

SEIQR-Network Model with Community Structure

S. ORANKITJAROEN, W. JUMPEN *
 P. BOONKRONG, B. WIWATANAPATAPHEE
 Mahidol University
 Department of Mathematics, Faculty of Science
 Centre of Excellence in Mathematics
 PERDO, CHE, Bangkok 10400
 THAILAND
 scwjp@mahidol.ac.th

Y. H. WU
 Curtin University of Technology
 Department of Mathematics and Statistics
 Perth, WA 6845
 AUSTRALIA
 yhwu@maths.curtin.edu.au

Abstract: This paper presents numerical simulation of the Susceptible-Exposed-Infectious-Quarantined-Recovered network model taking into account the community structure of the population. Numerical simulation of the disease transmission within and across the communities has been carried out. The effects of neighborhood radius, transmission rate and quarantine rate on the disease transmission are investigated.

Key-Words: complex network, genetic algorithm, community detection, disease transmission, multi-group epidemic model, network modularity, SEIQR network model

1 Introduction

It has been realized that the compartmental models do not reasonably reflect the real transmission of epidemic due to the uniform mixing assumption, i.e., all susceptible individuals are able to catch the disease with the same infection rate. The structure of human community has a direct impact to the disease proliferation. However, it has been ignored in many models. Recently, the existing models typically focus on the simulation using the combination of mathematical models and complex networks [9, 10]. It is shown that the contacts among people in the same community occur a lot more than those among people among different communities. The way to study the transmission of epidemic should consider the community structure and the interaction of people in the communities. For community establishment in the network, most of recent algorithms for constructing a network use the network modularity as a practical measure to find clear partitions (communities) of the network [1, 5]. For the application of complex network in epidemiology, many researchers have proposed epidemic network models in a discrete space domain [3, 8]. Jumpen et. al. (2011) proposed SEIQR-SIS epidemic network model and its stability [2]. Nodes in their network were classified into hubs and people, and they found that hubs had significant effect on the disease transmission. However, none of these extensions had considered the role of community structure to the disease transmission on the network. The aforementioned re-

searches show that the development of proper epidemic network model is still worldwide active.

We propose a social network using GA to detect a suitable community of a people node in order to mimic the real structure of human society. Section 2 presents an implemented GA-Network algorithm, a community detection method. In section 3, the SEIQR network model is introduced and used in conjunction with the constructed GA-Network to simulate the disease transmission. Then, some numerical results are given in section 4, followed by conclusion in section 5.

2 GA-Network algorithm

To mimic the social activities, a complex network with communities is proposed for the study of the SEIQR disease transmission. We have developed the following algorithm to generate the network in which any node can be reached by other nodes. The network consists of a number of people nodes with connecting links within a unit square region.

Let N , n and M be total numbers of people nodes, chromosomes and possible communities, respectively. The GA-Network algorithm is presented as follows.

Step 1. Locate the positions of N people nodes randomly in a unit square region by setting the shortest distance among people nodes.

Step 2. Determine neighbors of a people node using a

*Corresponding author

contact radius \mathcal{R}_c and generate connecting links between the node and its neighbors.

Step 3. Determine communities of all people nodes based on the connecting links using GA proposed by Tasgin [4] as follows:

(i) **Initialization:** There are n chromosomes and each chromosome contains N cells stored the community identification called ‘‘CommID’’ of N people nodes. These chromosomes of the node community are generated by selecting randomly CommID to some selected nodes and also to all their neighbors, and then to other nodes.

(ii) **Selection:** We select two chromosomes called as a source chromosome (C_S) and a destination chromosome (C_D) as follows:

(a) Calculate modularity (\mathcal{H}) of the chromosomes by the expression:

$$\mathcal{H} = \sum_g \left[\frac{e_g}{e} - \left(\frac{\text{deg}}{2e} \right)^2 \right], \quad (1)$$

where e and e_g denote respectively the total number of edges in the network and the number of edges in community g , and deg represents the sum of degrees k_i of the vertices of community g .

