



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ & ΣΥΣΤΗΜΑΤΩΝ
ΠΛΗΡΟΦΟΡΙΚΗΣ

Μοντέλα Διάδοσης Επιδημιών Σε Δίκτυα Υπολογιστών

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αριστομένης Β. Μανωλάς

Επιβλέπων : Ευστάθιος Δ. Συκάς
Καθηγητής Ε.Μ.Π.

Αθήνα, Φεβρουάριος 2011



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ

ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ & ΣΥΣΤΗΜΑΤΩΝ

ΠΛΗΡΟΦΟΡΙΚΗΣ

Μοντέλα Διάδοσης Επιδημιών Σε Δίκτυα Υπολογιστών

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αριστομένης Β. Μανωλάς

Επιβλέπων : Ευστάθιος Δ. Συκάς
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 10^η Φεβρουαρίου 2011.

.....
Ε. Συκάς
Καθηγητής Ε.Μ.Π

.....
Μ. Θεολόγου
Καθηγητής Ε.Μ.Π.

.....
Γ. Στασινόπουλος
Καθηγητής Ε.Μ.Π

Αθήνα, Φεβρουάριος 2011

.....

Αριστομένης Β. Μανωλάς

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Αριστομένης Β. Μανωλάς 2011

Με επιφύλαξη παντός δικαιώματος. All rights reserved

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Η μοντελοποίηση επιδημιών είναι ένα πεδίο έρευνας ήδη ανεπτυγμένο και μπορεί να έχει εφαρμογή και στα δίκτυα υπολογιστών. Σήμερα η ανταλλαγή αρχείων μεταξύ των χρηστών καθώς και η συμμετοχή σε κοινωνικά δίκτυα είναι αρκετά διαδεδομένη. Αρχικά θα περιγράψουμε τα πιο γνωστά μοντέλα δίνοντας την απαραίτητη έμφαση στο τρόπο παραγωγής τους με σκοπό να μπορέσουμε να τα χρησιμοποιήσουμε στα δίκτυα.

Έπειτα από την περιγραφή ενός μοντέλου είναι δυνατή η ανάλυση συνεπειών του καθώς και η προσπάθεια διατύπωσης αξιόλογων προβλέψεων. Στη συγκεκριμένη εργασία θα ασχοληθούμε με το μοντέλο SIR (Susceptible, Infected, Recovered) σε μια στοχαστική εκδοχή του μέσω μιας διαδικασίας Markov.

Κατά την προσπάθειά μας να ανάγουμε την θεωρία στα δίκτυα, θα διαφοροποιήσουμε το αρχικό μοντέλο εισάγοντας την δράση των κακόβουλων κόμβων. Η φύση της ασθένειας καθορίζει συχνά το συγκεκριμένο μοντέλο, παρόλα αυτά ο τρόπος εμφάνισής της ή ο αντιός θεραπείας δεν θα μας απασχολήσουν. Ένα θέμα που θα πρέπει να μας απασχολήσει είναι με ποιόν τρόπο η τοπολογία του δικτύου θα επηρεάσει την περιγραφή μας για την πιθανότητα μετάβασης στο στοχαστικό μοντέλο επιδημίας. Ένα σημαντικό εργαλείο θα αποδειχθεί ότι είναι η κατανομή βαθμού.

Λέξεις Κλειδιά: μοντέλα επιδημίας, κοινωνικά δίκτυα, SIR, στοχαστικό, Markov, κατανομή βαθμού

Abstract

Modeling an epidemic is a field of research already developed and can also be applied to computer networks. Today the exchanges of files between users as well participation in social networks are quite widespread. We will first describe the most popular models giving the necessary emphasis on their method of production with a view to achieve their use to the computer networks.

After describing a model is possible to analyse the consequences and the attempt to formulate meaningful predictions. In this paper we will deal with SIR (Susceptible, Infected and Recovered) model in a stochastic version through a Markov process.

During our approach to carry the theory to the communication networks we will differentiate between the original models by introducing the action of malicious nodes. The nature of the disease often determines the specific model; however this nature or the form of treatment should not concern us. One issue that should concern us in our study of a disease is how the network topology will affect our description of the transition probabilities in our stochastic epidemic model. A useful tool that will prove to be is the degree distribution.

Keywords: epidemic models, social networks, SIR, stochastic, Markov, degree distribution

Πρόλογος

Θα ήθελα να ευχαριστήσω τον Ε. Δ. Συκά, καθηγητή της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Ε.Μ.Π., για την παρακολούθηση και τη βοήθειά του σε σημαντικά σημεία της εργασίας.

Αφιερώνεται στην οικογένειά μου

Περιεχόμενα

Εισαγωγή	10
1. Γνωστά αιτιοκρατικά μοντέλα επιδημιών	12
1.1 Ντετερμινιστικά Μοντέλα	12
1.1.1 SIR (Susceptible, Infected, Recovered) Μοντέλο.....	12
1.1.2 Το μοντέλο SIR με γεννήσεις-θανάτους	15
1.1.3 Το μοντέλο SIS με γεννήσεις - θανάτους.....	15
1.1.4 Το SIRS μοντέλο	16
1.1.5 Το μοντέλο SEIS	16
1.1.6 Το μοντέλο SEIR.....	17
1.1.7 Το μοντέλο MSIR.....	17
1.1.8 Το μοντέλο MSEIR.....	18
1.2 Συνθήκες έναρξης επιδημίας – Χρονική εξέλιξη του μοντέλου SIR.....	18
2. Επίδραση της τοπολογίας στην διάδοση επιδημιών	21
2.1 Δομή Δικτύου	22
2.2 Κεντρικότητα	24
2.3 Ιδιότητες δικτύων απουσία κλίμακας και επιδημίες	28
2.3.1 Κατώφλι επιδημίας στο μοντέλο SIS	29
2.3.2 Κατώφλι επιδημίας στο μοντέλο SIR.....	34
2.4 Διάδοση επιδημιών σε δίκτυα με τη χρήση γεννητριών συναρτήσεων.....	38
2.4.1 Γεννήτριες Συναρτήσεις.....	40
2.4.2 Κατανομή Μεγέθους Ξεσπάσματος	42
2.4.3 Μεγέθη Ξεσπάσματος και η Μετάβαση Επιδημίας.....	43
2.4.4 Βαθμός Μολυσμένων Ατόμων	45
2.4.5 Συσχετισμένες Πιθανότητες Διάδοσης.....	46
3. Στοχαστική εκδοχή μοντέλου SIR σε δίκτυα υπολογιστών	48
3.1 Εξισώσεις μοντέλου	49
3.2 Λύση εξισώσεων.....	55
3.3 Εξάρτηση πιθανότητας μετάβασης από την τοπολογία	61
4. Αποτελέσματα – Συμπεράσματα.....	66
4.1 Γενικά χαρακτηριστικά της μετάβασης της επιδημίας	66
4.2 Επίδραση των ρυθμών μόλυνσης και θεραπείας-Εκθετική αύξηση της πιθανότητας τελικής κατάστασης	71

4.3 Επίδραση της τοπολογίας στην Επιδημία.....	73
4.4 Παραλλαγές του μοντέλου.....	76
4.4.1 Μείωση πλήθους μολύνσεων.....	76
4.4.2 Καθυστέρηση θεραπείας.....	78
4.5 Επίλογος.....	81
Βιβλιογραφία:	83

ΕΝΟΤΗΤΑ 1

Εισαγωγή

Οι επιδημίες έχουν αρχίσει να απασχολούν τους ερευνητές τουλάχιστον από την αρχή του προηγούμενου αιώνα. Σε κάθε περίπτωση ασθένειας που μπορεί να διαδοθεί υπάρχει το ενδιαφέρον της χρονικής εξέλιξης και του μεγέθους της, ένα θέμα που διακρίνεται από την καθαρά ιατρική αντιμετώπιση της ίδιας της ασθένειας. Αρχικά το ενδιαφέρον εστιάστηκε σε ασθένειες που οδήγησαν σε επιδημία εντός ενός πληθυσμού ανθρώπων. Παρόμοια είναι και η περίπτωση των άλλων ζωντανών οργανισμών που μέσω της κίνησής τους μπορούν να μεταδώσουν μια νόσο. Από το 1995 και μετά, το internet και το World Wide Web (www) άρχισαν να είναι μια πραγματικότητα για απλούς και μέσους χρήστες σε όλο τον κόσμο, οπότε ήρθαμε σε επαφή με ένα τελείως διαφορετικό είδος επιδημίας, αυτό που λαμβάνει δράση σε υπολογιστές που μπορούν να συνδεθούν σε ένα δίκτυο επικοινωνιών. Τον πρώτο αυτό καιρό τα μηχανήματα ήταν ιδιαίτερα ευάλωτα καθώς δεν υπήρχε διαδεδομένη η έννοια της πρόληψης όπως σήμερα. Σε αναλογία με τη βιολογική περίπτωση σε κάθε περίπτωση ιού αναπτυσσόταν ο κατάλληλος αντιός και η σταδιακή θεραπεία. Στα επόμενα χρόνια πολλοί ιοί ομαδοποιήθηκαν, ώστε μια κοινή θεραπεία ή ένα λογισμικό ασφαλείας μπορούσε να δράσει προληπτικά. Αν και σήμερα οι επιδημίες δεν είναι τόσο συχνές όπως τότε, παραμένει πάντα ανοιχτό το θέμα ενός νέου άγνωστου ιού που μπορεί να οδηγήσει σε επιδημία. Η ήδη γνωστή θεωρητική μελέτη της βιολογίας σχετικά με τις επιδημίες έρχεται να καλύψει αρχικά αυτό το πεδίο και στο χώρο των δικτύων υπολογιστών, μια επιλογή αρκετά δικαιολογημένη μιας και τα δύο είδη πληθυσμών μοιράζονται κοινά χαρακτηριστικά στην διάδοση μιας επιδημίας.

Με μια πιο προσεκτική μελέτη και έπειτα από την εμπειρία των δικτύων υπολογιστών μπορούν να προκύψουν νέες ιδιότητες και χαρακτηριστικά, οπότε υπάρχει η δυνατότητα να δημιουργηθεί ένα πιο ρεαλιστικό μοντέλο διάδοσης επιδημίας. Οι χρήστες του ιστού εκτελούν ποικίλες εφαρμογές όπως μετάδοση αρχείων αλλά και ενεργή συμμετοχή σε κοινωνικά δίκτυα, ένα πεδίο αρκετά διαδεδομένο σήμερα. Όπως προκύπτει, ένα δίκτυο εφαρμογών έχει να δείξει μια αρκετά διαφορετική συμπεριφορά από ένα αντίστοιχο βιολογικό και η μεταφορά των ιδιαίτερων χαρακτηριστικών του πρώτου σε ένα μοντέλο επιδημίας φαίνεται πλέον αναπόφευκτη. Η τοπολογία του φυσικού δικτύου μπορεί να επηρεάσει τη μετάδοση μιας ασθένειας, αλλά στην εργασία αυτή θα δώσουμε σημασία στην τοπολογία του λογικού δικτύου των εφαρμογών η οποία επηρεάζει κατά πολύ περισσότερο την διαδικασία μιας επιδημίας. Στην πρώτη ενότητα θα δούμε την εικόνα της

επιδημίας από την πλευρά της βιολογίας, ώστε να έχουμε μια πρώτη άποψη. Στην δεύτερη ενότητα θα δούμε πώς η τοπολογία ενός δικτύου σχετίζεται με τις επιδημίες, αν και υπάρχουν και άλλα σημαντικά χαρακτηριστικά που θα μπορούσε κανείς να λάβει υπόψη του. Στη συνέχεια θα χρησιμοποιήσουμε ένα κατάλληλο μαθηματικό μοντέλο με υποθέσεις οι οποίες ταιριάζουν σε ένα δίκτυο υπολογιστών. Προφανώς δεν θα είχε νόημα να εξετάσει κανείς θεωρητικά μια επιδημία μιας συγκεκριμένης ασθένειας, οπότε το επιλεγμένο μοντέλο θα είναι επιθυμητό να αφορά σε επιδημίες διαφορετικών ιών με χαρακτηριστικά κοινά, ώστε να περιγράφονται με ενιαίο τρόπο.

Στην τρίτη ενότητα συνδυάζοντας ένα γνωστό μοντέλο από την βιολογία και τα αποτελέσματα της δεύτερης ενότητας καταλήγουμε σε μια παραλλαγή του μοντέλου σε στοχαστική εκδοχή του. Μια διαφοροποίησή μας είναι ότι η δυνατότητα μετάδοσης της ασθένειας και τελικά η διάδοση της επιδημίας δεν συμβαίνει μόνο από τους μολυσμένους προς τους ευάλωτους κόμβους. Υπάρχει δυνατότητα ένας μικρός αριθμός από κακόβουλους κόμβους να έχει ενεργή δράση εντός τού πληθυσμού των ευάλωτων. Μια τέτοια επιλογή φαίνεται λογική στα δίκτυα υπολογιστών όπου αν και απαιτείται συγκατάθεση του χρήστη σε μια εφαρμογή, οι χρήστες έχουν πλέον πιο ενεργή δράση σε σχέση με το παρελθόν. Τελικά αναγνωρίζουμε τις προσεγγίσεις μας και το όριο ισχύος του μοντέλου διάδοσης επιδημιών.

1. Γνωστά αιτιοκρατικά μοντέλα επιδημιών

Στην πρώτη αυτή ενότητα θα εξετάσουμε τα γνωστά από τη βιολογία κυρίως, μερικά απλά, μοντέλα. Η επαφή μας με αυτά, εκτός από μια πρώτη γνωριμία με τις διαφορετικές περιγραφές που μπορεί να έχει μια επιδημία, θα μας εφοδιάσει με τα κατάλληλα εννοιολογικά εργαλεία, ώστε να έχουμε μια αρχική εκτίμηση των υποθέσεων που θα υιοθετήσουμε καθώς προσεγγίζουμε τα δίκτυα. Στα επόμενα μοντέλα έχουμε ως παραμέτρους τους ρυθμούς που ο κάθε πληθυσμός περνάει σε κάποιον άλλο. Επίσης μπορούμε να ενσωματώσουμε τους ρυθμούς γεννήσεων - θανάτων. Όσο πιο σύνθετο είναι το μοντέλο τόσο περισσότερους πληθυσμούς ή και παραμέτρους θα έχουμε, ώστε να ενσωματωθεί ένα ακόμα χαρακτηριστικό της ασθένειας.

1.1 Ντετερμινιστικά Μοντέλα

Μια κατηγοριοποίηση των μοντέλων είναι αν έχουν αιτιοκρατική ή στοχαστική μορφή. Αυτό καθορίζεται συνήθως από την ίδια τη φύση του φαινομένου που θα περιγράψουμε, αλλά υπάρχουν περιπτώσεις όπου η δυνατότητα να επιλέξουμε καθορίζεται από άλλους λόγους. Η συνήθης τακτική ευνοεί την περιγραφή με διαφορικές εξισώσεις όταν έχουμε μεγάλους πληθυσμούς και τότε η μέση περιγραφή πεδίου είναι ικανοποιητική.

1.1.1 SIR (Susceptible, Infected, Recovered) Μοντέλο

Σε αυτό το μοντέλο (A. G. McKendrick and W. O. Kermack) [1] ο συνολικός πληθυσμός N διαιρείται στις τρεις ομάδες $S(t)$, $I(t)$ και $R(t)$. Η $S(t)$ αντιπροσωπεύει το μέρος του πληθυσμού που είναι ευάλωτο στην ασθένεια, η $I(t)$ το μέρος που είναι μολυσμένο και τα μέλη του οποίου είναι ικανά να μεταδώσουν την ασθένεια στα μέλη της $S(t)$. Η ποσότητα $R(t)$ αντιπροσωπεύει το μέρος του πληθυσμού τα μέλη του οποίου είχαν μολυνθεί αλλά ξεπέρασαν την ασθένεια και δεν μπορούν να μολυνθούν ξανά ή να την μεταδώσουν. Έχουμε: $N = S(t) + I(t) + R(t)$. Οι εξισώσεις που περιγράφουν το δυναμικό σύστημα είναι:

$$dS/dt = -\beta SI \quad (1.1.1)$$

$$dI/dt = \beta SI - \gamma I \quad (1.1.2)$$

$$dR/dt = \gamma I \quad (1.1.3)$$

Περιγραφή του μοντέλου SIR [1], [2]

Το κάθε άτομο του πληθυσμού έχει την ίδια πιθανότητα με τα υπόλοιπα να μολυνθεί με ρυθμό β ο οποίος θεωρείται ο τρόπος μετάδοσης ή ο ρυθμός μόλυνσης από την ασθένεια. Η εξίσωση (1.1.1) αναπαριστά το γεγονός ότι ο ρυθμός με τον οποίο μολύνονται τα ευάλωτα άτομα είναι ανάλογος με το πλήθος τους καθώς και με το πλήθος των μολυσμένων. Η σταθερά β έχει λοιπόν αυτή την έννοια κλίμακας στον παραπάνω ρυθμό αύξησης. Επιπλέον, ένα άτομο μολυσμένο μπορεί να έρθει σε επαφή με βN άλλους ανά μονάδα χρόνου. Το κλάσμα των ευάλωτων προς τον συνολικό πληθυσμό είναι S/N . Ο αριθμός των νέων μολυσμένων στη μονάδα του χρόνου ανά μολυσμένο άτομο είναι τότε $(\beta N)(S/N)$ το οποίο δίνει τον ρυθμό των μολυσμένων ως $(\beta N)(S/N)I = \beta SI$.

Για τις (1.1.2), (1.1.3) εξισώσεις θεωρούμε τον πληθυσμό που αφήνει την ομάδα $S(t)$ ως ίσο με τον πληθυσμό που εισέρχεται στην μολυσμένη ομάδα. Παρόλα αυτά, ένας αριθμός ίσος με το κλάσμα αυτών που θεραπεύονται ανά μονάδα χρόνου αφαιρείται από την ομάδα των μολυσμένων και εισέρχεται στην ομάδα $R(t)$. (Το γ παριστάνει το μέσο ρυθμό ανάρρωσης ή το $1/\gamma$ το μέσο χρόνο μόλυνσης). Ο πληθυσμός αυτός δεν μπορεί να μεταδώσει την ασθένεια ούτε να την αποκτήσει ξανά. Έχει πλέον την ιδιότητα της ανοσίας για αυτή την ασθένεια.

Οι παραπάνω διαδικασίες οι οποίες συμβαίνουν ταυτόχρονα αναφέρονται ως νόμος δράσης των μαζών, μια ευρέως αποδεκτή ιδέα κατά την οποία ο ρυθμός της επαφής μεταξύ δύο ομάδων ενός πληθυσμού είναι ανάλογος του μεγέθους της κάθε ομάδας.

Τελικά υποθέτουμε ότι ο ρυθμός μόλυνσης και θεραπείας είναι πολύ μεγαλύτερος από την κλίμακα χρόνου γεννήσεων-θανάτων, οπότε αυτές οι παράμετροι αγνοούνται σε αυτό το μοντέλο. Μία ακόμα υπόθεση του μοντέλου είναι ότι ο συνολικός πληθυσμός έχει αναμειχθεί αρκετά, ώστε κάθε ζεύγος ατόμων έχει την ίδια πιθανότητα επαφής.

Θα εξετάσουμε με κάπως μεγαλύτερη λεπτομέρεια [2] πώς παράγονται οι εξισώσεις του μοντέλου, με σκοπό να κάνουμε κάποιες ενδιαφέρουσες παρατηρήσεις και να δούμε κάποια ποιοτικά χαρακτηριστικά των συνεπειών του. Θα φανεί χρήσιμο να θεωρήσουμε διακριτό χρόνο χωρίς βλάβη της γενικότητας.

Ένα ευάλωτο άτομο μπορεί να μεταβεί στον πληθυσμό I , αν έρθει σε επαφή με ένα μολυσμένο άτομο. Τι ορίζεται ως επαφή, εξαρτάται από τη φύση της ασθένειας. Θεωρούμε ότι κάθε μολυσμένο άτομο έρχεται με λ άτομα σε επαφή κατά μέσο όρο ανά περίοδο χρόνου. (Σε συνεχή χρόνο θα ήταν ανά μονάδα χρόνου). Έστω α το ποσοστό των επαφών με αποτέλεσμα μετάδοσης. Τότε ο αριθμός των υποψηφίων μεταδόσεων ανά μολυσμένο άτομο μπορεί να είναι $\alpha*\lambda$. Ορίζουμε $\beta = \alpha*\lambda$ το μέσο αριθμό των πιθανών μεταδόσεων ανά μολυσμένο άτομο σε κάθε περίοδο. Ομοίως ένα μολυσμένο

άτομο μπορεί να θεραπευτεί και ορίζουμε γ το μέσο αριθμό θεραπειών ανά μολυσμένο άτομο σε κάθε περίοδο.

Οι δυνατές περιπτώσεις επαφών μολυσμένων ατόμων οι οποίες οδηγούν την κατάσταση του συστήματος από την χρονική στιγμή t σε αυτήν της χρονικής στιγμής $t+1$ είναι:

1. Μετάδοση ενός μολυσμένου σε έναν ευάλωτο γεγονός το οποίο οδηγεί σε έναν ακόμα μολυσμένο. Εφόσον S_t το πλήθος των ευάλωτων, κάθε μολυσμένο άτομο παράγει βS_t μολύνσεις κάθε περίοδο.

2. Θεραπεία ενός μολυσμένου γεγονός το οποίο οδηγεί σε έναν ακόμα θεραπευμένο. Εφόσον έχουμε I_t μολυσμένους κάθε περίοδο μπορούν να θεραπευτούν γI_t άτομα.

Υπάρχουν και οι περιπτώσεις επαφής ενός μολυσμένου με έναν επίσης μολυσμένο ή με έναν απομονωμένο οι οποίες όμως δεν αλλάζουν την κατάσταση του συστήματος. Επίσης η επαφή των ατόμων από τον πληθυσμό S_t με άτομα του R_t δεν επιφέρει καμία αλλαγή προφανώς.

Δεδομένης της τωρινής κατάστασης σε χρόνο t του πληθυσμού η οποία περιγράφεται από τους S_t , I_t και R_t (ή αντίστοιχα από τα ποσοστά s_t, i_t και r_t τα οποία προκύπτουν με διαίρεση του N) καταλήγουμε σύμφωνα με τα παραπάνω στις επόμενες εξισώσεις για την περιγραφή του συστήματος στο χρόνο $t+1$:

$$S_{t+1} = S_t - \beta S_t I_t \quad (1.1.4)$$

$$I_{t+1} = I_t + \beta S_t I_t - \gamma I_t \quad (1.1.5)$$

$$R_{t+1} = R_t + \gamma I_t \quad (1.1.6)$$

Από αυτές τις εξισώσεις μπορούν να προκύψουν και οι αντίστοιχες συνεχούς χρόνου αν η περίοδος αντικατασταθεί με ένα μικρό χρονικό διάστημα δt και οι ρυθμοί β , γ θεωρηθούν ότι είναι ανά μονάδα χρόνου. Τότε θα πάρουμε από τις προηγούμενες εξισώσεις:

$$\delta S(t) = -\beta S(t)I(t)\delta t$$

$$\delta I(t) = \beta S(t)I(t)\delta t - \gamma S(t)\delta t$$

$\delta R(t) = \gamma R(t)\delta t$ από όπου διαιρώντας με δt και παίρνοντας το όριο $\delta t \rightarrow 0$ καταλήγουμε στο σύστημα διαφορικών εξισώσεων που αναφέρθηκαν στην αρχή.

Όπως θα δούμε, δεν υπάρχουν θετικές τιμές σταθερής κατάστασης μολυσμένων ατόμων για σταθερό συνολικό πληθυσμό N [2]. Αρχικά ας υποθέσουμε ότι $I(t)=0$. Από την εξίσωση (1.1.5)

βλέπουμε ότι $I_{t+1}=0$. Επομένως αν δεν ξεκινήσουμε έστω με έναν μολυσμένο δεν θα αλλάξουμε κατάσταση. Έστω $I(t)>0$. Θα εξετάσουμε αν το I_{t+1} αυξάνεται, μειώνεται ή διατηρείται σταθερό. $I_{t+1}=I_t(1+\beta S_t-\gamma)$. Ορίζουμε $\rho_t=1+\beta S_t-\gamma$. Αν $\rho_t>1$ τότε $I_{t+1}>I_t$. Ο αριθμός των αρχικά μολυσμένων αυξάνεται, ενώ αν $\rho_t<1$ τότε $I_{t+1}<I_t$ ο αριθμός τους μειώνεται. Ας θεωρήσουμε $\rho_t=1$ και θα δείξουμε ότι αυτό δεν μπορεί να διατηρηθεί. Είναι $\rho_{t+1}=1+\beta S_{t+1}-\gamma \Rightarrow \rho_{t+1}=1+\beta(S_t-\beta S_t I_t)-\gamma=1-\beta^2 S_t I_t < \rho_t=1$. Επομένως δεν θα έχουμε μια σταθερή κατάσταση εφόσον τότε τα μολυσμένα άτομα θα μειώνονται. Επιπλέον έχουμε ότι αν το I_t μειώνεται αυτό συμβαίνει με θετικό ρυθμό, ενώ αν αυξάνει συμβαίνει με αρνητικό ρυθμό. Συνδυάζοντας αυτά τα αποτελέσματα παίρνουμε ότι σε μεγάλο χρονικό διάστημα ο μολυσμένος πληθυσμός θα εξαφανιστεί για έναν συνολικά σταθερό πληθυσμό. Το $S(t)$ και $R(t)$ θα έλθουν σε μια οριακή κατάσταση ισορροπίας όταν $\lim_{t \rightarrow \infty} I(t) = 0$

1.1.2 Το μοντέλο SIR με γεννήσεις-θανάτους

Υποθέτοντας ότι οι γεννήσεις-θάνατοι έχουν τελικά μεγάλη επίδραση στη δυναμική του συστήματος, μπορούμε να καταλήξουμε [4] στις παρακάτω εξισώσεις:

$$\frac{dS}{dt} = -\beta SI + \mu(N - S)$$

$$\frac{dI}{dt} = \beta SI - \gamma I - \mu I$$

$$\frac{dR}{dt} = \gamma I - \mu R$$

Όπου μ είναι ο κοινός μέσος ρυθμός θανάτων-γεννήσεων. Το μοντέλο αυτό θεωρεί ότι ο ρυθμός γεννήσεων είναι ίσος με τον ρυθμό θανάτων οπότε ο συνολικός πληθυσμός διατηρείται σταθερός. Μπορεί να προκύψει εύκολα αν συλλογιστούμε ότι μια γέννηση στο σύστημα αυξάνει κατά ένα μόνο το πλήθος των ευάλωτων ατόμων ενώ ένας θάνατος μπορεί να συμβεί και στις τρεις ομάδες που θεωρούμε, οπότε ο ρυθμός του επί το πλήθος του αντίστοιχου πληθυσμού δίνει τα άτομα που αφήνουν τον πληθυσμό.

1.1.3 Το μοντέλο SIS με γεννήσεις - θανάτους

Το μοντέλο SIS μπορεί εύκολα να προκύψει από το SIR θεωρώντας ότι τα άτομα που θεραπεύονται δεν έχουν ανοσία στην ασθένεια, δηλαδή είναι άμεσα ευάλωτα από τη στιγμή που θεραπεύονται.

Αφαιρώντας την εξίσωση που αφορά στον πληθυσμό που θεραπεύτηκε και προσθέτοντας αυτούς που αφαιρούνται από τους μολυσμένους στους ευάλωτους, παίρνουμε τις παρακάτω διαφορικές εξισώσεις:

$$\frac{dS}{dt} = -\beta SI + \mu(N - S) + \gamma I$$

$$\frac{dI}{dt} = \beta SI - \gamma I - \mu I$$

1.1.4 Το SIRS μοντέλο

Αυτό το μοντέλο είναι απλώς μια επέκταση του μοντέλου SIR. Η μόνη διαφορά είναι ότι επιτρέπει μέλη του πληθυσμού $R(t)$ να εισέλθουν στην ευάλωτη ομάδα. Η παράμετρος f ορίζεται ως ο μέσος ρυθμός απώλειας ανοσίας των μελών που έχουν θεραπευτεί.

$$\frac{dS}{dt} = -\beta SI + \mu(N - S) + fR$$

$$\frac{dI}{dt} = \beta SI - \gamma I - \mu I$$

$$\frac{dR}{dt} = \gamma I - \mu R - fR$$

1.1.5 Το μοντέλο SEIS

Το μοντέλο αυτό λαμβάνει υπόψη του την περίοδο έκθεσης στην ασθένεια, γεγονός το οποίο δίνει τον επί μέρους πληθυσμό $E(t)$. Ο πληθυσμός E είναι εκτεθειμένος στην ασθένεια αλλά όχι ικανός ακόμα να μολύνει τον ευάλωτο πληθυσμό.

$$\frac{dS}{dt} = B - \beta SI - \mu S + \gamma I$$

$$\frac{dE}{dt} = \beta SI - (\varepsilon + \mu)E$$

$$\frac{dI}{dt} = \varepsilon E - (\gamma + \mu)I$$

Οπού B ο μέσος ρυθμός γεννήσεων και $1/\varepsilon$ η μέση λανθάνουσα περίοδος.

1.1.6 Το μοντέλο SEIR

Το μοντέλο SIR που συζητήθηκε παραπάνω λαμβάνει υπόψη του ασθένειες κατά τις οποίες ένα άτομο μπορεί να μεταδώσει σε άλλους κατευθείαν εφόσον είναι φορέας. Σε πολλές ασθένειες υπάρχει αυτό που περιγράφεται με τον όρο φάση έκθεσης όπου ένας φορέας έχει την ασθένεια αλλά δεν μπορεί να την μεταδώσει.

Σε αυτό το μοντέλο ο αρχικός πληθυσμός διαιρείται σε 4 επί μέρους: Τους ευάλωτους $S(t)$, τους φορείς $E(t)$, τους μεταδότες $I(t)$ και τους απομονωμένους (αυτοί που έχουν θεραπευτεί και δεν μπορούν να μεταδώσουν την ασθένεια) $R(t)$. Θα ισχύει και εδώ $N = S(t) + E(t) + I(t) + R(t)$.

$$\frac{dS}{dt} = B - \beta SI - \mu S$$

$$\frac{dE}{dt} = \beta SI - (\varepsilon + \mu)E$$

$$\frac{dI}{dt} = \varepsilon E - (\gamma + \mu)I$$

$$\frac{dR}{dt} = \gamma I - \mu R$$

Σχετικά με την ευστάθεια κάποιων από τα μοντέλα μπορούμε να κάνουμε σχετικές διατυπώσεις με τη βοήθεια των εξισώσεων Lyapunov. Από την αναφορά [5] και το σχετικό θεώρημα των Andrei Korobeinikov και Philip K. Maini διαπιστώνουμε ότι για τα μοντέλα SIR με γεννήσεις και θανάτους καθώς και SEIR οι καταστάσεις ισορροπίας είναι καθολικά ασυμπτωτικά ευσταθείς. Το θεώρημα μας λέει πολλά περισσότερα για εκδόσεις των μοντέλων αυτών όπου ο ρυθμός μόλυνσης έχει την μορφή $\beta I^p S^q$. Συγκεκριμένα εφόσον $p \leq 1$ το παραπάνω αποτέλεσμα για την ευστάθεια ισχύει και μάλιστα αυτό δεν εξαρτάται από την τιμή της παραμέτρου q .

1.1.7 Το μοντέλο MSIR

Υπάρχουν επιδημίες όπου ένα άτομο γεννιέται με παθητική ανοσία από τη μητέρα του. Αυτό περιγράφεται μαθηματικά με την εισαγωγή ενός ακόμα πληθυσμού $M(t)$.

$$\frac{dM}{dt} = B - \delta MS - \mu M$$

$$\frac{dS}{dt} = \delta MS - \beta SI - \mu S$$

$$\frac{dI}{dt} = \beta SI - \gamma I - \mu I$$

$$\frac{dR}{dt} = \gamma I - \mu R$$

Όπου δ η μέση προσωρινή περίοδος ανοσίας.

1.1.8 Το μοντέλο MSEIR

Εισάγοντας στο μοντέλο MSIR την λανθάνουσα περίοδο (αντίστοιχα με την κατάσταση όπου έχουμε φορείς αλλά όχι μεταδότες) παίρνουμε το μοντέλο MSEIR.

$$\frac{dM}{dt} = B - \delta MS - \mu S$$

$$\frac{dS}{dt} = \delta MS - \beta SI - \mu S$$

$$\frac{dE}{dt} = \beta SI - (\varepsilon + \mu)E$$

$$\frac{dI}{dt} = \varepsilon E - (\gamma + \mu)I$$

$$\frac{dR}{dt} = \gamma I - \mu R$$

1.2 Συνθήκες έναρξης επιδημίας – Χρονική εξέλιξη του μοντέλου SIR

Με σκοπό να εμβαθύνουμε στο μοντέλο SIR το οποίο θα μας απασχολήσει στη συνέχεια ως αντιπροσωπευτικό των δικτύων, θα συνεχίσουμε την ανάλυσή του [20]. Για το μοντέλο SIR δεν υπάρχει μια λύση κλειστής μορφής όπως είναι γνωστό. Σε κάθε περίπτωση μια αριθμητική λύση είναι δυνατή για κάθε πρακτικό σκοπό. Στα επόμενα θα δούμε κάποια αποτελέσματα των εξισώσεων τα οποία απαντούν σε ενδιαφέροντα ερωτήματα σχετικά με την επιδημία που περιγράφεται από το μοντέλο.

Αν και το μοντέλο αυτό είναι απλό, μπορούμε να καταλήξουμε σε ορισμένα ενδιαφέροντα αποτελέσματα. Αρχικά θα θεωρήσουμε τις αρχικές τιμές των εκτεθειμένων στην ασθένεια $I(0)=I_0$ και

ευάλωτων $S(0)=S_0$. Τα δεδομένα αυτά μαζί με τις τιμές των ρυθμών μόλυνσης β και θεραπείας γ μπορούν να μας δώσουν απάντηση στο ενδιαφέρον ερώτημα αν η επιδημία θα διαδοθεί.

Από την εξίσωση (1.1.2) έχουμε $\left. \frac{dI}{dt} \right|_{t=0} = I_0(\beta S_0 - \gamma)$ απ' όπου συμπεραίνουμε ότι $\left. \frac{dI}{dt} \right|_{t=0} > 0$ αν $\beta S_0 - \gamma > 0$ ή $S_0 > \frac{\gamma}{\beta} = \rho$. Επιπλέον από την (1.1.1) είναι για κάθε t $\frac{dS}{dt} \leq 0$ οπότε $S < S_0$. Σύμφωνα με τα παραπάνω διακρίνουμε τις περιπτώσεις :

A) $S_0 < \rho$ οπότε $\frac{dI}{dt} \leq 0$ για κάθε t . Σε αυτή την περίπτωση ο αριθμός των μολυσμένων παραμένει κάτω από I_0 και τείνει στο 0 καθώς αυξάνει ο χρόνος. Η επιδημία δεν μπορεί να αρχίσει.

B) $S_0 \geq \rho$ Σε αυτή την περίπτωση το $I(t)$ αρχίζει να αυξάνεται το οποίο είναι και ο ορισμός της επιδημίας.

Η παραπάνω συνθήκη για την έναρξη της επιδημίας δίνεται και μέσω του καθιερωμένου μεγέθους R_0 που καλείται ρυθμός αναπαραγωγής. Ορίζοντας

$$R_0 = S_0 \frac{\gamma}{\beta} = \frac{S_0}{\rho} \quad (1.2.1)$$

προκύπτει όπως και προηγουμένως η συνθήκη για την έναρξη της επιδημίας

$$R_0 > 1 \quad (1.2.2)$$

Η ασθένεια κατά μέσο όρο δεν διαδίδεται εντός του χρόνου μόλυνσης και σβήνει με την πάροδο του χρόνου για $R_0 < 1$. Στην ειδική περίπτωση όπου $R_0 = 1$ η ασθένεια γίνεται επιδημία με την έννοια ότι παραμένει στον πληθυσμό σε ένα σταθερό ρυθμό καθώς ένας μολυσμένος μπορεί να μεταδώσει την ασθένεια σε έναν ευάλωτο.

Στην συνέχεια θα εξετάσουμε την μέγιστη τιμή των ασθενών σε μια επιδημία της μορφής SIR. Η προηγούμενη διαπίστωσή μας για την αρχική αύξηση των ασθενών και την τελική τους εξαφάνιση σε μεγάλο χρονικό διάστημα στο τέλος της επιδημίας φανερώνει ότι σε κάποια χρονική στιγμή θα υπάρχει ένα μέγιστο. Από τις εξισώσεις (1.1.1) και (1.1.2), αν και δεν θα υπολογίσουμε ακριβώς την χρονική συνάρτηση του $I(t)$, έχουμε $\frac{dI}{dS} = -\frac{I(\beta S - \gamma)}{\beta S I} = -1 + \frac{\rho}{S}$, $I \neq 0$ απ' όπου $dI = -dS + \rho \frac{dS}{S}$ ολοκληρώνοντας την τελευταία με αρχικές συνθήκες $R_0=0$ και I_0 (οπότε $I_0+S_0= I_0+N- I_0=N$) παίρνουμε $I + S - \rho \ln S = c = I_0 + S_0 - \rho \ln S_0$, όπου c σταθερά. (1.2.3)

Διακρίνοντας την σχέση αυτή ως συνάρτηση του S , $I = c - S + \rho \ln S$ έχουμε $\frac{dI}{dS} = -1 + \frac{\rho}{S}$ το οποίο δίνει την μέγιστη τιμή για $S=\rho$ η οποία προκύπτει $I = I_0 + S_0 - \rho \ln S_0 - \rho + \rho \ln \rho$ και τελικά λόγω της αρχικής συνθήκης

$$I_{max} = N - \rho + \rho \ln \frac{\rho}{S_0} \quad (1.2.4)$$

Στη συνέχεια με χρήση και της εξίσωσης (1.1.3) των απομονωμένων σε συνδυασμό με την (1.1.1) μπορούμε να καταλήξουμε σε επιπλέον αποτελέσματα.

$$\frac{dS}{dR} = -\frac{\beta SI}{\gamma I} = -\frac{S}{\rho}, \text{ για } I \neq 0 \Rightarrow \frac{dS}{S} = -\frac{dR}{\rho} \Rightarrow \int_{S_0}^S \frac{dS}{S} = \int_0^R \frac{dR}{\rho} \Rightarrow \ln \frac{S}{S_0} = -\frac{R}{\rho} \Rightarrow$$

$$S = S_0 e^{-\frac{R}{\rho}} \geq S_0 e^{-\frac{N}{\rho}}, \text{ για } I \neq 0 \quad (1.2.5)$$

Η (1.2.5) ισχύει και για $I = 0$ εφαρμόζοντας τον κανόνα de' Hospital στην αρχική έκφραση. Είναι

$$\lim_{t \rightarrow \infty} \frac{dS}{dR} = \lim_{I \rightarrow 0} -\frac{\beta SI}{\gamma I} = -\frac{\beta}{\gamma} \lim_{I \rightarrow 0} \frac{(SI)'}{(I)'} = -\frac{\beta}{\gamma} \lim_{I \rightarrow 0} (S'I + S) = -S \frac{\beta}{\gamma} = -\frac{S}{\rho} \text{ οπότε θα έχουμε τελικά}$$

$$S_{\infty} = S_0 e^{-\frac{R_{\infty}}{\rho}} \text{ όπως και προηγουμένως. Εναλλακτικά είναι}$$

$$S_{\infty} = S_0 e^{-\frac{N-S_{\infty}}{\rho}} \quad (1.2.6)$$

Από την εξίσωση αυτή συμπεραίνουμε ότι θα είναι αυστηρά $S_{\infty} > 0$. Το αποτέλεσμα αυτό σημαίνει ότι στο τέλος της επιδημίας θα υπάρχουν ακόμα κάποιοι ευάλωτοι οι οποίοι δεν δέχτηκαν την ασθένεια. Η επιδημία δεν θα μολύνει τελικά ολόκληρο τον πληθυσμό. Από την (1.2.6) επίσης με αριθμητικές μεθόδους και γνωστές τις παραμέτρους μπορούμε να υπολογίσουμε τον τελικά απομείναντα αριθμό των ευάλωτων στο τέλος της επιδημίας. Επίσης ο αριθμός των ασθενών σε όλη τη διάρκεια της επιδημίας συνολικά θα ισούται προφανώς (για $R_0 = 0$) με:

$$R_{\infty} = I_{total} = N - S_{\infty} \quad (1.2.7)$$

ΕΝΟΤΗΤΑ 2

2. Επίδραση της τοπολογίας στην διάδοση επιδημιών

Στην δεύτερη αυτή ενότητα θα προσπαθήσουμε να προσεγγίσουμε από πιο κοντά την επιδημία σε ένα δίκτυο υπολογιστών. Αρχικά θα αναφέρουμε κάποια ιδιαίτερα χαρακτηριστικά των δικτύων σε σχέση με την επιδημία και στην συνέχεια θα εξετάσουμε τις τοπολογικές επιπτώσεις. Αρχίζοντας να εξετάζουμε την διάδοση ασθενειών σε δίκτυα υπολογιστών, κάνουμε μια πρώτη παρατήρηση η οποία τα διαφοροποιεί σε σχέση με τους πληθυσμούς κοινωνιών. Οι ζωντανοί οργανισμοί κινούνται και ως αποτέλεσμα υπάρχει η έννοια της συσσώρευσης. Η επιφανειακή πυκνότητα των ατόμων θεωρείται παράγοντας υψίστης σημασίας για την πιθανότητα μόλυνσης από επιδημία, εφόσον τότε έχουμε αύξηση των επαφών. Οι ροές κίνησης μπορούν να συμπεριληφθούν σε ένα μοντέλο όπου ο χώρος θα εισάγεται φυσιολογικά εφόσον ενδιαφερόμαστε και για τη χωρική ανάπτυξη της επιδημίας. Αν και οι υπολογιστές μας κινούνται μάλλον σε ένα μικρότερο ποσοστό ο τρόπος της μόλυνσης δεν έχει σχέση με αυτή την κίνηση στο χώρο. Παρόλα αυτά θα ήταν λάθος να θεωρήσουμε ότι δεν υπάρχει κίνηση σε ένα μοντέλο επιδημίας για δίκτυο υπολογιστών. Στη πραγματικότητα η κίνηση είναι εντονότατη στα δίκτυα, εφόσον σε ελάχιστο χρόνο ένα άτομο μπορεί να έρθει σε επαφή με πολλά άλλα και μάλιστα ταυτόχρονα. Εδώ λοιπόν η έννοια της κίνησης θα πρέπει να διαφοροποιεί την διαίσθησή μας σε σχέση με την βιολογική περίπτωση. Αν και τα προηγούμενα μοντέλα που περιγράψαμε δεν εισάγουν εμφανώς μια χωρική παράμετρο, στα δίκτυα αυτό πρέπει να ληφθεί είτε άμεσα είτε έμμεσα υπόψη μας. Εδώ οι υποψήφιοι δέκτες μιας ασθένειας έρχονται σχεδόν στιγμιαία σε επαφή με τον ιό ή γενικότερα το κακόβουλο λογισμικό.

Ως κακόβουλο λογισμικό για έναν υπολογιστή μπορούμε να θεωρήσουμε καταρχήν οποιοδήποτε ενεργεί εν αγνοία του χρήστη. Ο ορισμός αυτός χρήζει επιπλέον εξέτασης εφόσον ένας χρήστης ενδέχεται να συμφωνήσει στην εγκατάσταση ενός λογισμικού ακόμα και αν αυτό τρέξει έπειτα από σχετική ερώτηση αλλά κάποιες ανεπιθύμητες ενέργειες να συμβούν και να υπάρχει διάδοση αυτών. Προφανώς για να είναι ανεπιθύμητες, κάποιιοι άλλοι χρήστες δεν θα είναι σύμφωνοι στο τρέξιμο του λογισμικού. Ο παραπάνω ορισμός δεν θα μας απασχολήσει μιας και δεν είναι αντικείμενο της παρούσας εργασίας αλλά τον αναφέρουμε στο βαθμό που καθορίζει τη φύση μιας ασθένειας και κατ' επέκταση τον τρόπο παραγωγής ενός μοντέλου. Για παράδειγμα αναφέρουμε το είδος ιού όπου η εφαρμογή του έχει ως αποτέλεσμα την παύση λειτουργίας ενός υπολογιστή οπότε εδώ οι γεννήσεις-θάνατοι παίζουν καθοριστικό ρόλο, ενώ στο άλλο άκρο ένα απλό e-mail με διαφημιστικό περιεχόμενο μπορεί απλώς να διαδίδεται από άτομο σε άτομο. Το τελευταίο παράδειγμα κάνει

εμφανή και μια διάκριση των διαδόσεων, σε αυτές που ο χρήστης αποφασίζει και αυτές που συμβαίνουν με πιο έξυπνο τρόπο εν πλήρη άγνοιά του από μηχανήμα σε μηχανήμα και από δίκτυο σε δίκτυο χωρίς η ανθρώπινη παρουσία να εμπλέκεται στην περιγραφή του μολυσμένου ο οποίος μολύνει. Η προηγούμενη συζήτηση διαφοροποιεί την έννοια του ατόμου σε έναν πληθυσμό αφού όπως είδαμε για κάποιες ασθένειες που διαδίδονται η κρίσιμη στιγμή της μόλυνσης συμβαίνει από ανθρώπινη ενέργεια άμεσα σχετική με την ασθένεια. Σε άλλες οι άνθρωποι εκτελούν τελείως διαφορετικές ενέργειες αλλά μια ασθένεια διαδίδεται ακόμα και αν αυτοί απουσιάζουν και απλώς οι υπολογιστές είναι συνδεδεμένοι στο δίκτυο. Ο πληθυσμός λοιπόν δεν είναι ένα κλειστό σύστημα από υπολογιστές, αλλά η έννοια του ατόμου-υπολογιστή θα πρέπει σε κάποιες περιπτώσεις να συμπληρώνεται από την έννοια του χρήστη. Τότε η πιθανότητα μόλυνσης αποκτά ένα τελείως διαφορετικό νόημα.

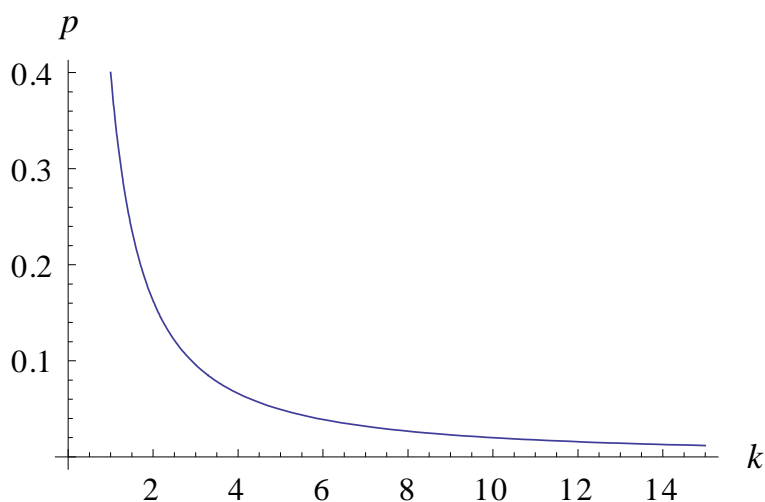
Στα προηγούμενα δεν αναφέραμε τα antivirus ή τα firewall μέσω των οποίων η ασθένεια θεραπεύεται και η διάδοση επιβραδύνεται. Προφανώς θα έχουν το ανάλογο τους με τον ρυθμό θεραπείας γ . Μια διαισθητική παρατήρησή μας σχετικά με τη θεραπεία γ σε υπολογιστές είναι μια μάλλον μεγάλη τιμή του ρυθμού θεραπείας όταν το "αντιβιοτικό" βρεθεί. Συνήθως οι υπολογιστές συνδεδεμένοι σε ένα δίκτυο λαμβάνουν τα updates και ξεπερνούν τους ιούς μαζικά. Ένα ρεαλιστικό λοιπόν μοντέλο θα πρέπει να ενσωματώνει με κάποιον τρόπο αυτή την περίοδο όπου δεν έχει βρεθεί θεραπεία και μια γρήγορη παύση της ασθένειας όταν τρέξει το κατάλληλο λογισμικό.

Θα πρέπει να εξετάσουμε την υπόθεση της ισοτιμίας των ατόμων σε έναν ευάλωτο πληθυσμό. Ειδικά για τα δίκτυα υπολογιστών θεωρούμε ότι αυτός ο πληθυσμός, έστω $S(t)$, θα πρέπει να χωριστεί σε τουλάχιστον δύο ομάδες για ένα σημαντικό πλήθος ασθενειών. Πραγματικά είναι μάλλον αδύνατο να φανταστούμε όλα τα άτομα του $S(t)$ να έχουν την ίδια εκτίμηση για το τι θεωρείται ασθένεια και να πραγματοποιήσουν την επαφή που θα τα κάνει μέλη του πληθυσμού μολυσμένων $I(t)$. Η παρατήρηση αυτή μαζί με την ιδιότητα του ατόμου ως υπολογιστή-χρήστη μπορεί να ορίσει με έναν πιο αποτελεσματικό τρόπο τον πληθυσμό $S(t)$ ως άτομα με διαφορετική πιθανότητα μόλυνσης. Από τις παραπάνω παρατηρήσεις και συνδυάζοντας στα επόμενα την τοπολογία θα καταλήξουμε στο μοντέλο SIR ως κατάλληλο προς μια στοχαστική περιγραφή επιδημίας σε δίκτυα υπολογιστών.

2.1 Δομή Δικτύου

Θα αναφέρουμε περιληπτικά κάποιες έννοιες που σχετίζονται με την τοπολογία δικτύου, με σκοπό να δούμε πώς θα επηρεάσουν την μοντελοποίηση της διάδοσης επιδημιών χωρίς να επεκταθούμε

σε όλες τις έννοιες της θεωρίας γράφων. Αρχικά μπορούμε να χωρίσουμε τα δίκτυα σε φυσικά και λογικά. Τα φυσικά δίκτυα απαρτίζονται από υπολογιστικές μηχανές (υπολογιστές, δρομολογητές κτλ.) οι οποίες συνδέονται με φυσικό τρόπο μεταξύ τους μέσω ζεύξεων. Στην ορολογία του διαγράμματος δικτύου θα τις ονομάζουμε ακμές και σε αυτό το είδος δικτύου το διάγραμμα θα είναι μη κατευθυνόμενο. Ένα σημαντικό παράδειγμα είναι βέβαια το internet. Η άλλη κατηγορία, τα λογικά δίκτυα, αντιπροσωπεύει τις συνδέσεις σε επίπεδο εφαρμογών μεταξύ των ομότιμων διαδικασιών που ορίζουν οι χρήστες. Ορίζονται και εδώ οι έννοιες του κόμβου και του διαγράμματος δικτύου, το οποίο μπορεί να είναι και κατευθυνόμενο. Ένα σημαντικό παράδειγμα εδώ είναι ο ιστός. Το σημαντικό στοιχείο είναι ότι και στις δύο περιπτώσεις οι ιδιότητες του διαγράμματος είναι κοινές, οπότε μελετώντας τις μπορούμε να βρούμε με ποιον τρόπο μπορούμε να τις εισάγουμε σε ένα μοντέλο διάδοσης επιδημίας. Μια σημαντική έννοια ενός δικτύου είναι η κατανομή βαθμού του. Ο κάθε κόμβος έχει έναν αριθμό από συνδέσεις με τους γειτονικούς του (βαθμός κόμβου). Η κατανομή βαθμού $P(k)$ είναι η κατανομή που περιγράφει την πιθανότητα να έχουμε έναν κόμβο με k συνδέσεις. Είναι γνωστό ότι οι κατανομές βαθμού στα δίκτυα είναι έντονα δεξιά επικλινείς, δηλαδή οι πιο πολλοί κόμβοι έχουν μικρό βαθμό, αλλά υπάρχει ένας μικρός αριθμός κόμβων οι οποίοι έχουν μεγάλο βαθμό. Όταν η κατανομή βαθμού ακολουθεί ένα νόμο δύναμης $\sim k^{-\gamma}$ όπου γ συνήθως μεταξύ 2 και 3 τότε χαρακτηρίζουμε το δίκτυο ως δίκτυο απουσία κλίμακας (free-scale network).



Σχήμα 2.1.1

Στο παραπάνω σχήμα, ανεξάρτητα από την αριθμητική κλίμακα, φαίνεται η scale-free φύση των δικτύων που όπως φαίνεται περιγράφει αρκετά ρεαλιστικά την πραγματική φύση των δικτύων που μας ενδιαφέρουν όπως ο ιστός και το Internet. Πιο συγκεκριμένα η ιδιότητα αυτή σημαίνει ότι η μέση περιγραφή του βαθμού κόμβου δεν επαρκεί για την καθολική περιγραφή του δικτύου και

επομένως αυτό δεν μπορεί να έχει μια χαρακτηριστική μέση μέτρηση, οπότε ορίζεται ως “scale-free” [17]. Μια όψη αυτής της ιδιότητας είναι ότι τα δίκτυα δεν έχουν τόσο τυχαία δομή και σε αυτό θα επανέλθουμε αργότερα.

Clusterization

Ένα κοινό χαρακτηριστικό των δικτύων, το οποίο αναιρεί εν μέρει την τυχειότητα και εισάγει επιπλέον δομή, είναι το φαινόμενο κατά το οποίο ένας κόμβος έχει κύκλο γειτόνων ο οποίος τείνει να καλύπτει μέρος των κύκλων των γειτόνων του. Σε κοινωνικά δίκτυα αυτό μεταφράζεται με την παρατήρηση ότι οι φίλοι του φίλου ενός ατόμου τείνουν να γίνουν και δικοί του. Το φαινόμενο αυτό αναγνωρίστηκε πρώτα από τους Watts και Strogatz και από τότε κίνησε την προσοχή και άλλων ερευνητών. Θα μπορούσαμε να αποδώσουμε την έννοια αυτή ως όμοιο-γεινίαση. Είναι δυνατόν να ορίσουμε έναν συντελεστή όμοιο-γεινίασης ως το μέσο κλάσμα των ζευγών των γειτόνων. Έστω ο κόμβος i έχει k_i ακμές. Αν είναι όλοι γείτονες του i και μεταξύ τους γείτονες τότε θα έχουμε $\frac{k_i(k_i-1)}{2}$ συνδέσεις μεταξύ τους. Αν όμως υπάρχουν έστω E_i τότε ο συντελεστής όμοιο-γεινίασης ορίζεται ως

$$C_i = \frac{2E_i}{k_i(k_i-1)} \quad (2.1.1)$$

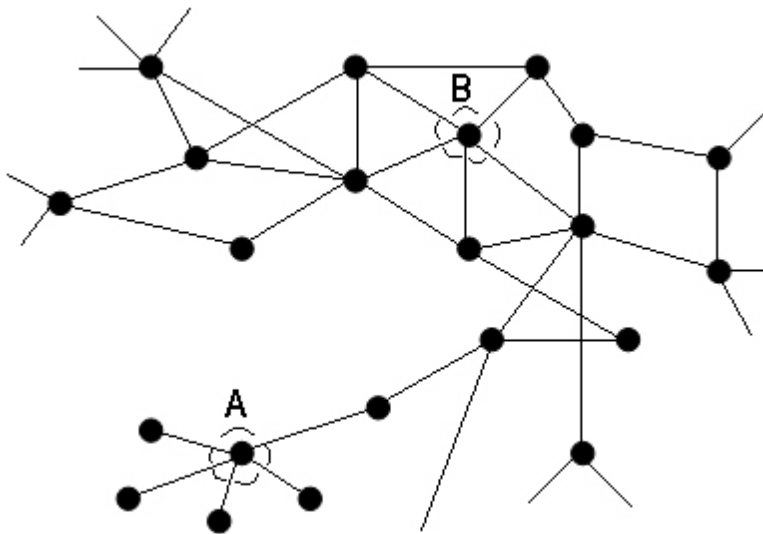
Ο συντελεστής όμοιο-γεινίασης ολόκληρου του δικτύου ορίζεται τότε ως ο μέσος όρος των C_i ως προς όλο το δίκτυο. Σε ένα πλήρως συνδεδεμένο δίκτυο θα είναι $C=1$. Μπορούμε να δούμε την έννοια αυτή ως την πιθανότητα κόμβοι που είναι γείτονες σε έναν κόμβο είναι και μεταξύ τους γείτονες. Σε ένα τυχαίο δίκτυο όπως αυτό των Erdos και Renyi είναι $C=p=\langle k \rangle/N$, το οποίο είναι πολύ μικρό για μεγάλα δίκτυα. Στην πραγματικότητα έχει βρεθεί ότι το C είναι σημαντικά μικρότερο της μονάδας αλλά μεγαλύτερο από N^{-1} .

2.2 Κεντρικότητα

Μία σημαντική ιδιότητά των κόμβων είναι η κεντρικότητα (centrality). Η έννοια αυτή εισάγει στην προσέγγισή μας το νέο στοιχείο ότι τα άτομα του πληθυσμού τα οποία διαδίδουν την ασθένεια δεν είναι πλέον ισότιμα. Το στοιχείο αυτό μπορεί να εισαχθεί και στον πληθυσμό των ευάλωτων ατόμων. Θα μπορούσαμε να θεωρήσουμε ότι τα τελευταία με καλή προσέγγιση είναι ισότιμα καθώς συνήθως οι κόμβοι ή οι υπολογιστές κάνουν ένα κοινό update κατά των ιών. Και όμως, όπως ήδη

αναφέραμε στην ενότητα 1, τα άτομα του πληθυσμού θα μπορούσαν να ιδωθούν και με την υπόσταση του χρήστη, όχι μόνο την υπολογιστική. Καθώς οι χρήστες αντιπροσωπεύουν ποικιλία συμπεριφορών και αντιδράσεων έχει νόημα και σε αυτούς μια τέτοια έννοια “βαθμού προδιάθεσης”. Εκτός από την παραπάνω θεώρηση η έννοια της κεντρικότητας έχει ένα συγκεκριμένο νόημα σχετικό με την τοπολογία του δικτύου το οποίο θα δούμε στη συνέχεια.

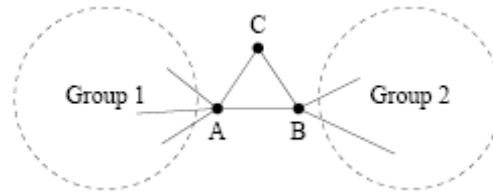
Διαισθητικά μπορεί να γίνει κατανοητό ότι η ικανότητα ενός κόμβου να έχει μεγάλη επιρροή στην διάδοση της πληροφορίας σχετίζεται με τη θέση του στο διάγραμμα δικτύου και τον τρόπο σύνδεσής του με άλλους κόμβους. Μια πρώτη μέτρηση της κεντρικότητας είναι ο βαθμός κόμβου ο οποίος ορίζεται ως ο αριθμός των συνδέσεών του με τους γειτονικούς του κόμβους. Είναι πράγματι ένα σημαντικό μέτρο για το πόσο εύκολα ένας κόμβος μπορεί να διαδώσει την πληροφορία, αλλά η συνολική διάδοση προς το πλήθος των κόμβων απαιτεί μάλλον μια γενίκευση της έννοιας. Στο επόμενο σχήμα φαίνεται παραστατικά η έννοια της κεντρικότητας όπου στη συνέχεια θα την ορίσουμε ακριβώς. Αυτό που διακρίνει την κεντρικότητα από το βαθμό κόμβου είναι η δυνατότητα ενός κόμβου να έχει μεγαλύτερη επιρροή σε άλλους σημαντικούς σε σύγκριση με έναν άλλο κόμβο ο οποίος έχει τον ίδιο βαθμό με τον πρώτο [11].



Σχήμα 2.2.1 : ο κόμβος B είναι πιο κεντρικός από τον A

Όπως φαίνεται στο σχήμα 2.2.1 ο κόμβος B είναι πιο κεντρικός στο δίκτυο από τον A παρόλο που και οι δύο έχουν βαθμό 5. Μια άλλη όψη της κεντρικότητας είναι και η διαμεσότητα (betweenness). Στην περίπτωση αυτή εξετάζουμε το βαθμό ελέγχου ενός κόμβου i στη διάδοση της πληροφορίας μεταξύ όλων των άλλων (εκτός του i) όταν η πληροφορία διαδίδεται μέσω του i [13]. Αν και αρχικά η έννοια αυτή είχε οριστεί μέσω της έννοιας των συντομότερων μονοπατιών, μπορούμε να

γενικεύσουμε την έκφραση που μπορεί να πάρει, αναγνωρίζοντας ότι οι συντομότερες διαδρομές δεν είναι πάντα οι πιο ρεαλιστικές, αλλά μπορούμε να θεωρήσουμε και άλλες υποθέσεις [13]. Σχηματικά η έννοια αυτή παρουσιάζεται στο παρακάτω σχήμα.



Σχήμα 2.2.2 : Οι κόμβοι A και B έχουν μεγάλη διαμεσότητα σε σχέση με τον κόμβο C

Μια πρώτη μέτρηση της κεντρικότητας, εκτός του βαθμού κόμβου, είναι αυτή του ιδιοδιανύσματος. Αν ένας κόμβος έχει μεγάλη επιρροή οι κόμβοι που συνδέονται σε αυτόν θα δανειστούν από αυτή σε μεγαλύτερο βαθμό από τυχόν άλλες συνδέσεις μικρότερης επιρροής. Έστω δίκτυο n κόμβων και λ θετική σταθερά. Συμβολίζοντας την κεντρικότητα κόμβου i ως x_i θεωρούμε αυτή ως ανάλογη προς τις κεντρικότητες των γειτόνων του ως εξής:

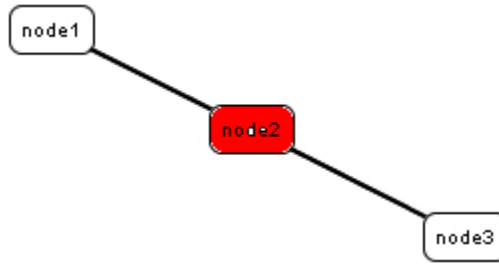
$$x_i = \frac{1}{\lambda} \sum_{j=1}^n A_{ij} x_j \quad (2.2.1)$$

Όπου ο συντελεστής A_{ij} ισούται με 1 αν υπάρχει σύνδεση μεταξύ i και j και με 0 διαφορετικά. Ορίζοντας το διάνυσμα κεντρικότητας $\mathbf{x} = (x_1, x_2, \dots, x_n)$ και τον πίνακα γειτνίασης $[A]$ τα στοιχεία του οποίου ορίζονται ως A_{ij} η εξίσωση 2.2.1 μπορεί να γραφεί ως εξίσωση πινάκων οπότε αναγνωρίζουμε ότι το \mathbf{x} είναι ένα ιδιοδιάνυσμα του πίνακα γειτνίασης και λ η αντίστοιχη ιδιοτιμή.

$$\mathbf{Ax} = \lambda \mathbf{x} \quad (2.2.2)$$

Θεωρώντας ότι επιθυμούμε να έχουμε θετικές κεντρικότητες, μπορεί να δειχθεί (μέσω του θεωρήματος Perron–Frobenius) ότι το λ πρέπει να είναι η μεγαλύτερη ιδιοτιμή και \mathbf{x} το αντίστοιχο ιδιοδιάνυσμα. Η κεντρικότητα ιδιοδιανύσματος όπως ορίστηκε με αυτό τον τρόπο αντιστοιχίζει μια τιμή σε κάθε κόμβο η οποία εξαρτάται από το πλήθος των συνδέσεων καθώς και την ποιότητά τους σε σχέση με τον κεντρικό ρόλο τους στο δίκτυο.

Θα δούμε τώρα εν συντομία την προσέγγιση όπως παρουσιάζεται από τον Akram [11] για τη χρήση των ιδιοδιανυσμάτων και το ρόλο της κεντρικότητας. Σαν ένα παράδειγμα παρουσίασης της κεντρικότητας υποθέτουμε ένα πολύ απλό δίκτυο με 3 κόμβους όπως στο σχήμα :



Σχήμα 2.2.3 Απλό δίκτυο 3 κόμβων

Ο πίνακας γειτνίασης είναι: $A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ Εφαρμόζοντας την εξίσωση (2.2.2) στη μορφή

$$(A - \lambda I)\mathbf{x} = 0 \text{ στον } A \text{ και αναγνωρίζοντας ότι } |A - \lambda I| = 0 \text{ παίρνουμε } \begin{vmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{vmatrix} = 0$$

Από την τελευταία εξίσωση έχουμε τις ιδιοτιμές $(\lambda) = (0, \sqrt{2}, -\sqrt{2})$. Ενδιαφερόμαστε να βρούμε το κυρίως ιδιοδιάνυσμα, αυτό που αντιστοιχεί στην μεγαλύτερη τιμή του λ , $\sqrt{2}$. Η εξίσωση (2.2.2)

$$\text{γράφεται τότε } \begin{bmatrix} -\sqrt{2} & 1 & 0 \\ 1 & -\sqrt{2} & 1 \\ 0 & 1 & -\sqrt{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \text{ από όπου έχουμε το κυρίως ιδιοδιάνυσμα } \mathbf{x} = \begin{bmatrix} 1 \\ \sqrt{2} \\ 1 \end{bmatrix}$$

Η εξίσωση αυτή αντιπροσωπεύει το κυρίως ιδιοδιάνυσμα του δικτύου μας και οι τιμές τις δείχνουν ότι ο πιο κεντρικός κόμβος είναι ο 2. Στη συνέχεια υπολογίζουμε το άθροισμα των στοιχείων του κυρίως ιδιοδιανύσματος και διαιρούμε κάθε στοιχείο του με αυτή την τιμή. Έτσι προκύπτει ένα ιδιοδιάνυσμα του οποίου η i -οστή τιμή θεωρείται ως η πιθανότητα του ρυθμού μόλυνσης του i -οστού κόμβου.

Επιλέγοντας την μεγαλύτερη από αυτές τις τιμές, έστω τ , τη θεωρούμε ως τον ρυθμό διάδοσης για το μοντέλο S.I. Το μοντέλο αυτό, απλό και εύκολα επιλύσιμο, είναι χρήσιμο κατά την ανάλυση κυρίως σε αρχικά στάδια μιας επιδημίας. Στην συγκεκριμένη ανάλυση ξεκινώντας από ένα μολυσμένο κόμβο καταλήγουμε στην εξίσωση $I(t) = \frac{N-1}{1+(N-2)e^{-\tau(N-1)t}}$ όπου από τους σταθερούς σε πλήθος συνολικά κόμβους N , οι I ερχόμενοι σε επαφή με την επιδημία μεταβάλλονται με το χρόνο όπως δείχνει η προηγούμενη εξίσωση.

2.3 Ιδιότητες δικτύων απουσία κλίμακας και επιδημίες

Στην υπό-ενότητα αυτή θα εξεταστεί πώς η ιδιότητα απουσίας κλίμακας των δικτύων επιδρά στην δυναμική επιδημίας, όπως παρουσιάζεται στην αναφορά [18]. Συγκεκριμένα θα δούμε εδώ την διαισθητικά αναμενόμενη μείωση του κατώφλιου επιδημίας με αύξηση του μεγέθους δικτύου. Το ίσως λιγότερο αναμενόμενο αποτέλεσμα είναι ότι συγκρίνοντας το κατώφλι σε ένα ανομοιογενές δίκτυο με το αντίστοιχο ενός ομογενούς θα βρούμε ότι το πρώτο παραμένει πάντα πιο χαμηλά.

Στην μελέτη των τοπολογικών ιδιοτήτων δικτύων έχει παρουσιαστεί ένας μεγάλος αριθμός από αυτά όπου παρουσιάζεται το φαινόμενο “μικρού κόσμου”, το οποίο σχετίζεται με το πολύ μικρό μέσο μονοπάτι μεταξύ κόμβων. Πιο εντυπωσιακά, σε μερικές περιπτώσεις αυτή η ιδιότητα συνδέεται με την κατανομή απουσίας κλίμακας $P(k) \sim k^{-2-\gamma}$, με $0 < \gamma \leq 1$ όπου k είναι ο αριθμός των ζεύξεων ενός κόμβου. Αυτή η φύση της απουσίας κλίμακας είναι συνδεδεμένη με μια μεγάλη ανομοιομορφία στις ιδιότητες συνδεσιμότητας ενός συστήματος. Εφόσον η δεύτερη ροπή της κατανομής συνδεσιμότητας $\langle k^2 \rangle$ αποκλίνει καθώς αυξάνεται το μέγεθος δικτύου, οι διακυμάνσεις συνδεσιμότητας σε ένα δίκτυο απουσίας κλίμακας δεν έχουν ένα εγγενές όριο και αποκλίνουν στο όριο καθώς το μέγεθος του συστήματος αυξάνει απεριόριστα. Ιδιότητες απουσίας κλίμακας έχουν παρατηρηθεί στο internet, στον ιστό και αλλού.

Εφόσον τα παραπάνω δίκτυα είναι απουσία κλίμακας, η μελέτη των επιδημιών και της δυναμικής διάδοσης αυτών είναι ένα σχετικό θεωρητικό θέμα. Σε ανομοιογενή δίκτυα είναι γνωστό ότι το κατώφλι επιδημίας μειώνεται με την τυπική απόκλιση της κατανομής συνδεσιμότητας και το χαρακτηριστικό αυτό ενισχύεται σε δίκτυα απουσίας κλίμακας όπου έχουν αποκλίνουσες διακυμάνσεις στο όριο που το μέγεθός τους αυξάνει απεριόριστα. Όντως, όπως έχει βρεθεί ήδη σε μη πεπερασμένα δίκτυα απουσία κλίμακας, η διαδικασία επιδημίας δεν παράγει ένα κατώφλι επιδημίας κάτω από αυτό που ένα ξέσπασμά της αποκτά ένα σημαντικό μέγεθος. Η απουσία ενός εγγενούς κατώφλιου επιδημίας έχει βρεθεί στο μοντέλο SIS καθώς και στο SIR σε μη πεπερασμένα δίκτυα. Οι τακτικές ανοσοποίησης επηρεάζονται κατά πολύ από τη φύση των απουσία κλίμακας κατανομών συνδεσιμότητας.

Όπως συνήθως συμβαίνει σε στατιστικά συστήματα μη ισορροπίας, έχει επίσης φανεί και σε πεπερασμένα συστήματα να παρακινείται ένα πεπερασμένο κατώφλι επιδημίας από πεπερασμένου μεγέθους αποτελέσματα. Πραγματικά συστήματα κατασκευάζονται από πεπερασμένο αριθμό ατόμων, γεγονός το οποίο είναι μακριά από το θερμοδυναμικό όριο. Αυτός ο πεπερασμένος πληθυσμός εισάγει μια μέγιστη συνδεσιμότητα k_c εξαρτημένη από το N , η οποία έχει το αποτέλεσμα να αποθηκεύει ένα όριο στις διακυμάνσεις συνδεσιμότητας, προκαλώντας με τον τρόπο αυτό ένα

μη μηδενικό κατώφλι. Πιο γενικά, μπορούμε να θεωρήσουμε μια κλάση δικτύων απουσία κλίμακας φραγμένων στα οποία η κατανομή συνδεσιμότητας έχει τη μορφή $P(k) \sim k^{-\rho-\gamma} f(\frac{k}{k_c})$, όπου η συνάρτηση $f(x)$ μειώνεται πολύ γρήγορα για $x > 1$. Θα αναφερόμαστε στην τιμή k_c ως αποκοπή συνδεσιμότητας με την έννοια ότι για $k > k_c$ η κατανομή φαίνεται σαν να μηδενίζεται. Η αποκοπή k_c μπορεί να προκύπτει λόγω του περιορισμένου μεγέθους του δικτύου ή λόγω περιορισμών στη προσθήκη επιπλέον δεσμών σε ένα κατά τα άλλα μη πεπερασμένο δίκτυο.

2.3.1 Κατώφλι επιδημίας στο μοντέλο SIS

Αρχικά θα παρουσιάσουμε μια αναλυτική μελέτη του μοντέλου SIS σε φραγμένα δίκτυα απουσία κλίμακας με έναν γενικό εκθέτη γ ($0 < \gamma \leq 1$) εστιάζοντας στα φαινόμενα που εισάγονται από μια πεπερασμένη αποκοπή k_c . Αναλύουμε την περίπτωση μιας έντονης αποκοπής $f(x) = \theta(1 - x)$, όπου θ είναι η συνάρτηση βήματος Heaviside, όπως συμβαίνει σε δίκτυα που αυξάνονται με έναν πεπερασμένο αριθμό στοιχείων. Θεωρούμε επίσης την περίπτωση μιας πιο ήπιας αποκοπής $f(x) = e^{-x}$, όπως συχνά υπάρχει σε συστήματα με φυσικούς περιορισμούς. Βρίσκουμε την συμπεριφορά του κατωφλίου επιδημίας ως συνάρτηση της αποκοπής συνδεσιμότητας και του μεγέθους δικτύου N και προκύπτει ότι ακόμα και για σχετικά μικρά δίκτυα το προκαλούμενο κατώφλι είναι αρκετά μικρότερο από το αντίστοιχο σε ομογενή δίκτυα. Αυτό επιβεβαιώνει ότι η φύση της απουσίας κλίμακας δεν μπορεί να αγνοηθεί σε πρακτικές εκτιμήσεις κατωφλίων επιδημίας ή θεραπείας σε πραγματικά δίκτυα. Θα δείξουμε επίσης την ακριβή αναλυτική μορφή εξάπλωσης επιδημίας (πυκνότητα μολυσμένων κόμβων) σε φραγμένα δίκτυα απουσία κλίμακας (BSF). Τα αποτελέσματα μπορούν να επεκταθούν άμεσα και στο μοντέλο SIR. Στη συνέχεια θα αναφερόμαστε στα δίκτυα απουσία κλίμακας ως SF δίκτυα.

Με σκοπό να εκτιμήσουμε την επίδραση του k_c σε επιδημίες σε BSF δίκτυα θα εξετάσουμε το καθιερωμένο μοντέλο SIS. Το μοντέλο βασίζεται στην αδρή περιγραφή κατά άτομο σε έναν πληθυσμό. Κάθε κόμβος του γράφου παριστάνει έναν κόμβο και κάθε ακμή μια σύνδεση δια μέσου της οποίας μπορεί να διαδοθεί μια ασθένεια. Κάθε υγιής κόμβος μολύνεται με ρυθμό ν αν είναι συνδεδεμένος σε έναν ή περισσότερους μολυσμένους κόμβους. Μολυσμένοι κόμβοι θεραπεύονται και γίνονται ξανά ευάλωτοι με ρυθμό δ , οπότε ορίζουμε έναν ενεργό ρυθμό διάδοσης $\lambda = \nu/\delta$ (χωρίς βλάβη της γενικότητας μπορούμε να θεωρήσουμε $\delta=1$). Το μοντέλο SIS δεν λαμβάνει υπόψη την δυνατότητα της απομόνωσης των κόμβων λόγω θανάτου ή ανοσίας και οι κόμβοι διατρέχουν τον κύκλο ευάλωτων – μολυσμένων - ευάλωτων. Αυτό το μοντέλο χρησιμοποιείται γενικά για να περιγράψει ασθένειες που οδηγούν σε καταστάσεις με στάσιμο μέσο όρο μολυσμένων ατόμων. Στην περίπτωση των δικτύων υπολογιστών, internet και ιστού που ενδιαφερόμαστε εδώ, μπορούμε

να δούμε την ανάλυση να αναφέρεται σε περιπτώσεις που όντως ισχύει ένα τέτοιο μοντέλο. Το μοντέλο βρίσκεται ανάμεσα στις περιπτώσεις των SI και SIR όπου η πρώτη βρίσκει εφαρμογή στα αρχικά στάδια της επιδημίας ενώ η δεύτερη στη φάση που έχει βρεθεί ο αντιός. Όπως αναφέρθηκε, τα γενικά αποτελέσματα ισχύουν και για το μοντέλο SIR το οποίο θα παρουσιαστεί στη συνέχεια. Με σκοπό να λάβουμε υπόψη μας την ανομοιογένεια των SF δικτύων, θεωρούμε τη σχετική πυκνότητα $\rho_k(t)$ των μολυσμένων ατόμων με δεδομένη τη συνδεσιμότητα k , δηλαδή την πιθανότητα ένας κόμβος με βαθμό k να είναι μολυσμένος. Οι μέσες δυναμικές πεδιακές εξισώσεις μπορούν να γραφούν ως :

$$\frac{d\rho_k(t)}{dt} = -\rho_k(t) + \lambda k p [1 - \rho_k(t)] \theta(\rho(t)) \quad (2.3.1)$$

Ο πρώτος όρος στο δεύτερο μέλος της εξίσωσης αυτής παριστάνει μολυσμένους κόμβους να γίνονται ευάλωτοι με μοναδιαίο ρυθμό. Ο δεύτερος όρος αντιπροσωπεύει τη μέση πυκνότητα των νέων μολυσμένων κόμβων που παράγονται, η οποία είναι ανάλογη του ρυθμού διάδοσης λ και της πιθανότητας ενός κόμβου με βαθμό k να είναι υγιής $[1 - \rho_k(t)]$ και να μολύνεται μέσω ενός συνδεδεμένου κόμβου. Ο ρυθμός του τελευταίου γεγονότος δίνεται από την πιθανότητα $\theta(\rho(t))$ μια δεδομένη σύνδεση να οδηγεί σε μολυσμένο κόμβο.

$$\theta(\rho(t)) = \langle k \rangle^{-1} \sum_k k P(k) \rho_k(t) \quad (2.3.2)$$

Από τις (2.3.1), (2.3.2) σε ισορροπία ($\frac{d\rho_k(t)}{dt} = 0$) παίρνουμε

$$\theta = \langle k \rangle^{-1} \sum_k k P(k) \frac{\lambda k \theta}{1 + \lambda k \theta} \quad (2.3.3)$$

όπου θ είναι μια συνάρτηση του λ μόνο. Η εξίσωση (2.3.3) επιτρέπει λύση με $\theta \neq 0$ και $\rho_k \neq 0$ μόνο αν ικανοποιείται η συνθήκη $\lambda \frac{\langle k^2 \rangle}{\langle k \rangle} \geq 1$. Η συνθήκη αυτή μπορεί να εξαχθεί αναγνωρίζοντας τα δύο μέλη της (2.3.3) ως συναρτήσεις του θ στο διάστημα $(0,1]$ και καταλήγοντας ότι $\frac{d}{d\theta} \left[\frac{1}{\langle k \rangle} \sum_k k P(k) \frac{\lambda k \theta}{1 + \lambda k \theta} \right]_{\theta=0} \geq 1$. Επομένως το κατώφλι επιδημίας ορίζεται από την ισότητα της συνθήκης :

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle} \quad (2.3.4)$$

Το αποτέλεσμα μας λέει ότι αν η τιμή του λ είναι πάνω από το κατώφλι $\lambda \geq \lambda_c$ η επιδημία εξαπλώνεται. Σε διαφορετική περίπτωση φθίνει εκθετικά [18], [20]. Το αποτέλεσμα αυτό δείχνει ότι σε ένα πεπερασμένο δίκτυο SF με $0 < \gamma \leq 1$, όπου $\langle k^2 \rangle \rightarrow \infty$ έχουμε $\lambda_c = 0$. Το γεγονός αυτό συνεπάγεται ότι για κάθε θετική τιμή του λ η μόλυνση μπορεί να διεισδύσει στο σύστημα με μια πεπερασμένη εξάπλωση σε ένα αρκετά μεγάλο δίκτυο. Ενώ αυτό ισχύει για μη πεπερασμένα

δίκτυα, μια πεπερασμένη τιμή της $\langle k^2 \rangle$ σε φραγμένα δίκτυα, ορίζει ένα ενεργό μη μηδενικό κατώφλι λόγω φαινομένων πεπερασμένου μεγέθους τα οποία συναντιούνται συνήθως σε μεταβάσεις μη ισορροπίας. Το κατώφλι αυτό παρόλα αυτά δεν είναι μια εγγενής ποσότητα όπως στα ομογενή συστήματα και χάνεται καθώς αυξάνεται το μέγεθος δικτύου ή η αποκοπή συνδεσιμότητας.

Με σκοπό να υπολογίσουμε τα ακριβή αποτελέσματα ενός πεπερασμένου k_c θεωρούμε δυο διαφορετικές περιπτώσεις αποκοπής συνδεσιμότητας. Σαν πρώτο παράδειγμα θεωρούμε μια ήπια εκθετική αποκοπή με χαρακτηριστική συνδεσιμότητα αποκοπής k_c . Αυτή η περίπτωση αντιστοιχεί σε αυτά τα πραγματικά δίκτυα όπου εξωτερικοί παράγοντες θέτουν ένα άνω όριο στη συνδεσιμότητα. Το δίκτυο μπορεί να έχει άπειρο αριθμό κόμβων αλλά η κατανομή συνδεσιμότητας του “νόμου δύναμης” φθίνει εκθετικά για μεγάλες τιμές του k . Με σκοπό να εκτελέσουμε ακριβείς υπολογισμούς χρησιμοποιούμε μια συνεχή προσέγγιση η οποία αντικαθιστά την συνδεσιμότητα από μια συνεχή μεταβλητή k στο εύρος $[m, \infty)$ όπου m είναι η ελάχιστη συνδεσιμότητα του δικτύου. Η κατανομή συνδεσιμότητας σε αυτή την περίπτωση είναι $P(k) = Ak^{-2-\gamma} e^{-\frac{k}{k_c}}$ όπου A είναι ένας παράγοντας κανονικοποίησης. Το ενεργό μη μηδενικό κατώφλι επιδημίας $\lambda_c(k_c)$ που προκύπτει από το εκθετικό κατώφλι είναι τότε λόγω της (2.3.4):

$$\lambda_c(k_c) = \frac{\int_m^\infty k^{-1-\gamma} e^{-k/k_c} dk}{\int_m^\infty k^{-\gamma} e^{-k/k_c} dk} \quad (2.3.5)$$

το οποίο έπειτα από ολοκλήρωση δίνει

$$\lambda_c(k_c) = k_c^{-1} \frac{\Gamma(-\gamma, \frac{m}{k_c})}{\Gamma(1-\gamma, \frac{m}{k_c})} \quad (2.3.6)$$

Όπου $\Gamma(x, y)$ είναι η μη ελλειπής συνάρτηση γάμμα. Για μεγάλο k_c μπορούμε να πάρουμε μια επέκταση Taylor και να κρατήσουμε τον πρώτο όρο, οπότε για κάθε $0 < \gamma < 1$ παίρνουμε

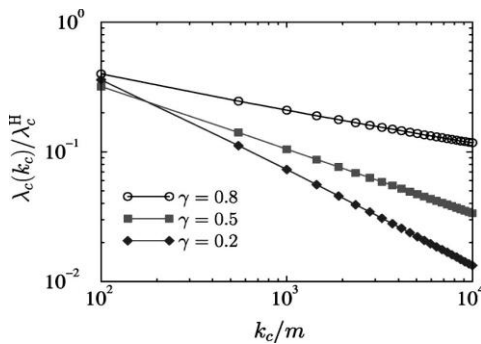
$$\lambda_c(k_c) \cong \frac{1}{m\gamma\Gamma(1-\gamma)} \left(\frac{k_c}{m}\right)^{\gamma-1} \quad (2.3.7)$$

Το όριο $\gamma \rightarrow 1$ στην εξίσωση (2.3.6) αντιστοιχεί σε μια λογαριθμική απόκλιση που οδηγεί σε πρώτη τάξη στο $\lambda_c(k_c) \cong [m \ln \left(\frac{k_c}{m}\right)]^{-1}$. Σε κάθε περίπτωση έχουμε ότι το κατώφλι επιδημίας μηδενίζεται όταν αυξάνεται η χαρακτηριστική αποκοπή (αποκοπή συνδεσιμότητας). Για μεγάλο k_c , η μέση συνδεσιμότητα είναι σταθερή και δίνεται από $\langle k \rangle = (\gamma + 1) \frac{m}{\gamma}$ για κάθε $\gamma > 0$. Είναι ενδιαφέρον να συγκρίνουμε το εγγενές κατώφλι που παρουσιάζουν τα ομογενή δίκτυα με αμελητέες διακυμάνσεις με το μη μηδενικό ενεργό κατώφλι των BFS δικτύων. Το εγγενές κατώφλι επιδημίας ενός ομογενούς δικτύου είναι $\lambda_c^H = \langle k \rangle^{-1}$ όπου $\langle k \rangle$ η σταθερή συνδεσιμότητα κόμβου. Συγκρίνοντας τα δύο

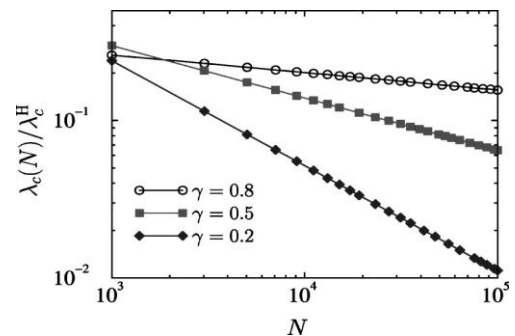
παραπάνω με την ίδια μέση συνδεσιμότητα $\langle k \rangle = (\gamma + 1) \frac{m}{\gamma}$ παίρνουμε ότι ο λόγος των κατωφλίων είναι :

$$\frac{\lambda_c(k_c)}{\lambda_c^H} = \frac{\gamma+1}{\gamma^2 \Gamma(1-\gamma)} \left(\frac{k_c}{m}\right)^{-1} \quad (2.3.8)$$

Αυτό δείχνει καθαρά ότι ακόμα και στην περίπτωση της αποκοπής συνδεσιμότητας το ενεργό κατώφλι επιδημίας ενός BSF δικτύου είναι πολύ μικρότερο από το εγγενές κατώφλι ενός συνηθισμένου. Στο σχήμα 1 έχουμε σχεδιάσει το λόγο που προκύπτει χρησιμοποιώντας την πλήρη έκφραση του $\lambda_c(k_c)$ από την εξίσωση (2.4.7)



Σχήμα 4.2.1 Λόγος μεταξύ του ενεργού κατωφλίου επιδημίας σε ένα BSF δίκτυο με κατώφλι συνδεσιμότητας εκθετικό, και του εγγενούς κατωφλίου επιδημίας ενός ομογενούς δικτύου με την ίδια μέση συνδεσιμότητα



Σχήμα 4.2.2 Λόγος μεταξύ του ενεργού κατωφλίου επιδημίας σε ένα BSF δίκτυο με πεπερασμένο μέγεθος N , και του εγγενούς κατωφλίου επιδημίας ενός ομογενούς δικτύου με ίδια μέση συνδεσιμότητα

Παρατηρούμε ότι ακόμα και για μικρές συνδεσιμότητες ($k_c = 10^2$ έως 10^3), για $\gamma=0,5$ περίπου, το ενεργό κατώφλι επιδημίας του BSF είναι μικρότερο κατά ένα παράγοντα κοντά στο $1/10$.

Ως δεύτερο παράδειγμα επίδρασης πεπερασμένου μεγέθους, θεωρούμε την παρουσία ενός απότομου κατωφλίου συνδεσιμότητας k_c . Εφόσον τα SF δίκτυα είναι συχνά δυναμικά αναπτυσσόμενα δίκτυα, η περίπτωση αυτή παριστάνει ένα δίκτυο που έχει αναπτυχθεί σε έναν πεπερασμένο αριθμό κόμβων. Η μέγιστη συνδεσιμότητα k_c του κάθε κόμβου σχετίζεται με την ηλικία του δικτύου μέσω της σχέσης

$$k_c \cong mN^{1/(1+\gamma)} \quad (2.3.9)$$

όπου m είναι η ελάχιστη συνδεσιμότητα του δικτύου. Σε αυτή την περίπτωση δεν μπορεί να υπάρχει κόμβος με βαθμό k μεγαλύτερο από k_c και μπορούμε να το θεωρήσουμε ως μια ισχυρή αποκοπή συνδεσιμότητας. Χρησιμοποιώντας ξανά την συνεχή προσέγγιση k , η κανονικοποιημένη συνδεσιμότητα έχει τώρα την μορφή:

$$P(k) = \frac{(1+\gamma)m^{1+\gamma}}{1-\left(\frac{k_c}{m}\right)^{-1-\gamma}} k^{-2-\gamma} \theta(k_c - k) \quad (2.3.10)$$

Οπού $\theta(x)$ η συνάρτηση βήματος Heaviside. Το πεπερασμένου μεγέθους κατώφλι είναι $\lambda_c(k_c)$ δίνεται από την έκφραση:

$$\lambda_c(k_c) = \frac{\int_m^\infty k^{-1-\gamma} dk}{\int_m^\infty k^{-\gamma} dk} \quad (2.3.11)$$

Σε πρώτη τάξη ως προς $\frac{k_c}{m}$ παίρνουμε :

$$\lambda_c(k_c) \cong \frac{1-\gamma}{\gamma m} \left(\frac{k_c}{m}\right)^{\gamma-1} \quad (2.3.12)$$

Στην περίπτωση αυτή μπορούμε από την εξίσωση (2.3.9) να το εκφράσουμε ως συνάρτηση του N οπότε παίρνουμε

$$\lambda_c(N) = \frac{1-\gamma}{\gamma m} N^{(\gamma-1)(\gamma+1)} \quad (2.3.13)$$

Η έκφραση αυτή ισχύει για $0 < \gamma < 1$, ενώ για $\gamma = 1$ λαμβάνουμε σε πρώτη τάξη την λογαριθμική συμπεριφορά $\lambda_c(N) \cong 2[m \ln(N)]^{-1}$. Επιπλέον έχουμε ότι το κατώφλι επιδημίας προσεγγίζει το μηδέν καθώς αυξάνεται το μέγεθος δικτύου και συγκρίνοντας όπως και στο πρώτο παράδειγμα με ένα ομογενές δίκτυο με ταυτόσημη μέση συνδεσιμότητα, καταλήγουμε σε ανάλογα αποτελέσματα για τον λόγο $\frac{\lambda_c(k_c)}{\lambda_c^H}$ ο οποίος για $N \cong 10^4$ και $\gamma = 0,5$ φανερώνει ότι στα SF δίκτυα το κατώφλι επιδημίας είναι μικρότερο κατά περίπου μια τάξη μεγέθους.

Με σκοπό να βρούμε την συμπεριφορά εξάπλωσης πρέπει να λύσουμε την εξίσωση (2.3.3) στη συνεχή περίπτωση

$$\theta = \langle k \rangle^{-1} \int_m^\infty k P(k) \frac{\lambda k \theta}{1 + \lambda k \theta} dk \quad (2.3.14)$$

και να χρησιμοποιήσουμε την τιμή του θ για να υπολογίσουμε την πυκνότητα ρ των μολυσμένων κόμβων ως

$$\rho = \sum_k \rho_k P(k) = \int_m^\infty P(k) \frac{\lambda k \theta}{1 + \lambda k \theta} dk \quad (2.3.15)$$

όπου το $P(k)$ δίνεται από την εξίσωση (2.3.10). Με απουσία κατωφλίου συνδεσιμότητας $k_c \rightarrow \infty$ και στο θερμοδυναμικό όριο $N \rightarrow \infty$ η εξάπλωση τείνει ως $\rho \sim \lambda^{1/(1-\gamma)}$ αν $0 < \gamma < 1$ και $\rho \sim e^{-\frac{1}{m\lambda}}$ αν $\gamma = 1$. Αντίστοιχα, με απουσία κατωφλίου επιδημίας η εξάπλωση είναι μηδενική μόνο αν ο ρυθμός

διάδοσης είναι $\lambda=0$. Στην περίπτωση μιας ισχυρής αποκοπής συνδεσιμότητας μπορούμε να ολοκληρώσουμε την (2.4.15) αγνοώντας τους όρους τάξης $(k_c/m)^{-\gamma}$ στην κατανομή $P(k)$ οπότε :

$$\theta \cong \gamma m^\gamma \lambda \theta \int_m^{k_c} \frac{k^{-\gamma}}{1+\lambda k \theta} dk = F(1, \gamma, 1 + \gamma, -[\lambda \theta m]^{-1}) - \left(\frac{k_c}{m}\right)^{-\gamma} F(1, \gamma, 1 + \gamma, -[\lambda \theta k_c]^{-1})$$
 όπου F είναι η υπεργεωμετρική συνάρτηση Gauss. Για σταθερό k_c μπορεί κανείς να επεκτείνει και τις δύο αυτές συναρτήσεις στο δεξί μέλος, διατηρώντας τους πιο σημαντικούς όρους της θ και θεωρώντας στη συνέχεια το όριο για μεγάλο k_c . Η τελική λύση για το θ δίνεται τότε σε πρώτη τάξη ως προς $\frac{k_c}{m}$ από:

$$\theta \cong \frac{1}{m \lambda^2} \frac{2-\gamma}{1-\gamma} \left(\frac{k_c}{m}\right)^{-1} \left[\lambda - \frac{1-\gamma}{\gamma m} \left(\frac{k_c}{m}\right)^{\gamma-1} \right] \quad (2.3.16)$$

Υπολογίζοντας το ολοκλήρωμα στην (2.3.15) και κρατώντας τον πρώτο όρο ως προς θ και k_c παίρνουμε τελικά την εξάπλωση επιδημίας ως: $\rho \cong \frac{\gamma+1}{\lambda \gamma} \frac{2-\gamma}{1-\gamma} \left(\frac{k_c}{m}\right)^{-1} \left[\lambda - \frac{1-\gamma}{\gamma m} \left(\frac{k_c}{m}\right)^{\gamma-1} \right]$. Εισάγοντας την σχέση κλίμακας ανάμεσα στην μέγιστη συνδεσιμότητα και το μέγεθος δικτύου N παίρνουμε την έκφραση :

$$\rho \sim N^{-\frac{1}{1+\gamma}} [\lambda - \lambda_c(N)] \quad (2.3.17)$$

Το αποτέλεσμα αυτό μας λέει ότι το πεπερασμένο μέγεθος δικτύου εισάγει μια τυπική μέση μετάβαση στο κατώφλι επιδημίας $\lambda_c(N)$ και δίνεται από την εξίσωση (2.3.13). Παρόλα αυτά η εξάπλωση συμπιέζεται από τον παράγοντα $N^{-\frac{1}{1+\gamma}}$ από την αντίστοιχη τιμή για ένα ομογενές δίκτυο. Οι παραπάνω υπολογισμοί μπορούν να επαναληφθούν σε παρόμοιες γραμμές και στην περίπτωση μιας ήπιας εκθετικής αποκοπής συνδεσιμότητας, καταλήγοντας σε παρόμοια αποτελέσματα.

Ανάλογα αποτελέσματα ισχύουν και για το μοντέλο SIR. Παρόλο που το μοντέλο αυτό αναφέρεται σε μόνιμη ανοσία και δεν επιτρέπει μόνιμη κατάσταση, το κατώφλι επιδημίας ως προς το οποίο ένα ξέσπασμα συμβαίνει έχει την ίδια αναλυτική μορφή $\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle}$

2.3.2 Κατώφλι επιδημίας στο μοντέλο SIR

Σε ένα ομογενές σύστημα το μοντέλο SIR μπορεί να περιγραφεί, κατά τα γνωστά, από τους πληθυσμούς $S(t)$, $\rho(t)$ και $R(t)$ των ευάλωτων, μολυσμένων και απομονωμένων ατόμων αντίστοιχα. Οι ποσότητες αυτές συνδέονται με την συνθήκη κανονικοποίησης

$$1=S(t)+\rho(t)+R(t) \quad (2.3.18)$$

και ικανοποιούν το σύστημα διαφορικών εξισώσεων:

$$\begin{aligned}\frac{dS}{dt} &= -\lambda\bar{k}\rho S \\ \frac{d\rho}{dt} &= -\mu R + \lambda\bar{k}\rho S \\ \frac{dR}{dt} &= \mu R\end{aligned}\tag{2.3.19}$$

Η ερμηνεία των εξισώσεων μας είναι γνωστή. Εδώ, λ είναι ο μικροσκοπικός ρυθμός διάδοσης της ασθένειας και \bar{k} είναι ο αριθμός συνδέσεων ανά μονάδα χρόνου, ο οποίος θεωρείται σταθερός για τον παραπάνω πληθυσμό. Τα παραπάνω είναι σε συμφωνία με την υπόθεση της ομοιογενούς ανάμειξης, η οποία ισχυρίζεται ότι η ισχύς της ασθένειας (ο καθαρός ρυθμός απόκτησης της ασθένειας από τους ευάλωτους) είναι ανάλογος της πυκνότητας των μολυσμένων. Η υπόθεση ομοιογενούς ανάμειξης είναι πράγματι ισοδύναμη με μια μέση περιγραφή πεδίου του μοντέλου στην οποία δεχόμαστε ότι ο ρυθμός επαφών μεταξύ μολυσμένων και ευάλωτων είναι σταθερός και ανεξάρτητος από κάθε πιθανή ετερογενή πηγή παρούσας στο σύστημα. Μια άλλη αυτονόητη υπόθεση του μοντέλου είναι ότι η χρονική κλίμακα της ασθένειας είναι αρκετά μικρότερη από την ζωή των ατόμων, οπότε δεν λαμβάνουμε υπόψη στις εξισώσεις όρους που αντιστοιχούν στη γέννηση ή το φυσικό θάνατο των ατόμων.

Η πιο σημαντική πρόβλεψη του μοντέλου είναι η παρουσία ενός μη μηδενικού κατωφλίου επιδημίας λ_c . Αν η τιμή του λ είναι μεγαλύτερη, $\lambda > \lambda_c$ η επιδημία εξαπλώνεται και μολύνει ένα πεπερασμένο κλάσμα του πληθυσμού. Αντίθετα, όταν $\lambda < \lambda_c$ ο συνολικός αριθμός μολυσμένων ατόμων $R_\infty = \lim_{t \rightarrow \infty} R(t)$ είναι μικρός στο όριο των πολύ μεγάλων πληθυσμών (θερμοδυναμικό όριο). Με σκοπό να δούμε το σημείο αυτό, θεωρούμε το σύστημα των εξισώσεων (2.3.19) στο οποίο χωρίς βλάβη της γενικότητας θέτουμε $\mu=1$. Ολοκληρώνοντας την εξίσωση για το $S(t)$ με αρχική συνθήκη $R(0)=0$ και $S(0) \cong 1$ (δηλαδή $\rho(0) \cong 0$, ένας πολύ μικρός αριθμός αρχικά μολυσμένων) παίρνουμε:

$$S(t) = e^{-\lambda\bar{k}R(t)}\tag{2.3.20}$$

Συνδυάζοντας την εξίσωση αυτή με την (2.3.18) παίρνουμε ότι ο συνολικός αριθμός απομονωμένων ατόμων R_∞ , ικανοποιεί την αναδρομική σχέση :

$$R_\infty = 1 - e^{-\lambda\bar{k}R_\infty}\tag{2.3.21}$$

Ενώ $R_\infty = 0$ είναι πάντα μια λύση της εξίσωσης αυτής, μια μη μηδενική λύση υπάρχει εφόσον ικανοποιείται η συνθήκη

$$\left. \frac{d}{dR_\infty} (1 - e^{-\lambda \bar{k} R_\infty}) \right|_{R_\infty=0} > 1 \quad (2.3.22)$$

Η συνθήκη αυτή είναι ισοδύναμη με τον περιορισμό $\lambda > \lambda_c$ όπου το κατώφλι επιδημίας παίρνει την τιμή $\lambda_c = (\bar{k})^{-1}$. Σημειώνουμε ότι η αναφορά μας για τους R_∞ ως τους απομονωμένους είναι ισοδύναμη με την διαπίστωση ότι R_∞ είναι και οι μολυσμένοι οι οποίοι απαραίτητα υπήρχαν πριν περάσουν στο στάδιο των απομονωμένων.

Το μοντέλο SIR σε σύνθετα δίκτυα

Με σκοπό να δούμε τις επιπτώσεις της ανομοιογένειας στην διάδοση επιδημίας, ας θεωρήσουμε το μοντέλο SIR σε ένα δίκτυο με γενική κατανομή συνδεσιμότητας $P(k)$ και πεπερασμένη μέση συνδεσιμότητα $\langle k \rangle = \sum_k kP(k)$. Σε κάθε βήμα χρόνου κάθε ευάλωτος κόμβος μολύνεται με πιθανότητα λ , αν συνδέεται με έναν ή περισσότερους μολυσμένους κόμβους και ταυτόχρονα κάθε μολυσμένος γίνεται απομονωμένος (μόνιμα υγιής) με πιθανότητα μ , την οποία χωρίς βλάβη της γενικότητας, θέτουμε ίση με μονάδα. Θεωρούμε τη χρονική εξέλιξη των μεγεθών $\rho_k(t), s_k(t)$ και $R_k(t)$ τα οποία είναι οι πυκνότητες των μολυσμένων, ευάλωτων και απομονωμένων κόμβων βαθμού k σε χρόνο t αντίστοιχα. Οι μεταβλητές αυτές συνδέονται με την συνθήκη κανονικοποίησης

$$\rho_k(t) + s_k(t) + R_k(t) = 1 \quad (2.3.23)$$

Καθολικές ποσότητες όπως το μέγεθος της επιδημίας εκφράζονται τότε από το μέσο όρο διαφορετικών κλάσεων συνδεσιμότητας. Στο επίπεδο μέσης τιμής οι πυκνότητες αυτές ικανοποιούν το ακόλουθο σύστημα διαφορικών εξισώσεων :

$$\frac{d\rho_k(t)}{dt} = -\rho_k(t) + \lambda k S_k(t) \theta(t) \quad (2.3.24)$$

$$\frac{ds_k(t)}{dt} = -\lambda k S_k(t) \theta(t) \quad (2.3.25)$$

$$\frac{dR_k(t)}{dt} = \rho_k(t) \quad (2.3.26)$$

Ο παράγοντας $\theta(t)$ παριστάνει την πιθανότητα μια δεδομένη ακμή να οδηγεί σε μολυσμένο κόμβο. Η πιθανότητα μια ακμή να οδηγεί σε έναν κόμβο με s ακμές είναι ανάλογη του $sP(s)$. Επομένως η πιθανότητα μια τυχαία επιλεγμένη ακμή να οδηγεί σε έναν μολυσμένο κόμβο δίνεται από :

$$\theta(t) = \frac{\sum_k k P(k) \rho_k(t)}{\sum_s s P(s)} = \frac{\sum_k k P(k) \rho_k(t)}{\langle k \rangle} \quad (2.3.27)$$

Σε αυτή την προσέγγιση αγνοούμε τις συσχετίσεις συνδεσιμότητας στο δίκτυο. Για παράδειγμα, η

πιθανότητα μια ακμή να οδηγεί σε έναν μολυσμένο κόμβο θεωρείται ανεξάρτητη της συνδεσιμότητας του κόμβου από τον οποίο ξεκινάει η ακμή. Μια πιο εκλεπτυσμένη προσέγγιση θα θεωρούσε τις συσχετίσεις δικτύου δεδομένες μέσω των υπό συνθήκη πιθανοτήτων $P(k/k')$ ενός κόμβου με δεδομένο k' να συνδέεται με κόμβο βαθμού k . Παρόλα αυτά, αυτή η μάλλον ακατέργαστη προσέγγιση είναι αρκετά ικανή να μας φανερώσει πολλές ιδιότητες όπως αυτές φαίνονται σε προσομοιώσεις του μοντέλου.

Οι εξισώσεις (2.3.24) έως και (2.4.26) σε συνδυασμό με τις αρχικές συνθήκες $R_k(t)=0$, $\rho_k(0)=\rho_k^0$ ($S_k(0)=1-\rho_k^0$) ορίζουν πλήρως το μοντέλο SIR σε κάθε σύνθετο δίκτυο με κατανομή συνδεσιμότητας $P(k)$. Θα θεωρήσουμε την ειδική περίπτωση μιας ομογενούς αρχικής κατανομής μολυσμένων κόμβων $\rho_k^0=\rho^0$. Σε αυτή την περίπτωση, στο όριο $\rho^0 \rightarrow 0$, μπορούμε να αντικαταστήσουμε $\rho_k \cong 0$ και $S_k(0) \cong 1$. Μέσω αυτής της προσέγγισης η εξίσωση (2.3.25) μπορεί άμεσα να ολοκληρωθεί καταλήγοντας

$$S_k(t) = e^{-\lambda k \varphi(t)} \quad (2.3.28)$$

Όπου έχουμε ορίσει την βοηθητική συνάρτηση

$$\varphi(t) = \int_0^t \theta(t') dt' = \frac{1}{\langle k \rangle} \sum_k k P(k) R_k(t) \quad (2.3.29)$$

και στην τελευταία εξίσωση κάναμε χρήση της (2.3.27)

Με σκοπό να πάρουμε μια κλειστή έκφραση της πυκνότητας των μολυσμένων ατόμων είναι πιο βολικό να εστιάσουμε στην χρονική εξέλιξη του φ . Είναι:

$$\frac{d\varphi}{dt} = \frac{1}{\langle k \rangle} \sum_k k P(k) \rho_k(t) \quad (2.3.30)$$

$$= \frac{1}{\langle k \rangle} \sum_k k P(k) (1 - \rho_k(t) - S_k(t)) \quad (2.3.31)$$

$$= 1 - \varphi(t) - \frac{1}{\langle k \rangle} \sum_k k P(k) S_k(t) \quad (2.3.32)$$

Εισάγοντας την (2.3.28) καταλήγουμε σε μια διαφορική εξίσωση για το $\varphi(t)$

$$\frac{d\varphi}{dt} = 1 - \varphi(t) - \frac{1}{\langle k \rangle} \sum_k k P(k) e^{-\lambda k \varphi(t)} \quad (2.3.33)$$

Έχοντας επιλύσει την (2.3.33) μπορούμε να πάρουμε το συνολικό μέγεθος επιδημίας R_∞ ως συνάρτηση του $\varphi_\infty = \lim_{t \rightarrow \infty} \varphi(t)$ εφόσον $R_k(\infty) = 1 - S_k(\infty)$ έχουμε

$$R_\infty = \sum_k P(k) (1 - e^{-\lambda k \varphi_\infty}) \quad (2.3.34)$$

Για μια γενική κατανομή $P(k)$, η εξίσωση (2.3.33) δεν μπορεί να λυθεί σε κλειστή μορφή. Παρόλα αυτά μπορούμε να πάρουμε χρήσιμη πληροφορία στο όριο χρόνου $t \rightarrow \infty$ στο τέλος της επιδημίας. Εφόσον $\rho_k(\infty) = 0$ και επομένως $\lim_{t \rightarrow \infty} \frac{d\varphi}{dt} = 0$, από την (2.3.33) έχουμε:

$$\varphi_\infty = 1 - \frac{1}{\langle k \rangle} \sum_k k P(k) e^{-\lambda k \varphi_\infty} \quad (2.3.35)$$

Η τιμή $\varphi_\infty = 0$ είναι μια λύση. Για να έχουμε μη μηδενική λύση πρέπει να ικανοποιείται η συνθήκη:

$$\left. \frac{d}{d\varphi_\infty} \left(1 - \frac{1}{\langle k \rangle} \sum_k k P(k) e^{-\lambda k \varphi_\infty} \right) \right|_{\varphi_\infty=0} > 1 \quad (2.3.36)$$

Από αυτό συνεπάγεται ότι

$$\frac{1}{\langle k \rangle} \sum_k k P(k) \lambda k = \lambda \frac{\langle k^2 \rangle}{\langle k \rangle} > 1 \quad (2.3.37)$$

Η συνθήκη αυτή ορίζει το κατώφλι επιδημίας

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle} \quad (2.3.38)$$

2.4 Διάδοση επιδημιών σε δίκτυα με τη χρήση γεννητριών συναρτήσεων

Στην υπό-ενότητα αυτή θα παρουσιάσουμε το στοχαστικό μοντέλο SIR του Newman [14] το οποίο κάνει χρήση γεννητριών συναρτήσεων.

Έστω ένα ζεύγος ατόμων τα οποία είναι συνδεδεμένα, ένα i από τους αρρώστους ικανό να μεταδώσει τον ιό και ένα j από τους ευάλωτους. Υποθέτουμε ότι ο μέσος ρυθμός μετάδοσης της επιδημίας είναι r_{ij} και το άρρωστο άτομο παραμένει άρρωστο για χρόνο τ_i . Τότε η πιθανότητα $1 - T_{ij}$ η επιδημία να μην μεταδοθεί από τον i στον j είναι:

$$1 - T_{ij} = \lim_{\delta t \rightarrow 0} (1 - r_{ij} \delta t)^{\frac{\tau_i}{\delta t}} = e^{-r_{ij} \tau_i} \quad (2.4.1)$$

και η πιθανότητα μετάδοσης είναι

$$T_{ij} = 1 - e^{-r_{ij} \tau_i} \quad (2.4.2)$$

Γενικά τα r_{ij} και τ_i διαφέρουν μεταξύ ατόμων, οπότε η πιθανότητα μετάβασης επίσης διαφέρει. Ας υποθέσουμε αρχικά ότι οι δύο αυτές ποσότητες είναι ανεξάρτητα και ταυτόσημα κατανομημένες τυχαίες μεταβλητές (α.τ.κ) επιλεγμένες από κάποιες κατάλληλες κατανομές $P(r)$ και $P(\tau)$ (Θα μετριάσουμε αυτή την υπόθεση αργότερα). Ο ρυθμός r_{ij} δεν χρειάζεται να είναι συμμετρικός. Η πιθανότητα μετάβασης μπορεί να μην είναι η ίδια σε κάθε κατεύθυνση. Εξάλλου η T_{ij} είναι γενικά μη συμμετρική εξαιτίας της παρουσίας του τ_i στην εξίσωση (2.4.2). Το τέχνασμα τώρα είναι το εξής: Επειδή οι r_{ij} , τ_i είναι (α.τ.κ), το ίδιο συμβαίνει και με την T_{ij} οπότε η a priori πιθανότητα μετάβασης της ασθένειας μεταξύ δυο ατόμων είναι απλά η μέση τιμή του T_{ij} ως προς τις κατανομές $P(r)$ και $P(\tau)$, η οποία είναι:

$$T = \langle T_{ij} \rangle = 1 - \int_0^\infty dr d\tau P(r)P(\tau)e^{-r\tau} \quad (2.4.3)$$

Καλούμε την T ικανότητα διάδοσης της ασθένειας. Είναι πάντα απαραίτητα στο διάστημα $0 \leq T \leq 1$. Επομένως το γεγονός ότι οι μεμονωμένες πιθανότητες διάδοσης διαφέρουν δεν έχει σημασία τελικά. Σε έναν πληθυσμό σαν σύνολο η ασθένεια θα εξαπλώνεται σαν όλες οι πιθανότητες διάδοσης να είναι ίσες με T . Είναι αυτό το αποτέλεσμα που κάνει το μοντέλο μας επιλύσιμο. Περιπτώσεις που οι μεταβλητές r_{ij} , τ_i δεν είναι (α.τ.κ), είναι δυσκολότερες αλλά όπως θα δείξουμε είναι μερικές φορές επίσης επιλύσιμες. Αναφέρουμε ότι πιο σύνθετα μοντέλα διάδοσης, όπως το SEIR, καλύπτονται επίσης από αυτό το φορμαλισμό. Η ικανότητα διάδοσης T είναι απλώς το ολοκλήρωμα πιθανότητας της μετάδοσης μεταξύ δυο ατόμων. Η ακριβής προσωρινή συμπεριφορά της νόσου και άλλων μεταβλητών δεν είναι σημαντικές. Πάντως το μοντέλο μπορεί να γενικευτεί ώστε να συμπεριλάβει κάθε προσωρινή διακύμανση στην ασθένεια και η διάδοση να εξακολουθεί να παριστάνεται με την μεταβλητή T όπως παραπάνω. Υποθέτουμε ότι παρακολουθούμε ένα ξέσπασμα επιδημίας η οποία αρχίζει με ένα μεμονωμένο άτομο το οποίο είναι ικανό να μεταδώσει την ασθένεια. Αν σημειώσουμε κάθε ακμή του γράφου διαμέσου του οποίου υπάρχει διάδοση, το οποίο συμβαίνει με πιθανότητα T , το τελικό μέγεθος της επιδημίας θα είναι ακριβώς το μέγεθος της ομάδας κόμβων στους οποίους μπορούμε να φτάσουμε από τον αρχικό κόμβο διασχίζοντας μόνο τις σημειούμενες ακμές. Επομένως το μοντέλο είναι ισοδύναμο με ένα γνωστό μοντέλο φιλτραρίσματος δεσμών με πιθανότητα κατάληψης δεσμού T στο γράφο που παριστάνει το δίκτυο.

Στο επόμενο τμήμα θα δείξουμε πώς το πρόβλημα του φιλτραρίσματος μπορεί να λυθεί σε τυχαίους γράφους με αυθαίρετη κατανομή βαθμού δίνοντας ακριβείς λύσεις για το τυπικό μέγεθος των ξεσπασμάτων επιδημίας, μέγεθος επιδημίας (αν υπάρχει) και έναν αριθμό άλλων ποσοτήτων ενδιαφέροντος.

Ένα από τα πιο σημαντικά αποτελέσματα στην εμπειρική δουλειά πάνω στα δίκτυα είναι ότι αυτά έχουν κατανομές βαθμών οι οποίες είναι έντονα δεξιά ετεροκλινείς. Με άλλα λόγια οι πιο πολλοί κόμβοι έχουν μικρό βαθμό, αλλά υπάρχει ένας μικρός αριθμός κόμβων που έχουν πολύ μεγάλο βαθμό. Είναι γνωστό ότι η παρουσία έντονα συνδεδεμένων κόμβων έχει δυσανάλογη επίδραση σε ορισμένες ιδιότητες ενός δικτύου. Επίσης έρευνες υποδεικνύουν ότι το ίδιο συμβαίνει στην εξάπλωση επιδημιών στα δίκτυα, οπότε είναι σημαντικό να εξετάσουμε μη τετριμμένες κατανομές βαθμών στα μοντέλα μας. Σαν πρώτη εικόνα της μεθόδου μας θα εξετάσουμε μια απλή κλάση ενιαίων γράφων των οποίων η κατανομή βαθμού είναι γνωστή αλλά κατά τα άλλα ο γράφος είναι αυθαίρετος.

Οι γράφοι αυτοί ορίζονται εύκολα. Ορίζουμε την κατανομή βαθμού δίνοντας τις κατάλληλα κανονικοποιημένες πιθανότητες P_k ένας κόμβος να έχει βαθμό k . Ένα σύνολο N βαθμών $\{k_i\}$, το οποίο καλείται επίσης ακολουθία βαθμού, επιλέγεται από αυτή την κατανομή και σε κάθε έναν από τους N κόμβους του γράφου δίνεται ο κατάλληλος αριθμός k_i των γειτόνων του, τέλη των ακμών που πηγάζουν από αυτόν. Ζεύγη από τέτοιους γείτονες στη συνέχεια επιλέγονται τυχαία και συνδέονται ώστε να σχηματιστούν οι συνολικές ακμές. Η διαδικασία συνεχίζεται μέχρι να μην απομείνει κανένας γείτονας. Αν κατά τύχη παραχθεί ένας περιττός αριθμός γειτόνων τότε εξαιρούμε ένα k_i και επιλέγουμε ένα άλλο μέχρι να επιτευχθεί ένας άρτιος αριθμός. Η τεχνική αυτή εγγυάται ότι ο γράφος που παράγεται επιλέγεται ομοιόμορφα τυχαία από το σύνολο όλων των γράφων με την επιλεγμένη ακολουθία βαθμού.

Για όλα τα αποτελέσματα που δίνονται στην ενότητα αυτή παίρνουμε το μέσο όρο ως προς την οικογένεια των πιθανών γράφων που μπορούν να παραχθούν όπως παραπάνω, στο όριο του μεγάλου μεγέθους γράφου.

2.4.1 Γεννήτριες Συναρτήσεις

Ορίζουμε την γεννήτρια συνάρτηση της κατανομής βαθμού ως

$$G_0(x) = \sum_{k=0}^{\infty} p_k x^k \quad (2.4.4)$$

Σημειώνουμε ότι $G_0(1) = \sum_k p_k = 1$ αν η p_k είναι μια σωστά κανονικοποιημένη κατανομή πιθανότητας. Η συνάρτηση αυτή περιλαμβάνει όλη την πληροφορία για την κατανομή βαθμού.

Δεδομένης αυτής, εύκολα ανακατασκευάζουμε την κατανομή με επαναλαμβανόμενη διαφόριση.

$$p_k = \frac{1}{k!} \frac{d^k G_0}{dx^k} \Big|_{x=0} \quad (2.4.5)$$

Είναι πιο εύκολο να εργαστούμε με την γεννήτρια συνάρτηση από την κατανομή απευθείας λόγω των παρακάτω κρίσιμων ιδιοτήτων :

Δυνάμεις: Αν η κατανομή μιας ιδιότητας K ενός αντικειμένου παράγεται από μια δεδομένη γεννήτρια συνάρτηση, τότε η κατανομή του αθροίσματος m ανεξάρτητων πραγματοποιήσεων της K παράγεται από την m -οστή δύναμη της γεννήτριας συνάρτησης. Για παράδειγμα αν επιλέξουμε m κόμβους στην τύχη από έναν μεγάλο γράφο τότε η κατανομή του αθροίσματος των βαθμών αυτών παράγεται από την $[G_0(x)]^m$.

Ροπές: Η μέση τιμή της κατανομής πιθανότητας που παράγεται από μια γεννήτρια συνάρτηση δίνεται από την πρώτη παράγωγο της συνάρτησης υπολογισμένη στην τιμή 1. Για παράδειγμα η μέση τιμή z ενός κόμβου στο δίκτυο μας δίνεται από:

$$z = \langle k \rangle = \sum_k k p_k = G_0'(1) \quad (2.4.6)$$

Ανώτερες ροπές της κατανομής μπορούν να υπολογιστούν από ανώτερες παραγώγους επίσης. Γενικά έχουμε:

$$\langle k^n \rangle = \sum_k k^n p_k = \left[\left(x \frac{d}{dx} \right)^n G_0(x) \right]_{x=1} \quad (2.4.7)$$

Μια επιπλέον παρατήρηση που θα αποδειχθεί σημαντική είναι η ακόλουθη. Ενώ η G_0 παραπάνω παράγει σωστά την κατανομή βαθμών τυχαία επιλεγμένων κόμβων στο γράφο μας, μια διαφορετική γεννήτρια συνάρτηση χρειάζεται για την κατανομή των βαθμών κόμβων στους οποίους φτάνουμε ακολουθώντας μια τυχαία επιλεγμένη ακμή. Αν ακολουθήσουμε μια ακμή προς έναν κόμβο στο ένα άκρο της, τότε ο κόμβος είναι πιο πιθανό να είναι υψηλότερου βαθμού από έναν τυχαίο κόμβο, εφόσον οι υψηλού βαθμού κόμβοι έχουν πιο πολλές ακμές συνδεδεμένες από αυτούς που έχουν χαμηλό βαθμό. Η κατανομή των βαθμών που φτάνουμε ακολουθώντας μια ακμή είναι ανάλογη του p_k , και έτσι η γεννήτρια συνάρτηση αυτών είναι

$$\frac{\sum_k k p_k x^k}{\sum_k k p_k} = x \frac{G_0'(x)}{G_0'(1)} \quad (2.4.8)$$

Γενικά θα ενδιαφερθούμε για διάφορους τρόπους που μπορούμε να απομονώσουμε έναν τέτοιο κόμβο εξαιρώντας την ακμή που ακολουθήσαμε το οποίο είναι ο βαθμός πλην 1. Για να επιτύχουμε αυτό απλά διαιρούμε την παραπάνω συνάρτηση με την πρώτη δύναμη του x , οπότε καταλήγουμε σε μια νέα γεννήτρια συνάρτηση

$$G_1(x) = \frac{G_0'(x)}{G_0'(1)} = \frac{1}{z} G_0'(x) \quad (2.4.9)$$

όπου z είναι η μέση τιμή βαθμού κόμβου όπως προηγουμένως.

Με σκοπό να λύσουμε το πρόβλημα φιλτραρίσματος θα χρειαστούμε γεννήτριες συναρτήσεις $G_0(x,T)$ και $G_1(x,T)$ για την κατανομή του αριθμού των κατεχόμενων ακμών που είναι συνδεδεμένες σε ένα κόμβο, ως μια συνάρτηση της πιθανότητας μετάβασης. Αυτές είναι εύκολο να προκύψουν. Η πιθανότητα ενός κόμβου να έχει m ακμές κατεχόμενες από τις k που αναδύονται από αυτόν δίνεται από την διωνυμική κατανομή $\binom{k}{m}T^m(1-T)^{k-m}$, οπότε η κατανομή πιθανότητας του m παράγεται από την:

$$G_0(x,T) = \sum_{m=0}^{\infty} \sum_{k=m}^{\infty} p_k \binom{k}{m} T^m (1-T)^{k-m} x^m = \sum_{k=0}^{\infty} p_k \sum_{m=0}^k \binom{k}{m} (xT)^m (1-T)^{k-m} = \sum_{k=0}^{\infty} p_k (1-T + xT)^k = G_0(1 + (x-1)T) \quad (2.4.10)$$

Ομοίως, η κατανομή πιθανότητας των κατειλημμένων ακμών από ένα κόμβο στον οποίο φθάνουμε ακολουθώντας τυχαία μια ακμή παράγεται από την:

$$G_1(x,T) = G_1(1 + (x-1)T) \quad (2.4.11)$$

Σημειώνουμε ότι σύμφωνα με τον συμβολισμό μας,

$$G_0(x,1) = G_0(x) \quad (2.4.12a)$$

$$G_0(1,T) = G_0(1) \quad (2.4.12b)$$

$$G'_0(1,T) = TG'_0(1) \quad (2.4.12c)$$

Και ομοίως για την G_1 . ($G'_0(x,T)$ αναπαριστά εδώ την παράγωγο του $G_0(x,T)$ ως προς το πρώτο όρισμα του)

2.4.2 Κατανομή Μεγέθους Ξεσπάσματος

Η πρώτη ποσότητα με την οποία θα ασχοληθούμε είναι η κατανομή $P_s(T)$ των μεγεθών ξεσπάσματος της επιδημίας στο δίκτυο, η οποία είναι επίσης η κατανομή αυτών των κόμβων οι οποίοι συνδέονται μεταξύ τους από τις κατειλημμένες ακμές στο αντίστοιχο μοντέλο φιλτραρίσματος. Έστω $H_0(x,T)$ η γεννήτρια συνάρτηση για αυτή την κατανομή

$$H_0(x,T) = \sum_{s=0}^{\infty} P_s(T) x^s \quad (2.4.13)$$

Σε αναλογία με το προηγούμενο τμήμα, ορίζουμε επίσης την $H_1(x,T)$ ως την γεννήτρια συνάρτηση των συνδεδεμένων κόμβων που φθάνουμε ακολουθώντας μια τυχαία ακμή. Η H_1 μπορεί να εκφραστεί ως ένα άθροισμα όρων όπως παρακάτω. Η ομάδα που μπορούμε να φτάσουμε ακολουθώντας μια ακμή μπορεί να είναι: (1) Ένας απλός κόμβος χωρίς ακμές συνδεδεμένες εκτός

αυτής που ακολουθήσαμε προς αυτόν. (2) Ένας απλός κόμβος που συνδέεται σε $m \geq 1$ ακμές, εκτός αυτής που ακολουθήσαμε, όπου η καθεμία οδηγεί σε μια άλλη ομάδα της οποίας η κατανομή μεγέθους επίσης παράγεται από την H_1 . Επιπλέον παρατηρούμε ότι η δυνατότητα δυο πεπερασμένων ομάδων οι οποίες είναι συνδεδεμένες στον ίδιο κόμβο να έχουν μια ακμή που να τις συνδέει απευθείας φθίνει ως N^{-1} ως προς το μέγεθος του γράφου, οπότε γίνεται μηδέν στο όριο $N \rightarrow \infty$. Με άλλα λόγια, δεν υπάρχουν βρόχοι στις ομάδες μας. Η δομή είναι εντελώς δενδρική. Χρησιμοποιώντας τα αποτελέσματα αυτά, μπορούμε να εκφράσουμε την $H_1(x,T)$ ως

$$H_1(x, T) = xG_1(H_1(x, T), T) \quad (2.4.14)$$

Τότε το μέγεθος της ομάδας που μπορούμε να φθάσουμε από έναν τυχαίο αρχικό κόμβο κατανέμεται σύμφωνα με

$$H_0(x, T) = xG_0(H_1(x, T), T) \quad (2.4.15)$$

Οι εξισώσεις αυτές παρέχουν την λύση για την πιο γενική περίπτωση πιθανότητας μετάβασης που εφαρμόζεται στα SIR μοντέλα. Με δεδομένη την $H_0(x,T)$ μπορούμε να αποσπάσουμε την κατανομή ξεσπάσματος $P_s(T)$ μέσω επαναλαμβανόμενης διαφορίσης της (2.4.13) στο $x=0$. Παρ' όλα αυτά στις πιο πολλές περιπτώσεις δεν είναι δυνατό να βρούμε αυθαίρετες παραγώγους της H_0 σε κλειστή μορφή. Απεναντίας εκτιμάμε αυτές αριθμητικά. Εφόσον η άμεση αριθμητική εκτίμηση παραγώγων είναι επιρρεπής σε προβλήματα ακριβείας των υπολογιστών, συνιστάται η εκτίμηση των παραγώγων με ολοκλήρωση χρησιμοποιώντας τον τύπο του Cauchy

$$P_s(T) = \frac{1}{s!} \frac{d^s H_0}{dx^s} \Big|_{x=0} = \frac{1}{2\pi i} \oint \frac{H_0(\zeta, T)}{\zeta^{s+1}} d\zeta \quad (2.4.16)$$

όπου το ολοκλήρωμα είναι ως προς το μοναδιαίο κύκλο.

Είναι δυνατόν να βρούμε τις πρώτες χίλιες παραγώγους μιας συνάρτησης χωρίς δυσκολία με τη μέθοδο αυτή [16]. Επομένως μπορούμε να βρούμε την ακριβή πιθανότητα P_s για ένα δεδομένο ξέσπασμα επιδημίας να μεταδοθεί αυτό σε s άτομα, ως συνάρτηση της πιθανότητας μετάβασης.

2.4.3 Μεγέθη Ξεσπάσματος και η Μετάβαση Επιδημίας

Παρόλο που γενικά πρέπει να χρησιμοποιήσουμε αριθμητικές μεθόδους για να βρούμε τη συνολική κατανομή P_s , μπορούμε να βρούμε το μέσο όρο της σε κλειστή μορφή. Χρησιμοποιώντας την εξίσωση (2.4.6) έχουμε

$$\langle s \rangle = H'_0(1, T) = 1 + G'_0(1, T)H'_1(1, T) \quad (2.4.17)$$

όπου έχουμε κάνει χρήση του γεγονότος ότι οι γεννήτριες συναρτήσεις είναι ίσες με 1 στο $x=1$ αν οι κατανομές τις οποίες παράγουν είναι κατάλληλα κανονικοποιημένες. Διαφορίζοντας την εξίσωση (2.4.14) έχουμε $H_1'(1, T) = 1 + G_1'(1, T)H_1'(1, T) = \frac{1}{1-G_1'(1, T)}$ (2.4.18)

Οπότε

$$\langle s \rangle = 1 + \frac{G_0'(1, T)}{1-G_1'(1, T)} = 1 + \frac{T G_0'(1)}{1-T G_1'(1)} \quad (2.4.19)$$

Δεδομένων των εξισώσεων (2.4.4), (2.4.9), (2.4.10) και (2.4.11) μπορούμε να υπολογίσουμε τότε το μέσο όρο του ξεσπάσματος για κάθε τιμή του T και κατανομή βαθμού.

Σημειώνουμε ότι η εξίσωση (2.4.19) αποκλίνει όταν $T G_1'(1) = 1$. Το σημείο αυτό δείχνει την εισβολή μιας επιδημίας. Είναι το σημείο που ένα τυπικό ξέσπασμα σταματά να είναι περιορισμένο σε έναν πεπερασμένο αριθμό ατόμων και εξαπλώνεται σε ένα εκτεταμένο κλάσμα του γράφου. Η μετάβαση πραγματοποιείται όταν το T είναι ίσο με την κρίσιμη τιμή T_c ,

$$T_c = \frac{1}{G_1'(1)} = \frac{G_0'(1)}{G_0''(1)} = \frac{\sum_k k p_k}{\sum_k k(k-1)p_k} \quad (2.4.20)$$

Για $T > T_c$ έχουμε μια επιδημία. Μπορούμε να υπολογίσουμε το μέγεθος της επιδημίας αυτής ως εξής: Πάνω από το κατώφλι της επιδημίας η εξίσωση (2.4.14) δεν ισχύει επειδή μπορεί να περιέχονται βρόχοι, γεγονός το οποίο καταστρέφει τις υποθέσεις στις οποίες η εξίσωση βασίστηκε. Η εξίσωση είναι σε ισχύ παρόλα αυτά, αν ορίσουμε ξανά την H_0 ως τη γεννήτρια συνάρτηση για ξεσπάσματα εκτός επιδημίας, δηλαδή απομονωμένες ομάδες κόμβων οι οποίες δεν καλύπτουν ολόκληρο το γράφο αλλά μόνο το κομμάτι που δεν επηρεάζεται από την επιδημία. Επομένως πάνω από την μετάβαση επιδημίας έχουμε:

$$H_0(1, T) = \sum_S P_S = 1 - s(T) \quad (2.4.21)$$

Όπου $s(T)$ είναι το κλάσμα του πληθυσμού που επηρεάζεται από την επιδημία. Εκτελώντας πράξεις στην τελευταία εξίσωση και κάνοντας χρήση της εξίσωσης (2.4.15) βρίσκουμε ότι το μέγεθος της επιδημίας είναι

$$s(T) = 1 - G_0(u, T) \quad (2.4.22)$$

όπου $u \equiv H_1(1, T)$ είναι η λύση της εξίσωσης

$$u = G_1(u, T) \quad (2.4.23)$$

Σημειώνουμε ότι ακόμα και άνω του T_c δεν οδηγούν όλα τα ξεσπάσματα σε επιδημία. Υπάρχουν πεπερασμένα ξεσπάσματα ακόμα και σε καθεστώς επιδημίας. Ενώ η προσέγγιση αυτή είναι πολύ

φυσική, βρίσκεται παρόλα αυτά σε αντίθεση με τα καθιερωμένα μοντέλα πλήρους ανάμειξης, στα οποία όλα τα ξεσπάσματα οδηγούν σε επιδημίες πάνω από το σημείο μετάβασης επιδημίας. Στην παρούσα περίπτωση η πιθανότητα ένα ξεσπάσμα να γίνει επιδημία είναι απλά ίση με $s(T)$.

2.4.4 Βαθμός Μολυσμένων Ατόμων

Η ποσότητα u που ορίστηκε στην εξίσωση (2.4.23) έχει μια απλή ερμηνεία: Είναι η πιθανότητα ένας κόμβος στο άκρο μιας τυχαίας επιλεγμένης ακμής να μην μολυνθεί κατά τη διάρκεια μιας επιδημίας. Θα είναι τότε για αυτόν τον κόμβο $v = 1 - T + Tu$ η πιθανότητα να μην μολυνθεί αυτός μέσω μιας ακμής. Η πιθανότητα αυτή ισούται με το άθροισμα $1 - T$, να είναι η ακμή μη κατεχόμενη και της Tu να είναι κατεχόμενη αλλά να οδηγεί σε έναν μη μολυσμένο κόμβο. Η ολική πιθανότητα ένας κόμβος να είναι μη μολυσμένος, αν έχει βαθμό k είναι v^k και η πιθανότητα να έχει βαθμό k δεδομένου ότι είναι μη μολυσμένος είναι $\frac{p_k v^k}{\sum_k p_k v^k} = \frac{p_k v^k}{G_0(v)}$ όπου η κατανομή παράγεται από την συνάρτηση $\frac{G_0(vx)}{G_0(v)}$. Διαφορίζοντας και θέτοντας $x=1$ βρίσκουμε ότι ο μέσος βαθμός z_{out} των μη μολυσμένων κόμβων είναι:

$$z_{out} = \frac{vG_0'(v)}{G_0(v)} = \frac{vG_1(v)}{G_0(v)} z = \frac{u[1-T+Tu]}{1-s} z \quad (2.4.24)$$

Παρομοίως η κατανομή βαθμού ενός μολυσμένου κόμβου παράγεται από την $\frac{G_0(x)-G_0(vx)}{1-G_0(v)}$ η οποία δίνει έναν μέσο βαθμό z_{in} μολυσμένων κόμβων

$$z_{in} = \frac{1-vG_1(v)}{1-G_0(v)} z = \frac{1-u[1-T+Tu]}{1-s} z \quad (2.4.25)$$

Σημειώνουμε ότι $1 - s = G_0(u, T) \leq u$, εφόσον όλοι οι συντελεστές της $G_0(x, T)$ είναι εξορισμού θετικοί (επειδή συνθέτουν μια κατανομή πιθανότητας) και έτσι η $G_0(x, T)$ έχει μόνο θετικές παραγώγους, το οποίο σημαίνει ότι είναι κυρτή παντού στο θετικό ημιάξονα εντός της περιοχής που συγκλίνει. Επομένως από την εξίσωση (2.4.24) $z_{out} \leq z$ και ομοίως $z_{in} \geq z$ όπως θα αναμέναμε ο μέσος αριθμός μολυσμένων ατόμων είναι μεγαλύτερος ή ίσος από αυτόν των μη μολυσμένων. Πράγματι η πιθανότητα ένας κόμβος να είναι μολυσμένος δεδομένου ότι έχει βαθμό k είναι ανάλογη του $1 - v^k = 1 - e^{-kln\frac{1}{v}}$ δηλαδή τείνει εκθετικά στη μονάδα καθώς ο βαθμός αυξάνει πολύ.

Στη συνέχεια θα παρουσιάσουμε το δεύτερο τμήμα της εργασίας του Newman στο οποίο εξετάζει την περίπτωση ο ρυθμός διάδοσης της επιδημίας και ο χρόνος παραμονής σε αυτή να μην είναι ανεξάρτητα κατανεμημένες τυχαίες μεταβλητές.

2.4.5 Συσχετισμένες Πιθανότητες Διάδοσης

Είναι δυνατόν να φανταστούμε αρκετές περιπτώσεις στις οποίες οι πιθανότητες διάδοσης μιας ασθένειας από έναν μολυσμένο προς αυτούς που έχει συνδέσεις δεν είναι ανεξάρτητα κατανομημένες τυχαίες μεταβλητές. Με άλλα λόγια, οι πιθανότητες μετάβασης από έναν δεδομένο κόμβο προς άλλους μπορούν να έχουν διαφορετικές κατανομές για διαφορετικούς κόμβους. Αυτό επιτρέπει, για παράδειγμα, τον χειρισμό περιπτώσεων στις οποίες οι πιθανότητες τείνουν όλες να είναι υψηλές ή όλες χαμηλές, αλλά υπάρχει ένα πολύ μικρό μίγμα και των δύο. Στο τμήμα αυτό θα δείξουμε πώς το προηγούμενο μοντέλο μπορεί να γενικευτεί προς το σκοπό αυτό.

Υποθέτουμε ότι οι ρυθμοί μετάδοσης r από ένα μολυσμένο κόμβο i προς κάθε άλλο k_i με τον οποίο συνδέεται έχει μια κατανομή $P_i(r)$, η οποία μπορεί να διαφέρει με όποιον τρόπο επιθυμούμε από κόμβο σε κόμβο. Τότε η a priori πιθανότητα της μετάδοσης από τον i προς κάθε έναν από τους γείτονές του στο δίκτυο είναι:

$$T_i = 1 - \int_0^\infty dr d\tau P_i(r) P_i(\tau) e^{-r\tau} \quad (2.4.26)$$

θα μπορούσαμε να επιτρέψουμε επίσης στην κατανομή $P(\tau)$ να διαφέρει από κόμβο σε κόμβο αν και αυτό δεν επιφέρει κάποια λειτουργική αλλαγή στη θεωρία, οπότε θα ήταν μάλλον άσκοπο. Σε κάθε περίπτωση ο φορμαλισμός που παρουσιάζεται μπορεί να χειριστεί εξίσου καλά και αυτή την εξάρτηση.

Από την εξίσωση (2.4.10) παρατηρούμε ότι στην εκδοχή φιλτραρίσματος του μοντέλου μας, η κατανομή των κατεχόμενων ακμών που ξεκινούν από έναν συγκεκριμένο κόμβο παράγεται τώρα από τη συνάρτηση

$$G_0(x, \{T_i\}) = \frac{1}{N} \sum_{i=0}^N \sum_{m=0}^{k_i} \binom{k_i}{m} T_i^m (1 - T_i)^{k_i - m} x^m = \frac{1}{N} \sum_{i=0}^N [1 + (x - 1)T_i]^{k_i} \quad (2.4.27)$$

Ομοίως, η κατανομή πιθανότητας των κατεχόμενων ακμών οι οποίες ξεκινούν από έναν κόμβο στον οποίο φτάνουμε ακολουθώντας μια τυχαία επιλεγμένη ακμή παράγεται από την :

$$G_1(x, \{T_i\}) = \frac{\sum_i k_i [1 + (x-1)T_i]^{k_i - 1}}{\sum_i k_i} \quad (2.4.28)$$

Όπως φαίνεται οι εξισώσεις αυτές ανάγονται στις (2.4.10), (2.4.11) όταν το T_i είναι ανεξάρτητο του i . Με αυτούς τους ορισμούς των βασικών γεννητριών συναρτήσεων η ανάλυση συνεχίζεται όπως προηγουμένως. Η συνολική κατανομή των μεγεθών ξεσπασμάτων επιδημίας, εξαιρώντας απλά ξεσπάσματα, αν αυτά υπάρχουν, παράγεται από την:

$$H_0(x, \{T_i\}) = x G_0(H_1(x, \{T_i\}), \{T_i\}) \quad (2.4.29)$$

όπου

$$H_1(x, \{T_i\}) = xG_1(H_1(x, \{T_i\}), \{T_i\}) \quad (2.4.30)$$

Το μέσο μέγεθος ξεσπάσματος όταν δεν υπάρχει επιδημία δίνεται από την εξίσωση (2.4.19) όπως προηγουμένως, και το μέγεθος επιδημιών άνω του ορίου μετάβασης από τις (2.4.22), (2.4.23). Η ίδια η μετάβαση συμβαίνει όταν $G_1'(1, \{T_i\})=1$ και αντικαθιστώντας την G_1 από την (2.4.28) μπορούμε να εκφράσουμε αυτό με τη μορφή :

$$\sum_{i=0}^N k_i [(k_i - 1)T_i - 1] = 0 \quad (2.4.31)$$

Όπως φαίνεται από την εξίσωση αυτή, αν το αριστερό μέλος είναι μεγαλύτερο του μηδενός συμβαίνει επιδημία, ενώ αν είναι μικρότερο από μηδέν δεν συμβαίνει.

Για παράδειγμα, ας θεωρήσουμε την ειδική περίπτωση στην οποία η κατανομή ρυθμών μετάβασης $P(r)$ εξαρτάται από το βαθμό μολυσμένου κόμβου. Μπορούμε να φανταστούμε ότι οι κόμβοι με μεγάλο αριθμό συνδέσεων τείνουν να έχουν χαμηλότερους ρυθμούς μετάβασης σχετικά με αυτούς που έχουν μικρό βαθμό. Σε αυτή την περίπτωση το T_i είναι μια συνάρτηση μόνο του k_i και επομένως έχουμε:

$$G_0(x, \{T_k\}) = \frac{1}{N} \sum_{i=0}^N [1 + (x - 1)T_{k_i}]^{k_i} = \sum_{k=0}^{\infty} p_k [1 + (x - 1)T_k]^{T_k} \quad (2.4.32)$$

$$G_1(x, \{T_k\}) = \frac{\sum_k k p_k [1 + (x - 1)T_k]^{T_k - 1}}{\sum_k k p_k} \quad (2.4.33)$$

όπου T_k είναι η μέση ικανότητα μετάβασης των κόμβων με βαθμό k .

Μια άλλη περίπτωση είναι αυτή όπου η ικανότητα μετάβασης είναι μια συνάρτηση των συνδέσεων που έχει ο μολυσμένος κόμβος. Αν η πιθανότητα μετάβασης σε έναν κόμβο με βαθμό k είναι U_k , τότε ορίζουμε :

$$G_0(x, \{U_k\}) = \sum_k p_k x^k \quad (2.4.34)$$

$$G_1(x, \{U_k\}) = \frac{\sum_k p_k [1 + (x^k - 1)U_k]}{\sum_k k p_k} \quad (2.4.35)$$

ΕΝΟΤΗΤΑ 3

3. Στοχαστική εκδοχή μοντέλου SIR σε δίκτυα υπολογιστών

Θα ξεκινήσουμε την αναζήτηση ενός στοχαστικού μοντέλου αντίστοιχου του SIR, το οποίο είναι ο οδηγός μας προς την κατανόηση της διάδοσης μιας επιδημίας, πρώτον επειδή είναι απλό στον εύκολο χειρισμό και δεύτερον επειδή δεν εισάγει πολλούς συγκεκριμένους περιορισμούς οι οποίοι πιθανόν θα μας απομάκρυναν από το σκοπό της διάδοσης επιδημιών στα δίκτυα υπολογιστών, εφόσον θέλουμε να διατηρήσουμε ένα απλό αλλά γενικό μοντέλο. Στην ενότητα 2 είδαμε διάφορες εκδοχές της τοπολογίας και ήδη παρατηρήσαμε ως αποδεκτό το μοντέλο SIR σε δίκτυα υπολογιστών. Συνήθως στους υπολογιστές ένας ιός ή worm απομακρύνεται, όταν το κατάλληλο update του antivirus ή firewall τρέξει. Τότε το μηχάνημα δεν κινδυνεύει πλέον από το συγκεκριμένο ανεπιθύμητο – κακόβουλο λογισμικό. Για παράδειγμα, η πιθανότητα ένας ιός να θεραπευτεί και το μηχάνημα να καταστεί ξανά ευάλωτο δεν ταιριάζει τόσο στα δίκτυα υπολογιστών. Σύμφωνα με αυτή τη φιλοσοφία είναι μάλλον προτιμότερο να ξεκινήσουμε με το SIR παρά με το SIS. Η υπόθεση της συγκεκριμένης ασθένειας είναι λογικό είτε να αναφέρεται σε μία ακριβώς ή καλύτερα, όταν είναι εφικτό, μάλλον σε μια ομάδα ασθενειών με ιδιότητες τέτοιες ώστε η διάδοση να περιγράφεται από κοινό μοντέλο. Μια αρχικά περίεργη υπόθεσή μας θα είναι ότι οι ευάλωτοι έχουν κάποιο ρόλο αρκετά πιο κοντινό σε αυτόν των μολυσμένων από ότι στην κλασική περίπτωση SIR. Όπως θα δούμε μπορούν να γεννήσουν έναν ιό και να προκύψουν μολυσμένοι από μηδενικό πλήθος αυτών. Η υπόθεση αυτή προκύπτει από την επιθυμία μας να εισάγουμε στο μοντέλο την εκ προθέσεως δημιουργία ιού από κάποιους hackers για παράδειγμα. Σύμφωνα με την υπόθεση αυτή ο πληθυσμός των ευάλωτων περικλείει και τον παραπάνω κακόβουλο πληθυσμό. Με αυτές τις παρατηρήσεις μπορούμε να διαφοροποιήσουμε το μοντέλο μας από ένα πιθανό μοντέλο IR (μολυσμένων – θεραπευμένων). Το μοντέλο που θα εξετάσουμε θα μοιράζεται την πεποίθηση ότι μόνο εφόσον ένας κόμβος έχει εγκαταστήσει το κατάλληλο «αντιβιοτικό» είναι πλέον ασφαλής, αλλά η υπόθεσή μας που το διαφοροποιεί από το κλασικό SIR φαίνεται λογική ειδικά σε μικρούς πληθυσμούς όπως στο παράδειγμα που θα εξετάσουμε ώστε να περιγράψουμε το μοντέλο.

Το παρακάτω μοντέλο δεν μπορούμε να θεωρήσουμε ότι περιγράφει πιστά τη διάδοση επιδημίας σε πραγματικά δίκτυα αλλά ενσωματώνει χαρακτηριστικά της χειρότερης κατάστασης. Η υπόθεση αυτή είναι πολύ βασική και όπως θα δούμε, θα μας υποχρεώσει να θεωρήσουμε διακριτό χρόνο. Παρόλα αυτά εισάγοντας τον πληθυσμό που έχει θεραπευτεί από την ασθένεια θεωρούμε το

μοντέλο αυτό να έχει εφαρμογή στις καταστάσεις που έχει βρεθεί θεραπεία, οπότε έτσι μπορεί να γίνει μελέτη των antivirus και firewalls. Σύμφωνα με αυτό το πνεύμα η κατεξοχήν διάδοση της επιδημίας στη χειρότερη περίπτωση μελετάται με το μοντέλο SI όπου βέβαια ενσωματώνονται τα χαρακτηριστικά των δικτύων[10]. Η περίπτωση αυτή περιγράφεται στην ενότητα 2 σε συνδυασμό με την κεντρικότητα. Έτσι το επόμενο μοντέλο μπορεί να θεωρηθεί ως μια οριακή περίπτωση για πολλά μοντέλα επιδημιών εκτός από μια εισαγωγή στα στοχαστικά μοντέλα. Επίσης μπορεί να εφαρμοστεί με αρκετή ακρίβεια σε ένα αρκετά συνδεδεμένο δίκτυο.

3.1 Εξισώσεις μοντέλου

Θεωρούμε το σταθερό πληθυσμό N καθώς και τους επιμέρους $S(t), I(t)$ και $R(t)$. Η κατάσταση του συστήματος ορίζουμε ότι είναι σε κάθε χρόνο t η διατεταγμένη τριάδα $\{S(t), I(t), R(t)\}$. Με δεδομένη την εξάρτηση των επιμέρους πληθυσμών ως το άθροισμα αυτών ίσο με το συνολικό πληθυσμό N οι ανεξάρτητες μεταβλητές κατάστασης θα είναι δυο. Μπορούμε να επιλέξουμε τις ομάδες $I(t)$ και $R(t)$. Έστω $K_{ij}(t) = \{I(t)=i, R(t)=j\}$ η κατάσταση του συστήματος την χρονική στιγμή t . Στο σημείο αυτό θα πρέπει να κάνουμε κάποιες υποθέσεις για τη στοχαστική διαδικασία. Μπορούμε να θεωρήσουμε ότι πρόκειται για μια διαδικασία Markov. Επίσης υποθέτουμε ότι κατά το χρονικό διάστημα δt (περίοδος χρόνου χωρίς βλάβη της γενικότητας) μπορούν να γίνουν ταυτόχρονα πολλαπλές γεννήσεις $\Gamma(\delta t)$ και θάνατοι $\Theta(\delta t)$. Προφανώς θα πρέπει $\Theta(\delta t) \leq I(t) + \Gamma(\delta t)$. Για ένα στοχαστικό σύστημα διακριτού χρόνου θα πρέπει να αναζητήσουμε την πιθανότητα $P_{ij}(t+1)$ το σύστημα να είναι στην κατάσταση ij την χρονική στιγμή $t+1$ αν την χρονική στιγμή t υπολογιστούν οι δυνατοί τρόποι για τη μετάβαση στην κατάσταση ij [7]. Συμβολικά θα έχουμε :

$$P_{ij}(t+1) = \sum_{\lambda=0}^j \sum_{\kappa=0}^{i+\lambda} P_{\kappa, j-\lambda}(t) Q_{\kappa, j-\lambda; ij}(\delta t) \quad (3.1.1)$$

Όπου $Q_{\kappa, j-\lambda; ij}(\delta t)$ συμβολίζουμε την πιθανότητα το σύστημα να μεταβεί από την κατάσταση $\kappa, j-\lambda$ στην οποία βρίσκεται στο χρόνο t , στην ij στην οποία θα βρίσκεται στο χρόνο $t+1$. Αν και χωρίς βλάβη της γενικότητας το βήμα χρόνου μπορεί να είναι 1, θα αναφέρουμε κάποιες φορές το βήμα χρόνου ως δt . Αναφέρουμε εδώ σχετικά με τον συμβολισμό μας ότι δύο ομάδες μεταβλητών θα διαχωρίζονται με ερωτηματικό ενώ οι δύο μεταβλητές της κάθε ομάδας θα διαχωρίζονται με κόμμα ή χωρίς κάποιο σύμβολο για απλότητα στις εξισώσεις όταν η κάθε μεταβλητή αναπαριστάνεται ως ένα μοναδικό σύμβολο. Για παράδειγμα στην προηγούμενη εξίσωση συμβολίζουμε απλά ως ij και όχι i, j . Για την ομάδα $\kappa, j-\lambda$ το κόμμα είναι αναγκαστικό ώστε να διαχωρίσει τις δύο μεταβλητές προς αποφυγή σύγχυσης. Όπως μπορεί να δειχθεί το πλήθος των δυνατών διαφορετικών καταστάσεων

είναι $\frac{N(N+3)}{2} + 1$. Επίσης το πλήθος των δυνατών μεταβάσεων προς μια συγκεκριμένη κατάσταση i, j σε διάστημα δt είναι $(j + 1)(i + \frac{j+2}{2})$, όσοι και οι όροι του αθροίσματος στην (3.1.1).

Θα περιγράψουμε τώρα την εξίσωση 3.1.1 ώστε να γίνουν φανερός ο τρόπος παραγωγής της και η σημασία της. Θεωρούμε ότι στον πληθυσμό των αρρώστων $I(t)=k$ συμβαίνουν πολλαπλές γεννήσεις και θάνατοι τα οποία αντιστοιχούν στις μολύνσεις και θεραπείες αντίστοιχα. Προφανώς ο αριθμός των θανάτων, έστω Θ , που συμβαίνουν κατά το χρονικό διάστημα δt είναι ίσος με τις γεννήσεις στον πληθυσμό $R(t)=\lambda$ κατά το ίδιο διάστημα. Αν κ, λ η κατάσταση στον χρόνο t τότε θα ισχύουν:

$$\kappa + \Gamma - \Theta = i \quad \text{και} \quad \lambda + \Theta = j \quad (3.1.2)$$

ώστε για $t+\delta t$ η κατάσταση να είναι i, j . Υποθέσαμε εδώ ότι κατά το χρονικό διάστημα δt μπορούν να γίνουν ταυτόχρονα πολλαπλές γεννήσεις και πολλαπλοί θάνατοι.

Έστω ότι θέλουμε να υπολογίσουμε την πιθανότητα $P_{i0}(t+1)$. Αυτή είναι η πιθανότητα το σύστημα να έχει i μολυσμένους και 0 θεραπευμένους σε χρόνο $t+1$ και αναζητάμε τα δυνατά γεγονότα της στιγμής t τα οποία μπορούν να οδηγήσουν σε αυτό. Εφόσον $j=0$ θα έχουμε $\lambda=0$. Αυτό προκύπτει και από την γενική παρατήρηση ότι ο πληθυσμός R δεν μειώνεται ποτέ. Από την (3.1.2) παίρνουμε $\Theta=0$ οπότε $\kappa+\Gamma=i$. Υπολογίζουμε διαδοχικά για $\kappa=0, \kappa=1, \dots, \kappa=i$ τις δυνατές τιμές των γεννήσεων Γ . Το σύνολο που προκύπτει είναι $(i, i-1, \dots, 0)$. Περιλαμβάνεται δηλαδή και η περίπτωση να είμαστε ήδη στην μεταβλητή κατάσταση i οπότε η κατάσταση παρέμεινε αμετάβλητη, καθώς και η κατάσταση να είμαστε στην κατάσταση $0,0$ και να έγιναν i γεννήσεις. Εδώ έχουμε και μια διαφοροποίηση σε σχέση με το αιτιοκρατικό SIR μοντέλο όπου εκεί χρειαζόταν τουλάχιστον ένας άρρωστος ώστε να συνεχιστεί η διαδικασία, ενώ στο стоχαστικό μοντέλο μας, επιτρέπεται γέννηση εκ του μηδενός (αν και αυτό μπορούμε με κατάλληλη τροποποίηση να το αποφύγουμε). Από τα παραπάνω για την απλή αυτή περίπτωση P_{i0} παίρνουμε $P_{i0}(t+1) = \sum_{\kappa=0}^i P_{\kappa 0}(t) Q_{\kappa 0; i 0}(\delta t)$. Έστω τώρα η πιθανότητα P_{i1} . Θα είναι $\Theta=\{0,1\}$. Για κάθε τιμή του Θ υπολογίζουμε το σύνολο γεννήσεων οπότε προκύπτει από την (3.1.2) $(i, i-1, \dots, 0)$ και $(i+1, i, i-1, \dots, 0)$ για $\Theta=0$ και $\Theta=1$ αντίστοιχα. Θα έχουμε δηλαδή είτε ακριβώς 0 θανάτους και $i+1$ γεννήσεις, είτε ακριβώς 1 θάνατο και $i+2$ γεννήσεις (περιλαμβάνεται και η περίπτωση 0 γεννήσεις).

Επομένως

$$P_{i1}(t+1) = \sum_{\kappa=0}^i P_{\kappa 1}(t) Q_{\kappa 1; i 1}(\delta t) + \sum_{\kappa=0}^{i+1} P_{\kappa 0}(t) Q_{\kappa 0; i 1}(\delta t) = \sum_{\lambda=0}^1 \sum_{\kappa=0}^{i+\lambda} P_{\kappa, 1-\lambda}(t) Q_{\kappa, 1-\lambda; i 1}(\delta t).$$

Εκτελώντας ένα ακόμα βήμα στην επαγωγική μεθοδολογία μας παίρνουμε :

$$P_{i2}(t+1) = \sum_{\kappa=0}^i P_{\kappa 2}(t) Q_{\kappa 2; i2}(\delta t) + \sum_{\kappa=0}^{i+1} P_{\kappa 1}(t) Q_{\kappa 1; i2}(\delta t) + \sum_{\kappa=0}^{i+2} P_{\kappa 0}(t) Q_{\kappa 0; i2}(\delta t) = \sum_{\lambda=0}^2 \sum_{\kappa=0}^{i+\lambda} P_{\kappa, 2-\lambda}(t) Q_{\kappa, 2-\lambda; i2}(\delta t)$$

Παρατηρώντας την παραπάνω εξίσωση καθώς και αυτήν της P_{i1} και συγκρίνοντας τις με την (3.1.1) είναι εύκολο να δει κανείς ότι η μόνη διαφορά είναι στην τιμή της μεταβλητής j . Γενικεύοντας με επαγωγή καταλήγουμε στην έκφραση (3.1.1). Η εξίσωση αυτή θα πρέπει να θεωρείται μια χαρακτηριστική του στοχαστικού μοντέλου που ενσωματώνει τις εξής υποθέσεις:

α) Έχουμε μια διαδικασία Markov

β) Κατά το χρονικό διάστημα $\delta t=1$ μπορούν να συμβούν ταυτόχρονα πολλαπλές γεννήσεις και πολλαπλοί θάνατοι. Συγκεκριμένα επιτρέπονται από 0 έως τις μέγιστες δυνατές γεννήσεις $N-i-j$ και από 0 έως τους μέγιστους δυνατούς θανάτους $N-j$. Στο σημείο αυτό πρέπει να αναφέρουμε ότι ο αριθμός $N-j$ προκύπτει από την υπόθεση ότι οι υποψήφιοι θάνατοι δεν είναι απλώς όσοι και ο μολυσμένος πληθυσμός $I(t)=i$, αλλά από 0 έως i +πλήθος δυνατών γεννήσεων στο διάστημα δt . Δηλαδή είναι δυνατό κατά το διάστημα δt να υπάρχει μετάδοση ασθένειας και αμέσως θεραπεία του ίδιου ατόμου.

γ) Το μέγεθος του συνολικού πληθυσμού διατηρείται σταθερό, ίσο με N το οποίο διαιρείται στους γνωστούς πληθυσμούς $S(t)$, $I(t)$ και $R(t)$.

Μια πλήρης στοχαστική μεταφορά του μοντέλου SIR θα πρέπει να ενσωματώνει και τις αναλογίες που υποδηλώνουν οι αιτιοκρατικές εξισώσεις αλλά αυτό μπορεί να εξεταστεί αργότερα. Επιπλέον με σκοπό να εμβαθύνουμε στα δίκτυα, θα πρέπει αργότερα να εισάγουμε κατάλληλους περιορισμούς και συνθήκες. Το παρόν μοντέλο ειδικά με την υπόθεση ότι σε κάθε βήμα μπορούν να συμβούν οι μέγιστες μεταδόσεις ασθένειας και μάλιστα οι ίδιες να θεραπευτούν, είναι η πιο ακραία περίπτωση επιδημίας. Υπάρχει πιθανότητα από την κατάσταση $(0,0)$ να μεταβούμε κατευθείαν στην $(0,N)$ όπου είναι το τέλος της επιδημίας. Θα περιγράψουμε τη διαδικασία με τις παρούσες συνθήκες αν και η αλλαγή σε άλλες είναι εφικτή με όμοια μεθοδολογία.

Η εξίσωση της κατάστασης $(0,0)$ είναι $P_{00}(t+1) = P_{00}(t) Q_{00;00}(\delta t)$ το οποίο απλώς δηλώνει την πιθανότητα το σύστημα να παραμείνει στην μηδενική κατάσταση η οποία θεωρείται μια αρχική κατάσταση. Αρχικές καταστάσεις μπορούμε να ορίσουμε και άλλες της μορφής P_{ij} με $i \neq 0$ και $j = 0$ ανάλογα με την εφαρμογή. Μπορούμε να ορίσουμε και όλες τις καταστάσεις ως αρχικές όταν δεν έχουμε ακριβή εικόνα. Αυτό θα εξυπηρετήσει το μοντέλο μας, όπως θα δούμε παρακάτω όπου θα

αναζητήσουμε την πιθανότητα μιας κατάστασης σε χρόνο n . Έτσι μπορούμε να έχουμε μια κατανομή πιθανοτήτων για τις αρχικές καταστάσεις παρά μία μόνο κατάσταση. Ως τελικές καταστάσεις δεν ορίζουμε αυτές της μορφής P_{ij} με $i = 0$ και $j \neq 0$, αλλά την $(0, N)$ σε συμφωνία με την παραδοχή μας ότι είναι δυνατή η γέννηση ασθένειας σε συνολικό πληθυσμό με μηδενικό πλήθος αρρώστων αλλά μη μηδενικό πλήθος ευάλωτων. Δηλαδή είναι δυνατή η μετάβαση $(0, j < N)_t \rightarrow (i, j)_{t+\delta t}$. Στο σημείο αυτό έχουμε και μια διαφοροποίηση με το αιτιοκρατικό μοντέλο SIR. Εκεί η αναγκαιότητα ύπαρξης τουλάχιστον ενός ασθενή σημαίνει παρόλα αυτά ότι κάποια στιγμή ένας σταθερού πλήθους πληθυσμός ήρθε σε επαφή με μία εξωτερική του συστήματος ασθένεια. Στην πιθανοτική θεώρησή μας, εδώ με την υπόθεσή μας ότι η ασθένεια μπορεί να γεννηθεί μέσα στο σύστημα, ενσωματώνουμε ένα αναγκαίο γεγονός, ώστε να ξεκινήσει μια επιδημία. Επίσης στην παρούσα εξέτασή μας στο τέλος της επιδημίας δεν θα υπάρχουν ευάλωτοι. Η παραπάνω θεώρηση μπορεί να ιδωθεί και ως ένα εργαλείο προς το φορμαλισμό μας αφού σε κάθε περίπτωση μπορούμε να θεωρήσουμε ως τέλος της επιδημίας την κατάσταση που δεν θα έχουμε κανέναν μολυσμένο.

Προφανώς θα ισχύει επίσης

$$\sum_{j=0}^{\infty} P_{i,j}(t) = 1 \quad (3.1.3)$$

Εξισώσεις Δυναμικής Συστήματος :

Στα προηγούμενα αναφέραμε το σύνολο των μεταβάσεων από μια κατάσταση στην ij . Εδώ θα εξάγουμε τις εξισώσεις διαφορών που παριστάνουν τη δυναμική του πιθανοτικού μας συστήματος εισάγοντας πρώτα τους ρυθμούς μετάβασης. Οι ρυθμοί αυτοί μπορούν να θεωρηθούν και ως γενικευμένοι ρυθμοί γεννήσεων-θανάτων, όπου επιτρέπονται δηλαδή πολλαπλές γεννήσεις-θάνατοι. Ο ρυθμός $\mu_{\kappa,\lambda;i,j}$ είναι αυτός ώστε, ενώ το σύστημα είναι στην κατάσταση ij , θα γίνουν κ γεννήσεις και λ θάνατοι, ώστε να μεταβούμε στην κατάσταση που ορίζεται από τις (3.1.2). Ο κάθε ρυθμός μετάβασης επί την πιθανότητα να είμαστε στην κατάσταση του ρυθμού δίνει ένα μέρος της ροής προς μια συγκεκριμένη κατάσταση. Για παράδειγμα, με $\mu_{1,1;0,0} P_{0,0}(t)$ παίρνουμε ένα μέρος της ροής προς την κατάσταση $i=0+1-1$ $j=0+1$ ή $(0,1)$. Ακολουθώντας παρόμοια διαδικασία με αυτή που προέκυψε η (3.1.1) καταλήγουμε στην εξίσωση που περιγράφει την συνολική ροή προς μια κατάσταση ij .

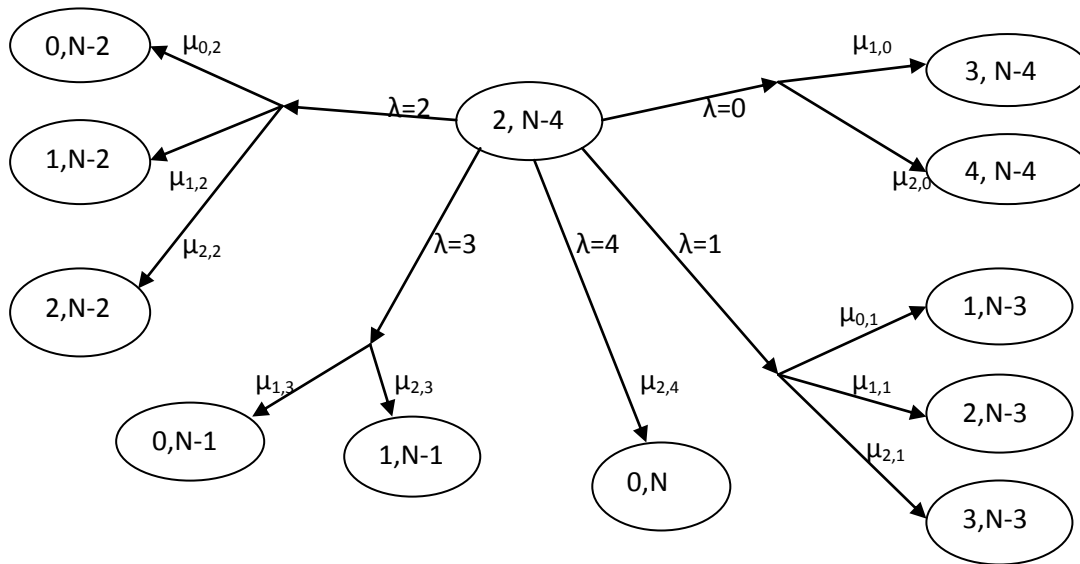
$$\text{Ροή προς την κατάσταση } i, j = -\mu_{0,0} P_{ij}(t) + \sum_{\lambda=0}^j \sum_{\kappa=0}^{i+\lambda} \mu_{\kappa,\lambda} P_{i-\kappa+\lambda, j-\lambda}(t) \quad (3.1.4)$$

Στην εξίσωση αυτή παραλείπουμε τον δείκτη κατάστασης $i-\kappa+\lambda, j-\lambda$ του ρυθμού ο οποίος μπορεί να εννοηθεί εφόσον εμφανίζεται πάντα δίπλα στην αντίστοιχη πιθανότητα κατάστασης. Επίσης, επειδή το άθροισμα περιέχει τη ροή από την κατάσταση ij προς τον εαυτό της, αφαιρούμε αυτή την

ποσότητα. Στη συνέχεια υπολογίζοντας τα δυνατά γεγονότα από την ροή i, j προς το σύνολο καταστάσεων έξω από αυτήν παίρνουμε την εξίσωση:

$$\text{Ροή έξω απο την κατάσταση } i, j = -\mu_{0,0}P_{ij} + P_{ij}(t) \sum_{\lambda=0}^{N-j} \sum_{\kappa=\kappa_{min}}^{N-i-j} \mu_{\kappa,\lambda} \quad (3.1.5)$$

όπου $\kappa_{min} = \begin{cases} 0, & \lambda \leq i \\ \lambda - i, & \lambda > i \end{cases}$. Το νόημα του κάτω άκρου κ_{min} μπορεί να αποδοθεί καλύτερα με ένα παράδειγμα. Έστω η ροή έξω από την κατάσταση $(2, N-4)$



Σχήμα 3.1.1 : Ροή έξω από την κατάσταση $(2, N-4)$

Στο σχήμα 3.1.1 φαίνονται οι δυνατές καταστάσεις μετάβασης από την $2, N-4$ με $\mu_{\kappa,\lambda}$ ρυθμούς των κ γεννήσεων και λ θανάτων. Έχουμε χωρίσει το σύνολο των μεταβάσεων σε ομάδες ανάλογα με το πλήθος των δυνατών θανάτων λ . Όπως φαίνεται σε κάθε ομάδα, αν $\lambda \leq i$ τότε αυτή έχει $N - i - j + 1 = -2 + 4 + 1 = 3$ μέλη. Εξαιρέση αποτελεί η ομάδα για $\lambda=0$ αλλά με σκοπό την παραγωγή του γενικού τύπου θεωρούμε ότι και αυτή έχει 3 μέλη. Στο τέλος όπως φαίνεται από την (3.1.5) αφαιρούμε τον περιττό όρο. Αν $\lambda > i$ τότε για $\lambda = i + 1, i + 2, \dots, N - j$ ο αριθμός γεννήσεων είναι $\lambda - i = 1, 2, \dots, N - j - i$. Επομένως αν και πάντα ο μέγιστος αριθμός γεννήσεων είναι $N-j-i$ ο ελάχιστος είναι $\lambda-i$ όταν $\lambda > i$. Μπορούμε εναλλακτικά να θέσουμε στην (3.1.5) $\kappa_{min} = \lambda - i$ για κάθε λ με την προϋπόθεση ότι $\kappa_{min} = 0$ για $\lambda - i < 0$.

Από τις (1.3.4) και (1.3.5) έχουμε την μεταβολή της πιθανότητας να βρισκόμαστε στην κατάσταση ij

$$\Delta P_{ij}(t) = \sum_{\lambda=0}^j \sum_{\kappa=0}^{i+\lambda} \mu_{\kappa,\lambda} P_{i-\kappa+\lambda, j-\lambda}(t) - P_{ij}(t) \sum_{\lambda=0}^{N-j} \sum_{\kappa=\kappa_{min}}^{N-i-j} \mu_{\kappa,\lambda} \quad (3.1.6)$$

Όπου στον πρώτο όρο οι ρυθμοί $\mu_{\kappa,\lambda}$ αναφέρονται στην κατάσταση $i-\kappa+\lambda, j-\lambda$ ενώ στο δεύτερο όρο στην κατάσταση ij .

Υπολογίζουμε την $\Delta P_{00}(t)$.

$$\begin{aligned} \Delta P_{00}(t) &= \mu_{0,0} P_{0,0}(t) - P_{0,0}(t) \sum_{\lambda=0}^N \sum_{\kappa=\kappa_{min}}^N \mu_{\kappa,\lambda} = \mu_{0,0} P_{0,0}(t) - P_{0,0}(t) \left(\sum_{\kappa=0}^N \mu_{\kappa,0} + \sum_{\lambda=1}^N \sum_{\kappa=\lambda}^N \mu_{\kappa,\lambda} \right) \\ P_{0,0}(t+1) &= P_{0,0}(t) \left(1 - \left(\sum_{\kappa=1}^N \mu_{\kappa,0} + \sum_{\lambda=1}^N \sum_{\kappa=\lambda}^N \mu_{\kappa,\lambda} \right) \right) \end{aligned}$$

Σε κάθε τέτοιο υπολογισμό από την (3.1.6) για κάθε ij η ροή $\mu_{0,0}P_{ij}(t)$ από την ij προς τον εαυτό της θα απαλειφθεί ως δύο αντίθετοι όροι. Συγκρίνοντας την τελευταία εξίσωση με την (3.1.1) έχουμε ότι $Q_{0,0;0,0} = (1 - \delta t (\sum_{\kappa=1}^N \mu_{\kappa,0} + \sum_{\lambda=1}^N \sum_{\kappa=\lambda}^N \mu_{\kappa,\lambda}))$ αλλά $Q_{0,0;0,0} = \mu_{0,0;0,0} \delta t$. Εξισώνοντας και έπειτα από εκτέλεση πράξεων παίρνουμε $1 = (\sum_{\kappa=0}^N \mu_{\kappa,0} + \sum_{\lambda=1}^N \sum_{\kappa=\lambda}^N \mu_{\kappa,\lambda}) \delta t$ και $P_{0,0}(t+1) = \mu_{0,0;0,0} P_{0,0}(t)$

Αξίζει να αναφέρουμε κάποια γενικά χαρακτηριστικά των διαδικασιών Markov και να δούμε αν ικανοποιούνται από το μοντέλο μας. Καταρχήν δεν έχουμε υιοθετήσει προς το παρόν μια κατανομή για τους χρόνους μεταξύ διαδοχικών μεταβάσεων οπότε αν αυτοί δεν έχουν την εκθετική κατανομή θα έχουμε μια διαδικασία semi-Markov. Ένα χαρακτηριστικό της διαδικασίας μας είναι ότι ο πληθυσμός $S(t)$ δεν αυξάνεται ποτέ, ενώ ο πληθυσμός $R(t)$ δεν μειώνεται ποτέ. Αυτό μας οδηγεί να αναμένουμε ότι ο πληθυσμός $I(t)$ αρχικά θα αυξάνεται και κάποια στιγμή θα φτάσει σε ένα μέγιστο οπότε στη συνέχεια θα μειώνεται έως το τέλος της διαδικασίας στην κατάσταση $(0, N)$. Η παραπάνω συμπεριφορά αποκλείει τη διαδικασία από την ιδιότητα μιας αμείωτης (irreducible) αλυσίδας. Ως επακόλουθο, το ερώτημα αν υπάρχει κατανομή μόνιμης κατάστασης μπορεί να αναχθεί στην εύρεση της οριακής πιθανότητας $P_{0,N}$. Θα πρέπει να θεωρείται προφανές ότι η ροή έξω από την κατάσταση $(0, N)$ είναι μηδέν αφού δεν υπάρχει κανένα δυνατό γεγονός από αυτή την κατάσταση προς κάποια άλλη. Η εξίσωση (3.1.5) μας δίνει το σωστό αποτέλεσμα 0 για τη ροή έξω από την $(0, N)$ αλλά δεν μπορεί να μας πληροφορήσει για την οριακή πιθανότητα $P_{0,N}$ (Ανάλογο αποτέλεσμα ισχύει και για τη ροή προς την κατάσταση $(0, 0)$ και την $P_{0,0}$). Από την εξίσωση (3.1.6) εκτελώντας τους υπολογισμούς για $\Delta P_{0N}(t)$ καταλήγουμε στο αποτέλεσμα

$$\Delta P_{0N}(t) = \sum_{\lambda=1}^N \sum_{\kappa=0}^{\lambda} \mu_{\kappa,\lambda; \lambda-\kappa, N-\lambda} P_{\lambda-\kappa, N-\lambda} \quad (3.1.7)$$

3.2 Λύση εξισώσεων

Έχοντας εξετάσει χαρακτηριστικά τοπολογίας δικτύων τα οποία εμπλέκονται στην διαδικασία εξάπλωσης μιας επιδημίας θα επιστρέψουμε στην αρχική μας παρουσίαση του μοντέλου SIR για έναν πεπερασμένο πληθυσμό. Επιθυμούμε να συνδυάσουμε τα αποτελέσματα της ενότητας αυτής με μια πιο συγκεκριμένη δομή δικτύου. Εξετάζοντας την εξίσωση (3.1.1) διαπιστώνουμε ότι αυτή μας δίνει την πιθανότητα για χρόνο $t+1$, αν γνωρίζουμε τις πιθανότητες σε χρόνο t , δηλαδή μας παρέχει πληροφορία για διάστημα χρόνου δt , ενώ ο σκοπός μας είναι να παράγουμε μια έκφραση για κάθε t αν γνωρίζουμε την αρχική κατανομή κατάστασης του συστήματος.

Παρατηρούμε αρχικά ότι είναι χρήσιμο να έχουμε έναν πίνακα όπου θα φαίνεται ποιες καταστάσεις μπορούν να οδηγήσουν σε ένα βήμα χρόνου σε μια συγκεκριμένη. Ως παράδειγμα στην αναζήτησή μας θα πάρουμε το απλό δίκτυο 4 κόμβων το οποίο παρά το μικρό μέγεθός του είναι ικανό να μας αποκαλύψει αρκετές ιδιότητες του μοντέλου επιδημίας.

Κατάσταση	i,j	Καταστάσεις προς i,j σε δt	M
1	0,0	1	1
2	1,0	1,2	2
3	0,1	1,2,3	3
4	2,0	1,2,4	3
5	1,1	1,2,3,5	4
6	0,2	1,2,3,4,5,6	6
7	3,0	1,2,4,7	4
8	2,1	1,2,3,4,5,7,8	7
9	1,2	1,2,3,4,5,7,8,9	8
10	0,3	1,2,3,4,5,6,7,8,9,10	10
11	4,0	1,2,4,7,11	5
12	3,1	1,2,3,4,5,7,8, 11,12	9
13	2,2	1,2,3,4,5,6,7,8,9,11,12,13	12
14	1,3	1,2,3,4,5,6,7,8,9,10,11,12,13,14	14
15	0,4	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15	15

Πίνακας 3.2.1

Στον πίνακα 3.1 φαίνονται οι $\frac{4(4+3)}{2} + 1 = 15$ καταστάσεις και για κάθε μία το σύνολο των δυνατών προηγούμενων από όπου σε ένα βήμα χρόνου μπορούμε να φτάσουμε σε μια συγκεκριμένη. Εναλλακτικά αντί για t θα συμβολίζουμε με n το διακριτό χρόνο. Έχουμε αριθμήσει τις καταστάσεις οπότε από την αναφορά τους με 2 σύμβολα χρειάζεται πλέον 1. Η αρίθμηση μπορεί να είναι αυθαίρετη, αν και όπως φαίνεται στον πίνακα έχουμε ακολουθήσει αυτή ώστε να αυξάνεται με αύξηση του $i+j$ και για σταθερό $i+j$ να αυξάνεται με το j . Η τρίτη στήλη προκύπτει από τις εξισώσεις (3.1.2) $\kappa + \Gamma - \theta = i$ και $\lambda + \theta = j$ ώστε η κατάσταση i,j να προκύπτει από την κ,λ σε ένα βήμα χρόνου. Όπως φαίνεται κάθε κατάσταση μπορεί να προέλθει από τον εαυτό της και από την 0,0. Η πιο σημαντική ιδιότητα μιας κατάστασης είναι ότι ανεξάρτητα από το χρόνο που αυτή θα υπάρξει, θα περάσει από συγκεκριμένες καταστάσεις οι οποίες είναι ακριβώς το σύνολο αυτών στην τρίτη στήλη του πίνακα στο σχήμα 3.1. Λόγω αυτής της ιδιότητας θα ονομάσουμε το παραπάνω σύνολο ομάδα κατάστασης. Αν γνωρίζουμε την ομάδα μιας κατάστασης μπορούμε να περιγράψουμε πλήρως το παρελθόν της σε συνδυασμό με την αρχική ή αρχικές καταστάσεις. Ένα χαρακτηριστικό της ομάδας κατάστασης είναι ότι αυτή έχει διάταξη στο χρόνο. Σε κάθε ομάδα η διέλευση των καταστάσεων της στο χρόνο είναι μοναδική. Η ιδιότητα αυτή σχετίζεται άμεσα και με την αρίθμηση των καταστάσεων. Για τη συγκεκριμένη αρίθμηση όπως θα δείξουμε η κίνηση στο χρόνο είναι μόνο κάθετη στο σχήμα, δηλαδή προς μεγαλύτερου αύξοντα αριθμού καταστάσεις. Έχουμε αριθμήσει τις καταστάσεις ανάλογα με το $i+j$ οπότε προκύπτουν ομάδες των 1,2,3,4 και 5 στοιχείων η κάθε μία καθώς κοιτάμε κατακόρυφα. Σε σταθερό $i+j$ για παράδειγμα ίσο με 3 έχουμε τις καταστάσεις 7, 8, 9, και 10 οι οποίες αυξάνονται με το j . Έστω ότι εξετάζουμε την πιθανότητα σε χρόνο $n\delta t$ μιας κατάστασης ij . Θα λάβουμε υπόψη μας τις καταστάσεις της ομάδας της και με την σειρά που υποδηλώνει η αρίθμηση, επειδή η ομάδα της κατάστασης αποκλείεται να αναφέρεται σε γραμμή του πίνακα κάτω από την ij . Αν σε χρόνο n έχουμε την κατάσταση 7 και σε χρόνο $(n-1)\delta t$ έχουμε την 4 τότε για όλους τους προηγούμενους χρόνους οι δυνατές καταστάσεις είναι αυτές της ομάδας 4 και οι οποίες περιέχονται στην 7. Ο τρόπος κατασκευής μιας ομάδας προκύπτει από τις παρακάτω δύο συνθήκες.

Αν για χρόνο $n\delta t$ έχουμε την κατάσταση i,j , τότε οι καταστάσεις κ,λ για χρόνο μικρότερο ή ίσο με $n\delta t$ ικανοποιούν τις:

- a) $\lambda \leq j$
- b) $\kappa + \lambda \leq i + j$

Συμβολίζοντας πλέον ως $Q_{\alpha,\beta}$ την πιθανότητα μετάβασης από την α για χρόνο t στην β για χρόνο $t+1$ αναμένουμε για την έκφραση της $P_{ij}(t+1)$ να υπεισέρχονται σε αυτήν γινόμενα της μορφής: $Q_{\alpha,\beta}Q_{\beta,\gamma}\dots Q_{\gamma',\tau}$ το καθένα n παραγόντων όσο και το πλήθος των διαστημάτων χρόνου. Σε κάθε

έκφραση μιας πιθανότητας σε χρόνο $t+n\delta t$ με γνωστές τις πιθανότητες για t μπορούμε χωρίς βλάβη της γενικότητας να θέσουμε $t=0$. Θα εξετάσουμε τώρα την ιστορία της κατάστασης 7 σύμφωνα με την αρίθμηση του πίνακα 3.1. Αν για $t=n\delta t$ θεωρούμε την πιθανότητα της κατάστασης 7 θα ισχύει από την 3.1.1

$$P_{7|t=n} = P_1 Q_{1;7} + P_2 Q_{2;7} + P_4 Q_{4;7} + P_7 Q_{7;7} \Big|_{t=(n-1)} \quad (3.2.1)$$

Εκτελώντας την διαδικασία για ένα ακόμα βήμα χρόνου θα πάρουμε

$$P_{7|t=n} = P_1 Q_{1;1} Q_{1;7} + (P_1 Q_{1;2} + P_2 Q_{2;2}) Q_{2;7} + (P_1 Q_{1;4} + P_2 Q_{2;4} + P_4 Q_{4;4}) Q_{4;7} + (P_1 Q_{1;7} + P_2 Q_{2;7} + P_4 Q_{4;7} + P_7 Q_{7;7}) Q_{7;7} \Big|_{t=(n-2)} \quad (3.2.2)$$

Η (3.2.2) προέκυψε από την (3.2.1) με εφαρμογή της (3.1.1) στην τελευταία. Η διαδικασία μπορεί να επαναληφθεί για όσα βήματα χρόνου επιθυμούμε και να εκτελέσουμε τους υπολογισμούς με δεδομένα την αρχική κατανομή της ομάδας 7 και τις πιθανότητες μετάβασης μεταξύ των καταστάσεων της ομάδας. Όπως γίνεται φανερό, αν δεν έχουμε μια πιο συγκεκριμένη γνώση, ως αρχικές καταστάσεις θεωρούνται όλες οι καταστάσεις της ομάδας εν γένει και αυτό είναι ανεξάρτητο του χρόνου n . Εξάιρεση αποτελεί η κατάσταση 15 η οποία οδηγεί μόνο στον εαυτό της και η ομάδα της θα περιέχει και τις 15 καταστάσεις. Στο άλλο άκρο, αν γνωρίζουμε ακριβώς την αρχική κατάσταση σε μια έκφραση όπως η (3.1.2) και ανεξάρτητα του χρόνου n , θα μηδενίζονται όλες οι αρχικές πιθανότητες των καταστάσεων εκτός από αυτή που είναι γνωστή με βεβαιότητα και προφανώς παίρνει την τιμή 1. Ανάλογα με την εκτίμησή μας για την αρχική κατάσταση του συστήματος σε μια ενδιάμεση περίπτωση των δύο ανωτέρω, κάποιες πιθανότητες μπορούν να μηδενιστούν. Με σκοπό όπως προαναφέραμε να μην εφαρμόζουμε αναδρομικά την σχέση (1.3.1) μπορούμε να παρατηρήσουμε ότι η ιστορία μιας κατάστασης μπορεί να γραφεί πιο απλά ως εξής :

Έχοντας αριθμήσει τις $\frac{N(N+3)}{2} + 1$ καταστάσεις του δικτύου αρχίζουμε να καταγράφουμε τις μεταβάσεις προς την τελική κατάσταση που μας ενδιαφέρει σε χρόνο n σύμφωνα με τον κανόνα ότι η επόμενη κατάσταση κάθε φορά είναι η ίδια ή κάποια μεγαλύτερή της το πολύ έως αυτήν σε χρόνο n . Στο παράδειγμά μας για την κατάσταση 7 σε χρόνο n θα έχουμε τις παρακάτω δυνατότητες :

t=0	t=1	t=2	t=3
1	1	1	7
1	1	2	7
1	1	4	7
1	1	7	7
1	2	2	7

1	2	4	7
1	2	7	7
1	4	4	7
1	4	7	7
1	7	7	7
2	2	2	7
2	2	4	7
2	2	7	7
2	4	4	7
2	4	7	7
2	7	7	7
4	4	4	7
4	4	7	7
4	7	7	7
7	7	7	7

Στο διπλανό σχήμα βλέπουμε κάθε δυνατή ιστορία προς την κατάσταση 7 η οποία έχει ομάδα την (1,2,4,7) 4 στοιχείων και για 3 βήματα χρόνου. Η γενίκευση προς μεγαλύτερα βήματα χρόνου είναι άμεση αρκεί να ακολουθήσουμε τον κανόνα που αναφέραμε. Κάθε γραμμή αντιστοιχεί σε ένα γινόμενο πιθανοτήτων μεταβάσεων με n παράγοντες. Η συνολική πιθανότητα της κατάστασης 7 προκύπτει από το άθροισμα αυτών των γινομένων το καθένα πολλαπλασιασμένο με την αρχική πιθανότητα για $t=0$, η κατάσταση της οποίας φαίνεται στον πίνακα 3.2.

Η αναφορά μας στον τρόπο αρίθμησης είναι ουσιαστική ώστε να ακολουθήσουμε τον κανόνα προς μεγαλύτερες καταστάσεις. Μια άλλη αρίθμηση που επιτυγχάνει το ίδιο είναι να χωρίσουμε τις καταστάσεις ανάλογα με το j πρώτα και για σταθερό j να αυξάνεται με αύξηση του $i+j$.

Σχήμα 3.2

Η παραπάνω προσέγγιση μπορεί να παρασταθεί πιο κομψά μέσω πινάκων. Θεωρούμε τον διαγώνιο πίνακα P_0 του οποίου τα στοιχεία είναι οι αρχικές πιθανότητες των καταστάσεων. Για την κάθε κατάσταση που ενδιαφερόμαστε θεωρούμε τον πίνακα μετάβασης Q του οποίου το στοιχείο $q_{a,b}$ ισούται με την πιθανότητα μετάβασης από την κατάσταση a στη b . Τότε η πιθανότητα μιας κατάστασης σε χρόνο k δίνεται από

$$P_k = P_0 Q^k \quad (3.2.3)$$

όπου

$$Q = \begin{bmatrix} q_{1,1} & q_{1,2} & 0 & q_{1,4} & 0 & 0 & q_{1,7} \\ 0 & q_{2,2} & 0 & q_{2,4} & 0 & 0 & q_{2,7} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & q_{4,4} & 0 & 0 & q_{4,7} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & q_{7,7} \end{bmatrix}$$

Ως παράδειγμα για την κατάσταση 7 όπως προηγουμένως θα έχουμε:

$$P_{|k=2} = \begin{bmatrix} P_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & P_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & P_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & P_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & P_5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & P_6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & P_7 \end{bmatrix} Q^2. \quad \text{Εκτελώντας τις πράξεις θα καταλήξουμε στο}$$

αποτέλεσμα της (3.2.2) παίρνοντας το άθροισμα της τελευταίας στήλης του $P_{|k=2}$. Ανάλογα, κατασκευάζοντας τον αντίστοιχο πίνακα για τυχούσα κατάσταση σύμφωνα με την ταξινόμηση του πίνακα 3.1 ,μπορούμε να υπολογίσουμε την πιθανότητα της για αυθαίρετο χρόνο.

Για το μοντέλο μας σύμφωνα με τις επιτρεπτές μεταβάσεις και την απεικόνισή τους στον πίνακα 3.2.1 ο πίνακας μετάβασης του συστήματος θα είναι 15x15 και άνω τριγωνικός με όλα τα άνω δεξιά στοιχεία στη μορφή $q_{a,b}$ αν και κάποια θα είναι μηδενικά.

$$Q = \begin{bmatrix} q_{1,1} & q_{1,2} & q_{1,3} & q_{1,4} & q_{1,5} & q_{1,6} & q_{1,7} & q_{1,8} & q_{1,9} & q_{1,10} & q_{1,11} & q_{1,12} & q_{1,13} & q_{1,14} & q_{1,15} \\ 0 & q_{2,2} & q_{2,3} & q_{2,4} & q_{2,5} & q_{2,6} & q_{2,7} & q_{2,8} & q_{2,9} & q_{2,10} & q_{2,11} & q_{2,12} & q_{2,13} & q_{2,14} & q_{2,15} \\ 0 & 0 & q_{3,3} & 0 & q_{3,5} & q_{3,6} & 0 & q_{3,8} & q_{3,9} & q_{3,10} & 0 & q_{3,12} & q_{3,13} & q_{3,14} & q_{3,15} \\ 0 & 0 & 0 & q_{4,4} & q_{4,5} & q_{4,6} & q_{4,7} & q_{4,8} & q_{4,9} & q_{4,10} & q_{4,11} & q_{4,12} & q_{4,13} & q_{4,14} & q_{4,15} \\ 0 & 0 & 0 & 0 & q_{5,5} & q_{5,6} & 0 & q_{5,8} & q_{5,9} & q_{5,10} & 0 & q_{5,12} & q_{5,13} & q_{5,14} & q_{5,15} \\ 0 & 0 & 0 & 0 & 0 & q_{6,6} & 0 & 0 & q_{6,9} & q_{6,10} & 0 & 0 & q_{6,13} & q_{6,14} & q_{6,15} \\ 0 & 0 & 0 & 0 & 0 & 0 & q_{7,7} & q_{7,8} & q_{7,9} & q_{7,10} & q_{7,11} & q_{7,12} & q_{7,13} & q_{7,14} & q_{7,15} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{8,8} & q_{8,9} & q_{8,10} & 0 & q_{8,12} & q_{8,13} & q_{8,14} & q_{8,15} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{9,9} & q_{9,10} & 0 & 0 & q_{9,13} & q_{9,14} & q_{9,15} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{10,10} & 0 & 0 & 0 & q_{10,14} & q_{10,15} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{11,11} & q_{11,12} & q_{11,13} & q_{11,14} & q_{11,15} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{12,12} & q_{12,13} & q_{12,14} & q_{12,15} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{13,13} & q_{13,14} & q_{13,15} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{14,14} & q_{14,15} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Γνωρίζοντας την αρχική μήτρα P_0 μπορούμε σε αυθαίρετο χρόνο k να υπολογίσουμε μέσω του γενικού τύπου $P_k = P_0 Q^k$ την πιθανότητα κατάστασης j από το άθροισμα των στοιχείων της j στήλης του P_k

Το επόμενο ερώτημα ενδιαφέροντος είναι να βρούμε το πλήθος των ιστοριών για αυθαίρετα n και m , βήματα χρόνου και πλήθος στοιχείων ομάδας αντίστοιχα. Προς το σκοπό αυτό θα φανεί χρήσιμο να κατασκευάσουμε ένα γράφο προς την κατάσταση που μας ενδιαφέρει και για ορισμένο πλήθος βημάτων χρόνου. Για το παράδειγμα μας προς την κατάσταση 7 και για 4 βήματα αυτό φαίνεται στο σχήμα 3.2.

Στο σχήμα 3.3 φαίνεται ότι και στον πίνακα μεταβάσεων του σχήματος 3.2 συν ένα επιπλέον βήμα χρόνου. Μπορούμε να εκφράσουμε το ζητούμενο άθροισμα ιστοριών για χρόνο n ως ένα άθροισμα το οποίο έχει όρους τα πλήθη των κλάδων προς κάθε ξεχωριστή κατάσταση. Θα έχουμε δηλαδή :

$$S(n) = f_1(n) + f_2(n) + f_3(n) + \dots + f_m(n) \quad (3.2.4)$$

Γενικά το πλήθος $S(n)$ έχει και κάτω δείκτη της κατάστασης που αναφέρεται. Στα παρακάτω εξετάζουμε την κατάσταση 7 και παραλείπουμε τον δείκτη χάριν απλότητας.

Ο συντελεστής $f_k(n)$ δείχνει πόσο συνεισφέρει ένα στοιχείο k της ομάδας σε χρόνο n . Ο συντελεστής ισούται ακριβώς με τις φορές που εμφανίζεται το στοιχείο. Στο παράδειγμά μας το στοιχείο 2 έχει πλήθος 3. Αυτή η συστάδα των 2 εμφανίζεται 3 φορές όπως φαίνεται στο σχήμα μεταξύ του $n=2$ και $n=3$. Η (3.2.3) θα γραφεί για $n=3$ λοιπόν:

$$S(3) = 10 + 6 + 3 + 1 = 20. \text{ Μια γενική παρατήρηση είναι ότι } f_1(n)=S(n-1)$$

Στη συνέχεια για τους υπόλοιπους συντελεστές από το σχήμα παρατηρούμε ότι ικανοποιούν τις εκφράσεις : $f_2(n)=S(n-1)-f_1(n-1)$ και $f_3(n)= S(n-1)- f_1(n-1)- f_2(n-1)$. Η συμπεριφορά αυτή μας οδηγεί στην υπόθεση ότι γενικά για τους συντελεστές στο άθροισμα θα ισχύει :

$$f_k(n) = S(n - 1) - \sum_{i=1}^{k-1} f_i(n - 1) , \quad \text{για κάθε } k \text{ από } 1 \text{ έως } m \quad (3.2.5)$$

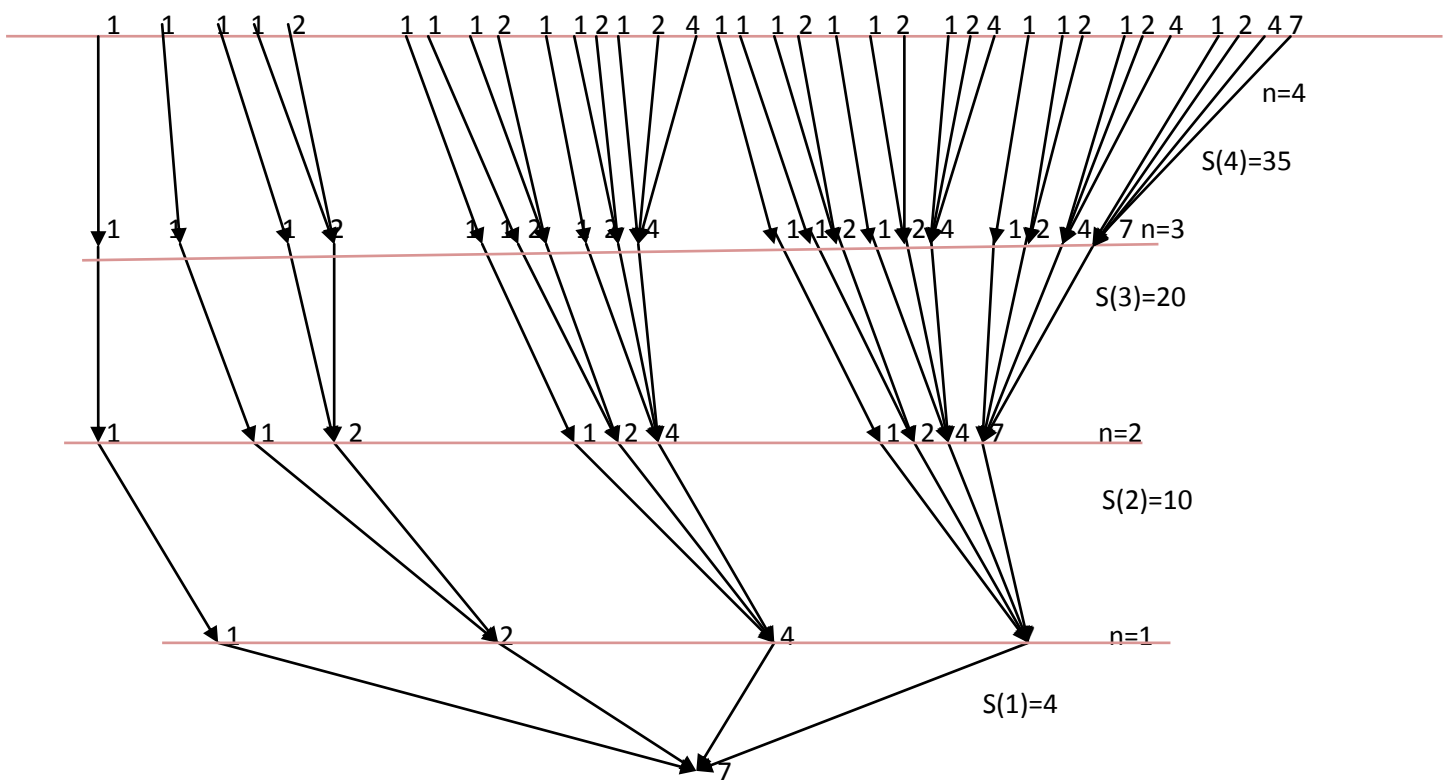
Για $k=1$ θα πρέπει να πάρουμε $S(n-1)$. Η (3.1.4) δίνει $f_1(n) = S(n - 1)$ όπως θα έπρεπε (για $k < 1$ $f_k=0$)

Στις παραπάνω εκφράσεις είναι $S(0)=1$. Εκτελώντας πράξεις καταλήγουμε για το παράδειγμα μας ότι :

$$S(n) = 4S(n - 1) - 6S(n - 2) + 4S(n - 3) - S(n - 4) \quad (3.2.6)$$

Για τυχούσα κατάσταση y με ομάδα m στοιχείων όπως μπορούμε να δείξουμε μπορούμε να εκφράσουμε το πλήθος $S_y(n)$ ως μια γραμμική σχέση των $S(n - 1), (n - 2) \dots$ έως και $(n - m)$.

Υποθέτουμε την λύση λ^n και αντικαθιστούμε στην (3.1.5). Προκύπτει η ρίζα $\lambda=1$ πολλαπλότητας 4. Εκτελώντας τις πράξεις καταλήγουμε για την κατάσταση 7 ότι σε χρόνο n οι δυνατές ιστορίες της έχουν πλήθος $S(n) = \frac{1}{6}n^3 + n^2 + \frac{11}{6}n + 1$



Σχήμα 3.3

Εκτιμούμε ότι το πλήθος των μεταβάσεων θα έχουν αυτή τη μορφή πολυωνύμου για κάθε ομάδα και χρόνο t . Μια περαιτέρω ανάλυση θα μπορούσε να εξηγήσει τις αριθμητικές τιμές των συντελεστών καθώς και τον βαθμό που για παράδειγμα, εδώ είναι κατά ένα μικρότερος από την τάξη της ομάδας.

3.3 Εξάρτηση πιθανότητας μετάβασης από την τοπολογία

Στη συνέχεια θα προσπαθήσουμε να εισάγουμε το χαρακτηριστικό της κατανομής βαθμού στην πιθανότητα μετάβασης. Σημειώνουμε αρχικά ότι η διαδικασία θεραπείας είναι ανεξάρτητη από αυτή της μόλυνσης. Η κατανομή βαθμού επηρεάζει τον σχετικό παράγοντα της μόλυνσης στην έκφραση της πιθανότητας μετάβασης σε χρόνο $\delta t=1$ την οποία εξετάσαμε σε γενική μορφή προηγουμένως. Η πιθανότητα θεραπείας θεωρούμε ότι εξαρτάται κυρίως από άλλα θέματα και όχι από τη δομή του δικτύου.

Αρχικά θα βρούμε μια έκφραση για την πιθανότητα μόλυνσης ενός ευάλωτου κόμβου. Θεωρούμε τις γνωστές πυκνότητες πληθυσμών $s_k(t)$, $\rho_k(t)$ και $r_k(t)$ των ευάλωτων, μολυσμένων και απομονωμένων κόμβων. Έχουμε σημειώσει και το δείκτη βαθμού, ώστε να κάνουμε εμφανή την κατανομή τους ως προς k . Οι ίδιες αυτές οι ποσότητες είναι ίσες με τις πιθανότητες ένας κόμβος να ανήκει στον αντίστοιχο πληθυσμό. Όπως έχουμε αναφέρει και προηγουμένως, η πιθανότητα μια ακμή να οδηγεί σε έναν κόμβο με s ακμές είναι ανάλογη του $sP(s)$. Επομένως η πιθανότητα μια τυχαία επιλεγμένη ακμή να οδηγεί σε έναν μολυσμένο κόμβο δίνεται από :

$$P_{\rho,link} = \frac{\sum_k k p(k) \rho_k(t)}{\sum_s s P(s)} = \frac{\sum_k k p(k) \rho_k(t)}{\langle k \rangle} \quad (3.3.1)$$

Η ποσότητα $\rho_k(t)$ είναι το ποσοστό των k ακμών που οδηγούν σε μολυσμένους κόμβους καθώς και σε κακόβουλους οι οποίοι είναι ικανοί να μολύνουν. Όπως είναι επόμενο η ποσότητα αυτή πρέπει να αυξάνεται με την πιθανότητα $\rho = \frac{P}{N}$ ένας κόμβος να είναι μολυσμένος, όπου P οι μολυσμένοι και N το μέγεθος του δικτύου. Επιπλέον θα πρέπει να εισάγουμε την εξάρτηση από k . Διαισθητικά περιμένουμε το $\rho_k(t)$ να αυξάνεται με το k .

Έστω ένας ευάλωτος κόμβος. Μεταφέροντας τον παραπάνω συλλογισμό από την περίπτωση μιας ακμής στο σύνολο των ευάλωτων ακμών ενός κόμβου παρατηρούμε ότι τα ενδεχόμενα της (3.3.1) μπορούν να θεωρηθούν ανεξάρτητα αλλά όχι και ξένα μεταξύ τους. Εφόσον θα αναζητήσουμε την πιθανότητα ένας ευάλωτος κόμβος να μολυνθεί μέσω τουλάχιστον μιας ακμής του, δεν θα ήταν τόσο εύχρηστο μάλλον να την περιγράψουμε με τον γενικευμένο αθροιστικό νόμο, οπότε θα κάνουμε χρήση της ανεξαρτησίας. Η πιθανότητα καμία ακμή να μην μολυνθεί από τους συνδεδεμένους k κόμβους είναι

$$(1 - \lambda P_{\rho,link})^k \quad (3.3.2)$$

όταν ο ευάλωτος κόμβος έχει k ακμές. Έχουμε εισάγει εδώ και τον μικροσκοπικό ρυθμό λ μέσω του οποίου μια ακμή που οδηγεί σε μολυσμένο κόμβο αποκτά τελικά την ασθένεια με τον ρυθμό λ . Η συνολική πιθανότητα λοιπόν να μην μολυνθεί ένας ευάλωτος κόμβος προκύπτει με άθροιση

$\sum_k (1 - \lambda P_{\rho,link})^k p(k)$ και τελικά η ζητούμενη, συμπληρωματική, πιθανότητα ο κόμβος να μολυνθεί μέσω τουλάχιστον μιας ακμής του είναι

$$P_1 = 1 - \sum_k \left(1 - \lambda \frac{\sum_k k p(k) (\rho_k(t) + h)}{\langle k \rangle}\right)^k p(k) \quad (3.3.3)$$

Όπως βλέπουμε δεν έχουμε αντικαταστήσει ακριβώς την (3.3.1) η οποία θα ίσχυε στο γνωστό μοντέλο SIR όπου, όταν η πυκνότητα μολυσμένων είναι μηδέν, η παραπάνω πιθανότητα πρέπει να

είναι μηδέν. Στην περίπτωση κακόβουλα ενεργών κόμβων πρέπει να θεωρήσουμε αντί της ποσότητας $\rho_k(t)$ έναν μικρό όρο ακόμα, έστω h , ο οποίος αντιπροσωπεύει την πυκνότητά τους και δίνει την δυνατότητα να μολύνουν, ενώ φαινομενικά συνυπάρχουν στο σύνολο των ευάλωτων ακόμα και όταν $\rho_k(t) = 0$. Στην (3.3.3) έχουμε λάβει υπόψη μας αυτή την δυνατότητα. Η ακριβής έκφραση του όρου h είναι αντικείμενο των υποθέσεών μας σχετικά με το ρόλο του σε συγκεκριμένα παραδείγματα. Μπορεί να εξαρτάται ή μη από το χρόνο και την κατανομή. Μπορεί επίσης να θεωρηθεί ότι είναι ανάλογος με τον πληθυσμό των ευάλωτων και τελικά ίσος με ένα μικρό ποσοστό αυτών.

Συμβολίζουμε με λ τον μικροσκοπικό ρυθμό μετάδοσης της ασθένειας μεταξύ ενός ευάλωτου και ενός μολυσμένου κόμβου μέσω μιας ακμής. Στην (3.3.3) φαίνεται ότι η πιθανότητα μόλυνσης αυξάνεται με τον ρυθμό λ . Επίσης όπως ήδη θα περιμέναμε, κατανομές όπως αυτή της απουσίας κλίμακας δεν ωφελούν την διάδοση της επιδημίας σε σχέση με άλλες που θα έδιναν περισσότερη συνδεσιμότητα στο δίκτυο.

Η πιθανότητα να μολυνθούν x νέοι κόμβοι σε χρονικό διάστημα ενός βήματος χρόνου μπορεί να δοθεί από τη διωνυμική κατανομή

$$\binom{S_k(t)}{x} P_1^x (1 - P_1)^{S_k(t)-x} \quad (3.3.4)$$

Όπου $S_k(t) = s_k(t)N$ ο απόλυτος αριθμός των ευάλωτων στο συνολικό αριθμό κόμβων N του δικτύου.

Για την πιθανότητα να θεραπευτούν y κόμβοι σε διάστημα δt όταν έχουμε $P_k(t)$ μολυσμένους έχουμε

$$\binom{P_k(t)+x}{y} P_2^y (1 - P_2)^{(P_k(t)+x)-y} \quad (3.3.5)$$

όπου P_2 η πιθανότητα ενός μολυσμένου κόμβου να θεραπευτεί.

$$P_2 = (\rho_k(t) + \frac{x}{N})\theta \quad (3.3.6)$$

όπου θ ο μικροσκοπικός ρυθμός θεραπείας. Στην περίπτωση αυτή της θεραπείας ο δείκτης k δεν παίζει κάποιο ρόλο, το $\rho_k(t)$ ισούται απλώς με την πιθανότητα ένας κόμβος να είναι μολυσμένος.

Θεωρώντας τη διωνυμική κατανομή αποδεχόμαστε αυτόματα ότι η διαδικασία μετάδοσης της ασθένειας γίνεται ανεξάρτητα και με την ίδια πιθανότητα για τους x νέους μολυσμένους. Η διωνυμική κατανομή προσφέρει μια μέτρηση πιθανότητας μακροσκοπικά πλέον. Πρέπει να σημειώσουμε εδώ ότι η διωνυμική κατανομή χρησιμοποιείται εδώ ως τμήμα μιας γενικότερης

κατανομής. Έτσι δεν χρειάζεται να απαιτήσουμε να έχει άθροισμα μονάδα. Εξάλλου για $x=0$ καθώς και για $y=0$ δεν χρησιμοποιούμε τις εξισώσεις (3.3.4) και (3.3.5). Όπως θα δείξουμε παρακάτω η γενικότερη κατανομή μας ικανοποιεί το αίτημα να έχει άθροισμα μονάδα.

Οι εξισώσεις (3.3.3) έως και (3.3.6) αποτελούν τα δομικά στοιχεία ώστε να έχουμε μια περιγραφή της μετάβασης σε συνδυασμό με την τοπολογία. Οι εξισώσεις αυτές δίνουν μια εικόνα για την μέτρηση πιθανότητας αλλά θα πρέπει να χρησιμοποιηθούν κατάλληλα ώστε να ενσωματωθούν στο μοντέλο μας με συνέπεια.

Θα πρέπει να διακρίνουμε περιπτώσεις ώστε να χρησιμοποιήσουμε τις εκφράσεις (3.3.3) έως και (3.3.6) ικανοποιητικά στο μοντέλο επιδημίας. Η περιγραφή της πιθανότητας μετάβασης από την κατάσταση a στην b θα ορίζεται παραμετρικά

$$q_{a,b} = \begin{cases} \binom{S(t)}{x} P_1^x (1 - P_1)^{S(t)-x} \binom{P(t)+x}{y} P_2^y (1 - P_2)^{P(t)+x-y} & , \text{όταν } x \text{ και } y \neq 0 \\ \binom{P(t)+x}{y} P_2^y (1 - P_2)^{P(t)+x-y} & , \text{όταν } x = 0 \\ \binom{S(t)}{x} P_1^x (1 - P_1)^{S(t)-x} & , \text{όταν } y = 0 \end{cases} \quad (3.3.8)$$

$$P_1 = s(t) \left[1 - \sum_k \left(1 - \lambda \frac{\sum_k k p(k) (\rho_k(t) + h)}{\langle k \rangle} \right)^k p(k) \right] \quad P_2 = (\rho_k(t) + \frac{x}{N}) \theta \quad (3.3.9)$$

Έχοντας μια έκφραση για την πιθανότητα μετάβασης σε κάθε περίπτωση εκτός της παραμονής στην ίδια κατάσταση, η εύρεση της τελευταίας αφενός δεν μπορεί να βρεθεί από τις παραπάνω εξισώσεις, αφετέρου πρέπει να κανονικοποιήσουμε τις πιθανότητες ώστε η συνολική πιθανότητα μετάβασης από κάθε κατάσταση προς όλες τις δυνατές να είναι μονάδα. Συνεπώς η πιθανότητα παραμονής σε μία κατάσταση είναι

$$q_{a,a} = 1 - \sum_{b \neq a} q_{a,b} \quad (3.3.10)$$

Όταν ο πληθυσμός των ευάλωτων είναι μηδέν, προφανώς η γέννηση μολυσμένου είναι αδύνατη και θα είναι $P_1=0$. Η πυκνότητα $s(t)$ δηλώνει ακόμα πόσο πιθανό είναι να γεννηθεί ένας μολυσμένος. Θα είναι πιο πιθανή η γέννηση ενός μολυσμένου όταν υπάρχουν πολλοί ευάλωτοι παρά όταν είναι λίγοι. Στις καταστάσεις όπου $s(t)=0$, εκτός από μια παροδική παραμονή σε μια κατάσταση, αυτό που απομένει να αλλάξει με την πάροδο του χρόνου είναι η θεραπεία μολυσμένων. Επομένως οι δυνατές αυτές μεταβάσεις θα εκφραστούν από την (3.3.5). Από αυτή την εξίσωση θα περιγράφεται η μετάβαση και στην περίπτωση που η νέα κατάσταση δεν απαιτεί μόλυνση κάποιου ατόμου αλλά

απλώς θεραπεία. Και στις δύο αυτές περιπτώσεις είναι $x=0$. Αντίστοιχες παρατηρήσεις ισχύουν όταν στην μετάβαση συμβαίνει μόνο μόλυνση και όχι θεραπεία. Τότε η πιθανότητα περιγράφεται από την (3.3.4). Σε αυτή την περίπτωση είναι $\gamma=0$.

Για την μοναδική κατάσταση όπου $P(t) = 0$ και $S(t) = 0$ είναι προφανές ότι πρόκειται για την τελική κατάσταση (έχουν όλοι θεραπευθεί) όπου δεν μπορεί να μεταβεί σε καμία άλλη παρά να παραμείνει στον εαυτό της με πιθανότητα μονάδα.

Έχει ενδιαφέρον να παρατηρήσουμε ότι μέσω της διωνυμικής κατανομής, την οποία επιλέξαμε για να εκφράσουμε την επέκταση μόλυνσης, οι πιθανότητες γενικά μειώνονται προς μεγάλες διαφορές καταστάσεων, το οποίο πρακτικά σημαίνει ότι είναι όλο και λιγότερο πιθανό να μολύνεται σε ένα βήμα χρόνου μεγάλος αριθμός κόμβων.

4. Αποτελέσματα - Συμπεράσματα

Στην τελευταία αυτή ενότητα παρουσιάζουμε τα πιο κύρια αποτελέσματα για το μοντέλο επιδημίας το οποίο θεωρήσαμε. Παράλληλα, αναφέρουμε τις παρατηρήσεις και τα συμπεράσματά μας εξετάζοντας την εξάρτηση της επιδημίας από την κατανομή βαθμού καθώς και από παραμέτρους του μοντέλου. Στην υπό-ενότητα 4.4 θα δούμε δυο πιθανές παραλλαγές του μοντέλου επιδημίας.

4.1 Γενικά χαρακτηριστικά της μετάβασης της επιδημίας

Συνεχίζοντας με το παράδειγμά μας των $N=4$ κόμβων, αρχικά θα κατασκευάσουμε τον πίνακα Q . Υποθέτουμε αρχικά την επόμενη κατανομή βαθμού για το δίκτυό μας.

$$p(k) = \begin{cases} 0,7, & \text{για } k = 1 \\ 0,2, & \text{για } k = 2 \\ 0,1, & \text{για } k = 3 \end{cases}$$

Οπότε προκύπτει $\langle k \rangle = 1,4$ και $\rho_k(t) = \frac{k \cdot P}{N}$. Θεωρούμε $\lambda = \theta = 0,1$ για τους ρυθμούς μόλυνσης και θεραπείας. Για τον όρο h θεωρούμε ότι είναι 0,25 δηλαδή ένας από τους 4 κόμβους είναι κακόβουλος. Όπως φαίνεται στον πίνακα 4.4.1 από μια κατάσταση a γίνεται μετάβαση στην b με πιθανότητα ίση με το στοιχείο $q_{a,b}$.

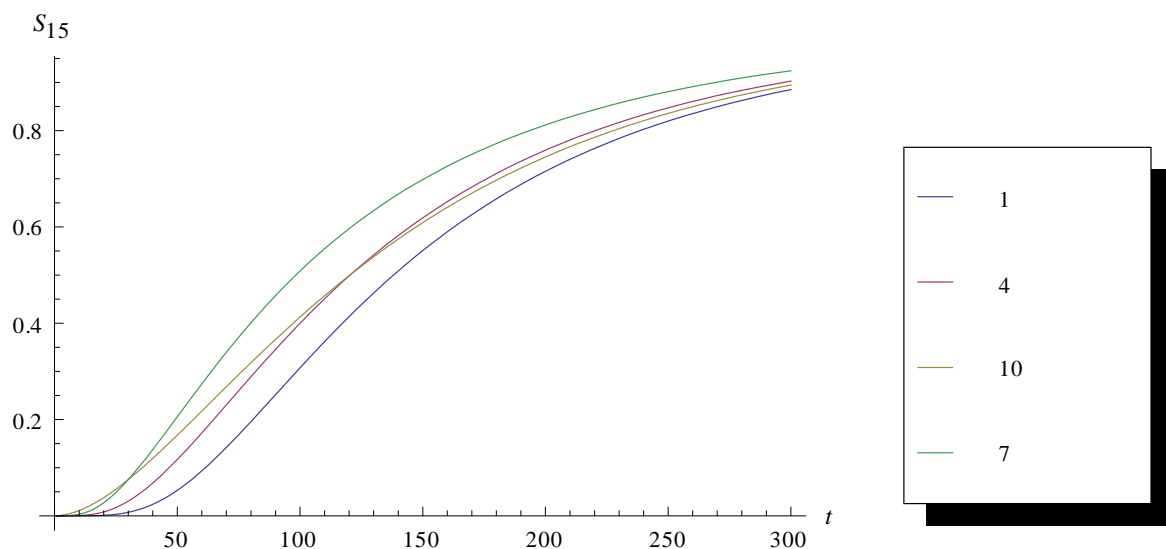
8,64	1,24	3,12	6,72	6,39	1,68	1,61	3,10	2,51	6,79	1,44	4,22	7,03	5,21	1,44
487	811	028	776	137	194	177	292	588	968	800	238	730	282	800
9E-01	4E-01	5E-03	5E-03	7E-04	1E-05	7E-04	2E-05	3E-06	3E-08	6E-06	4E-07	7E-08	0E-09	6E-10
0	8,57	2,50	1,03	9,78	2,57	3,96	7,62	6,18	1,67	5,07	1,48	2,46	1,82	5,07
	098	000	004	540	510	045	450	202	081	591	013	689	733	591
	2E-01	0E-02	3E-01	8E-03	7E-04	9E-03	3E-04	9E-05	9E-06	6E-05	7E-05	5E-06	0E-07	6E-09
0	0	9,21	0	7,40	7,40	0	1,97	1,87	4,94	0	1,76	3,39	2,74	7,42
		915		419	419		779	890	447		100	021	882	925

		1E-01		8E-02	8E-02		1E-03	1E-04	7E-06		9E-05	7E-06	4E-07	5E-09
0	0	0	8,263293E-01	9,500000E-02	2,500000E-03	6,189035E-02	1,191486E-02	9,660697E-04	2,610999E-05	1,021893E-03	2,979841E-04	4,966402E-05	3,678816E-06	1,021893E-07
0	0	0	0	9,214199E-01	2,500000E-02	0	4,814919E-02	4,574173E-03	1,203730E-04	0	6,092943E-04	1,172987E-04	9,510703E-06	2,570460E-07
0	0	0	0	0	9,647303E-01	0	0	3,408740E-02	8,521849E-04	0	0	3,008328E-04	2,857911E-05	7,520819E-07
0	0	0	0	0	0	7,651164E-01	1,925156E-01	1,560938E-02	4,218750E-04	1,959720E-02	5,714550E-03	9,524250E-04	7,055000E-05	1,959720E-06
0	0	0	0	0	0	0	8,831831E-01	9,500000E-02	2,500000E-03	0	1,598353E-02	3,077080E-03	2,494930E-04	6,743053E-06
0	0	0	0	0	0	0	0	9,614547E-01	2,500000E-02	0	0	1,234194E-02	1,172485E-03	3,085486E-05
0	0	0	0	0	0	0	0	0	9,911109E-01	0	0	0	8,672266E-03	2,168066E-04
0	0	0	0	0	0	0	0	0	0	6,561000E-01	2,916000E-01	4,860000E-02	3,600000E-03	1,000000E-04

0	0	0	0	0	0	0	0	0	0	0	7,91 453 1E- 01	1,92 515 6E- 01	1,56 093 8E- 02	4,21 875 0E- 04
0	0	0	0	0	0	0	0	0	0	0	0	9,02 500 0E- 01	9,50 000 0E- 02	2,50 000 0E- 03
0	0	0	0	0	0	0	0	0	0	0	0	0	9,75 000 0E- 01	2,50 000 0E- 02
0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Πίνακας 4.4.1 Πίνακας μετάβασης Q

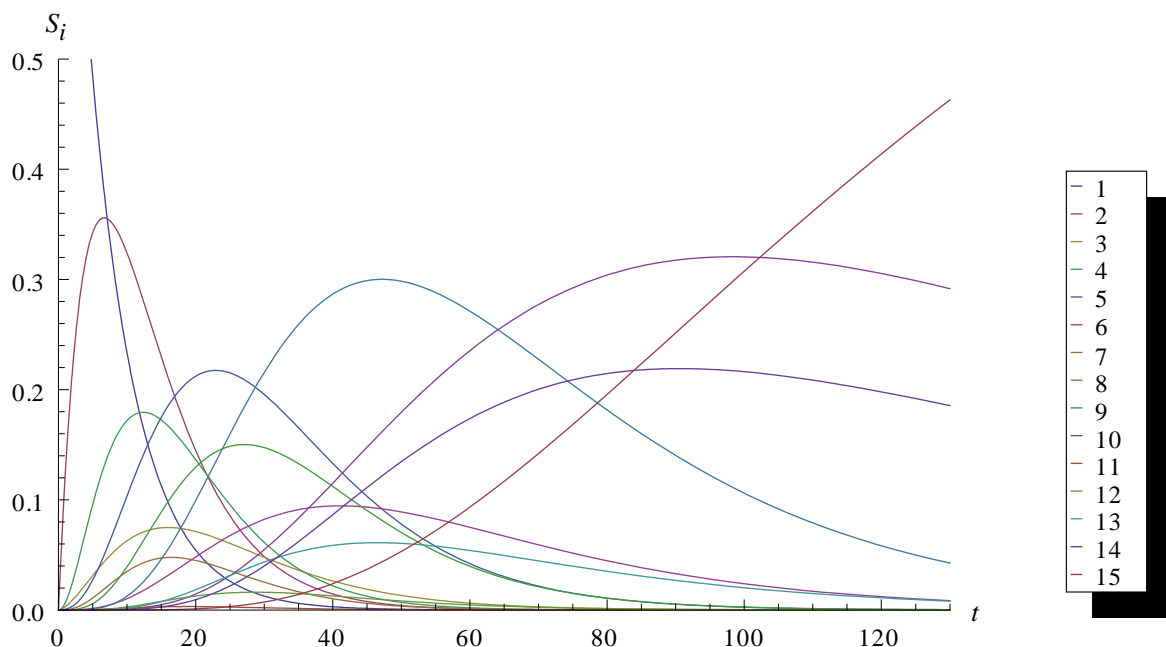
Ένα πρώτο χαρακτηριστικό που μας ενδιαφέρει, είναι το πώς μεταβάλλεται η τελική κατάσταση του συστήματος με το χρόνο. Στην πρώτη γραφική παράσταση παρουσιάζεται αυτή η μεταβολή για διάφορες αρχικές καταστάσεις.



Σχήμα 4.1.1 :Μεταβολή τελικής κατάστασης σε συνάρτηση με το χρόνο για πιθανότητες αρχικών καταστάσεων $S(0)_1, S(0)_4, S(0)_{10}$ και $S(0)_7 = 1$

Όπως φαίνεται, η τελική κατάσταση αυξάνεται μετά από ένα αρχικό μεταβατικό στάδιο εκθετικά προς την μονάδα. Το σημείο αυτό θα το παρουσιάσουμε και αργότερα συγκρίνοντας τις καμπύλες με μια εκθετική εξίσωση. Στο σχήμα 4.4.1 παρατηρούμε ότι το σύστημα προσεγγίζει την τελική κατάσταση πιο γρήγορα ή πιο αργά ανάλογα με την αρχική κατάσταση που βρίσκεται. Αν και στο σχήμα φαίνονται κάποιες μόνο περιπτώσεις, με μια συνολική γραφική παράσταση θα βλέπαμε ότι ο πιο μεγάλος ρυθμός αύξησης για αρχικές καταστάσεις 1 έως 10 προκύπτει για την αρχική κατάσταση 7. Όπως μπορούμε να παρατηρήσουμε από ένα τέτοιο συνολικό σχήμα ή από διαδοχική σύγκριση παραστάσεων, το σύστημα τείνει στην τελική του κατάσταση με ρυθμό που αυξάνεται με το πλήθος των αρχικά μολυσμένων κόμβων για σταθερό αριθμό ευάλωτων. Καθώς οι τελευταίοι μειώνονται, ο ρυθμός αύξησης αυξάνεται. Είναι αναμενόμενο το σύστημα να τείνει πιο γρήγορα στην τελική κατάσταση όταν αυτό ξεκινά από μια κατάσταση ήδη προετοιμασμένο με λιγότερες επόμενες γεννήσεις μολυσμένων. Τα προηγούμενα αποτελούν γενικές παρατηρήσεις αφού οι καμπύλες μπορούν να τέμνονται (περίπτωση της 4 και 10 στο σχήμα 4.4.1), γεγονός που δηλώνει ότι καθώς περνά ο χρόνος μπορούμε να έχουμε αποκλίσεις από τα παραπάνω χαρακτηριστικά.

Ένα σχήμα που φαίνονται όλες οι πιθανότητες σε σχέση με το χρόνο θα βοηθούσε πολύ για μια εποπτική εικόνα της διάδοσης της επιδημίας. Στο επόμενο σχήμα φαίνεται αυτό για ένα κατάλληλο διάστημα χρόνου, ώστε να φανεί όσο το δυνατόν καλύτερα το μέγιστο πλήθος πιθανοτήτων.



Σχήμα 4.1.2 :Μεταβολή πιθανότητας των καταστάσεων S_i για $i=1,2$, έως και 15 σε συνάρτηση με το χρόνο για την ίδια πιθανότητα αρχικής κατάστασης $S(0)_1 = 1$

Αν και το σχήμα που προκύπτει είναι αρκετά μπερδεμένο γενικά, μπορούμε να κάνουμε αρκετές παρατηρήσεις. Οι διάφορες πιθανότητες έχουν ένα μέγιστο στον κατάλληλο χρόνο, αλλά πολλές από αυτές επικαλύπτονται μερικά ή ολικά από άλλες. Σε κάθε σημείο του χρόνου το άθροισμα των πιθανοτήτων είναι μονάδα, οπότε αυτές που έχουν μεγάλη τιμή περιορίζουν τις υπόλοιπες σε μικρότερες. Αν “ρωτήσουμε” το σύστημα πού είναι πιο πιθανό να βρίσκεται σε μια συγκεκριμένη χρονική στιγμή, αυτό θα βρεθεί από την πιθανότητα κατάστασης που έχει τη μεγαλύτερη τιμή τη στιγμή αυτή. Από την παρατήρηση αυτή και το σχήμα συμπεραίνουμε ότι η πιο πιθανή ιστορία της επιδημίας δεν περνάει από όλες τις καταστάσεις αλλά από αυτές που οριοθετούν προς τα κάτω τις υπόλοιπες. Για την περίπτωση του παραδείγματός μας, εκτός από την αρχική και την τελική, οι πιο πιθανές καταστάσεις που θα περάσει το σύστημα είναι οι 2,5,9 και 10 με την σειρά αυτή στο χρόνο. Εκ των υστέρων βλέπουμε ότι παρόλο που δώσαμε στο σύστημα τόση μεγάλη ελευθερία μεταβάσεων, το πιο πιθανό είναι ότι αυτό θα ακολουθήσει ένα μικρό αριθμό από αυτές. Το αντίστροφο είναι αρκετά αναμενόμενο αφού οι πιθανότητες δεν θα επικαλύπτονται τόσο οπότε διαδοχικά θα αυξάνονται σε ξεχωριστούς χρόνους.

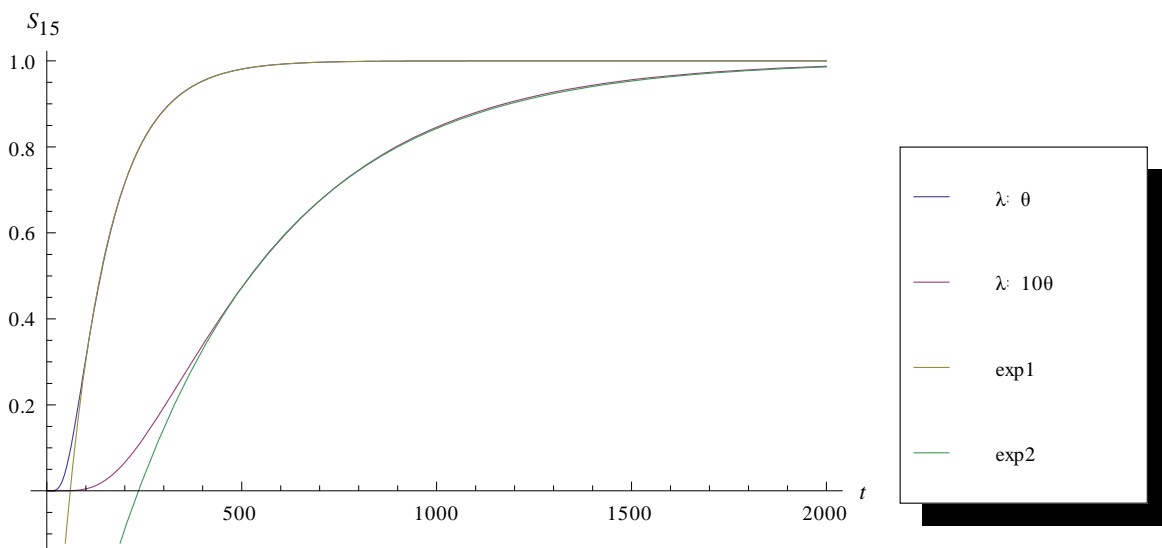
Η κάθε καμπύλη εκτός από την προφανή της ερμηνεία για την πιθανότητα της κατάστασης σε χρόνο t έχει και μια άλλη, αυτή της πιθανότητας παραμονής στον εαυτό της. Καθώς η πιθανότητα μιας κατάστασης είναι αύξουσα αρχικά, είναι πιο πιθανό να μείνει το σύστημα στην κατάσταση αυτή, όταν βρίσκεται στην κατάσταση πριν από το χρόνο μέγιστου σε σχέση με το ίδιο ενδεχόμενο παραμονής σε αυτήν για χρόνο μεγαλύτερο του χρόνου μέγιστου.

Το σχήμα 4.1.2 μας δείχνει σε γενικές γραμμές 3 χρονικά διαστήματα για την επιδημία. Στην αρχή υπάρχουν λίγες καταστάσεις αρκετά πιθανές, ενώ αρκετές αρχίζουν να αυξάνουν. Στη συνέχεια έχουμε μια περίοδο όπου ένα δεύτερο σύνολο καταστάσεων αρκετά πιο μεγάλο σε πλήθος κυριαρχεί αν και με μικρότερο πλάτος. Στο τελευταίο στάδιο υπάρχουν λίγες καταστάσεις πιθανές και η τελευταία τελική κατάσταση έχει αυξηθεί σημαντικά. Τελικά η διάρκεια της επιδημίας, εφόσον την ορίσουμε ως το χρόνο που η τελική κατάσταση έχει μια πιθανότητα πολύ κοντά στη μονάδα, μπορούμε να υποθέσουμε και να περιμένουμε ότι θα εξαρτάται από τους ρυθμούς μόλυνσης και θεραπείας

4.2 Επίδραση των ρυθμών μόλυνσης και θεραπείας-Εκθετική αύξηση της πιθανότητας τελικής κατάστασης

Έχουμε ήδη υποθέσει ότι οι ρυθμοί μόλυνσης και θεραπείας θα αμβλύνουν την διάρκεια της επιδημίας. Όπως θα δούμε, κάτι τέτοιο καθώς και συγγενικές παρατηρήσεις πράγματι συμβαίνουν στα αποτελέσματά μας.

Θα υποθέσουμε τώρα τα ίδια δεδομένα με το προηγούμενο παράδειγμα εκτός από τον ρυθμό θεραπείας όπου θα είναι τώρα $\theta=10\lambda$. Αρχικά συγκρίνουμε την πιθανότητα της τελικής κατάστασης με την προηγούμενη περίπτωση $\theta=\lambda$. Στο επόμενο σχήμα φαίνεται αυτή η αλλαγή λόγω του θ στον τρόπο που το σύστημα τείνει στην τελική κατάσταση. Επιπλέον εδώ κάνουμε μια προσπάθεια περιγραφής μέσω της εκθετικής συνάρτησης και τελικά βλέπουμε ότι όντως έχουμε μια εκθετική συμπεριφορά στην αύξηση της πιθανότητας.

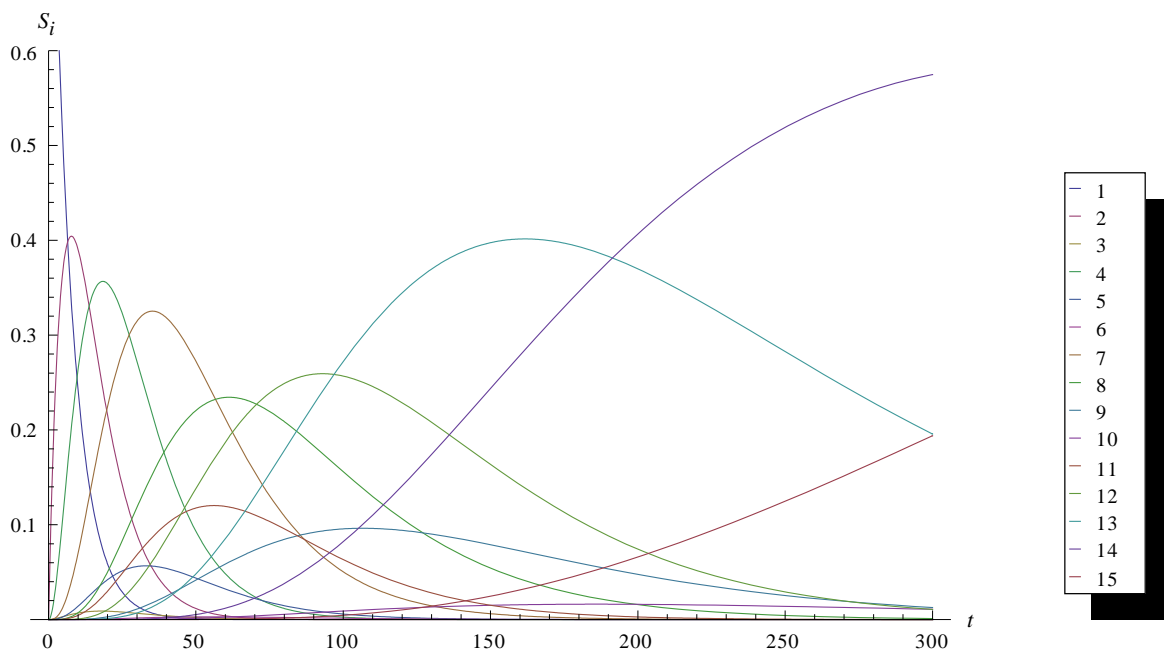


Σχήμα 4.2.1 : Εκθετική μεταβολή πιθανότητας της τελικής κατάστασης με το χρόνο για $\lambda=\theta$ και για $\lambda=10\theta$

Όπως φαίνεται ξεκάθαρα από το σχήμα 4.2.1 όταν ο ρυθμός θεραπείας είναι μικρότερος, το σύστημα θέλει περισσότερο χρόνο για να φτάσει στην τελική κατάσταση. Σε συνδυασμό με την παρατήρηση της υπό-ενότητας 4.1 μπορούμε εδώ να διακρίνουμε την εκθετική συμπεριφορά για δύο διαφορετικές εκδοχές της τελικής κατάστασης. Όπως παρατηρούμε, η εκθετική αύξηση ταυτίζεται με την κατάσταση 15. Συγκεκριμένα, όταν $\lambda=\theta$ η εκθετική συνάρτηση που προσεγγίζει την τελική κατάσταση είναι $exp_1(t) = 1 - e^{-0.009(t-60)}$ για $t>80$. Η ακριβής τιμή του t που αρχίζει

να ισχύει η εκθετική συμπεριφορά, εξαρτάται από την προσέγγισή μας για το πόσο πρέπει να διαφέρουν οι τιμές των δύο καμπυλών. Για την περίπτωση $\lambda=10\theta$ η εκθετική συνάρτηση που θεωρούμε προκύπτει $\exp_2(t) = 1 - e^{-0.00242(t-235)}$ για $t>400$. Οι τιμές των παραμέτρων για τις δύο εκθετικές συναρτήσεις προέκυψαν έπειτα από την προσπάθεια να ταυτίσουμε τις αντίστοιχες καμπύλες για το μεγαλύτερο τμήμα τους. Μιας και η προσέγγιση αυτή είναι γραφική δεν θα εξάγουμε κάποιο αυστηρό συμπέρασμα για τη σύνδεση των τιμών με παραμέτρους του μοντέλου επιδημίας. Συγκρίνοντας τις δύο εκθετικές αυξήσεις με τους ρυθμούς θεραπείας μπορούμε να διακρίνουμε έστω και αδρά ότι ο ρυθμός που η τελική κατάσταση τείνει στην μονάδα μεταβάλλεται αντιστρόφως ανάλογα και αντίστοιχα ανάλογα για τη φάση κατά έναν παράγοντα περίπου 4 για αύξηση του ρυθμού θεραπείας κατά 10 φορές.

Στη συνέχεια φαίνεται η μεταβολή της πιθανότητας όλων των καταστάσεων για την περίπτωση αυτή $\lambda=10\theta$.



Σχήμα 4.2.2 :Μεταβολή πιθανότητας των καταστάσεων S_i για $i=1,2$, έως και 15 σε συνάρτηση με το χρόνο για την ίδια πιθανότητα αρχικής κατάστασης $S(0)_1 = 1$ και $\lambda=10\theta$

Το σχήμα 4.2.2 μπορεί να συγκριθεί με το αντίστοιχο 4.1.2 και να διακριθούν πρόσθετα αποτελέσματα. Εκτός από την αύξηση στην διάρκεια επιδημίας που ήδη παρατηρήσαμε, φαίνεται πως οι πιθανότητες των διαφόρων καταστάσεων επικαλύπτονται λιγότερο πλέον ως αποτέλεσμα του μειωμένου ρυθμού θεραπείας. Στο σύστημα πλέον δίνεται ο χρόνος να περάσει σε περισσότερες καταστάσεις με μέγιστη πιθανότητα. Παρατηρώντας όπως και προηγουμένως την πιο

πιθανή ιστορία της επιδημίας έχουμε ότι αυτή θα περνάει από τις καταστάσεις 2,4,7,8,12,13 και 14 ,εκτός των 1 και 15 που είναι η αρχική και τελική κατάσταση.

Στην αντίθετη πλευρά βρίσκονται οι καταστάσεις 3 ,6 και 10 όπου για αυτή την περίπτωση ρυθμού θεραπείας παρουσιάζουν την ελάχιστη πιθανότητα να συμβούν. Πράγματι αυτό που παρατηρούμε όταν μεταβάλλεται ο ρυθμός θεραπείας είναι ότι το σύνολο καταστάσεων με τη μικρότερη πιθανότητα αλλάζει. Αυτό δεν είναι αρχικά αναμενόμενο, αν δούμε την περίπτωση $\lambda=\theta$ όπου εκεί η κατάσταση 11 ήταν η πιο μικρή (δεν φαίνεται στο σχήμα 4.1.2) και θα δικαιολογούσαμε την τιμή αυτή όχι τόσο από τις τιμές των παραμέτρων αλλά ενδεχομένως από την φύση της κατάστασης αυτής. Θα μπορούσαμε να υποθέσουμε ότι είναι γενικά απίθανο το σύστημα να κάνει μεταβάσεις μόνο προς την αύξηση μολυσμένων χωρίς καμία θεραπεία (κατάσταση 11 = 4 μολυσμένοι, 0 θεραπευμένοι) και να αποδώσουμε την μικρή πιθανότητα της 11 στη φύση του μοντέλου και όχι στις τιμές των παραμέτρων. Η παρατήρησή μας αυτή αναιρείται εν μέρει μόνο, αν σκεφτούμε ότι όταν μειώνεται ο ρυθμός θεραπείας η κατάσταση με μηδέν θεραπευμένους πρέπει να αυξηθεί. Παρόλα αυτά όμως, αυξάνεται τόσο ώστε να γίνει μια από τις πιο πιθανές.

4.3 Επίδραση της τοπολογίας στην Επιδημία

Στην προσπάθειά μας να εξετάσουμε την επίδραση της τοπολογίας, το παράδειγμα των 4 κόμβων φαίνεται υπερβολικά μικρό και απλοϊκό ώστε να έχουμε μια ρεαλιστική εικόνα του τι συμβαίνει. Θα υπενθυμίσουμε εδώ ότι η κατανομή βαθμού δεν χρειάζεται να συμφωνεί με μία από τις δυνατές φυσικές τοπολογίες του δικτύου μας αφού τα συμπεράσματά μας προορίζονται για το λογικό δίκτυο των κόμβων, αυτό που δημιουργούν οι κόμβοι λόγω των δικτυακών τους εφαρμογών. Σε αυτή επίσης την περίπτωση η ιδιότητα απουσίας κλίμακας έχει νόημα. Όπως έχουμε ήδη παρατηρήσει η φυσική συνδεσιμότητα δεν θα παίξει τόσο ρόλο στο στην εξάπλωση της επιδημίας.

Στα μεγάλα δίκτυα η υπόθεσή μας ότι σε κάθε βήμα χρόνου μπορεί να μολυνθεί το μέγιστο πλήθος ευάλωτων φαίνεται υπερβολική. Στο δίκτυο των 4 κόμβων του παραδείγματός μας αντίθετα, φαίνεται αρκετά ρεαλιστικό. Μπορούμε να επιτρέψουμε ως ένα όριο το πλήθος των μεταβάσεων στο σύστημα με τρόπο που να εξαρτάται από τον βαθμό κόμβου όπως θα δούμε στην επόμενη υπό-ενότητα.

Υποθέσουμε μια κατανομή βαθμού για το δίκτυο των 4 κόμβων με έντονο το χαρακτηριστικό της απουσίας κλίμακας. Η προσέγγισή μας ακολουθεί αυτή της υπό-ενότητας 4.1 όπου ενδιαφερόμαστε

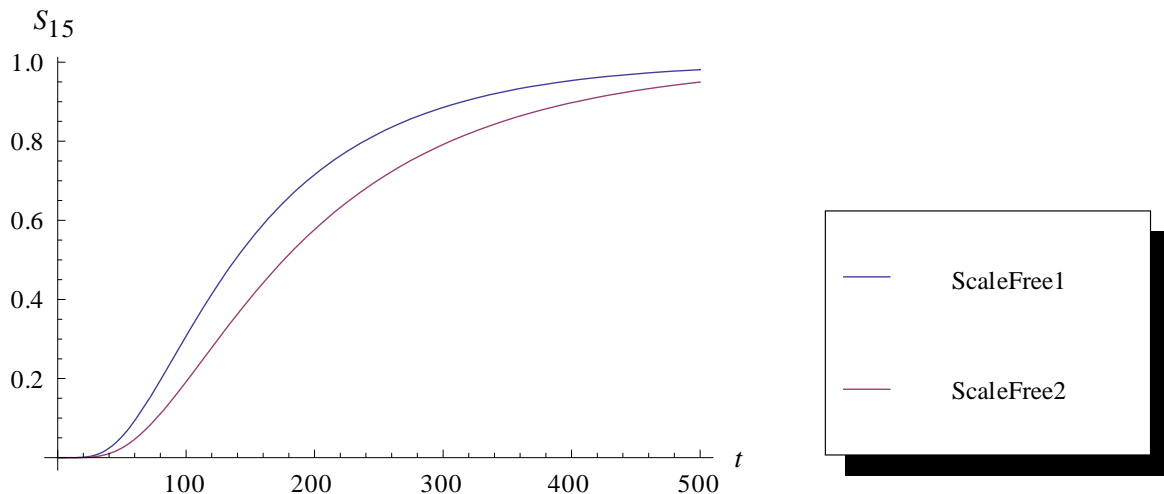
για την ιδιότητα απουσίας κλίμακας στην κατανομή βαθμού και εδώ εξετάζουμε μια πιο απότομη κατανομή.

Μπορούμε να φανταστούμε δίκτυα όπου οι χρήστες για παράδειγμα σε έναν χώρο εργασίας γενικά έχουν λίγες και συγκεκριμένες ταυτόχρονες εφαρμογές στο δίκτυο αλλά για ένα πολύ μικρό ποσοστό του χρόνου μπορεί να κάνουν χρήση και άλλων εφαρμογών.

Έστω η παρακάτω κατανομή βαθμού

$$p(k) = \begin{cases} 0,9, & \text{για } k = 1 \\ 0,1, & \text{για } k = 2 \\ 0, & \text{για } k = 3 \end{cases}$$

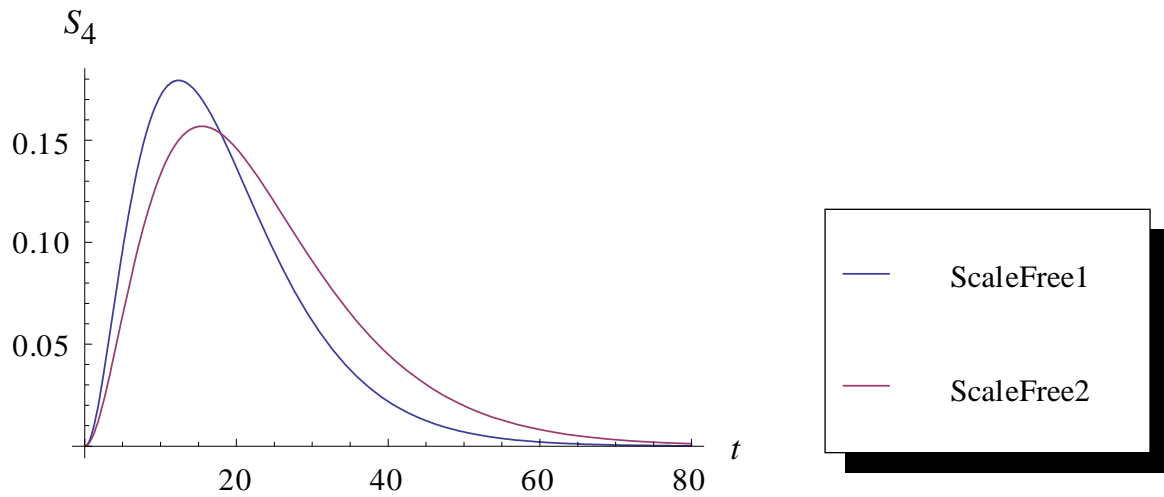
Με τις υπόλοιπες παραμέτρους ίδιες όπως στην 4.1 συγκρίνουμε πρώτα την μεταβολή πιθανότητας της τελικής κατάστασης με την αντίστοιχη εκείνης της κατανομής.



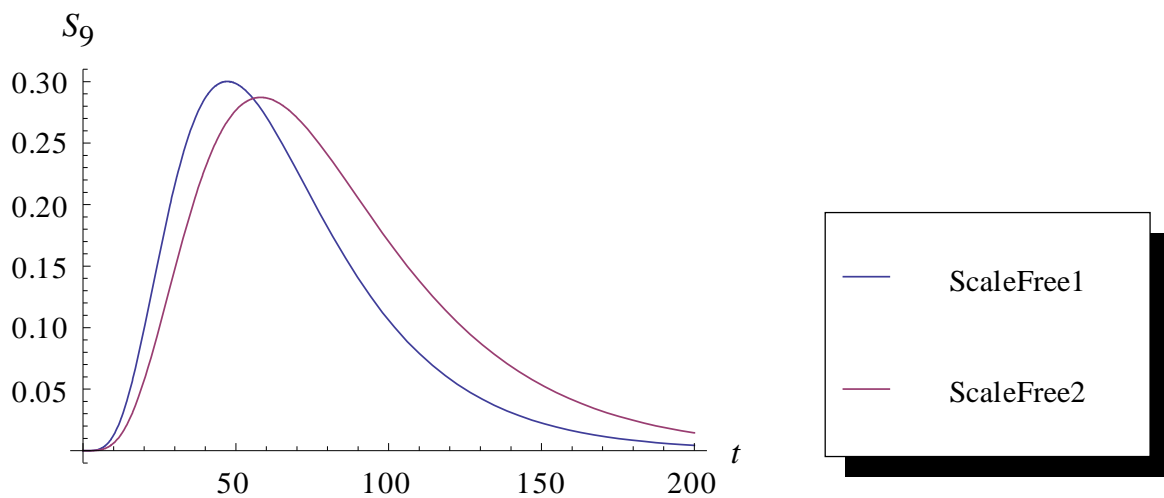
Σχήμα 4.3.1: Μεταβολή πιθανότητας της τελικής κατάστασης με το χρόνο για δύο διαφορετικές κατανομές απουσίας κλίμακας με την "scale free2" πιο έντονη από την "scale free1".

Όπως παρατηρούμε το σύστημα προσεγγίζει πιο αργά την τελική του κατάσταση όταν η συνδεσιμότητα μειώνεται.

Δυο τυπικές πιθανότητες καταστάσεων στη συνέχεια συγκρίνονται για τις δύο διαφορετικές κατανομές.



Σχήμα 4.3.2: Μεταβολή πιθανότητας της κατάστασης 4 με το χρόνο για δύο διαφορετικές κατανομές απουσίας κλίμακας με την "scale free2" πιο έντονη από την "scale free1"



Σχήμα 4.3.3: Μεταβολή πιθανότητας της κατάστασης 9 με το χρόνο για δύο διαφορετικές κατανομές απουσίας κλίμακας με την "scale free2" πιο έντονη από την "scale free1"

Τα δύο προηγούμενα σχήματα μας δείχνουν ότι έχουμε μια σχετική μετατόπιση του μέγιστου για μια πιθανότητα κατάστασης καθώς και μια αύξηση στο χρόνο που έχουν αυτές ένα ενεργό πλάτος εντός της διάρκειας της επιδημίας. Τελικά μια πιο απότομη κατανομή βαθμού αυξάνει την διάρκεια της επιδημίας καθώς οι πιθανότητες των μεταβάσεων μειώνονται λόγω της πιο περιορισμένης συνδεσιμότητας.

Από ένα αντίστοιχο του 4.1.2 σχήματος θα βλέπαμε ότι οι πιθανότητες παρουσιάζουν παρόμοια επικάλυψη για μια τέτοια μεταβολή στην κατανομή βαθμού. Μπορεί λοιπόν να εξεταστεί το ερώτημα αν η κατανομή βαθμού μπορεί να προκαλέσει το χαρακτηριστικό της μείωσης επικάλυψης όπως το προκαλεί η μείωση του ρυθμού θεραπείας.

Ένα γενικότερο δίδαγμα είναι ότι μειώνοντας τις συνδέσεις μειώνουμε την πιθανότητα να μολυνθεί ένας κόμβος αλλά από την άποψη της επιδημίας δεν αλλάζει δραματικά ο τρόπος που αυτή τελικά θα επικρατήσει παρά μονό η διάρκεια της.

4.4 Παραλλαγές του μοντέλου

Στη συνέχεια εξετάζουμε κάποιες παραλλαγές του μοντέλου μας δεχόμενοι την ίδια πιθανότητα μόλυνσης ενός κόμβου, καθώς και την ίδια κατανομή πιθανότητας για το σύνολο των νέων μολυσμένων και θεραπευμένων σε ένα βήμα χρόνου. Στις δύο εκδοχές που θα παρουσιάσουμε κάποιες μεταβάσεις γίνονται αδύνατες, αλλά η πιθανότητα παραμονής σε μια κατάσταση προκύπτει και εδώ ως η συμπληρωματική της συνολικής πιθανότητας μετάβασης.

4.4.1 Μείωση πλήθους μολύνσεων

Μπορούμε να υποθέσουμε ένα πλήθος μολύνσεων x μικρότερο από το μέγιστο δυνατό που είναι ίσο με το πλήθος των ευάλωτων, ώστε να προσομοιώσουμε την συμπεριφορά αυτή στο μικρό δίκτυο. Το χαρακτηριστικό αυτό δεν έχει σχέση με την κατανομή βαθμού αλλά με την ελευθερία που αφαιρούμε από το σύστημα να κάνει μεταβάσεις σε μεγάλη κλίμακα σε κάθε βήμα χρόνου. Στην έως τώρα μελέτη μας δεν έχουμε δώσει στην συνδεσιμότητα κόμβου τη δυνατότητα να επηρεάσει το πλήθος x των νέων μολύνσεων. Μπορούμε να θεωρήσουμε τη δυνατότητα να μολύνονται όλοι σε ένα βήμα χρόνου και μάλιστα να θεραπεύονται στο ίδιο διάστημα σε ένα τεράστιο δίκτυο όπως ο ιστός να φαίνεται υπερβολική. Παρόλα αυτά κάποιος μπορεί να θεωρήσει μία υποθετική, ανεπίσημη και δημοφιλή νέα ενημέρωση να μεταφέρεται από χρήστη σε χρήστη και να θεωρείται ασθένεια με την ευρεία έννοια του όρου.

Έως τώρα η κατανομή βαθμού επηρέασε την μετάβαση του συστήματος από μια κατάσταση σε μια άλλη, μέσω της πιθανότητας μόλυνσης ενός κόμβου. Όπως έχουμε ήδη αναφέρει, δεν είναι η συγκεκριμένη μορφή της κατανομής βαθμού που θα επηρεάσει ποιες μεταβάσεις μπορούν να γίνουν αλλά οι υποθέσεις μας για το μοντέλο. Η κατανομή βαθμού είναι μια πρώτη μέτρηση της

τοπολογίας και δεν θα μπορούσε να περικλείει τα συνολικά τοπολογικά χαρακτηριστικά του δικτύου. Στα μεγάλα δίκτυα θεωρούμε ότι στο βήμα χρόνου είναι υπερβολικό να αφήσουμε το σύστημα να έχει τόσο μεγάλη ελευθερία αλλά η κατανομή μπορεί και πάλι να προσφέρει ένα μέτρο στο πλήθος των επιτρεπτών μεταβάσεων.

Στην προσπάθεια μας να δώσουμε στην κατανομή βαθμού την δυνατότητα να επιδράσει στο σύνολο x των νέων γεννήσεων μολυσμένων, μπορούμε να εισάγουμε τον περιορισμό για παράδειγμα

$$x \leq \left[S \frac{\langle k \rangle}{N} + c \right] \quad (4.3.2.1)$$

Η εξίσωση αυτή (όπου με αγκύλη θεωρούμε τον γνωστό συμβολισμό του ακέραιου μέρους) μας δείχνει ότι ένα κλάσμα μόνο από τους ευάλωτους μπορεί να μολυνθεί σε ένα βήμα χρόνου. Το γεγονός οι νέες γεννήσεις μολυσμένων στο σύστημα να είναι ίσες με το πλήθος ευάλωτων θα μπορεί να θεωρηθεί δυνατό μόνο αν η μέση συνδεσιμότητα είναι ίση με το πλήθος του δικτύου. Όπως βλέπουμε κάτι τέτοιο ισχύει κατά προσέγγιση αφού έχουμε εισάγει μια σταθερά $c < 1$, ώστε να αποφύγουμε την μηδενική τιμή του x παρόλο που μπορεί να υπάρχει ένας μικρός αριθμός ευάλωτων. Η έκφραση (4.3.2.1) μπορεί να βελτιωθεί αλλά εδώ μας αρκεί το γενικό της μήνυμα.

Για το δίκτυο μας και με την υπόθεση κατανομής της υπό-ενότητας 4.1 έχουμε $\langle k \rangle = 1,4$. Για $c = 0,65$ καταλήγουμε ότι όταν οι ευάλωτοι είναι από 1 έως και 3 οι δυνατές μεταβάσεις είναι αυτές που προκύπτουν από την γέννηση ενός το πολύ μολυσμένου ($x=1$). Για $S=4$ επιτρέπεται να γίνουν το πολύ 2 μολύνσεις ενώ για μηδενικό πλήθος ευάλωτων $x=0$ όπως θα έπρεπε. Στη περίπτωση αυτή βέβαια έχουμε ήδη θεωρήσει ότι αλλάζει η περιγραφή της μετάβασης.

Σε αντιστοιχία με το αρχικό μοντέλο οι μεταβάσεις ορίζονται πλέον όπως στον επόμενο πίνακα.

$$Q = \begin{bmatrix} q_{1,1} & q_{1,2} & q_{1,3} & q_{1,4} & q_{1,5} & q_{1,6} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & q_{2,2} & q_{2,3} & q_{2,4} & q_{2,5} & q_{2,6} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & q_{3,3} & q_{3,4} & q_{3,5} & q_{3,6} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & q_{4,4} & q_{4,5} & q_{4,6} & q_{4,7} & q_{4,8} & q_{4,9} & q_{4,10} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & q_{5,5} & q_{5,6} & q_{5,7} & q_{5,8} & q_{5,9} & q_{5,10} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & q_{6,6} & q_{6,7} & q_{6,8} & q_{6,9} & q_{6,10} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & q_{7,7} & q_{7,8} & q_{7,9} & q_{7,10} & q_{7,11} & q_{7,12} & q_{7,13} & q_{7,14} & q_{7,15} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{8,8} & q_{8,9} & q_{8,10} & 0 & q_{8,12} & q_{8,13} & q_{8,14} & q_{8,15} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{9,9} & q_{9,10} & 0 & 0 & q_{9,13} & q_{9,14} & q_{9,15} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{10,10} & 0 & 0 & 0 & q_{10,14} & q_{10,15} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{11,11} & q_{11,12} & q_{11,13} & q_{11,14} & q_{11,15} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{12,12} & q_{12,13} & q_{12,14} & q_{12,15} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{13,13} & q_{13,14} & q_{13,15} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{14,14} & q_{14,15} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

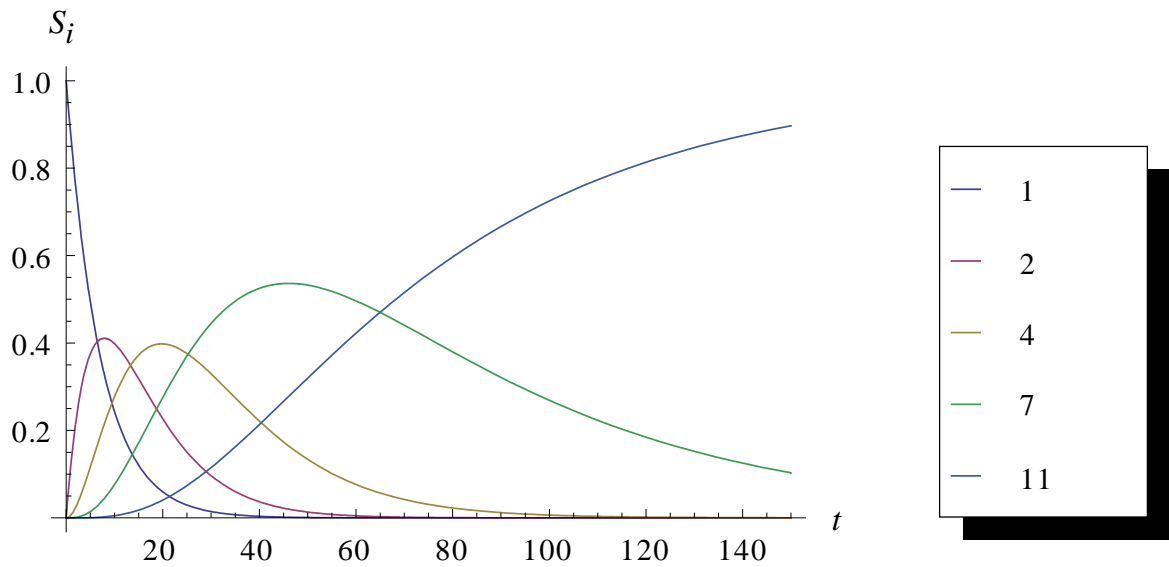
Εκτελώντας τους υπολογισμούς θα παρατηρούσαμε για το παράδειγμα αυτό ότι δεν έχουμε καμία ουσιαστική διαφορά με το αντίστοιχο του αρχικό. Μια αιτία θα φανεί αν κοιτάξουμε τον αρχικό πίνακα μετάβασης. Καθώς οι μεταβάσεις προς πιο μακρινές καταστάσεις γίνονται όλο και πιο απίθανες το ποσό πιθανότητας που χάνουμε όταν τις αφαιρούμε δεν είναι ικανό να αλλάξει την χρονική εξέλιξη σε μεγάλη κλίμακα. Στα μεγαλύτερα δίκτυα ενδέχεται να έχουμε αισθητή διαφοροποίηση. Στην περίπτωση μας εδώ το αποτέλεσμα μπορεί να θεωρηθεί ως μια επιβεβαίωση της διωνυμικής κατανομής ως αποτρεπτικής ώστε να μολύνονται με μεγάλη πιθανότητα πολλοί κόμβοι σε ένα διάστημα χρόνου.

4.4.2 Καθυστέρηση θεραπείας

Μια από τις υποθέσεις μας έως τώρα, αυτή του χρόνου έναρξης της θεραπείας, φαίνεται να μην είναι τόσο ρεαλιστική τουλάχιστον σε ένα μεγάλο πλήθος περιπτώσεων. Γενικά οι ενημερώσεις ασφαλείας κατά των κακόβουλων λογισμικών μπορούμε να θεωρήσουμε ότι γίνονται γνωστές και τελικά τρέχουν σε όλους τους χρήστες σε ένα πολύ μικρό χρονικό διάστημα σε μια χρονική στιγμή μετά την έναρξη της επιδημίας και όχι αμέσως με την εμφάνιση της.

Μπορούμε να χρησιμοποιήσουμε δύο μακροβιανές περιγραφές του μοντέλου μας για το σκοπό αυτό διαδοχικά στο χρόνο. Η πρώτη θα αποτελείται από όλες αυτές τις καταστάσεις που συμβαίνει μόνο μόλυνση δηλαδή θα έχουμε ένα μοντέλο επιδημίας SI. Τη χρονική στιγμή που θα θεωρήσουμε ότι έγινε γνωστή η θεραπεία μπορούμε να χρησιμοποιήσουμε τις πιθανότητες των καταστάσεων του μοντέλου SI στο μοντέλο SIR ως αρχικές. Η μελέτη της επιδημίας θα γίνει διαδοχικά λοιπόν.

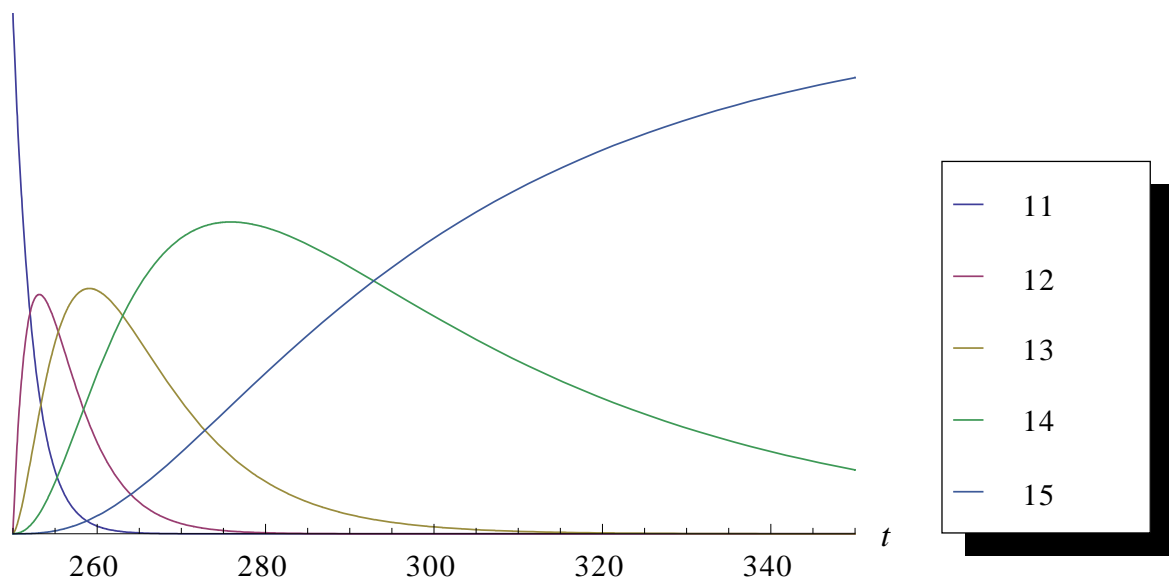
Στο αρχικό παράδειγμα μας των 4 κόμβων και για $\lambda=\theta$ με τις ίδιες υπόλοιπες υποθέσεις η παρούσα περίπτωση περιγράφεται με 5 καταστάσεις οι οποίες αντιστοιχούν στις 1,2,4,7,11 του αρχικού παραδείγματος. Η χρονική μεταβολή των πιθανοτήτων για τις καταστάσεις αυτές φαίνεται στο παρακάτω σχήμα, για αρχική κατάσταση 1, ενώ διατηρούμε για λόγους σύγκρισης την αρίθμηση όπως προηγουμένως.



Σχήμα 4.4.1 : Μεταβολή πιθανότητας καταστάσεων για το μοντέλο SI 4 κόμβων.

Όπως φαίνεται από το παραπάνω σχήμα οι καταστάσεις με κανέναν, έναν, δύο, τρεις και τέσσερις μολυσμένους έχουν μέγιστη πιθανότητα σε διαδοχικά χρονικά διαστήματα. Στη συνέχεια είναι θέμα υπόθεσης η εισαγωγή του χρόνου που αρχίζει η θεραπεία. Ας θεωρήσουμε έναν χρόνο $t=250$ στο παράδειγμα αυτό όταν δηλαδή ο συνολικός πληθυσμός έχει αποκτήσει την ασθένεια, κάτι που συμβαίνει με πιθανότητα 0,98, μια υπόθεση χρόνου έναρξης θεραπείας λογική συνήθως για μικρά δίκτυα. Στην περίπτωση αυτή έναρξης της θεραπείας στην περιγραφή που απομένει μέσω του μοντέλου SIR θα κυριαρχεί η κατάσταση 11.

Στο επόμενο σχήμα φαίνεται η διαδικασία αυτή, της θεραπείας.



Σχήμα 4.4.2: Μεταβολή πιθανότητας καταστάσεων 11 έως 15 για το μοντέλο SIR 4 κόμβων με κυρίαρχη αρχική κατάσταση την 11, την χρονική στιγμή $t=250$

Το σχήμα αυτό θα πρέπει να θεωρηθεί ως συνέχεια του 4.4.2.1 για $t=250$ και μετά. Έχει ενδιαφέρον να συγκρίνουμε τον συνολικό χρόνο επιδημίας στο αρχικό μοντέλο με τον αντίστοιχο της εκδοχής μας εδώ. Ως συνολικό χρόνο επιδημίας μπορούμε να ορίσουμε την διάρκεια από $t=0$ έως το χρόνο που η πιθανότητα της τελικής κατάστασης φτάνει σε μια πιθανότητα πολύ κοντά στην μονάδα, για παράδειγμα 0,99. Στην παρούσα εκδοχή μας της καθυστερημένης θεραπείας ο συνολικός χρόνος προκύπτει από το άθροισμα των δύο χρονικών διαρκειών.

Από την συγκεκριμένη επιλογή της έναρξης θεραπείας $t=250$ και για διάρκεια 200 μονάδων για την επίτευξη πιθανότητας τελικής κατάστασης με πιθανότητα 0,99 παρατηρούμε ότι η προκύπτει μικρότερη συνολική διάρκεια επιδημίας στην περίπτωση καθυστερημένης θεραπείας. Συγκεκριμένα στο αρχικό μοντέλο χρειάζεται χρόνος $t=580$ μονάδων ώστε να φτάσουμε στην τελική κατάσταση με πιθανότητα 0,990652. Θα πρέπει να εξετάσουμε με προσοχή τα επιλεγμένα κριτήρια για τους χρόνους μας πριν βγάλουμε ένα συμπέρασμα όπου στην περίπτωση της καθυστερημένης θεραπείας η επιδημία διαρκεί λιγότερο. Παρόλα αυτά ακόμα και για έναρξη επιδημίας στο $t=300$ όπου η τελική κατάσταση της πρώτης φάσης έχει φτάσει με πιθανότητα 0,994693 συν 200 μονάδων χρόνου για την δεύτερη φάση βλέπουμε ότι η διάρκεια επιδημίας στο αρχικό μοντέλο είναι μεγαλύτερη. Σαφώς ένας λόγος που θα μπορούσε να αποκαταστήσει ένα τέτοιο αποτέλεσμα, ίσως αντίθετο στην διαίσθηση μας, είναι ότι η πιθανότητα μιας τελικής κατάστασης τείνει στη μονάδα με πολύ μικρό ρυθμό από ένα σημείο χρόνου και μετά όπου το όριο προσεγγίζεται με όλο και μεγαλύτερη ακρίβεια. Αλλά και τότε ένας υπερβολικά μεγάλος χρόνος για την ολοκλήρωση της πρώτης φάσης μοιάζει αφύσικος. Αντίθετα αν βιαστούμε να θεωρήσουμε ότι έληξε η πρώτη φάση θα αγνοήσουμε έστω ένα μικρό ποσό πιθανότητας το οποίο αντιστοιχεί σε ένα αισθητό χρονικό διάστημα.

Ένα σημείο που πρέπει να δώσουμε σημασία για να δικαιολογήσουμε το παραπάνω αποτέλεσμα είναι το πλήθος των καταστάσεων. Σε κάθε μια από τις δύο φάσεις υπάρχουν 5 καταστάσεις και αυτό που φαίνεται, είναι η μείωση των καταστάσεων σε κάθε φάση σε σχέση με τις 15 του αρχικού μοντέλου να έχει ως επακόλουθο την μείωση της χρόνου όπου η κάθε πιθανότητα τελικής κατάστασης τείνει στην μονάδα. Σύμφωνα με τα παραπάνω και για λογική παραδοχή του χρόνου λήξης της πρώτης φάσης μπορούμε να δούμε ότι η συνολική διάρκεια της επιδημίας είναι περίπου η ίδια στις δύο εκδοχές των μοντέλων. Έχει ενδιαφέρον να εξετάσουμε τον ίδιο υπολογισμό για μεγαλύτερα δίκτυα όπου μπορεί να φανεί τυχόν όμοια συμπεριφορά ώστε να καταλήξουμε και σε αυτές τις περιπτώσεις ότι αν η θεραπεία αρχίσει μόλις όλοι έχουν μολυνθεί ο συνολικός χρόνος επιδημίας δεν θα αλλάξει αισθητά. Σίγουρα θα επωφεληθούμε βέβαια όταν η θεραπεία αρχίζει

νωρίτερα αν αναλογιστούμε ότι τότε υπάρχει πιθανότητα διάφορη του μηδενός να έχουμε θεραπευμένα μέλη.

4.5 Επίλογος

Έχοντας δει και αριθμητικά τη διαδικασία του μοντέλου που περιγράφουμε μπορούμε να κάνουμε κάποιες επιπλέον παρατηρήσεις. Σχετικά με τη διαδικασία της δημιουργίας ασθένειας μέσω των κακόβουλων κόμβων έχουμε παρατηρήσει στην αρχή της 3.1 ότι διακρίνεται η ενεργή δράση τους κάπου μέσα στον ευάλωτο πληθυσμό και ταυτόχρονα τους τοποθετεί εκεί. Κάτι τέτοιο θα μπορούσε να θεωρηθεί όχι τόσο ρεαλιστικό μια και η γνώση της θεραπείας είναι ήδη γνωστή για αυτούς. Παρόλα αυτά η περιγραφή μας εφόσον διατηρεί τρία είδη πληθυσμών, αν και δεν μπορεί να τους τοποθετήσει κάπου αλλού, παρουσιάζει τη δράση τους. Στην πραγματικότητα ο κακόβουλος πληθυσμός αποτελεί ένα τέταρτο είδος και δρα στον ευάλωτο πληθυσμό, αλλά ανήκει σε έναν τέταρτο πληθυσμό με ιδιότητες ανοσίας όπως και οι θεραπευμένοι. Μια άλλη αντιμετώπιση θα μπορούσε να είναι πιο ακριβής σύμφωνα με αυτή τη λογική. Είναι γεγονός ότι κατά τη διάρκεια της επιδημίας δεν είναι γνωστός ο τρόπος δράσης των κακόβουλων οπότε θα πρέπει να προχωρήσουμε με προσοχή εφόσον θέλουμε να τους εντάξουμε σε ένα μοντέλο ως ένα ξεχωριστό πληθυσμό. Στην περίπτωση μας οι κακόβουλοι κόμβοι θεωρούν τον εαυτό τους ως ευάλωτο για το συνολικό δίκτυο και κάποια στιγμή αποφασίζουν ότι χρειάζεται να αλλάξουν την ταυτότητά τους σε μολυσμένους και τελικά θεραπευμένους. Θα ήταν επιθυμητό να υπάρχει ένας μηχανισμός ώστε να μην μολύνουν οι κακόβουλοι κόμβοι ολοκληρωτικά τους ευάλωτους. Αν και σκοπός τους είναι να το καταφέρουν μπορούμε να φανταστούμε περιπτώσεις όπου ο ευάλωτος κόμβος είναι υποψιασμένος ώστε να προχωρήσει σε μια σύνδεση που θα αποβεί μοιραία για αυτόν. Εδώ υπάρχει μετάβαση από ευάλωτους σε θεραπευμένους και τελικά στο τέλος της επιδημίας δεν θα έχουν όλοι μολυνθεί παρόλο που οι ευάλωτοι θα είναι μηδέν. Στο κλασικό μοντέλο SIR στο τέλος της επιδημίας υπάρχει δυνατότητα να μην έχουν μολυνθεί όλοι, αλλά οι ευάλωτοι που παρέμειναν δεν συνειδητοποίησαν τον κίνδυνο, απλά έτυχε και θεραπεύτηκαν όλοι οι μολυσμένοι. Για παράδειγμα σε μια επικίνδυνη εφαρμογή ενός κοινωνικού δικτύου ο χρήστης μπορεί να αρνηθεί την συμμετοχή του αν υποψιαστεί κάτι τέτοιο. Το θέμα αυτό όμως εμπλέκει και το χαρακτηριστικό της ταυτότητας των κόμβων το οποίο είναι πολύ σημαντικό. Δεν είναι όλοι οι κόμβοι το ίδιο ευάλωτοι. Στην εξέτασή μας ο κάθε κόμβος είναι ισότιμος σε όλα με τους υπόλοιπους που ανήκουν στον ίδιο πληθυσμό. Θα μπορούσαμε να ορίσουμε ένα μέτρο για την τιμή προδιάθεσης για ασθένεια και μια κατανομή

πιθανότητας στο σύνολο των ευάλωτων ώστε να πάρουν την τιμή αυτή. Παρόμοιες παρατηρήσεις ισχύουν και για την αντιμετώπιση θεραπείας.

Όπως φαίνεται από την θεωρητική ανάπτυξη του μοντέλου μας η κατανομή βαθμού βοήθησε αρκετά αλλά έως το σημείο που έπρεπε να συσχετίσουμε πιθανότητες πέρα από τους γείτονες ενός κόμβου. Πραγματικά στο σημείο αυτό η κατανομή βαθμού αδυνατεί να μας πληροφορήσει για την κατανομή της γενικότερης τοπολογίας. Αυτό φαίνεται συγκεκριμένα στην υποθετική μας διωνυμική κατανομή για την πιθανότητα x νέων μολύνσεων στο συνολικό δίκτυο. Μια άλλη κατανομή που θα ήταν υποψήφια είναι η πολυωνυμική. Επίσης στην κατανομή των μολυσμένων ακμών ενός κόμβου έπρεπε να κάνουμε μια υπόθεση που αφορούσε μεγάλη κλίμακα. Η έννοια της κεντρικότητας είναι σημαντική αλλά δεν μπορεί να εφαρμοστεί σε στοχαστική περίπτωση αφού απαιτεί την ακριβή τοπολογία, στην παρούσα μορφή της.

Βιβλιογραφία:

- [1] Kermack W. O. and A. G. McKendrick, Contributions to the mathematical Theory of epidemics-I, Proceedings of the Royal Society 115A: 700-721, 1927 (Reprinted in Bulletin of Mathematical Biology 53(1/2): 33-55, 1991)
- [2] Troy Tassier, SIR Model of Epidemics, Lesson Notes at Fordham University_Rose_Hill, 2005
- [3] Kermack and McKendrick, The Variable Susceptibility Model for Infectious Diseases, Japan Journal of Industrial and Applied Mathematics - Volume 18, 2001.
- [4] <http://en.wikipedia.org>: Epidemic Model
- [5] Andrei Korobeinikov and Philip K. Maini, Mathematical Biosciences and Engineering - Volume 1, 2004
- [6] Fred Brauer - Pauline Van den Driessche - Jianhong Wu, Mathematical Epidemiology, Springer-Verlag Berlin and Heidelberg GmbH & Co. KG, 2008
- [7] Ε. Συκάς, Σημειώσεις Θεωρίας Τηλεφωνικής Κίνησης, ΕΜΠ, 2003.
- [8] Β. Μάγκλαρης - Τ. Χιώτης, Σημειώσεις Συστημάτων Αναμονής, ΕΜΠ, 2000
- [9] Nikolaos Demiris, Bayesian Inference for Stochastic Epidemic Models using Markov chain Monte Carlo Methods, Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy, 2004
- [10] Romualdo Pastor-Satorras and Alessandro Vespignani, Immunization of complex networks, Phys.Rev. E 65 036104, 2002
- [11] Akram H. Rustam, Epidemic Network and Centrality, Master Thesis, 2006
- [12] M. E. J. Newman, The mathematics of networks, The New Palgrave Encyclopedia of Economics 2nd edition, 2008.
- [13] M. E. J. Newman, A measure of betweenness centrality based on random walks, Social Networks 27, 2005
- [14] M. E. J. Newman, Spread of epidemic disease on networks, Phys. Rev. E 66 016129, 2002
- [15] P. Grassberger, Math. Biosci. 63, 157, 1983.
- [16] M. E. J. Newman - S.H. Strogatz and D.J. Watts, Random graphs with arbitrary degree distributions and their applications, Phys. Rev. E64, 026118, 2001
- [17] Albert-László Barabási, Scale-Free Networks: A Decade and Beyond, Science Magazine Volume 35, 2009
- [18] Romualdo Pastor-Satorras - Alessandro Vespignani, Epidemic dynamics in finite size scale-free networks. Phys. Rev. E 65, 2002

[19] Romualdo Pastor-Satorras - Alexei Vázquez and Alessandro Vespignani, Dynamical and Correlation Properties of the Internet, Phys. Rev. Lett.87, 2001

[20] Guillermo Abramson, Mathematical modeling of the spread of infectious diseases, A series of lectures given at PANDA UNM, 2001.