ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

“Spreading and blocking of rumor in complex networks”

Διπλωματική Εργασία

ΝΕΡΑΝΤΖΑΚΗ ΕΥΑΓΓΕΛΙΑ

Βόλος, 2014

«Εξάπλωση και περιστολή φήμης σε σύνθετα δίκτυα»

Επιβλέπων: Κατσαρός Δημήτριος, Λέκτορας

Συνεπιβλέπων: Μποζάνης Παναγιώτης, Αναπληρωτής Καθηγητής

Spreading and blocking of rumor in complex networks”

Supervisor: Katsaros Dimitrios, Lecturer

Co-advisor: Bozanis Panayiotis, Associate Professor

Ευχαριστίες

Αρχικά, θα ήθελα να ευχαριστήσω τον καθηγητή του τμήματος Μηχανικών Η/Υ Τηλεπικοινωνιών και Δικτύων και βασικό επιβλέποντα της πτυχιακής αυτής εργασίας κ.Δημήτριο Κατσαρό που μου έδωσε την ευκαιρία να πραγματοποιήσω αυτή την μελέτη. Η υποστήριξή του, η αμέριστη συμπαράστασή του, αλλά και οι διαρκείς και εύστοχες υποδείξεις του βοήθησαν στην έγκαιρη ολοκλήρωση αυτής της μελέτης.

Επιπρόσθετα, θα ήθελα να ευχαριστήσω τον έτερο επιβλέποντα καθηγητή κ.Παναγιώτη Μποζάνη και τον διδακτορικό φοιτητή του τμήματος Παύλο Μπασαρά για την προγενέστερη δουλειά του, την μεταλαμπάδευση της γνώσης αυτής καθώς και τη συνεχή καθοδήγησή του έως της περάτωση της εργασίας.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου και τους φίλους μου που μου συμπαραστάθηκαν σε όλη την διάρκεια της εκπόνησης αυτής της εργασίας.

Περίληψη

Τα online κοινωνικά δίκτυα εδραιώνουν την κυριαρχία τους ώστε να γίνουν το πιο δημοφιλές μέσω πληροφόρησης καθώς μεγαλώνουν τόσο σε μέγεθος όσο και σε δημοτικότητα. Ο έλεγχος των κακόβουλων δεδομένων (όπως φήμη, διαφημίσεις, λογισμικό, κτλ) απαιτεί άμεση δράση λόγω των τεράστιων επιπτώσεων τους σε παγκόσμια κοινωνική κλίμακα. Σε αυτή την εργασία προτείνεται μία μέθοδος για την περιστολή της εξάπλωσης της κακόβουλης πληροφορίας (παραπληροφόρηση) με αξιοποίηση της τοπικής πληροφορίας του κάθε χρήστη-κόμβου και των γειτόνων του. Βασισμένη στη μετρική *μ-Pci*, υπολογίζεται η τιμή της *ncPci* του κάθε κόμβου, που είναι το μέγεθος της συμβολής του κάθε γείτονα στον υπολογισμό *της μ-Pci*. Ο πειραματισμός μας πάνω σε πραγματικά σύνθετα δίκτυα αποδεικνύει ότι η προσέγγισή μας είναι σε θέση να αποτρέψει την διάδοση των κακόβουλων δεδομένων σε μεγαλύτερο βαθμό απ’ ότι οι ανταγωνιστές της στα πλαίσια του μοντέλου διάδοσης SIR (Susceptible Infected Recovered).

Abstract

Online social networks (OSNs) consolidate their sovereignty in becoming the most popular medium for information dissemination as they grow in both size and popularity. Controlling malicious data (rumors, advertisements, software, etc.) traversing through those networked people, requires immediate action due to its enormous impact in global social scale. In this article we propose our method for blocking the spread of misinformation by utilizing local information from node-users environs, in accordance to its own significance as originally indexed by *μ-Pci*, namely*, neighbor contribution on Power Community Index, ncPci*. Our experimentation in real complex networks shows that our approach is able to deter malicious propagation at a greater extent than its competitors under the Susceptible Infected Recovered, SIR, spreading model.

Contents

1 Introduction & Related Work

In the last decade, influence maximization in complex networks has received increased attention from the research society. In complex-social networks this problem is interpreted as locating those node-users for the spreading of messages (products, rumors, etc.) through which the largest possible portion of the network can be influenced by diverse and specific ways. Quantifying and ultimately restraining the spreading of false rumors or malicious software among those people networks is a major and urgent necessity of equivalent importance. This reverse-dual problem is formulated as follows: find the minimum *set of key players* or *key connections* in a social network whose removal (immunization) will hinder or stop further false (rumor) propagation to the greatest possible extent.

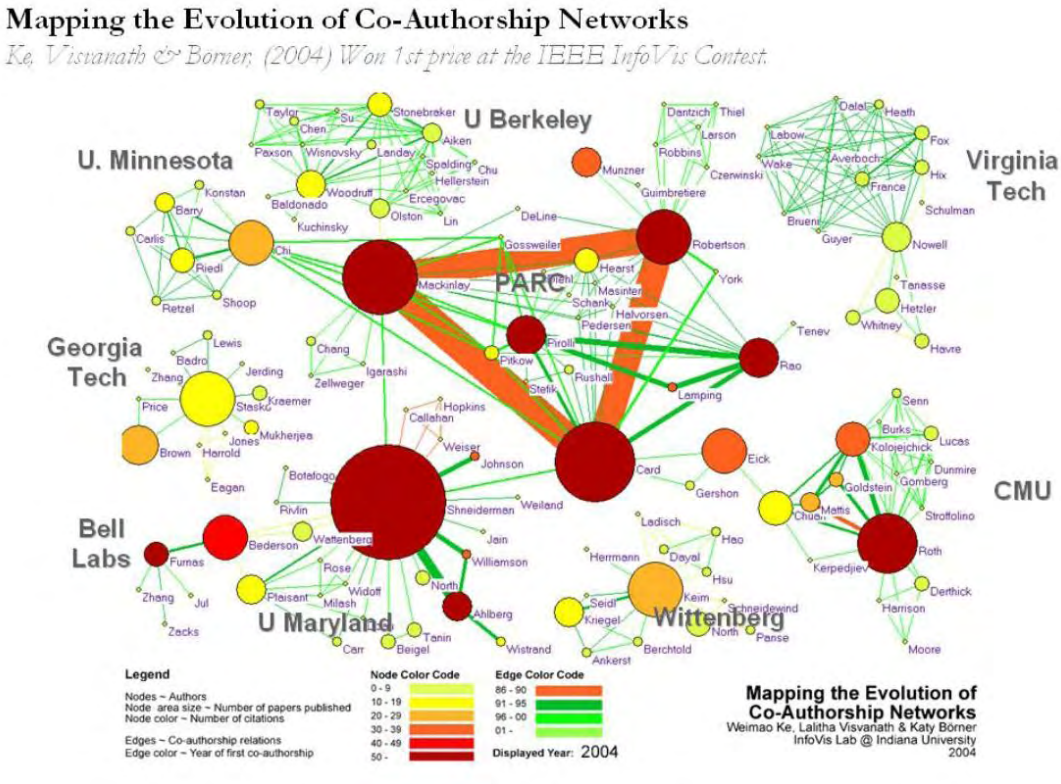

Fig.1: An example of a complex network.

Most of the so far proposed techniques focus on a specific characteristic of social networks, community. Social networks are community networks, i.e. dense connections between nodes within the same community and sparser connections between nodes residing in different communities. Immunizing the bridge nodes, i.e. nodes which connect different communities (or nodes-connections which lead to those bridges) are the primary objective of these studies and thus isolating malicious rumors in the community of origin.

In [1] the authors illustrate a greedy approach based on shortest paths between the bridge nodes and rumor originators to stop the false propagation. This article takes advantage of the community structure of social networks and their primary objective is to stop the malicious rumors from reaching other communities by immunizing individuals who are intermediates in these paths. Bridge nodes and bridge hubs (i.e. nodes with many connections in different communities) are noted as critical components in information dissemination and thus researchers focus on community detection techniques for immunization. However identifying community structures in real social networks (e.g. Facebook, Twitter) is a challenging task due to their size and magnitude. Random walks for detecting small sized communities and dealing with the lack of global structural network information are reviewed in [2][3][4], where bridge hubs constitute the targets for immunization.

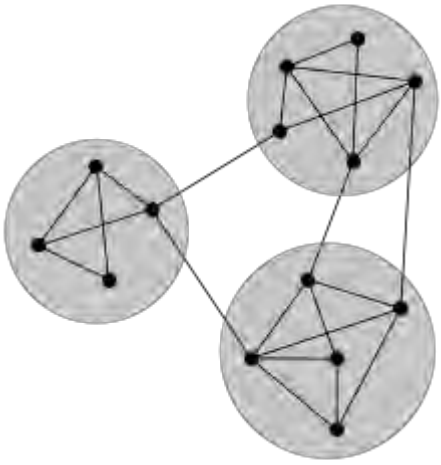

Fig.2: Community Structure

[5] reports that local information of nodes surroundings is usually sufficient to characterize its role in an epidemic, and proves via experimental results under different spreading models that different immunizations strategies need to be applied in order to efficiently block the infection. Leaving aside communities, removing the hub nodes (the highly connected individuals) is a widely used and low cost approach which is proven efficient in blocking the spreading of malicious contents. Other strategies constitute those trying to immunize highly influential nodes as identified by centrality measures such as betweenness or eigenvector [6], however, such measures become impractical in large-scale networks.