(b) Select two chromosomes having the highest modularity.

(iii) **Crossover:** We randomly select CommID named as ID-select. We then check the values in any cells in C_S . If there are ID-select, then the values in the corresponding cells of C_D are changed to be ID-select.

(iv) **Mutation:** A node may be put into a random community in the network.

(v) **Clean-up:** A node having the highest community variance (CV) is eliminated from the network.

(a) Calculate community variance for each node i in the network by the expression:

$$CV(i) = \frac{\sum_{(i,j) \in \mathcal{E}} f(i,j)}{\text{deg}(i)}, \quad (2)$$

where \mathcal{E} is the set of all edges in the network and

$$f(i,j) = \begin{cases} 1, & \text{CommID}(i) \neq \text{CommID}(j) \\ 0, & \text{otherwise.} \end{cases}$$

(b) Delete a node with the highest community variance.

(vi) **Repetition:** The process is repeated from (ii) to (v) until the network modularity is between 0.3 and 0.7 exhibiting a good community structure. It is noted that $\mathcal{H} = 0$ if a community has no within-community edges whilst $\mathcal{H} = 1$ when all nodes are put into a single community [4, 7].

Finally, the chromosome with the highest modularity representing the suitable network structure is selected for studying the disease transmission.

3 The SEIQR-Network model

The SEIQR-Network model has been developed to study the spread of the infectious disease based on SEIQR dynamics and multi-group structure in the complex network. This section illustrates the SEIQR-Network model and proposes an algorithm for the simulation of SEIQR-Network model on GA-Network presented in section 2.

In the model, there are M communities where the i^{th} community has N_i people nodes. As can be seen in Figure 1 considering the i^{th} community, the total population size $N_i(t)$ is divided into five distinct epidemiological subclasses of individuals which are susceptible $S_i(t)$, exposed $E_i(t)$, infectious $I_i(t)$, quarantined $Q_i(t)$ and recovered $R_i(t)$.

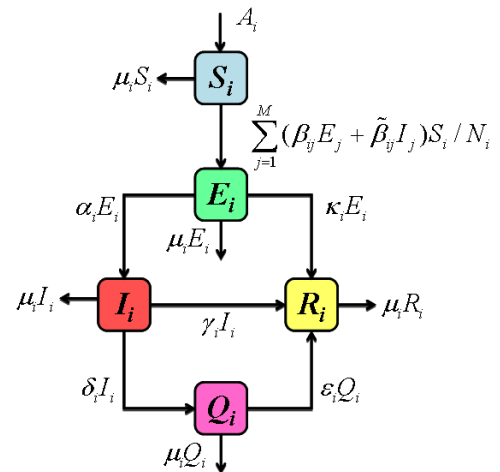


Figure 1: Progression diagram for the SEIQR disease transmission.

Figure 1 can be inferentially interpreted to the system of ordinary differential equations as follows.

For $i = 1, 2, \dots, M$,

$$\begin{aligned} \frac{dS_i}{dt} &= A_i - \sum_{j=1}^M \left[\frac{\beta_{ij}E_j + \tilde{\beta}_{ij}I_j}{N_i} \right] S_i - \mu_i S_i, \\ \frac{dE_i}{dt} &= \sum_{j=1}^M \left[\frac{\beta_{ij}E_j + \tilde{\beta}_{ij}I_j}{N_i} \right] S_i - (\alpha_i + \kappa_i + \mu_i)E_i, \\ \frac{dI_i}{dt} &= \alpha_i E_i - (\gamma_i + \delta_i + \mu_i)I_i, \\ \frac{dQ_i}{dt} &= \delta_i I_i - (\varepsilon_i + \mu_i)Q_i, \\ \frac{dR_i}{dt} &= \kappa_i E_i + \gamma_i I_i + \varepsilon_i Q_i - \mu_i R_i \end{aligned} \quad (3)$$

where $N_i = S_i + E_i + I_i + Q_i + R_i$. The outbreak of the disease on the complex network occurs when an infectious and exposed individuals transmit the disease to its susceptible neighbors via the edges with an infection rate β . It has been recognized that the transmission probability in the same community is higher than that between communities. Thus, the transmission probability within a community is assigned to be higher the transmission probability between communities. After the susceptible individuals receive an amount of virus, their status become exposed.

