Influence of Individual on Social Networks

Team Members:

Aida Amini **810189239**

Aryaz Eghbali **810189445**

Behnood moemenzadeh **810188324**

Navid Salehnamadi **810190541**

Problem definition: In our daily life there are so many event going on around us. Our decisions about attending the event are not purely dependent to the event itself. It also depends on the community attending that event. We are influenced by others to manage our social life even though we are not aware of that. These kinds of influences are worked in other aspects of out life. People in the power are able to do many things by their social influences. For example information diffusion is one kind of the phenomena that this kind of influences causes. The other example is computer virus attacks that if the attackers find the influential ones they are able to make their attack work. Our project is developing a way to find the most influential ones in the given community according to the decisions people have made.

Previous works: The general form of our problem is how to make a data diffusion and how to make people adapt to new habits of their friends. One example is smoking or using narcotics. It is more probable that people with smoker friends starts to smoke or on the other hand friends getting out of smoking after their friends quit too.

Not all the friends have the same impact on an individual. In some communities change in more than one party is needed for the enormous movement.

There are many previous works in finding the individual influential. Here is the summary of one of these works.

This work previously is done by Stony Broke University. They have defined influenced game and introduce one approach in order to find most influenced community. They have defined a graph and make this rule that every node of this graph is getting the color gray or black and they should match with at least half of their neighbors. In the sample graph there were 2 nodes that can rule every body in the graph to have the same color as theirs, if they were aware of their power and happen to choose the same colors. The graph of the game that they introduced in that paper was an incomplete graph with some links representing the influences on nodes. They then calculated some weights for the edges and they have introduced the linear influenced game called LIG. In that game each impact from the outside world is calculate as : $\sum_{i!=j} w_{ij} * x_i - b_i$. In there defense, if the result of the calculated influence of the

outside world on each node is positive the node did somehow adapt the behavior from the neighbors and if this calculated value is less than zero it did have some influence over the others. In the formula shown above w_{ij} stands for the parameter called "influence factor" and the parameter called b_i represents the biased parameter that they have chosen to have balance in the whole system.

They continue their calculation for finding the group of most influential nodes in the community. There is a brute force approach for doing so that is not applicable in current over weighted social networks. As the problem space is exponential taking a look at every arrangement of these decisions will take so much time and memory. And it is solvable if p=np.

They have called the influential group PSNE that is stands for pure-strategy Nash equilibrium. Our goal is to find PSNE of each community.

They have announced that there is an algorithm available with the order of O(nd) that n stands for the number of nodes and d stands for the maximum degree of the graph. But this order is achieved with 2 conditions: first the game is LIG and second the graph of the game should be tree called Nash tree. But having the tree is less probable in most of the communities. The basic approach, which is backtracking, is not practical, but they found that if they choose nodes in some order, they can implement some pruning on the nodes and they can reduce the time of execution. To get the proper order they first find the maximum outdegree and first select the node that has the maximum outdegree as they believe that this node is most constraining node in the graph as is have the most influential edges. In each step they choose one new node. Their approach for doing so is to select the node that has the minimum edges to the nodes in the selected group. Each node should be assigned with one action on accepting or rejecting the event. If they cannot form PSNE out of new group they can prune. The answer of wither they can form PSNE is derived by a formula using the influenced factor parameter.

For testing their approach they have generated random graphs according to the preferential attachment algorithm and by adding nodes randomly to the initiated triangles.

Our approach: We should have some communities in the large graph of the social interactions and also we need to gather data from different individuals in our social network. The kind of data needed for this kind of problem is a matrix of decisions one individual made with the knowledge of other's decisions.

In our project we are going to implement an approach for finding the influential group in the community.

For doing so we need to have the graph of the network we are dealing with. The improvement that we accomplished is that we are not using some random graphs in this project. We have decided to work with real data gathered from different individuals. We consider social network in Facebook as our data source. Communities are defined as a group of people tagged in a common photo. So the nodes of our graph will be the people who attend the polling and the ones who are tagged in the photos of them. In the next section we will describe how the data is gathered. After that we should just convert the complete graph in the result to the incomplete version according to the answers and calculate the PSNE according to that. In each photo, some question asked about people who were tagged. So there is a clique in every picture but there is no influence in each relation between all two persons in a photo. We add weight to edges according to user's decision. If the states are the same (go-go or don't go-don't go) between two user (voter and friend), the edge's weight increases. After iterating all votes, there is a weighted graph which shows the influence factor in every relations. The people who has influence on others, have more weighted in-edge. We need an unweight graph, so we remove some edges according to ratio of weight to total count of votes. If the ratio is greater than a constant value (from voter U to friend V), it means V has influence on U, so an edge is established from V to U. we generate this graphs with constant value (0.05, 0.1, 0.2, 0.4) and check those on "Gephi". The results are attached.