All the aforementioned techniques apply node removal as immunization strategies to block the infection. However with such a way we *isolate* the related individuals from the remaining structure and thus significantly reduce the connectivity of a network, especially when a large number of nodes must be removed. Recently edge blocking approaches appeared. As reported in [7] the removal of nodes is equivalent to cutting off their connections and thus removing a user's links is a more profound approach. Here the authors provide a greedy method for selecting edges to block the contagion by utilizing the Bond Percolation Method [8] to estimate the influence degree of edges. In [9] the authors quantify the significance of an edge by multiplying the degrees of its adjacent nodes, and illustrate that their approach is equal or better to node removing techniques, under the *SIS* spreading model, while also maintaining the network integrity. However their experimentation is limited only in artificial networks. [10] is an interesting approach that considers two problems. First the removal of user connections to hinder false propagations, and second the addendum of edges to deal with the opposite issue. This article focuses on the eigenvalue of the graph structure inspired by the findings in [11][12], to minimize the leading eigenvalue and deter misinformation, or maximize it and boost the dissemination process by removing or adding edges respectively, within a specified budget. A similar approach attempting to reduce the spectrum of the adjacency matrix to slow down an epidemic by removing edges is investigated in [13].

In [14] the authors illustrate their algorithm for blocking a contagion and study its performance under threshold progressive and deterministic models. The algorithm allows the contagion to move step by step and at each ”next time step” (*step i+1*) each infected node is assigned a cost value determined by the least number of edges needed to be shut down in the ”previous step” (*step i*) in order to save the node. If the total cost of salvaging all the next infected nodes is within the constraining budget at *step i*, then this edge set is the final immunization set and the algorithms stops. Otherwise the computation is repeated in subsequent steps were at each step either all next infected nodes are saved or none. If this condition is never met, then the solution with the minimum impact of infection is selected as the target edge set.

In [15][16][17] the authors investigate on competitive influence when several opposite products or opinions coexist and compete over influence propagation. This case can be formulated as two opposing campaigns of different actors, one spreading delusive information, for instance about a certain event, and another trying to warn the community in a social network about the falsity of this particular rumor. As other examples consider political rivalry or marketing affairs. These techniques focus on blocking one’s influence potential and thus this problem is similar to our case study. Some of these issues focus on game theory aspects, others select positive seeds to bate and subside negative rumor spreading [18][19][20], while others propose variations in the diffusion models to account for opposing propagations and then solve the influence maximization problem.

To provide a summary on the aforementioned related works, and portray a general framework for the formulation of the problems that these studies solve; some of these tasks are constrained by some maximum available budget (≤ β), and within this threshold value they apply either node or edge immunization techniques to minimize the impact of the contagion, while others try to locate the *minimum set*, *among many sets of edges or nodes* that block the infection to the largest possible extent and address the problem. The so far algorithms used to identify the **target nodes or edges** for immunization vary. Many use traditional centrality measures such as betweenness or

eigenvector, some focus on the highly connected individuals or discount degree approaches, others utilize community detection techniques to detect the bridge nodes or edges, while others apply greedy based approaches to select their target set. The majority of earlier works involves a fixed cost for the removal of *any* edge or *any* node of a network. However can we assume the expenditure of removing a single node or edge to be of a static amount? What if some nodes or connections due to their nature (consider army communication networks) cannot be removed or what if when considering different costs for nodes or edges the bridge connectors cannot be immunized? Even if we can afford to remove bridges, malicious products originating from a large social community or maybe the largest community, will still harm the network environment at a great deal. Undoubtedly there are many unexplored aspects, key parameters and limitations that need to be further studied in order to provide a broad formula and solution to the problem.

This project further addresses the problem of blocking the spread of *misinformation* by removing node-users, i.e. closing down schools and thus removing all interactions between children in case of example of a disease outbreak or shutting down routers to block the spreading of viruses among networked computed. We propose a variation of well studied measure in the literature of identifying influential users in complex networks, namely, *neighbor contribution on power community index, ncPci,* and prove via experimental results that the proposed measure is able to deter malicious diffusion at a greater extent than its competitors. Evaluation was performed in several real complex networks under the Susceptible Infectious Recovered spreading model, SIR.

2 Problem Formulation

In this study we apply the widely used *SIR* spreading model to simulate the diffusion of malicious contents (viruses, rumors, etc.). SIR models three possible states:

- The susceptible state *S*, in which the S nodes are vulnerable to infection.
- The infected state *I*, in which the *I* nodes try to infect their susceptible neighbors and succeed with probability λ.
- The recovered state *R*, in which nodes have recovered from infection and cannot be reinfected.

Starting the simulation an initial seed set of nodes, I_s, is elected as the starting point, comprised of nodes in *I* state for the dissemination. In our experimentation this initial set is chosen from either the *hub* nodes (highly connected individuals) or the *core* nodes as identified by the k-shell decomposition algorithm. Our performance metric is the total number of nodes in the infected state, I_f, at the end of the propagation.

Finally to formulate the addressed problem, given a network *G = (V,E)* of *V* vertices and *E* connections among them, find a set of removable nodes within a minimum budget -β number of nodes- whose removal will deter the malicious propagation to the largest possible extent i.e. will minimize I_f.

There are two approaches to the problem; the first one is characterized as offline and the second one as online. In the first approach we start the immunization process before the infection diffusion starts, meaning that we remove β number of nodes before the first step of the SIR process and after this we let the diffusion propagate without other blocking moves. The second approach presents a new way to deal with the problem and no similar work could be found throughout our research. In this approach, we immunize nodes or edges after each SIR step. The process, in detail, consists of

removing β number of nodes (or edges) after each sir step completes. The infection directs us to the specific nodes (or edges) that are going to be deleted.

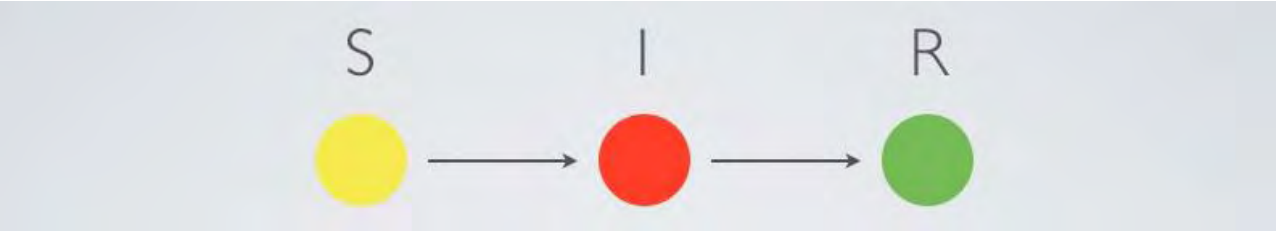

Fig.3: SIR, diffusion process states

3 Offline Methodologies

3.1 Original Power Community Index

Definition 1 (*μ*-**Power Community Index**, *μ*-**pci**). *The μ-PCI of a node v is equal to k, such that there are up to μ × k nodes in the μ-hop neighborhood of v with degree greater than or equal to k, and the rest of the nodes in that neighborhood have a degree less than or equal to k. The goal is to detect nodes located in dense areas of the network and thus likely influential spreaders.*

μ-pci was initially presented in for caching decision in ad-hoc networks and later used for the identification of influential spreaders in dynamic complex networks. It is a local index that characterizes nodes according to the density of their vicinities. Due to its locality it becomes feasible for implementation in real time applications and dynamic networks.

3.2 Neighbor Contribution on pci, ncPci

In this section we discuss the notion of similarity in connections between node *u*, and its one hop vicinity, in order to decide whether or not to include each individual neighbor in our new ranking method. Through our experimentation we realized that *pci* values collected from neighbors that formed their initial *pci*'s from similar environs, as illustrated in Figure 1, should not be included fully but rather partially. For instance, if an epidemic starts from node *x* who shares common vicinity with *y*, some of *y*'s neighbors will be influenced from *x* directly, and thus *y*'s significance on *x* should drop.

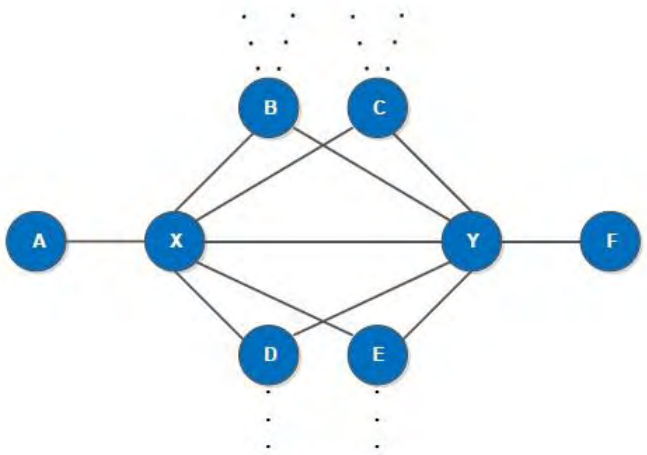

Fig.4: Toy network for building ncPci. Nodes x and y form their initial pci's from almost identical neighboring nodes thus the contribution of one to the other is minimal.

We also deducted that including neighbors with very low *pci* values does not improve our algorithms ranking performance. To account for such occasions we incorporate a tunable parameter *γ* between 0 and 1 as a threshold value, and include those of *u*'s neighbors that satisfy: *pci (j) / pci (u)* > *γ* .When *γ* = 0 all of *u*'s neighbors are included whereas if γ = 1 only those with greater or equal *pci* are considered.

