ROYAUME DU MAROC
Université Mohammed V - RABAT -
جامعة محمد الخامس – الرباط –
Faculté des sciences
كلية العلوم
CENTRE D'ETUDES DOCTORALES - SCIENCES ET TECHNOLOGIES

N° d'ordre : **3355**

THESE

En vue de l'obtention du : ***DOCTORAT***

Structure de Recherche : ***Laboratoire de Matière Condensée et Sciences Interdisplinaires (LaMCScI)***
Discipline : ***Physique Informatique***
Spécialité : ***Informatique et Analyse des Réseaux complexes***

Présentée et soutenue le 24/10/2020 par :

Ahmed ALWEIMINE

Modeling and simulation of routing protocols and computer virus spreading in complex networks

JURY

Mr. Abdelilah BENYOUSSEF	PES, Membre résident à l'académie Hassan II des Sciences et Techniques, Rabat	Président/ Rapporteur
Mr. Hamid EZ-ZAHRAOUY	PES, Université Mohammed V, Rabat, Faculté des Sciences	Directeur de Thèse
Mr. Abdeljalil RACHADI	PA, Université Mohammed V, Rabat, Faculté des Sciences	Co-Directeur de Thése
Mr. Abdelmajid AINANE	PES, Université Moulay Ismail, Meknès, Faculté des Sciences	Rapporteur / Examinateur
Mr. Youssef EL AMRAOUI	PES, Université Mohammed V, Rabat, Faculté des Sciences	Examinateur

Année Universitaire : 2020-2021

Acknowledgements

All works presented in this thesis have been carried out in the Laboratory of Condensed Matter and Interdisciplinary Sciences (LaMCScI). Faculty of Sciences - Mohammed V of Rabat, under the supervision of the professor Hamid EZ-ZAHRAOUY and the co-supervision of the professor Abdeljalil RACHADI.

So to begin, I'm deeply grateful to my thesis reporter the professor Abdelilah BENYOUSSEF , who did me honor to head the chair of jury for this thesis.

I would like to acknowledge my thesis supervisor the professor Hamid EZ-ZAHRAOUY for his support and help during these years. He consistently allowed this paper to be my own work, meanwhile steered me in the right direction whenever needed it.

My sincere thanks go also to my co-supervisor and reviewer the professor Abdeljalil RACHADI. I'm very grateful for all the time he devoted to provide useful suggestions and many relevant discussions throughout my research.

I would like also to express my appreciation to my thesis reporter and jury member the professor, Abdelmajid AINANE for accepting to evaluate my thesis and research works.

I'm very thankful to the professor Youssef El AMRAOUI who was also jury member for accepting to evaluate this manuscript and all the works that it includes.

Special thanks to all of my friends and my classmates for their backup and their endless friendship.

Abstract

This thesis focuses on modeling and simulation of routing protocols to reduce congestion and on the problem of the virus propagation in networks. In this thesis, we have introduced two main contributions:

First, for the purpose of alleviating the congestion in traditional shortest path (SP) strategy and to deal more with priority traffic concept in internet, we have proposed a new prioritization model of traffic flow where packets under shortest path strategy are prioritized according to their destination. We found that the prioritization of nodes with high degree (hubs) is always more efficient than the prioritization of nodes with small degree or the random prioritization of nodes.

Second, we studied the effectiveness of local routing protocols and their additional algorithms; next-nearest neighbors (NNN) and restrictive queue-length algorithm (RQL) in term of robustness in computer virus spreading. It is found that, under the additional algorithm RQL, local routing protocols become highly secured and overcome surprisingly the efficient path (EP) routing protocol.

Key Words (5): Complex networks, traffic engineering, routing protocols, computer virus, computer sciences.

Résumé

Cette thèse porte sur la modélisation et la simulation de protocoles de routage pour réduire la congestion, et sur le problème de la propagation des virus dans les réseaux complexes. Dans cette thèse, nous avons introduit deux contributions principales:

Premièrement, dans le but de réduire la congestion dans la stratégie de routage de plus court chemin (SP) et de mieux prendre en compte le concept de trafic prioritaire sur Internet, nous avons proposé un nouveau modèle de hiérarchisation des flux de trafic dans lequel les paquets sous la stratégie de plus court chemin sont hiérarchisés en fonction de leur destination. Nous avons constaté que la priorisation de trafic destiné aux nœuds avec degré de connexion élevé (hubs) est toujours plus efficace que la priorisation de trafic destiné aux nœuds avec un faible degré ou la priorisation aléatoire de traffic

Deuxièmement, nous avons étudié l'efficacité des protocoles de routage locaux et de leurs algorithmes supplémentaires. Voisins les plus proches (NNN) et algorithme restrictif de longueur de la file d'attente (RQL) en termes de robustesse dans la propagation de virus informatiques. Il s'avère que, sous l'algorithme supplémentaire RQL, les protocoles de routage locaux sont hautement sécurisés et dépassent de manière surprenante le protocole de routage EP (voie efficace).

Mot clés : Réseaux complexes, ingénierie du trafic, protocoles de routage, virus informatique, informatiques.

Résumé détaillé

Vue le rôle crucial que jouent les réseaux des données dans la vie humaine, il n'est plus possible d'ignorer le problème de la congestion et de la propagation des virus informatique dans ces systèmes pratiques.

Dans la dernière décennie, l'accroissement du nombre d'utilisateurs actifs dans ces réseaux ne cesse pas d'augmenter de façon redoutable, ce qui déclenche des problèmes comme la congestion du trafic et la propagation des virus informatique et ce qui influence aussi la qualité des services fournis aux clients. Pour surmonter ces problèmes, les études scientifiques sont focalisées essentiellement sur des méthodes capables de limiter la propagation des virus, et sur la conception des stratégies efficaces de routage capables de rester à la hauteur des progrès rapides des réseaux en satisfaisant les demandes des utilisateurs.

L'intérêt de cette thèse est d'étudier les phénomènes qui apparaissent dans les réseaux complexes et notamment la propagation des virus informatique, le trafic et le rapport entre les deux dans les réseaux de type sans échelle « scale-free ». L'objectif du trafic est d'augmenter sa capacité sans frais supplémentaires au niveau technique et économique. En outre, l'étude de la propagation des virus est faite pour connaître sa dynamique afin de trouver des méthodes efficaces pour contrôler la prévalence de l'infection. Alors, nous proposons ci-dessous les résumés des chapitres constituant cette thèse.

Chapitre I

Au début de ce chapitre, nous avons introduit quelques notions importantes qui sont fréquemment utilisées dans la science des réseaux, également elles sont empruntées de la théorie des graphes. Ensuite, nous avons étudié les caractéristiques les plus intéressantes qui sont détectées dans des réseaux réels, et nous avons exposé une étude concernant la modélisation des réseaux complexes existés dans la réalité, afin de justifier le choix d'un réseau sans échelle modélisé par Barabási-Albert, comme une plateforme du trafic et propagation des virus.

Chapitre II

Dans le deuxième chapitre, nous avons présenté les phénomènes qui apparaissent dans les réseaux complexes. En effet, il existe deux types de processus dynamiques qui ont été étudiés profondément, à savoir la propagation de virus informatique et la dynamique du trafic. Nous avons défini les principaux concepts du trafic. Ensuite, nous avons entamé une revue globale sur les stratégies de routage les plus utilisées pour améliorer la capacité du trafic des données, tout en distinguant la politique derrière chaque approche et en précisant les avantages et les inconvénients de chaque stratégie.

Chapitre III

Dans ce chapitre, nous avons proposé un nouveau modèle de hiérarchisation des flux de trafic dans lequel les paquets sous la stratégie de plus court chemin sont hiérarchisés en fonction de leur destination. Nous avons constaté que la priorisation de trafic destiné aux nœuds avec degré de connexion élevé (hubs) est toujours plus efficace que la priorisation de trafic destiné aux nœuds avec un faible degré ou la priorisation aléatoire de trafic

Chapitre IV

Dans ce chapitre, Nous avons étudié la propagation des virus. D'abord, nous avons commencé par les définitions fondamentales de la propagation des virus informatique. Ensuite, nous avons présenté les méthodes les plus mémorables et les plus utilisables par les chercheurs pour comprendre la dynamique de la propagation des virus dans les réseaux complexes.

Chapitre V

Dans ce chapitre, nous nous sommes focalisés sur l'étude de lien entre la propagation de virus informatique et le processus de transport. Également nous avons étudié les effets des stratégies de routage sur la propagation des virus dans les réseaux complexes. nous avons étudié l'efficacité des protocoles de routage locaux et de leurs algorithmes supplémentaires,

Voisins les plus proches (NNN) et algorithme restrictif de longueur de la file d'attente (RQL) en termes de robustesse dans la propagation de virus informatiques. Il s'avère que, sous l'algorithme supplémentaire RQL, les protocoles de routage locaux sont hautement sécurisés et dépassent de manière surprenante le protocole de routage EP (voie efficace).

ملخص

بشكل عام ، فإن معظم الدراسات في مجال بروتوكولات التوجيه وانتشار فيروسات الكمبيوتر في الشبكات المعقدة لها هدفين رئيسيين. الهدف الأول هو تصميم بروتوكولات توجيه فعالة لتقليل الازدحام وتحسين أداء نقل المعلومات دون إغفال التكاليف التقنية والمالية. الهدف الثاني هو تحقيق مستوى عال من الأمان لحماية المعلومات دون التأثير على قدرة الشبكة الإستيعابية وعلى عملية تدفق المعلومات.

في هذا الصدد، تتناول هذه الأطروحة نمذجة ومحاكاة بروتوكولات التوجيه لتقليل الازدحام ومشكلة انتشار الفيروسات في الشبكات المعقدة الإنترنت كنموذج. في هذه الأطروحة ، قدمنا مساهمتين رئيسيتين:

أولاً ، من أجل تقليل الازدحام في بروتوكول توجيه المسار الأقصر المعمول بها في شبكة الأنترنت ومراعاةً لمفهوم تدفق المعلومات ذات الأولوية القصوى على الإنترنت ، اقترحنا استراتيجية جديدة لتحديد أولويات المعلومات بناءا على أهمية الوجهة المستقبلة للمعلومة. حسب النتائج فقد وجدنا أن استراتيجية أولوية المعلومات المتجهة صوب العقد ذات الدرجة العالية (محاور) تكون دائمًا أكثر كفاءة من استراتيجية أولوية المعلومات المتجهة صوب العقد ذات الدرجة الصغيرة أواستراتيجية الأولوية العشوائية للمعلومات.

ثانياً ، بحثنا في فعالية بروتوكولات التوجيه المحلية وخوارزمياتها الإضافية،خوارزمية الجار من الدرجة الثانية وخوارزمية طول قائمة الانتظار التقييدية من حيث مقاومتها في انتشار فيروسات الكمبيوتر بالشبكات الضخمة كالإنترنت. حسب النتائج اتضح أنه ، باستخدام الخوارزمية الإضافية (تقييد قائمة الإنتظار) ، تكون بروتوكولات التوجيه المحلية آمنة للغاية وتتجاوز مقاومتها للفيروسات بشكل مذهل بروتوكول التوجيه (القناة الفعالة).

.

الكلمات المفتاحية : الشبكات المعقدة ، هندسة حركة المعلومات ، بروتوكولات التوجيه ، فيروسات الكمبيوتر ، علوم الكمبيوتر.

Contents

Acknowledgements **2**
Abstract……. **3**
Résumé…… **3**
Résumé détaillé **5**
ملخص……. **8**
Contents……..... **9**
List of figures. **11**
List of tables.. ..... **14**
Acronym List. **15**
General introduction **16**
Chapter I. Introduction to network sciences and network modeling **23**
I.1 Introduction:..... 24
I.2 Network and graph..... 25
I.2.1 Network types 26
I.2.2 Networks characteristics 27
I.3 Random network models 33
I.3.1 Erdös-Rényi model..... 34
I.3.2 Small world network 36
I.3.3 Real Networks are Not Random 40
I.4 Scale free networks 42
I.4.1 Scale free property in real networks 42
I.4.2 Barabasi-Albert model..... 43
I.5 Conclusion: 46
Chapter II. Introduction to traffic sciences and traffic routing modeling **47**
II.1 Introduction to traffic routing: 48
II.1.1 Basic concepts : 48
II.1.2 Routing protocol:..... 48
II.2 Traffic routing protocol modeling: 50
II.2.1 Traffic Models :..... 50
II.2.2 Traffic measurements: 50
II.2.3 Routing protocols: 53

II.3 Conclusion: 63
Chapter III. Prioritization of traffic flow in complex networks 64
III.1 Introduction: 65
III.2 The shortest path problem (Dijkstra algorithm) 66
III.2.1 Dijkstra algorithm: 66
III.2.2 Exemple: 69
III.3 Implementing beneficial prioritization of traffic flow in complex networks 74
III.3.1 Description and simulation steps 74
III.3.2 Simulation results and discussion 75
III.4 Conclusion: 81
Chapter IV. Introduction to computer virus and propagation modeling 82
IV.1 Introduction: 83
IV.1.1 Defintion: 83
IV.1.2 Virus propagation 84
IV.1.3 Anti-virus or Immunization 86
IV.2 Virus propagation modeling 87
IV.2.1 Susceptible Infected Model (SI): 87
IV.2.2 Susceptible Infected Recovered (SIR) Model: 88
IV.3 Conclusion: 90
Chapter V. Local routing protocols performance 91
V.1 Introduction 92
V.2 Models and routing strategies 93
V.2.1 Network types 93
V.2.2 The traffic model: 93
V.2.3 Epidemic model: 93
V.2.4 The routing strategies: 94
V.3 Results and Analysis: 95
V.4 Conclusion: 105
General conclusion 106
References…. 109
List of publications & communications 119

List of figures

Figure I-1 : A contemporary map of Königsberg (now Kaliningrad, Russia) during Euler's time. 24
Figure I-2 : The figure shows a small subset of (a) the Internet, where routers (specialized computers) are connected to each other; (b) the Hollywood actor network, where two actors are connected if they played in the same movie; (c) a protein-protein interaction network. 26
Figure I-3 : Example of the degree distribution of a network with N=4 nodes. 29
Figure I-4 : The adjacency matrices of two different types of networks with N=4; (a) undirected graph, (b) directed network. 30
Figure I-5 (a) Distance or path shown in orange between nodes 1 and 6 (b) The shortest paths between nodes 1 and 7 , There can be multiple paths of the same length, as illustrated by the two paths shown in orange and grey. 31
Figure I-6 : Example of calculation of the clustering coefficient in a network. 33
Figure I-7 : Top Row: Three realizations of a random network generated with the same parameters p=1/6 and N=12. Despite the identical parameters, the networks not only look different, but they have a different number of links as well (L=10, 10, 8). Bottom Row:Three realizations of a random network with p=0.03 and N=100. 35
Figure I-8: Six Deegree of Separation: According to six degrees of separation two individuals, anywhere in the world, can be connected through a chain of six or fewer acquaintances 36
Figure I-9 : The figure shows, (a)the length distribution of the completed chains In Milgram's experiment, the mean of the distribution was 5.2 [25]. (b) The distance distribution, pd, for all pairs of Facebook users worldwide and within the US only [18]. 37
Figure I-10: The distance distribution for all pairs of Microsoft Messenger network users [38]. 38
Figure I-11: The random rewiring procedure of Watts and Strogatz model; we start with N=16 nodes, each node is linked to its first K=4 neighbors. 39
Figure I-12: Degree Distribution of Real Networks: The degree distribution of the (a) Internet, (b) science collaboration network, and (c) protein interaction network. 41
Figure I-13: Networks are not static, but grow via the addition of new nodes: The evolution of the number of WWW hosts, documenting the Web's rapid growth. After http://www.isc.org/solutions/ survey/history. 42
Figure I-14: The sequence of images shows nine subsequent steps of the Barabási-Albert model. Empty circles mark the newly added node to the network, which decides where to connect its two links (m=2) using preferential attachment [26]. 43
Figure I-15: The degree distribution of a network generated by the Barabási-Albert model [26]. 44
Figure I-16 : The topology of the Internet An iconic representation of the Internet topology at the beginning of the 21st century. The image was produced by CAIDA, an organization based at University of California in San Diego, devoted to collect, analyze, and visualize Internet data. The map illustrates the Internet's scale-free nature: A few highly connected hubs hold together numerous small nodes. 45
Figure II-1 : Determination of a path between a source host and destination host with the lowest cost in a computer network. 49
Figure II-2 : transition from free flow to congested phase. 51
Figure II-3: (Color online) Critical Rc vs Lc under global routing strategy with network parameters (a) N=1000, m=2 and 5; (b) N =5000, m=2 and 5. The data are obtained by averaging Rc over ten network realizations. ... 54
Figure II-4: (a) the traffic capacity Rc and (b) the average shortest path length $Lave$ versus the fraction of new added links over the total L existing links $f\alpha$ under the shortest path routing strategy. In each figure, the gray and black colors are for the cases with m0 =m=3 and m0=m =5, respectively. The network size is set to N =500. The circles, asterisks, and plus signs are the results of the strategy in [102, the random strategy, and the strategy that adds links among hub nodes, respectively. 55

Figure II-5: The order parameter η vs R under the three routing strategies: shortest path, efficient path, and global dynamic. Network size N=500; delivering capacity C=1[72] 57
Figure II-6: The critical Rc versus α with network size N=1000 and constant node delivering capacity C=10. The maximum of Rc corresponds to α =−1 marked by a dotted line. 58
Figure II-7: The order parameter η as a function of generating rate R for different value of parameter β. Other parameters are delay = 0, C = 5 and N = 1000 [75]. 59
Figure II-8 : Order parameter η versus R for different routing strategies. Local static routing strategy, local dynamic routing strategy and local pheromone routing strategy. N = 1024, the mean degree is K = 10 and the delivering capability of each node is C = 5. 60
Figure II-9: The Rc vs network size N in the shortest path protocol for BA scale-free networks with size N=1000 and average degree <k> = 8, the delivery resource is allocated with uniform (red color) and (betweenness allocation) (black), respectively. 62
Figure III-1 : Example of the calculation of a length of a given path and the distance between two nodes. 67
Figure III-2: Illustration of the relaxation proceeding performed when discovering new path with less cost. 69
Figure III-3: Step zero of Dijkstra algorithm; (a) initialization of the tables and (b) graphic representation of the estimated distances 70
Figure III-4: Step one of Dijkstra algorithm exploration of the source s and updating initial information. 70
Figure III-5: Step two of Dijkstra algorithm exploration of the first neighbors of the source s; we start from the closest neighbor to the source s (node a) and update the information of step one. 71
Figure III-6: Step two of Dijkstra algorithm exploration of the first neighbors of the source s; we move to the next closest neighbor to the source s (node b) and continue the update of the information from step one. 71
Figure III-7: Step tree of Dijkstra algorithm exploration of the next neighbors of the source s; we start from the closest neighbor to the source s (node c) and update the information of step two. 72
Figure III-8: End of Dijkstra algorithm; by visiting the last node in the graph, we have processed all the nodes and no more relaxation is possible. 72
Figure III-9: The SP tree related to the node s. 73
Figure III-10 : The corresponding SP routing table established using Dijkstra algorithm 73
Figure III-11 : (a)-(c) The order parameter as a function of packets generation rate R for each type of traffic High priority traffic (HPT), Low Priority Traffic (LPT) , and shortest path without priority (WP) for different values of prioritized fraction from hubs nodes $f = 0.1$, $f = 0.4$, and $f = 0.8$. N=500, m=2 and <k>=4. 77
Figure III-12: (a)-(c) The order parameter as a function of packets generation rate R for each type of traffic High priority traffic (HPT), Low Priority Traffic (LPT) , and shortest path without priority (WP) for different values of prioritized fraction from peripheral nodes $f = 0.1$, $f = 0.4$, and $f = 0.8$. N=500, m=2 and <k>=4. 78
Figure III-13: (a)-(b) The order parameter as a function of prioritized fraction nodes f for each type of traffic High priority traffic (HPT), Low Priority Traffic (LPT), for different values of α . α=1 Hub are prioritized, α=-1 Peripheral nodes are prioritized, and $\alpha = 0$ for the random prioritization. $N = 500, m = 2$ and $< k >= 4$ 79
Figure III-14: (a)-(c) The traveling time as a function of packets generation rate R for each type of traffic High priority traffic (HPT), Low Priority Traffic (LPT) , and shortest path without priority (WP) for different values of prioritized fraction from hubs nodes $f = 0.1$, $f = 0.4$, and $f = 0.8$. N=500, m=2 and <k>=4. 80
Figure IV-1 : Mobile phone viruses 84
Figure IV-2: Life Cycle of Computer Virus. 85
Figure IV-3: Transmission diagram of the SI model. 88
Figure IV-4: SIR Model 89
Figure V-1: the proportion of infected nodes in the whole network for different values of probability infection as a function of time in the free flow phase for a fixed packet generating rate R=3. With an initial proportion of

infected nodes f_0=10% in the SI model using (a) Local static routing protocol, (b) Local dynamic routing protocol. $N = 1000, < k >= 10$..... *95*
Figure V-2: the proportion of infected nodes in the whole network for different values of probability infection β as a function of time in the free-flow phase for a fixed R=3, with an initial proportion of infected nodes f0=10% in the SI model using the addition additional Next-Nearest- Neighbor algorithm in (a) Local static routing protocol, (b) Local dynamic routing protocol. $N = 1000, < k >= 10.$ *97*
Figure V-3: the proportion of infected nodes in the whole network for different values of probability infection β as a function of time in the free-flow phase for a fixed R=3, with an initial proportion of infected nodes f_0=10% in the SI model using the additional Restrictive- Queue-Length algorithm in (a) Local static routing protocol, (b) Local dynamic routing protocol. $N = 1000, < k >= 10.$ *99*
Figure V-4: the proportion of infected nodes in the whole network for fixed probability infection β=0.4, as a function of time in the free-flow phase for a fixed R=3, with an initial proportion of infected nodes f_0=10% in the SI model using the additional Restri Restrictive-Queue-Length algorithm with Node-duplication-avoidance algorithm in local static routing protocol (a) sparse network $< k >= 4$*, (b) dense network* $< k >= 10.$ $N = 1000.$ *101*
Figure V-5: the proportion of infected nodes in the whole network for different values of probability infection β as a function of time in the free-flow phase for a fixed R=3, with an initial proportion of infected nodes f_0=10% in the SI model using (a) Shortest path, (b) Efficient path. $N = 1000, < k >= 10$ *103*
Figure V-6: the proportion of infected nodes in the whole network in local routing protocol under RQL, Shortest path (SP) and efficient path (EP). For a fixed probability infection rate β=0.4 as a function of time in the free-flow phase for a fixed R=3, and with an initial proportion of infected nodes $f0 = 10\%.$ *in the SI model using* $N = 3000. < k >= 10.$ *104*

List of tables

Table V-1 : The average path length under LSRQL, SP and EF routing strategies in dense and sparse network 102

Acronym List

BA	**Barabási-Albert**
BC	**Betweenness Centrality**
ER	**Erdös-Rényi**
EP	**Efficient Path**
FIFO	**First-In-First-Out**
GD	**Global Dynamic**
HPT	**High prioritized traffic**
LS	**Local Static**
LD	**Local Dynamic**
LPT	**Low Prioritized Traffic**
NNN	**Next-Nearest-Neighbor**
NDA	**Node duplication Avoidence**
RQL	**Restrictive-Queue-Length**
SP	**Shortest Path**
SF	**Scale-Free**
SW	**Small World**
SI	**Susceptible-Infected**

General introduction

Complex networks have come to penetrate many aspects of our lives, such as the internet, World Wide Web (WWW), mobile network and social networks. Complex network are the power behind the most revolutionary technologies of 21st century, from Google to Facebook, and twitter. Therefore in recent years, great amounts of research have investigated the structure of such complex networks [5, 26,154, 25].

Complex networks science has a long history, in the 1930s, José-Luis Moreno was the first to systematically record and analyze social interactions in small groups, particularly classrooms and working groups. What gave birth to sociometry [165]. Then in 1940, Alfred Radcliffe-Brown took a stand, in favor of the systematic study of social networks [166]. At the same time, a group of Harvard researchers, led by Loyd Warner and Elton Mayo, explored interpersonal relationships in the workplace [167]. We can say that these anthropologists are the founders of interest in network analysis.

This new discipline then developed ramifications in other branches such as communications, sociolinguistics or even biology and geography.

From 1954, J. A. Barnes uses the term network more specifically, speaking of small groups, such as family or tribe, or social categories. Then, in the 1960s, many academics worked together to combine different studies. In 1967, the psychologist Stanley Milgram [21, 22] experimented on the existence of a small number of links between two people. He asked a group of US citizens to send a letter to a person, while passing by their knowledge. Milgram found that couriers arrived safely had passed through only a small number of people, stopping at an average of five intermediaries, or six links. Although studies have shown that this number may be higher depending on the case [12].

Recently, a similar study was conducted by Columbia University and showed that five to seven degrees are enough to connect two people by e-mail [168]. In 1973, Mark Granovetter developed a new theory on the importance of weak links [169]. A weak link links an individual to a distant knowledge and is nevertheless a structural basis for exchange. As a result, having many weak links is important when seeking information or disseminating an innovation. If an individual tends to create strong bonds with people, sharing opinions and common traits [170], it also limits the knowledge that he can acquire through them because this knowledge is similar to each of them. To increase their knowledge, members of such a group will have to seek information from individuals with

whom they have created less important links. This is what Granovetter calls the strength of weak ties.

Network analysis has thus shifted from metaphor to a field of research with its own right, with its own theories, methods, researchers and tools, mainly computer and statistical. However, for a very long time, scientists used to ignore that the majority of practical networks have actually complex structure.

They used to believe that these networks have random and relatively simple architecture [12]. This idea was the consequence of the absence of sufficient real data acquisition which makes it impossible to determine with certitude the structure of these networks.

Therefore, since the creation of internet in 1990s, there is no doubt that we are living in a networked world today. On the one hand, networks bring us convenience and benefits, improve our efficiency of work and quality of life, and create tremendous advantages and opportunities which we never had before. Thus, the Internet revolution, offering effective and fast data sharing methods and cheap digital storage, fundamentally changed our ability to collect, assemble, share, and analyze data pertaining to real networks.

Thanks to these technological advances, the analysis of huge amount of available data collected from internet crowned with significant discoveries in graph theory and network science. This resides essentially in the discovery of the small-world phenomenon by Watts and Strogatz in 1998 [5] and the scale-free characteristic by Barabási and Albert in 1999 [26].

The small-world property includes two popular manifestations. The first one is for example, in a friendship network, a friend of my friend is more likely a friend of mine too. The second is “the six degree of separation” which means that two individuals can be linked in relatively short chain [5, 21, 22].

The scale-free (SF) property means that the number of connections that got each node of the network is not homogeneous as random graphs [11, 12]. Considering the same example of friendship network, this is reflected by the fact that we don’t have the same number of friends. More than that, the number that has each one of us doesn’t oscillate around an average. The reality is that few of us have very large number of friends and the majority has few numbers of friends. In network science, those very popular people are called “hubs”. In general, with respect to the network nature, a hub is a node with large number of

connections, greatly above the average. The presence of such nodes as long with other properties is literally a signature of a complex topology or SF networks [26].

The first observations of scale free property behavior in internet were carried out at the end of the 1980s [171-173], on Ethernet data, of high quality, collected in a Bellcore laboratory [174,175].

Furthermore, the Internet is nowadays changing in terms of its uses. From a single-service network to transport binary or text files, twenty years ago, to a multiservice network for the transport of diverse and varied data such as audio and video data.

Therefore, in order to avoid the problem of congestion in traffic on internet and to offer users a service with good quality, reliable and fast, a lot of research has been done for this purpose.

However, network traffic modeling consists in establishing realistic models including the topology and the type of network, the properties of the nodes and links, the routing protocols, the problems that can affect the network (breakdown, virus ...) and the manner in which packets circulate in the network. This leads to understanding and studying the properties of the internet network.

In order to improve Internet traffic capacity, several routing strategies have been proposed [70-88]. These routing strategies govern how information is routed through the network; their role is to pass packets of a node to their destinations, according to a well-defined routing algorithm. In general routing strategies can be devised in two categories; the first one resides in changing the topological features of the network so that it can handles more traffic load, by adding or removing nodes or links. This is called “hard strategy” because it operates on the underlying network structure. In, the second strategy which is referred to as “soft strategy”, the network structure is kept intact and only routing protocols are meant to alleviate the traffic flow and improve the transmission efficiency [176]. It is easily predictable that hard strategies can remarkably enhance the network performance, nevertheless, they are very expensive and not practical speaking of the implementation problem. In the other side, soft strategies are more flexible and not much costly. For these reasons soft strategies have drawn more attention from scientists [1,176, 112]. In general, to design effective routing protocols, researchers rely on using properly information related to either topological properties or instantaneous traffic condition of the network, or both at the same time.

However, unfortunately, many disadvantages could affect the packets of information during their passage through the nodes. Example: computer viruses. A computer virus is a malicious mobile code such as viruses, worms, Trojan horses, logic bomb and so on [177-179]. Each code has its own way, to spread in the internet. Computer viruses have the same specificities, namely: infectivity, invisibility, latency, and unpredictability [180]. They are considered the most distracting weapon in the internet, and their spread has a detrimental effect on the world economics and business [181-184].

Computer virus has a recent history, in 1988; a student from Cornell University compiled a code called the “Endless Life Virus”. The spread of this code has put an end to thousands of computers [185]. There are also the "Code Red" and "Nimda" which caused several billion economic losses, when they are broadcast on the Internet [186,187]. In 2003, a virus called "2003 worm King" spread quickly and attacked computers around the world, this led to the blocking of the Internet and the paralyzing of servers [188].

Moreover, today computer virus spreading prediction is one of the most active applications of network science [189, 190]. In January 2010 network science tools have predicted the conditions necessary for the emergence of viruses spreading through mobile phones [191]. The first major mobile epidemic outbreak that started in the fall of 2010 in China, infecting over 300,000 phones each day, closely followed the predicted scenario.

Contribution of the thesis

Following this way, and inspired by some previous works, in this thesis, we will focus on the study of traffic routing strategies and computer virus propagation in complex network, namely:

Traffic routing strategies: for the purpose of alleviating the congestion in traditional shortest path strategy and to deal more with packets transport on real protocol in internet, we proposed a priority policy based on packet destination [192]. This procedure is applied to a fraction of nodes; packets are classified as High Priority if their destinations are among the fraction, otherwise the packets are treated as Low Priority. We found that the prioritization of nodes with high degree (hubs) is always more efficient than the prioritization of nodes with small degree or the random prioritization of nodes. Moreover, we observed three regimes based on the prioritized fraction in the network: the first one is

characterized by an improvement of the High Priority Traffic (HPT) flow without any degradation of the flow of the Low Priority Traffic (LPT). In the second regime, the HPT gains some performance at the expense of a loss of the performance in the LPT flow. While in the last regime the LPT experiences a low performance without any noticeable improvement in HPT compared to the normal flow. This result is very useful for traffic engineers who try to implement a traffic prioritizing policy and at the same time want to be sure that their systems operate as far as possible in the first regime or in the second one in the worst case because they don't want the non prioritized traffic to be impacted. Furthermore based on our prioritization model the companies in the internet can prioritize the traffic coming in or going to specific servers.

Computer virus propagation: due to the fact of fast internet growth, the need in global routing protocols (EP, SP, and GD) of whole network information becomes a huge monetary and computing cost. Therefore, the local routing protocols (LS, LD) with their need only of network local information overcomes global routing strategies and remain highly promising for large and real networks. Hence, for the purpose of alleviating the computer virus spreading in network under soft strategies, we studied the effectiveness of local routing protocols and their additional algorithms [193]; next-nearest neighbors (NNN) and restrictive queue- length algorithm (RQL) in term of robustness in computer virus spreading. It is found that, the local routing protocols without additional algorithms favor the virus spreading, due to the blind transmission of packets between source and destination, while in local routing protocols under RQL, the virus propagation reduced remarkably. Moreover, in comparison with shortest path (SP) and efficient path (EP) global routing strategies ,local routing protocols under the restrictive queue-length algorithm (RQL), the virus propagation is highly reduced and becomes unexpectedly comparable to the traditional shortest path strategy and overcomes the efficient path strategy which shows a high vulnerability and sensitivity to computer virus propagation.

Regardless of the network architecture, we believe that the results of our works may be applied in several traffic network systems. For instance, biological network, communication networks such as the Internet, intelligent transportation networks, and airlines and so on.

After this introduction where we have exposed the motivation of our research and the contributions, the contents of the manuscript will be treated in five chapters planned as follows:

Chapter 1: Begins with defining some basic concepts from graph theory which also used later in our study of traffic and virus propagation in complex network such as node degree and graph matrix. Then it exposes a study of the evolution of complex networks modeling in order to justify the choice of Barabási–Albert model as platform for the traffic routing strategies.

Chapter 2: Starts with basic definitions with respect to the traffic study. Then, it presents a global review of the most important strategies of improving the performance traffic system. We distinguish the policy behind each approach; meanwhile we explain its strengths and weaknesses.

Chapter 3: devoted to introduces prioritization traffic models. Then, explains our proposed model of priority based on packet‘s destination.

Chapter 4: this chapter deals with the introduction and explanation of virus propagation models (SI and SIR) which are the most used models in the field of computer science.

Chapter 5: In this chapter, the computer virus propagation in internet under different proposed local routing strategies (LS, LD) is studied; using Model SI we found that the local routing protocols under theirs additional algorithms are very robust in term of virus propagation.

At the end we address a conclusion and announce some perspectives and future works

Chapter I. Introduction to network sciences and network modeling

I.1 Introduction:

The study of networks has had a long history in mathematical sciences. Graph theory [29-32] is the most rigorous and efficient mathematical tool behind complex networks sciences. The birth of graph theory go back to 1735 in Königsberg (a town which now in territory of Russia), there is a river named Pregel passing through the town Königsburg, and there are seven bridges over the river as shown in Fig I.1.

Therefore, in the old days, the residents always wondered whether someone could walk through single path that across all seven bridges and then return to the starting point without going over any bridge more than once. Despite the fact that the people of Königsberg tried many times to find a such single path, the popular puzzle remained unsolved until the great mathematician Leonhard Euler (1707–83) studied the famous seven-bridge problem and offered a rigorous proof that such single path does not exist [1,2],thereby proving that the Königsburg seven-bridge problem has no solutions.

In order to resolve seven-bridge problem, in 1736, the mathematician swiss Euler comes with a great idea to describe the seven-bridge problem by an abstract graph, using four points A, B, C, D to represent each of the four land areas connected by lines Fig I.1. Thus he succeed for the first time to convert and model a real problem by a graph, for this reason Euler was considered as the father of graph theory.

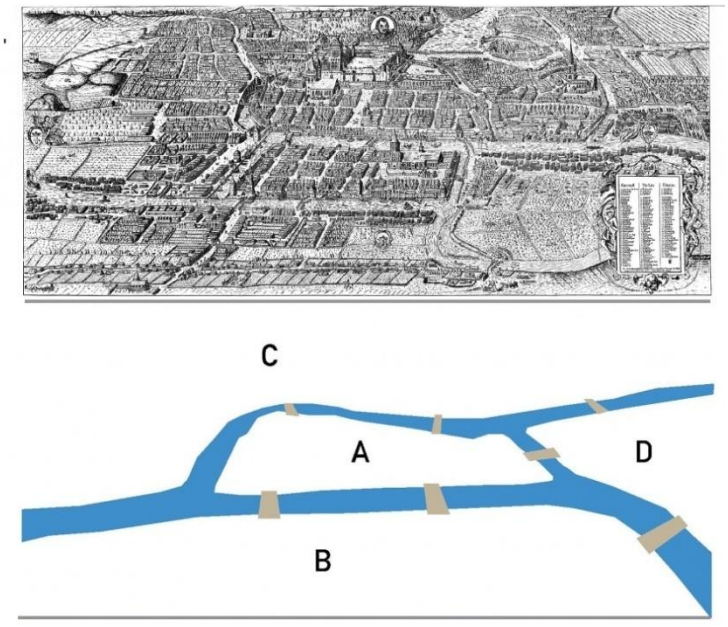


Figure I-1 : A contemporary map of Königsberg (now Kaliningrad, Russia) during Euler's time.

In addition to the developments in mathematical graph theory, the study of networks has seen important achievements in some specialized contexts, as for instance in the social sciences.

With the growing public fascination with complex connectedness that comes to penetrate many aspects of our lives. At the heart of this fascination comes the idea of a network. How should we define networks?

I.2 Network and graph

A graph $G = (V, E)$ is defined as the finite set $V = \{v_1, v_2, v_3 \dots v_n\}$ whose elements are called vertices, and the finite set $E = \{e_1, e_2, e_3 \dots e_n\}$ whose elements are called edges. An edge e of the finite set E represents a link between a pair of vertices, if e connects two vertices a and b, we say that a and b are adjacent or neighboring and also called the ends of e.

A network (describing a real complex system) differs from a graph (that is the abstract mathematical object formed by vertices and edges) because it is a specific graph describing the interactions present in a specific real world system. As example, the WWW is a network of web documents linked by URLs. In contrast, the web graph is the mathematical representation of the WWW network. Therefore in network, we use the terminology {node, link}.While in graph we talk about {vertex, edge}.

Indeed many different real systems may have exactly the same network representation. Yet, this distinction is rarely made, so these two terminologies are often synonyms of each other.

As shown in Figure I.2, three rather different networks have exactly the same graph representation. While the nature of the nodes and the links differs, these networks have the same graph representation, consisting of $N = 4$ nodes and $L = 4$ links, shown in Figure I.2 (d).

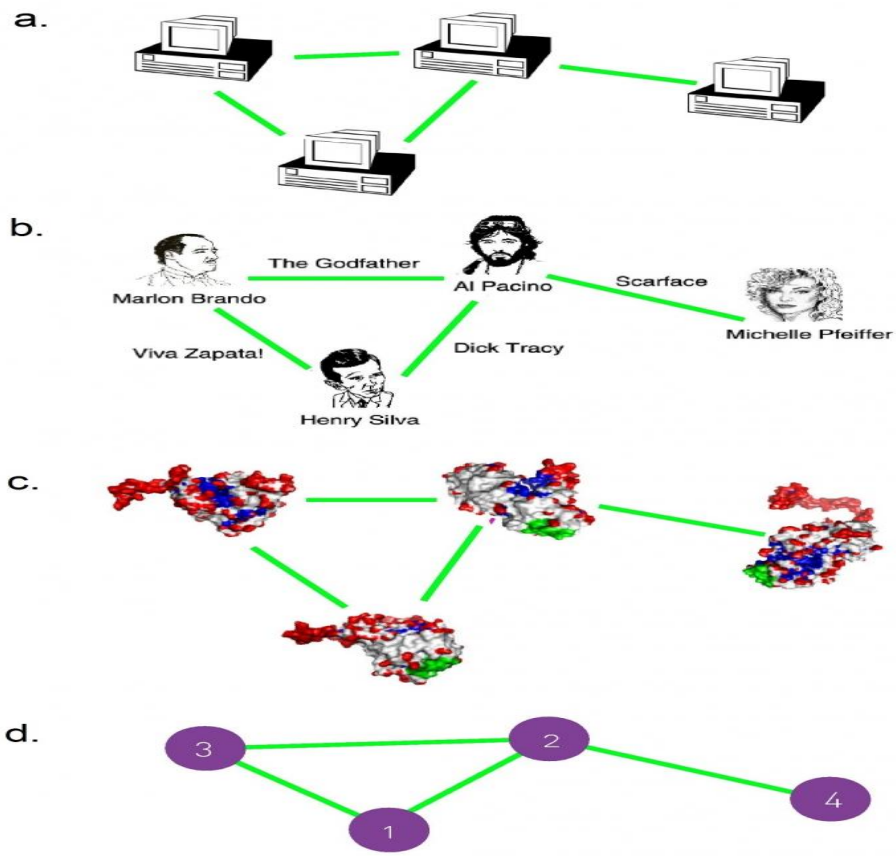


Figure I-2 : The figure shows a small subset of (a) the Internet, where routers (specialized computers) are connected to each other; (b) the Hollywood actor network, where two actors are connected if they played in the same movie; (c) a protein-protein interaction network.

I.2.1 Network types

I.2.1.1 Undirected and directed network

The network can be undirected or directed. In undirected network, all the links are undirected. If the nodes a and b are connected, then there is also a link between b and a, like transmission lines on the power grid, on which the electric current can flow in both directions. While In directed network, all the links are directed. If the link between a and b is: l_{ab}, then l_{ab} is different than l_{ba}, like the WWW, whose uniform resource locators (URL) point from one web document to the other. Some networks simultaneously have directed and undirected links. For example in the metabolic network some reactions are reversible (i.e., bidirectional or undirected) and others are irreversible, taking place in only one direction (directed).

I.2.1.2 Weighted and unweighted networks

In many applications, it is necessary to assign a numerical value w_{ab} to each link l_{ab} of a network, which called the weight. The weight of links is usually positive integer. Depending on the context, in mobile call networks the weight can represent the total number of minutes two individuals talk with each other on the phone.

I.2.2 Networks characteristics

In order to characterize the various topological features of complex networks, we have a set of measures, the most commonly used of which are presented in this section.

I.2.2.1 Degree

The concept of degree (or connectivity) is the most fundamental character and measure of a node in a network, which may be defined in different ways. We denote with k_i the degree of the node i in the network. For example, for the undirected networks shown in Figure I.2 we have $k_1 = 2$, $k_2 = 3$, $k_3 = 2$, $k_4 = 1$. In an undirected network the total number of links, L, can be expressed as the sum of the node degrees:

$$L = \frac{1}{2}\sum_{i=1}^{N} k_i$$

Here the $\frac{1}{2}$ factor corrects for the fact that in the sum each link is counted twice. For example, the link connecting the nodes 2 and 4 in Figure I.2 will be counted once in the degree of node 2 and once in the degree of node 4.

While, in directed networks, the degree of the node i has two components: incoming degree, k_i^{in}, representing the number of links that point to node i, and outgoing degree, k_i^{out}, representing the number of links that point from node i to other nodes. Finally, the total degree k_i is then defined as

$$k_i = k_i^{in} + k_i^{out}$$

For example, on the WWW the number of pages a given document points to represents its outgoing degree,k^{out}, and the number of documents that point to it represents its incoming degree, k^{in},.

The total number of links in a directed network is:

$$L = \sum_{i=1}^{N} k_i^{in} = \sum_{i=1}^{N} k_i^{out}$$

The $\frac{1}{2}$ factor seen in undirected networks is now absent, as for directed networks the two sums separately count the outgoing and the incoming degrees.

I.2.2.2 Average degree

An important property of a network is its average degree, which for an undirected network is:

$$< k > = \frac{1}{N} \sum_{i=1}^{N} k_i = \frac{2L}{N}$$

While for a directed network, the average degree is:

$$< k^{in} >= \frac{1}{N} \sum_{i=1}^{N} k_i^{in} =< k^{out} >= \frac{1}{N} \sum_{i=1}^{N} k_i^{out} = \frac{L}{N}$$

Where N is the network and L is the network total number of links.

I.2.2.3 Degree distribution

The most network central and basic topological characterization can be obtained in terms of the degree distribution p_k. One reason is that the calculation of most network properties requires to know p_k. For example, the average degree of a network can be written as

$$< k >= \sum_{k=1}^{\infty} k p_k$$

Where p_k defined as the probability that a node chosen uniformly at random has degree k or, equivalently, as the fraction of nodes in the graph having degree k. Since p_k is a probability, it must be normalized, i.e.

$$\sum_{k=1}^{\infty} p_k = 1$$

Numerically, we can plot the degree distribution probability p_k by establishing the fraction of nodes with the degree k; we calculate the total number of nodes N_k that have the same degree k divided by the total number of the nodes N . For network with N nodes the degree distribution is the normalized histogram in Fig I.3 is given by:

$$p_k = \frac{N_k}{N}$$

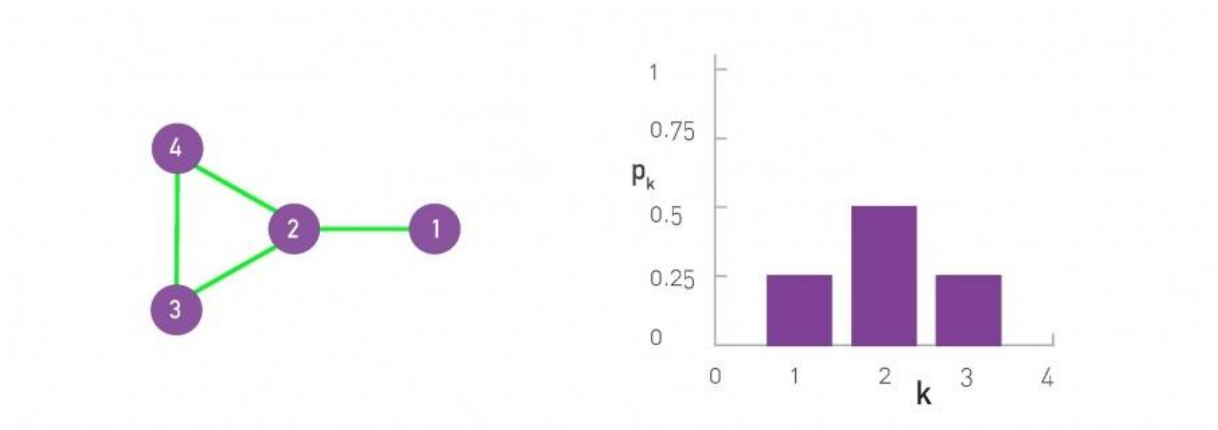


Figure I-3 : Example of the degree distribution of a network with N=4 nodes.

I.2.2.4 Adjacency matrix

A network can be completely described by giving the binary matrix adjacency. In a network the adjacency matrix A is a matrix representation exactly equivalent to the graph. For undirected network the matrix has two entries for each link $A_{ij} = 1$ and $A_{ji} = 1$, if there is a link between the nodes i and j, $A_{ij} = 0$ otherwise. Note that it is necessarily symmetrical according to our definition of undirected graph, and its diagonal elements are null. As shown in undirected network in (Fig I.4 a), link (1, 2) is represented as $A_{12} = 1$ and $A_{21} = 1$.

In the context of directed networks, we will also define an adjacency matrix. The lines will correspond to the source nodes and the columns to the destinations, its elements

being: $A_{ij} = 1$ If there is a link pointing from node i to node j , $A_{ij} = 0$ otherwise. Unlike the previous case of undirected graphs, this matrix is not symmetrical (Fig I.4 b).

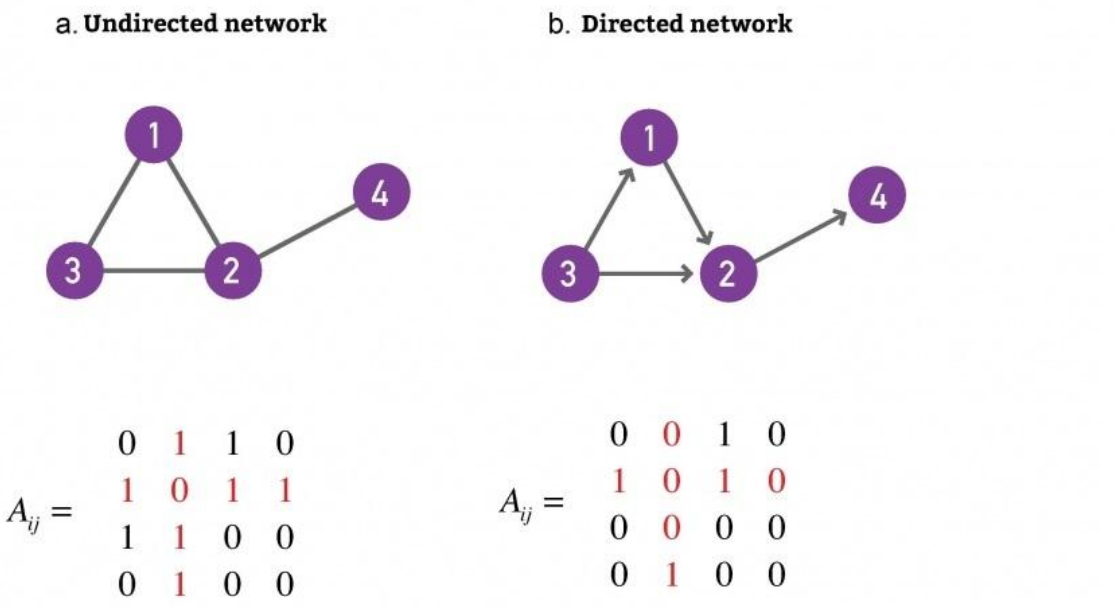


Figure I-4 : The adjacency matrices of two different types of networks with N=4; (a) undirected graph, (b) directed network.

I.2.2.5 Distances

The distance or path length plays a central role in determining the interactions between nodes in a network. Path length is a route that runs along the links of the network and represents the number of links n that the path contains between source and destination. For example, in (Fig I.5 a), the path shown in orange between nodes 1 and 6 follows the route 1→2→5→7→4→6, hence its length is $n = 5$.

I.2.2.6 Shortest path

Shortest paths play an important role in the transport and communication within a network. Suppose one needs to send a data packet from one computer to another through the Internet: the geodesic provides an optimal path way, since one would achieve a fast transfer and save system resources [28]. The shortest path between nodes i and j is the

path with the fewest number of links. The shortest path never contains loops or intersects itself.

Like in (Fig I.5 b), the shortest paths between nodes 1 and 7, correspond to the path with the fewest number of links that connect nodes 1 to 7. We can have multiple shortest paths of the same length between a pair of nodes as illustrated by the two paths shown in orange and grey. The shortest path between two vertices i and j can be calculated by the Dijkstra algorithm [116, 117].

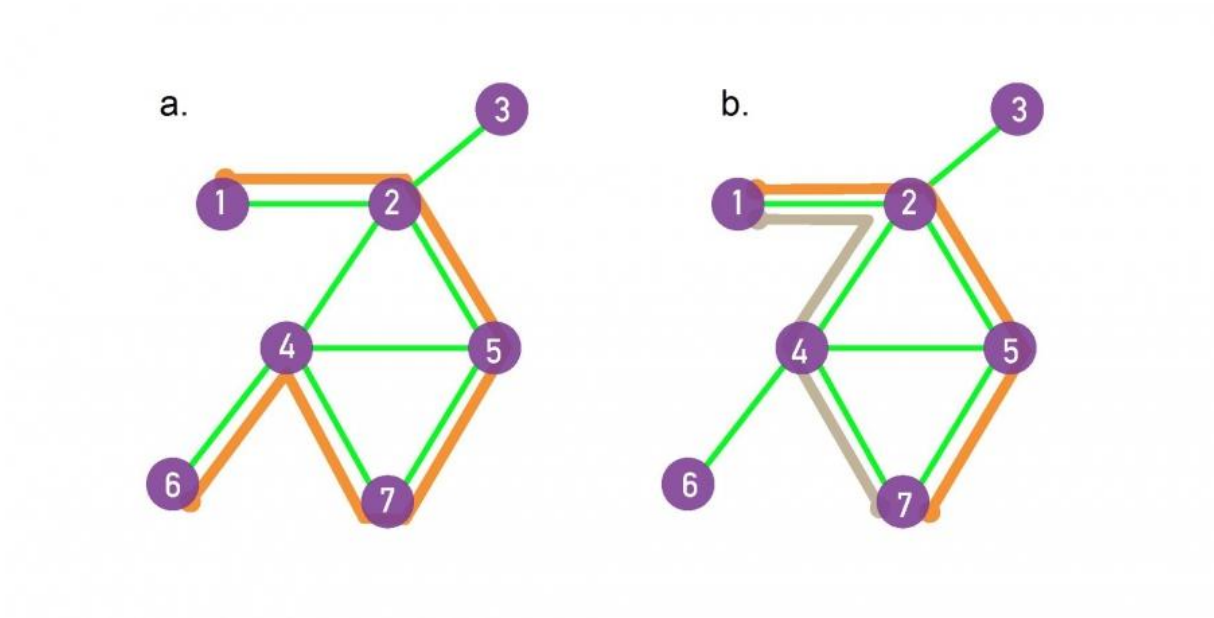


Figure I-5 (a) Distance or path shown in orange between nodes 1 and 6 (b) The shortest paths between nodes 1 and 7 , There can be multiple paths of the same length, as illustrated by the two paths shown in orange and grey.

I.2.2.7 Average path length

A measure of the typical separation between two nodes in a network is given by the average length of the path, $L(G)$ [3 ,5,35]:

$$L(G) = \frac{1}{N(N-1)} \sum_{i \neq j} d(i,j)$$

Where $d(i,j)$ is value of the shortest path between nodes i and j and N being the number of vertices of the graph.

I.2.2.8 Betweeness centrality:

The betweenness centrality (BC) attaches the importance of a vertex v to the number of the shortest paths crossing it [55–57]. In other term it is the number of times that the node v in question acts like a bridge between every pair of the network nodes. We define more precisely the BC of a vertex v as:

$$BC(v) = \sum_{i \neq v \neq j \neq \in N} \frac{\sigma_{ij}(v)}{\sigma_{ij}}$$

Where $\sigma_{ij}(v)$=$\sigma_{ji}(v)$ is the number of the shortest paths that the vertex v lies on and σ_{ij} denotes the total number of the shortest paths or also the total number of non distinct pairs of nodes that we can have excluding the node v.

I.2.2.9 Clustering coefficient

The calculation of the clustering coefficient comes from a simple observation. If a person A knows a person B then there is a high probability that a person C, knowing A, also knows B. In terms of network, we will say that if a node i is connected to a node j and that latter is connected to a node, then there is a high probability that the node i is related to k. This characteristic can be measured for each node of the graph through the clustering coefficient defined by Watts [4, 5]. The latter is calculated for a node i, as a function of the number of links between the various neighbors of i and the maximum number of possible links between all the neighbors of i. The local coefficient of clustering has the value:

$$C_i = \frac{2e_i}{k_i(k_i - 1)}$$

Where e_i being the number of links between the neighbors of i and k_i corresponding to the degree of i.

Note that C_i is between 0 and 1. $C_i = 0$ if none of the neighbors of node i link to each other. $C_i = 1$ if the neighbors of node i form a complete graph, i.e. they all link to each other. C_i is the probability that two neighbors of a node link to each other. In summary C_i

measures the network's local link density: The more densely interconnected the neighborhood of node i, the higher is its local clustering coefficient.

The clustering coefficient of a whole network [5] is captured by the average clustering coefficient,$\langle C \rangle$, representing the average of C_i over all nodes $i = 1, .. N$:

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^{N} C_i$$

In line with the probabilistic interpretation,$\langle C \rangle$ is the probability that two neighbors of a randomly selected node link to each other. As the undirected network, the clustering coefficient is generalized to directed and weighted networks [6-9].

In Fig I.6, the value of local clustering coefficient C_i of each node is calculated and shown next to it. We also list the network's average global clustering coefficient $\langle C \rangle$.

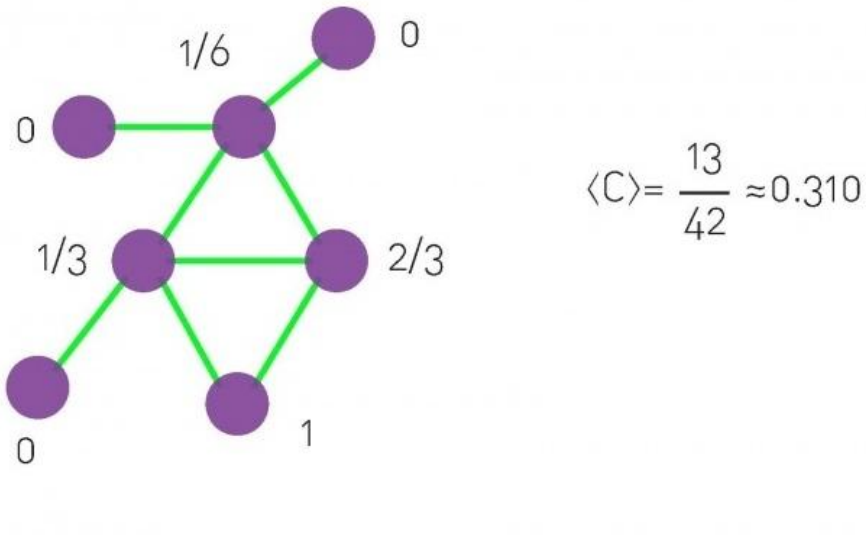


Figure I-6 : Example of calculation of the clustering coefficient in a network.

I.3 Random network models

The First classical model of a network that we would call random network was described by the Russian mathematician and biologist Anatol Rapoport and Ray Solomonoff in 1951 [10]. Rapoport introduces the random graph for the first time and define a quantity called

the weak connectivity, which is the expected number of vertices reachable through the network from a randomly chosen vertex. In the modern terminology of networks, the weak connectivity is the average component size in the network. Rapoport, demonstrated that if we increase the average degree of a network, we observe an abrupt transition from disconnected nodes to a graph with a giant component.

Despite the early contribution of Rapoport, random graph theory reached prominence thanks to sequence of several papers[11-18] published by Paul Erdös and Alfréd Rényi between 1959 -1968; they merged probability theory and combinatory with graph theory. In [12], Erdös showed that many properties of random graphs emerge not gradually but suddenly, when enough edges are added to the graph. However, based on small world theory [21], watts and Strogatz in 1998 proposed an extension of the random network model [5, 23]. This model was the first successful model manifesting both the strong clustering and the small average path characteristics encountered in many real world networks.

I.3.1 Erdös-Rényi model

Throughout this section, we will explore the Erdös-Rényi random model [12], for the ease that it allows to calculate key network characteristics. This model consists of N nodes where each node pair is connected with probability p.

I.3.1.1 Construction:

To generate Erdös-Rényi network, we follow these steps:

- Start with N isolated nodes.
- Select a node pair and generate a random number between 0 and 1. If the number exceeds p, connect the selected node pair with a link, otherwise leave them disconnected.
- Repeat the precedent step for each of the $N\ (N-1)/2$ node pairs.

The advantage of this model is that it is relatively simple to understand and to study. Note that , due to the randomness of the connectivity probability, for each random network generated with the same parameters N, p looks slightly different (see Fig I.7), not only the detailed wiring diagram changes between realizations, but so does the number of links L.

Nevertheless, all resulting networks are similar in some properties as the degree distribution.

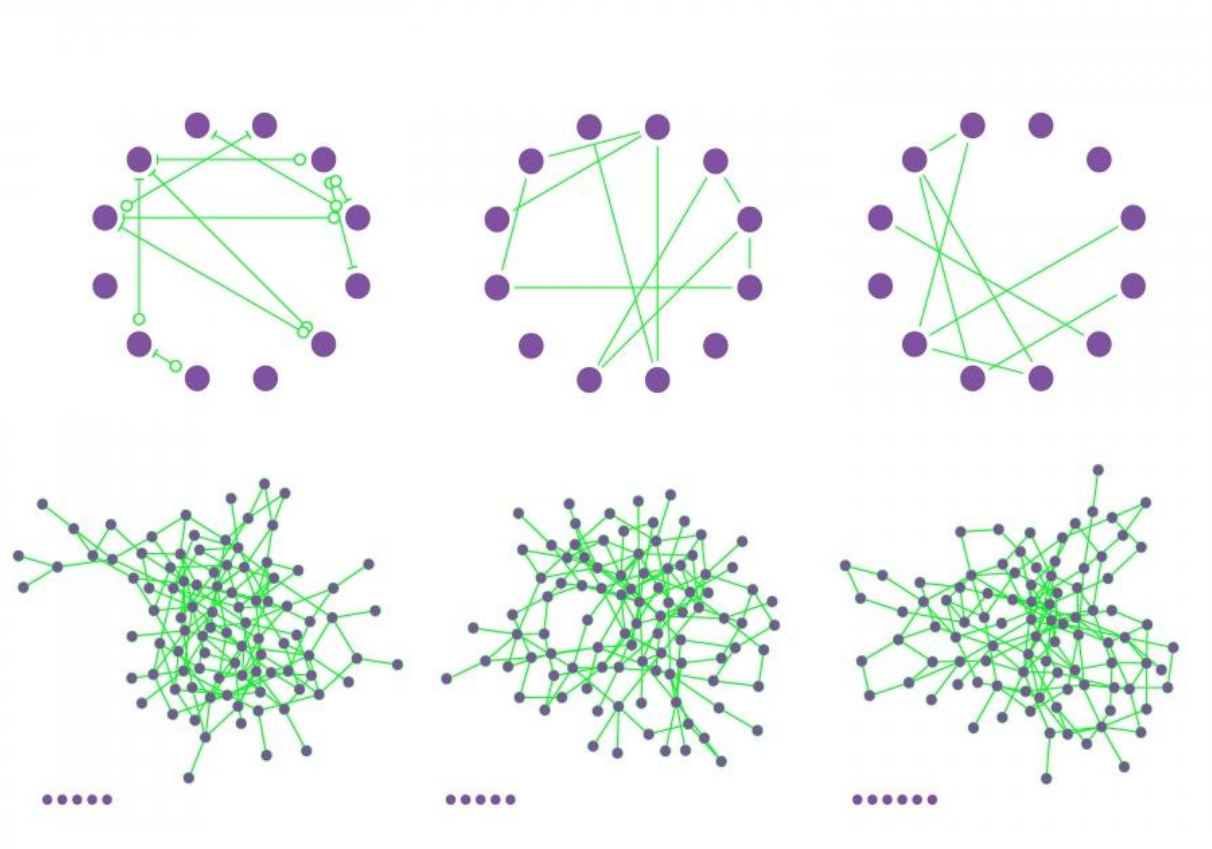


Figure I-7 : Top Row: Three realizations of a random network generated with the same parameters p=1/6 and N=12. Despite the identical parameters, the networks not only look different, but they have a different number of links as well (L=10, 10, 8). Bottom Row:Three realizations of a random network with p=0.03 and N=100.

I.3.1.2 Degree distribution:

In a given realization of Erdös-Rényi network some nodes gain numerous links, while others acquire only a few or no links (see Fig I.7). These differences are captured by the degree distribution, p_k, which is the probability that a randomly chosen node has degree k.

It has been proved that the degree distribution of a random network follows the binomial distribution [36]

$$p_k = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$

I.3.2 Small world network

I.3.2.1 Small world phenomena

The Small world (SW) phenomenon or six degree of separation (Fig I.8) is another way to say that most real networks have a relatively short average path length in spite of their huge sizes. The first empirical study of the small world phenomena took place in 1967, when Stanley Milgram, based on the work of Pool and Kochen [20], performed an experiment to find out the distances in social networks. The experiment consisted in asking some volunteers from Nebraska and Kansas to send some information letters to a recipient in Boston, Massachusetts using as intermediate people that they know [21, 22].

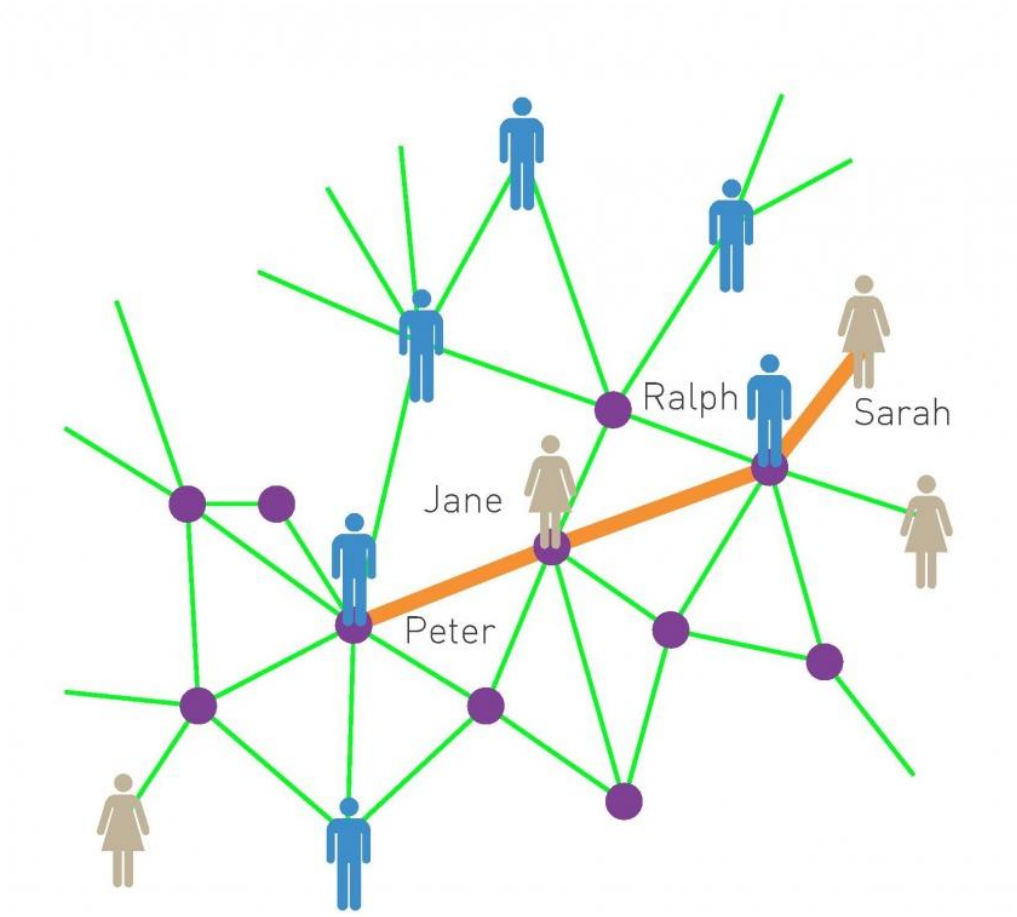


Figure I-8: Six Deegree of Separation: According to six degrees of separation two individuals, anywhere in the world, can be connected through a chain of six or fewer acquaintances

Milgram was surprised to find out that the average of chains formed by peoples from the sender to the recipient was only 5.2 which is a short distance compared to the total size of the network [22, 37]. These completed chains allowed Milgram to conclude that two randomly chosen individuals in the North American population can be connected by a very short chain of relations as shown in Fig I.9 a.

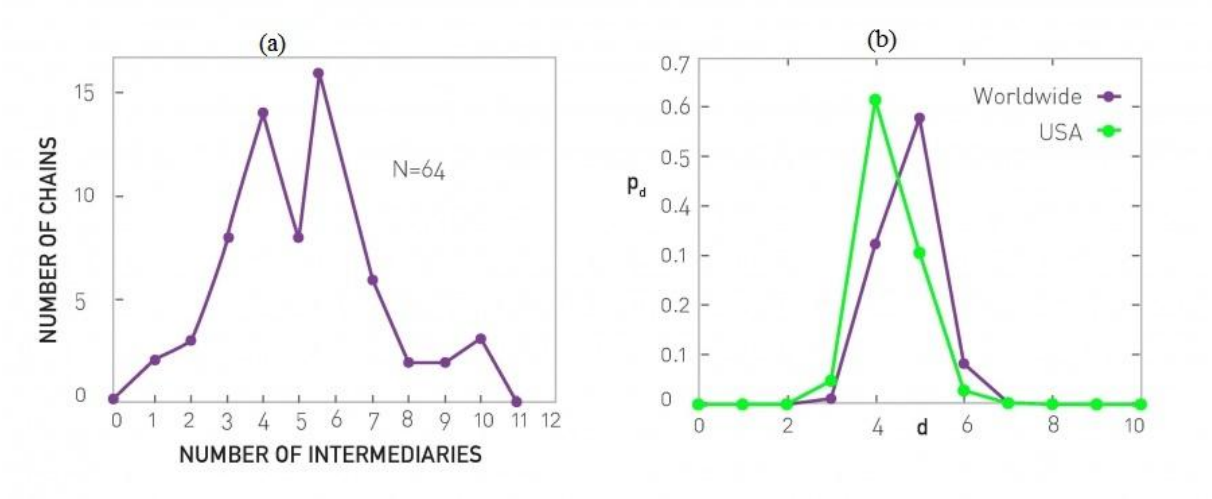


Figure I-9 : The figure shows, (a)the length distribution of the completed chains In Milgram's experiment, the mean of the distribution was 5.2 [25]. (b) The distance distribution, pd, for all pairs of Facebook users worldwide and within the US only [18].

While in the reality, Milgram lacked real algorithms like the ones we use and database of friendship links as in facebook network or in social media networks that we have in our days. Thus, Milgram was measuring the average length of a routing path on a social network, which is truly only an upper bound on the average distance (as the people involved in the experiment were not necessarily sending the postcard to an acquaintance on a shortest path to the destination) because the shortest path routing was an unintended finding of these experiments, and largely went unremarked until much later [116,117]. Hence Milgram experiment could not detect the true distance between his study's participants. Based on this fact, two recent studies [25, 38] explored the real distance on Microsoft messenger and Facebook social network.

From Milgram to Facebook:

Today Facebook has the most extensive social network map ever assembled. Lars Backstrom [25], studied the distance distribution of Facebook network. Using Facebook's

social graph of May 2011, consisting of 721 million active users and 68 billion symmetric friendship links, researchers found an average distance 4.74 between the users (see Fig I.9 b). Therefore, the study detected only 'four degrees of separation', closer to the prediction of Milgram's six degrees [21, 22].

From Milgram to snapshot of Microsoft Messenger :

In 2008, Jure Leskovec [38], studied the distance distribution of Microsoft Messenger network; the authors considered a communication graph with 180 million nodes and 1.3 billion undirected edges extracted from a snapshot of the Microsoft Messenger network, they find that the average path length among Messenger users is 6.6 closer to the prediction of Milgram's six degrees [21, 22] (see Fig I.10).

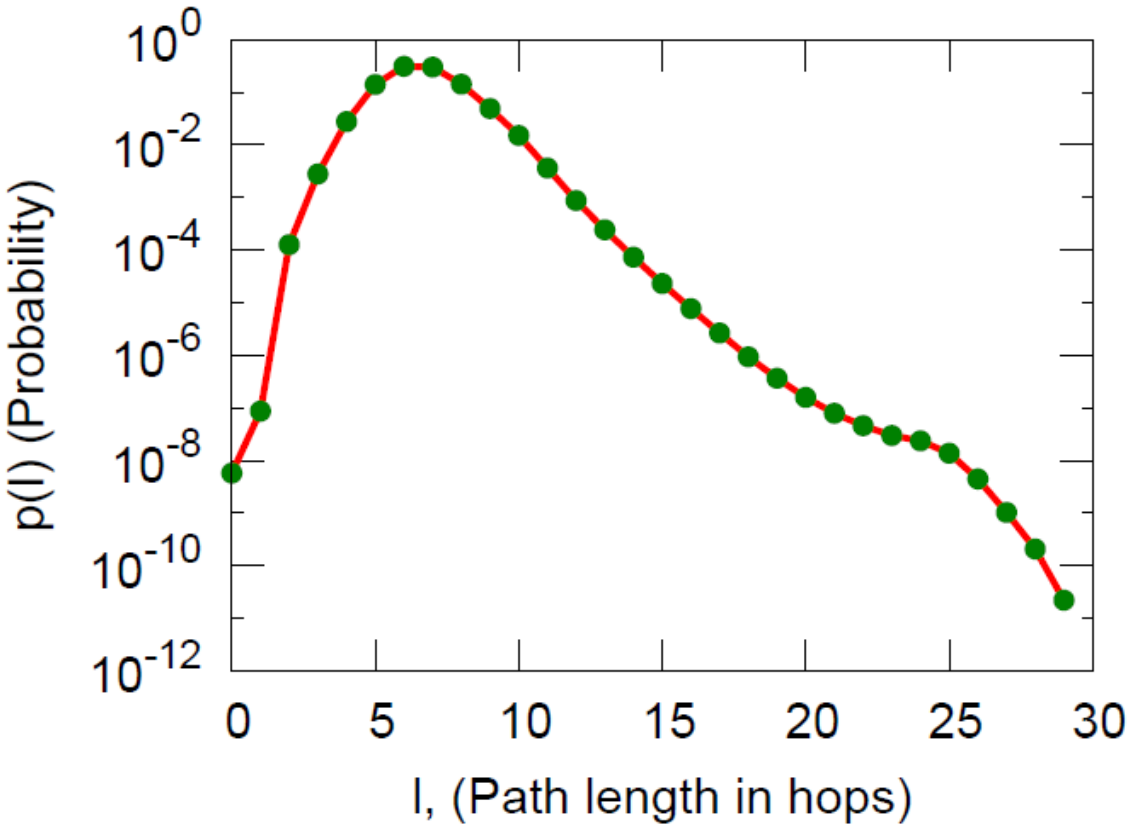


Figure I-10: The distance distribution for all pairs of Microsoft Messenger network users [38].

The main difference between Facebook and Microsoft Messenger experiments and Milgram's is that the notion of friendship in Facebook and Messenger is hardly comparable to the idea of friendship in real life.

I.3.2.2 Watts-Strogatz model

After the small world phenomena discovery in real world, Watts and Strogatz [5] proposed a new extension of the random network model to generate a small world network characterized by a small average distance and a high coefficient of transitivity.

Construction:

To generate The Watts-Strogatz Model as shown in Fig I.11, we follow these steps:

- We start from a ring of nodes, each node being connected to their immediate and next neighbors.
- With probability p each link is rewired to a randomly chosen node.
- For $p = 1$ all links have been rewired, so the network turns into a random network.

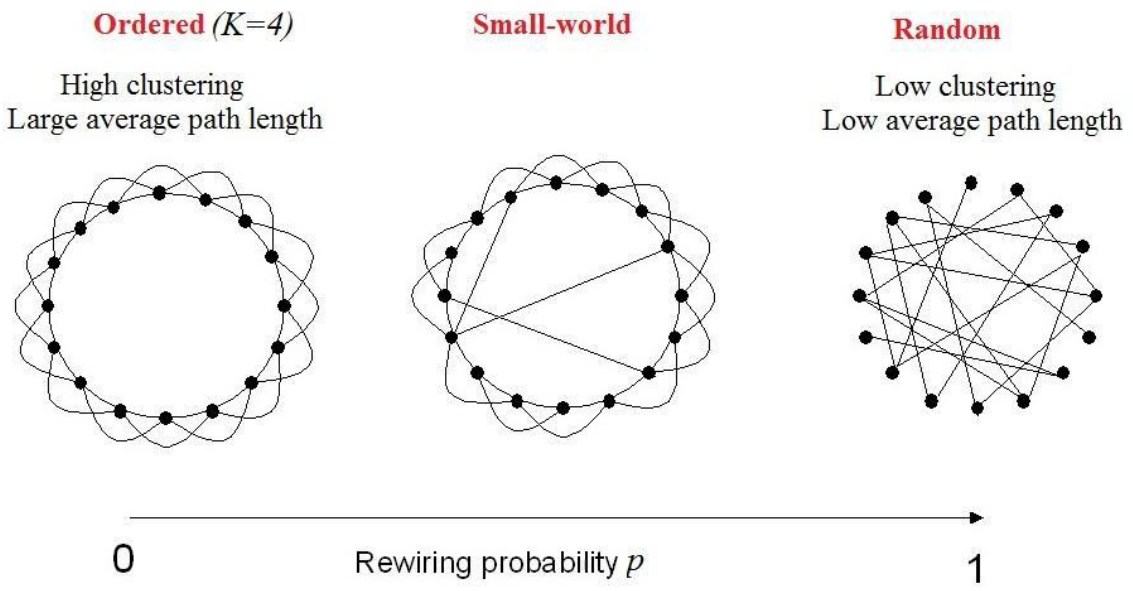


Figure I-11: The random rewiring procedure of Watts and Strogatz model; we start with N=16 nodes, each node is linked to its first K=4 neighbors.

For small probability p values, many nodes retain their connections with their initial neighbors, that is, those of the ordered network. While the closer p to 1, the more the network tends to a random structure, since all the links are rewritten randomly. Thus the

small world networks are random networks whose average shortest path is weak but whose clustering coefficient is high. They get closer to random networks and ordered networks and take advantage of the properties of both.

Degree distribution:

The small world networks are extension of random networks whose average shortest path is weak but whose clustering coefficient is high. Thus, the Watts- Strogatz model predicts a Poisson- degree distribution [5]. Consequently high degree nodes, like those seen in Fig I.12, are absent from it.

I.3.3 Real Networks are Not Random

Le assumes that real networks are well described by the random network models. Therefore, in a random interconnection in Internet, all routers are expected to have comparable number of connection. Of the same in random WWW network, all the hosts or website will have the same degrees of popularity: there is no highly popular website as Google or Facebook, and no website is left behind, having only a low popularity (page ranking). Furthermore, a study [25] on social network on Facebook social network has found numerous famous individuals with 5,000 Friends, while others have much less facebook friends. To see clearly how random networks are different from real networks, we must compare the degree distribution of real and random networks.

Fig I.12 shows the degree distribution of the (a) Internet, (b) science collaboration network, and (c) protein interaction network. The green line corresponds to the Poisson prediction. The significant deviation between the data and the Poisson fit indicates that the random network model underestimates the frequency of the high degree nodes, as well as the number of low degree nodes. Instead the random network model predicts a larger number of nodes in the vicinity of $\langle k \rangle$ than seen in real networks. Hence random network are enable to model real networks which have a degree distribution far from Poissonian with heavy tails which rather follows a log-normal distribution or alternatively a power law.

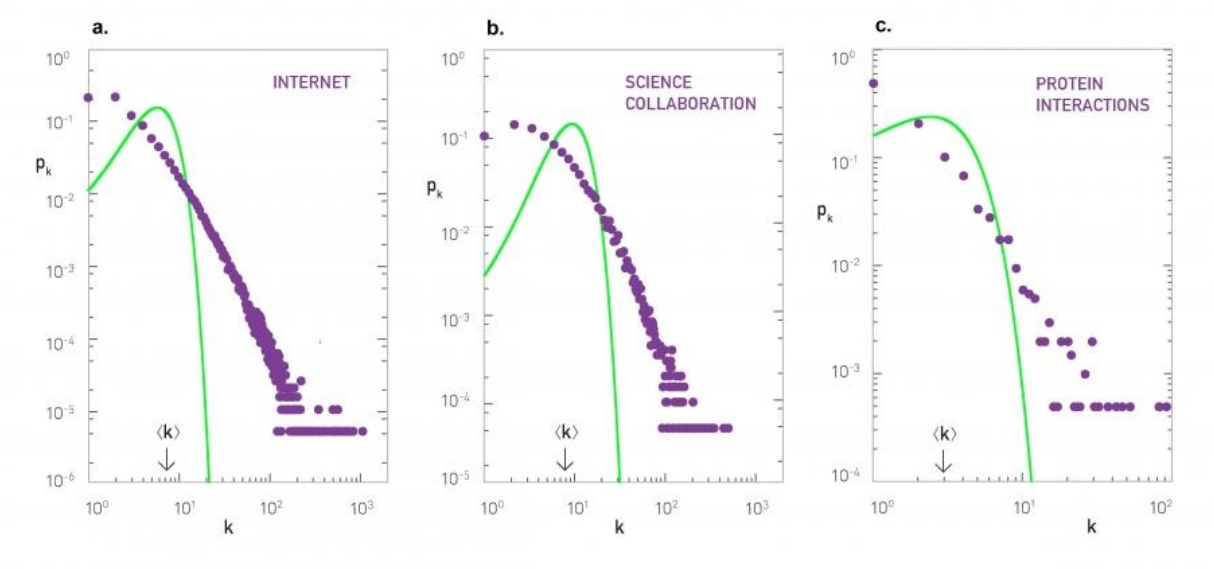


Figure I-12: Degree Distribution of Real Networks: The degree distribution of the (a) Internet, (b) science collaboration network, and (c) protein interaction network.

In general, the random network model failed to model the real networks due to the absence of two important characteristics observed in real networks, the Growth and the preferential attachment:

Growth: Real networks are the result of a growth process that continuously increases size N. In contrast the random network model assumes that the number of nodes, N, is fixed. While in real networks the number of nodes continually grows thanks to the addition of new nodes. As example

- In 1991 the WWW had a single node, the first webpage build by Tim Berners-Lee, the creator of the Web. Today the Web has over a trillion documents, an extraordinary number that was reached through the continuous addition of new documents by millions of individuals and institutions (see Fig I.13).

Preferential Attachment: In real networks new nodes tend to link to the more connected nodes (hubs). In contrast nodes in random networks randomly choose their interaction partners. For example in Facebook social network , a famous individual having many friends have always more probability to get a new friendship invitation from other individuals.

In a nutshell, all of the models discussed so far attempt to create networks that incorporate the properties of real networks. The models do not, however any serious approach of modeling real networks must take in account the two characteristics explained above.

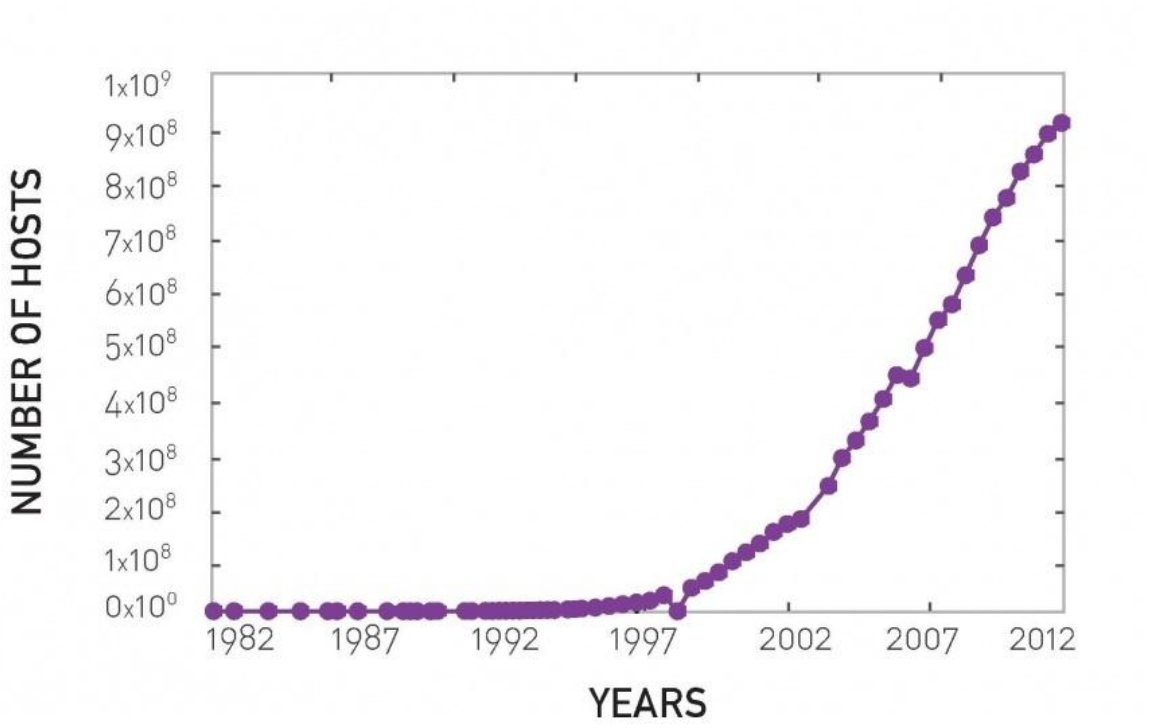


Figure I-13: Networks are not static, but grow via the addition of new nodes: The evolution of the number of WWW hosts, documenting the Web's rapid growth. After http://www.isc.org/solutions/ survey/history.

I.4 Scale free networks

I.4.1 Scale free property in real networks

The name "scale free" comes from the fact that there is no characteristic value of k. While in random networks, the characteristic k is the average degree $\langle k \rangle$, i.e., there is no typical degree in scale-free networks. In the scale free model, the networks typically grow by the gradual addition of nodes and links in some manner intended to reflect growth processes that might be taking place on the real networks.

The scale-free (SF) property means that the number of connections that got each node of the network is not homogeneous as in random networks [11, 12].

Probably, the earliest published example of a scale-free network is probably Price's network of citations between scientific papers [40]. Price's tried to explain real network citation between papers properties and discus for the first time the idea of cumulative advantage [39].

However, the study of scale free networks reached prominence in 1999 thanks to the fundamental work of Barabási-Albert [26]. Barabási proposed a scale free model (BA), which overcomes the drawbacks of all random network models discussed before in this chapter. In this model, the networks typically grow by the gradual addition of nodes and links in some manner intended to reflect growth processes that might be taking place on the real networks. Since there, the scale free networks have been observed in a host of other networks, including notably other citation networks [41, 42], the World Wide Web [43-45], the Internet [46, 27, 48], metabolic networks [49, 50], telephone call graphs [51, 52], and the network of human sexual contacts [53, 54].

I.4.2 Barabasi-Albert model

The scale free property observed in real networks has inspired Barabási and his colleagues to propose the famous BA model [26].

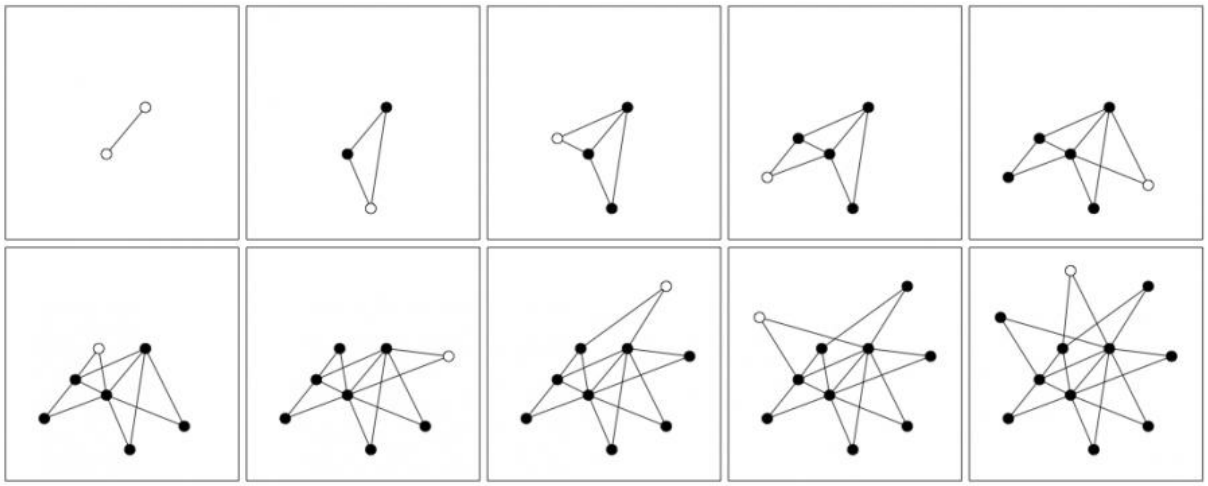


Figure I-14: The sequence of images shows nine subsequent steps of the Barabási-Albert model. Empty circles mark the newly added node to the network, which decides where to connect its two links (m=2) using preferential attachment [26].

I.4.2.1 Construction:

In BA model, the network in Fig I.14 is constructed as follow:

- **growth**: starting from a small amount of m_0 fully connected nodes, the number of nodes increases throughout the lifetime of the network by the subsequent addition of new nodes
- **preferential attachment** :Each new node is connected preferentially to m ($m \leq m_0$) old ones in such a way that the probability of connecting to an existing node i is proportional to its degree ():

$$P_i = {k_i}/{\sum_j k_j} \quad (1)$$

I.4.2.2 Degree distribution:

As we have already seen, the scale free networks as the real networks are characterized by power law distribution. First, the degree distribution curve must be a power law and secondly, the value of λ recovered from the curve results must be close to -3, which is the value specific to the BA network (see Fig I.15).

$$y = a.k^{\lambda} \text{ With } \lambda \approx -3$$

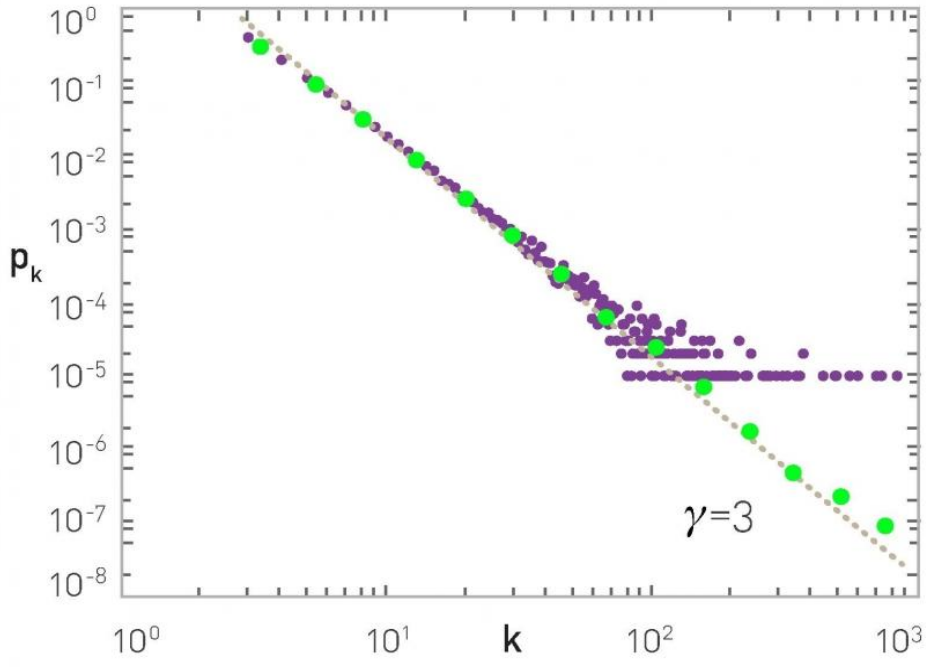


Figure I-15: The degree distribution of a network generated by the Barabási-Albert model [26].

This power law degree distribution represents exactly the real network behavior that can be explicated by preferential attachment mechanism .like, the signatures of the Internet's scale-free nature in Fig I.16 [26-28], showing that a few high-degree routers hold together a large number of routers with only a few links.

From all these results on complex network models studies, we deduced that the most appropriate among all the previous models and the closest to the reality is the Barabási-Albert (BA) model [26]. For this reason, next this model will be used to emulate the infrastructure where the traffic and dynamic spreads will take place in the next chapters.

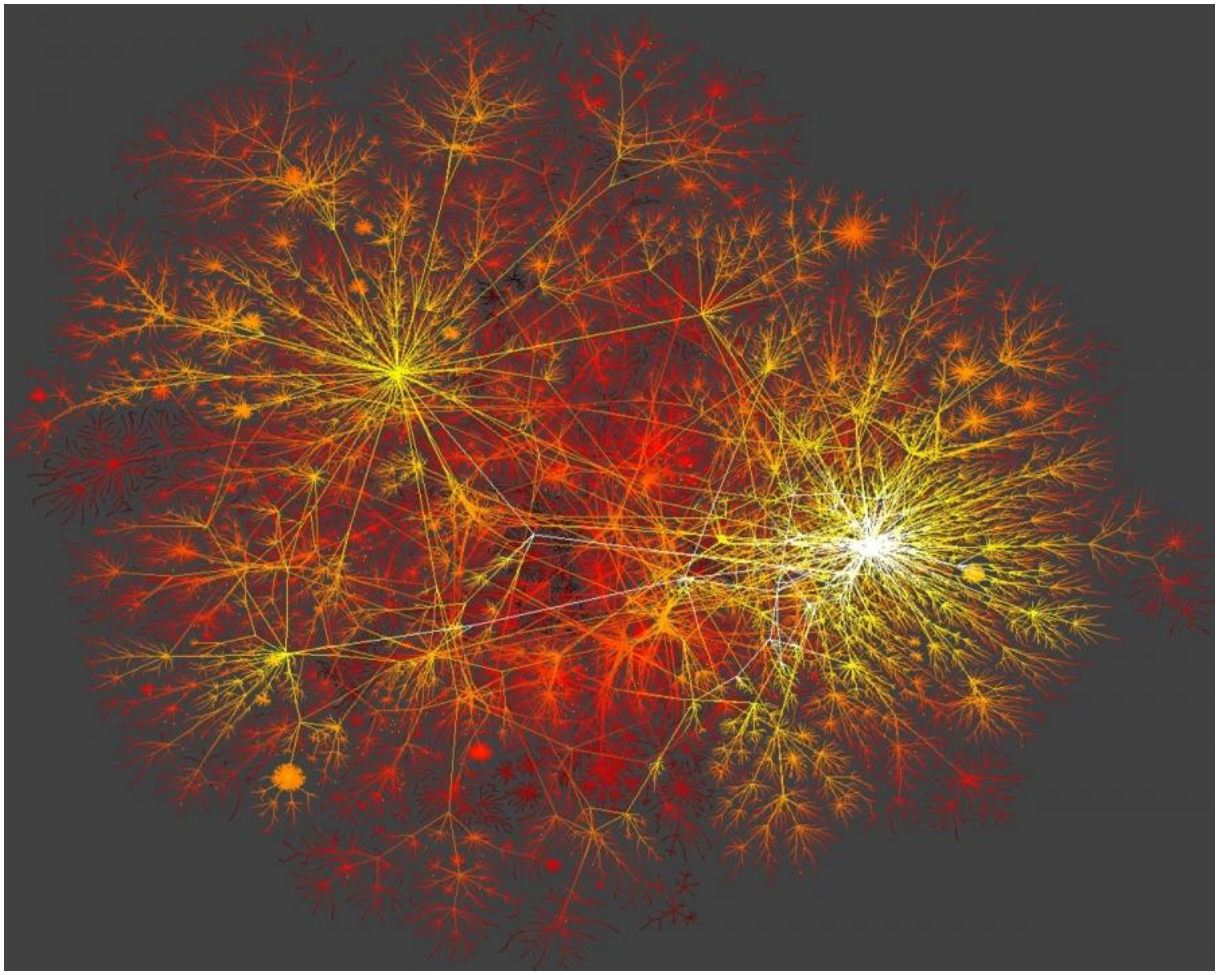


Figure I-16 : The topology of the Internet An iconic representation of the Internet topology at the beginning of the 21st century. The image was produced by CAIDA, an organization based at University of California in San Diego, devoted to collect, analyze, and visualize Internet data. The map illustrates the Internet's scale-free nature: A few highly connected hubs hold together numerous small nodes.

I.5 Conclusion:

This first chapter is a general representation where we presented the general context of our research. We have started by preliminary definitions in graph theory. Then we have focused on the structure of complex networks, mainly the definitions and the characteristics that influence the network behavior. This characteristics and measures are very useful for the implementation of routing protocols and virus propagations studied in the next chapters.

Chapter II. Introduction to traffic sciences and traffic routing modeling

II.1 Introduction to traffic routing:

The traffic consists of the set of phenomena due to the displacement of information sent between the users or nodes in real complex network. It is characterized by the performance of the network as seen by users of network services that is truly paramount. This crucial point should be considered throughout the development of traffic engineering mechanisms and policies. For the optimization of traffic engineering several basic concepts must be explained.

II.1.1 Basic concepts :

a) Data packtes :

Typically, a data packet is a unit of data made into a single package that navigates along a digital network path such as the Internet. However, it can represent travelling items in other kinds of networks. Indeed, regarding the networks nature, that could be vehicles in roadways, airplanes in airlines electrical signals in neural networks and so on.

b) Packets generation rate:

This represents the number of packets R added to the network per time step.

c) Packets delivering capacity:

The delivering capacity or also the transmission capability is maximal number of packets that a node can transmit at each time step. For a basic traffic simulation this quantity is uniform and it is equal to $C = 1$ for all the network routers or nodes.

II.1.2 Routing protocol:

Routing protocol refers to the mechanism of selecting a path for traffic items in networks (links to follow in order to reach the destination) [106,104,130,143]. The process is governed by the routing protocol used. Fig II.1 shows an example of routing in a computer network. The algorithm selects the path with the minimum cost. According to the routing strategy adopted, the cost could be for instance, the distance, the waiting time, the realization cost and so on.

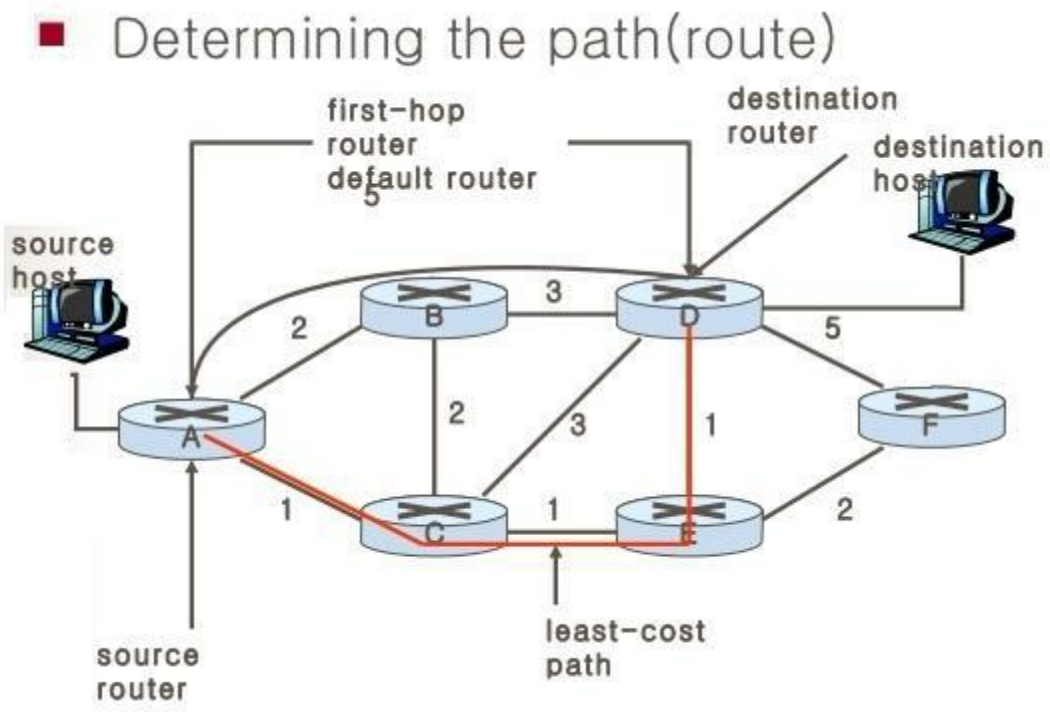


Figure II-1 : Determination of a path between a source host and destination host with the lowest cost in a computer network.

In general, there are two distinct modes of routing when we want to address the implementation of a routing protocol.

II.1.2.1 Static routing:

Static routing also known as non-adaptive routing does not involve any change in routing table unless the network administrator changes or modify them manually. The routing decisions are not made by current topology or traffic because the static routing systems can not react to network changes hence it doesn't require extra resources to learn the changes. That is the reason, static routing is considered as an efficient method for a small and simple network that does not change frequently and as inappropriate for large and constantly changing networks.

II.1.2.2 Dynamic routing

Dynamic routing or otherwise called adaptive routing is a superior routing technique which alters the routing information according to the altering network circumstances by examining the arriving routing update messages. When the network change occurs, it sends out a message to the router to specify that change, then the routes are recalculated and sent as a new routing update message. These messages pervade the network, enabling the router to change their routing tables correspondingly. The larger network requires

dynamic routing because with static routing larger networks could not be manageable and result in loss of connectivity.

II.2 Traffic routing protocol modeling:

II.2.1 Traffic Models :

Under the background of complex network, a basic traffic dynamics model has been proposed and frequently used to mimic the traffic transport in communication networked systems [58, 59,60,61,62].

In basic model, every node is considered as both host and router which can either generate packets or forward packets on the network. At each time step, R packets are generated in the network with randomly chosen sources and destinations. Every node can deliver at most $C = 1$ packets to its immediate neighbors based on a given routing protocol. Packets are sent by first-in-first-out (FIFO) procedure. The packet is removed immediately from the system once it arrived at its destination.

This basic model has also been generalized to more realistic models [66-68], which incorporate the fact that hub nodes usually have high delivering capacities or can generate more packets at each unit time step than those low-degree nodes can do [59,63,64].

II.2.2 Traffic measurements:

II.2.2.1 Traffic capacity:

Experimental calculation

Traffic capacity is one of the most important measurements for transport performance. In this context and in order to investigate the traffic capacity, Arenas. et al [69] proposed an order parameter for traffic capacity defined as:

$$\eta(R) = \lim_{t\to\infty} \frac{C\Delta N_p(t)}{R\Delta t}$$

Where, C is the delivering capacity of the nodes, R is the packet generating rate, and $\Delta N_p(t = N_p(t + \Delta t) - N_p(t)$ denotes the change of total number of queuing packets in the network at the interval Δt. Generally, the order parameter represents the balance between the inflow and outflow of packets. When $\eta(R) = 0$, the system is in the free flow

state, due to the balance of created and removed packets. With increasing packet generation rate R, there will be a critical value of R_c that characterizes the phase transition from free flow to congestion. When $R > R_c$, the order parameter increases: $\eta(R) > 0$, which leads to the accumulation of queuing packets and the congestion of the entire networks. Thus as clear in Fig II.2, R_c is a threshold at which a traffic phase transition occurs from free-flow state to congested state.

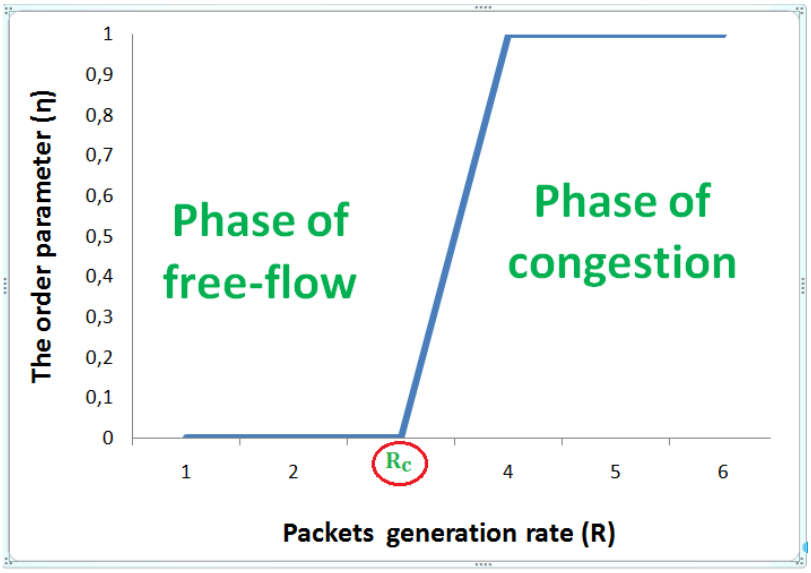


Figure II-2 : transition from free flow to congested phase.

Theoretical estimation:

For a topology based routing algorithm, it is possible to provide an analytical estimation of the network capacity using the concept of the betweenness centrality BC introduced earlier in chapter I (page 29). Originally, as we have seen, this measure referred to the number of total SP crossing a vertex v. It has been proved that if each node forwards only one packet at each time step through the SP, the network capacity would be estimated [61] by:

$$R_c = \frac{N(N-1)}{(BC)_{max}}$$

Where N the network is size and $(BC)_{max}$ is the largest BC value in the network.

II.2.2.2 Average traveling time:

The travel time or also the transport time of packet is define as the total time that the packet spends in the network from the creation until reaching its destination [71]. It could be also defined as the total waiting time at the queues of nodes along the packet trajectory plus the path length (number of the nodes along the path). We characterize the effectiveness of a routing strategy by calculating the average transmission time of all packets in the network, that is

$$< T >= \frac{1}{N}\sum_{i=1}^{N} T_i$$

Where N is the arrived packet number in the network, T_i is the overall transmission time of packet i, which consists of the transfer time from source to destination and the waiting time at the congested nodes.

II.2.2.3 Average path length:

As for the average travelling time, the average path length concerns only the packets arrived and removed from the system. It is equal to the sum of the packets path lengths divided by the number of removed packets. The path length means the number of hops that the packet performs from the source to the destination. Notice that the path length is nothing but the travel time minus the waiting time [72]. The average path length is calculating as:

$$< L >= \frac{1}{N}\sum_{i=1}^{N} L_i$$

Where N the arrived packet number in the network is, L_i is the overall transmission path length of packet i, which consists of the number of jumps from source to destination.

II.2.2.4 Traffic load of nodes

Traffic load of nodes in the network under different routing strategies is essential, for it reveals the utilization of all queue resources. Based on this quantity we can determine the nodes with high load and those with low load, In general congestion start to spread in network from excessive utilization of some nodes instead of all networks nodes. The node buffer load distribution is defined in [73] as follows :

Let $V = \{v_1, v_2, .., v_n\}$ be the set of nodes in entire network, the degree of each node is denoted as $D_e(v_i)$, for $i = 1, .., N$. Let $B_u(v_t,t)$ be the buffer load size of node v_i at time t. Then the average buffer load size of node v_i is defined as:

$$\overline{B_u}\,(v_t) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} B_u(v_t, t)$$

Let V^j be the subset of V, which consists of nodes whose degree is j, such that $V^j = \left\{V_1^j, \ldots, V_{|V^j|}^j\right\} \in V$, where $|V^j|$ is the number of nodes in V^j. Then the average buffer size of nodes, whose degree is j is defined as:

$$\widehat{B_u}(j) = \frac{1}{|V^j|} \sum_{i=1}^{|V^j|} \overline{B_u}\,(v_i)$$

II.2.3 Routing protocols:

Without routing strategies or protocols, packets in internet have no fast routes to arrive at their destinations and heavy traffic congestion may occur frequently. There is common demand on navigation or routing in these artificial systems. The main goal is to achieve higher traffic capacity leading to a best packets transport with high quality and without congestion spreading in the network. Therefore two general types of scenarios have been proposed to alleviate traffic congestion and improve traffic performance; namely, hard routing strategies [99–103], and soft routing strategies [149, 150, 74–98].

II.2.3.1 Hard routing strategies (optimizing network structure)

Hard strategies mean network topological structure is appropriately optimized so that network capacity can be enhanced. Adding or rewiring links is hard task; it's neither practical nor economic and usually has to consume much financial. On the contrary, removing links from networks is usually easy to be implemented at lower cost. For example in internet, network administrators can easily isolate some connections among computers through computer software. However the network optimizing structure strategies and due to their high cost have not received adequate attention compared with soft strategies (i.e., designing efficient routing strategies).

Removing links:

Enhancing transport capacity by removing links in communication networks is more economical compared with links addition. An effective method to enhance the traffic capacity of scale-free networks by closing some key links at heavily loaded times is proposed in [93],the value of $(k_m \times k_n)$ is used instead of the links' betweenness centrality, which measures exactly the expected number of packages flowing through the link, first because it has been found that the betweenness centrality of links has strong correlation almost linear with $(k_m \times k_n)$ [105]; second it is easier to rank the links by using the local information, since the calculation of betweenness centrality needs system-wide information.

In the proposed method, first ranks the links according to the value of the product $(k_m \times k_n)$, where k_m and k_n are the links' end node degrees. Then, links are closed according to this order from large to small. Because hub nodes are usually more important and bear more traffic loads, the links with larger values of $(k_m \times k_n)$ are easier to jam. Hence, removing highly congested links can lead to the redistribution of traffic loads along links so as to enhance the overall packet handling and delivering capacity.

Fig II.3 shows the increment of network traffic capacity R_c versus the number of closed links L_c under the shortest path routing strategy. Results show that on closing the links according to the order of $(k_m \times k_n)$, traffic capacity can be remarkably enhanced.

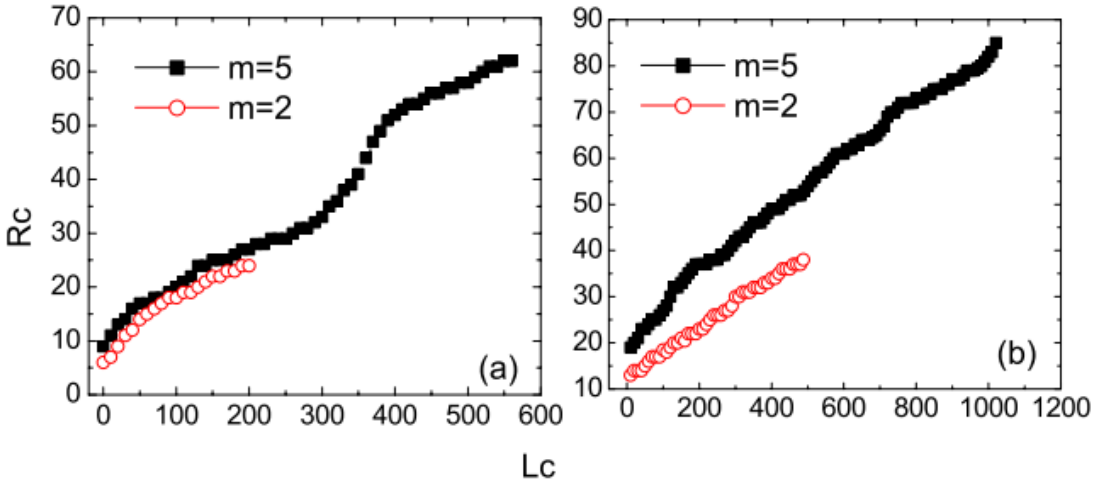


Figure II-3: (Color online) Critical R_c vs L_c under global routing strategy with network parameters (a) N=1000, m=2 and 5; (b) N =5000, m=2 and 5. The data are obtained by averaging R_c over ten network realizations.

Adding links:

Adding links to network to enhance traffic capacity is considered as the more expensive. In [102], the authors based on shortest path strategy proposed a strategy that can effectively enhance the traffic capacity via the process of adding nodes and links. It is proved that nodes with high degree (hub) under the shortest path routing strategy become congested more than nodes with small degree due to the extensive use of hubs. In the strategy of [102], shortcut links are added among nodes that have the longest shortest path lengths. The shortcut links are placed in proper positions to avoid packets flowing through hub nodes so that there are not too many packets accumulated on hub nodes. As a result, the traffic capacity is reduced. Therefore heavy accumulation of packets on hub nodes should be avoided as much as possible. Thus the network capacity increases. Fig II.4 shows the traffic capacity R_c and the average shortest path length l_{ave} versus the fraction of new added links over the total L existing links f_α under the shortest path routing strategy. From Fig II. 4 (a), it is found that both the strategy of [102] and the random strategy can enhance the traffic capacity R_c when links are added into networks. But the traffic capacity R_c is enhanced more using the strategy of [102] than using the random strategy.

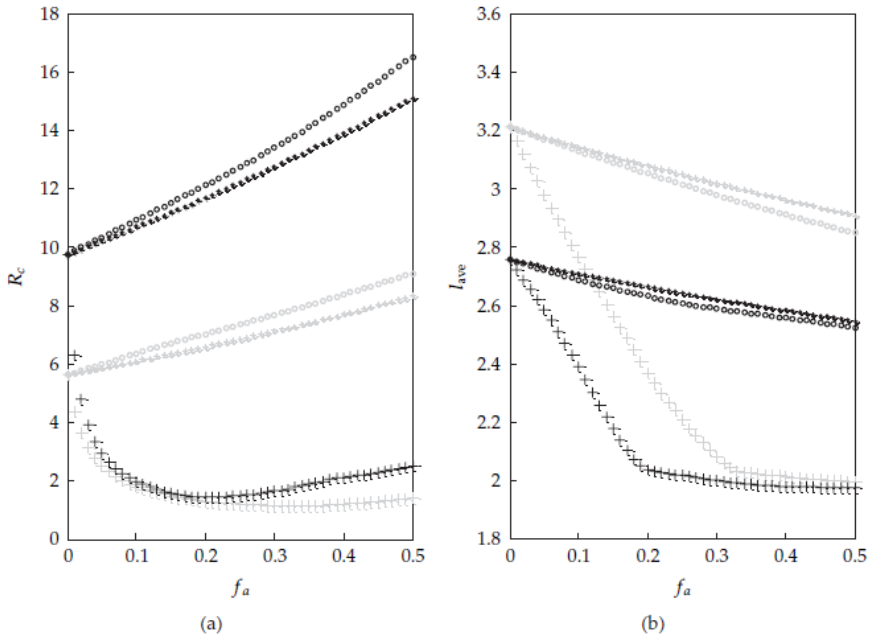


Figure II-4: (a) the traffic capacity R_c and (b) the average shortest path length L_{ave} versus the fraction of new added links over the total L existing links f_α under the shortest path routing strategy. In each figure, the gray and black colors are for the cases with m0 =m=3 and m0=m =5, respectively. The network size is set to N =500. The circles, asterisks, and plus signs are the results of the strategy in [102, the random strategy, and the strategy that adds links among hub nodes, respectively.

II.2.3.2 Soft routing strategies

In most cases and compared with high developing cost of changing network structure and reallocating the network resources of well established networked systems, designing efficient routing strategy is always more economic and practical for enhancing the network capacity. In general these soft strategies can be implemented by using, global or local network information, node's delivering capacity and priority policies.

Using global informations:

In global routing strategies the global information is needed, such as the characteristics of each node or the real-time information of each transmitted packet [72, 77, 79].

Shortest path routing strategy (SP): Among numerous different kinds of proposed routing protocols, the shortest path strategy [79], is widely used in real communication systems [107] because of its simplicity and its low implementation cost; in the shortest path routing strategy, each packet is transported through the topological shortest path between the packet's sender and receiver and its outgoing is determined by the current routing table. While dynamic routing strategies make use of dynamic and stochastic factors and the network routing table is recomputed at each time step thus consuming computation time and available bandwidth. Moreover, the shortest path routing cannot be easily replaced by the dynamic routing on the Internet currently due to high economic cost and implementation technology. However the network under the shortest strategy have very low capacity and the congestion spread easily to nodes with high degree due to their high exploitations .Therefore others routing strategies where proposed in order to overcomes this shortest path drawbacks.

Efficient path routing strategy (EP) : in this strategy [77], the global static information is needed where the path between sources and destination is determined as follow:

$$P_{i\to j} = min \sum_{e=0}^{L} {k_e}^{\theta}$$

Where k_e represent the degree of node e on the path between i and j and θ is an adjustable parameter. It's obvious that the shortest path strategy is a particular case of the efficient path strategy when $\theta = 0$. The optimal value of $\theta = 1$. R_c First increases with θ and then decreases, with the maximum of R_c corresponds to $\theta \approx 1.0$. In comparison with the

shortest path routing case (i.e. $\theta = 0$), the capability of the network in freely handling information is greatly improved, from Fig II.5 $R_c \approx 3.0$ when $\theta = 0$ to $R_c \approx 20$ when $\theta = 1.0$; more than six times. This result suggests us the effectiveness of the efficient routing strategy.

Global dynamic strategy (GD): An efficient network routing strategy should not only consider the topology of the network, but also the effects of the queue length of nodes. The author in [72] introduced a global dynamic routing strategy for the networks. In this routing strategy, the packets are delivered along the path in which the sum queue length of nodes is a minimum. From the many paths between the source and destination, the path in which the sum of node queue lengths is a minimum is selected. Therefore, the path between nodes source i and destination j can be denoted as:

$$P_{i \to j} = \min \sum_{e=0}^{L} (1 + q_e)$$

Where q_e represent the queue length of node e on the path between i and j. Fig II.5 compares the relation of order parameter η versus the packet generation R under the three global routing strategies (SP, EP and GD). It can be seen that the traffic capacity under the shortest path strategy [79] is $R_c = 3$. The efficient strategy in [77] has $R_c = 20$ and the global dynamic strategy can reach up to $R_c = 41$. Therefore, the global dynamic strategy can achieve the highest traffic capacity.

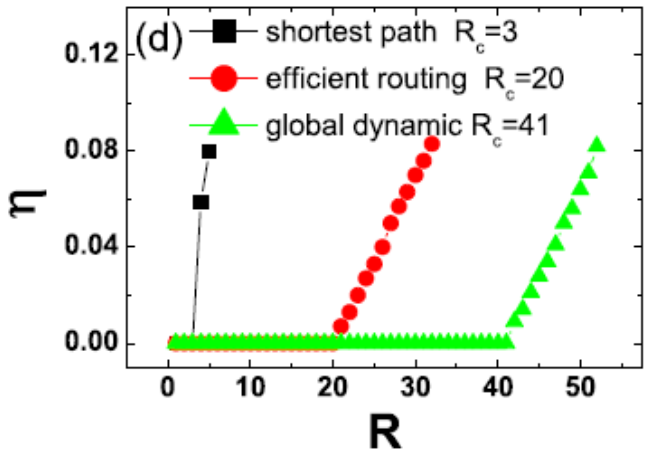


Figure II-5: The order parameter η vs R under the three routing strategies: shortest path, efficient path, and global dynamic. Network size N=500; delivering capacity C=1[72]

Using local informations:

Designing global routing protocol in which every node has the global or dynamic network information is very efficient for small and medium networks, but not in the case of large and huge evolving real networks such as the Internet, WWW; due to technical cost of computing and memorizing much information per time step [43]. Therefore, strategies based on local information [74, 75, 76] where each node only knows the information of its neighbors are very interesting and practical in large real networks.

Local static routing strategy (LS): In this strategy [74], each node performs a local search among its neighbors. If the packet's destination is found within the searched area, it is delivered directly to its target; otherwise, it is forwarded to a neighbor node according to the preferential probability: $P_i = \frac{k_i^{\alpha}}{\sum_{j=1}^{N} k_j^{\alpha}}$

where the sum runs over the neighbors of node i and α is an adjustable parameter. As shown in Fig II.6. This result indicates that when $\alpha = -1$, the system's capacity can be enhanced maximally.

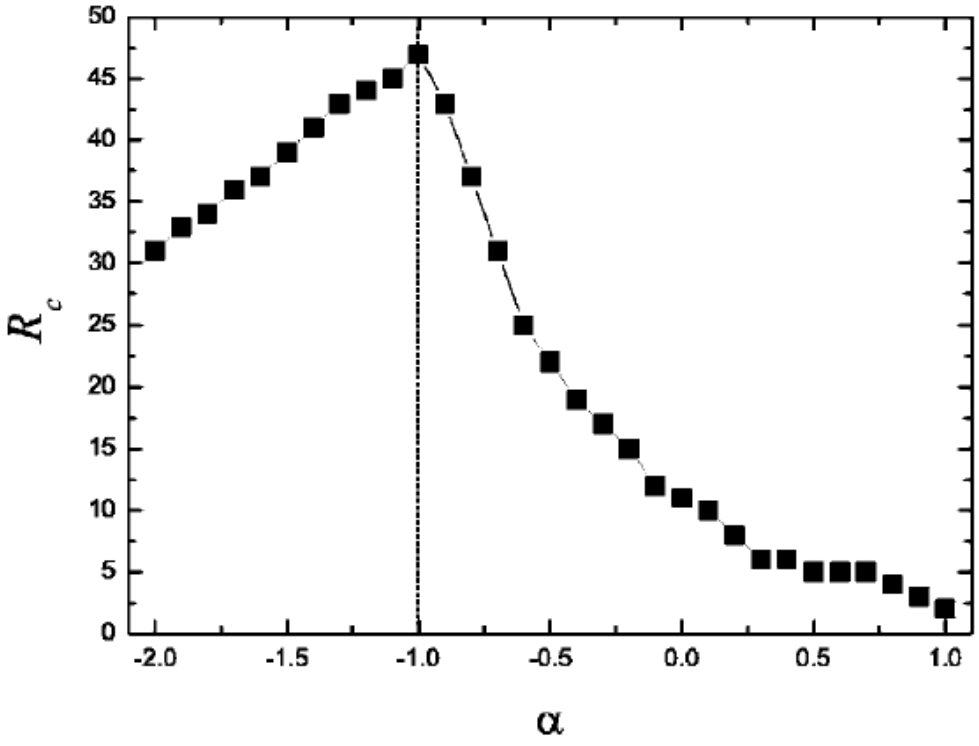


Figure II-6: The critical R_c versus α with network size N=1000 and constant node delivering capacity C=10. The maximum of R_c corresponds to α =−1 marked by a dotted line.

Local dynamic routing strategy (LD): in this strategy [75], the local search is based on both neighbor's degree k_i and queue length q_i:

$$P_i = \frac{k_i(1+q_i)^\beta}{\sum_j k_j(1+q_j)^\beta}$$

Where the sum runs over the neighbors of node i and β is an adjustable parameter.

In Fig II.7, the order parameter η as a function of generating rate R for different parameter β. For each β, when generation packet rate R is less than a specific value R_c , η is zero; it suddenly increases when R is slightly larger than R_c. Moreover, in this figure, different β corresponds to different R_c , thus the authors investigate the network capacity R_c depending on β for finding the optimal value of parameter β. We can see clearly the impact of adding the dynamic knowledge in the protocol. Indeed, when $\beta = 0$ the routing is done with respect to only local topological features the network capacity is very limited. The optimal value of β maximizing the network capacity is -3.

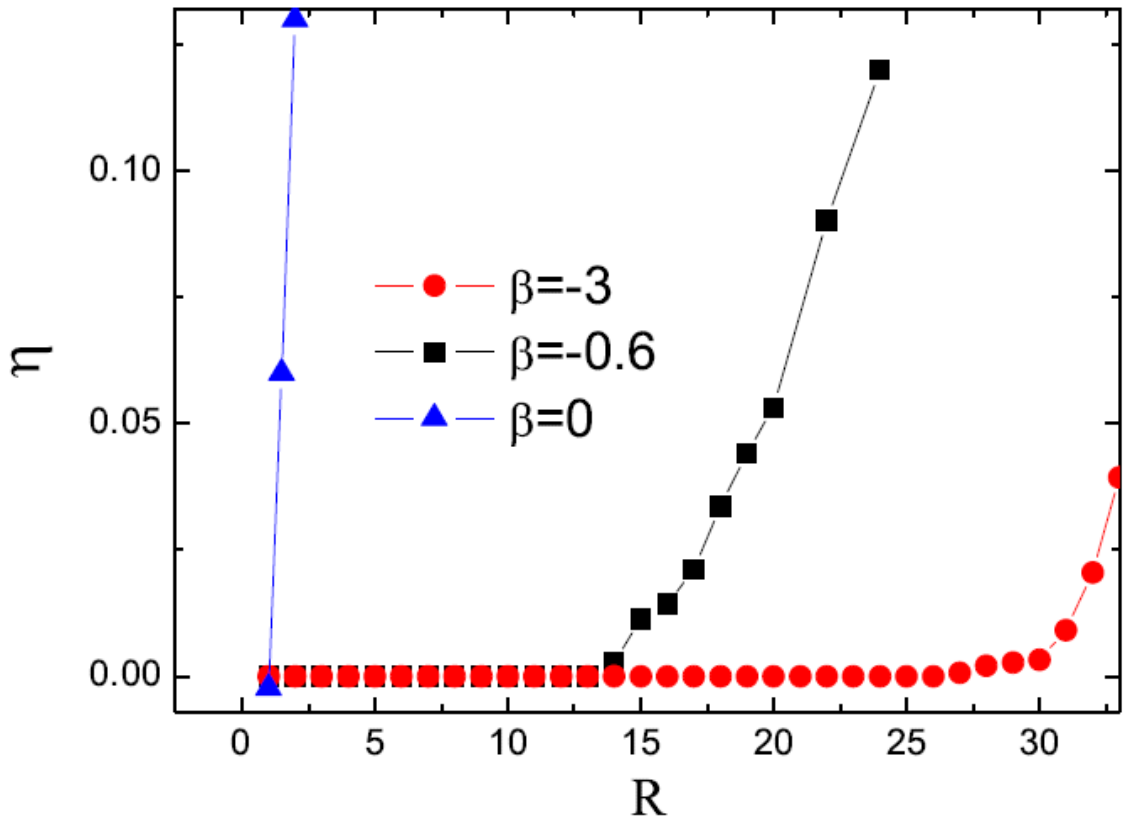


Figure II-7: The order parameter η as a function of generating rate R for different value of parameter β. Other parameters are delay = 0, C = 5 and N = 1000 [75].

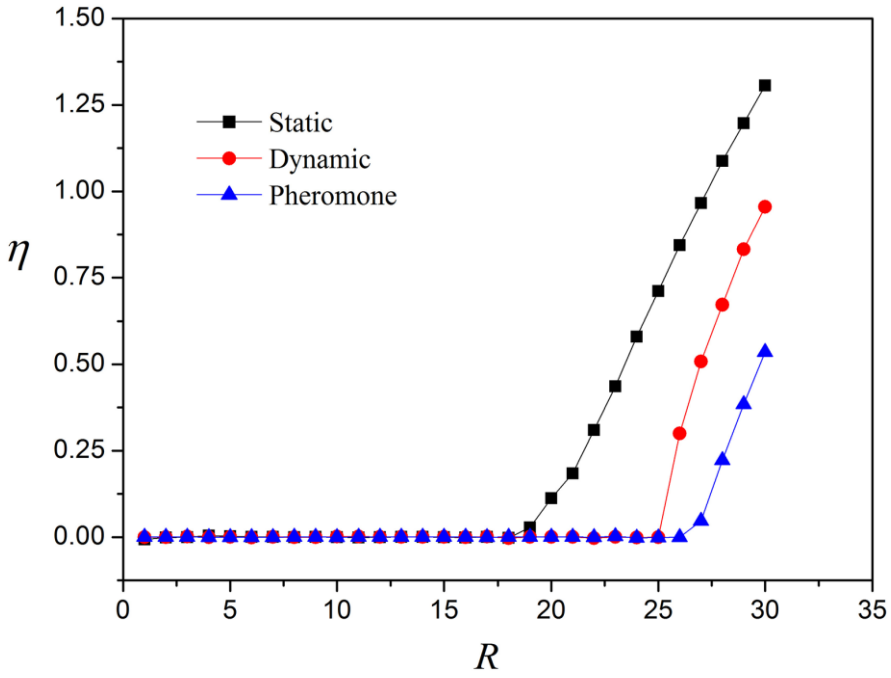


Figure II-8 : Order parameter η versus R for different routing strategies. Local static routing strategy, local dynamic routing strategy and local pheromone routing strategy. N = 1024, the mean degree is K = 10 and the delivering capability of each node is C = 5.

Local pheromone routing strategy: Ling et al. [76] proposed the local pheromone routing strategy, enlightened by a chemical substance which is laid by ants so that they can pick up their trails. To mimic the production and evaporation of the pheromone to iterate to the best route, the algorithm defines as:

$$\prod i \to j = \frac{P_{ij}^{\alpha}}{\sum_{j=1}^{N} P_{ij}^{\alpha}}$$

where $\prod i \to j$ is the probability of the information packets transmitted from node i to neighbor node j. P_{ij} is the pheromone of the link pointing from node i to j, and α is a tunable parameter. Initially, the value of the pheromone concentration on each link δp is set to a small unit value (set to be 0.001). Let L_c be the critical queue length of a node.

When a packet is delivered successfully from node i to node j, the pheromone of link decreases by a unit if $L_j > L_c$, i.e.

$$P_{ij} = max\{P_{ij} - \delta p, \delta p\}$$

Otherwise, if $L_j \leq L_c$, the pheromone of link increases by a unit, i.e.,

$$P_{ij} = P_{ij} + \delta p$$

L_c is set as $L_c = \beta C$ and β is a tunable parameter and C is the node delivery capacity. The best performance emerges at $\alpha = 1$ and $\beta = 2$ [76]. Fig II.8 compares the relation of

order parameter η versus R under the three different local routing strategies. It can be seen that the traffic capacity under the pheromone routing strategy can achieve the highest traffic capacity from all proposed local strategies.

Using Hybrid information:

Hybrid strategies are considered as another important way to enhance the performance of traffic in complex networks. The name hybrid comes from the fact that a different kind of information is considered in this category of strategies. Within this context, many models have been suggested [19, 24, 33, 34, 47, 70]. In [34], Tan et al. proposed a hybrid strategy which combines node degree and traveling time information's together and therefore can balance the traffic between hubs and peripheral nodes more effectively. Simulation results show that the network capacity can be enhanced considerably, and the average traveling time is also shortened sharply. In [108], a new hybrid routing to enhance traffic capacity strategy was proposed based on combining the shortest path [79] and the global dynamic [72] routing strategies.

Using node's delivering capacity

Optimizing the allocation of the limited traffic resources is another way to maximize network capacity in complex networks. In this context many researchers have been proposed to do so despite the high economical cost of this type of strategies. Yang et al. [109] proposed a strategy in which node capacities are assigned based on their degrees. While Ling Xiang [110] proved that when the node capacity is set to be proportionally to the node betweenneess centrality; the more paths going through a node the more delivering capacity will be assigned to it. With this betweenneess allocation strategy the network capacity increase and reach very high value more than the case where the allocation is based on nodes degree or uniform (all nodes have the same capacity). Fig II.9 depicts the network capacities for networks with different sizes N. We can see that with the allocation strategy based on nodes betweenness centrality in shortest path strategy, network capacity R_c is the largest for different network sizes in comparison with the uniform delivering capacity.

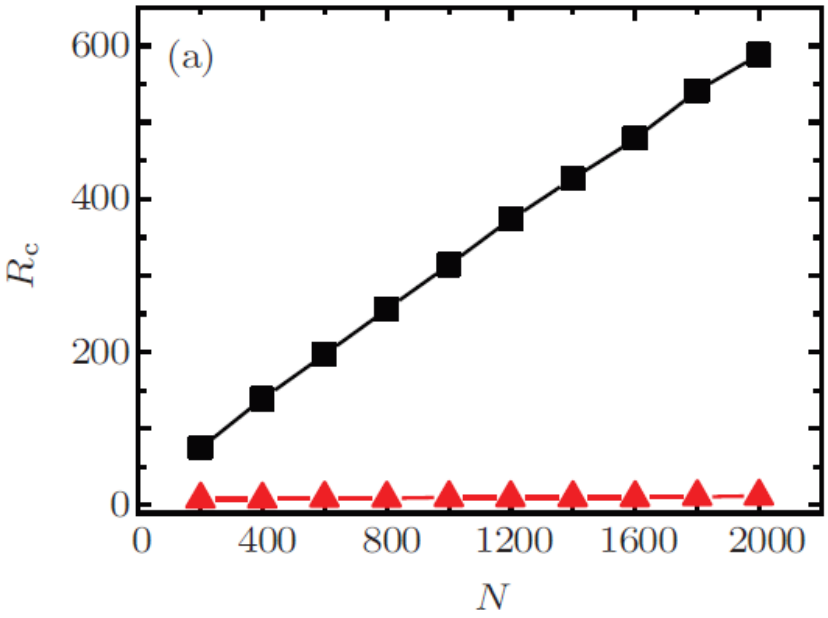


Figure II-9: The R_c vs network size N in the shortest path protocol for BA scale-free networks with size N=1000 and average degree <k> = 8, the delivery resource is allocated with uniform (red color) and (betweenness allocation) (black), respectively.

Using priority policy:

Priority policy is another way for increasing networks capacities. For example using FIFO policy is fair and minimizing both cost and time but on the expense of network capacity. Therefore doesn't give the best performance [111, 112]. For this reason many studies proposed to apply other priority policies in order to enhance traffic capacity. These studies would be further discussed in chapter III. We will propose also another priority policy based on packet destination instead of those priority policies used previously in literature.

II.3 Conclusion:

In this chapter we have introduced tools which are mandatory to understand and measure traffic on networks. We went over the most memorable methods that researchers had used to implement efficient routing protocols.

We explained their Principe and also shed some light on the differences between each routing strategy in order to understand which one is the most appropriate according to the context. However we brought to a conclusion that there are many factors that decide about the efficiency of each method other than the network capacity, such as the implementation facility and cost, the network size and type etc.

In the next chapter we will come up with new priority policy based on packet destination instead of given probability distribution independently of the node source or destination under shortest path leading to reduce the congestion and maximize the prioritized traffic without affecting both the unless traffic priority and the whole network capacity.

Chapter III. Prioritization of traffic flow in complex networks

III.1 Introduction:

In traffic engineering, another idea is used in conjunction with a given routing protocol. This is prioritization. In the internet, the network traffic is prioritized and is classified into different types according to the payments, or depending on time sensitive packets: give important network traffic precedence over less important network traffic; namely, packets with relatively higher payment or sensitivity need to be delivered faster than others.

An example of traffic which is in general prioritized is real-time video conferences and telephone communication which have to be sent with high priority compared with other packets such as data files for download or mails which can tolerate longer traveling time.

Despite its importance in routing studies, prioritization has not received adequate attention, until recently. In [113], Zhong et al. based on global dynamic routing proposed a dynamic source routing (DS), in which flows of two priorities levels are treated differently in two aspects. Although traffic capacity can be improved, Zhang et al. in [73] and based on efficient routing designed a new strategy with multiple priorities concept; a new optimal path depending on nodes degree is found for the packets with high priority that maximizes the network capacity.

In [114], Li et al. proposed routing protocol model based on the different priority of traffic in which the capacity of the network can be improved. The authors in [115] proposed a priority queuing discipline on networks of mobile agents; the proposed discipline improves remarkably in both the network throughput and the packet arrival rate.

However, in all the previous works, prioritization is attached to packets following a given probability distribution independently of packets' information, until recently Du et al. [112] introduced a shortest-remaining-path-first queuing strategy into traffic network model. They found that the traffic behavior of the system is improved in the congestion region but in the free flow region, there are no evident changes.

In this context in this chapter, in order to improve the network capacity, we propose a new prioritization model of traffic flow where packets under shortest path strategy are prioritized according to their destination [192]. The practical importance of this model is

the possibility to companies in the internet to pay for the prioritization of traffic sent to their servers.

Before giving further explanation about the methodology followed in the creation of our priority model, first we evoke thoroughly the shortest path routing strategy used in this study.

III.2 The shortest path problem (Dijkstra algorithm)

The most popular algorithm that allows the determination of a unique shortest path (SP) between two nodes in directed/undirected weighted graphs (with positive cost) is Dijkstra algorithm. However, many other algorithms are meant to solve the problem of the SP in a given network. For instance, the Bellman-Ford algorithm that tolerates negative weights unlike Dijkstra algorithm [116,117], the Floyd-Warshall Algorithm and the Genetic Algorithm which is based on biological evolution and that may give different solution at each execution which might be an advantage over the others [118-122].

Since we are not interested in negative weight costs, and due to the acceptable execution time of Dijkstra algorithm we have implemented this algorithm to find when needed the SP in all our simulation programs.

III.2.1 Dijkstra algorithm:

To explain how Dijkstra algorithm does works, let's consider the weighted graph $G(E,V)$ and w: $E \rightarrow R$ is the corresponding weights function $w(e)$ associating each edge $e = (u;v)$ between two nodes u and v, to a real positive value weight. We define the two following terms:

- **The length of a path p** $=< \mathbf{v_0}, \mathbf{v_1}, .., \mathbf{v_k} >$ is the sum of the weights of its constituent edges:

$$lengh\ (p) = \sum_{i=1}^{k} w(u,v)$$

- **The distance**

The shortest distance or simply the distance from a vertex u to another vertex v, denoted $\delta(u,v)$ is the length of the SP between u and v if it there is one, if u and v

cannot be linked (disconnected graph) $\delta(u,v)$ is equal to infinity. See the example in Fig. III.1.

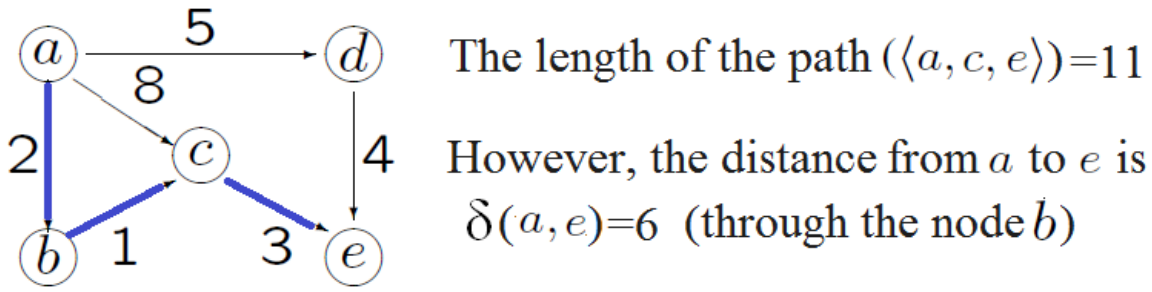


Figure III-1 : Example of the calculation of a length of a given path and the distance between two nodes.

- **The Problem:**

Given a weighted graph with positive edge weights $G = (V, E)$ and a distinguished source vertex ϵ V , we aim to determine the distances and the SP from the source s to every other node in the graph G.

a) Initialization:

To begin, we have to set the following tables:

d[v]: Which is an estimation of the length $\delta(s,u)$ of the SP for each vertex v from s, in other words it is the estimation of the distance between the source s and the vertex v.

Initially, $d[s] = 0$ and $d[v] = \infty$ for all the other vertices, and at the end of the algorithm we should have $d[v] = \delta\ (s,v)\ \forall\ v \in V$.

$Prev[v]$: The previous, the predecessor or parent array; It is the previous node to reach u from s through the shortest known (discovered) path. Notice that $Prev[v]$ is also the next step from v if we are going in the opposite direction i.e. from the node v to the node s.

Initially, we don't have this information so we set $Prev[v] \equiv None\ v\ \in V$.

$Visit[v]$: When the algorithm is seeking for the SP from, this table indicates whether the node v was visited or not visited yet; it takes the value "Yes" if it has been already visited and "No" if not. Initially $Visit[v] \equiv No\ \forall\ v\ \in V$.

b) The progress

In order to establish the SP and the distance to any v from s, the algorithm will visits and treats the nodes one by one in some order; we start the proceeding of the vertices from the closest vertex to the source. It is simply the neighbor $\in Ad[s]$ which is linked to s with the minimum weight. Then we treat all remaining neighbors from the nearest to the furthest (minimum to the maximum weight).

Once all the adjacent vertices have been explored, we move to the next neighbors and treat them following the same order priority i.e. beginning always from the closest. We continue until all the nodes are visited.

Visiting a node u means that when its turn comes, we compare the old known estimated path with the path passing through that node u. If it is better than the current value (lower cost), we update the distance $d[v]$ and that of all the other neighbors of u if necessary. The process by which an estimated distance is updated is called relaxation.

The relaxation :

To decide whether we update $d[v]$ or not, we have to compare the current value of $d[v]$ and total weight of the path from s to u, plus the weight of the link $[u, v]$.We do the relaxation only if the new path from s to v is shorter than $d[v]$. If it is the case, then we replace the old path $\langle s, :::, x, v\rangle$ with the new shorter path $\langle s, :::, u, v\rangle$ and u become the previous node to v from s instead of x (see Fig. III.2).

$$If\ d[u] + w(u, v) < d[v]$$

$$Then\ d[v] = d[u] + w(u, v)$$

$$And\ \ Prev[v] = u(formerly\ Prev[v] == x)$$

Remark: During the algorithm process, always the estimation $d[v] \geq \delta(s, v)$ and $d[v]$ equals to the length of the shortest known path.

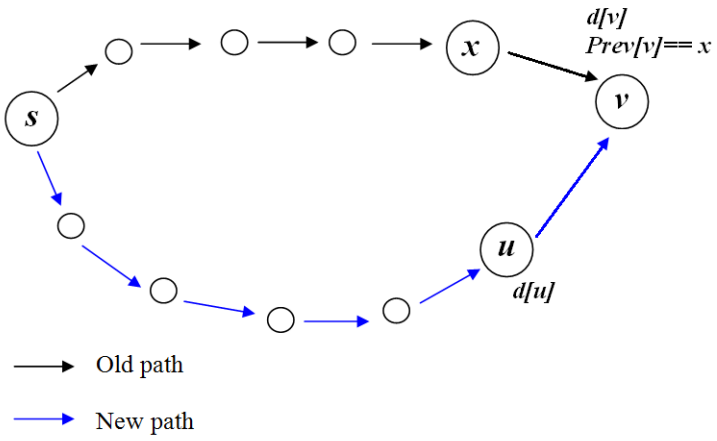


Figure III-2: Illustration of the relaxation proceeding performed when discovering new path with less cost.

c) End of the algorithm

When all vertices have been treated, for all the vertices the estimation is validated as being the real shortest distance. Thus, $d[v] = \delta(s, v)$ and no further relaxations are done. Of course if s and v aren't connected $d[v] = \delta(s, v) = \infty$.

$Prev[v]$ now indicates the closest neighbor of u to the source s, which means that the table $Prev[v]$ gives the shortest sequence of nodes from s to u and vice versa. Using this table we can build the SP tree of the node s and the SP routing table of the network as we will see in the following example.

III.2.2 Exemple:

We consider the graph in Fig. III.3-b. We aim to build the shortest tree related to the node s (source). In order to do so, we have to find the shortest paths from the node s to every other nodes of the graph.

a) Step zero

The initialization of table regrouping:

$d[v]$ The shortest estimated distance between s and v, $Prev[v]$ and $visit[v]$. (Fig. III.3-a)

u	s	a	b	c	d
d[u]	0	∞	∞	∞	∞
Prev[u]	None	None	None	None	None
Visit[u]	No	No	No	No	No

(a)

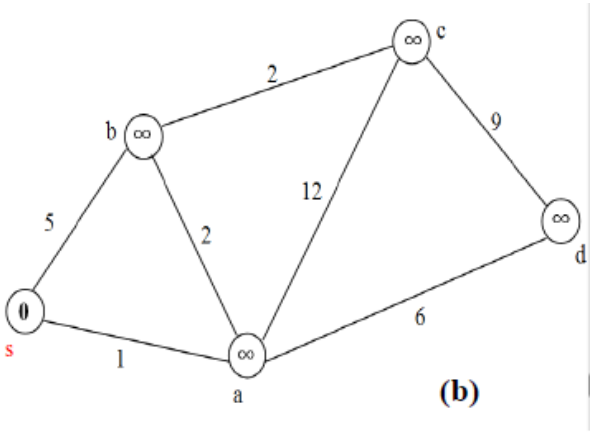


Figure III-3: Step zero of Dijkstra algorithm; (a) initialization of the tables and (b) graphic representation of the estimated distances

b) Step 1

We begin visiting the node with less known distance $d[u]$ from s. It is the node s itself. We examine the distances from s to its adjacent nodes and update the routing tables. In this example s has two neighbors: $Adj\ (s) = \{a, b\}$. At the moment, all the shortest distances are infinity; we have then to do the update of the distances of the neighbors.

The new distances are nothing else but the weights of the links between the s and its neighbors ($d[a] = w(s, a) = 1; d[b] = w(s, b) = 5$). The previous node is s. The node s is visited, now we explore its neighbors. $Visit[u] \equiv Yes$. Fig. III.4 shows in tables and in graphic representation the progress of the algorithm.

u	s	a	b	c	d
d[u]	0	1	5	∞	∞
Prev[u]	None	s	s	None	None
Visit[u]	Yes	No	No	No	No

(a)

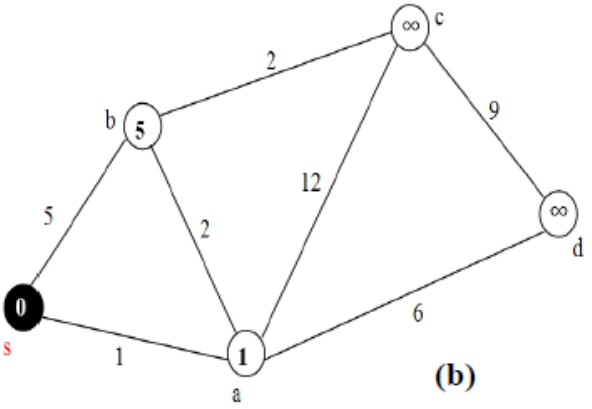


Figure III-4: Step one of Dijkstra algorithm exploration of the source s and updating initial information.

c) Step two

We spot the neighbor with least distance from s. In this case it is a. For this current node, we calculate the distances of its neighbors $Adj\ (a) = \{b, c, d\}$ from the starting node s. We only update the distances table if the calculated distance is less than the known distance, which means only $If\ d[Adj\ (a)] > d[a] + w(a, Adj\ (a))$. The Fig. III.5 is an update of The Fig. III.4.

u	*s*	*a*	*b*	*c*	*d*
d[u]	*0*	*1*	*3*	*13*	*7*
Prev[u]	*None*	*s*	*a*	*a*	*a*
Visit[u]	*Yes*	*Yes*	*No*	*No*	*No*

(a)

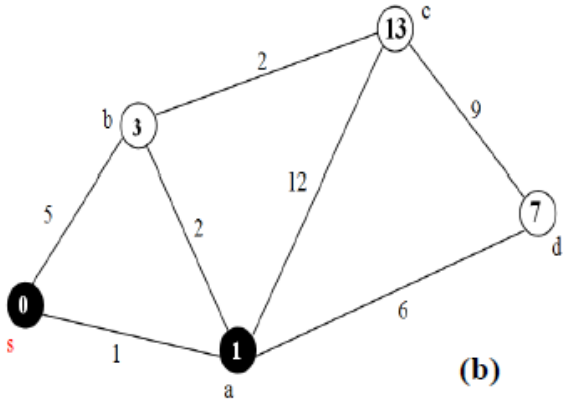


Figure III-5: Step two of Dijkstra algorithm exploration of the first neighbors of the source s; we start from the closest neighbor to the source s (node a) and update the information of step one.

Henceforth, the node a is visited, we can move to b. We do the same thing as we did in the node a. We noticed that although the node b is a neighbor of the node s, the shortest path to get to b via s is not direct. In fact the shortest distance $d[b] = 3$ via the node a. Considering this distance, have to do the relaxation and change the distance of the node b The Fig.III.6.

u	*s*	*a*	*b*	*c*	*d*
d[u]	*0*	*1*	*3*	*5*	*7*
Prev[u]	*None*	*s*	*a*	*b*	*a*
Visit[u]	*Yes*	*Yes*	*Yes*	*No*	*No*

(a)

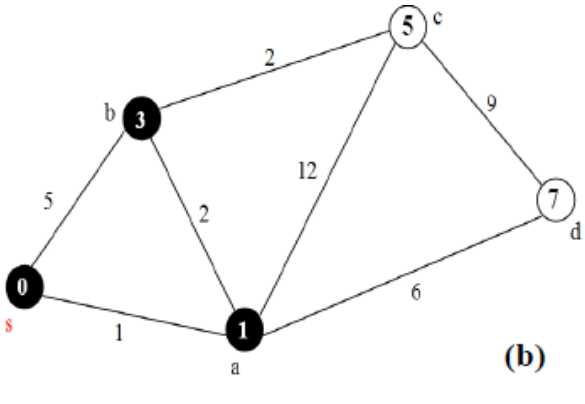


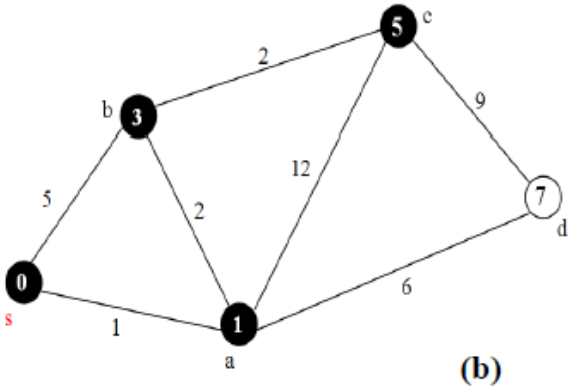
Figure III-6: Step two of Dijkstra algorithm exploration of the first neighbors of the source s; we move to the next closest neighbor to the source s (node b) and continue the update of the information from step one.

d) Step three

We have processed the first neighbors, now we explore the next neighbors of the starting node s, $Adj\ (a) = \{b, c, d\}$ and $Adj\ (b) = \{a, c\}$. As we did earlier for the first neighbors, we begin with the one that has the shortest known distance. From $d[u]$ table above, it is the node c (The Fig. III.7).

u	s	a	b	c	d
$d[u]$	0	1	3	5	7
$Prev[u]$	None	s	a	b	a
$Visit[u]$	Yes	Yes	Yes	Yes	No

(a)



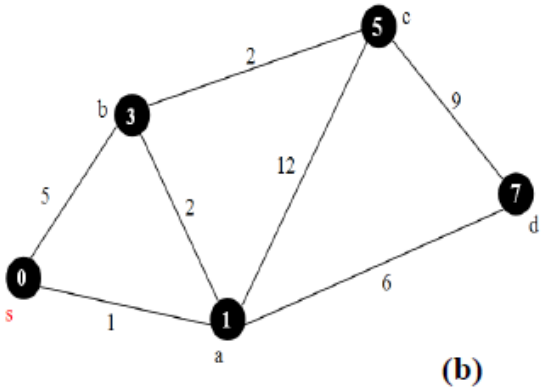
(b)

Figure III-7: Step tree of Dijkstra algorithm exploration of the next neighbors of the source s; we start from the closest neighbor to the source s (node c) and update the information of step two.

We repeat the same thing as before; we compare the distance of the remaining node $d[d]$ to $d[c] + w(c, d)$. since this time $d[d] < d[c] + w(c, d)$, we keep this distance (Fig.III.8).

u	s	a	b	c	d
$d[u]$	0	1	3	5	7
$Prev[u]$	None	s	a	b	a
$Visit[u]$	Yes	Yes	Yes	Yes	Yes

(a)



(b)

Figure III-8: End of Dijkstra algorithm; by visiting the last node in the graph, we have processed all the nodes and no more relaxation is possible.

now that we have visited all the vertices, we draw the shortest distance tree related to the node s. To do so, we delete the unused edges that don't figure in any of the SP (Fig.III.9).

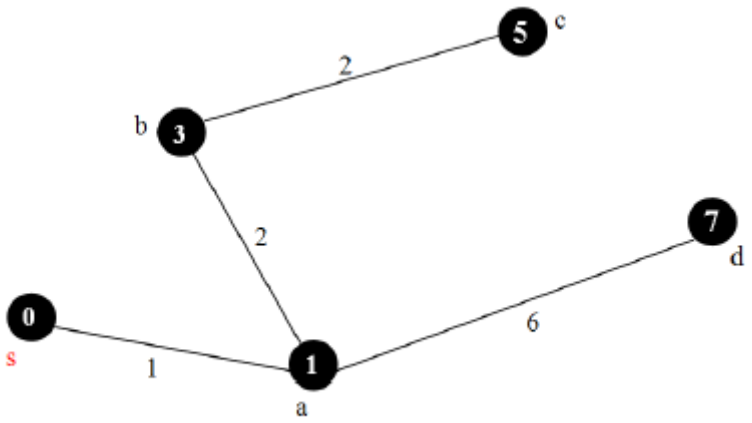


Figure III-9: The SP tree related to the node s.

e) The SP routing table

In order to establish the routing table of the shortest path between every two chosen nodes among the network, we have to change the source s and repeat the algorithm for all the network's nodes. For each node we stock the corresponding $Prev[u]$ and write the routing matrix. In the example above the routing table would be as shown in Fig.III.10.

Source \ Destination	s	a	b	c	d
s	0	a	a	a	a
a	s	0	b	b	d
b	a	a	0	c	a
c	b	b	b	0	d
d	a	a	a	c	0

Figure III-10 : The corresponding SP routing table established using Dijkstra algorithm

Notice that the intersection between a line i and a column j gives the next node from the source i toward the destination j . For example to reach b starting from d via the SP, the intersection between the line d and the column b indicates that next we have to go to a. Once in a, the intersection between the line a and the column b indicates that next we go to b (see Fig.III.10).

In all what follows, the principle of finding the SP remains the same; the only difference is that all the weights of the links are equal to one.

III.3 Implementing beneficial prioritization of traffic flow in complex networks

As we described earlier in this chapter, we have stated that we are going to assume that packets are prioritized based on their destination and routed through the SP. In our model, when a packet is generated it is labeled as, High or Low Priority based on its destination.

III.3.1 Description and simulation steps

We describe here the steps involved in our priority model:

i. First a fraction f of nodes is chosen from network, then subsequently, any generated packet with destination within the fraction f will be labeled as High Priority; otherwise it is labeled as Low Priority.

ii. The problem is which nodes should we prioritize its coming traffic? To answer this question, we test in independent simulations three different selections criteria and compare between them:

- Random selection: we prioritize the traffic coming to nodes chosen randomly among all network nodes.
- High degree selection: we prioritize the traffic coming to nodes with the highest connectivity degree.
- Low degree selection: we prioritize the traffic coming to nodes with the lowest connectivity degree.

We use the following probability distribution for the selection of a given node to be in the fraction f:

$$P_i = \frac{k_i^\alpha}{\sum k_j^\alpha}$$

Where k_i is the degree of node i and the sum is taken over all nodes. α is a tunable parameter. For, $\alpha > 0$, the prioritized packets are mostly among those having hubs as destination. *For* $\alpha < 0$, the traffic going to peripheral nodes are prioritized instead, and for $\alpha = 0$, the packets are randomly prioritized, and the procedure reduces to the random prioritization of packets regardless of their destination. This process results in dealing with two types of traffic: High Priority Traffic (HPT) (privileged-traffic) and Low Priority Traffic (LPT).

During the simulation time we repeat the following steps:

1) R packets are created with random sources and destinations.

2) Each node delivers one packet ($C = 1$) to its neighboring nodes according to the SP. Non prioritized traffic (LPT) packets are sent by first-in-first-out (FIFO) procedure. Prioritized traffic (HPT) packets are sent before the LPT packets.

3) The packet is removed from the system once it reaches its destination.

4) We assume that the queue length of every node is unlimited.

III.3.2 Simulation results and discussion

For each case we use as platform the well known BA network [26] which is as we have seen a SF network. Let's us recall briefly the steps of the BA algorithm; we start with m_0 full connected nodes then at each time step a new node is added and established m new links with the existing nodes according to the preferential attachment process. In our case the network size $N = 500$, m_0=3 and $m = 2$.

III.3.2.1 Evolution of the network capacity

In order to investigate the traffic behavior under prioritization, we use the order parameter defined in chapter II as [69]:

$$\eta(R) = \lim_{t\to\infty} \frac{C\Delta N_p(t)}{R\Delta t}$$

Where, C is the delivering capability of the nodes, R is the packet generating rate, and $\Delta N_p(t) = N_p(t + \Delta t) - N_p(t)$ denotes the change of total number of queuing packets in the network at the interval Δt.

Generally, the order parameter represents the balance between the inflow and outflow of packets. When $\eta(R) = 0$, the system is in the free flow state, due to the balance of created and removed packets. With increasing packet generation rate R, there will be a critical value of R_c that characterizes the phase transition from free flow to congestion. When $R > R_c$, the order parameter increases: $\eta(R) > 0$, which leads to the accumulation of queuing packets and the congestion of the entire networks. Thus, R_c is a threshold at which a traffic phase transition occurs from free-flow state to congested state.

As traffic types are treated differently by our priority policy, we will define an order parameter for each traffic type: η_{HPT} for high priority traffic and η_{LPT} for low priority traffic. We compute also the normal traffic parameter η_{WP} of SP strategy without priority. These three parameters are computed using the respective packet number N_p of the corresponding traffic type. As we have explained above in the priority policy, we proceed to prioritize a node fraction f from the network.

First, we target nodes with high degree by fixing $\alpha = 1$. Fig.III.11 shows the results of the order parameter of HPT, LPT, and WP traffic. From Fig.III.11-a , we find that the prioritization of hubs nodes is more beneficial when $f = 0.1$ since the High Priority Traffic order parameter η_{HPT} reaches very large capacity $R_c = 46$ without affecting the low priority traffic capacity $R_c = 3$ which remains the same as in the traditional shortest path strategy without priority (WP). Moreover the values of the order parameter for LPT: η_{LPT} in the congested phase are remarkably close to those of WP meaning no noticeable degradation of the traffic of less important packets.

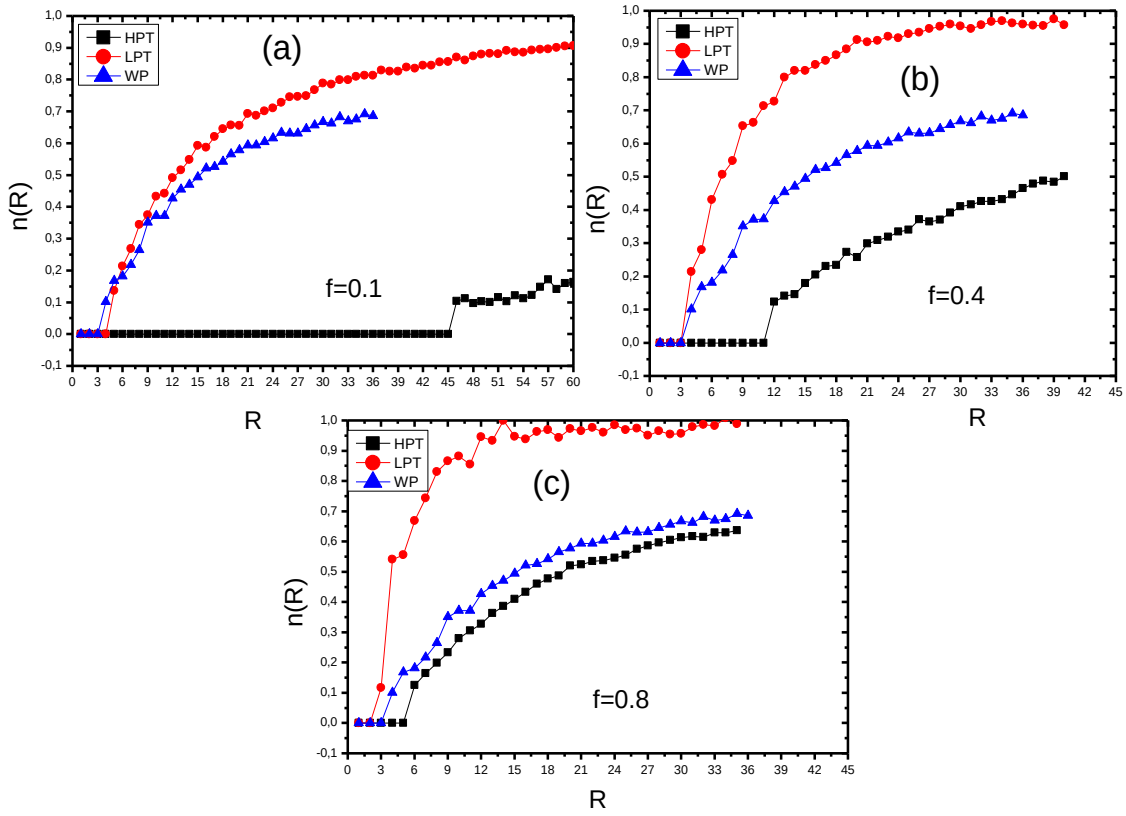


Figure III-11 : (a)-(c) The order parameter as a function of packets generation rate R for each type of traffic High priority traffic (HPT), Low Priority Traffic (LPT) , and shortest path without priority (WP) for different values of prioritized fraction from hubs nodes $f = 0.1$,$f = 0.4$,and $f = 0.8$. N=500, m=2 and <k>=4.

Fig.III.11-b, shows that for $f = 0.4$ the HPT reaches less capacity $R_c = 11$ compared to the case $f = 0.1$ and with a loss of performance of the LPT especially in the congested phase where the values of η_{LPT} increase rapidly with R and reach very high value compared to those of WP and their capacity remains the same: $R_c = 3$.

While in Fig.III.11-c for $f = 0.8$, the HPT reaches less capacity $R_c = 5$ without any noticeable improvement compared to the normal flow WP. In contrast, in the congested phase the value of η_{HPT} is almost close to the WP meaning more spread of congestion in HPT. Furthermore the LPT experiences a low performance both in free flow phase where the capacity decreases to $R_c = 2$ and the congested phase where the value of η_{LPT} increases towards very high values.

For the case where peripheral nodes are prioritized $\alpha = -1$, the result for the different order parameters η_{HPT} , η_{LPT} and η_{WP} are shown in Fig.III.12(a)-(c). The figure shows similarly the presence of the three regimes discussed in Fig.III.11, but the general remark is that the network capacity (measured by R_c) is less than the case where hubs are targeted.

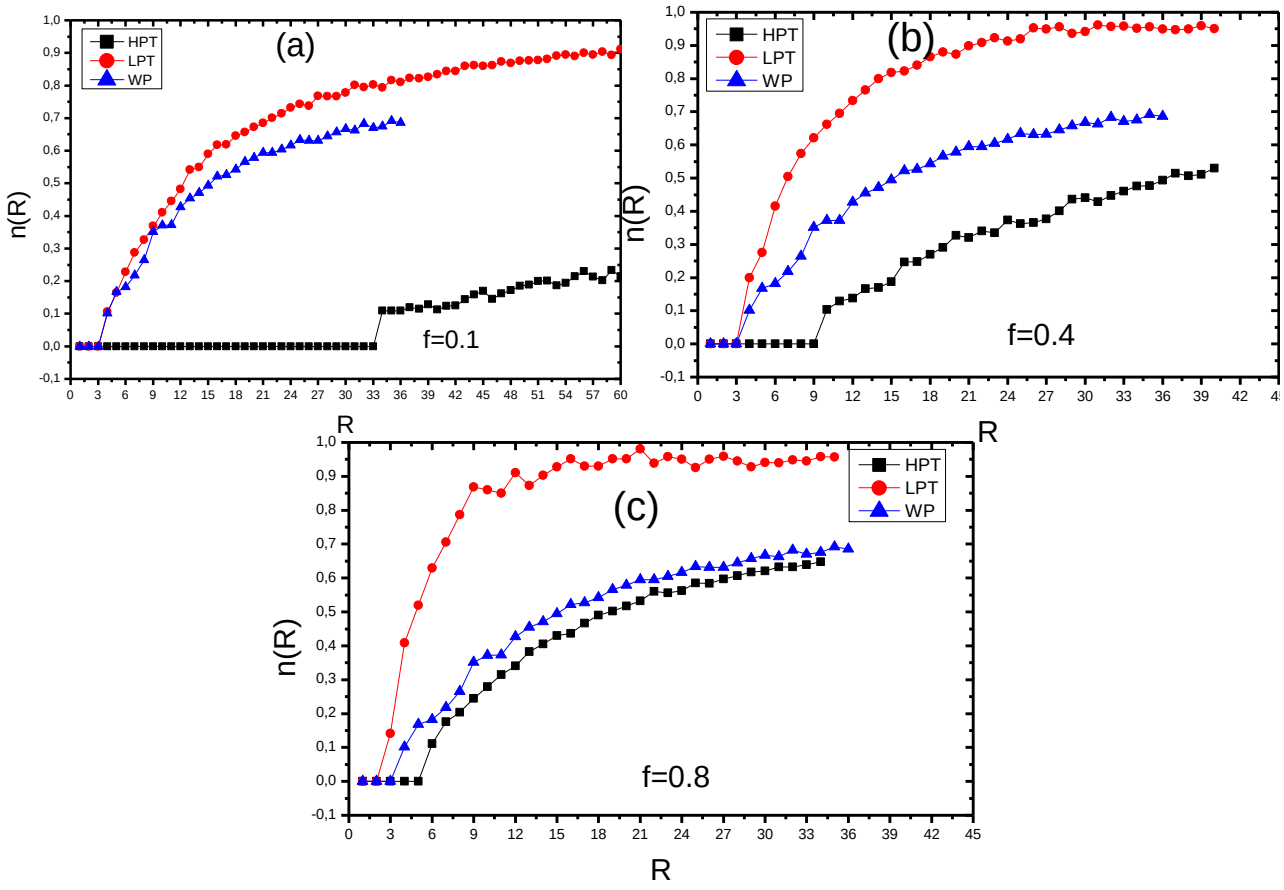


Figure III-12: (a)-(c) The order parameter as a function of packets generation rate R for each type of traffic High priority traffic (HPT), Low Priority Traffic (LPT) , and shortest path without priority (WP) for different values of prioritized fraction from peripheral nodes $f = 0.1$,$f = 0.4$,and $f = 0.8$. *N=500, m=2 and <k>=4.*

Fig.III.13(a)-(b) shows HPT and LPT capacity R_c when increasing prioritized fraction f. In Fig.III.13-a the different values of R_c of High Priority for different value of α. As we can see the prioritization of nodes with high degree $\alpha = 1$ is always more efficient and the random prioritization $\alpha = 0$ gave almost comparable capacity to the small degree nodes prioritization $\alpha = -1$. While in Fig.III.13-b the Low Priority Traffic still have the same capacity of the traditional shortest path strategy without priority (WP) $R_c = 3$ until $f = 0.8$ where its capacity decreases to $R_c = 2$.

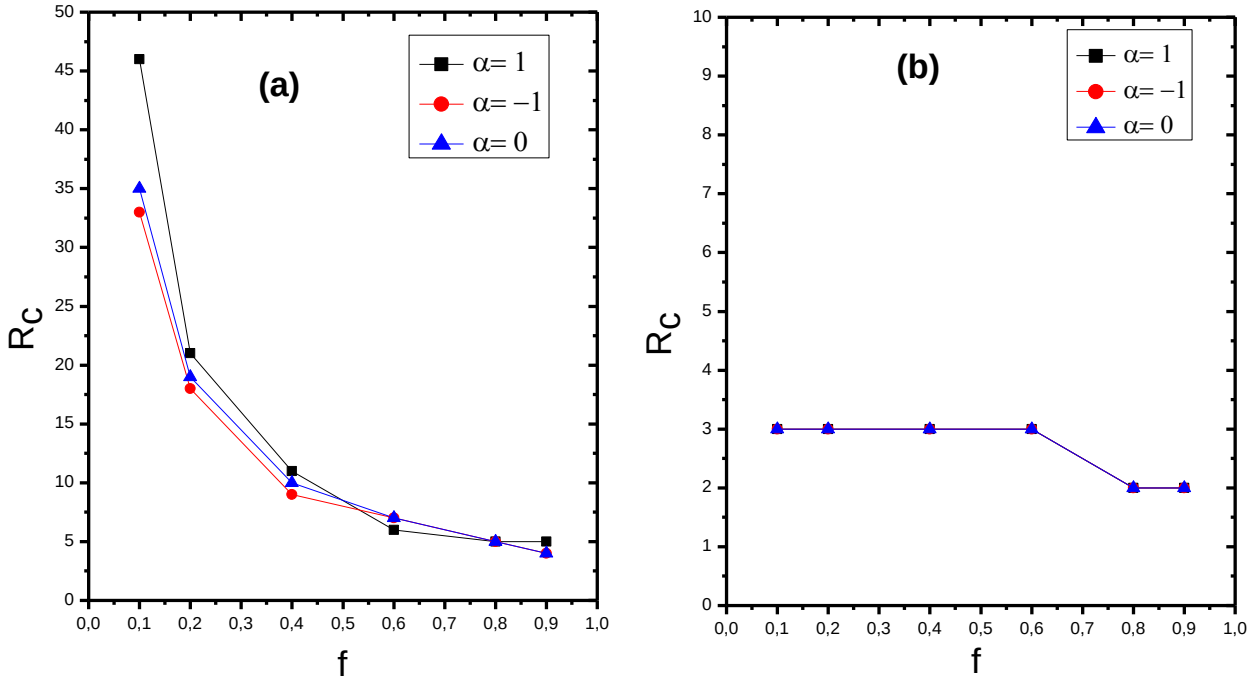


Figure III-13: (a)-(b) The order parameter as a function of prioritized fraction nodes f for each type of traffic High priority traffic (HPT), Low Priority Traffic (LPT), for different values of α . α=1 Hub are prioritized, α=-1 Peripheral nodes are prioritized, and $\alpha = 0$ for the random prioritization. $N = 500, m = 2$ and $< k >= 4$.

III.3.2.2 The average traveling time

Another important quantity that measures network traffic efficiency is the average packet traveling time $< T >$ that is the time spent by a packet between its source and its destination. Similarly we define the traveling time for each traffic type: $< T >_{HPT}$ for HPT, $< T >_{LPT}$ for LPT, and $< T >_{WP}$ the normal traveling time in SP where no prioritization is used.

Fig.III.14 (a)-(c), shows $< T >_{HPT}$ and $< T >_{LPT}$ traveling time when increasing prioritized fraction f from hubs nodes. We find as in case of traffic capacity that the prioritization of hubs nodes is more beneficial when $f = 0.1$ since the High Priority Traffic traveling time $< T >_{HPT}$ reaches very low value of time without affecting the low priority traffic traveling time which remains the same as in the traditional shortest path strategy without priority (WP).

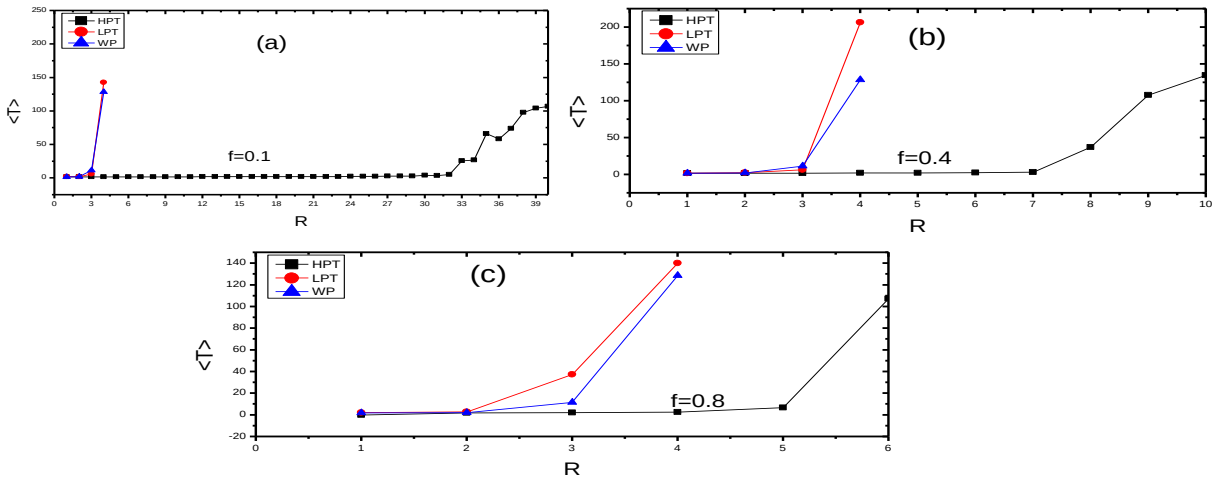


Figure III-14: (a)-(c) The traveling time as a function of packets generation rate R for each type of traffic High priority traffic (HPT), Low Priority Traffic (LPT) , and shortest path without priority (WP) for different values of prioritized fraction from hubs nodes f $= 0.1$,f $= 0.4$,and f $= 0.8$. N=500, m=2 and <k>=4.

In summary, we can say that we have three important regimes of traffic when our priority policy is applied:

- Gain with no loss regime: for lower prioritized fraction f; where the HPT benefits from a higher capacity and lower traveling time without affecting the capacity and the traveling time of less prioritized traffic.
- Gain with some loss: for intermediate fraction f,. In this regime, HPT still benefits from capacity increase and lower traveling time but this time at the expense of a degradation of LPT traffic performance.
- No gain with loss: for high prioritized fraction f, . In this regime, the HPT traffic does not gain substantial performance and the prioritization policy results only on lowing the performance of the LPT traffic.

III.4 Conclusion:

For the purpose of alleviating the congestion in traditional shortest path strategy and to deal more with packets transport on real protocol in internet, in this chapter we have proposed a priority policy based on packet destination, it's found that for maximizing the High Traffic capacity the traffic coming into nodes with high degree should be prioritized and the fraction of prioritized nodes shouldn't be very high.

This result is very useful for traffic engineers who try to implement a traffic prioritizing policy and at the same time want to be sure that their systems operate as far as possible in the first regime or in the second one in the worst case because they don't want the non prioritized traffic to be impacted. Furthermore based on our prioritization model the companies in the internet can prioritize the traffic coming in or going to specific servers.

However enhancing the network capacity is not the only main goal of routing designers and complex network researches; it worth nothing to alleviate the congestion spreading in network without be able to protecting packet information from computer virus infection. In this context, in the next chapter we will introduce and explain the virus propagation and its modeling.

Chapter IV. Introduction to computer virus and propagation modeling

IV.1 Introduction:

In comparison to other epidemics in nature [152], computer viruses are much easier to spread over the huge Internet, and have much longer lifetimes [123].

Currently, due to the great expansion of computers and the huge growth of internet, computer virus, has a very large number of potential targets to affect an impressive number of people.

Therefore, it is not uncommon today to see a virus swept across the planet via the network in a few days, even a few hours. Viruses have become highly publicized and the attacks are of ever greater magnitude. More recently, we have also begun to witness viruses that can spread on social networks. These viruses spread by infecting the accounts of users of the social network. By clicking on any option, it can run the virus and share the capabilities of the user. This results in the spread of these malicious programs without the knowledge of the user.

However, if the general public begins to know viruses terms, and to be sensitized, the level of global knowledge about viruses remains low for the greatest number.

This fact has created several challenges, mainly network security. A reliable defense system is therefore necessary to safeguard both the valuable information stored on a system and the information in transit. To achieve this goal, it becomes imperative to understand and study the nature of various forms of malicious entities. It also becomes necessary to understand how these propagate through computer networks.

To better understand this phenomenon, in this chapter, we propose to explain the different types of existing viruses and their actions, and we will explain how an anti-virus works.

IV.1.1 Defintion:

Computer viruses or digital viruses generally refer to instruction or series of parasitic instructions, introduced in a program that can spread by making copies of itself and likely to cause various disturbances in the computer, disrupt the normal operation, and cause damage to data and programs.

However digital viruses are similar to biological viruses. A biological virus enters a body, damages the body, spreads to other bodies, and eventually is eradicated by the internal immune system or by external means. Similarly, the digital virus enters the computer or phone system and gets attached with a program (or set of programs or applications). As the application(s) is invoked, the virus becomes activated and spreads to the other parts of the system. Then it continuously grows and spread out thereby damaging or even completely destroying the regular functioning of the device and their programs. Like in Figure IV. 1, viruses can be carried by Multimedia Messaging Services (MMS) or Bluetooth services from the infected phone (red color) to other susceptible phones.

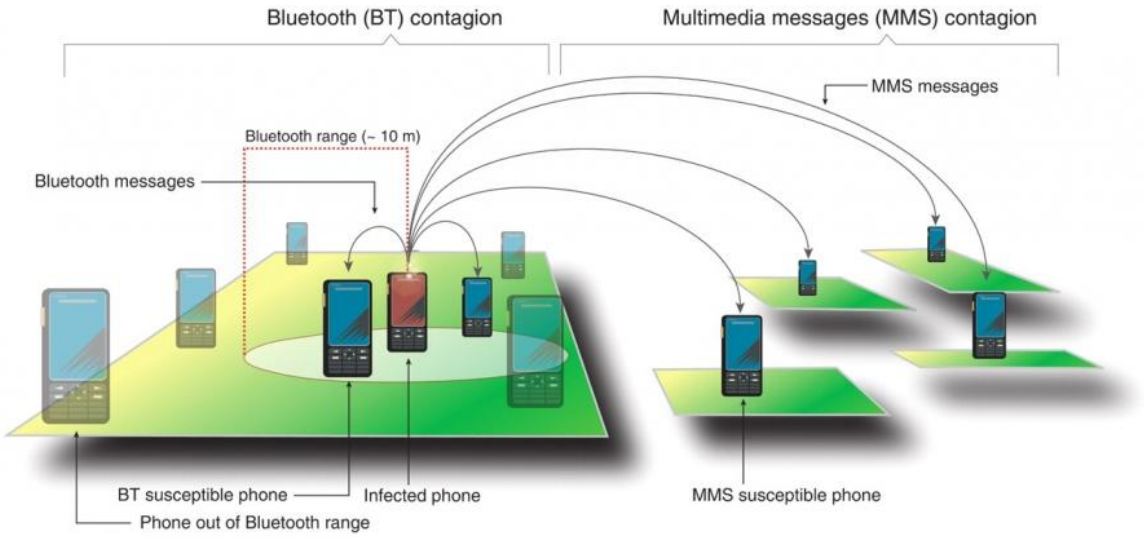


Figure IV-1 : Mobile phone viruses

IV.1.2 Virus propagation

As a technical term invented by Cohen [124], computer viruses can spread between computers by self-replication and, during their hatching periods, they can perform detrimental operations, even destroy the entire system computer.

There are many media for spreading viruses, especially with the explosion of the Internet over the last decade, which provides the largest circulation highway to the latter. Initially, removable media were the means of spreading viruses. Floppies at first, then burned CDs, external hard drives, and why not today USB sticks. If the virus is in a file that someone wants to transfer to another computer through removable media, it can infect the

destination computer. Domestic local networks, and even more so, business networks are also a major vector of propagation. Computers are all connected to each other, making it easier for the virus to spread to all machines.

Regarding the Internet, emails containing the virus attachments are a classic. By several more or less subtle systems, the designer of the virus makes sure to push the user who receives the email to execute the attachment to infect the computer. The virus can manage to send itself to all the people in the address book of the first victim and so on. And spread exponentially. Finally, there is propagation due to downloading, either directly by uploading an infected file to a site, or on peer-to-peer networks.

However, computer viruses follow a life cycle that starts when they're created and ends when they're completely eradicated. The Figure IV. 2 resumes virus life cycle.

Figure IV-2: Life Cycle of Computer Virus.

Stage 1 - Creation – In this period a virus as ferocious as possible are developed by misguided individuals who wish to cause widespread, random damage to computers. The programming is generally done in assembly code or Visual Basic or sometimes in C or C ++.

Stage 2 -Replication - As mentioned, the virus must reproduce. A properly crafted virus will reproduce a significant number of times before activating. This is the best way to ensure the durability of a virus.

Stage 3 -Activation - Viruses, which have a destructive routine (portions of code intended to cause damage to the host), activate only when certain conditions are met. Some are activated in a specific date (set by the developer), others have an internal countdown system. Activation can also take place remotely, by the developer. Even viruses, which do not have such routines and do not require a specific activation procedure, can cause system damage by gradually appropriating all resources.

Stage 4 -Discovery – This phase doesn't always come after activation, but it usually does. This is the moment when the user notices, that his system shows strange behaviors and suspects the presence of viruses, or the powerful anti-virus discover some viruses before they have had time to wreak havoc.

Stages 5 -Assimilation - Once the discovery is made, developers of anti-virus software update their viral database, so that users can detect the presence of viruses on their computer. They also develop the patch (or antidote) to eradicate the virus (if possible).

Stage 6 -Eradication - This is the death of the virus. At least, it is the death of the copy of the virus on a user station. This is the moment when the anti-virus, having discovered the virus, proposes to the user to remove it. If enough users install up-to-date virus protection software, any virus can be wiped out. So far no viruses have disappeared completely, but some have long ceased to be a major threat.

IV.1.3 Anti-virus or Immunization

Immunization or antivirus strategies specify how vaccines are distributed in network, in the field of computer viruses; long-term vaccination is often used to describe the process of installing the latest version of the antivirus software into an uninfected computer. As a result, a newly vaccinated computer may acquire temporary immunity.

Generally, anti-virus software can detect all types of known viruses, and it needs to be updated frequently for its effectiveness. Basically, there are four means of virus detection: signature based scanning, emulation, heuristics, behavioral analysis and check summing.

Signature-based detection: This scheme searches for unique strings of code, i.e., the virus's signature specific to particular viruses. The signature could represent a series of bytes in the file. It could also be a cryptographic hash of the file or its sections.

Heuristics-based detection: aims at generically detecting new malware by statically examining files for suspicious characteristics without an exact signature match. Heuristic method often employs generalized signature scanning geared to detect families of viruses. If the virus is related to a known family, heuristics will detect it and report it as suspicious or infected with an unknown virus. The biggest down-side of heuristics is it can inadvertently flag legitimate files as malicious.

Behavioral –based detection: This approach attempts to identify malware by looking for suspicious behaviors. For example, if a file attempts to write to the system registry, modifying the hosts file or observing keystrokes, the action can be blocked, either by the user or automatically, depending on configuration. Noticing such actions allows an antivirus tool to detect the presence of previously unseen malware on the protected system.

IV.2 Virus propagation modeling

Due to the high similarity between computer virus and biological virus [123], Cohen [124] and Murray [125] have proposed exploiting the techniques developed in the dynamics of the biological epidemic to study the laws governing the spread of viruses. From there, many multiple epidemic models of computer viruses, ranging from classical models, such as SIS models [126,127], SIR models [128], SIRS models [129],models SEIR [131], SEIRS models [132], SEIQRS models [133], SLBS models [134,135], SICS models [136], and some other models [137-140], to unconventional models such as delayed models [141-142], impulse models [134,144], and models stochastic [147], have been proposed.

These models are used in different fields such as medicine; spread of diseases within a population, computer science; spread of viruses, and on any type of network (random, small world, without scale).

IV.2.1 **Susceptible** Infected Model (SI):

It is a simple propagation model where one considers that the number of nodes studied N can be decomposed into two categories:

- Nodes susceptible to be infected (S);
- The infected nodes (I).

The infection spread by direct contact (via a link) of a susceptible S with one of the infected ones I. So, the susceptible S become infected I with a factor of probability β (also called infection rate) see Figure IV.3. Where β expresses that contacts do not necessarily lead to virus spreading, a contact not necessarily leading to infection.

A node, when infected, remains in this status until the end of his life.

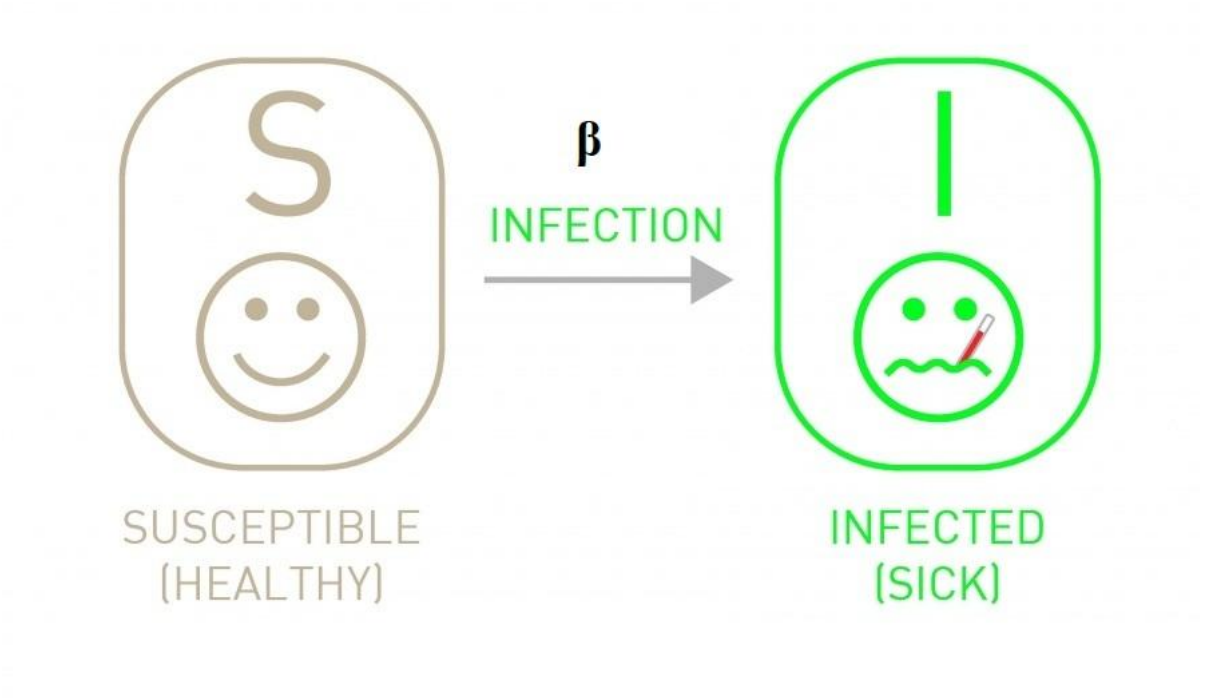


Figure IV-3: Transmission diagram of the SI model.

This hypothesis is reasonable for many viruses in the early stages of infection. It is assumed that the number of nodes is fixed, then:

$$S + I = N$$

Where N is the size of the network.

IV.2.2 Susceptible Infected Recovered (SIR) Model:

This model [128,148] (see Figure IV.4) divides the network into three groups, the susceptible S, the infected I and the repaired or recovered R.

Susceptible S: all nodes in this group can become infected if they come into contact with an infected node.

Infected I: all nodes in this group carry viruses to others. They spread the infection.

Vaccinated or recovered R: the nodes to which the antivirus has treated the infection.

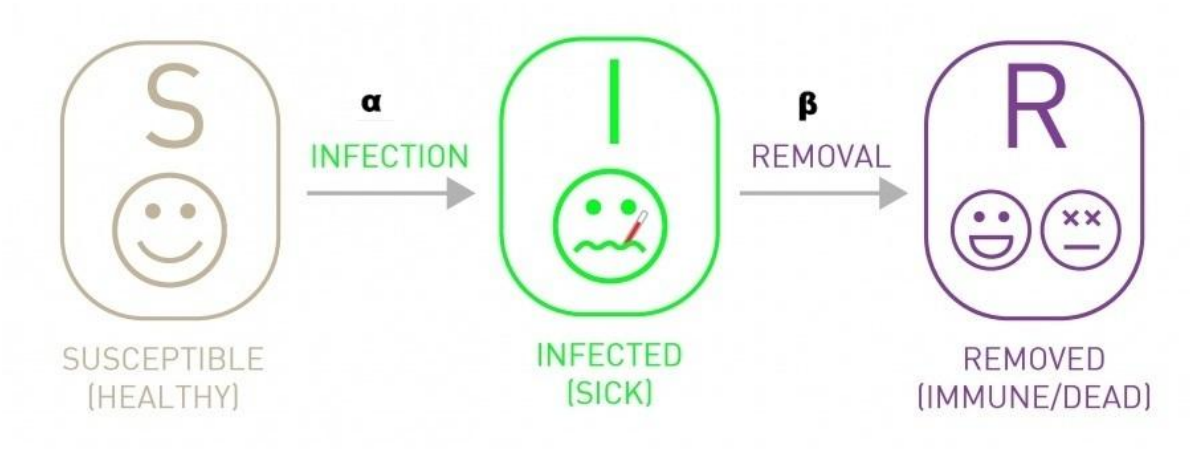


Figure IV-4: SIR Model

The model contains only two parameters: α the infection rate and β the repair rate. The size of the network is assumed to be constant:

$$N = S + I + R$$

α represents the infection rate, the probability of transmission of viruses between an infected person and people likely to be infected.

β represents the immunized rate. The epidemic establishes a chain of transmission: Infected nodes or computer will infect susceptible nodes before being corrected by an antivirus.

IV.3 Conclusion:

We introduced here some models of virus propagation. The probabilistic models are a natural way of modeling the evolution of a virus: each node has a certain probability of being infected by a virus (without consider its type). Here we have introduced the following models: SI which is the simplest model which represents the basis for introducing other models. SIR which classifies the nodes of the network into three categories namely those susceptible to infection, infected and repaired.

Furthermore, in reality, the spread of computer viruses in internet is motivated by routing strategies. The next chapter, we will propose a new study about the computer virus spreading under specified routing strategies; we will use the SI model to measure the efficiency and the robustness of routing strategies in term of computer virus propagation.

Chapter V. Local routing protocols performance

V.1 Introduction

Complex networks have come to penetrate many aspects of our lives, such as the internet, World Wide Web (WWW) and social networks. The internet is the infrastructure of transferring information packets which are the basic objects of various and useful network applications such as the Email, and online social activities. Therefore in recent years, great amounts of research have investigated the structure of such complex networks [5, 26, 36, 61,153,145,154].

However, epidemic spreading is another dynamical process in internet network that have been studied extensively. The focus of previous works on spreading was mainly on how the network topology affects the epidemics [155, 156]. While in the reality the virus propagation is motivated by the packets routing between computers in complex networks [146,151].

In a recent work, Meloni et al. [157], attempt to incorporate traffic dynamics in epidemic spreading, and proposed another traffic-driven epidemic spreading model; in particular, they integrate the susceptible-infected-susceptible (SIS) model [158] into the shortest path routing protocol, the epidemic can spread between nodes by the transmission of packets information.

At each time step a susceptible node will probably be infected if it receives new packets from its infected neighbor's nodes. Yang et al. in [159, 160] tried to alleviate virus epidemic propagation by changing the optimal routing path in static local and global routing strategies which badly affects the network capacity. While In [161], the authors suppress the traffic-driven epidemic spreading by cutting off some edges in the network.

Furthermore, O. Bamaarouf et al. [162] showed that the global routing algorithms: Efficient path [77] and global dynamics [72], favor the virus spreading more than the case where the shortest path algorithm is used and proposed a new vaccination methods to eradicate the virus propagation in these performant algorithms.

However, due to the low transmission capacity of the local routing strategies, this category of protocols has not received a great attention for a while, until recently Lin B. et al. [163] proposed a restrictive queue length algorithm (RQL) based on the idea of next-nearest

neighbors (NNN) [164]. Local strategies with RQL can reach very high capacity and overcome the global routing strategies under certain circumstances.

In this chapter, we propose a new study about the virus propagation in local routing protocol under the additional Thiers algorithms [193]. Our results have shown that the virus propagation in local protocols without additional algorithms (NNN and RQL) is very fast, due to the high roundabout of packets. While in local routing protocol under the additional algorithms (NNN and RQL) the virus propagation is reduced remarkably.

Moreover in comparison with global routing protocols, the local routing protocols under RQL surprisingly overcome the efficient path strategy regarding computer virus propagation.

V.2 Models and routing strategies

V.2.1 Network types

We adopt the Barabasi-Albert (BA) network model [26] with a degree distribution following the power-law distribution $P(k) \sim k^{-\gamma}$, with $\gamma \sim 3$ as explained in chapter I.

V.2.2 The traffic model:

Every node is considered as both host and router which can either generate packets or forward packets on the network. At each time step, R packets are generated in the network with randomly chosen sources and destinations. Every node can deliver at most C (here, we set $C = 1$) packets to its immediate neighbors based on a chosen routing protocol [74, 75, 77, 79]. Packets are sent either by first-in-first-out (FIFO) procedure or via a priority policy described in the following subsection. The packet is removed immediately from the system once it arrived at its destination.

V.2.3 Epidemic model:

After a transient time of traffic, an initial fraction of nodes f_0 is randomly set to be infected (e.g., we set $f_0 = 10\%.$) we adopt the Susceptible-Infected (SI) model to see how virus

propagates in the network with the different routing protocols. The infection spreads in the network through packet navigation. All packets in an infected node are also infected, while in a susceptible node, all packets are uninfected. At every time step a susceptible node has the probability β of being infected, if it receives an infected packet from infected neighbor nodes.

V.2.4 The routing strategies:

Local static routing protocol: in this strategy [74], the packet is forwarded to a neighbor node according to the probability:

$$P_i = \frac{k_i^{\alpha}}{\sum_{j=1}^{N} k_j^{\alpha}}$$

Where the sum runs over the neighbors of node i and α is an adjustable parameter. The optimal value of $\alpha = -1$.

Local dynamic routing strategy: in this strategy [75], the local search is based on both neighbor's degree k_i and queue length q_i:

$$\mathrm{P_i} = \frac{k_i(1+q_i)^{\beta}}{\sum_j k_j(1+q_j)^{\beta}}$$

Where the sum runs over the neighbors of node i and β is an adjustable parameter. The optimal value of $\beta = -3$.

Global efficient path routing strategy: in this global strategy[77], the global information is needed where the path between sources and destination is determined as follow:

$$\mathrm{P_{i\to j}} = \min \sum_{e=0}^{L} \mathrm{k_e}^{\theta}$$

Where e represent the node on the path between i and j and θ is an adjustable parameter. The optimal value of $\theta = 1$.

Global shortest path routing strategy: this global strategy [79] is a particular case of the efficient path strategy when $\theta = 0$.

Each strategy has been well explained and described earlier in chapter II.

V.3 Results and Analysis:

We adopt a network with size $N = 1000$ and average degree $< k >= 10$.We assume that the network is in the free flow state by fixing the packet generation rate as $R = 3$. All the global and local routing protocols are realized with the optimal adjustable parameters. The model SI is implemented in all our simulations.

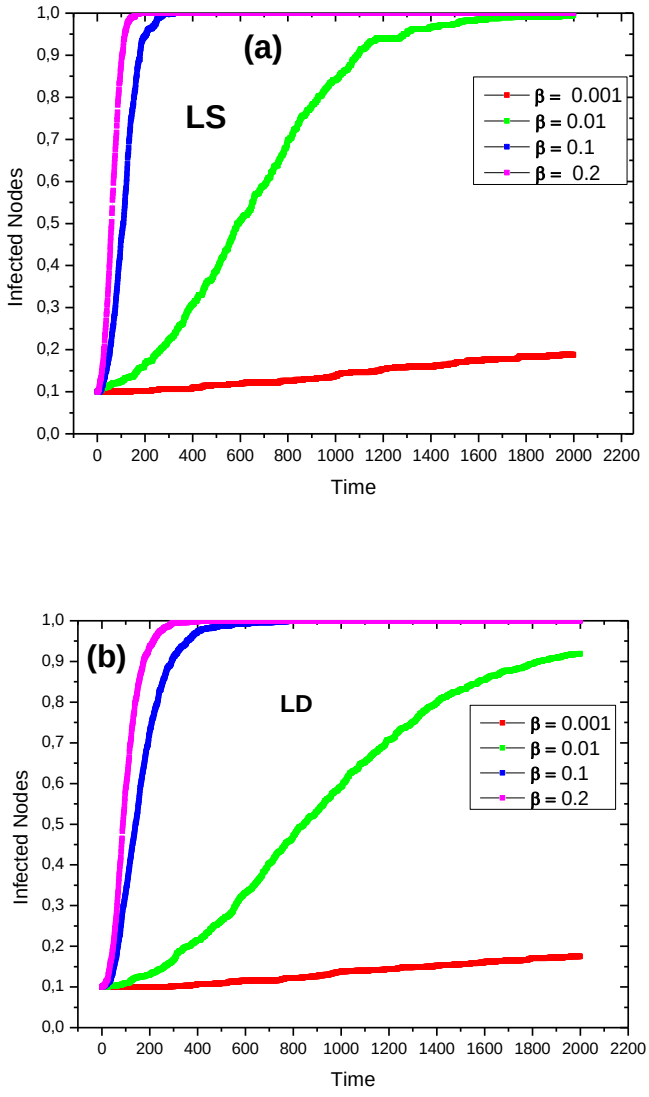


Figure V-1: the proportion of infected nodes in the whole network for different values of probability infection as a function of time in the free flow phase for a fixed packet generating rate R=3. With an initial proportion of infected nodes f_0=10% in the SI model using (a) Local static routing protocol, (b) Local dynamic routing protocol. $N = 1000$, $< k >= 10$.

Figure V.1 (a)-(b) show the infected nodes as function of time for different values of infection probability β under respectively local static routing (LS) and local dynamic (LD)

routing protocols. One can see that the computer virus propagation is very fast and increase sharply with time indicating a high sensitivity to computer virus. This is due to the roundabout and the blind transmission of the packet in local routing protocols; because the packet is transmitted to the neighbor blindly and without taking in consideration that packet could be near to its destination which could be just at two steps far from the packet source (next neighbor). This results in longer path length ($L = 171.82$ for LS and $L = 120.08$ for LD) and more computer virus spreading in the whole network.

In order to limit the roundabout and reduce packet traveling path length in the local protocols, we applied the additional next-nearest-neighbor (NNN) algorithm.

Next-nearest-neighbor (NNN) algorithm:

- When a node receives a new packet, it first searches the destination among its neighbors. If it exists, the packet will be delivered to it.
- Otherwise, the node searches the destination from its next-nearest-neighbors. if it is found the packet will be delivered to the neighbor linked to it.
- If the packet fails to transmit after the first two steps, it will be transmitted based on a chosen local routing protocol (static or dynamic).

When the additional next-nearest-neighbors (NNN) is applied to the local routing protocols, the computer virus spreading is remarkably reduced as shown in Figure V.2 (a)-(b); the computer virus is reduced by the searching of the packet destination among the next neighboring nodes. This mechanism means more traffic load on hubs, because nodes with high degree are more likely to be a neighbor of the searched destination. Then the majority of packets will reach their destination with the minimum jump on high degree nodes minimizing the path length. Moreover it is clear that the virus spreading in LS with NNN (LSNNN) is faster than in LD protocol under NNN (LDNNN).

This can be explained as follow: In LDNNN the nodes search first the neighbor linked with the packet destination or the neighbor with minimum queue length which is leading to small value of path length $L = 4.69$. While in LSNNN the node searches first the neighbor

linked with the packet destination or the neighbor with small degree leading to a longer path length $L = 11.19$. and traveling time.

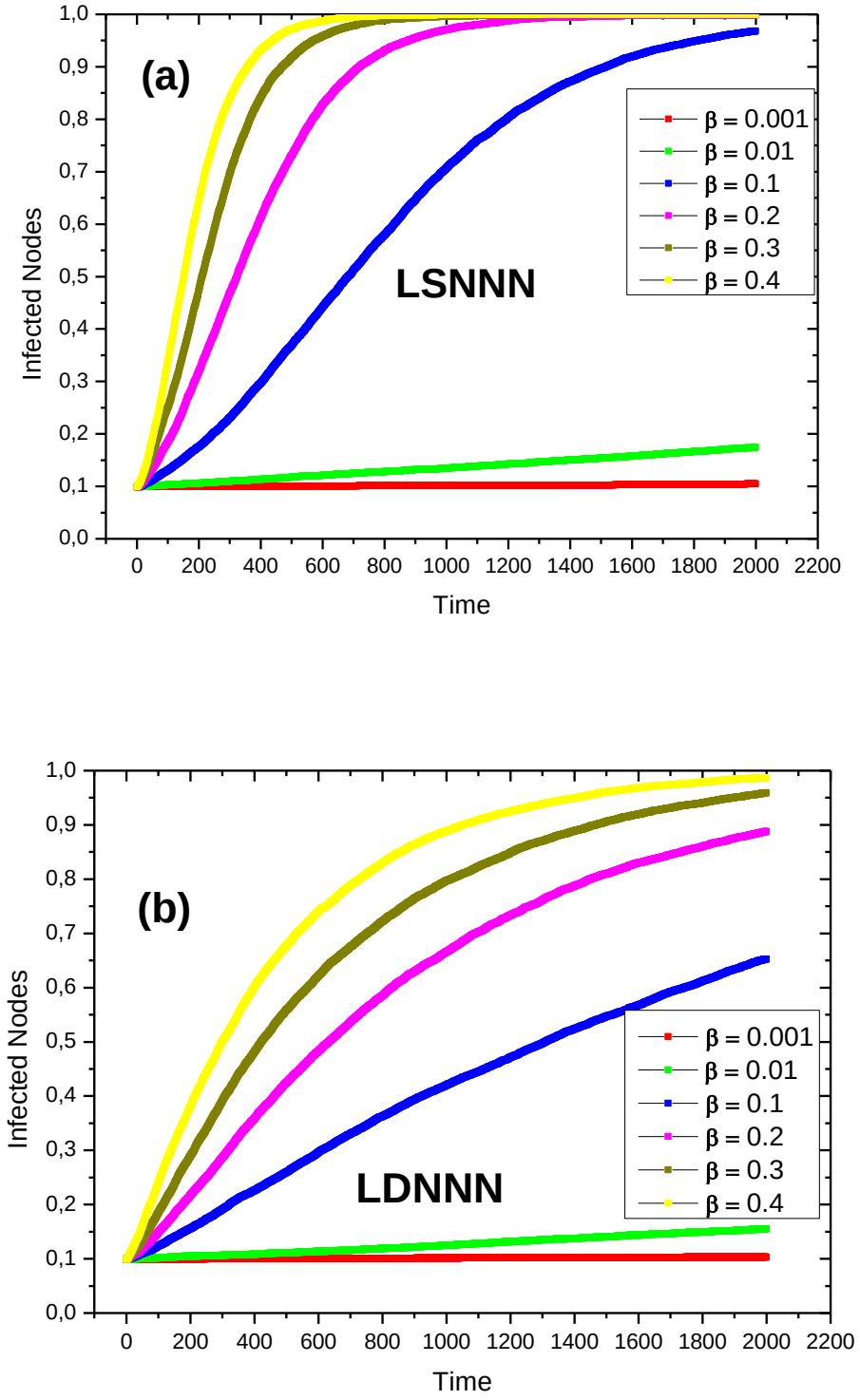


Figure V-2: the proportion of infected nodes in the whole network for different values of probability infection β as a function of time in the free-flow phase for a fixed R=3, with an initial proportion of infected nodes f0=10% in the SI model using the addition additional Next-Nearest- Neighbor algorithm in (a) Local static routing protocol, (b) Local dynamic routing protocol. N = 1000, $< k >= 10$.

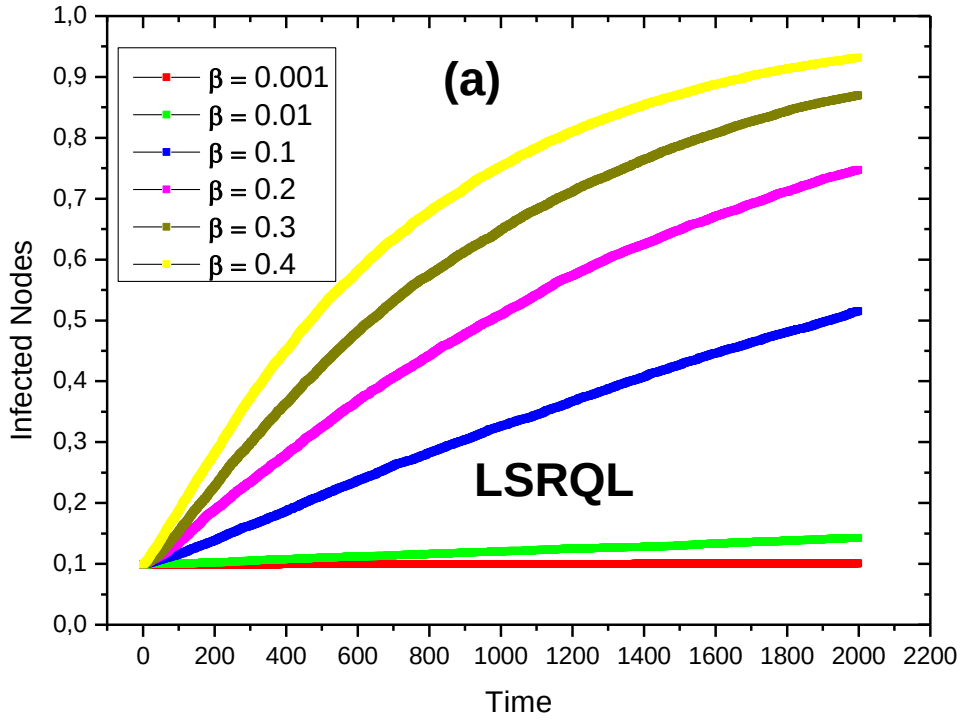
Despite the fact that, the virus propagation is reduced in local routing protocols with the NNN additional algorithm, the nodes with high load are extensively used; the probability of hub nodes to be infected increases with time. It is clear that the more hubs are infected the more packets are infected in the network and the computer virus propagation becomes very high. To overcome this anomaly, we use the additional restrictive queue length algorithm (RQL) proposed in [163] first without Node duplication avoidance (NDA).

RQL algorithm without Node duplication avoidance (NDA):

- When a node receives a new packet, it first searches the destination among its neighbors. If there exists such a node, the packet will be delivered to it.
- Otherwise, we set a threshold $s + t$. If a neighbor's queue length does not exceed this value, its next-nearest-neighbors will be searched. Once the packet's destination is found, the information packet is delivered to the neighbor linked with it.
- If the packet fails to transmit after the previous steps, we set a lower threshold s. Among the neighbors with the queue length being not more than s, the information packet is delivered based on the local routing protocol chosen.
- If all previous steps are not satisfied, the packet is delivered based on the chosen local routing protocol.

In Figure V.3 (a)-(b) the infected nodes percentage in respectively local static (LSRQL), and dynamic (LDRQL) routing protocols under the RQL algorithm, is plotted as a function of time. Although the virus spreads in the network over time, it takes lower values compared to those found for the NNN (see Figure V.2). This can be interpreted as a result of the restriction queue length imposed in the RQL algorithm; the packets are forced to avoid the nodes with high load as much as possible.

We can think that such hub avoidance may result in large path length and thus resulting more virus propagation contrary to the virus reduction observed. But this apparent contradiction can be resolved by hypothesizing that the packets visit practically the same few nodes more than new once around the destination.



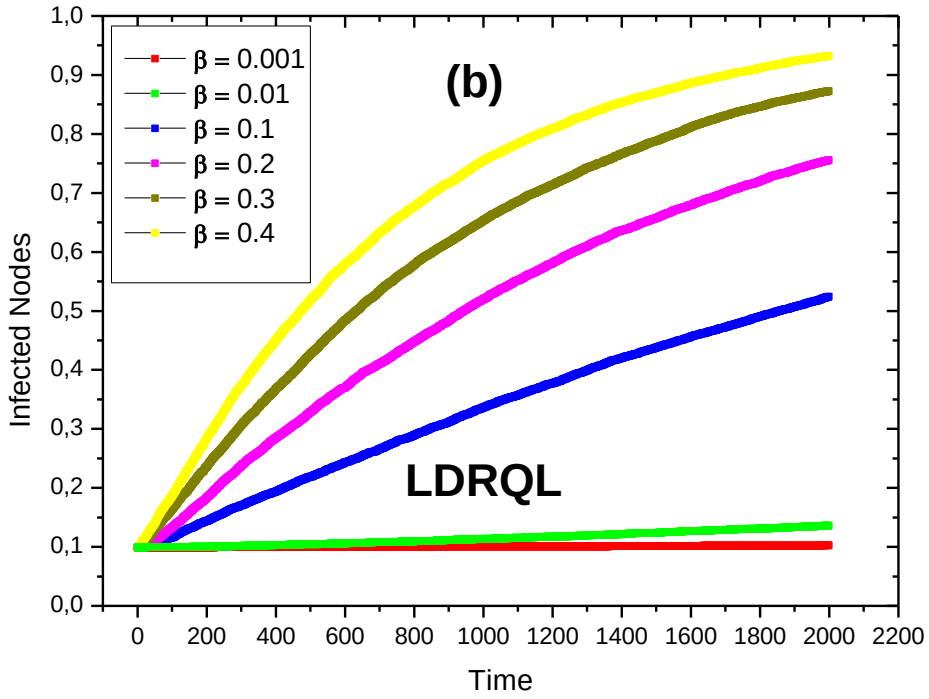


Figure V-3: the proportion of infected nodes in the whole network for different values of probability infection β as a function of time in the free-flow phase for a fixed R=3, with an initial proportion of infected nodes f_0=10% in the SI model using the additional Restrictive- Queue-Length algorithm in (a) Local static routing protocol, (b) Local dynamic routing protocol. $N = 1000, < k >= 10$.

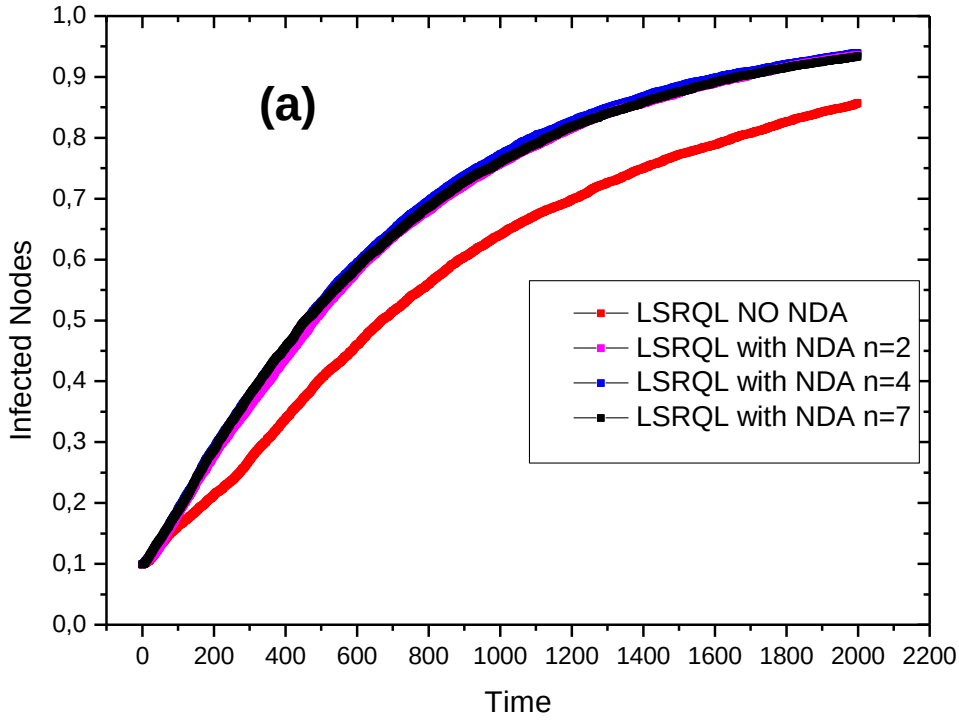
To prove that the same packets visit the few same nodes several times in local routing protocols under RQL algorithm, we implement this time the RQL with the node duplication avoidance (NDA) algorithm as proposed in [163]. This algorithm tries to alleviate the roundabout of the packet on the same nodes by memorizing the last n visited nodes and not allow the packets to revisit those nodes in the next packet delivery.

RQL algorithm with Node duplication avoidance (NDA):

- When a node receives a new packet, it first searches the destination among its neighbors. If there exists such a node, the packet will be delivered to it.
- Otherwise, we set a threshold $s + t$. If a neighbor's queue length does not exceed this value, its next-nearest-neighbors will be searched. Once the packet's destination is found, the information packet is delivered to the neighbor linked with it.
- If the packet fails to transmit after the previous steps, we set a lower threshold s. Among the neighbors with the queue length being not more than s, we record the latest n nodes that the information packet has visited so that the packet would not be passed on to these nodes in the next transmission. In the unrecorded neighbors, information packets choose their next propagation node depending on the local routing strategy.
- If all previous steps are not satisfied, the packet is delivered based on the chosen local routing protocol.

The performance of this algorithm may depend strongly on the structure of the network especially the average degree $< k >$. For this purpose, we will use in this part, two types of networks: a dense network with average degree $< k >= 10$ and a sparse one with $< k >= 4$.

We compute the packets traveling path length in local routing protocols under RQL with NDA. From Table V-1, one can see that for the dense network with average degree $< k >= 10$, the path length value decreases from 3.97 to 3.68 when memorizing the last 4 nodes. However, the effect of NDA algorithm is much clearer under the sparse network; the path length decreases clearly from 48.2 to 26.02 which proves our interpretation.



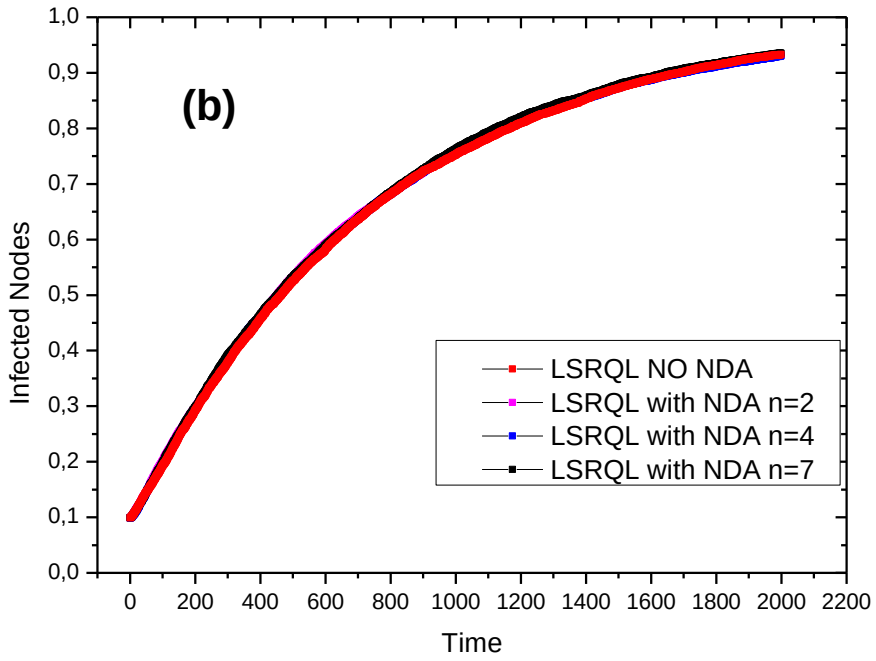


Figure V-4: the proportion of infected nodes in the whole network for fixed probability infection β=0.4, as a function of time in the free-flow phase for a fixed R=3, with an initial proportion of infected nodes f_0=10% in the SI model using the additional Restri Restrictive-Queue-Length algorithm with Node-duplication-avoidance algorithm in local static routing protocol (a) sparse network $< k >= 4$, (b) dense network $< k >= 10$. N = 1000.

Moreover In Figure V.4 (a)-(b), we plot the percentage of infected nodes as a function of time in local static protocol under RQL with NDA for different value of n last visited nodes ($n = 2,4$). One can see that with NDA algorithm the virus spreading increases in the sparse network (Figure V.4 (a)) when increasing the number of memorized nodes. While in dense scale free network (Figure V.4 (b)) the NDA seems to be without any influence on virus propagation, due to small effect on path length as explained above.

After the virus spreading in local respectively static and dynamic routing protocols with and without NNN and RQL algorithms is investigated, we proceed to the comparison between local with RQL and the global static routing protocols in term of virus spreading. Figure V.5 (a)-(b) show the virus spreading in Shortest Path (SP) and Efficient Path (EP) protocols.

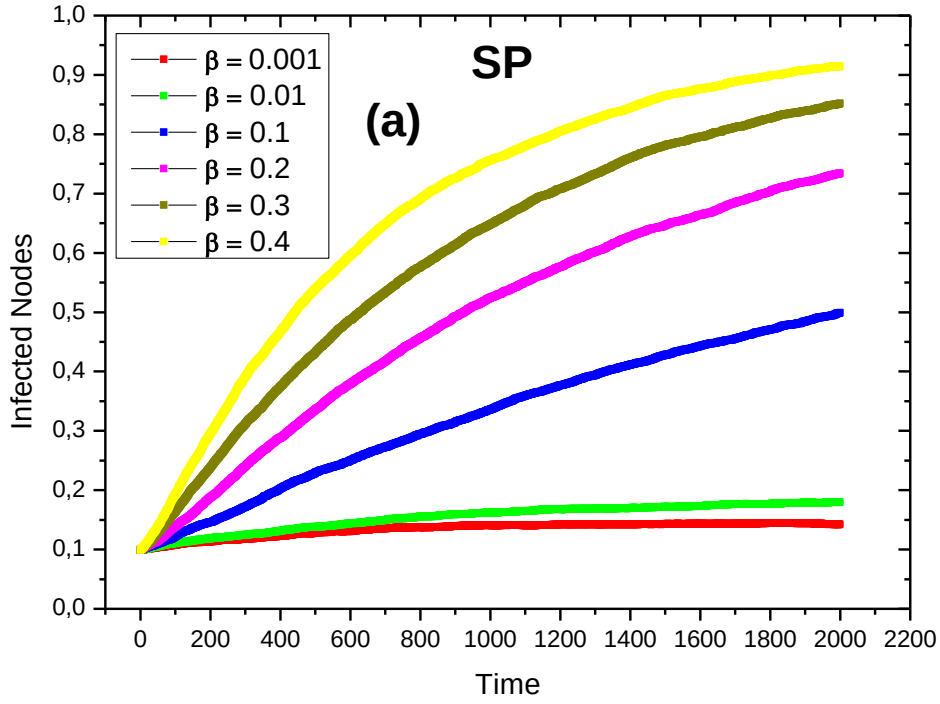
We found that the values of infected nodes in local static, dynamic routing protocols with RQL (Figure V.3) are comparable to those values when the shortest path strategy is adopted. This can be explained by the closeness of path length between the two strategies as can be seen in Table V-1. Furthermore, we found that the local protocols with RQL overcome the EP strategy in terms of security against the virus spreading.

The reason is that the efficient path strategy searches to avoid the hubs by passing on nodes with small degree to alleviate the congestion in the network; so the efficient path will pass most the time on different nodes with small degree which leads to more jump on new nodes accelerating virus spreading.

While in local protocols with RQL, the packets can pass on the same nodes several times to search a path to its destination leading to less contact with new nodes and the virus spreads just on a few nodes in the whole network.

Routing Protocol	LSRQL		LSRQL-NDA $n = 2$		LSRQL-NDA $n = 2$		SP	EP
Average degree $< k >$	10	4	10	4	10	4	10	10
Traveling Path Length (L)	3.97	48.24	3.93	26.02	3.68	25.5	2.97	4.05

Table V-1 : The average path length under LSRQL, SP and EF routing strategies in dense and sparse network



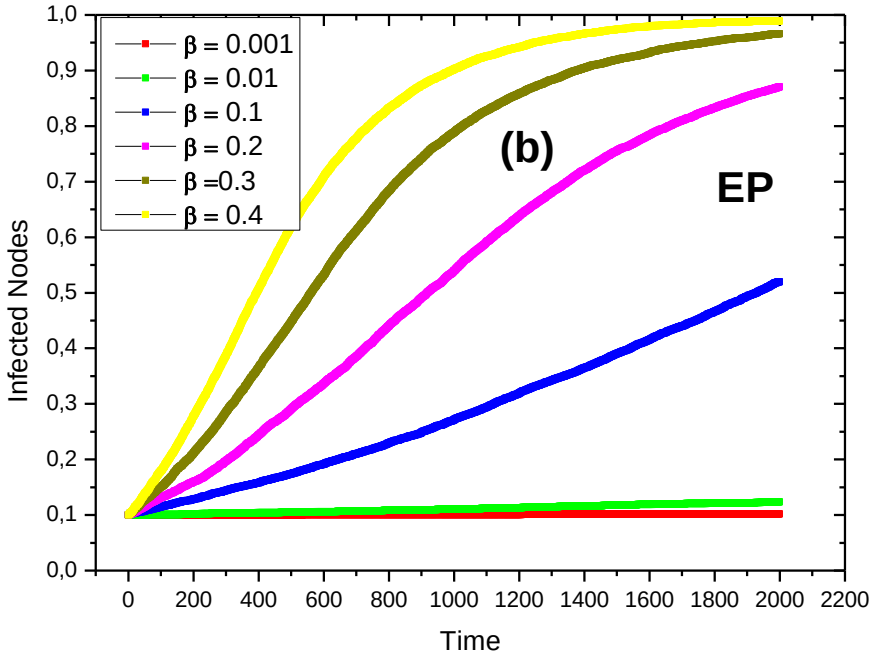


Figure V-5: the proportion of infected nodes in the whole network for different values of probability infection β as a function of time in the free-flow phase for a fixed R=3, with an initial proportion of infected nodes f_0=10% in the SI model using (a) Shortest path, (b) Efficient path. N = 1000, $< k >= 10$

Finally and in order to study the effect of network size on the robustness of ours results. We have computed the percentage of infected nodes in network as function of time for different network size N up to N=3000 (the data for N=3000 is shown in (Figure V.6) and for a fixed probability infection rate β =0.4. We have found that these results remain robust for different network sizes and the local routing under RQL stand always efficient in term of computer virus spreading in comparison with the efficient routing strategy EP, and remain always comparable with the shortest path routing strategy.

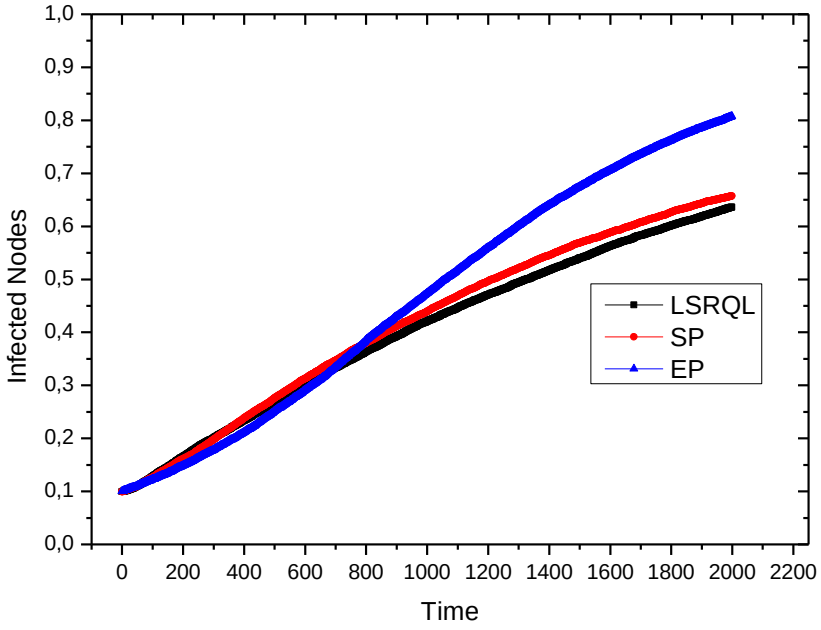


Figure V-6: the proportion of infected nodes in the whole network in local routing protocol under RQL, Shortest path (SP) and efficient path (EP). For a fixed probability infection rate β=0.4 as a function of time in the free-flow phase for a fixed R=3, and with an initial proportion of infected nodes $f_0 = 10\%$. in the SI model using $N = 3000$. $< k >= 10$.

V.4 Conclusion:

In conclusion, in this chapter, we have studied the virus propagation in respectively local static, dynamic routing protocols with and without the additional algorithms next-nearest-neighbors (NNN), restrictive-queue-length (RQL) and node-duplication-avoidance (NDA).

We have shown that the algorithms NNN and RQL which were designed to overcome network congestion in local routing strategies consolidate at the same time the robustness of local routing strategies against the computer virus propagation. While the additional NDA algorithm favors the virus spreading in the sparse networks.

Finally and in comparison with global routing strategies(SP,EP), the local routing protocol under the RQL unexpectedly overcome the efficient path protocol regarding to computer virus propagation and show a high similarity with the traditional shortest path strategy, due to the closeness in packet traveling path length.

Moreover, the rapid grow in time of real networks as the internet, is a huge challenge for the global routing strategy in terms of resource consuming which requires the global network information. Thus the shortest path protocol can easily be replaced by local routing strategies with RQL to benefit from the high capacity and the high level of security plus the adaptability with the large scale free networks.

General conclusion

Research on network dynamic spreading namely; routing protocols and computer virus propagation could not be effective without a better knowledge of the different types of networks, and the choice of a model that complies with the reality of networks in general.
Indeed, the choice of network was based on Barabási and Albert model of scale-free networks [26], to represent the structure of Internet. This model having the advantage of taking in consideration two interesting properties in internet, which did not exist in the previous random network models, namely:

- The Internet is not fixed in size but expands continuously.
- The new node attachment is always preferential and not random and uniform.

However, in order to overcome the congestion issue and reduce computer network vulnerably against viruses, scientists were focused essentially on designing efficient routing strategies capable to deal with the increasing demand for traffic network systems and in the same time able to reduce the probability of computer virus to spread across the network.
In this context, this thesis was centered on designing efficient routing strategy to overcome congestion under shortest path strategy and on other hand to study the robustness of global and local routing strategies in term of computer virus spreading.

Inspired by shortest path strategy and to more deal with time sensitive traffic in real networks [73], in chapter 3, we have suggested a new priority model based on packet destination; Instead of uniform prioritization of packet used in previous models, we have shown that the priority should be attributed to a given packet based on its destination. Moreover, using our proposed priority model, the traffic capacity is clearly reduced in congested phase compared with traditional famous shortest path routing strategy.

However, the information packets sent from one router to another, and following any routing strategy, may be unintentionally infected and may also infect other information packets in network. Furthermore, recognizing that an efficient routing is not only about the augmentation of the traffic capacity but it also about packets security and robustness in term of computer virus, in chapter 5, using model SI we have investigate the virus spreading in local routing protocols with and without their additional algorithms, then we performed a comparison with the global routing strategies. We found that the additional

algorithms[163] designed for alleviating congestion from local routing strategies consolidate surprisingly at the same time the network more than the global efficient routing strategy.

We also believe that, based on these results and due to the rapid grow in time of real networks as the internet wich remain a huge challenge for the global routing strategy in terms of resource consuming which requires the global network information, These results could be very helpful for network routing protocols designers to give more attention to local routing strategies and their additional algorithms in order to benefit from high capacity and high level of security plus the adaptability with the large scale free networks as the internet. Thus the shortest path protocol can easily be replaced by local routing strategies with additional algorithms to benefit from the high capacity and the high level of security plus the adaptability with the large scale free networks

Future works expectation

Thanks to the increasing attention given to this subject from network science community, up till now many intriguing questions have been answered. Nevertheless, there are yet many open questions that raise up.

- In the majority of designed routing strategy, the only main goal is to achieve a high network capacity. While in reality its worth nothing to overcome the congestion and favor the virus spread; in this context and based on our results ,we propose to study the possibility of designing new hybrid routing strategy based on both local and global routing strategy mainly shortest path two gain both security and economic sides.
- In our knowledge and in general due the low capacity of local routing strategies compared with global routing strategies, this categories have received less interest from scientists .While as show in chapter 5, designing additional algorithms for local routing strategies remain highly promising in term of network capacity and security. Thus this category of routing strategies should gain more attention from routing strategies designers.

References

[1] L. Euler, Solutio Problemat is ad Geometriam Situs Pertinentis. Commentarii Academiae Scientiarum Imperialis Petropolitanae 8:128-140, 1741.

[2] G. Alexanderson. Euler and Königsberg's bridges: a historical view. *Bulletin of the American Mathematical Society* 43: 567, 2006.

[3]Vito Latora and Massimo Marchiori. The Architecture of complexe Systems. Oxford University Press, 2002.

[4] Duncan J.Watts. A simple model of global cascades on random networks. Proceedings of the National Academy of Sciences, 99 : 5766-5771, 2002.

[5] D. J. Watts and S. H. Strogatz. Collective dynamics of "small-world" networks. Nature, 393:440–442, 1998.

[6] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex weighted networks. PNAS, 101:3747–3752, 2004.

[7] J. P. Onnela, J. Saramäki, J. Kertész, and K. Kaski. Intensity and coherence of motifs in weighted complex networks. Physical Review E, 71:065103, 2005.

[8] B. Zhang and S. Horvath. A general framework for weighted gene coexpression network analysis. Statistical Applications in Genetics and Molecular Biology, 4:17, 2005.

[9] P. Holme, S. M. Park, J. B. Kim, and C. R. Edling. Korean university life in a network perspective: Dynamics of a large affiliation network. Physica A, 373:821–830, 2007.

[10] R. Solomonoff and A. Rapoport. Connectivity of random nets. Bulletin of Mathematical Biology, 13:107-117, 1951.

[11] P. Erdős and A. Rényi. On random graphs, I. Publicationes Mathematicae (Debrecen), 6:290-297, 1959.

[12] P. Erdős and A. Rényi. On the evolution of random graphs. Publ. Math. Inst. Hung. Acad. Sci., 5:17-61, 1960.

[13] P. Erdős and A. Rényi. On the evolution of random graphs. Bull. Inst. Internat. Statist., 38:343-347, 1961.

[14] P. Erdős and A. Rényi. On the Strength of Connectedness of a Random Graph, Acta Math. Acad. Sci. Hungary, 12: 261–267, 1961.

[15] P. Erdős and A. Rényi. Asymmetric graphs. Acta Mathematica Acad. Sci. Hungarica, 14:295-315, 1963.

[16] P. Erdős and A. Rényi. On random matrices. Publ. Math. Inst. Hung. Acad. Sci., 8:455-461, 1966.

[17] P. Erdős and A. Rényi. On the existence of a factor of degree one of a connected random graph. Acta Math. Acad. Sci. Hungary, 17:359-368, 1966.

[18] P. Erdős and A. Rényi. On random matrices II. Studia Sci. Math. Hungary, 13:459-464, 1968.

[19] Shoubridge, P. Hybrid Routing in Dynamic Networks. 1381–1386

[20] I. de Sola Pool and M. Kochen. Contacts and Influence. Social Networks, 1: 5-51, 1978.

[21] S. Milgram. The Small World Problem. Psychology Today, 2: 60-67, 1967.

[22] J. Travers and S. Milgram. An Experimental Study of the Small World Problem. Sociometry, 32:425-443, 1969.

[23] Wang, X. F. & Chen, G. Complex Networks : Scale-Free and Beyond. (2003).

[24] Jiang, Z., Ma, J. & Jing, X. Enhancing traffic capacity of scale-free networks by employing hybrid routing strategy. Physica A 422, 181–186 (2015).

[25] L. Backstrom, P. Boldi, M. Rosa, J. Ugander, and S. Vigna. Four degrees of separation. In ACM Web Science 2012: Conference Proceedings, pages 45−54. ACM Press, 2012.

[26] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. Science, 286:509-512, 1999.

[27]M. Faloutsos, P. Faloutsos, and C. Faloutsos, On power-law relationships of the internet topology, Computer Communications Rev., 29 (1999), pp. 251–262.

[28] R. Pastor-Satorras and A.Vespignani. *Evolution and Structure of the Internet: A Statistical Physics Approach*. Cambridge University Press, Cambridge, 2004.

[29] B. Bollobas, Random Graphs, Academic Press, London, 1985.

[30] B. Bollobas, Modern Graph Theory, Graduate Texts in Mathematics, Springer, New York, 1998.

[31] D.B. West, Introduction to Graph Theory, Prentice-Hall, Englewood Cliffs, NJ, 1995.

[32] Janson, S., Luczak, T., and Rucinski, A., 2000, Random Graphs (New York: Wiley).

[33] Li, G. et al. Optimal transport exponent in spatially embedded networks. 42810, 1–8 (2013).

[34] Tan, F. & Xia, Y. Hybrid routing on scale-free networks. Physica A 392, 4146–4153 (2013).

[35] D.J. Watts, Small Worlds: The Dynamics of Networks between Order and Randomness, Princeton University Press, Princeton, NJ, 1999.

[36] Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. Rev. Mod. Phys. 74, 47–97 (2002).

[37] Barabási, A. & Bonabeau, E. Scale-free networks. Sci. Am. 288, 50–59 (2003).

[38] Jure Leskovec and Eric Horvitz. Planetary-scale views on a large instant-messaging network. In WWW, 915–924, 2008.

[39]D. J. de S. Price, A general theory of bibliometric and other cumulative advantage processes, J. Amer. Soc. Inform. Sci., 27 (1976), pp. 292–306.

[40] D. J. de S. Price, Networks of scientific papers, Science, 149 (1965), pp. 510–515

[41] S. Redner, How popular is your paper? An empirical study of the citation distribution, Eur. Phys. J. B, 4 (1998), pp. 131–134.

[42] P. O. Seglen, The skewness of science, J. Amer. Soc. Inform. Sci., 43 (1992), pp. 628–638.

[43] R. Albert, H. Jeong, and A.-L. Barab´asi, Diameter of the world-wide web, Nature, 401 (1999), pp. 130–131.

[44] A.-L. Barabasi, R. Albert, and H. Jeong, Scale-free characteristics of random networks: The topology of the World Wide Web, Phys. A, 281 (2000), pp. 69–77.

[45] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, Graph structure in the web, Computer Networks, 33 (2000), pp. 309–320.

[46] Q. Chen, H. Chang, R. Govindan, S. Jamin, S. J. Shenker, and W. Willinger, The origin of power laws in Internet topologies revisited, in Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Societies, IEEE Computer Society, Los Alamitos, CA, 2002.

[47] Ratanchandani, P. & Kravets, R. A Hybrid Approach to Internet Connectivity for Mobile Ad Hoc Networks. 0, 1522–1527 (2003).

[48] A. Vazquez, R. Pastor-Satorras, and A. Vespignani, Large-scale topological and dynamical properties of the Internet, Phys. Rev. E, 65 (2002), art. no. 066130.

[49] H. Jeong, S. Mason, A.-L. Barabasi, and Z. N. Oltvai, Lethality and centrality in protein networks, Nature, 411 (2001), pp. 41–42

[50] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabasi, The large-scale organization of metabolic networks, Nature, 407 (2000), pp. 651–654

[51] W. Aiello, F. Chung, and L. Lu, A random graph model for massive graphs, in Proceedings of the 32nd Annual ACM Symposium on Theory of Computing, Association of Computing Machinery, New York, 2000, pp. 171–180.

[52] W. Aiello, F. Chung, and L. Lu, Random evolution of massive graphs, in Handbook of Massive Data Sets, J. Abello, P. M. Pardalos, and M. G. C. Resende, eds., Kluwer Academic, Dordrecht, 2002, pp. 97–122.

[53] J. H. Jones and M. S. Handcock, An Assessment of Preferential Attachment as a Mechanism for Human Sexual Network Formation, preprint, University of Washington, Seattle, 2003.

[54] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Aberg, The web of human sexual contacts, Nature, 411 (2001), pp. 907–908.

[55] Dolev, S., Elovici, Y. & Puzis, R. Routing Betweenness Centrality. (2009).

[56] Brandes, U. A Faster Algorithm for Betweenness Centrality. 25, 163–177

[57] Freeman, L. C. A Set of Measures of Centrality Based on Betweenness. Sociometry 40, 35 (1977).

[58] R. Guimera, A. Arenas, A. D´ıaz-Guiler`a, and F. Giralt, "Dynamical properties of model communication networks," Physical Review E, vol. 66, no. 2, Article ID 026704, 8 pages, 2002.

[59] L. Zhao, Y. C. Lai, K. Park, and N. Ye, "Onset of traffic congestion in complex networks," Physical Review E, vol. 71, no. 2, Article ID 026125, 8 pages, 2005.

[60] B. Tadic, S. Thurner, and G. J. Rodgers, "Traffic on complex networks: towards understanding global statistical properties from microscopic density fluctuations," Physical Review E, vol. 69, no. 3, Article ID 036102, 5 pages, 2004.

[61] R. Guimera, A. Dıaz-Guilera, F. Vega-Redondo, A. Cabrales, and A. Arenas, "Optimal network topologies for local search with congestion," Physical Review Letters, vol. 89, no. 24, Article ID 248701,4 pages, 2002.

[62] G. Mukherjee and S. S. Manna, "Phase transition in a directed traffic flow network," Physical Review E, vol. 71, no. 6, Article ID 066108, pp. 1–6, 2005.

[63] Z. Wu, G. Peng, W. Wong, and K. Yeung, "Improved routing strategies for data traffic in scale-free networks," Journal of Statistical Mechanics, vol. 2008, no. 11, Article ID P11002, 2008.

[64] H. Zhang, Z. H. Liu, M. Tang, and P. M. Hui, "An adaptive routing strategy for packet delivery in complex networks," Physics Letters, Section A, vol. 364, no. 3-4, pp. 177–182, 2007.

[65] Z. Chen and X. Wang, "Effects of network structure and routing strategy on network capacity," Physical Review E, vol. 73, no. 3, Article ID 036107, pp. 1–5, 2006.

[66] M. B. Hu, B. H.Wang, R. Jiang, Q. S.Wu, and Y. H.Wu, "The effect of bandwidth in scale-free network traffic," Europhysics Letters, vol. 79, no. 1, Article ID 14003, 2007.

[67] M. B.Hu, R. Jiang, Y. H. Wu, and Q. S. Wu, "The effect of link and node capacity on traffic dynamics in weighted scale-free networks," in Proceedings of the International Conference on Complex Sciences, vol. 4, part 1 of Lecture Notes of the Institute for Computer Sciences and Telecommunications, pp. 580–588, Springer, Berling Heidelberg, Germany, 2009.

[68] R. Jiang, M. B. Hu, W. X. Wang, G. Yan, Q. S. Wu, and B. H. Wang, "Traffic dynamics of packets generated with none-homogeneously selected sources and destinations in scale-free networks," Dynamics of Continuous, Discrete and Impulsive Systems B Supplement, vol. 14, supplement 7, pp. 51– 54, 2007.

[69] A. Arenas, A. Daz-Guilera and R. Guimera, Phys. Rev. Lett. 86, 3196 (2001).

[70] He, K., Xu, Z. & Wang, P. A hybrid routing model for mitigating congestion in networks networks. Physica A (2015). doi:10.1016/j.physa.2015.02.087

[71] Thurner, S. TRANSPORT ON COMPLEX NETWORKS : FLOW ,. 17, 2363–2385 (2007).

[72] Ling, X., Hu, M., Jiang, R. & Wu, Q.-S. Global dynamic routing for scale-free networks. Phys. Rev. E 81, 16113 (2010).

[73] X. Zhang, Z. Zhou and D. Cheng, PLoS ONE 12, e0172035 (2017).

[74] W. X. Wang et al., Phys. Rev. E 73, 026111 (2006).

[75] W. X. Wang, C. Y. Yin and G. Yan, Phys. Rev. E 74, 016101 (2006).

[76] X. Ling, M. B. Hu and R. Jiang, Phys. Rev. E 80, 066110 (2009).

[77] G. Yan et al., Phys. Rev. E 73, 046108 (2006).

[78] R. V. Sole and S. Valverde, "Information transfer and phase transitions in a model of internet traffic," Physica A, vol. 289, no. 3-4, pp. 595–605, 2001.

[79] M. E. J. Newman, Phys. Rev. E 64, 016132 (2001).

[80] S. Valverde and R. V. Sol´e, "Self-organized critical traffic in parallel computer networks," Physica A,vol. 312, no. 3-4, pp. 636–648, 2002.

[81] M. E. Crovella and A. Bestavros, "Self-similarity in world wide web traffic: evidence and possible causes," IEEE/ACM Transactions on Networking, vol. 5, no. 6, pp. 835–846, 1997.

[82] X. Jia, W. Zhao, and J. Li, "An integrated routing and admission control mechanism for real-time multicast connections in ATM networks," IEEE Transactions on Communications, vol. 49, no. 9, pp. 1515–1519, 2001.

[83] X. Ling et al., Physica A, Stat. Mech. Appl. 387, 47094715 (2008).

[84] P. Echenique, J. Gomez-Gardenes and Y. Moreno, Phys. Rev. E 70, 056105 (2004).

[85] J. M. Kleinberg, Nature 406, 845 (2000).

[86] B. J. Kim et al., Phys. Rev. E 65, 027103 (2002).

[87] C. P. Herrero, Phys. Rev. E 71, 016103 (2005).

[88] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of Ethernet traffic," Computer Communication Review, vol. 23, article 283, 1993.

[89] A. T. Lawniczak and X. Tang, Eur. Phys. J. B-Condens. Matter Complex Syst. 50, 231 (2006).

[90] B. Danila et al., Phys. Rev. E 74, 046106 (2006).

[91] B. Danila et al., Chaos 17, 026102 (2007).

[92] Z. Y. Jiang and M. G. Liang, Physica A 392, 1894 (2013).

[93] Z. Liu et al., Phys. Rev. E 76, 037101 (2007).

[94] A. Rachadi, M. Jedra and N. Zahid, Chaos 23, 013114 (2013).

[95] A. Rachadi, M. Jedra and N. Zahid, J. Complex Netw. 3, 291 (2015).

[96] Y. Yang et al., Int. J. Mod. Phys. C 287, 1750087 (2017).

[97] J. Ma et al., Int. J. Mod. Phys. B 3213, 1850155 (2018).

[98] J. Ma et al., Physica A 456, 281 (2016).

[99] G. Q. Zhang, D. Wang and G. J. Li, Phys. Rev. E 76, 017101 (2007).

[100] W. Huang and T. W. S. Chow, J. Stat. Mech., Theory Exp. 1, P01016 (2010).

[101] K. Hu et al., J. Phys. A, Math. Theor. 43, 175101 (2010).

[102] W. Huang and T. W. Chow, Chaos, Interdiscip. J. Nonlinear Sci. 20, 033123 (2010).

[103] Z. Jiang, M. Liang and D. Guo, Int. J. Mod. Phys. C 22, 1211 (2011).

[104] Farooq, M. & Caro, G. A. Di. Routing Protocols for Next Generation Networks Inspired by Collective Behaviors of Insect Societies : An Overview.

[105] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han, Phys. Rev. E 65, 056109 (2002).

[106] Technology, N., Thorenoor, S. G. & Technologies, W. Background Using OPNET Modeler. 191–195 (2010). doi:10.1109/ICCNT.2010.66 122. There, D. R. et al. - Static vs . Dynamic Routing -. 1–4 (2007).

[107] R. Pastor-Satorras, A. V_azquez and and A. Vespignani, Phys. Rev. Lett. 87, 258701 (2001).

[108] Haddou, N. Ben & Rachadi, A. Implantation of the global dynamic routing scheme in scale-free networks under the shortest path strategy. Phys. Lett. A 1, 1–5 (2016).

[109] Yang H X, Wang W X, Wu Z X and Wang B H 2008 Physica A 387 6857

[110 traffic allocation] Ling Xiang, Hu Mao-Bin, Long Jian-Ch, Ding Jian-Xun, and Shi Qin Chin. Phys. B Vol. 22, No. 1 (2013) 018904.

[111] Kim, K., Kahng, B. & Kim, D. Jamming transition in traffic flow under the priority queuing protocol. EPL (Europhysics Lett. 86, 58002 (2009).

[112] Du, W., Wu, Z. & Cai, K. Effective usage of shortest paths promotes transportation efficiency on scale-free networks. Physica A (2013). doi:10.1016/j.physa.2013.03.032.

[113] Z.-Y. Jiang, M.-G. Liang and J.-J. Wu, PLoS ONE 8, e82162 (2013).

[114] S. B. Li et al., Int. J. Mod. Phys. C 28, 1750117 (2017).

[115] G. H. Wu, H. J. Yang and J. H. Pan, Mod. Phys. Lett. B 32, 1850137 (2018).

[116] Johnson, D. B. & B., D. A Note on Dijkstra‟ Shortest Path Algorithm. J. ACM 20, 385–388 (1973).

[117] Dijkstra, E. W. A note on two problems in connexion with graphs. Numer. Math. 1, 269–271 (1959).

[118] Magzhan, K. & Jani, H. M. A Review And Evaluations Of Shortest Path Algorithms. 2, 99–104 (2013).

[119] Zhan, F. B. & Noon, C. E. Shortest Path Algorithms: An Evaluation using Real Road Networks.

[120] Perkins, C. E., Bhagwat, P., Perkins, C. E. & Bhagwat, P. Highly dynamic Destination-Sequenced Distance-Vector routing (DSDV) for mobile computers. in Proceedings of the conference on Communications architectures, protocols and applications - SIGCOMM '94 24, 234–244 (ACM Press, 1994).

[121] Hougardy, S. The Floyd-Warshall Algorithm on Graphs with Negative Cycles.

[122] Chang Wook Ahn & Ramakrishna, R. S. A genetic algorithm for shortest path routing problem and the sizing of populations. IEEE Trans. Evol. Comput. 6, 566–579 (2002).

[123] http://www.avira.com/en/threats/section/wildlist_intro/index.html (last accessed August 10, 2014).

[124] F. Cohen, Computer virus: theory and experiments, Comput. Secur. 6 (1) (1987) 22–35.

[125] W.H. Murray, Comput. Secur. 7 (2) (1988) 130–150.

[126] J.O. Kephart, S.R. White, in: Proc. 1991 IEEE Symp. Security Privacy, 1991, pp. 343–359.

[127] L. Billings, W.M. Spears, I.B. Schwartz, Phys. Lett. A 297 (3–4) (2002) 261–266.

[128]Q. Zhu, X. Yang, J. Ren, Commun. Nonlinear Sci. Numer. Simul. 17 (12) (2012) 5117–5124.

[129] C. Gan, X. Yang, W. Liu, and Q. Zhu, Communications in Nonlinear Science and Numerical Simulation, vol. 19, no. 1, pp. 92–100, (2014).

[130] Lee, J. Traffic-Aware Hybrid Routing Algorithm for Mobile Ad-hoc Networks. 219–224.

[131] H. Yuan, G. Chen, Appl. Math. Comput. 206 (1) (2008) 357–367.

[132] B.K. Mishra, S.K. Pandey, Appl. Math. Comput. 217 (21) (2011) 8438–8446.

[133] B.K. Mishra, D.K. Saini, Appl. Math. Model. 34 (3) (2010) 710–715.

[134] X. Yang, L.-X. Yang, Discrete Dyn. Nat. Soc. 2012 (2012). Article ID 259671.

[135] L.-X. Yang, X. Yang, L. Wen, J. Liu, Discrete Dyn. Nat. Soc. 2012 (2012). Article ID 693695.

[136]Q. Zhu, X. Yang, L. X. Yang, and X. Zhang, Nonlinear Dynamics, vol. 73, no. 3, pp. 1433–1441, (2013).

[137] P.Owezarski et N.Larrieu. Techniques et outils de métrologie pour l'Internet et son trafic. (2007).

[138] W. Leland, M. Taqqu, W. Willinger, and D. Wilson. On the self-similar nature of ethernet traffic. IEEE/ACM Trans. Net. (1994).

[139] H.J. Fowler and W.E. Leland, IEEE J. Sel. Areas in Commun, (1994).

[140] N. Hohn, D. Veitch, and P. Abry, in ACM SIGCOMM Internet Measurement Workshop (IMW-2002):63–68, Marseille, France, Nov 6–8 (2002).

[141] J.O. Kephart, S.R. White,in: Proc. 1993 IEEE Symp. Security Privacy, pp. 2–15, (1993).

[142] J.R.C. Piqueira, A.A. de Vasconcelos, C.E.C.J. Gabriel, V.O. Araujo, Comput. Secur. 27 (7–8) 355–359(2008).

[143] Zheng, Y. & Liu, B. Fuzzy vehicle routing model with credibility measure and its hybrid intelligent algorithm. 176, 673–683 (2006).

[144] C. Zhang, Y. Zhao, Y. Wu, Discrete Dyn. Nat. Soc. 2012 (2012). Article ID 260962.

[145] Benyoussef, M., Ez-zahraouy, H., Benyoussef, A. & Magn, L. De. New behavior of degree distribution in connected communication networks. 25, 1–9 (2014).

[146] Lazfi, S., Lamzabi, S., Rachadi, A. & Ez-Zahraouy, H. The impact of neighboring infection on the computer virus spread in packets on scale-free networks. Int. J. Mod. Phys. B 1750228 (2017).

[147] C. Zhang, Y. Zhao, Y. Wu, S. Deng, Discrete Dyn. Nat. Soc. 2012 (2012). Article ID 264874.

[148] J. Ren, X. Yang, Q. Zhu, L.X. Yang, C. Zhang, Nonliear Anal. – Real 13 (1) (2012) 376–384.

[149] Ben Haddou, N., Ez-Zahraouy, H. & Benyoussef, A. An adaptive routing scheme in scale-free networks. Int. J. Mod. Phys. C 26, 1550138 (2015).

[150] Lamzabi, S. et al. Pair-dependent rejection rate and its impact on traffic flow in a scale-free network. Int. J. Mod. Phys. C 25, 1450019 (2014).

[151] S. Lamzabi, S. Lazfi, A. Rachadi, H. Ez-Zahraouy, A. Benyoussef, Modeling the spread of virus in packets on scale free network , Int. J. Mod. Phys. C, Vol. 27, No. 6 (2016).

[152] W.O. Kermack and A.G. Kendrick. A contribution to the mathematical theory of epidemics. In the royal Society, 1927

[153] Cholvi V, Laderas V, López L, Fernández A. Physical Review E. 71,035103 (2005).

[154] A. Ould Baba, O. Bamaarouf, A. Rachadi. H. EZ-Zahraouy. International Journal of Modern Physics C. 28, 1750064 (2017).

[155] O. Bamaarouf. A. Ould Baba, S. Lamzabi, A. Rachadi, H. Ez Zahraouy. International Journal of Modern Physics B. 31, 1750182 (2017).

[156] G. Ayalvadi, MASSOULIÉ, Laurent, et TOWSLEY, Don In : INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE. IEEE 1455-1466 (2005).

[157] S.Meloni, A. Arena, and Y. Moreno. Proc. Natl. Acad. Sci. USA 106, 16897 (2009).

[158] N. T. J. Bailey, The Mathematical Theory of Infectious Diseases ,(1975).

[159] Han-Xin Y, Wen-Xu W, Ying-Cheng L, Yan-Bo X, Bing- Hong W. Physical Review E. 84, 045101 (2011).

[160] Han-Xin Y and Zhi-Xi W. J. Stat. Mech. 03018 (2014).

[161] Han-Xin Y, Zhi-Xi W and Bing-Hong Wang. Physical Review E. 87, 064801 (2013).

[162] O. Bamaarouf, A.O.B. Alweimine, A. Rachadi, H. EZ-Zahraouy. Physica A: Statistical Mechanics and its Applications. 496,209-219 (2018).

[163] Lin B, Chen B, Gao Y, Tse CK, Dong C, Miao L. PLoS ONE. 11 (7): e0156756 (2016).

[164] Adamic LA, Lukose RM, Puniyani AR, Huberman BA. Physical Review E. 64: 046135 (2001).

[165] J.L Moreno, Fondement de la sociométrie, Presses Universitaires de France, (1934).

[166] A Radcliffe-Brown, Journal of the Royal Anthropological Institute, (70), (1940).

[167] E Mayo, Harvard University Press, (1945).

[168] DWatts, Six degrees: The science of connected age,Barnes & Noble, (2004).

[169] M Granovetter, American journal of sociology, (78), (1973).

[170] M McPherson, L Smith-Lovin, and J Cook, Annual review of sociology, (27), (2001).

[171] W.E. Leland, M.S. Taqqu, W. Willinger, and D.V. Wilson, IEEE/ACM Transactions on Networking, vol. 2, no. 1, pp. 1–15, Feb (1994).

[172] W. Willinger, M.S. Taqqu, R. Sherman, and D.V. Wilson, in Proceedings of the ACM/SIGCOMM'95, (1995).

[173] P.Owezarski et N.Larrieu. Techniques et outils de métrologie pour l'Internet et son trafic. (2007).

[174] W. Leland, M. Taqqu, W. Willinger, and D. Wilson. On the self-similar nature of ethernet traffic. IEEE/ACM Trans. Net. (1994).

[175] H.J. Fowler and W.E. Leland, IEEE J. Sel. Areas in Commun, (1994).

[176]Chen, S., Huang, W., Cattani, C. & Altieri, G. Traffic Dynamics on Complex Networks : A Survey. 2012, (2012).

[177] J.L. Aron, M. O'Leary, R.A. Gove, S. Azadegan, M.C. Schneider,Computers and Security 21 ,142–163,(2002).

[178] J.O. Kephart, S.R. White, in IEEE Computer Society Symposium On Research In Security and Privacy: Proceedings, pp. 2–15, (1993).

[179] R. Perdisci, A. Lanzi, W. Lee, Pattern Recognition Letters 29,1941–1946, (2008).

[180] Y.B. Kafai, Journal of Science Education and Technology 6, 523–529, (2009).

[181] L. Billings, W.M. Spears, I.B. Schwartz, Physics Letters A 297, 261–266, (2002).

[182] C. Nachenberg, Computer virus-Coevolution, Communications of the ACM 40, 46–51, (1997).

[183] H. Thimbleby, S. Anderson, P. Cairns, Computer Journal 41, 444–458, (1998).

[184] J.C. Wierman, D.J. Marchette, Computational Statistics and Data Analysis 45, 3–23, (2004).

[185] T. Li, Science in China Series F-Information Sciences 51, 1475–1486, (2008).

[186] A. Bissett, International Journal Of Human-Computer Studies 52, 899–913, (2000).

[187] E. Makinen, Computer Journal 44, 321–323, (2004).
[188] J.E. Sawyer, M.C. Kernan, D.E. Conlon, H. Garland, Journal of Applied Social Psychology 29, 23–51, (1999).

[189] D. Balcan, H. Hu, B. Goncalves, P. Bajardi, C. Poletto, J. J. Ramasco, D. Paolotti, N. Perra, M. Tizzoni, W. Van den Broeck, V. Colizza, and A. Vespignani. Seasonal transmission potential and activity peaks of the new influenza A(H1N1): a Monte Carlo likelihood analysis based on human mobility. BMC Medicine, 7: 45, 2009.

[190] L. Hufnagel, D. Brockmann, and T. Geisel. Forecast and control of epidemics in a globalized world. PNAS, 101: 15124, 2004.

[191] P. Wang, M. Gonzalez, C. A. Hidalgo, and A.-L. Barabási. Understanding the spreading patterns of mobile phone viruses. Science, 324: 1071, 2009.

[192] A. Ould Baba Alweimine, O. Bamaarouf, A. Rachadi. H. EZ-Zahraouy : Implementing beneficial prioritization of traffic flow in complex networks. International Journal of Modern Physics B (2018) 1850273

[193] A. Ould Baba Alweimine, O. Bamaarouf, A. Rachadi. H. EZ-Zahraouy : Local routing protocols performance for computer virus elimination in complex networks. Physica A: Statistical Mechanics and its Applications (2019) https://doi.org/10.1016/j.physa.2019.04.220.

List of publications & communications

Articles:

[1] A. Ould Baba, O. Bamaarouf, A. Rachadi. H. EZ-Zahraouy: Effect of connection on transport between scale free networks. International Journal of Modern Physics C. 28, 1750064 (2017).

[2] A. Ould Baba Alweimine, O. Bamaarouf, A. Rachadi. H. EZ-Zahraouy : Implementing beneficial prioritization of traffic flow in complex networks. International Journal of Modern Physics B (2018) 1850273

[3] A. Ould Baba Alweimine, O. Bamaarouf, A. Rachadi. H. EZ-Zahraouy : Local routing protocols performance for computer virus elimination in complex networks. Physica A: Statistical Mechanics and its Applications (2019).

[4] A. Ould Baba Alweimine, O. Bamaarouf , A. Rachadi. H. EZ-Zahraouy : Optimal hybrid routing strategy . We still work on.

[5] O. Bamaarouf. A. Ould Baba, S. Lamzabi, A. Rachadi, H. Ez Zahraouy: Effects of maximum node degree on computer virus spreading in scale-free networks. International Journal of Modern Physics B. 31, 1750182 (2017).

[6] O. Bamaarouf, A. Ould Baba Alweimine, A. Rachadi, H. EZ-Zahraouy: Selective epidemic vaccination under the performant routing algorithms. Physica A: Statistical Mechanics and its Applications. 496,209-219 (2018).

[7] O. Bamaarouf, A. Ould Baba Alweimine, A. Rachadi. H. EZ-Zahraouy : Border routing in interconnected networks . We still work on.

National & international communications:

[7] O. Bamaarouf, A. Ould Baba, A. Rachadi. H. EZ-Zahraouy : effect of maximum node degree on computer virus spreading in scale free networks . the xxxii edition of international conference of physics Students ICPS 2017, Torino, Italy, August 7-14, 2017.

[8] O. Bamaarouf, A. Ould Baba Alweimine, A. Rachadi, S. Lamzabi, H. EZ-Zahraouy : epidemic spreading in growing network by coping .the third Edition of the international Conference on electrical and information technologies ICIET2017, Rabat, Morocco, November 15-18, 2017.

[9] O. Bamaarouf, A. Ould Baba Alweimine, A. Rachadi, S. Lamzabi, H. EZ-Zahraouy : Traffic-driven epidemic spreading in interconnected networks . 2éme edition de la conference international technologies, in innovation & Systéme d'information Journée National de Vibration et Acoustique CITIST2018, Faculté des Sciences de Kénitra, Maroc,July 14-15, 2018.

[10] A. Ould Baba Alweimine, O. Bamaarouf, A. Rachadi, S. Lamzabi, H. EZ-Zahraouy : Time sensitive packets navigation routing protocol in comlex networks . 2éme edition de la conference international

technologies, in innovation & Systéme d'information Journée National de Vibration et Acoustique CITIST2018, Faculté des Sciences de Kénitra, Maroc,July 14-15, 2018.

[11] A. Ould Baba Alweimine, O. Bamaarouf, A. Rachadi, S. Lamzabi, H. EZ-Zahraouy : new beneficial prioritization policy of the traffic in complex networks. The first annual meeting of LaMCScI,Faculty of Sciences of rabat, Morocco, December 13-14,2018.

[12] O. Bamaarouf, A. Ould Baba Alweimine, A. Rachadi, S. Lamzabi, H. EZ-Zahraouy : Border routing in interconnected networks. The first annual meeting of LaMCScI,Faculty of Sciences of rabat, Morocco, December 13-14,2018

Résume :

Cette thèse porte sur la modélisation et la simulation de protocoles de routage pour réduire la congestion, et sur le problème de la propagation des virus dans les réseaux complexes. Dans cette thèse, nous avons introduit deux contributions principales:

Premièrement, dans le but de réduire la congestion dans la stratégie de routage de plus court chemin (SP) et de mieux prendre en compte le concept de trafic prioritaire sur Internet, nous avons proposé un nouveau modèle de hiérarchisation des flux de trafic dans lequel les paquets sous la stratégie de plus court chemin sont hiérarchisés en fonction de leur destination. Nous avons constaté que la priorisation de trafic destiné aux nœuds avec degré de connexion élevé (hubs) est toujours plus efficace que la priorisation de trafic destiné aux nœuds avec un faible degré ou la priorisation aléatoire de trafic

Deuxièmement, nous avons étudié l'efficacité des protocoles de routage locaux et de leurs algorithmes supplémentaires. Voisins les plus proches (NNN) et algorithme restrictif de longueur de la file d'attente (RQL) en termes de robustesse dans la propagation de virus informatiques. Il s'avère que, sous l'algorithme supplémentaire RQL, les protocoles de routage locaux sont hautement sécurisés et dépassent de manière surprenante le protocole de routage EP (voie efficace).

Mots-clefs (5) : Réseaux complexes, ingénierie du trafic, protocoles de routage, chemin le plus court, virus informatique, propagation du virus, informatiques, physique informatique, systèmes complexes.

Abstract :

This thesis focuses on modeling and simulation of routing protocols to reduce congestion and on the problem of the virus propagation in networks. In this thesis, we have introduced two main contributions:

First, for the purpose of alleviating the congestion in traditional shortest path (SP) strategy and to deal more with priority traffic concept in internet, we have proposed a new prioritization model of traffic flow where packets under shortest path strategy are prioritized according to their destination. We found that the prioritization of nodes with high degree (hubs) is always more efficient than the prioritization of nodes with small degree or the random prioritization of nodes.

Second, we studied the effectiveness of local routing protocols and their additional algorithms; next-nearest neighbors (NNN) and restrictive queue-length algorithm (RQL) in term of robustness in computer virus spreading. It is found that, under the additional algorithm RQL, local routing protocols become highly secured and overcome surprisingly the efficient path (EP) routing protocol.

Key Words (5) : Complex networks, traffic engineering , routing protocols, shortest path, computer virus, virus spreading, computer sciences, computational physics, complex systems.