$$\frac{Wu \rightarrow v}{Total\ Votes\ of\ u} > CONST$$

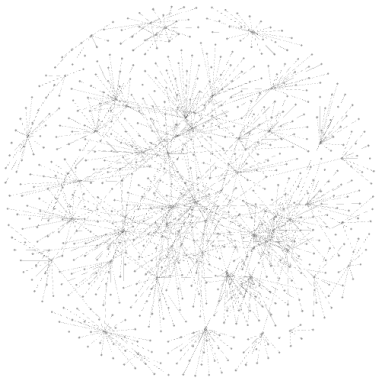


Figure 1 Influenced Graph with CONST = 0.4

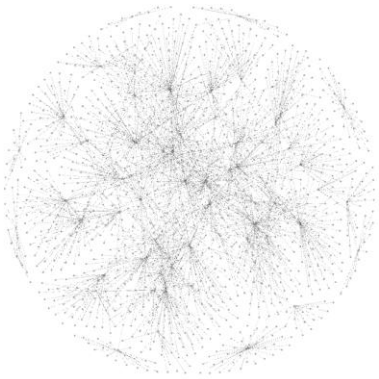


Figure 2 Influenced Graph with CONST = 0.05

Some people (who were tagged in picture) didn't vote, so there is no any evidence to determine their ideas. We can merge them with their local friend.

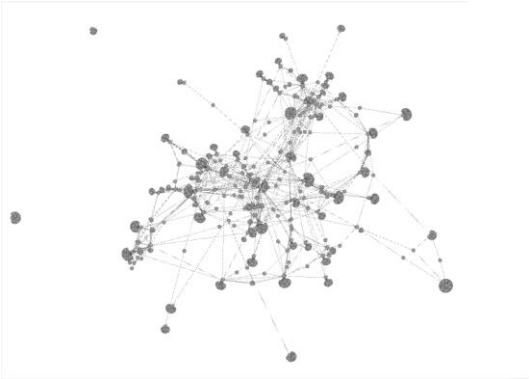
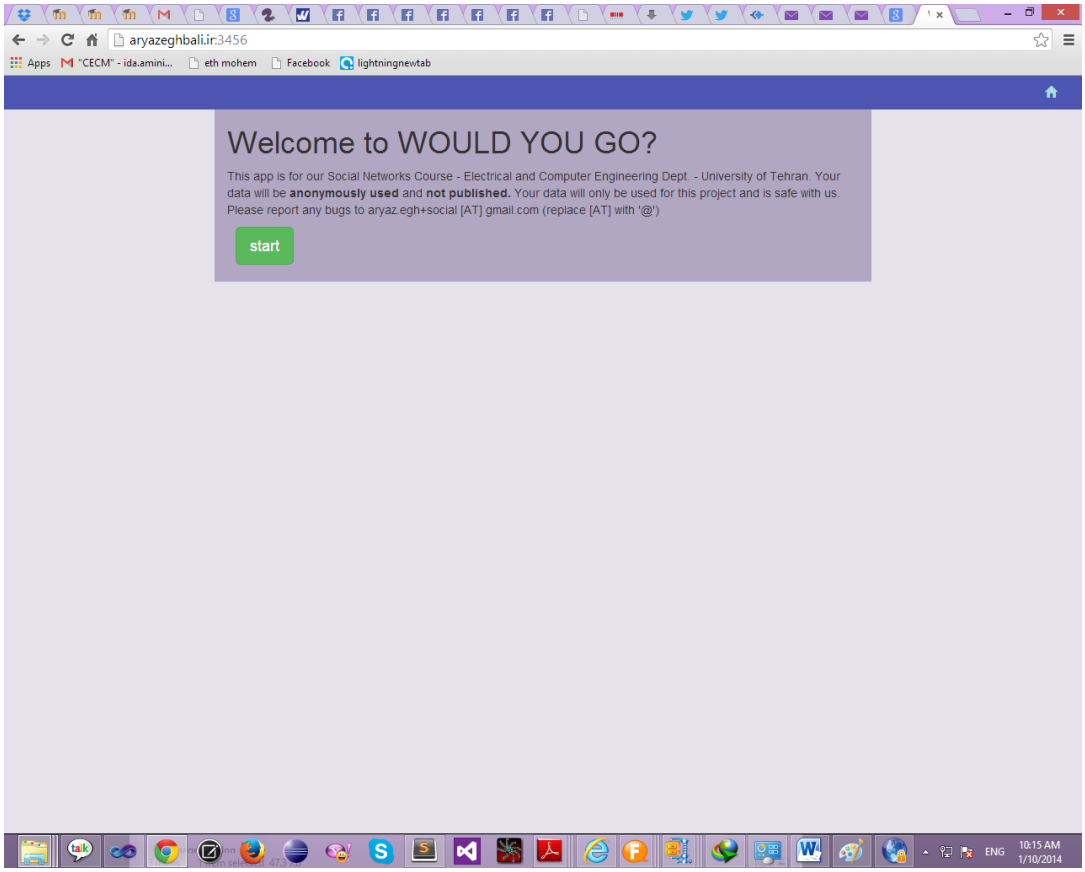