Based on these observations we built our approach, *neighbor contribution on Power community index, ncPci* as follows:

$$nImp = pci(j) \cdot (1 - \frac{N_1}{N_2})$$

$$ncPci(u) = \sum_{i=1}^{N'} \sum_{j=1}^{N'} nImp + \sum_{i=1}^{M'} \sum_{j=1}^{M'} nImp \cdot distance$$

where N_1 is the number of common neighbors between *u* and *j*, N_2 stands for the number of *j*`s neighbors, *N'* for neighbors to which *pci (j) / pci (u) > γ* applies and *M'* for neighbors to which *pci (j) / pci (u) ≤ γ* applies. The *distance* is the difference of the two *pci's, abs (pci (u) - pci (j)).*

3.3 Centralities

In graph theory and network analysis, **centrality** of a vertex measures its relative importance within a graph. Applications include how influential a person is within a social network, how important a room is within a building (space syntax), and how well-used a road is within an urban network. There are four main measures of centrality: degree, betweenness, closeness, and eigenvector.

3.3.1 Degree

In graph theory, the **degree** of a vertex of a graph is the number of edges incident to the vertex, with loops counted twice. The degree of a vertex v is $\deg(v)$. denoted The **maximum degree** of a graph *G*, denoted by Δ(*G*), and the **minimum degree** of a graph, denoted by δ(*G*), are the maximum and minimum degree of its vertices.

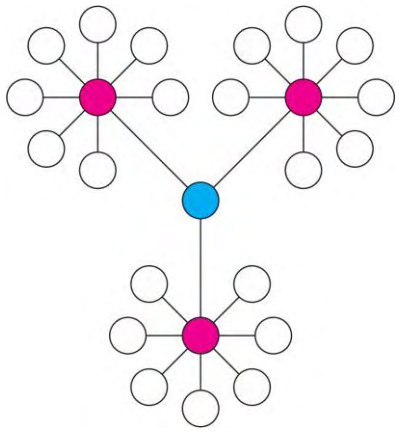

Fig.5: The blue node has degree deg (b) =3 and each red node has degree deg (r) =8.

3.3.2 Betweenness

Betweenness centrality is a measure of a node's centrality in a network. It is equal to the number of shortest paths from all vertices to all others that pass through that node. *Betweenness centrality* is a more useful measure (than just connectivity) of both the load and importance of a node. The former is more global to the network, whereas the latter is only a local effect.

The betweenness centrality of a node v is given by the expression:

$$g(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

where σ_{st} is the total number of shortest paths from node *s* to node *t* and $\sigma_{st}(v)$ is the number of those paths that pass through *v*.

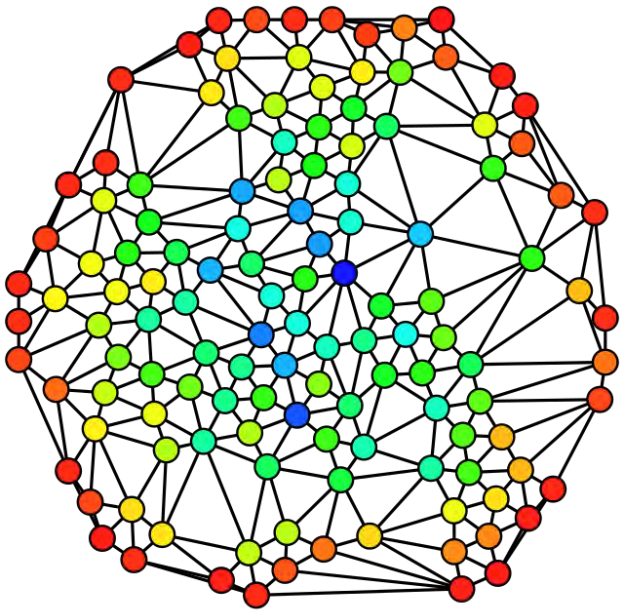

Fig.6: Hue (from red=0 to blue=max) shows the node betweenness. Hue scale representing node betweenness on a graph.

3.3.3 Closeness

In connected graphs there is a natural distance metric between all pairs of nodes, defined by the length of their shortest paths. The **farness** of a node *s* is defined as the sum of its distances to all other nodes, and its closeness is defined as the inverse of the farness. Thus, the more central a node is the lower its total distance to all other nodes. Closeness can be regarded as a measure of how long it will take to spread information from *s* to all other nodes sequentially.

In general directed graphs:

$$C_H(x) = \sum_{y \neq x} \frac{1}{d(y,x)}$$

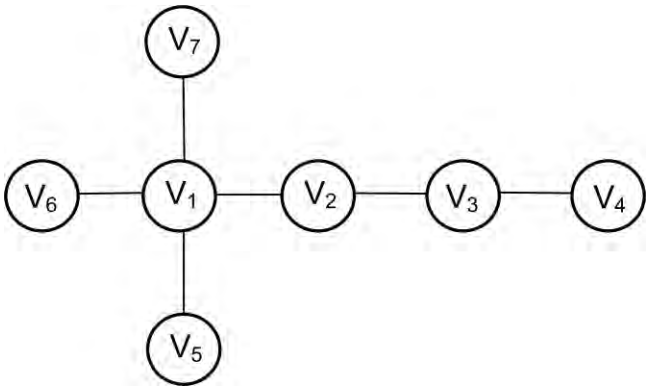

Fig.7: Closeness centrality. V_1: $d_1 = 4 \times 1 + 1 \times 2 + 1 \times 3 = 9$, $C_{clo}(1) = 6/9$. V_1 accesses 4 nodes (V_2, V_5, V_6, V_7) with step 1, 1 node (V_3) with step 2 and 1 node (V_4) with step 3. 6 nodes can be accessed in total by V_1. $\boldsymbol{V_2}$: $d_2 = 2 \times 1 + 4 \times 2 = 10 > d_1$, $C_{clo}(2) = 6/10$. V_2 accesses 2 nodes (V_1, V_3) with step 1 and 4 nodes (V_4, V_5, V_6, V_7) with step 2. 6 nodes can also be accessed in total by V_2. As a result, V_1 is more central than node V_2 since $d1 > d_2$.

Betweenness and closeness centrality require global knowledge of the network and it is difficult to compute them for large networks. In our work, we use *Gephi,* and *R-Library* in order to compute centralities for all networks.

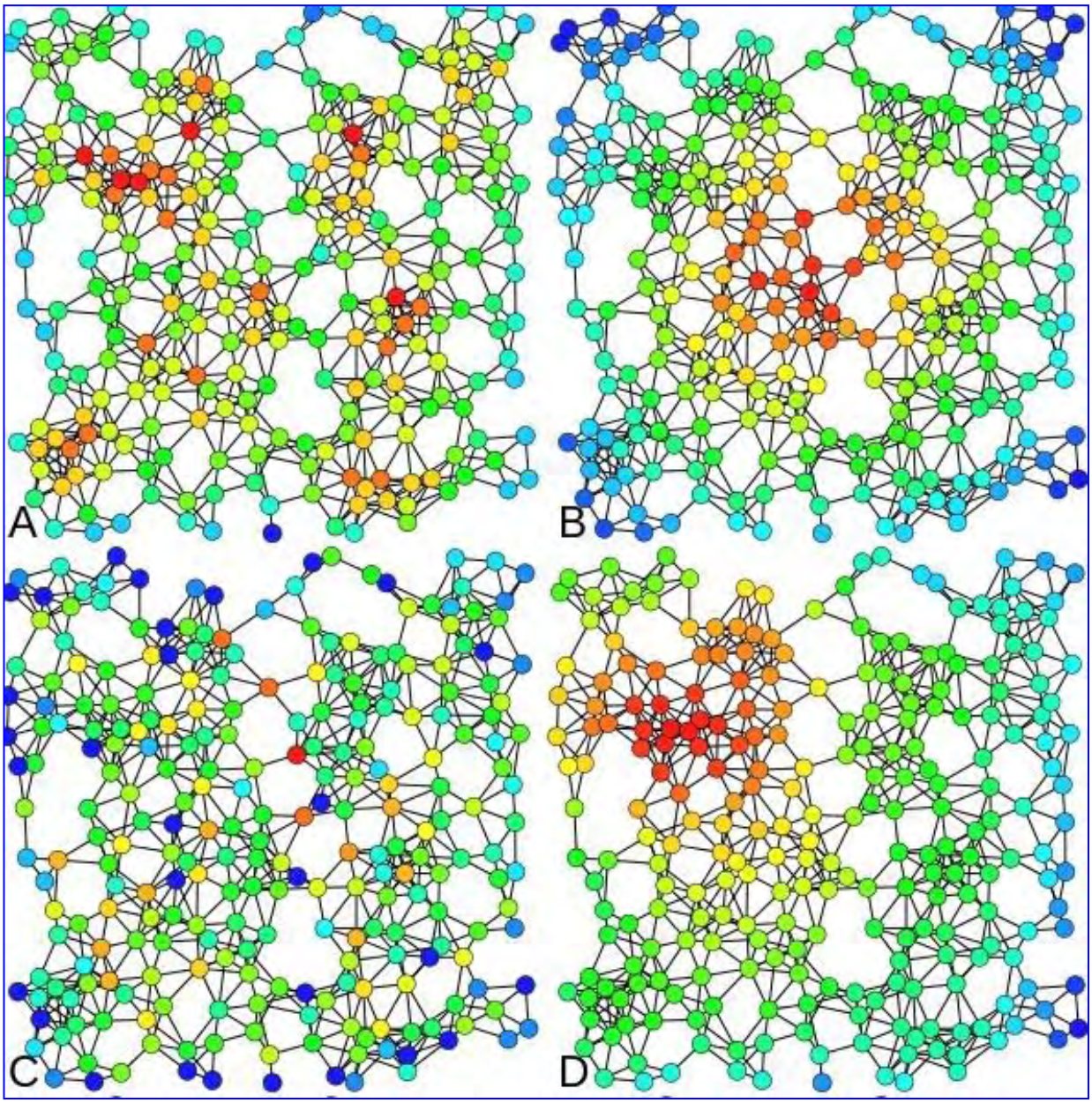

Fig.8: Examples of A) Degree centrality, B) Closeness centrality, C) Betweenness centrality D) Eigenvector centrality

3.4 K-shell Decomposition

In graph theory, a ***k*-degenerate graph** is an undirected graph in which every subgraph has a vertex of degree at most *k*: that is, some vertex in the subgraph touches *k* or fewer of the subgraph's edges. The **degeneracy** of a graph is the smallest value of *k* for which it is *k*-degenerate. The degeneracy of a graph is a measure of how sparse it is, and is within a constant factor of other sparsity measures such as the arboricity of a graph.