From the system (3), the particular assumptions for the model are described as follows:

- (i) Indices $i, j, k = 1, 2, \dots, M$ present the $i^{\text{th}}, j^{\text{th}}$ and k^{th} communities of the sub-classes S, E, I, Q and R .
- (ii) All parameters are non-negative constants defined as follows:
 - (a) α_i is the rate at which an exposed individuals $E_i(t)$ becomes infected individual $I_i(t)$;
 - (b) δ_i is the rate that individuals $I_i(t)$ moves to the quarantined individual $Q_i(t)$;
 - (c) κ_i, γ_i and ε_i are the rates at which individuals in the $E_i(t), I_i(t)$ and $Q_i(t)$ classes change their status to be in $R_i(t)$.
- (iii) The model (3) is considered with a recruitment-death demographic structure such that

$$\sum_{i=1}^M \frac{dN_i}{dt} = \sum_{i=1}^M (A_i - \mu_i N_i), \quad (4)$$

A_i is a constant recruitment in the i^{th} community such that $A_i = b_i S_i(0)$, where b_i and μ_i are the natural birth and death rates.

- (iv) β_{ij} and β_{ik} are the probabilities of catching the disease per contact to the infectious or exposed person, separately considered in two cases:

$$\beta_{ij} = \begin{cases} \beta_{\text{high}} & \text{when } i = j \text{ (intra-community)} \\ \beta_{\text{low}} & \text{when } i \neq j \text{ (inter-community)} \end{cases} \quad (5)$$

and β_{ik} is defined as the same sense as β_{ij} .

The system (3) is now considered in a social complex network. Let τ be the maximum time for the simulation, and Δt be the time step. Using the forward difference technique, we simulate the spread of the disease based on the SEIQR dynamics by the following algorithm.

- Step 1.** At initial time step $t = 0$, there are N people nodes including a few infected nodes say η in the infectious class (I -class) and $N - \eta$ people nodes in the susceptible class (S -class).
- Step 2.** Set $t = t + \Delta t$.
- Step 3.** Updating stage of people nodes based on their neighbor status according to the progression diagram as shown in Figure 1.
- Step 4.** Repeating Steps 2 and 3 until the τ reaches or there is no infectious and quarantined node in the network.

4 Numerical results and discussion

A test example is given to study the spreading of SEIQR-influenza infections in the GA network having 1,000 people nodes in 20 communities ($M = 20$). To understand how the disease spreads on the real-world population network, we first set five infectious nodes randomly and then simulate the SEIQR epidemic on the network over time using the SEIQR network algorithm presented in section 3, and set $\Delta t = 1$ day, $\beta_{\text{high}} = 2\beta_{\text{low}}$ and $\beta = \tilde{\beta}$. All other model parameters are the same for every communities. For the disease having incubation period of $1/\alpha$ days and sick period of $1/\gamma$ days, an exposed individual becomes an infectious individual at the transfer rate α and an infectious individual recovers at the transfer rate γ . From then, some infectious individuals are quarantined with the rate δ in order to reduce an infection. Finally, infectious and quarantined individuals recover when they reach the sick period. In Figure 2, the disease transmission at three different times on the complex network is presented in which square nodes represent people in the largest community and circle nodes are in other communities.

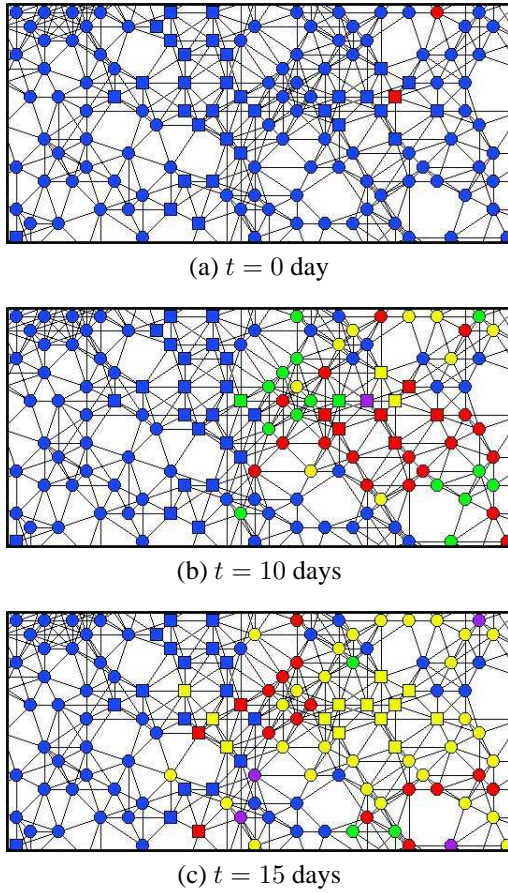


Figure 2: The disease transmission on the complex network at three different times.

Table 1: Values of parameters used in simulation.

| Biological Description | Value |
|--|-------|
| \mathcal{R}_c Neighborhood contact radius | 0.06 |
| α Transfer rate at which E becomes I | 1/4 |
| β Transmission rate | 0.02 |
| γ Transfer rate at which I becomes R | 1/6 |
| δ Quarantine rate | 1/10 |
| κ Transfer rate at which E becomes R | 0.02 |
| ε Transfer rate at which Q becomes R | 1 |
| μ Natural death rate | 0 |

Figure 3 presents the S , E , I and Q profiles corresponding to the neighborhood contact radius \mathcal{R}_c of 0.06, the transmission rate β of 0.002 and the quarantine rate δ of 0.1. The dark blue line represents the average proportion of each individual and a number of gray lines illustrate the proportion of each individual obtained from 30 runs.

The effects of model parameters including \mathcal{R}_c , β and δ on the disease transmission are investigated. Four neighborhood contact radii including \mathcal{R}_c of 0.060, 0.065 and 0.070 are considered.

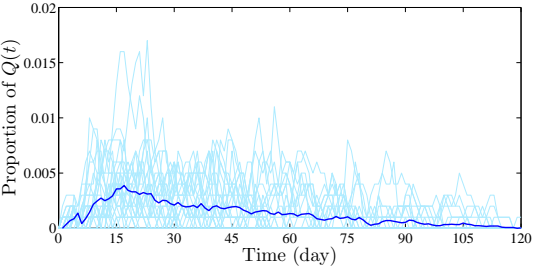
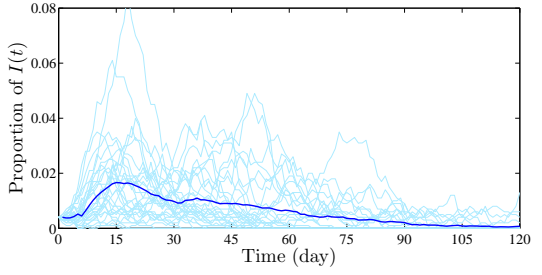
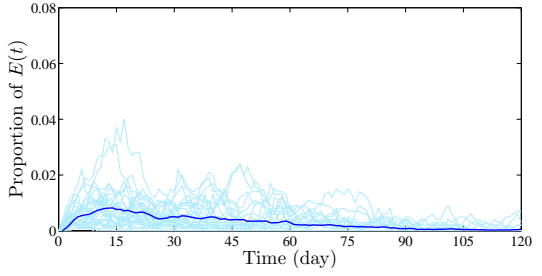
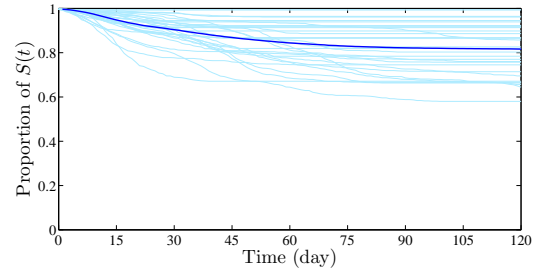


Figure 3: The profiles of the proportion of S , E , I and Q against time (day) from 30 runs of simulation.

Table 2: Effect of \mathcal{R}_c on Network Properties

| \mathcal{R}_c | 0.055 | 0.060 | 0.065 | 0.070 |
|-----------------|--------|--------|--------|--------|
| $ E $ | 4422 | 4841 | 6066 | 6879 |
| \bar{k} | 9 | 10 | 12 | 14 |
| C_l | 0.5164 | 0.5306 | 0.5461 | 0.5589 |
| \bar{d} | 8.5643 | 8.0927 | 7.1170 | 6.6882 |
| \mathcal{H} | 0.4255 | 0.4499 | 0.4548 | 0.4677 |

It is found that \mathcal{R}_c has significant effects on the properties of the network. That is an increase of \mathcal{R}_c

heightens the number of edges ($|E|$), the average degree (\bar{k}), the clustering coefficient (C_l), the network modularity (\mathcal{H}) but lessens the average vertex-vertex distance (\bar{d}) as shown in Table 2. As can be seen in Figure 4, a little change in \mathcal{R}_c gives a small effect on the disease transmission.

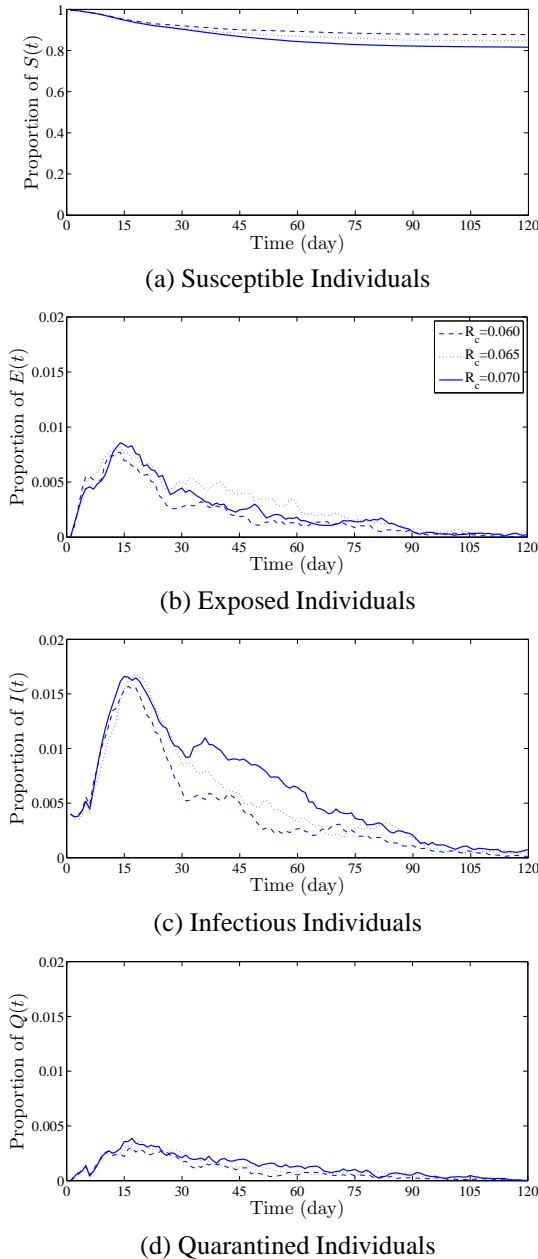


Figure 4: The effect of the neighborhood contact radius \mathcal{R}_c on the disease spread.

To investigate the impact of transmission rate, β , on the spreading behavior of the disease, we choose three values of β including 0.002, 0.003 and 0.004 and other values of parameters are fixed. As shown in Figure 5, higher value of β yields higher proportion of infectious individuals. An increase of β from 0.002

to 0.004 causes the growth in the proportion of infectious individuals, e.g., at 30 days from approximately 0.01 to 0.12.

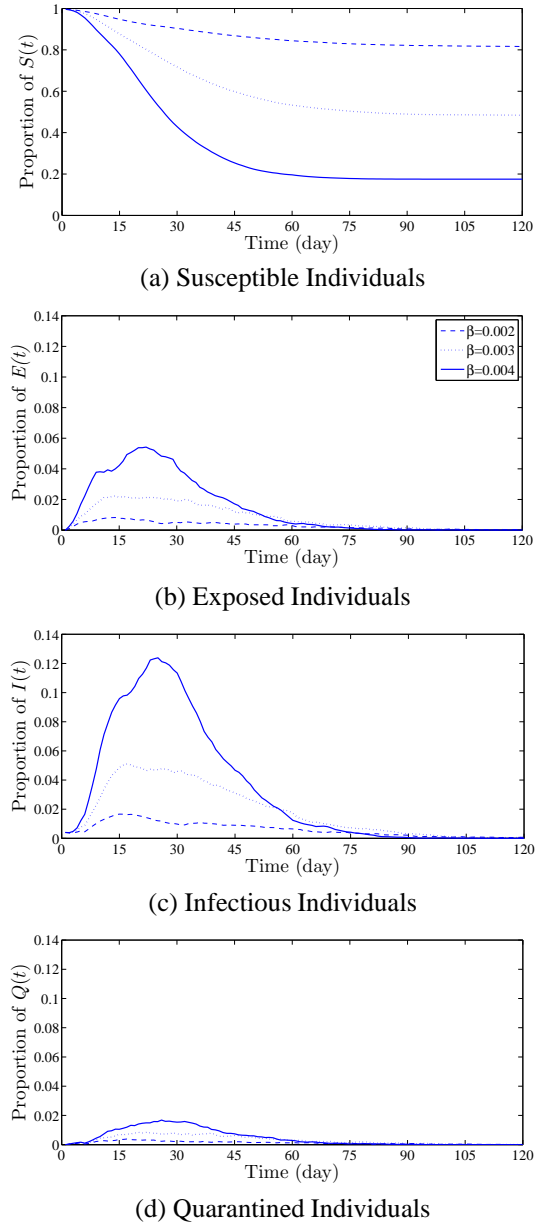


Figure 5: The effect of transmission rate β on the disease spread.

Isolation strategy is a vital approach to decrease the number of infectious individuals. Four values of the quarantine rate are chosen to investigate its impact on the disease transmission with β value of 0.003. The results in Figure 6 indicate that an increase of the quarantine rate decreases the proportion of infectious individuals. As δ increases from 0.05 to 0.20, the maximum proportion of infectious individuals drops from 0.065 to 0.030. In this study, we validate the results obtained from the model correspond-

ing to $\beta = 0.0035$, $\mathcal{R}_c = 0.065$, $\delta = 0.1$ with the actual influenza cases in Thailand during July-October, 2011. Figure 7 indicates that our model can capture the spreading behavior of the influenza in Thailand.

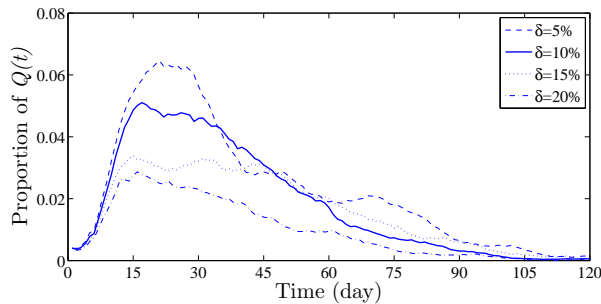


Figure 6: The effect of quarantine rate δ to the proportion of infectious individuals.

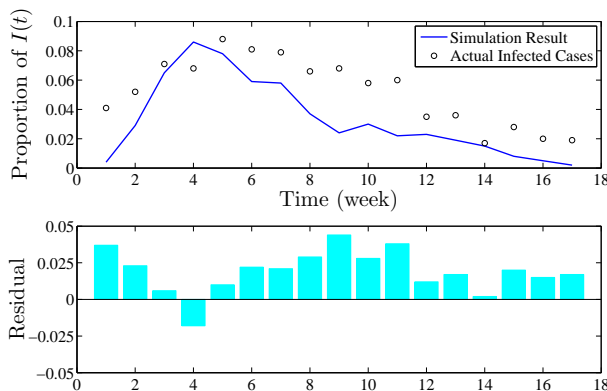


Figure 7: Comparison of disease spread between the simulation result and actual data reported by Centre of Epidemiological Information Bureau of Epidemiology, Ministry of Public Health of Thailand from July-October, 2011.

5 Conclusion

The appropriate structure of social complex network have been organized by the community detection method based on genetic algorithm. A constructed network is then cooperated with the SEIQR-Network model to study the spreading behavior of an epidemic. The simulated results point that our network model with the community structure can capture and reasonably describe the essential feature of disease transmission as expected. It shows that the transmission rate and the quarantine rate have significant effects on the disease transmission. An increase of the quarantine rate leads to a decrease of the number of the infection.

Acknowledgements: This research was supported by Centre of Excellence in Mathematics, PERDO, CHE, Ministry of Education, THAILAND.

References:

- [1] P. Chen and S. Redner, Community structure of the physical review citation network, *Journal of Informetrics* 4, 2010, pp. 278–290.
- [2] W. Jumpen, S. Orankitjaroen, P. Boonkrong and B. Wiwatanapataphee, SEIQR-SIS Epidemic Network Model and Its Stability, *International Journal of Mathematics and Computers in Simulation* , 4(5), 2011, pp. 326–333.
- [3] K. Li, M. Small, H. Zhange and X. Fu, Epidemic outbreaks on networks with effective contacts, *Nonlinear Analysis: Real World Applications.*, 11, 2010, pp. 1017–1025.
- [4] M. Tasgin, *Community Detection Model using Genetic Algorithm in Complex Networks and Its Application in Real-Life Networks*, MS Thesis, Graduate Program in Computer Engineering, Bogazici University, 2005, available at URL: <http://www.cmpe.boun.edu.tr/sosl/ab/publications.html>
- [5] Y. Qi, H. Ge, Modularity and dynamics of cellular networks. *PLoS. Comput. Biol.* 2(12), 2006, e174. doi:10.1371/journal.pcbi.0020174.
- [6] HJ. Suna and ZY. Gao, Dynamical behaviors of epidemics on scale-free networks with community structure, *Physica A* 381, 2007, pp. 491–496.
- [7] U. Brandes, D. Delling, M. Gaertler, R. Görke, M. Hoefer, Z. Nikoloski and D. Wagner *On Finding Graph Clusterings with Maximum Modularity*, Springer-Verlag Berlin Heidelberg 2007.
- [8] E. Voltz and LA. Meyers, Epidemic thresholds in dynamic contact networks, *J. R. Soc. Interface.* 6, 2009, pp. 233–41.
- [9] A. Vazquez, Epidemic outbreaks on structured populations, *em Journal of Theoretical Biology* 245, 2007, pp. 125–129.
- [10] V. Carchiol, A. Longheu, M. Malger and G. Mangioni, Search for overlapped communities by parallel genetic algorithms, *International Journal of Computer Science and Information Security* ,6(2) , 2009, pp. 113–118.