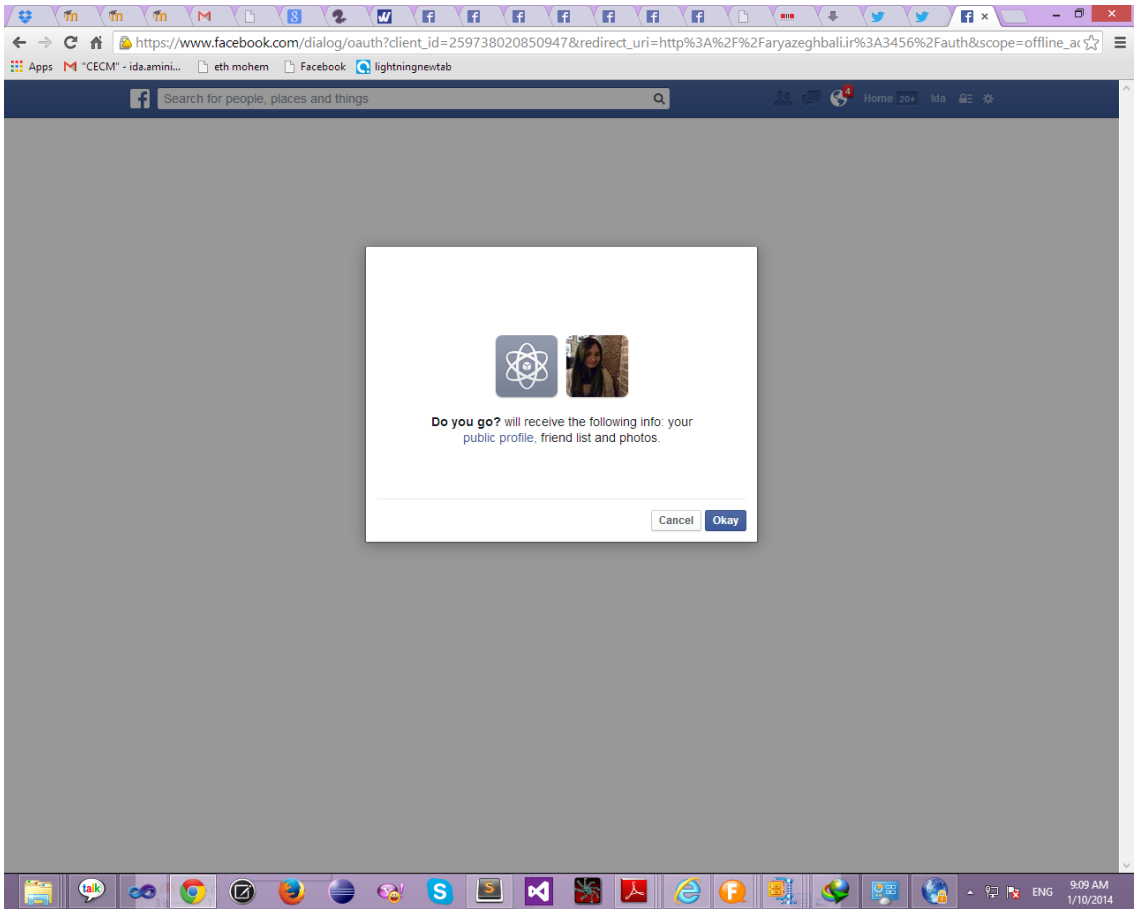
Figure 3 Influenced Graph with cluster users who didn't vote

Gathering data: The data that we should use in this project is people's decision to attend some event with the knowledge of the decisions of their other friends. In order to have this kind of information and as we couldn't find any source for this kind of data, we have implemented an application over Facebook. This application is available at **aryazeghbali.ir:3456**. This application works in this manner that it gathers all one's profile and it reaches photos in that profile. Then we monitor all the tags in the picture and we ask questions in order to fill our database. For the simplicity purpose for the people to attend this questioning, we decide to just extract the pictures that have between 3 and 8 tags. The question are in this manner that they name a place and they give information of decisions of other names which are tagged in the photo. And answers are either yes or no.

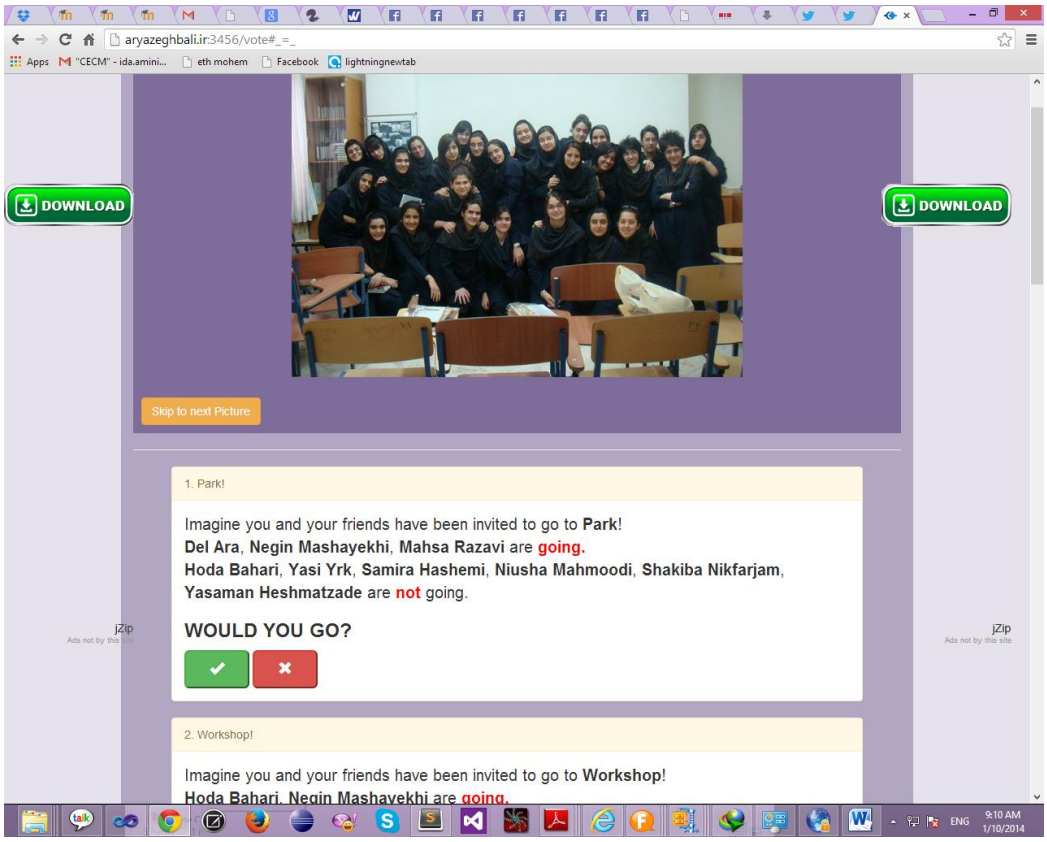
This procedure id describe below. First the person goes to the website aryazeghbali.ir:3456.

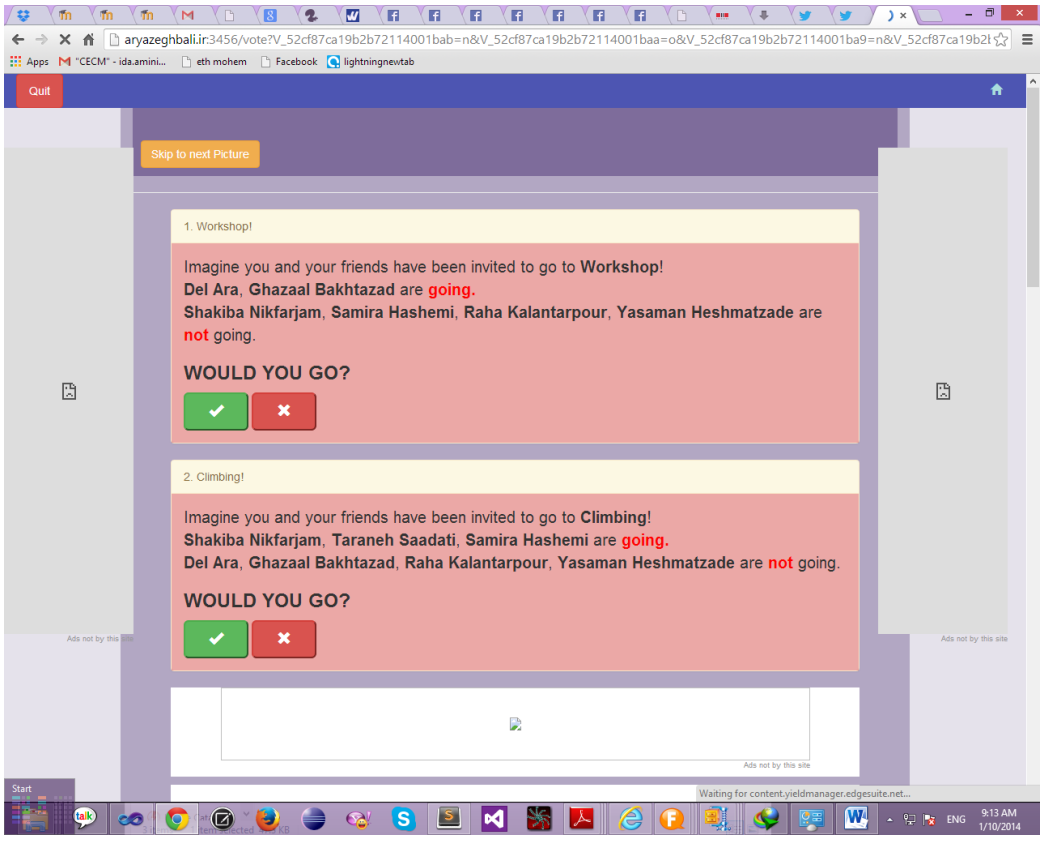


Then if he/she decides to continue the process the page is redirected to Facebook in order to ask for the person's permission. After getting the permission we are half way there.

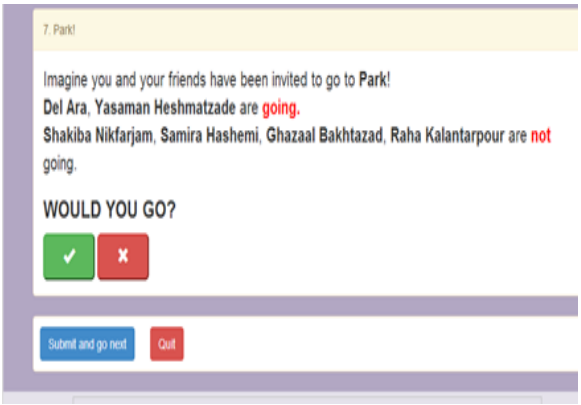


The only remaining job for us is to extract the pictures and for the person is to answer them. All the answers she/he gives for the question is recorded in our database. We ensure the person that all of his/her choices remain private and no one is going to use them in any way.


aryazeghbali.ir:3456/vote#_=_
Apps
"CECM" - ida.amini...
eth mohem
Facebook
lightningnewtab
DOWNLOAD
DOWNLOAD
Skip to next Picture
1. Park!
Imagine you and your friends have been invited to go to Park!
Del Ara, Negin Mashayekhi, Mahsa Razavi are going.
Hoda Bahari, Yasi Yrk, Samira Hashemi, Niusha Mahmoodi, Shakiba Nikfarjam, Yasaman Heshmatzade are not going.
WOULD YOU GO?
jZip
Ads not by this site
2. Workshop!
Imagine you and your friends have been invited to go to Workshop!
Hoda Bahari, Negin Mashayekhi are going.
ENG
9:10 AM
1/10/2014


aryazeghbali.ir:3456/vote?V_52cf87ca19b2b72114001bab=n&V_52cf87ca19b2b72114001baa=o&V_52cf87ca19b2b72114001ba9=n&V_52cf87ca19b2b
Apps
"CECM" - ida.amini...
eth mohem
Facebook
lightningnewtab
Quit
Skip to next Picture
1. Workshop!
Imagine you and your friends have been invited to go to Workshop!
Del Ara, Ghazaal Bakhtazad are going.
Shakiba Nikfarjam, Samira Hashemi, Raha Kalantarpour, Yasaman Heshmatzade are not going.
WOULD YOU GO?
2. Climbing!
Imagine you and your friends have been invited to go to Climbing!
Shakiba Nikfarjam, Taraneh Saadati, Samira Hashemi are going.
Del Ara, Ghazaal Bakhtazad, Raha Kalantarpour, Yasaman Heshmatzade are not going.
WOULD YOU GO?
Ads not by this site
Start
Waiting for content.yieldmanager.edgesuite.net...
ENG
9:13 AM
1/10/2014

If the bottom quit is pressed the questioning is over and the data until that time is saved.



After this questioning process we will have the array of n*k per photo. n stand for the number of the questions that asked for one question and k stands for the number of the people who are tagged in those photos. Each cell of this array is either 1 or 0 that represents the person's decision wither to attend the event or not. Now we got all of our data and we can run the algorithm described below on it.

Future Works: As it has been mentioned before, we focused on using different people's opinion based on whether their friends choose to go somewhere or not. But there can be more arguments to this question. The most obvious one can be the place this people are invited too. For example people are more likely to enjoy spending time with a sportsman in a gym or by going to hiking than going to cinema. Also it can be known that some people are not interested in attending some sort of programs, regardless of the people who participate in that program too. For example some people don't enjoy going to movies, so they won't go to cinema regardless of the people who are invited and are going or not going to this program. Also some people don't get biased by the fact that someone is attending a program or not. All of the mentioned indicators can be taken into account to assign a more precise "influence factor" to anyone.

Although our project has been built around a simple questionnaire and the resulting dataset offers information regard people's friendship relationships, we can gather information in different formats and apply them to different situations. Some simple examples are mentioned below:

1- Consider you want to persuade some people to vote in your favor in a particular voting system. You should choose to persuade people who have the most "influence factor". This dataset can be gathered by observing the

people who had been in more winning parties in different voting situations. This can be easily done in parliaments.

2- Companies can use "influence factor" of popular people to persuade people to buy their products, the data gathering method can be through gathering the selling statistics after using different celebrities in advertisements, or even by checking these people's amicability in their popularity topic.

3- If one can persuade a friend to quit smoking, the more influence he has in his smoking community, the more people are likely to quit smoking after him. Gathering data for this situation can be easily done by checking friends who usually smoke together, or even friends who had been the cause of others starting to smoke.

All in all, there are many different ways to continue this project, ranging from using different indexes for calculating influence factor, to using our approaches or suggested ideas to find out people's influence factor in different communities, in which the relation between the people is not necessarily a friendship relationship, and use this factor according to the needs and goals.

References:

1- Mohammad T. Irfan and Luis E. Ortiz. A Game-Theoretic Approach to Influence in Networks. 25th AAAI Conference on Artificial Intelligence (AAAI), San Francisco, CA, 2011.

2- Kohli, Pushmeet, et al. Behavioral Game Theory on Online Social Networks: Colonel Blotto is on Facebook. mimeo, 2011.

3- Gomez, Daniel, et al. "Centrality and power in social networks: a game theoretic approach." Mathematical Social Sciences 46.1 (2003): 27-54.

4- Xu, Haiqing. "Social Interactions: A Game Theoretical Approach." UT-Austin mimeo (2011).

5- Easley, David, and Jon Kleinberg. Networks, crowds, and markets. Vol. 8. Cambridge: Cambridge University Press, 2010.

6- