δικτύου. Ο έβδομος όρος υπολογίζει την πιθανότητα να έχουμε αντίδραση σε μία ουρά του δικτύου και για να γίνει αυτό θα πρέπει την προηγούμενη χρονική στιγμή να είχαμε και τα δύο είδη αυξημένα κατά 1 και το όρο που μετρά τις αντιδράσεις μειωμένο κατά 1, και θα πρέπει να έχουν εξυπηρετηθεί και τα δύο είδη. Τέλος έχουμε την πιθανότητα να μην συμβεί τίποτα από τα προηγούμενα που αναφέραμε, και για να γίνει αυτό θα πρέπει σε αυτό το dt να μην εισέλθει τίποτα σε καμία ουρά, να μην εξυπηρετηθεί τίποτα (οι περιπτώσεις όπου έχουμε κίνηση μεταξύ δύο υπολογιστών εντός του δικτύου η που εξέρχεται κάτι από το δίκτυο περιέχονται σε αυτόν τον όρο αφού για να υπάρξει κίνηση θα πρέπει πρώτα να εξυπηρετηθεί το πακέτο) και να μην γίνει αντίδραση που περιγράφεται με τον όρο $mj^2 Pr(a)Pr(v)$. Στην συνέχεια μετακινούμε τον όρο

$$P((a_1, v_1, d_1), (a_2, v_2, d_2), ..., (a_N, v_N, d_N) : t)$$

στο αριστερό μέλος της εξίσωσης και διαιρούμε όλα τα μέλη με dt. Υπολογίζοντας το όριο καθώς το $dt \to 0$, στο αριστερό μέλος παίρνουμε την παράγωγο

$$\frac{d}{dt}P((a_1, v_1, d_1), (a_2, v_2, d_2), ..., (a_N, v_N, d_N) : t)$$

ενώ στα αριστερά μέλη έχει απαληφθεί ο παράγοντας dt.

$$\begin{aligned}
&\frac{d}{dt}P((a_1, v_1, d_1), (a_2, v_2, d_2), ..., (a_N, v_N, d_N) : t) = \\
&\sum_{j=1}^{N} P((a_1, v_1, d_1), ...(a_j - 1, v_j, d_j), ..., (a_N, v_N, d_N) : t) l_{0j} Pr(a) \\
&+ \sum_{j=1}^{N} P((a_1, v_1, d_1), ...(a_j, v_j - 1, d_j), ..., (a_N, v_N, d_N) : t) l_{0j} Pr(v) \\
&+ \sum_{i=1}^{N} P((a_1, v_1, d_1), ...(a_i + 1, v_i, d_i), ..., (a_N, v_N, d_N) : t) m_i q_{i0} Pr(a) \\
&+ \sum_{i=1}^{N} P((a_1, v_1, d_1), ...(a_i, v_i + 1, d_i), ..., (a_N, v_N, d_N) : t) m_i q_{i0} Pr(v) \\
&+ \sum_{i=1}^{N}\sum_{j=1}^{N} P((a_1, v_1, d_1), ...(a_i{+}1, v_i, d_i), .., (a_j{-}1, v_j, d_j).., (a_N, v_N, d_N) : t) m_i q_{ij} Pr(a) \\
&+ \sum_{i=1}^{N}\sum_{j=1}^{N} P((a_1, v_1, d_1), ...(a_i, v_i{+}1, d_i), .., (a_j, v_j{-}1, d_j).., (a_N, v_N, d_N) : t) m_i q_{ij} Pr(v) \\
&+ \sum_{i=1}^{N} P((a_1, v_1, d_1), ...(a_i + 1, v_i + 1, d_i - 1), ..., (a_N, v_N, d_N) : t) m_i^2 Pr(a)Pr(v) \\
&- \sum_{j=1}^{N} P((a_1, v_1, d_1), (a_2, v_2, d_2), ..., (a_N, v_N, d_N) : t)(l_{0j} + m_j + m_j^2 Pr(a)Pr(v))
\end{aligned}$$

Σε συνθήκες μόνιμης κατάστασης δεν έχουμε μεταβολές στους πληθυσμούς του συστήματος. Οπότε

$$\frac{d}{dt}P((a_1,v_1,d_1),(a_2,v_2,d_2),....,(a_N,v_N,d_N):t)=0$$

και άρα έχουμε την παρακάτω εξίσωση:

$$\sum_{j=1}^{N}P((a_1,v_1,d_1),(a_2,v_2,d_2),....,(a_N,v_N,d_N))(l_{0j}+m_j+m_j^2Pr(a)Pr(v))=$$

$$\sum_{j=1}^{N}P((a_1,v_1,d_1),...(a_j-1,v_j,d_j),...,(a_N,v_N,d_N))l_{0j}Pr(a)$$

$$+\sum_{j=1}^{N}P((a_1,v_1,d_1),...(a_j,v_j-1,d_j),...,(a_N,v_N,d_N))l_{0j}Pr(v)$$

$$+\sum_{i=1}^{N}P((a_1,v_1,d_1),...(a_i+1,v_i,d_i),...,(a_N,v_N,d_N))m_iq_{i0}Pr(a)$$

$$+\sum_{i=1}^{N}P((a_1,v_1,d_1),...(a_i,v_i+1,d_i),...,(a_N,v_N,d_N))m_iq_{i0}Pr(v)$$

$$+\sum_{i=1}^{N}\sum_{j=1}^{N}P((a_1,v_1,d_1),...(a_i+1,v_i,d_i),..,(a_j-1,v_j,d_j)..,(a_N,v_N,d_N))m_iq_{ij}Pr(a)$$

$$+\sum_{i=1}^{N}\sum_{j=1}^{N}P((a_1,v_1,d_1),...(a_i,v_i+1,d_i),..,(a_j,v_j-1,d_j)..,(a_N,v_N,d_N))m_iq_{ij}Pr(v)$$

$$+\sum_{i=1}^{N}P((a_1,v_1,d_1),...(a_i+1,v_i+1,d_i-1),...,(a_N,v_N,d_N))m_i^2Pr(a)Pr(v)$$

Αντικαθιστώντας τον όρο που παίρνουμε από το θεώρημα στην εξίσωση μόνιμης κατάστασης του συστήματος παίρνουμε την ακόλουθη εξίσωση αθροισμάτων:

$$\sum_{j=1}^{N}(l_{0j}+m_j+m_j^2Pr(a)Pr(v))=\sum_{j=1}^{N}\frac{l_{0j}Pr(a)}{p_j}+\sum_{j=1}^{N}\frac{l_{0j}Pr(v)}{p_j}$$

$$+\sum_{i=1}^{N}m_iq_{i0}p_iPr(a)+\sum_{i=1}^{N}m_iq_{i0}p_iPr(v)+\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{p_i}{p_j}m_iq_{ij}Pr(a)$$

$$+\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{p_i}{p_j}m_iq_{ij}Pr(v)+\sum_{i=1}^{N}m_i^2Pr(a)Pr(v)$$

Επειδή μελετάμε την κίνηση των πακέτων μόνο αυτών που περιέχουν λογισμικό προστασίας ή ιό έχουμε $Pr(a)+Pr(v)=1$. Από τη παραπάνω εξίσωση παίρνουμε

$$\sum_{j=1}^{N}(l_{0j}+m_j+m_j^2Pr(a)Pr(v))=\sum_{j=1}^{N}\frac{l_{0j}}{p_j}+\sum_{i=1}^{N}m_iq_{i0}p_i$$

$$+\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{p_i}{p_j}m_i q_{ij}+\sum_{i=1}^{N}m_i^2 Pr(a)Pr(v) \tag{3.3}$$

Θα μελετήσουμε τώρα κάθε όρο από τους τέσσερις στο δεξιό μέλος ξεχωριστά.

Πρώτος όρος: Αντικαθιστάμε το πηλίκο $p_i = l_i/m_i$ (utilization of i queue) και έχουμε

$$\sum_{j=1}^{N}\frac{l_{0j}}{p_j}=\sum_{j=1}^{N}\frac{l_{0j}m_j}{l_j}$$

Δεύτερος όρος:

$$\sum_{i=1}^{N}m_i q_{i0}p_i=\sum_{i=1}^{N}l_i q_{i0}=\sum_{i=1}^{N}l_i(1-\sum_{j=1}^{N}q_{ij})=\sum_{i=1}^{N}l_i-\sum_{i=1}^{N}\sum_{j=1}^{N}l_i q_{ij}$$

Θα μελετήσουμε τώρα τον τελευταίο όρο $\sum_{i=1}^{N}\sum_{j=1}^{N}l_i q_{ij}$ που είναι όλα τα πακέτα που έρχονται στα i επί την πιθανότητα να πάνε στα j, που αυτό ισούται με όλα αυτά που έρχονται στα i εκτός από αυτά που εξέρχονται του δικτύου και εκτός από αυτά που αντιδρούν. Άρα

$$\sum_{i=1}^{N}\sum_{j=1}^{N}l_i q_{ij}=\sum_{i=1}^{N}l_i-\sum_{i=1}^{N}l_{i0}-\sum_{i=1}^{N}m_i^2 Pr(a)Pr(v) \tag{3.4}$$

Από την αρχή διατήρησης των σωματιδίων έχουμε ότι

$$l_i=\sum_{j=0}^{N}l_{ij}+m_i^2 Pr(a)Pr(v)$$

Σε όλο το δίκτυο ισχύει:

$$\sum_{i=1}^{N}l_i=\sum_{i=1}^{N}\sum_{j=0}^{N}l_{ij}+\sum_{i=1}^{N}m_i^2 Pr(a)Pr(v)\Rightarrow$$

$$\sum_{i=0}^{N}\sum_{j=1}^{N}l_{ij}=\sum_{i=1}^{N}\sum_{j=0}^{N}l_{ij}+\sum_{i=1}^{N}m_i^2 Pr(a)Pr(v)\Rightarrow$$

$$\sum_{i=1}^{N}\sum_{j=1}^{N}l_{ij}+\sum_{j=1}^{N}l_{0j}=\sum_{i=1}^{N}\sum_{j=1}^{N}l_{ij}+\sum_{i=1}^{N}l_{i0}+\sum_{i=1}^{N}m_i^2 Pr(a)Pr(v)\Rightarrow$$

$$\sum_{j=1}^{N}l_{0j}=\sum_{i=1}^{N}l_{i0}+\sum_{i=1}^{N}m_i^2 Pr(a)Pr(v)$$

Οπότε η εξίσωση (3.4) γίνεται

$$\sum_{i=1}^{N}\sum_{j=1}^{N}l_i q_{ij}=\sum_{i=1}^{N}l_i-\sum_{j=1}^{N}l_{0j}$$

και όλος ο δεύτερος όρος τελικά γίνεται

$$\sum_{j=1}^{N} l_{0j}$$

.

Τρίτος όρος:
Ο τρίτος όρος φράφεται ως εξής:

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{p_i}{p_j}m_i q_{ij} = \sum_{i=1}^{N}\sum_{j=1}^{N} l_i q_{ij}\frac{m_j}{l_j} = \sum_{j=1}^{N}\frac{m_j}{l_j}\sum_{i=1}^{N} l_i q_{ij} = \sum_{j=1}^{N}\frac{m_j}{l_j}(l_j - l_{0j})$$

$$= \sum_{j=1}^{N} m_j - \sum_{j=1}^{N}\frac{m_j l_{0j}}{lj}$$

Ο τέταρτος όρος δεν χρειάζεται καμία μετατροπή.

Αθροίζοντας τους παραπάνω τέσσερις όρους βλέπουμε ότι ισχύει η ισότητα μεταξύ των δύο μελών της εξίσωσης (3.3), και καταλήγουμε στο συμπέρασμα ότι η συνάρτηση κατανομής (3.1) που έχουμε υποθέσει στο θεώρημα είναι σωστή, αφού ικανοποιεί την συνθήκη της μόνιμης κατάστασης του συστήματος, ολοκληρώνοντας την απόδειξη.

3.4 Ιδιότητες και αποτίμηση του μοντέλου

Το προτεινόμενο μοντέλο υποθέτει ότι οι ιοί και τα λογισμικά προστασίας εισέρχονται στο δίκτυο ως κανονικές "καθαρές" διεργασίες, όπως μέσω e-mail, αναβαθμίσεις λογισμικού ή και κατέβασμα δεδομένων και χρειάζονται να εκτελεστούν πρώτα από τους χρήστες για να αρχίσει η διάδοση τους. Ακόμα, η μόνη αλληλεπίδραση που επιτρέπεται από το μοντέλο είναι η αλληλοεξουδετέρωση ενός ιού και ενός λογισμικού προστασίας. Με αυτό επιτυγχάνουμε α) με την καταστροφή του ιού εμποδίζουμε την διάδοσή τους και β) με την καταστροφή του λογισμικού προστασίας επιτυγχάνουμε να μην υπερφορτώσουμε το δίκτυο. Παρόλο που το μοντέλο θα μπορούσε να έχει πιο δραστίκα μέτρα για την αντιμετώπιση των ιών, έτσι έχει δύο βασικά πλεονεκτήματα, α) μπορεί πολύ εύκολα να αναλυθεί μαθηματικά, με τις ίδιες τεχνικές που αναλύεται και το μοντέλο Jackson για τα ανοικτά δίκτυα ουρών και β) μπορούμε να το δούμε σαν το χειρότερο δυνατό σενάριο βάζοντας ένα άνω φράγμα στο πως περιορίζουμε μια διάδοση ιών σε ένα δίκτυο που έχει χαρακτηριστικά δικτύου ουρών.

Το προτεινόμενο μοντέλο επίσης αποφεύγει την χρήση των συστημάτων μη γραμμικών διαφορικών εξισώσεων που περιγράφουν την εξέλιξη δύο ειδών πληθησμού. Ακόμα η συνάρτηση κατανομής που παίρνουμε είναι απλής μορφής σαν γινόμενο, και είναι πολύ εύκολο να υπολογιστεί εφόσον γνωρίζουμε τις παραμέτρους του δικτύου. Επιπλέον το μοντέλο λαμβάνειν υπόψην την τοπολογία του δικτύου, αφού έχουμε ξεχωριστές παραμέτρους $l_{i,j}$ για κάθε δυάδα i, j υπολογιστών μέσα στο δίκτυο που χαρακτηρίζουν την συνδεσμολογία του δικτύου. Αυτές οι παράμετροι μπορεί να βρεθούν σαν συναρτήσεις των ρυθμών με τους οποίους οι διεργασίες καταφθάνουν στο δίκτυο, αφού καταγράψουμε τις εξισώσεις που δηλώνουν την διατήρηση των ρυθμών αποστολής μέσα στο δίκτυο [3]. Αυτές οι εξισώσεις είναι γραμμικές και μπορούν να επιλυθούν πολύ εύκολα και να μας

δώσουν όλα τα $l_{i,j}$ για κάθε κόμβο i. Οι λύσεις αυτές οδηγούν στον υπολογισμό του p_i, μία παράμετρος της εξίσωσης του μοντέλου μας (3.1).

Ο κανόνας αλληλοεξουδετέρωσης που αναλύσαμε μπορεί να γενικευτεί πολυ εύκολα, για παράδειγμα αν υποθέσουμε ότι ένα antivirus μπορεί να εξουδετερώσει s αριθμό virus, τότε η συνάρτηση κατανομής πιθανότητας μπορεί να γραφεί ανάλογα με την εξίσωση (3.1).

$$P((a_1, v_1, d_1), (a_2, v_2, d_2), ..., (a_N, v_N, d_N)) = \prod_{i=1}^{N} (1 - p_i) p_i^{a_i + v_i + (1+s)d_i} \quad (3.5)$$

Η εξίσωση (3.5) συνδέει ακριβώς την συνάρτηση πιθανότητας της κατάστασης του συστήματος, με την ικανότητα των antiviruses να εξουδετερώνουν viruses, που μεταφράζεται με την παράμετρο s.

Κεφάλαιο 4

Μοντελοποίηση ταυτόχρονης εξέλιξης DNS ιών και λογισμικών προστασίας σε IPv6 δίκτυα

4.1 Εισαγωγή

Ένα νέο ερευνητικό ενδοιαφέρον έχει προκύψει από τον εντοπισμό απειλών που διαδίδονται μέσω πληροφοριών που παίρνουν από τους εξυπηρετητές (Domain Name Servers (DNS)), με σκοπό την εύρεση ευπαθών υπολογιστών για να μολύνουν. Αυτοί οι ιοί χρησιμοποιούν έναν γεννήτορα τυχαίων αλφαριθμητικών (strings) για την εύρεση πιθανών πραγματικών διευθύνσεων και στην συνέχεια θέτουν ερώτημα στους εξυπηρετητές DNS με σκοπό την αποκόμιση της πραγματικής ηλεκτρονικής διεύθυνσης, αν το τυχαίο string που έχει παραχθεί τυγχάνει να συμπίπτει με ένα πραγματικό. Σε αυτό το κεφάλαιο προτείνουμε και αναλύουμε μοντέλα για την ταυτόχρονη εξέλιξη των ιών και των λογισμικών προστασίας που κινούνται στο δίκτυο με σκοπό τον εντοπισμό και την διακοπή της μόλυνσης. Τα προτεινόμενα μοντέλα θεωρούν ότι τα λογισμικά προστασίας γνωρίζουν το δίκτυο και συνεπώς τις πραγματικές IP διευθύνσεις, αλλά και λογισμικά προστασίας τα οποία δεν το γνωρίζουν και χρησιμοποιούν και αυτά τους εξυπηρετητές DNS με σκοπό να διαδοθούν. Τέλος έχουμε επεκτείνει το μοντέλο με την προσθήκη των dummy honeypot servers, οι οποίοι είναι εικονικοί εξυπηρετητές και προκαλούν καθυστερήσεις στις απαντήσεις των ερωτημάτων για αποκόμιση IP διευθύνσεων. Θα δείξουμε ότι με αυτή την προσθήκη δεν πετυχαίνουμε καλύτερη αντιμετόπιση της μόλυνσης του δικτύου, αφού οι dummy servers έχουν μικρή επίδραση στο ρυθμό διάδοσης των ιών.

4.2 IPv6 DNS απειλές

Ένας DNS ιός είναι ένας ιός ο οποίος χρησιμοποιεί τους εξυπηρετητές DNS προκειμένου να διαδοθεί. Ειδικότερα υπάρχει ένας γεννήτορας ηλεκτρονικών διευθύνσεων ο οποίος παράγει strings, (string generator) που είναι πιθανές πραγ-

ματικές διευθύνσεις (actual hosts) του διαδικτύου. Οι περισσότερες ηλεκτρονικές διευθύνσεις αποτελούνται τυπικά από κοινές λέξεις που χωρίζονται απο τελείες, π.χ www.yahoo.com. Οι πιο πολλές από αυτές τις λέξεις δύναται να βρεθούν σε ένα λεξικό. Οπότε ένας "έξυπνος" ιός μπορεί να κάνει επιθέσεις σε πιθανές ηλεκτρονικές διευθύνσεις που παράγει ο γεννήτορας του μέσα από λεξικό (dictionary attack). Με αυτήν την τεχνική ο γεννήτορας μπορεί να παράγει πραγματικές διευθύνσεις με μεγάλη πιθανότητα, και στην συνέχεια χρησιμοποιεί τους εξυπηρετητές DNS με σκοπό την ανάκτηση των αντίστοιχων IP adresses. Αν η ηλεκτρονική διεύθυνση που επιστρέφεται από τους εξυπηρετητές DNS είναι ευπαθής, ο ιός επιτίθεται και την μολύνει. Η πρώτη ιδέα για μοντελοποίηση διάδοσης ιών που παίρνουν πληροφορίες από εξυπηρετητές DNS βρίσκεται στην εργασία [13] του Κερομύτη. Στην εργασία αυτή παρουσιάζεται ένα απλό μοντέλο βασισμένο στην εξίσωση διάδοσης

$$v' = \sigma \xi v(1 - v) \tag{4.1}$$

όπου v είναι το ποσοστό των μολυσμένων υπολογιστών του δικτύου που επιτίθεται ο ιός, σ είναι η πιθανότητα πετιχυμένης εύρεσης διεύθυνσης από τον ιό χρησιμοποιώντας τον εξυπηρετητή DNS και ξ είναι ο μέσος ρυθμός που οι μολυσμένοι υπολογιστές σαρώνουν το δίκτυο για εύρεση ευπαθών υπολογιστών. Για να μοντελοποιήσουμε την παράμετρο ξ μελετάμε τις αναμενόμενες καθυστερήσεις που μπορεί να υπάρξουν στο δίκτυο χρησιμοποιώντας θεωρία ουρών.

Έστω χ το σύνολο όλων των πιθανών strings που παράγει ο γεννήτορας του ιού (string generator). χ^{target} το υποσύνολο του χ των πραγματικών ηλεκτρονικών διευθύνσεων. Για παράδειγμα ένας DNS ιός που χρησιμοποιεί το string generator για να βρει πιθανές ηλεκτρονικές διευθύνσεις, μπορεί να μολύνει μόνο ιστοσελίδες από το σύνολο χ^{target}. Φυσικά, υπάρχουν και ενεργές ηλεκτρονικές διευθύνσεις που δεν μπορούν να παραχθούν από τον γεννήτορα και βρίσκονται εκτός του υποσυνόλου χ. Αυτές οι διευθύνσεις δεν μπορούν να μολυνθούν από έναν τέτοιο ιό. Από την μεριά του DNS ιού, οι ευπαθείς διευθύνσεις είναι αυτές που το όνομα τους περιέχεται σαν string μέσα στο σύνολο χ^{target}. Η πιθανότητα για ένα string που παράγεται από ένα string generator να είναι ενεργή ηλεκτρονική διεύθυνση είναι

$$\sigma = \frac{\chi^{target}}{\chi}$$

Οι εξυπηρετητές DNS παρέχουν μια απεικόνιση από αλφαβητικά ονόματα σε αριθμητικές διευθύνεις IP address που χαρακτηρίζουν τους υπολογιστές στο διαδίκτυο. Σε μία τυπική ερώτηση στον DNS, ο πελάτης client χρειάζεται να αποκομίσει την IP address, προκειμένου να μπορεί να συνδεθεί. Αρχικά συνδέεται με έναν τοπικό εξυπηρετητή (local resolver), ο οποίος είναι ένας εξυπηρετητής DNS στην ίδια περιοχή με τον πελάτη. Στην συνέχεια αυτός ο τοπικός εξυπηρετητής επικοινωνεί με τους root name-servers μέχρι να απαντηθεί το αρχικό ερώτημα για την IP address. Έπειτα ο root name-server απαντάει στον τοπικό εξυπηρετητή, και αυτός με την σειρά του στέλνει την απάντηση στον πελάτη, αφού πρώτα κρατήσει ένα αντίγραφο στην μνήμη του για πιθάνη μετέπειτα επανάληψη του ίδιου ερωτήματος, από πελάτη στην ίδια περιοχή με τον τοπικό εξυπηρετητή. Ο χρόνος που χρειάζεται για να απαντηθεί ένα ερώτημα σε έναν εξυπηρετητή DNS περιέχει καθυστερήσεις μεταξύ του πελάτη και του τοπικού εξυπηρετητή d_{local}, και καθυστερήσεις μεταξύ του τοπικού εξυπηρετητή και των root name-servers, $d_{internet}$. Συνεπώς

έχουμε

$$d = d_{local} + d_{internet} \tag{4.2}$$

Η καθυστέρηση $d_{internet}$ δημιουργείτε από παράγοντες όπως τα retransmissions τα Timeouts και τα DNS cache hit/miss. Αν ένα πακέτο ερωτήματος χαθεί, ο πελάτης τυπικά περιμένει για ένα χρονικό διάστημα timeout T, πριν στείλει ξανά το ίδιο πακέτο retransmission. Αν Pr είναι η πιθανότητα να χαθεί το πακέτο, (οπότε έχουμε retransmission), η αναμενόμενη DNS καθυστέρηση για ένα ερώτημα είναι

$$d_{av} = d_{local} + d_{internet} + \frac{p_r T}{1 - p_r} \tag{4.3}$$

Κάποια ερωτήματα απαντώνται αμέσως από τον τοπικό εξυπηρετητή όταν η απάντηση υπάρχει ήδη μέσα στην μνήμη του, και δεν υπάρχει ανάγκη να φτάσει το ερώτημα στους εξυπηρετητές DNS. Σε αυτήν την περίπτωση η καθυστέρηση είναι μόνο η d_{local}. Άν p_{DCH} είναι η πιθανότητα η απάντηση να υπάρχει στην μνήμη DNS cache hit, τότε η μέση καθυστέρηση είναι

$$d_{av}^{cached} = p_{DCH} d_{local} + (1 - p_{DCH}) d_{av} \tag{4.4}$$

Έστω τ_f ο μέσος χρόνος για να γίνει η μόλυνση. Τότε η μέση καθυστέρηση για ένα ερώτημα που οδηγεί σε μόλυνση είναι

$$d_{av}^{cached} + \tau_f \ .$$

Οπότε ο αποτελεσματικός μέσος χρόνος καθυστέρησης για έναν DNS ιό είναι:

$$d_{eff} = \sigma(d_{av}^{cached} + \tau_f) + (1 - \sigma)(d_{av}) \tag{4.5}$$

και ο αποτελεσματικός χρόνος σάρωσης ενός DNS ιού είναι $\xi = \frac{1}{d_{eff}}$. Για να μελετήσουμε την συμπεριφορά των εξυπηρετητών τους μοντελοποιούμε σαν συστήματα ουρών $M/M/1/K$. Αυτά τα συστήματα όπως έχουμε αναφέρει και στο προηγούμενο κεφάλαιο έχουν εκθετικό ρυθμό εξυπηρέτησης μ, K είναι ο μέγιστος αριθμός εργασιών που μπορεί να δεχτεί μία ουρά (σε αυτό το κεφάλαιο σαν εργασία εννοούμε ένα ερώτημα), και λ είναι ο ρυθμός αφίξεων των εργασιών. Η πιθανότητα μία ουρά να έχει i ερωτήματα να περιμένουν να απαντηθούν είναι

$$\pi(i) = \frac{(1 - p)p^i}{1 - p^{K+1}}$$

όπου p είναι ο ρυθμός χρησιμοποίησης της ουράς και ισούται με $p = \frac{\lambda}{\mu}$. Κάποιες φορές όταν υπάρχει πολύ κίνηση στο δίκτυο, πολλά ερωτήματα δεν μπορούν να εξυπηρετηθούν εξ' αιτίας της εξάντλησης της μνήμης της ουράς. Η αναμενόμενη πιθανότητα για να απορριφθεί ένα ερώτημα είναι

$$E[loss] = \pi(K) = \frac{(1 - p)p^K}{1 - p^{K+1}}$$

Η πιθανότητα αυτή είναι ένα καλό κριτήριο για να μοντελοποιήσουμε την περίπτωση όπου έχουμε retransmission σε ένα ερώτημα, οπότε $p_r = E[loss]$. Για να μοντελοποιήσουμε την καθυστέρηση $d_{internet}$, χρησιμοποιούμε τον αναμενόμενο χρόνο

απάντησης $E[X_a]$ από όλα τα ερωτήματα, και αυτό συμβαίνει όταν οι ουρές δεν είναι γεμάτες.

$$E[X_a] = E[X|accepted] = \frac{E[X]}{1 = E[loss]}$$

$$= \frac{1}{\mu}[\frac{1}{1-p} - \frac{Kp^K}{1-p^K}]$$

Για την παράμετρο λ μπορούμε να πούμε ότι είναι ίση με τον συνολικό αριθμό των πελατών που κάνουν ερωτήματα, σε αυτή την περίπτωση μόνο των μολυσμένων χρηστών επί τον ρυθμό σάρωσης, οπότε $p = \frac{aN\xi}{\mu}$.

4.3 DNS ιοί και λογισμικά προστασίας που γνωρίζουν το δίκτυο

Από τις εξισώσεις της προηγούμενης ενότητας μοντελοποιήσαμε κάποιες παραμέτρους από την απλή διαφορική εξίσωση (4.1) για την διάδοση των DNS ιών. Τώρα θα προσπαθήσουμε να επεκτείνουμε το μοντέλο σε ένα σύστημα δύο διαφορικών εξισώσεων που περιγράφουν την μεταβολή των πληθυσμών των ιών και των λογισμικών προστασίας σε ένα IPv6 δίκτυο.

Τώρα παράγουμε τις διαφορικές εξισώσεις των v και a, όπου είναι το ποσοστό των μολυσμένων και των καθαρών υπολογιστών στο δίκτυο. Ας κάνουμε πρώτα μερικές υποθέσεις. Έστω ότι το δίκτυο αποτελείται από N υπολογιστές. Κάθε υπολογιστής μπορεί να βρίσκεται σε μία από τις τρεις ακόλουθες πιθανές καταστάσεις: καθαρή, ευπαθή ή μολυσμένη. Οι ιοί διαδίδονται με ρυθμό ανάλογο του $\sigma\xi$. Ο πληθυσμός των ιών αυξάνεται κάθε φορά που ένας μολυσμένος υπολογιστής κάνει πετυχημένο ερώτημα σε ευπαθή υπολογιστή. Ο ρυθμός μόλυνσης είναι ανάλογος του ποσοστού των ευπαθών υπολογιστών $1 - v - a$. Ο πληθυσμός των ιών μειώνεται κάθε φορά που ένα λογισμικό προστασίας εγκαθίσταται σε έναν μολυσμένο υπολογιστή και τον "καθαρίζει". Τα λογισμικά προστασίας γνωρίζουν το δίκτυο οπότε διαδίδονται στα έγκυρα IP adresses. Η αύξηση τους εξαρτάται από δύο παραμέτρους: την λ_u όπου είναι ο ρυθμός ενημέρωσης των λογισμικών προστασίας, και την h όπου είναι παράμετρος που εξαρτάται στον ανθρώπινο παράγοντα να προστατεύεται από τις επιθέσεις ιών. Ο πληθυσμός των καθαρών υπολογιστών αυξάνεται κάθε φορά που ένα λογισμικό προστασίας εγκαθίσταται σε έναν υπολογιστή, είτε είναι μολυσμένος είτε ευπαθής, όπου αυτό το ποσοστό αυτών των υπολογιστών είναι $1 - a$. Στο μοντέλο μας περιέχεται και μία παράμετρο f που μοντελοποιεί την αδιαφορία των χρηστών να προστατεύουν τους υπολογιστές τους.

Έστω $I(t)$ ο πληθυσμός των μολυσμένων υπολογιστών και $A(t)$ ο πληθυσμός των καθαρών υπολογιστών την χρονική στιγμή t. Θεωρούμε ένα απειροελάχιστο χρονικό διάστημα dt. Οι νέες μολύνσεις κατά την διάρκεια αυτή της χρονικής περιόδου είναι

$$I(t+dt) - I(t) = I(t)\sigma\xi(1-v-a)dt - A(t)\lambda_u hvdt \Rightarrow$$

$$I'(t) = I(t)\sigma\xi(1-v-a) - A(t)\lambda_u hv$$

Οπότε έχουμε:

$$v' = v\sigma\xi(1-v-a) - va\lambda_u h \tag{4.6}$$

Η αντίστοιχη διαφορική εξίσωση για το ποσοστό των υπολογιστών που περιέχουν λογισμικά προστασίας παράγεται ως εξής:

$$A(t+dt) - A(t) = A(t)\lambda_u h(1-a)dt - fA(t)dt \Rightarrow$$

$$A'(t) = A(t)\lambda_u h(1-a) - fA(t) \Rightarrow$$

και παίρνουμε την αντίστοιχη διαφορική εξίσωση:

$$a' = a\lambda_u h(1-a) - fa \tag{4.7}$$

4.4 DNS ιοί και DNS λογισμικά προστασίας

Τώρα παρουσιάζουμε μια μεταβολή του προηγούμενου μοντέλου όπου έχουμε λογισμικά προστασίας τα οποία δεν γνωρίζουν ούτε αυτά το δίκτυο και κάνουν ερωτήματα στους εξυπηρετητές DNS για να πάρουν έγκυρες διευθύνσεις. Τα λογισμικά προστασίας προσπαθούν να καθαρίσουν τους μολυσμένους υπολογιστές ή να προστατεύσουν τους ευπαθής για να γνωστοποιήσουν την διεύθυνση τους. Υποθέτουμε ότι τα λογισμικά προστασίας δεν έχουν το ίδιο ρυθμό επιτυχίας με τους DNS ιούς επειδή οι ιοί προσπαθούν να διαδοθούν πολύ γρήγορα και σε πολύ μικρό χρονικό διάστημα. Ακόμα βρίσκουν πιο έξυπνο τρόπο να εξαπλώνονται και να αποσπούν πληροφορίες από τους εξυπηρετητές DNS. Είναι γενικότερα αποδεκτό ότι τα λογισμικά προστασίας είναι λίγο πιο "πίσω" από τους ιούς. Για να μοντελοποιήσουμε αυτή την υπόθεση θεωρούμε ότι ο γεννήτορας των strings που έχουν τα λογισμικά προστασίας έχει πιθανότητα $\sigma_1 < \sigma$ μικρότερη από αυτή των ιών για να αποκομίσει έγκυρη διεύθυνση. Τα ερωτήματα των λογισμικών προστασίας απαντώνται από τους ίδιους εξυπηρετητές DNS που απαντάνε και στα ερωτήματα των ιών. Οπότε η μέση καθυστέρηση που υπάρχει για τις απαντήσεις είναι η ίδια και για τα δύο είδη. Δηλώνουμε αυτή την μέση χρονική καθυστέρηση με ξ.

Τώρα θα παράγουμε τις διαφορικές εξισώσεις για τα a και v. Ο ρυθμός διάδοσης των ιών και οι παράμετροι είναι ίδιοι με το προηγούμενο μοντέλο αλλά ο ρυθμός εξολόθρευσης είναι διαφορετικός εξ' αιτίας κάποιων πολύπλοκων μηχανισμών: i) τα λογισμικά προστασίας ενημερώνονται από τους χρήστες και ii) τα λογισμικά προστασίας μετακινούνται με πληροφορίες από τους εξυπηρετητές DNS. Από αυτούς τους δύο παράγοντες παίρνουμε έναν ολικό ρυθμό εξολόθρευσης $\lambda_u h + \sigma_1 \xi$, όπου είναι και ο ρυθμός που διαδίδονται οι ιοί. Σε ένα απειροελάχιστο χρονικό διάστημα dt συμβαίνουν οι εξής μεταβολές στους πληθυσμούς:

$$I(t+dt) - I(t) = I(t)\sigma\xi(1-v-a)dt - A(t)v(\lambda_u h + \sigma_1\xi)dt \Rightarrow$$

$$I'(t) = I(t)\sigma\xi(1-v-a) - A(t)v(\lambda_u h + \sigma_1\xi)$$

και παίρνουμε

$$v' = v\sigma\xi(1-v-a) - av(\lambda_u h + \sigma_1\xi) \tag{4.8}$$

Για τα antivirus παίρνουμε την αντίστοιχη διαφορική εξίσωση:

$$A(t+dt) - A(t) = A(t)(1-a)(\lambda_u h + \sigma_1\xi)dt - fA(t)dt \Rightarrow$$

$$A'(t) = A(t)(1-a)(\lambda_u h + \sigma_1\xi) - fA(t)$$

$$a' = a(1-a)(\lambda_u h + \sigma_1\xi) - fa. \tag{4.9}$$

4.5 Εισαγωγή dummy "honeypot" servers

Το τελευταίο μοντέλο που θα παρουσιάσουμε βασίζεται στην ιδέα ότι μπορούμε να εμποδίσουμε την διάδοση των ιών, με την εισαγωγή των dummy "honeypot" servers. Αυτοί οι εξυπηρετητές μοιάζουν με πραγματικούς DNS servers εξωτερικά, αλλά δεν δίνουν απαντήσεις σε ερωτήματα για διευθύνσεις. Το μόνο που κάνουν είναι να προκαλούν καθυστερήσεις, αφού γίνονται συνέχεια retransmissions μέχρι να αποκομισθεί η διεύθυνση.

Οι dummy servers δεν μπορούν να ανιχνευθούν αφού δεν φαίνεται ότι έχουν διαφορά με τους πραγματικούς εξυπηρετητές. Πρακτικά αυτοί οι εικονικοί servers αυξάνουν την πιθανότητα να έχουμε retransmission. Αυτό επιδρά στην μείωση του ρυθμού διάδοσης των ιών. Η αρνητική πλευρά αυτής της ιδέας είναι ότι προκαλούνται και καθυστερήσεις σε "νόμιμες" ερωτήσεις των χρηστών, κάτι που είναι υποφερτό αφού ο σκοπός είναι να προστατεύεται το δίκτυο. Για να εισάγουμε τους dummy servers στο μοντέλο μας, τροποποιούμε την παράμετρο p_r στην εξίσωση της καθυστέρησης d_{av}, με έναν όρο $p_{rd} = p_r + p_d$, όπου p_d είναι το ποσοστό των dummy servers που υπάρχουν στο δίκτυο.

$$d_{av} = d_{local} + d_{internet} + \frac{p_{rd}T}{1 - p_{rd}} \tag{4.10}$$

Η βασική διαφορά αυτής της εξίσωσης με αυτή του προηγούμενου μοντέλου, βρίσκεται στον όρο του ρυθμού σάρωσης ξ. Τώρα θα κάνουμε μία υπόθεση ότι τα DNS λογισμικά προστασίας γνωρίζουν τους πραγματικούς εξυπηρετητές και δεν έχουν παραπάνω καθυστερήσεις για να πάρουν απαντήσεις. Οι καθυστερήσεις που προκαλούνται από τους εικονικούς εξυπηρετητές έχουν επίδραση στις καθυστερήσεις των απαντήσεων από τους ιούς. Οι διαφορικές εξισώσεις που παράγονται είναι οι ακόλουθες:

$$v' = v\sigma\xi_d(1 - v - a) - av(\lambda_u h + \sigma_1 \xi) \tag{4.11}$$

$$a' = a(1 - a)(\lambda_u h + \sigma_1 \xi) - fa \tag{4.12}$$

όπου ξ_d είναι ο ρυθμός σάρωσης των DNS ιών, που ενσωματώνει την πιθανότητα να "χτυπήσουν" έναν dummy server.

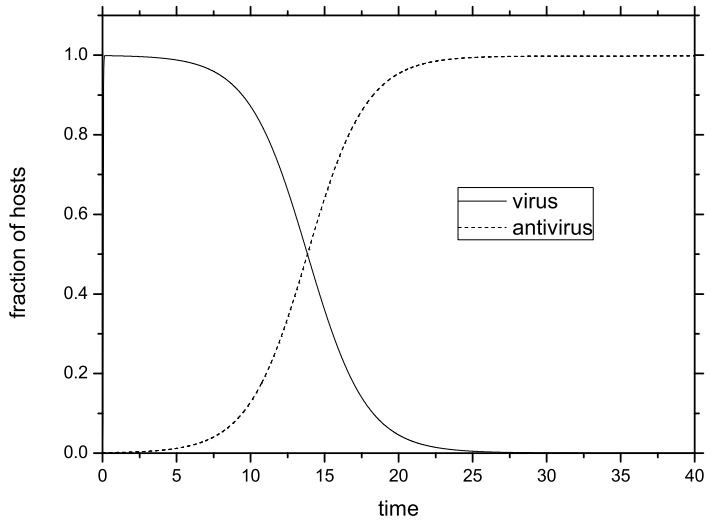
4.6 Αριθμητικά αποτελέσματα και συμπεράσματα

Σε αυτήν την ενότητα παρουσιάζουμε κάποια αριθμητικά αποτελέσματα από την λύση των διαφορικών εξισώσεων των μοντέλων μας. Για τις αριθμητικές προσεγγίσεις χρησιμοποιούμε μέθοδο Runge-Kutta τέταρτης τάξης. Για όλα τα ακόλουθα αριθμητικά αποτελέσματα έχουμε χρησιμοποιείσει τις ίδιες αρχικές τιμές για τους πληθυσμούς των ιών και των λογισμικών προστασίας, και όλες οι παραμέτροι έχουν τιμές από πραγματικά στοιχεία IPv6 δικτύων. Όπως για παράδειγμα για τον ρυθμό διάδοσης των ιών παίρνουμε τιμές από παρατηρήσεις του ιού slammer. Ακόμα υποθέτουμε ότι όλα τα ερωτήματα καλύπτουν ένα χώρο διευθύνσεων της τάξης του 2^{128}.

Στο πρώτο διάγραμμα 4.1 έχουμε αποτελέσματα από το πρώτο μοντέλο (εξισώσεις (4.6),(4.7» όπου έχουμε DNS ιούς και λογισμικά προστασίας που γνωρίζουν

Parameter	Description	Value
σ	virus probability of successful scan	0.02
ξ	scan rate	$4000/sec$
$\lambda_u h$	growth coefficient	0.5
f	decay of antivirus	0.001
σ_1	antivirus probability of successful scan	0.0002
ξ_d	scan rate in net with dummy servers	$3000/sec$

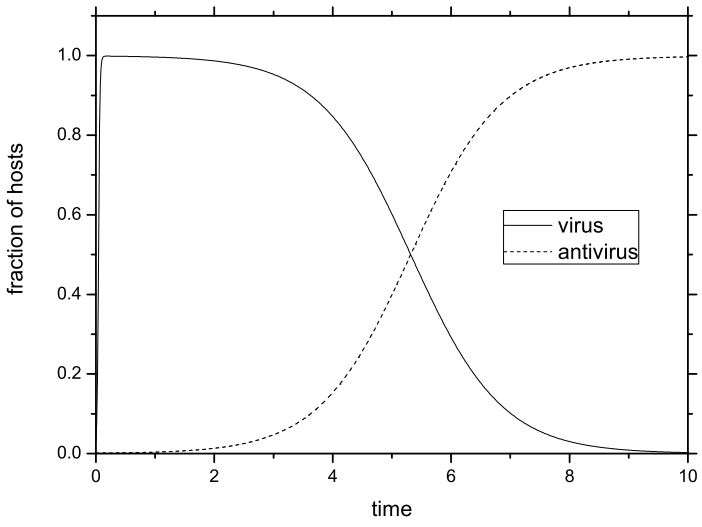
Πίνακας 4.1: Πίνακας παραμέτρων



Σχῆμα 4.1:

το δίκτυο και τις πραγματικές διευθύνσεις. Παρατηρούμε ότι ο πληθυσμός των ιών παρουσιάζει μια απότομη αύξηση στην αρχή της διάδοσης και σταδιακά καλύπτουν όλο το δίκτυο (ο slammer και ο witty ιός παρουσίασαν τόσο άμεση αύξηση στο πλήθος τους σε χρονικό διάστημα που δεν μπορούσε να υπάρξει ανθρώπινη αντίδραση). Μετά από κάποιο χρονικό διάστημα τα λογισμικά προστασίας αρχίζουν να αυξάνονται (όπου η αύξηση τους εξαρτάται από τον ρυθμό αύξησης των ιών [23, 24]), ενώ την ίδια χρονική στιγμή ο πληθυσμός των ιών μειώνεται. Ο πληθυσμός των antivirus αυξάνεται μέχρι τεικά να καλύψουν όλο το δίκτυο που μελετάμε. Σε αυτό το σημείο λέμε ότι ο ιός έχει πλέον εξαλειφθεί.

Στο δεύτερο διάγραμμα 4.2 έχουμε αποτελέσματα από το μοντέλο που έχουμε επεκτείνει (εξισώσεις (4.6), (4.9)) όπου έχουμε δύο είδη λογισμικών προστασίας, ένα που γνωρίζει το δίκτυο και ένα που χρεισιμοποιεί τους εξυπηρετητές DNS. Φυσικά ενδοιαφερόμαστε μόνο για το συνολικό πλήθος και των δύο ειδών, αφού και τα δύο μαζί προστατεύουν το δίκτυο. Παρατηρούμε πάλι ότι υπάρχει απότομη αύξηση του πλήθους των ιών στην αρχή της διάδοσης, αλλά η εξολόθρευση τους αρχίζει λίγο αργότερα από την προηγούμενη περίπτωση. Αυτό συμβαίνει από το γεγονός ότι τώρα έχουμε δύο είδη λογισμικών προστασίας όπου το ένα εξαρτάται

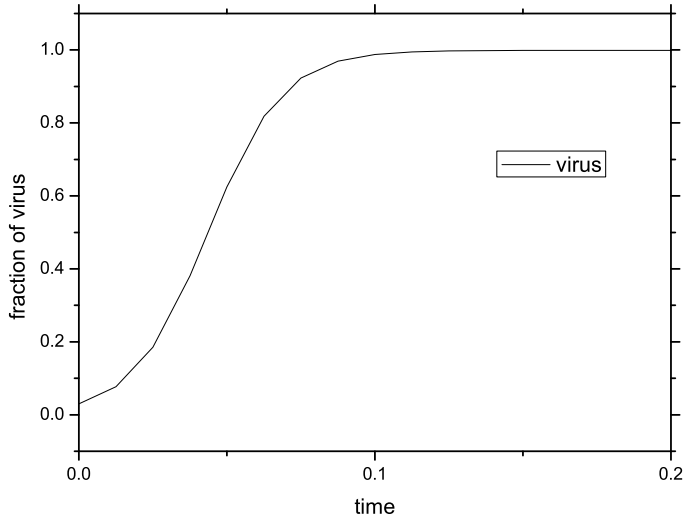


Σχῆμα 4.2:

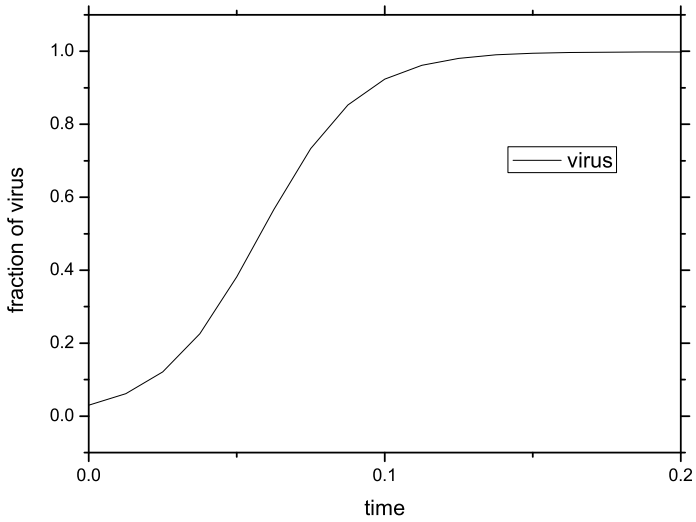
από τους χρήστες και το άλλο επηρεάζεται από την ταχύτητα που δίνουν απαντήσεις οι εξυπηρετητές DNS, οι οποίοι τώρα επιβαρύνονται από τα ερωτήματα των λογισμικών προστασίας και υπάρχει υπερφόρτωση στο δίκτυο.

Στα διαγράμματα 4.3, 4.4 συγκρίνουμε τους ρυθμούς της αύξησης των πληθυσμών των ιών όταν έχουμε εισάγει τους dummy servers. Παρατηρούμε ότι και στις δύο περιπτώσεις οι ιοί καλύπτουν όλο το δίκτυο που μελετάμε ενώ η εξολόθρευση αρχίζει την ίδια χρονική στιγμή αφού οι dummy servers δεν επηρεάζουν την αύξηση του πλυθυσμού. Στην δεύτερη περίπτωση πάντως παρατηρείται μια πιο ομαλή αύξηση του πλήθους των ιών. Αυτό συμβαίνει διότι οι dummy servers προκαλούν καθυστερήσεις αφού δεν επιστρέφουν απαντήσεις και έχουμε retransmissions.

Πάντως δεν φαίνεται ότι οι dummy servers μπορούν να εμποδίσουν την διάδοση των ιών. Παρουσιάζουν βέβαια μια μικρή καθυστέρηση στον ρυθμό αύξησης της διάδοσης που δεν αφήνει όμως περιθώρια αντίδρασης στους χρήστες. Το γενικό συμπέρασμα είναι ότι οι dummy servers δεν αποτελούν σοβαρό αντίμετρο για την καταπολέμηση των ιών σε IPv6 δίκτυα.



Σχῆμα 4.3:



Σχῆμα 4.4:

Κεφάλαιο 5

Μοντέλο κυνηγών - θηράματος με αλληλεπιδράσεις κυνηγών

5.1 Εισαγωγή

Στην εργασία [4] εξετάζεται ο ρόλος της επικοινωνίας σε τεχνητές κοινωνίες. Στην εργασία [20] εξετάζονται οι δυναμικές δημιουργίες ομάδων και δικτύων επικοινωνίας σε περιβάλλοντα τα οποία χαρακτηρίζονται από έλλειψη πόρων. Στην εργασία [14] αναφέρεται ότι στα θερμοδυναμικά συστήματα δημιουργείται τάξη η οποία διακρίνεται σε τρία επίπεδα. Το επίπεδο της κοινότητας ή του οικοσυστήματος, το επίπεδο της ταυτόχρονης εξέλιξης των ειδών και τέλος την εξέλιξη της ίδιας της εξελικτικής διαδικασίας. Το πρώτο επίπεδο ειδικότερα αναφέρεται στην δυναμική πληθυσμών σε οικοσυστήματα που συμπεριλαμβάνονται κυνηγοί και θηράματα. Η πρωτοπόρα εργασία των Lotka- Volterra έθεσε τις βάσεις για την δημιουργία απλών μοντέλων που καθορίζουν την αυξομείωση των πληθυσμών μέσα από την αλληλεπίδραση τους. Το δεύτερο επίπεδο το οποίο εμφανίζεται σε μεγαλύτερη κλίμακα χρόνου απ΄ την κλίμακα που συντελούνται οι μεταβολές στο οικοσύστημα είναι αυτό της ταυτόχρονης εξέλιξης. Τα είδη δεν εξελίσσονται μόνο σύμφωνα με τις δυνατότητες τους αλλά και μέσα από αλληλεπιδράσεις με άλλα είδη. Για παράδειγμα, οι θηρευτές και τα θηράματα αλληλεπιδρούν. Το κυρίαρχο πλαίσιο για την κατανόηση της αλληλεπίδρασης των ειδών είναι αυτό της θεωρίας παιγνίων. Ο John Maynard Smith γενίκευσε την έννοια ισορροπίας Nash με την έννοια της εξελικτικά σταθερής στρατηγικής. Τέλος στις εργασίες [26, 17], το παίγνιο των γερακιών και περιστεριών μελετάται σε σχέση με την εξέλιξη των στρατηγικών που επιλέγουν οι πληθυσμοί, όχι όμως ως συνάρτηση και ενός τρίτου πληθυσμού της λείας. Σε αυτό το κεφάλαιο προτείνουμε ένα γενικό μοντέλο για την μελέτη της ταυτόχρονης εξέλιξης των πληθυσμών πολλαπλών κυνηγών, καθώς συναγωνίζονται για το ίδιο θήραμα. Αυτό το μοντέλο ενσωματώνει και την επίδραση που έχουν όλες οι αλληλεπιδράσεις μεταξύ των ειδών των κυνηγών, πάνω στον πληθυσμό του θηράματος. Το προτεινόμενο μοντέλο επεκτείνει τα κλασσικά βιολογικά μοντέλα όπως το Lotka-Voltera και το May-Leonard με στοιχεία από την θεωρία παιγνίων που μοντελοποιούν την επίδραση των κυνηγών πάνω στο θήραμα. Το μοντέλο περιέχει έναν μεγάλο αριθμό από παραμέτρους, οι

οποίοι είναι οι συντελεστές των αλληλεπιδράσεων των κυνηγών, καθώς συναντώνται με σκοπό να επιτεθούν στο θήραμα. Αυτή η μεγάλη ποικιλότητα μας επιτρέπει να κάνουμε πολλά υποθετικά σενάρια, γι' αυτό και παρουσιάζουμε διάφορα αριθμητικά αποτελέσματα (αφού το μοντέλο μας είναι μη- γραμμικό), προσπαθώντας να προσομοιώσουμε καλύτερα την συμπεριφορά των ειδών, με διαφορετικές αρχικές συνθήκες για τους πληθυσμούς και διάφορες τιμές για τις παραμέτρους αλληλεπίδρασης.

5.2 Το γενικό μοντέλο πολλαπλών κυνηγών- θηράματος

Το πιο γνωστό μοντέλο για την εξέλιξη δύο ειδών που αλληλεπιδρούν μεταξύ τους είναι το Lotka-Voltera. Αυτό το μοντέλο αποτελείται από δύο μη- γραμμικές διαφορικές εξισώσεις που μοντελοποιούν την δυναμική της εξέλιξης των δύο πληθυσμών. Μία μορφή του μοντέλου αυτού με έναν παράγοντα που ελέγχει τον ρυθμό αύξησης των πληθυσμών των ειδών είναι η παρακάτω [2]:

$$\begin{aligned} x_1' &= x_1(1 - x_1 - \alpha_1 x_2) \\ x_2' &= x_2(1 - \beta_2 x_1 - x_2). \end{aligned} \tag{5.1}$$

Στον πραγματικό κόσμο όμως περισσότερα από δύο είδη ανταγωνίζονται το ένα το άλλο. Ένα μοντέλο που γενικεύει το μοντέλο Lotka-Voltera είναι το May-Leonard για τρία ανταγωνιστικά είδη:

$$\begin{aligned} x_1' &= x_1(1 - x_1 - \alpha_1 x_2 - \beta_1 x_3) \\ x_2' &= x_2(1 - \beta_2 x_1 - x_2 - \alpha_2 x_3) \\ x_3' &= x_3(1 - \alpha_3 x_1 - \beta_3 x_2 - x_3) \end{aligned} \tag{5.2}$$

με $x_1(0) > 0, x_2(0)$ και $x_3(0) > 0$, με την υπόθεση ότι $0 < \alpha_i < 1 < \beta_i$.

Ένα στοιχείο που λείπει από τα δύο προηγούμενα μοντέλα είναι πως η αλληλεπίδραση δύο ή περισσοτέρων ειδών, επιδρά στην μεταβολή του πληθυσμού του άλλου είδους. Αυτά τα δύο συστήματα των διαφορικών εξισώσεων μοντελοποιούν την αλληλεπίδραση μεταξύ κάθε δύο ειδών, αλλά όχι την αλληλεπίδραση περισσοτέρων από δύο ειδών, όταν όλα μαζι συναντώνται. Με μια τέτοια στοιχειώδη αλληλεπίδραση μπορούμε να μοντελοποιήσουμε καταστάσεις όπου για παράδειγμα δύο είδη συνεργάζονται με σκόπο να πιάσουν την λεία, ή το αντίθετο.

Έστω $S = \{s_1, \ldots, s_n\}$ ένα σύνολο από n ανταγωνιζόμενα είδη και $X(t) = \{x_1(t), \ldots, x_n(t)\}$ ο πληθυσμός κάθε είδους την χρονική στιγμή t, όπου x_i είναι ο πληθυσμός του είδους s_i. Με S^{i-} συμβολίζουμε το σύνολο $S - \{s_i\}$ και με X^{i-} το σύνολο $X_i - \{x_i\}$. Ακόμα με A_i θα συμβολίζουμε όλα τα υποσύνολα πλήθους i ενός σύνολο A πλήθους n, $1 \leq i \leq n$. Σύμφωνα με αυτούς τους συμβολισμούς μπορούμε να γενικεύσουμε το μοντέλο May-Leonard ως εξής:

$$\begin{aligned} x_i' = x_i[A_i - x_i - \sum_{j=1, j\neq i}^{n} a_{\{j,j\}} x_j x_j - \\ \sum_{j=1}^{n-1} \sum_{\{k_1,\ldots,k_j\} \in X_j^{i-}} a^i_{\{k_1,\ldots,k_j\}} x_{k_1} \cdots x_{k_j}] \end{aligned} \tag{5.3}$$

με $a^i_{\{k_1,\ldots,k_j\}} > 0, 1 \leq i \leq n$.

Το νέο στοιχείο στην εξίσωση (5.3) σε σχέση με την (5.2) είναι η ύπαρξη των πολλαπλασιαστικών όρων στα γινόμενα παραπάνω από δύο ειδών. Έστω για παράδειγμα η εξίσωση (5.3) για $i = 1$ και $n = 3$, (δηλαδή για τρία είδη):

$$\begin{aligned} x_1' &= x_1(A_1 - x_1 - a^1_{\{2,2\}}x_2x_2 - \\ &\quad a^1_{\{3,3\}}x_3x_3 - a^1_{\{2\}}x_2 - a^1_{\{3\}}x_3 - a^1_{\{2,3\}}x_2x_3) \Rightarrow \\ x_1' &= \underbrace{x_1(A_1 - x_1)}_{1} - \underbrace{a^1_{\{2\}}x_1x_2}_{2} - \underbrace{a^1_{\{3\}}x_1x_3}_{3} - \\ &\quad \underbrace{a^1_{\{2,3\}}x_1x_2x_3}_{4} - \underbrace{a^1_{\{2,2\}}x_1x_2^2}_{5} - \underbrace{a^1_{\{3,3\}}x_1x_3^2}_{6}. \end{aligned} \tag{5.4}$$

Στην εξίσωση (5.4), οι τρεις πρώτοι όροι από τους τέσσερις υπογραμμισμένους είναι οι όροι που εμφανίζονται και στο κλασσικό μοντέλο Lotka-Voltera αλλά και στο May-Leonard. Ο πρώτος όρος μοντελοποιεί την αύξηση του πληθυσμού x_1 του είδους s_1 που δεν επηρεάζεται από τα άλλα είδη. Ο δεύτερος και ο τρίτος όρος μοντελοποιούν την επίδραση των ειδών s_2 και s_3 αντίστοιχα, πάνω στην αύξηση του είδους s_1. Ο νέος όρος (σε σχέση με τα άλλα δύο μοντέλα), ο τέταρτος, μοντελοποιεί την από κοινού επίδραση των είδων s_2 και s_3, πάνω στο είδος s_1, δηλαδή πως ενεργούν (μάχονται ή συνεργάζονται) και επηρεάζουν τον πληθυσμό του πρώτου είδους. Αν θεωρήσουμε το πρώτο είδος ως θήραμα και τα άλλα δύο ως θηρευτές, η τελευταία εξίσωση μοντελοποιεί την μεταβολή του πληθυσμού του θηράματος καθώς επηρεάζεται ξεχωριστά από τους θηρευτές, αλλά και μαζί. Με την ίδια λογική οι όροι 4,5 και 6 στην εξίσωση (5.4) μοντελοποιούν την επίδραση των ειδών που ανταγωνίζονται για την ίδια λεία. Στο υπόλοιπο αυτού του κεφαλαίου θα επικεντρωθούμε σε συστήματα τριών ειδών, με δύο θηρευτές και ένα θήραμα και θα προσπαθήσουμε να προβλέψουμε την συμπεριφορά τους.

5.3 Θεωρητική μελέτη τριών πλυθησμών

Για $n = 3$ το γενικό μοντέλο είναι το εξής:

$$x_i' = q_i x_i[1 - (\sum_{j=1, j\neq i}^{3} a_{(i,j)}x_j) - w_i x_1x_2x_3],\ i = 1, 2, 3$$

με $q_1 = 1, q_2, q_3 > 0$ και $a_{(i,j)}, w_i \in \mathbb{R}$ για $i, j = 1, 2, 3$. Μπορούμε να γράψουμε το δεξιό μέλος της παραπάνω εξίσωσης ως εξής:

$$f_i(x_1, x_2, x_3) = q_i x_i[1 - \sum_{j=1, j\neq i}^{3} a_{(i,j)}x_j) - w_i x_1x_2x_3] = q_i x_i g_i$$

όπου $g_i = g_i(x_1, x_2, x_3)$, $i = 1, 2, 3$. Με γραμμικοποίηση παίρνουμε τον αντίστοιχο Ιακωβιανό πίνακα για το σύστημα μας:

$$\partial F = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_3} \end{pmatrix}_{(\bar{x_1}, \bar{x_2}, \bar{x_3})} =$$

$$\begin{pmatrix} g_1 - \bar{x_1} & -a_{(1,2)}\bar{x_1} - w_1\bar{x_1}\bar{x_3} & -a_{(1,3)}\bar{x_1} - w_1\bar{x_1}\bar{x_2} \\ -q_2a_{(2,1)}\bar{x_2} - q_2w_2\bar{x_2}\bar{x_3} & q_2g_2 - q_2\bar{x_2} & -q_2a_{(2,3)}\bar{x_2} - q_2w_2\bar{x_1}\bar{x_2} \\ -q_3a_{(3,1)}\bar{x_3} - q_3w_3\bar{x_2}\bar{x_3} & -q_3a_{(3,2)}\bar{x_3} - q_3w_3\bar{x_1}\bar{x_3} & q_3g_3 - q_3\bar{x_3} \end{pmatrix}$$

Τα πρόσημα των πραγματικών μερών των ιδιοτιμών του Ιακωβιανού πίνακα στα σημεία ισσοροπίας, καθορίζουν όπως έχουμε αναφέρει την ευστάθεια των αντίστοιχων σημείων. Στο σημείο ισσοροπίας $(0,0,0)$ έχουμε τις ακόλουθες ιδιοτιμές από τον πίνακα του γραμμικοποιημένου συστήματος $1, w_2, w_3$. Οπότε αυτό το σημείο ισσοροπίας είναι ασταθές, που σημαίνει ότι δεν υπάρχει περίπτωση να εξολοθρευτούν όλοι οι πληθυσμοί. Υπάρχουν άλλες τρεις πιθανές περιπτώσεις όπου έχουμε σημεία ισοοροπίας, αλλά με ισορροπία μόνο στο ένα είδος. Για $i = 1, 2, 3$ τα σημεία είναι

$$\bar{x_i} = 1,\ x_j = 0,\ j = 1, 2, 3 \text{ και } j \neq i$$

και είναι ευσταθή αν

$$a_{(j,i)} > 1,\ j = 1, 2, 3,\ j \neq i$$

Για παράδειγμα το σημείο ισορροπίας $(1,0,0)$ είναι ευσταθές αν $a_{(2,1)} > 1$ και $a_{(3,1)} > 1$. Αυτή η περίπτωση είναι πιθανή, δηλαδή να παραμείνει μόνο ένας πληθυσμός.

Τώρα θα μελετήσουμε σημεία ισσοροπίας όπου έχουμε δύο από τα τρία είδη. Έστω

$x_i = \bar{x_i},\ i = 1, 2, 3,$

$x_j = \bar{x_j},\ j = 1, 2, 3,\ j \neq i,$

$x_k = \bar{x_k},\ k = 1, 2, 3,\ k \neq i,\ k \neq j.$

Τότε το σημείο ισσοροπίας είναι το:

$$\bar{x_i} = \frac{1 - a_{(i,j)}}{1 - a_{(i,j)}a_{(j,i)}}, \bar{x_j} = \frac{1 - a_{(j,i)}}{1 - a_{(i,j)}a_{(j,i)}}, \bar{x_k} = 0$$

Αυτό το σημείο ανήκει στον χώρο $\mathbb{R}^+$ όταν ισχύει $a_{(i,j)} < 1$ και $a_{(j,i)} < 1$, αν $a_{(i,j)}a_{(j,j)} < 1$, ή $a_{(i,j)} > 1$ και $a_{(j,i)} > 1$, αν $a_{(i,j)}a_{(j,j)} > 1$. Από τις παραπάνω δύο περιπτώσεις έχουμε ευστάθεια αν:

$$a_{(i,j)}a_{(j,j)} < 1 \text{ και } g_k(\bar{x_i}, \bar{x_j}, \bar{x_k}) < 0$$

Για παράδειγμα το σημείο ισσοροπίας $(\bar{x_1}, \bar{x_2}, \bar{x_3})$ όπου

$$\bar{x_1} = \frac{1 - a_{(1,2)}}{1 - a_{(1,2)}a_{(2,1)}}, \bar{x_2} = \frac{1 - a_{(2,1)}}{1 - a_{(1,2)}a_{(2,1)}}, \bar{x_3} = 0$$

ανήκει στον $\mathbb{R}^+$ όταν ισχύει $a_{(1,2)} < 1$ και $a_{(2,1)} < 1$, αν $a_{(1,2)}a_{(2,1)} < 1$ ή $a_{(1,2)} > 1$ και $a_{(2,1)} > 1$, αν $a_{(1,2)}a_{(2,1)} > 1$, και έχουμε ευστάθεια αν $a_{(1,2)}a_{(2,1)} < 1$ και $g_3(\bar{x_1}, \bar{x_2}, \bar{x_3}) < 0$.

Για να βρούμε σημεία ισορροπίας αναλυτικά στο εσωτερικό του $\mathbb{R}^+$ είναι πολύ δύσκολο λόγω των μη- γραμμικών όρων στις εξισώσεις του συστήματος μας. Για αυτό θα εξετάσουμε ειδικά συστήματα όπου έχουμε ένα θήραμα (x_1) και δύο κυνηγούς (x_2, x_3). Οι δύο κυνηγοί όταν συναντώνται, μάχονται μεταξύ τους και έτσι έχουμε απώλειες $k_2,\ k_3$ στους πληθυσμούς τους, αλλά όταν βρίσκουν μαζί θήραμα, συνεργάζονται για να το πιάσουν. Μπορούμε ακόμα να υποθέσουμε ότι οι δύο πληθυσμοί έχουν τις ίδιες χωρητικότητες και ρυθμούς αύξησης λ, και

την ίδια ικανότητα m να πιάνουν το θήραμα. Για το θήραμα υποθέτουμε ότι έχει απώλειες k από τους κυνηγούς, και η παράμετρος που αναφέρεται στην συνεργασία είναι ανάλογη της παραμέτρου k. Έστω $q_2 = q_3 = \lambda$, $a_{(1,2)} = a_{(1,3)} = k$, $a_{(2,1)} = a_{(3,1)} = m$, $w_1 = \gamma k$, $a_{(2,3)} = k_2$, $a_{(3,2)} = k_3$ και $w_2 = w_3 = 0$. Οπότε το γενικό μοντέλο παίρνει την ακόλουθη μορφή:

$$x_1' = x_1(1 - x_1 - kx_2 - kx_3 - \gamma k x_2 x_3)$$

$$x_2' = \lambda x_2(1 - x_2 - mx_1 - k_2 x_3)$$

$$x_3' = \lambda x_3(1 - x_3 - mx_1 - k_3 x_2)$$

Σε αυτό το μοντέλο έχουμε $k, k_2, k_3 > 0$ επειδή αυτοί οι όροι επιδρούν στην μείωση του πληθυσμού του θηράματος και $m < 0$ αφού επιδρά στην αύξηση των αντίστοιχων πληθυσμών των θηρευτών.

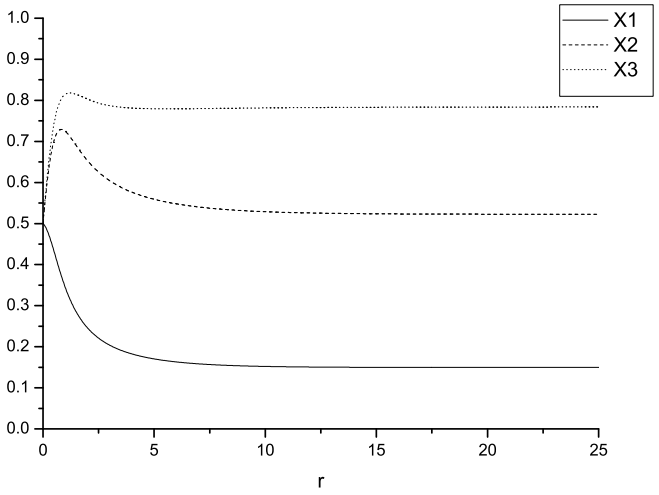
Θα αναλύσουμε το παραπάνω μοντέλο για τιμές στις παραμέτρους $k = 0.4$, $k_2 = 0.8$, $k_3 = 0.7$, $m = -1$, $g = 2$ και $\lambda = 2$. Από προηγούμενη ανάλυση μας βλέπουμε ότι δεν υπάρχει κανένα ευσταθές σημείο ισορροπίας με μόνο ένα είδος. Για παράδειγμα το σημείο $(1, 0, 0)$ είναι ασταθές αφού $m < 1$. Για σημεία ισορροπίας με δύο είδη έχουμε τα ακόλουθα. Εξετάζουμε ισσοροπία για τα πρώτα δύο είδη. Υπάρχει τέτοιο σημείο αφού $km < 1$ και $k, m < 1$. Αυτό το σημείο είναι το $(0.4285, 1.4285, 0)$ και είναι ασταθές επειδή $g_3(\bar{x_1}, \bar{x_2}, \bar{x_3}) = 1 - mx_1 - k3x_2 = 0.426 > 0$. Για σημεία ισορροπίας μεταξύ των άλλων ειδών x_1, x_3 και x_2, x_3 έχουμε παρόμοια αποτελέσματα και ανάλογα ασταθή σημεία ισορροπίας στον $\mathbb{R}^+$.

Για σημείο ισορροπίας με $x_i > 0$, $i = 1, 2, 3$, βρίσκουμε ότι υπάρχει το σημείο $(0.14969, 0.52258, 0.78388)$ και είναι ευσταθές αφού τα πραγματικά μέρη των ιδιοτιμών του αντίστοιχου Ιακωβιανού πίνακα έχουν αρνητικό πρόσημο. Στο παρακάτω διάγραμμα 5.1 επιβεβαιώνονται οι προβλέψεις μας από τα αριθμητικά αποτελέσματα που παίρνουμε από την επίλυση των διαφορικών εξισώσεων με την μέθοδο Runge-Kutta. Βλέπουμε ότι οι τιμές των x_i, $i = 1, 2, 3$ πηγαίνουν πάνω στο σημείο ισορροπίας και παραμένουν σε αυτό στην διάρκεια του χρόνου.

Τώρα θα αλλάξουμε λίγο την τιμή της παραμέτρου k. Έστω $k = 0.6$. Βρίσκουμε ότι αυτό το σύστημα δεν έχει σημείο ισορροπίας στον $\mathbb{R}^+$, αλλά έχουμε ευσταθές σημείο ισορροπίας από τα είδη x_2, x_3. Αυτό είναι το σημείο $(0, 0.4545, 0.6818)$ και παρακάτω έχουμε τα αντίστοιχα αριθμητικά αποτελέσματα στο διάγραμμα 5.2.

5.4 Μελέτη συμπεριφοράς θηρευτών με χρήση θεωρίας παιγίων

Σύμφωνα με το γενικό μοντέλο μας (5.3), ένα στοιχείο που επηρεάζει τον πληθυσμό της λέιας είναι οι από κοινού ενέργειες των δύο θηρευτών. Θα μοντελοποιήσουμε αυτήν την κοινή ενέργεια χρησιμοποιώντας θεωρία παιγνίων, ορίζοντας ένα απλό παιχνίδι μεταξύ των θηρευτών, όταν ανταγωνίζονται για την ίδια λεία [22]. Αυτή είναι και η διαφορά του μοντέλου μας σε σχέση με τα άλλα [7, 27, 21]. Ένα παράδειγμα που είναι κατάλληλο για το μοντέλο μας είναι το γνωστό παιχνίδι γερακιών- περιστεριών. Ένα παιχνίδι μεταξύ δύο ατόμων P_1 και P_2, όπου κάθε άτομο έχει να επιλέξει μεταξύ δύο συμπεριφορών, όταν μάχονται για την λέια. Ο πίνακας κερδών είναι ο παρακάτω (πίνακας 5.1) όπου v είναι η τιμή της λέιας και c το κέρδος απόκτησής της. Αυτός ο πίνακας απεικονίζει τα κέρδη (μπορεί να είναι είτε θετικά είτε αρνητικά) μεταξύ δύο παικτών, τον παίκτη στην σειρά και



Σχῆμα 5.1:

τον παίκτη στην στήλη, όταν επιλέγουν κάθε μία πιθανή κίνησή τους. Στο μέρος όπου οι επιλογές τους διασταυρώνονται, υπάρχει ένα ζευγάρι τιμών. Το πρώτο είναι το κέρδος του παίκτη στην σειρά και το δέυτερο το αντίστοιχο κέρδος του παίκτη στην στήλη. Σύμφωνα με τον πίνακα(5.1), κάθε παίκτης έχει δύο πιθανές

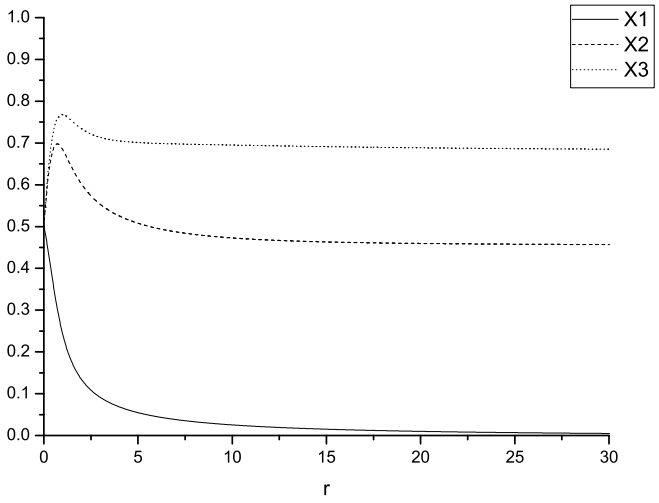
	Η	Δ
Η	$\frac{v-c}{2}, \frac{v-c}{2}$	$v, 0$
Δ	$0, v$	$\frac{v}{2}, \frac{v}{2}$

Πίνακας 5.1: Πίνακας κερδών για το παίγνιο γεράκια- περιστέρια

κινήσεις. Από την μεριά του παίκτη P_1, αν παίξει ως γεράκι και ο όλλος παίκτης ως περιστέρι, τότε θα τα κερδίσει όλα, με ώφελος v. Ακόμα και αν ο παίκτης P_2, παίξει και αυτός ως γεράκι, τότε το κέρδος τους θα είναι μικρότερο του v. Ακόμα και αν το κέρδος απόκτησης της λείας ήταν παραπάνω από την τιμή της, πάλι θα είχαν κάποια απώλεια. Να τονίσουμε ότι και αν το κέρδος της λείας εξαρτάται από τα αποθέματα της, μπορούμε να πιέσουμε τους παίκτες να είναι πιο προσεκτικοί όταν αποφασίσουν να παίξουν το παιχνίδι γεράκια- περιστέρια, να μην επιλέξουν τα γεράκια αν το κόστος είναι μεγάλο, ελπίζοντας ότι και ο άλλος παίκτης θα σκεφτεί λογικά, με σκοπό να μην φτάσουμε σε σημείο να εξαφανιστεί η λεία. Μπορούμε να έχουμε και την τιμή της λείας να είναι ανεξάρτητη από τα αποθέματα της, αλλά σε αυτήν την περίπτωση κάθε παίκτης θα έπαιζε ως γεράκι για να αποκομίσει τέτοιο κέρδος ανεξάρτητα με το κόστος.

Μπορεί να βρεθεί ότι το παίγνιο γεράκια- περιστέρια έχει τρεις ισορροπίες Nash, [8], επιλογές από κάθε αλλαγή κίνησης που δεν αλλάζουν το αποτέλεσμα του κέρδους για τον παίκτη που έκανε την αλλαγή:

(i) Ο παίκτης στην σειρά επιλέγει τα γεράκια ενώ ο παίκτης στην στήλη επιλέγει τα περιστέρια.



Σχῆμα 5.2:

(ii) Ο παίκτης στην σειρά επιλέγει τα περιστέρια ενώ ο παίκτης στην στήλη επιλέγει τα γεράκια.

(iii) Και οι δύο παίκτες επιλέγουν μία μικτή στρατηγική, (οι επιλογές τους βασίζονται σε κάποιες πιθανότητες) με τα γεράκια να παίζονται με πιθανότητα p και τα περιστέρια με πιθανότητα $1-p$. Μπορεί να αποδειχτεί ότι $p = \frac{V}{C}$ και $p = \frac{b-d}{b+c-a-d}$, στην πιο γενική εκδοχή (πίνακας 5.2).

	Η	Δ
Η	a, a	b, c
Δ	c, b	d, d

Πίνακας 5.2: Πίνακας κερδών για το γενικό παίγνιο γεράκια- περιστέρια, με $a < c$ και $d < b$

5.5 Περιγραφή του προβλήματος κατά την θεωρία μέσου πεδίου (mean field)

Σε αυτήν την ενότητα αναλύουμε το μοντέλο μας με χρήση της θεωρίας μέσου πεδίου, σαν πρώτο βήμα για την προσέγγιση της συμπεριφοράς του. Η θεωρία μέσου πεδίου αγνοεί χωροταξικές αλληλεπιδράσεις και θεωρεί ότι ολόκληρος ο πληθυσμός βρίσκεται σε έναν ενιαίο χώρο, με άπειρη χωρητικότητα. Θα ξαναγράψουμε τις εξισώσεις του μοντέλου μας αλλά για λόγους ευκολίας και κατανόησης θα αντικαταστήσουμε τις μεταβλητές x_1, x_2 και x_3, σε r (resource), h (hawck) και

d (dove) αντίστοιχα.

$$\begin{aligned} r' &= r(A_r - r - a^r_{\{h\}}h - a^r_{\{d\}}d - \\ &\quad a^r_{\{h,d\}}hd - a^r_{\{d,d\}}d^2 - a^r_{\{h,h\}}h^2) \\ h' &= h(A_h - h - a^h_{\{d\}}d - a^h_{\{r\}}r - \\ &\quad a^h_{\{d,r\}}dr - a^h_{\{d,d\}}d^2 - a^h_{\{r,r\}}r^2) \\ d' &= d(A_d - d - a^d_{\{h\}}h - a^d_{\{r\}}r - \\ &\quad a^d_{\{h,r\}}hr - a^d_{\{r,r\}}r^2 - a^d_{\{h,h\}}h^2). \end{aligned} \tag{5.5}$$

Τώρα θα παράγουμε ένα σύστημα τριών διαφορικών εξισώσεων που μοντελοποιεί τρία ανταγωνιζόμενα είδη. Οι εξισώσεις είναι οι ακόλουθες όπου για $p_1, p_2 \in \{H, D\}$ ορίζουμε ως $|\mathcal{F}[p_1, p_2]|$ το άθροισμα των κερδών για τους παίκτες p_1 και p_2:

$$\begin{aligned} r' &= r(A_r - r) - |\mathcal{F}[D,H]|rdh - |\mathcal{F}[H,H]|rhh - \\ &\quad |\mathcal{F}[D,D]|rdd - C^H rh - C^D rd \end{aligned} \tag{5.6}$$

$$\begin{aligned} h' &= h(A_h - h) - C^H_{R^-}h(C^H_{R^-,0} - r) - C_{H,H}rhh - \\ &\quad h \cdot C^H_{H\to D} - d \cdot C^H_{D\to H} \end{aligned} \tag{5.7}$$

$$\begin{aligned} d' &= d(A_d - d) - C^D_{R^-}d(C^D_{R^-,0} - r) - \\ &\quad d \cdot C^D_{D\to H} - h \cdot C^D_{H\to D}. \end{aligned} \tag{5.8}$$

Τώρα θα ερμηνεύσουμε αυτές τις εξισώσεις. Η εξίσωση (5.6) μοντελοποιεί την εξέλιξη του πληθυσμού της λείας. Ο πρώτος όρος απεικονίζει την μεταβολή του πληθυσμού, υποθέτωντας ότι το μέγιστο πλήθος της λείας μπορεί να φτάσει μέχρι το A_r (χωρητικότητα του συστήματος). Οι επόμενοι τρεις όροι μοντελοποιούν την επίδραση που έχει πάνω στον πληθυσμό της λείας οι αλληλεπιδράσεις των θηρευτών. Αυτή η επίδραση έτσι όπως έχουμε ορίσει το παίγνιο μας, απεικονίζεται στο παίγνιο μεταξύ γερακιών και περιστεριών. Οι τελευταίοι δύο όροι μοντελοποιούν την επίδραση των θηρευτών πάνω στον πληθυσμό της λείας. Η εξίσωση (5.7), μοντελοποιεί την εξέλιξη του πληθυσμού ενός από των δύο θηρευτών, τον πληθυσμό των γερακιών. Ο πρώτος όρος μοντελοποιεί την εγγενή εξέλιξη, χωρίς αλληλεπιδράσεις με άλλα είδη. Ο δεύτερος όρος $C^H_{R^-}h(C^H_{R^-,0} - r)$, μοντελοποιεί την "πείνα" των γερακιών, όταν υπάρχουν χαμηλά αποθέματα πόρων. Ο συντελεστής $C^H_{R^-}$ αναπαριστά την σημαντικότητα του παράγοντα της "πείνας", όσον $C^H_{R^-,0} \geq 0$ σημαίνει ότι υπάρχει έλλειμα πόρων και είναι μία επικίνδυνη κατάσταση για τον πληθυσμό της λείας, μέχρι ο όρος $C^H_{R^-}h(C^H_{R^-,0} - r)$ να γίνει θετικός (στην εξίσωσή μας εμφανίζεται με αρνητικό πρόσημο). Ο τρίτος όρος αναπαριστά το κόστος που έχει πάνω στον πληθυσμό των γερακιών, οι διαμάχες μεταξύ τους για την απόκτηση της λείας. Ο τέταρτος όρος, όπου $C^H_{H\to D} > 0$, αναπαριστά την ελλάτωση του πληθυσμού των γερακιών που απορρέει από την συνύπαρξη με το άλλο είδος κυνηγού, τα περιστέρια. Τέλος στην εξίσωση (5.8) μοντελοποιούμε την εξέλιξη του πληθυσμού του άλλου θηράματος για τα γεράκια, των περιστεριών. Η ερμηνεία αυτής της εξίσωσης είναι παρόμοια με αυτήν για τον πληθυσμό των γερακιών.

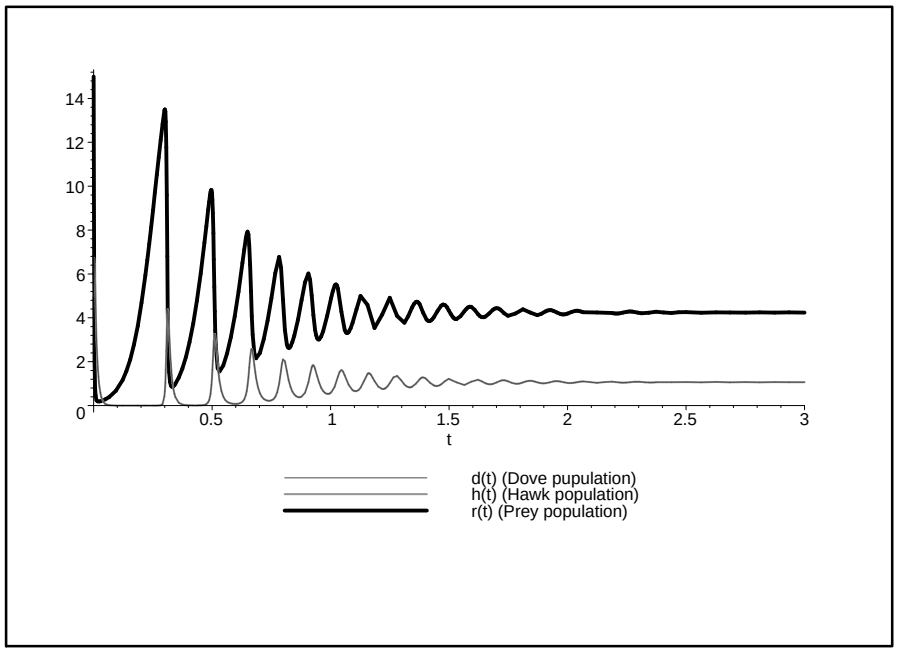
5.6 Διάφορα σενάρια εξέλιξης και αριθμητικά αποτελέσματα

Στα διαγράμματα 5.3, 5.4 και 5.5 βλέπουμε αποτελέσματα από την αριθμητική επίλυση των διαφορικών εξισώσεων (5.6), (5.7) και (5.8), με τιμές παραμέτρων που φαίνονται στους πίνακες 5.3, 5.4, και 5.5 αντίστοιχα.

A_r	20
$\|\mathcal{F}[D,H]\|$	3
$\|\mathcal{F}[H,H]\|$	3
$\|\mathcal{F}[D,D]\|$	3
C^H	1
C^D	1
A_h	20
$C^H_{R^-}$	25
$C^H_{R^-,0}$	5
$C_{H,H}$	0
$C^H_{H\to D}$	0
$C^H_{D\to H}$	0
A_d	20
$C^D_{R^-}$	25
$C^D_{R^-,0}$	5
$C^D_{D\to H}$	0
$C^D_{H\to D}$	0

Πίνακας 5.3: Συμπεριφορά με ταλαντώσεις που οδηγεί σε συνύπαρξη των ειδών

Γενικά, για μεγάλο εύρος τιμών των παραμέτρων του μοντέλου, παρατηρούμε μέτριες ταλαντώσεις μέχρι τα είδη να συνηπάρχουν ομαλά. Αλλάζοντας ανεξάρτητα κάποια υποσύνολα των παραμέτρων, πετυχαίνουμε διαφορετικές συμπεριφορές που εξηγούνται και φυσικά. Για παράδειγμα η διαφορά μεταξύ των διαγραμμάτων 5.3 και 5.4 είναι στην τιμή της παραμέτρου $C_{H,H}$, το κόστος στα γεράκια από εσωτερικές διαμάχες. Στο διάγραμμα 5.3 έχει τιμή 0 ενώ στο διάγραμμα 5.4 έχει υψηλή τιμή. Σε αυτές τις δύο περιπτώσεις που αναφέραμε, οι πληθυσμοί των γερακιών και των περιστεριών εξελίσσονται παρόμοια. Στις επόμενες ο πληθυσμός των γερακιών τείνει στην εξαφάνιση εξ' αιτίας των εσωτερικών διαμαχών. Στο δίαγραμμα 5.5 όπου έχουμε αποτελέσματα με τιμές από τον πίνακα 5.5, η διαφορά με τα προηγούμενα αποτελέσματα είναι στις χαμηλές τιμές του πίνακα κερδών. Παρατηρούμε ότι η λέια καταναλώνεται με μικρότερο ρυθμό. Στο διάγραμμα 5.6 έχουμε προσθέσει ένα στοιχείο περιοδικής μεταβολής στο κόστος από τις διαμάχες μεταξύ των γερακιών. Έχουμε μοντελοποιήσει αυτήν την παράμετρο με μία συνάρτηση συνημιτόνου. Στα σημεία όπου η συνάρτηση συνημιτόνου είναι αρνητική, τα γεράκια αποκτούν περισσότερο κέρδος όταν ανακατεύονται σε διαμάχες (αφού το κόστος είναι αρνητικό), και εξαπλώνονται αφού πρώτα έχει μειωθεί ο πληθυσμός τους, στο διάστημα όπου η συνάρτηση συνημιτόνου είναι θετική (αφού και το κόστος είναι θετικό). Μετά την πρώτη κορυφή υπάρχει μια άλλη υψηλή κορυφή που δείχνει την ισχυροποίηση του πληθυσμού των γερακιών που έχει ήδη ξεπεράσει την επίδραση στο διάστημα μετά την πρώτη κορυφή. Τέλος, στο διάγραμμα 5.7 έχουμε το εξής βασικό χαρακτηριστικό. Οι αρχικές συνθήκες για το κόστος των διαμαχών των γερακιών, εξαρτάται από τα αποθέματα της λείας. Παρατηρούμε ότι όσον η λεία παραμένει κοντά στην μέγιστη τιμή, το

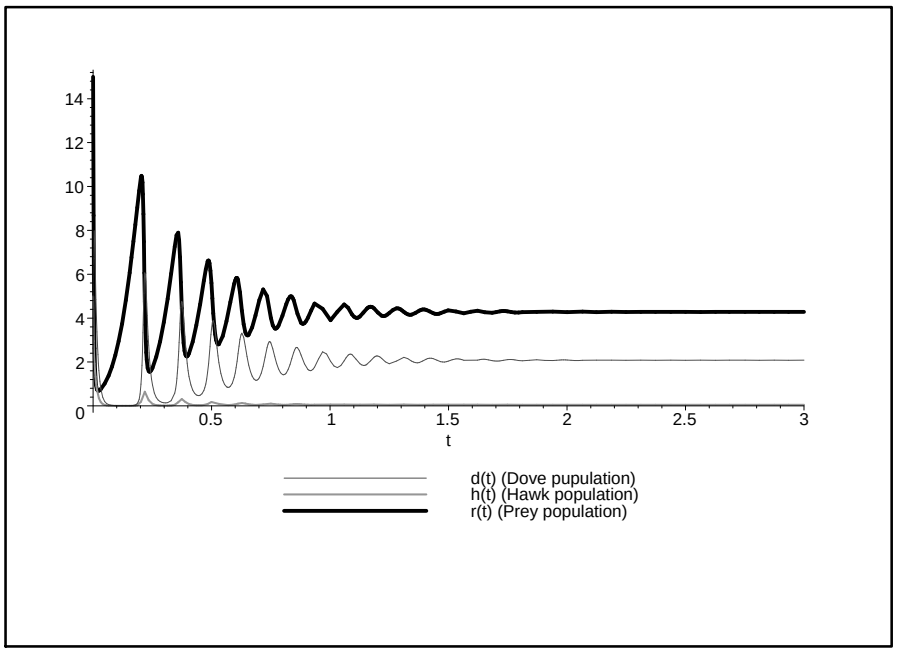


Σχῆμα 5.3: Εξέλιξη των πληθυσμών σύμφωνα με τον πίνακα 5.3

κόστος από τις διαμάχες είναι περίπου στο 0, και η συμπεριφορά τους είναι παρόμοια με αυτήν των περιστεριών. Όταν οι πόροι απέχουν από την μέγιστη τιμή, το κόστος από τις διαμάχες μεγαλώνει, και αποφέρει μείωση στον πληθυσμό των γερακιών σε σύγκριση με τα περιστέρια. Αυτό φαίνεται ξεκάθαρα στο διάγραμμα στα διαστήματα όπου ο πληθυσμός της λείας φτάνει τοπικά στο 0 και το κόστος από τις διαμάχες είναι και αυτό στο 0, τότε η συμπεριφορά των γερακιών είναι παρόμοια με των περιστεριών. Αυτό δείχνει την τάση των γερακιών να αποφεύγουν τις διαμάχες όταν η λεία διακινδυνεύει. Στα διαστήματα όπου η λεία φτάνει σε τοπικά μέγιστα, τα γεράκια έχουν κόστος από τις διαμάχες και αυτό ερμηνεύεται ως εξής: όταν υπάρχει αφθονία τροφής τα γεράκια στερούνται της προσεκτικότητας τους.

A_r	20
$\|\mathcal{F}[D,H]\|$	3
$\|\mathcal{F}[H,H]\|$	3
$\|\mathcal{F}[D,D]\|$	3
C^H	1
C^D	1
A_h	20
$C^H_{R^-}$	25
$C^H_{R^-,0}$	5
$C_{H,H}$	10
$C^H_{H\to D}$	0
$C^H_{D\to H}$	0
A_d	20
$C^D_{R^-}$	25
$C^D_{R^-,0}$	5
$C^D_{D\to H}$	0
$C^D_{H\to D}$	0

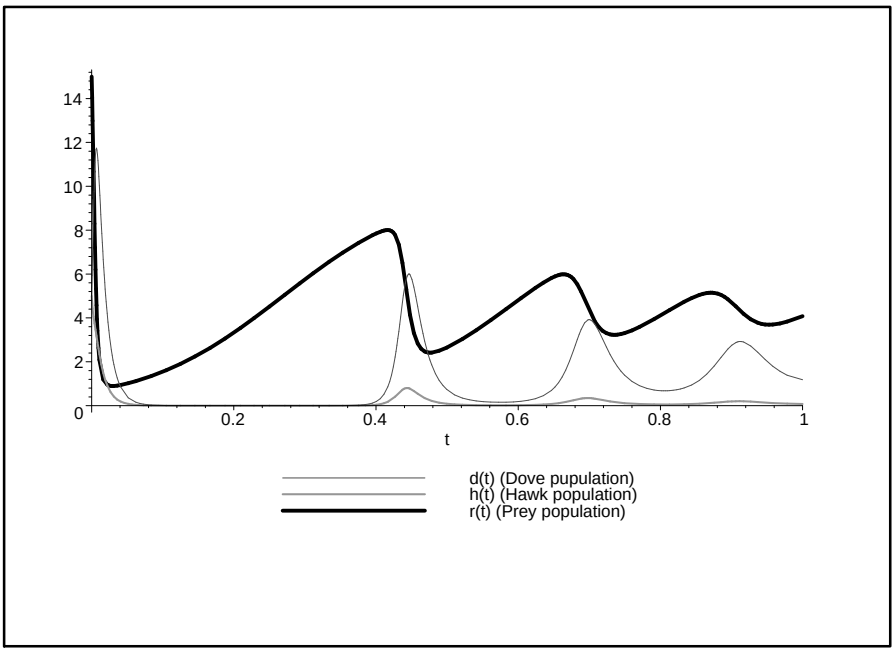
Πίνακας 5.4: Συμπεριφορά με ταλαντώσεις που οδηγεί σε συνύπαρξη των ειδών



Σχῆμα 5.4: Εξέλιξη των πληθυσμών σύμφωνα με τον πίνακα 5.4

A_r	10
$\|\mathcal{F}[D,H]\|$	1
$\|\mathcal{F}[H,H]\|$	1
$\|\mathcal{F}[D,D]\|$	1
C^H	1
C^D	1
A_h	20
$C^H_{R^-}$	25
$C^H_{R^-,0}$	5
$C_{H,H}$	10
$C^H_{H\to D}$	0
$C^H_{D\to H}$	0
A_d	20
$C^D_{R^-}$	25
$C^D_{R^-,0}$	5
$C^D_{D\to H}$	0
$C^D_{H\to D}$	0

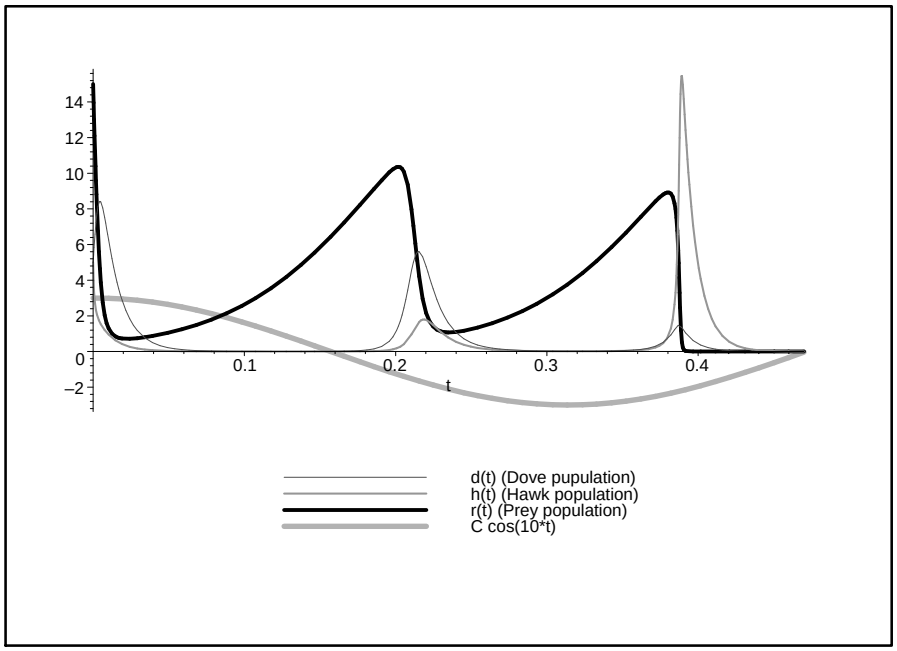
Πίνακας 5.5: Έλλειψη πόρων για μεγάλη κατανάλωση



Σχῆμα 5.5: Εξέλιξη των πληθυσμών σύμφωνα με τον πίνακα 5.5

A_r	20
$\|\mathcal{F}[D,H]\|$	3
$\|\mathcal{F}[H,H]\|$	3
$\|\mathcal{F}[D,D]\|$	3
C^H	1
C^D	1
A_h	20
$C^H_{R^-}$	25
$C^H_{R^-,0}$	5
$C_{H,H}$	3cos(10τ)
$C^H_{H\to D}$	0
$C^H_{D\to H}$	0
A_d	20
$C^D_{R^-}$	25
$C^D_{R^-,0}$	5
$C^D_{D\to H}$	0
$C^D_{H\to D}$	0

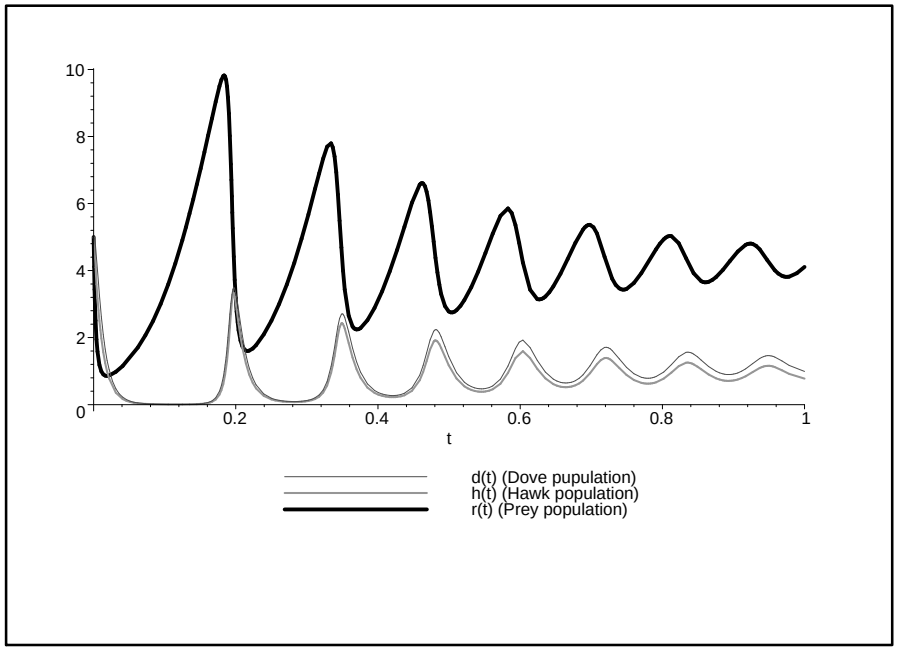
Πίνακας 5.6: Το κόστος από τις διαμάχες μεταξύ των γερακιών είναι περιοδική συνάρτηση του χρόνου



Σχῆμα 5.6: Εξέλιξη των τριών πληθυσμών όταν οι διαμάχες μεταξύ των γερακιών είναι περιοδική συνάρτηση του χρόνου

A_r	20
$\|\mathcal{F}[D,H]\|$	3
$\|\mathcal{F}[H,H]\|$	3
$\|\mathcal{F}[D,D]\|$	3
C^H	1
C^D	1
A_h	20
$C^H_{R^-}$	25
$C^H_{R^-,0}$	5
$C_{H,H}$	$3\frac{h(t)}{5}\frac{h(t)}{5}(1-\frac{r(t)}{5})$
$C^H_{H\to D}$	0
$C^H_{D\to H}$	0
A_d	20
$C^D_{R^-}$	25
$C^D_{R^-,0}$	5
$C^D_{D\to H}$	0
$C^D_{H\to D}$	0

Πίνακας 5.7: Το κόστος από τις διαμάχες είναι συνάρτηση των αποθεμάτων της λείας



Σχῆμα 5.7: Παρόμοια συμπεριφορά των γερακιών και περιστεριών όταν το κέρδος τους είναι συνάρτηση των αποθεμάτων της λείας

Κεφάλαιο 6

Επεκτάσεις και μελλοντική έρευνα

Σε αυτό το κεφάλαιο θα παρουσιάσουμε κάποιες διαφορές των μοντέλων μας με τα ήδη υπάρχοντα μοντέλα διαδόσεων απειλών και θα προτείνουμε κάποιες επεκτάσεις που θα μπορούσαν να γίνουν σε μελλοντική έρευνα. Οι κύριες διαφορές των μοντέλων μας, με το μεντέλο του Zou [34] για την διάδοση του worm code red, είναι ότι στην εργασία [34] δεν θεώρησαν ότι τα λογισμικά προστασίας μπορούν να κινούνται και αυτά στο δίκτυο και να διαδίδονται απο υπολογιστή σε υπολογιστή που με αυτό τον τρόπο μοντελοποιούμε καταστάσεις όπου έχουμε ενημερώσεις λογισμικών ή εγκαταστάσεις νέων. Ακόμα δεν θεωρούν ότι υπάρχει φθορά στο πλήθος των καθαρών υπολογιστών, λόγω παλαιότητας των λογισμικών προστασίας, ή μη ενημέρωσης των ήδη υπαρχόντων. Γι' αυτο τον λόγο θεωρούν ότι η αύξηση των καθαρών όπως και η μείωση των μολυσμένων, δεν επέρχεται σαν αλληλεπίδραση και των δύο πληθυσμών. Τέλος, στις τιμές των παραμέτρων για τον ρυθμό αύξησης του πληθυσμού των ιών, χρησιμοποίησαν τιμές ώστε οι αριθμητικές τους λύσεις να ταιριάζουν με τις πραγματικές παρατηρήσεις. Στη συγκεκριμένη διατριβή οι τιμές των παραμέτρων για το ρυθμό διάδοσης είναι πραγματικά στοιχεία απ' τις διαδόσεις των γνωστών ιών, slammer και code red και τα αριθμητικά μας αποτελέσματα συμπίπτουν με τα πραγματικά. Ακόμα στην δική μας διατριβή, τα μοντέλα μας αναλύθηκαν και θεωρητικά με μαθηματικά εργαλεία των δυναμικών συστημάτων για την ύπαρξη ευσταθών καταστάσεων ισορροπίας. Στο μοντέλο μας θεωρούμε ότι η διάδοση της απειλής είναι συνεχής. Αυτό όμως σε κάποιες περιπτώσεις ιών, δεν είναι ρεαλιστικό αφού κάποιοι ιοί εμφανίζονται μόνο για λίγες ώρες και στην συνέχεια σταματάει η διάδοση και ξαναρχίζει μετά από λίγες μέρες ή ακόμα και μήνες όπως στην περίπτωση του code red worm όπου μετά την πρώτη του εμφάνιση για πέντε ώρες, ξαναεμφανίστηκε μετά από έντεκα ημέρες. Μια προέκταση λοιπόν του μοντέλου μας σε μετέπειτα έρευνα είναι να συμπεριλαμβάνει όρους που θα μοντελοποιούν την παύση της διάδοσης του ιού, και την συνέχεια μετά από κάποιο χρονικό διάστημα. Προφανώς η χρονική στιγμή της επανεμφάνισης το ιού επιλέγεται τυχαία από τον κατασκευαστή του κώδικα της απειλής. Ακόμα στα μοντέλα μας που αποτελούνται από διαφορικές εξισώσεις έχουμε θεωρήσει ότι όλοι οι πληθυσμοί βρίσκονται σε ενιαίο χώρο, όπου στην πράξη αυτό σημαίνει ότι όλοι οι υπολογιστές συνδέονται ανά δύο μεταξύ τους. Το γράφημα ενός τέτοιου δικτύου θα ήταν πλήρες. Αυτό όμως δεν ισχύει για το διαδίκτυο όπου έρευνες έ-

χουν καταλήξει ότι η τοπολογία του διαδικτύου είναι scale free, όπου η πιθανότητα σύνδεσης δύο τυχαίων υπολογιστών ακολουθεί μια power law κατανομή. Μια ενδιαφέρουσα μελλοντική εργασία είναι να επαναδιατυπωθούν οι διαφορικές εξισώσεις που περιγράφουν τα μοντέλα μας έχοντας συμπεριλάβει όρους που θα χαρακτηρίζουν την τοπολογία του διαδικτύου. Στο τρίτο κεφάλαιο της διατριβής, για πρώτη φορά παρουσιάστηκε κατανομή για δύο ειδών πληθυσμών στους κόμβους ενός δικτύου. Για τον υπολογισμό αυτής της κατανομής χρησιμοποιούμε τις ίδιες παραδοχές όπως και ο Jackson για την μοντελοποίηση ενός κόμβου με μία $M/M/1$ ουρά. Αυτό όμως δεν ανταποκρίνεται στην πραγματικότητα, αφού οι υπολογιστές μπορούν να δεχτούν σαφώς μεγάλο πλήθος πακέτων εργασιών, αλλά πεπερασμένο. Ακόμα το πλήθος αφίξεων και οι χρόνοι εξυπηρέτησης δεν ακολουθούν ακριβώς Poisson και εκθετική κατανομή αντίστοιχα. Μια πιθανή επέκταση του μοντέλου μας είναι να υπολογίσουμε την κατανομή για πιο γενικές μορφές για τις κατανομές των αφίξεων και των εξυπηρετήσεων. Μια άλλη επέκταση που μπορεί να γίνει είναι, αν θεωρήσουμε μια πιο πολύπλοκη αλληλεπίδραση μεταξύ του ιού και του λογισμικού προστασίας αντί για την αντίδραση που θεωρήσαμε, για την μελέτη μιας πιο αποδοτικής εξάλειψης του ιού. Για παράδειγμα, μπορούμε να θεωρήσουμε ότι ένα λογισμικό προστασίας, εξουδετερώνει με κάποια πιθανότητα τον ιό, όπως υποθέσαμε στο δεύτερο κεφάλαιο και στην συνέχεια παράγει με μια άλλη πιθανότητα ένα αντίγραφο του εαυτού του, ώστε να μην εξολοθρεύεται και αυτό όπως έχουμε υποθέσει, και με αυτό τον τρόπο μπορεί να επιτευχθεί ταχύτερος καθαρισμός του δικτύου. Στο τέταρτο κεφάλαιο της διατριβής που επεκτείναμε το μοντέλο που προτάθηκε από τον Κερομύτη στην εργασία [13] που περιγράφεται από μία διαφορική εξίσωση για την διάδοση μιας DNS απειλής σε ένα σύστημα δύο διαφορικών εξισώσεων, για πρώτη φορά σε μοντελοποίηση διάδοσης απειλών με βάση επιδημιολογικά μοντέλα, χρησιμοποιήθηκαν παράμετροι για να μοντελοποιήσουν την ανθρώπινη συμπεριφορά ενάντια στις απειλές. Τα αριθμητικά μας αποτελέσματα βρίσκονται και εδώ σε απόλυτη αρμονία με πραγματικές παρατηρήσεις των γνωστών worms code red, slammer απ΄ τον οργανισμό caida.org. Επίσης σε αυτό το κεφάλαιο μελετάμε έναν έξυπνο ιό, που δεν κάνει τυχαία σάρωση για εύρευση ευπαθών υπολογιστών. αλλά σαρώνει το δίκτυο με συγκεκριμένο τρόπο. Μία πιθανή επέκταση της μελέτης που παρουσιάστηκε σε εκείνο το κεφάλαιο, είναι να χαρίσουμε το δίκτυο σε ένα γενικό και σε τοπικά δίκτυα. Με αυτό τον τρόπο θα μοντελοποιούμε τον ρυθμό διάδοσης των ιών, που γνωρίζουμε ότι έχουν ταχύτερο ρυθμό διάδοσης σε τοπικά δίκτυα, παρά στο διαδίκτυο, λόγω των μικρότερων καθυστερήσεων απ΄ τις καλύτερες συνδέσεις ενός τοπικού δικτύου. Ακόμα στα τοπικά δίκτυα, οι απαντήσεις στα ερωτήματα για ηλεκτρονικές διευθύνσεις παίρνονται άμεσα από τον τοπικό εξυπηρετητή και δεν υπάρχει καθόλου η καθυστέρηση που μοντελοποιούμε με τον όρο $d_{internet}$. Ξαναγράφοντας τις εξισώσεις του μοντέλου μας αλλά αυτήν την φορά με δύο εξισώσεις για κάθε μία του μοντέλου, μία για το διαδίκτυο και μία για τα τοπίκα δίκτυα, θα μπορέσουμε να προσομειώσουμε καλύτερα την συμπεριφορά των ιών. Στο πέμπτο κεφάλαιο που μοντελοποιούμε αλληλεπιδράσεις πολλαπλών κυνηγών καθώς ανταγωνίζονται ή συνεργάζονται για τι ίδιο θήραμα, μπορούμε να θεωρήσουμε πιο πολύπλοκες αλληλεπιδράσεις, όπου ένα είδος μπορεί να είναι ταυτόχρονα και θηρευτής αλλά και θήραμα. Για παράδειγμα, μπορούμε να υποθέσουμε ότι τα γεράκια εκτός ότι ανταγωνίζονται τα περιστέρια για την αποκόμιση της λείας, βλέπουν και αυτά ώς θήραμα, κάτι που συμβαίνει και στην πραγματικότητα. Ακόμα μπορούμε να κάνουμε μια αναγωγή του μοντέλου μας από τον βιολογικό κόσμο, στον κόσμο των υπολογιστών, αν θεωρήσουμε ότι ένας κυνηγός, είναι κάποιο είδος

κακόβουλου λογισμικού, ενώ ένας παθητικός κυνηγός όπως είναι τα περιστέρια είναι κάποιο είδος καλόβουλου λογισμικού που κινείται με σκοπό την προστασία του δικτύου. Βεβαίως ως λεία, μπορούμε να θεωρήσουμε το πλήθος των ευπαθών υπολογιστών, που δεν έχουν μέτρα αντιμετόπισης των κυνηγών.

Περίληψη

Ένας ιός υπολογιστών είναι ένα κακόβουλο κομμάτι κώδικα που διαδίδεται ταχύτατα στους υπολογιστές ενός δικτύου. Παρόλο που ένας μεγάλος αριθμός ερευνητών έχει εστιάσει την προσπάθεια του στην επινόηση νέων τεχνικών για την ανίχνευση και την εξάλειψη των ιών, δεν δίνει και τόση μεγάλη σημασία στην ανάπτυξη θεωρητικών μοντέλων που είναι ικανά να προβλέψουν το μέγεθος της εξάπλωσης των απειλών σε ευπαθή δίκτυα. Στην παρούσα διατριβή προτείνουμε και αναλύουμε μαθηματικά μοντέλα για την ταυτόχρονη εξέλιξη των πληθυσμών των ιών και των λογισμικών προστασίας σε ένα μολυσμένο δίκτυο υπολογιστών. Αρχικά μοντελοποιούμε ένα σύστημα δύο πλυθησμών που περιλαμβάμνει τους ιούς και τα λογισμικά προστασίας και αναλύοντας το θεωρητικά δείχνουμε ότι οι προβλέψεις μας συμπίπτουν με πραγματικές παρατηρήσεις διαδόσεων ιών. Επεκτείνουμε αυτό το μοντέλο με την προσθήκη των παγίδων, και καταλήγουμε σε ένα σύστημα τριών μη- γραμμικών διαφορικών εξισώσεων που περιγράφουν τις αντιδράσεις τους. Στην συνέχεια μελετάμε το πρόβλημα διάδοσης και αναχαίτησης ιών σε δίκτυα από μία άλλη οπτική γωνία η οποία αποφεύγει την χρήση των μη γραμμικών διαφορικών εξισώσεων. Αυτό το μοντέλο λαμβάνειν υπ' όψιν τα χαρακτηριστικά των εξυπηρετητών και της κίνησης του δικτύου υπολογιστών. Προτείνουμε ένα είδος αντίδρασης μεταξύ των ιών και των λογισμικών προστασίας που οδηγεί σε μία κατανομή μορφής γινομένου για τα πλήθη και των δύο ειδών στους κόμβους του δικτύου, όπως την κατανομή για τα πακέτα εργασίων σε δίκτυα ουρών Jackson. Έπειτα μελετάμε DNS ιούς, οι οποίοι χρησιμοποιούν έναν γεννήτορα τυχαίων αλφαριθμητικών για την εύρεση πιθανών πραγματικών διευθύνσεων και στην συνέχεια θέτουν ερωτήματα στους εξυπηρετητές με σκοπό την αποκόμιση της πραγματικής ηλεκτρονικής διεύθυνσης. Παρουσιάζουμε μοντέλα για την ταυτόχρονη εξέλιξη των ιών και των λογισμικών προστασίας, που κινούνται στο δίκτυο προσπαθώντας να σταματήσουν την διάδοση των ιών. Στην συνέχεια επεκτείνουμε τα μοντέλα με την προσθήκη των dummy honeypot servers, που προσπαθούν να προσελκύσουν τους ιούς να θέσουν ερωτήματα προκαλώντας καθυστερήσεις αφού δεν επιστρέφουν απαντήσεις. Στο τέλος προτείνουμε ένα μοντέλο κυνηγού θηράματος για την μελέτη της ταυτόχρονης εξέλιξης του πληθυσμού της λείας που συνυπάρχει με κάποιο πλήθος κυνηγών. Αυτό το μοντέλο επιτρέπει τον ορισμό πολλών διαφορετικών συμπεριφορών των πληθυσμών των ειδών, που απορρέουν από τις διαφορετικές αλληλεπιδράσεις μεταξύ των ειδών των κυνηγων, και της λέιας. Επιπλέον μοντελοποιούμε την συμπεριφορά δύο κυνηγών, χρησιμοποιώντας το απλό παίγνιο "γεράκια και περιστέρια", σύμφωνα με το οποίο το ένα είδος κυνηγού είναι επιθετικό (γεράκια) και το άλλο είναι υποχωρητικό (περιστέρια).

Abstract

A *computer virus* is a malicious, self-propagating piece of code that is able to spread fast in computer networks. Although a growing number of researchers focus their efforts on devising new techniques for detecting and eliminating worms, there seems to be less intense activity towards the development and evaluation of theoretical models able to account of how worms exploit vulnerabilities of computer networks and propagate, accordingly. In our work, we propose and analyze mathematical models for the co-evolution of the populations of virus and antivirus software agents across an infected computer network. We first model a two populations system comprised of worm and antivirus agents,

and analyze it theoretically showing that its prediction are in harmony with observations from propagation of real worms. We extend this model, adding the traps population, using a system of three non-linear differential equations that describe their interactions. Then we view the attack propagation/elimination problem in networks from another perspective that avoids the use of non-linear evolution dynamics. This model takes into account the traffic and server characteristics of the network computers. We propose a kind of interaction between virus and antivirus agents that results in a product form steady state distribution of the agent numbers for each network node, much like the product form solution for the distribution of network tasks for Jackson open networks of queues. Afterwards we study DNS worms, these worms generate random strings, as possible network domain names, and then query Domain Name Servers in order to discover the IP address. We present models for capturing the dynamics of the co-evolution of worm agents in the presence of anti-worm agents that move in the network in order to locate and stop worm propagation. We further enhance the model with "honeypot" domain name servers that attempt to lure worm agents to issue queries, introducing only a delay and providing no answer. Finally we propose a general predator-prey model for studying the joint evolution of a prey population that coexists with a number of predator populations. The model allows the definition of different behaviors between individuals from these populations as well as different interaction patterns between each of these populations and the prey population. Moreover, we model the behaviour between individuals of the two predator populations using as simple game, called *Hawks and Doves* game, according to which one predator population is aggressive (Hawks) and one predator population is submissive (Doves).

Βιβλιογραφία

[1] CERT advisory CA-2001-26 Nimda Worm.

[2] M.S. Bartlett, *Stochastic Population Models - in ecology and epidemiology*, John Wiley & Sons, 1960.

[3] D. Bundy, *Basic Queuing Theory*, Edward Arnold (Publishers) Ltd., 1986.

[4] P.C Buzing, A.E. Eiben, and M.C. Schut, Emerging communication and cooperation in evolving agent societies, *Journal of Artificial Societies and Social Simulation*, vol. 8, no. 1, 2005.

[5] V. Capasso, "Mathematical Structures of Epidemic Systems". Berlin, Heidelberg: Springer-Verlag, 1993.

[6] James C. Frauenthal. "Mathematical Modeling in Epidemiology". Springer-Verlag, 1980

[7] H. I. Freedman and P. Waltman, *Persistence in a model of three interacting predator-prey populations*, Math. Biosc. 68, pp. 213-231, 1984.

[8] D. Fudenberg and J. Tirole, *Game Theory*, The MIT Press, 1991.

[9] P. Kammas, C. Manolopoulos and Y.C. Stamatiou, 'A multiple-predator prey evolution model with game-theoretic description of predator interactions', submited

[10] P. Kammas, Y.C. Stamatiou and T. Komninos, 'Modeling the coevololution DNS worms and anti-worms in IPv6', Journal of Information Assurance and Security, Volume 1 2010.

[11] P.Kammas, Y.C. Stamatiou and T. Komninos, 'Queuing theory based models for studying intrusion evolution and elimination in computer networks', Journal of Information Assurance and Security, Volume 4, Issue 3, pp 200–208 June 2009

[12] P.Kammas, Y.C. Stamatiou and T. Komninos 'A three population computer worm propagation model and its theoretical analysis', submited

[13] A. Kamra, H. Feng, V. Misra, and A. Keromytis, "The effect of DNS delays on eorm propagation in an IPv6 internet"

[14] S. Kauffman, *At home in the universe: The search for the Laws of Self-Organization and Complexity*, Oxford University Press, 1995.

[15] J.O. Kephart and S.R. White, "Measuring and Modeling Computer Virus Prevalence," in *Proc. 1993 IEEE Computer Society Symposium on Research in Security and Privacy*, Oakland, California, 1993.

[16] L. Kleinrock, On the Modeling and Analysis of Computer Networks, *in Proc. of the IEEE*, Vol. 81, No. 8, August 1993.

[17] S. Kontogiannis and P. Spirakis, Evolutionary Games: An Algorithmic View, in Proc. *Self-star Properties in Complex Information Systems*, pp. 97–111, 2004.

[18] A.J. Lotka, *The extinction of families*, J. Wash. Acad. Sci. **21**, 1931.

[19] M. Mannan and P. Oorschot,"On Instant Messaging Worms, Analysis and Countermeasures," in *Proc. of the 2005 ACM workshop on Rapid malcode (WORM'05)*.

[20] C. Manolopoulos, Emerging Networks in Evolving Agent Societies, *Anthology of the B&ESI Conference*, Vol. 2, 2006.

[21] M. May and W.J. Leonard, Nonlinear Aspects of Competition Between Three Species, *SIAM Journal on Applied Mathematics*, Vol. 29, No. 2., pp. 243–253, 1975.

[22] J. Maynard Smith. Evolution and the Theory of Games. Cambridge University Press, Cambridge, 1982.

[23] D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford, and N. Weaver, "Inside the slammer worm," *IEEE security and privacy*, **1(4)**,33–39, July 2003.

[24] D. Moore, C. Shannon, "The Spread of the Witty worm" *IEEE security and privacy*,vol 2, no 4, pp 46-50, July/August 2004.

[25] J.D. Murray, Mathematical Biology I. An Introduction.

[26] S. Nikoletseas, C. Raptopoulos, and P. Spirakis, The survival of the weakest in networks, in Proc. *4th International Workshop on Online Algorithms (WAOA 2006)*, 316–329, 2006.

[27] G. Seifert, *A Lotka-Volterra predator-prey system involving two predators*, Methods and Applications of Analysis **2** (2), pp. 248–255, 1995.

[28] S. Staniford, V. Paxson, and N.Weaver, "How to 0wn the Internet in Your Spare Time", Proceedings of the 11th USENIX Security Symposium, 2002.

[29] S.H. Strogatz, Nonlinear dynamics and chaos.

[30] T. Timmreck, "An Introduction to Epidemiology". Sadbury, MA: Jones and Bartlett, 1998.

[31] V. Volterra, Variazioni e fluttuazioni del numero d' individui in specie animali conviventi, *Mem. Acad. Lineci Roma, 2, 31*, 1926.

[32] C. Wang, J. Knight, and M. Elder, "On computer viral infection and the effect of immunization," in *Proc. of the 16th annual computer security applications conference (ACSAC '00)* , New Orleans, LA, Dec. 2000.

[33] S. Wiggins, Introduction to Applied Nonlinear Dynamical Systems and Chaos.

[34] C.C Zou, W. Gong, and D. Towsley,"Code-red worm propagation modeling and analysis.," in *Proc. of the 9th ACM conference on Computer and Communications Security*, ACM Press, pp. 138–147, 2002.