K-shell decomposition of a network graph is performed iteratively. The first step involves removing all degree-1 nodes, along with their link, and indexing these as $k = 1$. In the resulting graph, all nodes of degree 1 are also considered to have $k = 1$ and are again pruned. The process is repeated until there are no nodes of degree 1. Similarly, all nodes with i or fewer connections are iteratively removed; these nodes are indexed as $k = i$.

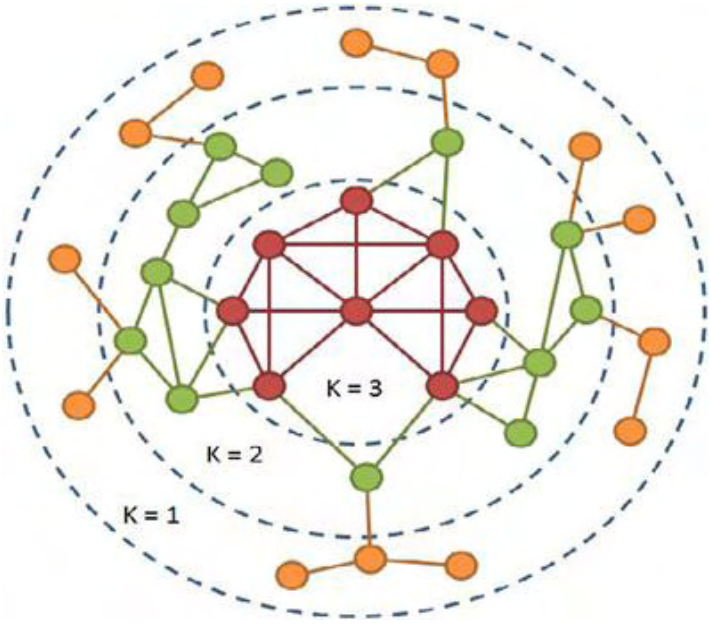

Fig.9: Examples of k-shell decomposition.

4 Online Methodologies

In order to deal with the online approach of the problem, apart from node removal we implement methodologies which face it by removing edges. This approach is new, so no competitors could be used.

4.1 Node Removal

The first methodology removes one hop neighbor nodes with the highest degree while the second removes one hop neighbor nodes with the highest ego betweenness. Ego betweenness is the betweenness centrality measure in an ego network.

4.1.1 Ego Network

Ego networks consist of a focal node ("ego") and the nodes to whom ego is directly connected to (called "alters") plus the ties, if any, among the alters. Each alter in an ego network has his/her own ego network.

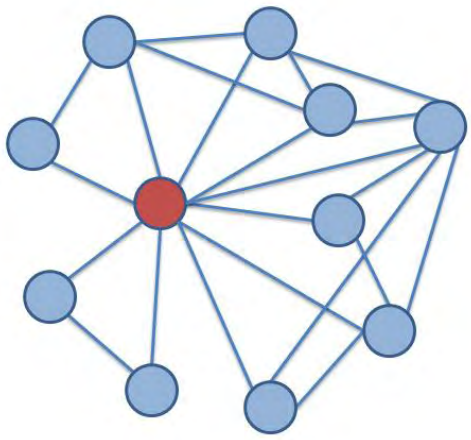

Fig.10: An example of an ego Network. The red node is the “ego” node and blue nodes are alters.

4.2 Edge Removal

We implement and compare two methods used for edge removal. The first includes removing edges which drives to nodes with the highest degree and the second incorporates removing edges which drives to nodes with the highest rdegree.

4.2.1 rDegree

rDegree is a metric we use to describe the number of the one-hop neighbors in state *S*.

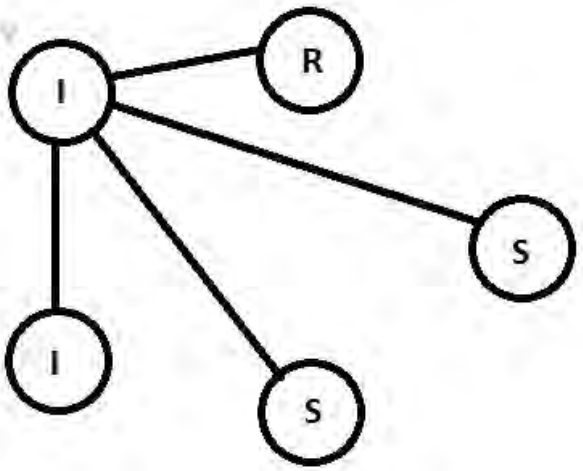

Fig.11: The degree of the v node is deg (v) =4 and the rdegree is rdeg (v) =2.

5 Performance Evaluation

In this section we present the results and the experimentation analysis of the methods we use. We also provide a description of the Online Social Networks we used.

5.1 Dataset Collection

The networks we use for our experiments are real networks. We use 4 networks with different sizes and attributes. All of them are collaboration networks from the e-print arXiv and covers scientific collaborations between authors' papers. For each network we use different infection probability which is computed from the infection spread threshold , T_C, over which infection epidemic exists.

Network	Nodes	Edges	Inf. Prob
CA-CondMat	23133	93497	0.05
CA-AstroPh	18772	198110	0.02
CA-GrQc	5242	14496	0.07
pgp	24316	186936	0.06

Network attributes

5.2 Experimental Settings

We performed two different types of experiments for the offline approach. In the first type we start the infection propagation from one node (*single spreader)* and in the second there are five or ten nodes from which the infection begins (*multiple spreaders)*. There are two kind of initial seed sets, one consists of the nodes with the highest

degree of the network and the other consists of the highest k-core values. Several amounts of budgets were tried in order to observe the infection propagation. At first we remove 0.25% of the network's size, and we continue with 0.5%, 0.75%, 1%, 2%, 3% and 4% of the network's size. All of the removals are performed before the infection propagation begins.

For the online approach we perform experiments using only a single spreader for both node and edge removal. The initial seed set consist of the top degree node. As we described before, in this scenario we remove β number of nodes (or edges) at each SIR step. For the node removal case the budget varies from 20% to 60%. As far as the edge removal case is concerned the budget varies from 10% to 100%.

We simulated 15000 diffusion instances for every scenario and there was a deviation in results less than 1%.

5.3 Experimental Results

In this section we present the results of our experiments. The first set of results concerns the offline approach and the second one is about the online approach.

5.3.1 Offline approach

5.3.1.1 Single Spreader

In the experiments below we present the propagation of infection without any blocking. We can easily observe that *betweenness centrality* and *k-shell* metric have the worst performance. This occurs because nodes with high *betweenness centrality* have many connections with leaf nodes and nodes with high *k-shell* metric have many connections with nodes near to the core of the network. This means that the nodes we remove have

many one-hop neighbors and no *n*-hop, (*n>1*) neighbors and they are not influential spreaders. *Closeness centrality* up till 1% of the removed nodes has low-performance; after this amount of budget it has similar behavior to the *degree, pci* and *ncPci* measures. We observe that this similarity in performance of the last three metrics is caused by the common removed nodes. Nodes with high *degree*, has also high *pci* and *ncPci* metrics. *Pci* gives us the density of the one-hop neighborhood and so does *ncPci*. However, *ncPci* counts the real contribution of each node in contrast to *pci* which may count the contribution of a node twice or more. So, from all the above it is clear that *ncPci* has the best performance and the results confirm its significance.

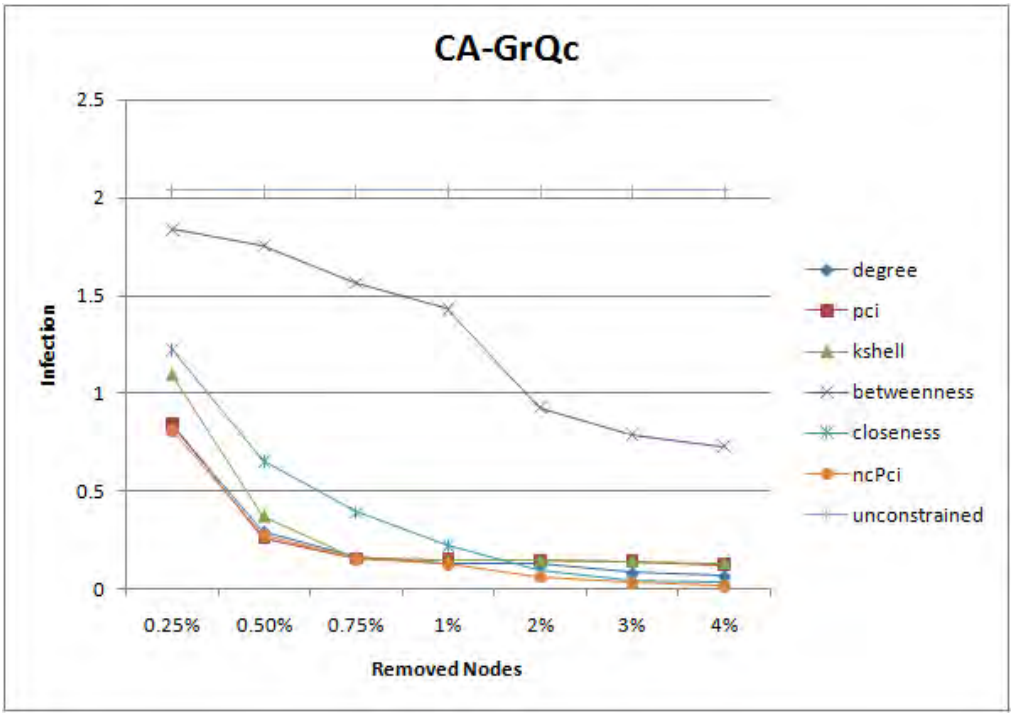

Fig.12: CA-GrQc, infection probability 7%, initial seed set : top degree

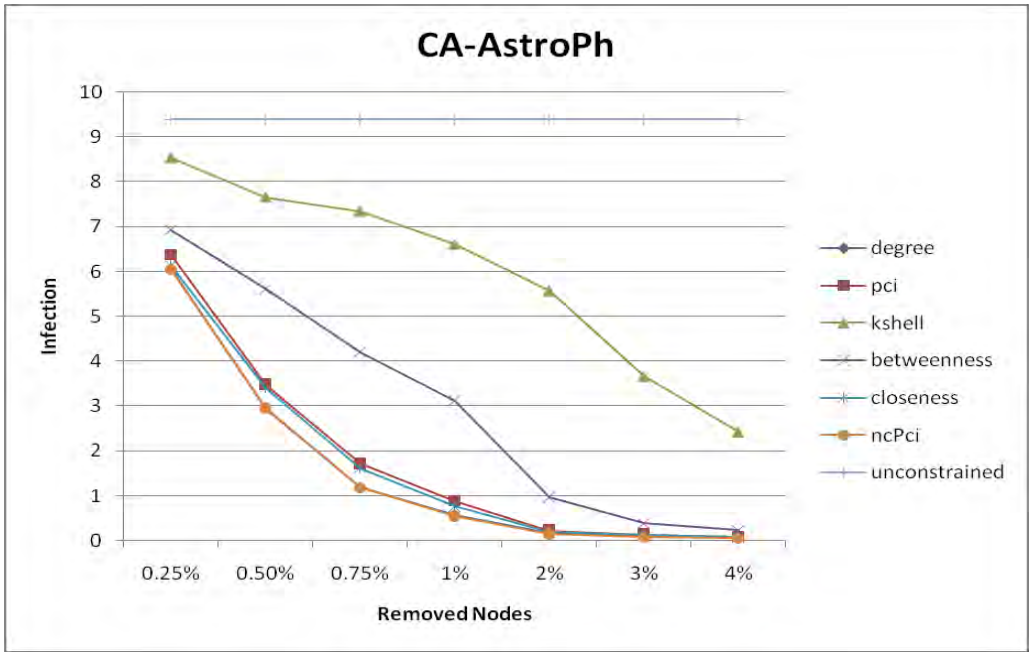

Fig.13: CA-AstroPh, infection probability 2%, initial seed set : top degree

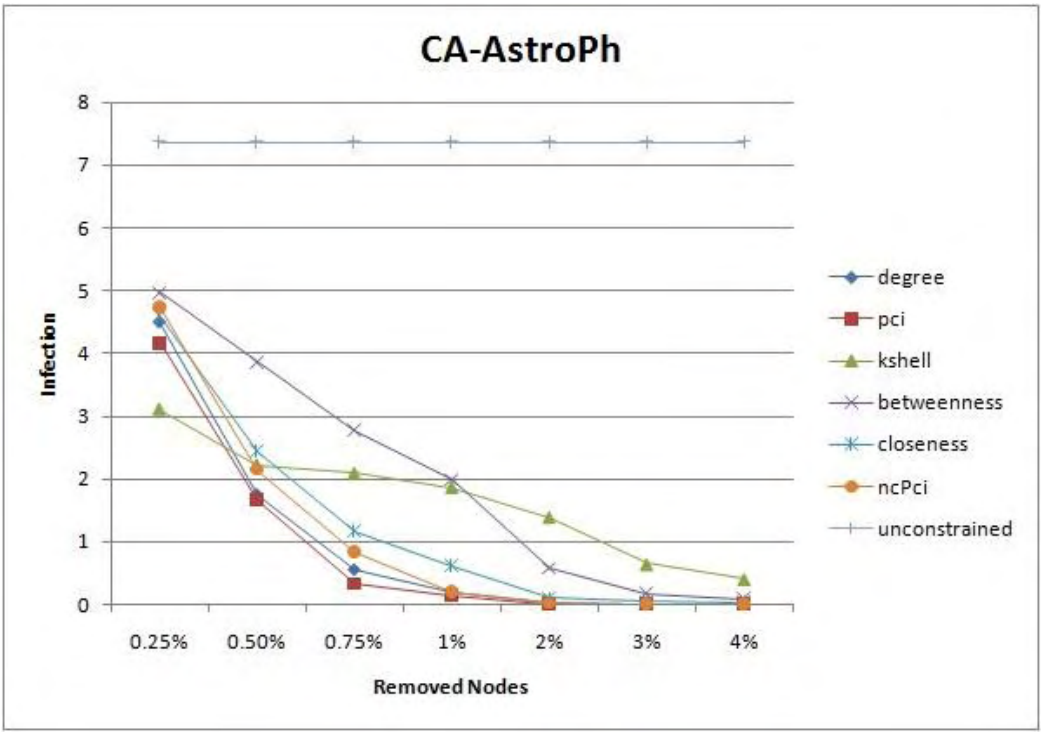

Fig.14: CA-AstroPh, infection probability 2%, initial seed set : top k-shell

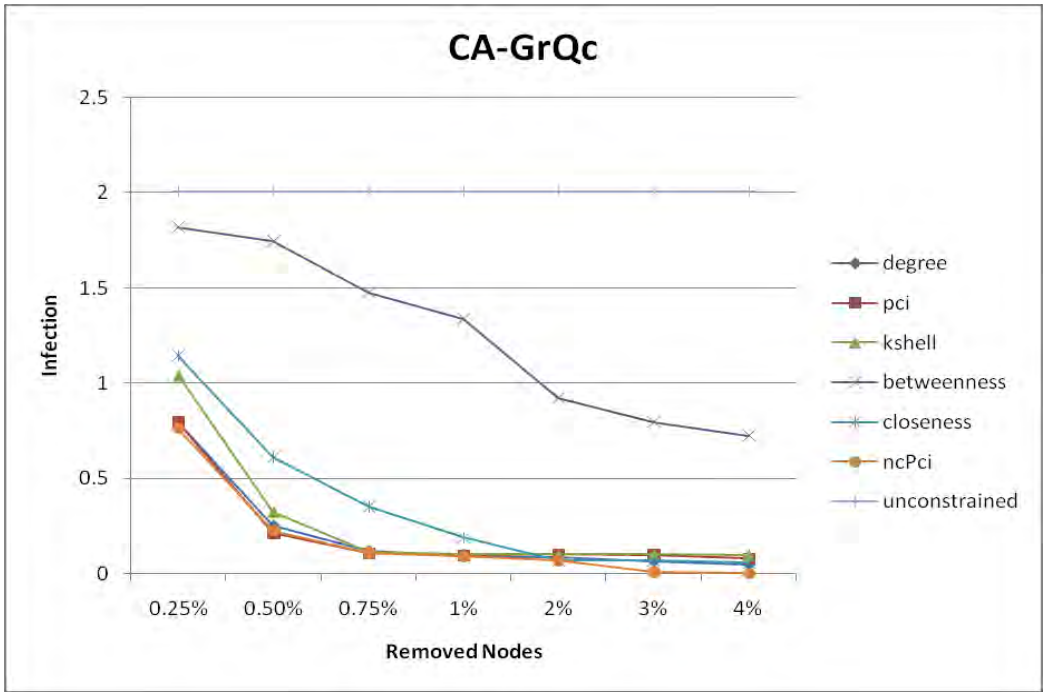

Fig.15: CA-GrQc, infection probability 7%, initial seed set : top k-shell

5.3.1.2 Multiple Spreader

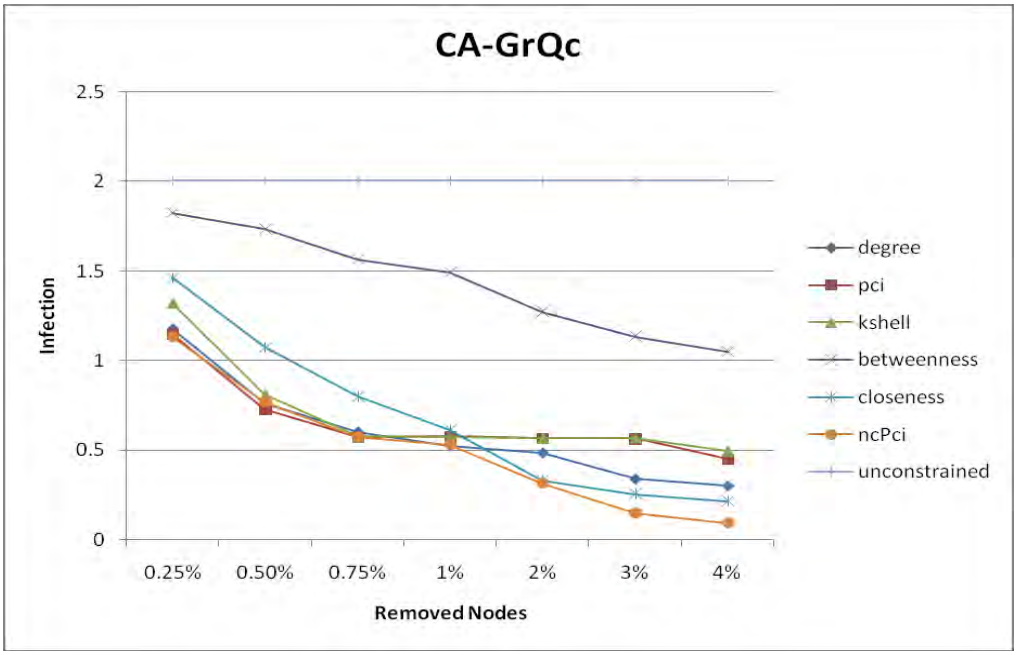

Fig.16: CA-GrQc, infection probability 7%, initial seed set : top 5 degree

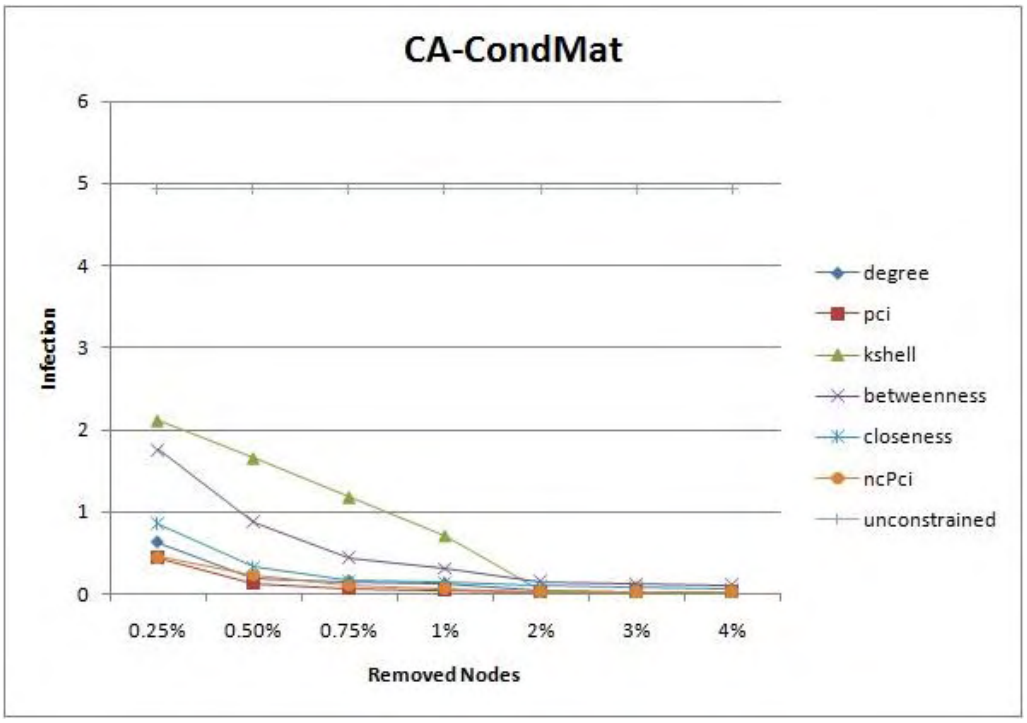

Fig.17: CA-CondMat, infection probability 5%, initial seed set : top 5 k-shell

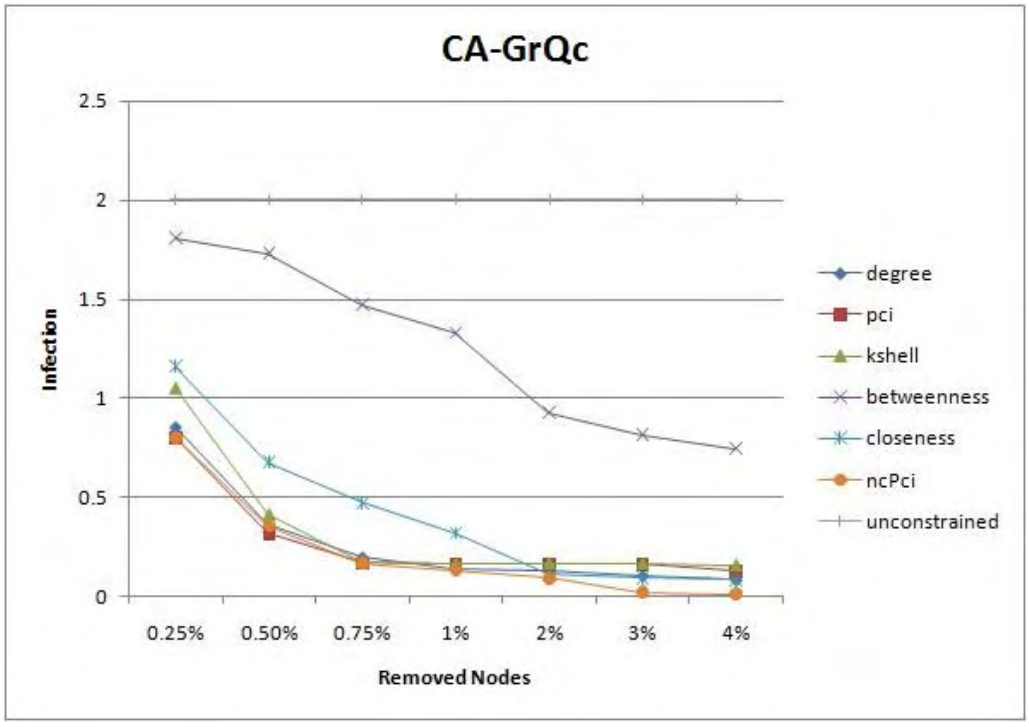

Fig.18: CA-GrQc, infection probability 7%, initial seed set : top 5 k-shell

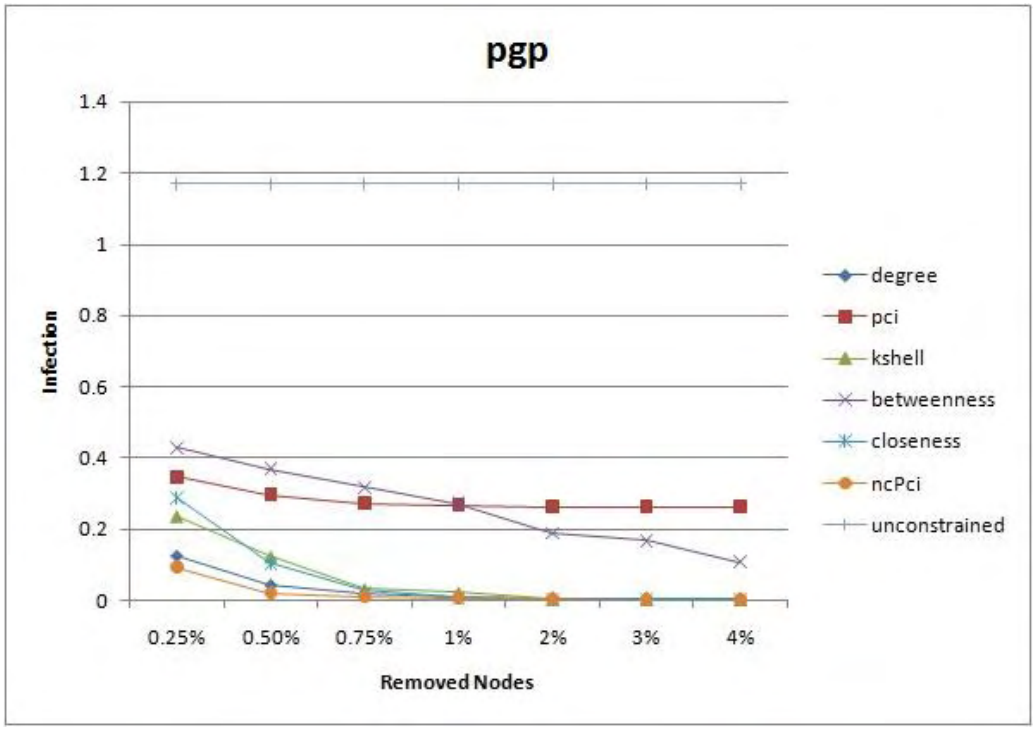

Fig.19: pgp, infection probability 6%, initial seed set : top 5 k-shell

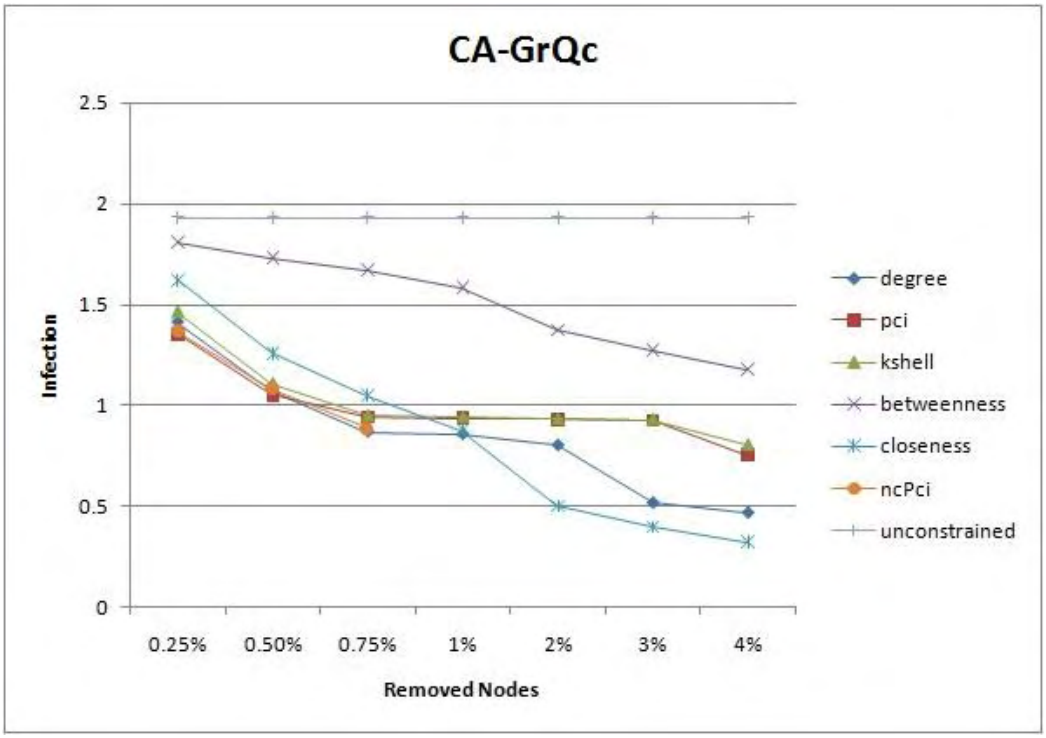

Fig.20: CA-GrQc, infection probability 7%, initial seed set : top 10 degree

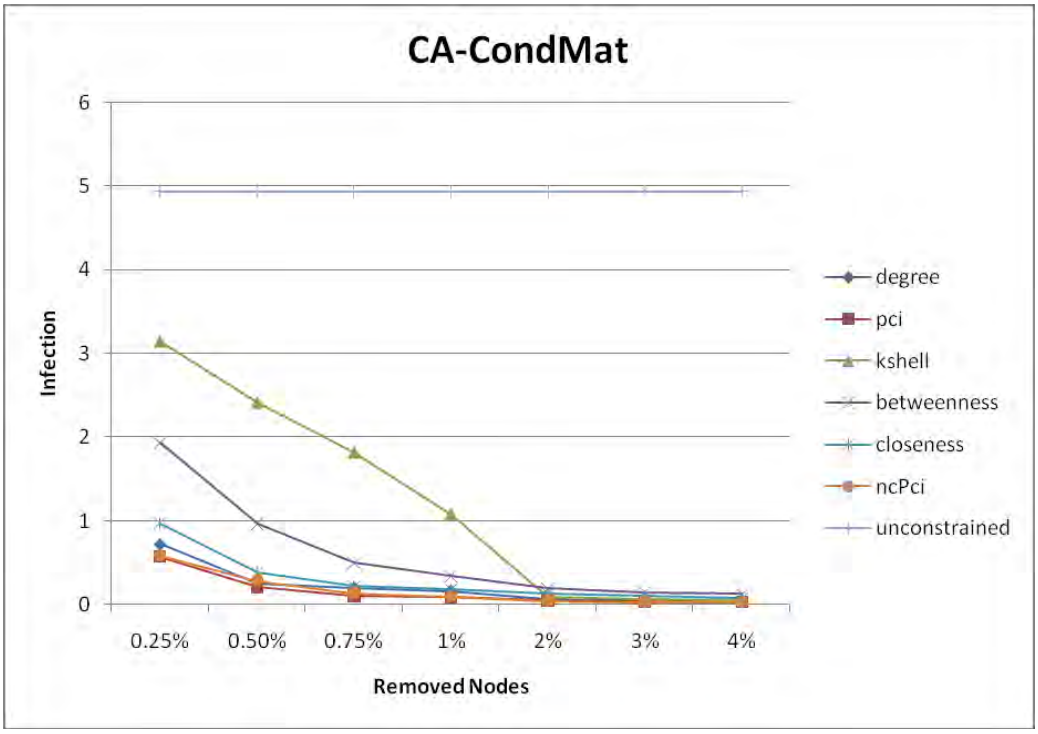

Fig.21: CA-CondMat, infection probability 5%, initial seed set : top 10 k-shell

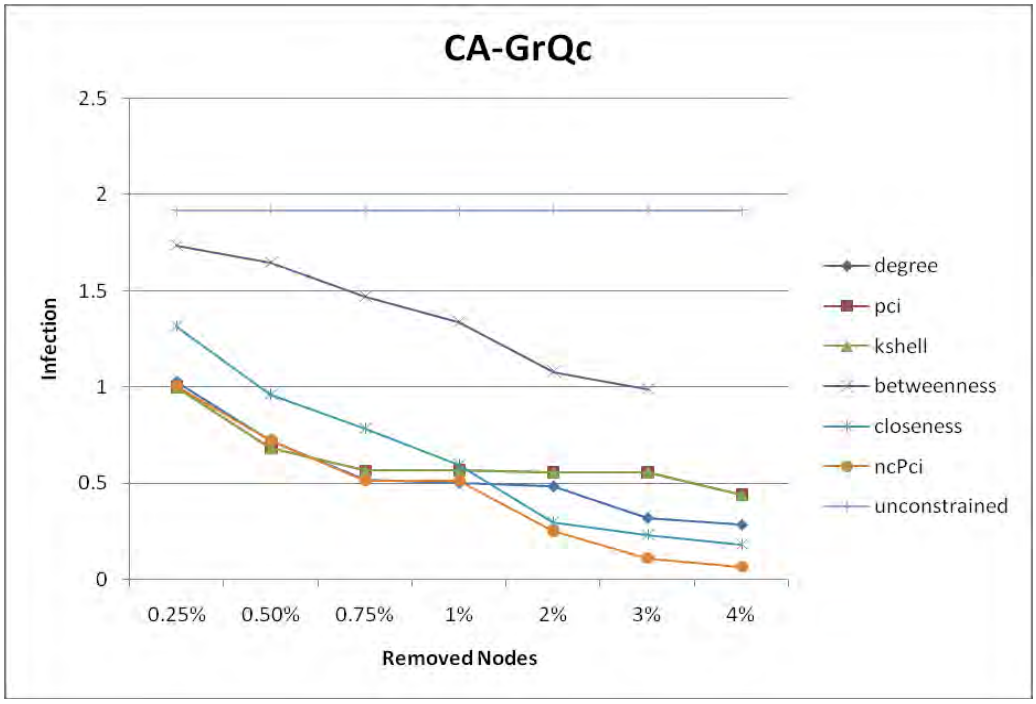

Fig.22: CA-GrQc, infection probability 7%, initial seed set : top 10 k-shell

5.3.2 Online approach

Here we present the results from the online approach experiments. rDegree looks to have better performance at first, however after 50% of the removed nodes degree has the same results. This occurs because the percentage of the removed nodes is large enough and we have common deleted nodes. As for node removal, degree centrality has much greater performance from the egoBetweenness and this makes sense as egoBetweenness takes into account only some of the one-hop neighbors in contrast to degree centrality which takes into account all of them.

5.3.2.1 Edge Removal

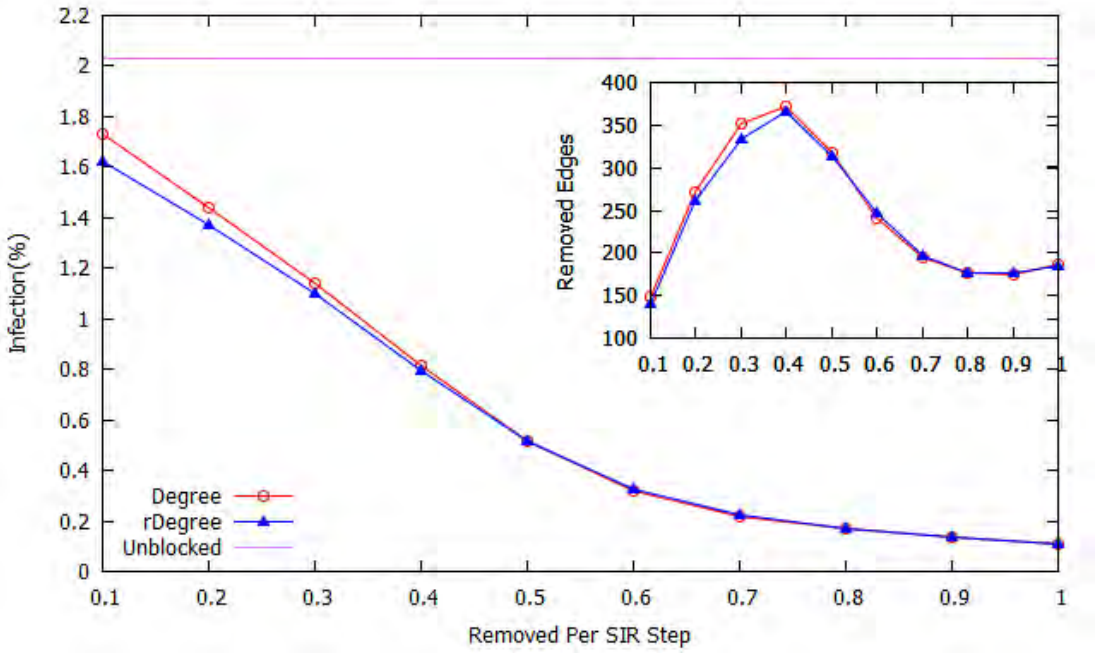

Fig.23: CA-GrQc, infection probability 7%, initial seed set : top degree

5.3.2.2 Node Removal

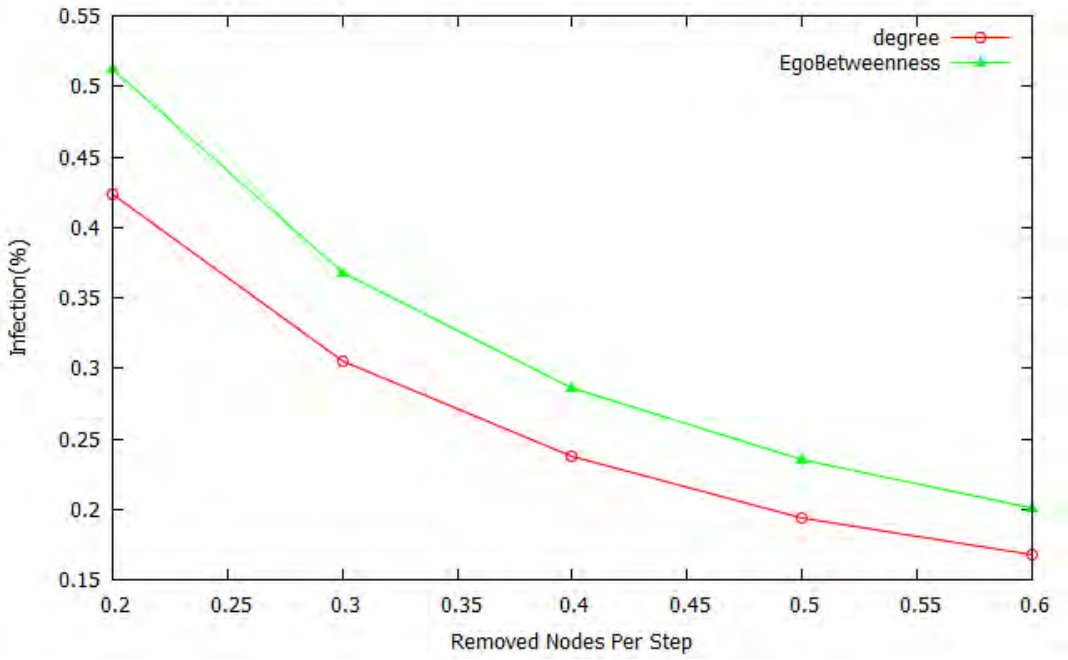

Fig.24: CA-GrQc, infection probability 7%, initial seed set : top degree

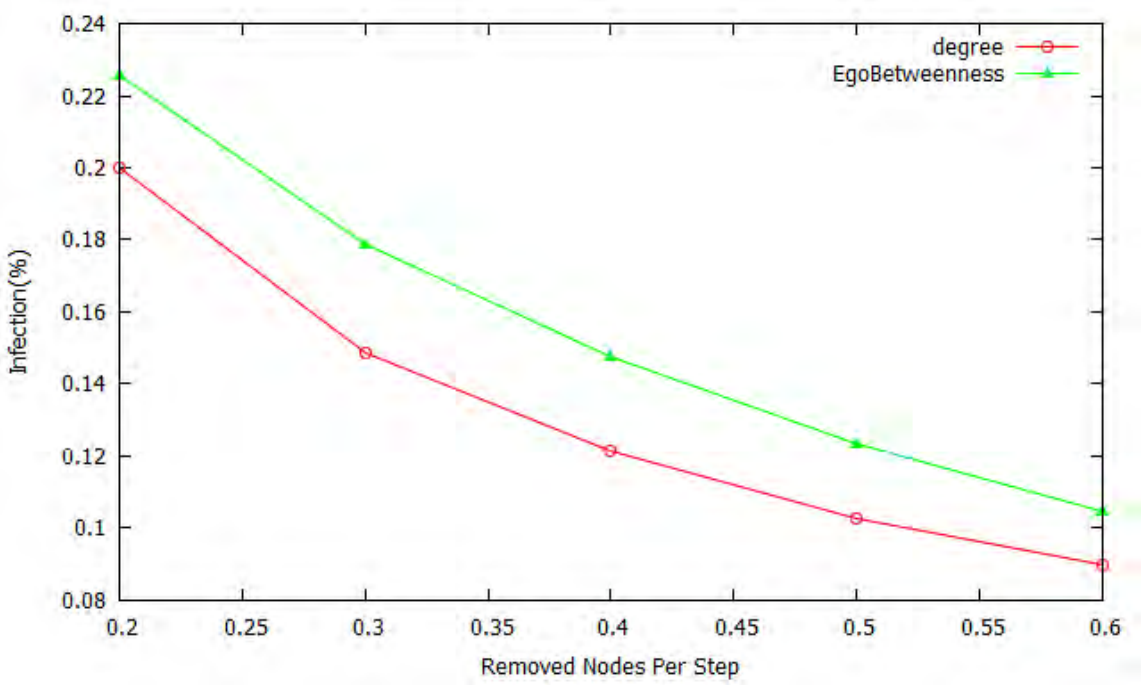

Fig.25: CA-GrQc, infection probability 7%, initial seed set : top degree

6 Conclusion & Future Work

In this work we try to constrain infection propagation in complex networks. We face the problem with two different approaches, offline and online, and with many methodologies. The results for the offline case, present the great performance of *ncPci, pci* and *degree centrality* with the first having the less infected nodes at the end of the propagation. *Degree* and *pci* are the most important competitors as in some cases succeed tiny infection percentage too. The results for the online case, present the great performance of the degree centrality. However, this case is a new approach and an unexplored region.

So as a future work, we plan to work on blocking the false rumor by removing edges or nodes while aiming to a better performance considering the way the aforementioned rumor is propagated through the complex network. This is going to be achieved by adjusting methods that have been used in the offline approach to the online one or implement new greedy algorithms.

References

[1] F. Lidan, L. Zaixin, W. Weili, T. Bhavani, M. Huan, and B. Yuanjun, “Least cost rumor blocking in social networks,” *ICDCS ’13 Proceedings of the 2013 IEEE 33rd International Conference on Distributed Computing Systems*, pp. 540–549, 2013.

[2] K. Gong, M. Tang, H. Ming, H. Zhang, D. Younghae, and Y.-C. Lai, “An efficient immunization strategy for community networks,” *PLOS ONE*, 2013.

[3] R. Cohen, S. Havlin, and A. D, “Efficient immunization strategies for computer networks and populations,” *Phys Rev Lett 91: 247901*, 2003.

[4] M. Salathe and J. J.H., “Dynamics and control of diseases in networks with community structure,” *PLOS Computational Biology*, 2010.

[5] H.-D. Laurent, A. Allard, J.-G. Young, and L. Dube, “Global efficiency of local immunization on complex networks,” *Physics and Society (physics.soc-ph); Populations and Evolution (q-bio.PE)*, 2013.

[6] Y. Yamada and T. Yoshiba, “A comparative study of community structure based node scores for network immunization,” *PLOS Computational Biology*, pp. 328–337, 2012.

[7] K. Masahiro, S. Kazumi, and M. Hiroshi, “Minimizing the spread of contamination by blocking links in a network,” *AAAI’08 Proceedings of the 23rd national conference on Artificial intelligence, vol. 2*, pp. 1175–1180, 2010.

[8] K. Masahiro, S. Kazumi, and N. Ryohei, “Extracting influential nodes for information diffusion on a social network,” *AAAI’07 Proceedings of the 22nd national conference on Artificial intelligence, vol. 2*, pp. 1371–1376, 2007.

[9] H.-F. Zhang, K.-Z. Li, X.-C. Fu, and B.-H. Wang, “An efficient control strategy of epidemic spreading on scale-free networks,” *Chinese Phys. Lett. 26 068901*, vol. 6, 2009.

[10] T. Hanghang, B.-A. Prakash, T. Elisiassi-Rad, M. Faloutsos, and C. Faloutos, “Gelling, and melting, large graphs by edge manipulation,” *CIKM ’12 Proceedings of the 21st ACM international conference onInformation and knowledge management*, vol. 6.

[11] B.-A. Prakash, D. Chakrabarti, M. Faloutsos, N. Valler, and C. Faloutsos, “Threshold conditions for arbitrary cascade models on arbitrary networks,” *ICDM ’11 Proceedings of the 2011 IEEE 11th International Conference on Data Mining*, vol. 33, pp. 549–575, 2011.

[12] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, “Epidemic spreading in real networks: An eigenvalue viewpoint,” *Reliable Distributed Systems. Proceedings. 22nd International Symposium*, pp. 25–34, 2003.

[13] A.-N. Bishop and I. Shames, “Link operations for slowing the spread of disease in complex networks,” *(EPL) Europhysics Letter*, vol. 95, 2011.

[14] C. Kuhlman, G. Tuli, S. Swarup, M. Marathe, and S. Ravi, “Blocking simple and complex contagion by edge removal,” *Data Mining (ICDM), 2013 IEEE 13th International Conference*, pp. 399 – 408, 2013.

[15] W. Chen, A. Collins, R. Cummings, T. Ke, Z. Liu, D. Rincon, X. Sun, Y. Wang, W. Wei, and Y. Yuan, “Influence maximization in social networks when negative opinions may emerge and propagate,” *Proceedings of the Eleventh SIAM International Conference on Data Mining, SDM 2011*, 2011.

[16] J. Kostka and R. Oswald, Y-A.and Wattenhofer, “Word of mouth: Rumor dissemination in social networks,” *Structural Information and Communication Complexity*, vol. 5058, pp. 185–196, 2008.

[17] N. Pathak, A. Banerjee, and J. Srivastava, “A generalized linear threshold model for multiple cascades,” *Data Mining (ICDM), 2010 IEEE 10th International Conference*, pp. 965–970, 2010.

[18] H. Xinran, S. Guojie, W. Chen, and J. Qingye, “Influence blocking maximization in social networks under the competitive linear threshold model technical report,” *Social and Information Networks (cs.SI); Physics and Society (physics.soc-ph)*, 2011.

[19] N.-P. Nguyen, Y. Guanhua, M.-T. Thai, and E. S., “Containment of misinformation spread in online social networks,” *WebSci ’12 Proceedings of the 4th Annual ACM Web Science Conference*, pp. 213–222, 2012.

[20] C. Budak, D. Agrawal, and A.-E. Abbabi, “Limiting the spread of misinformation in social networks,” *WWW ’11 Proceedings of the 20th international conference on World wide web*, pp. 665–674, 2011.

[21] http://snap.stanford.edu/data/

[22] http://konect.uni-koblenz.de/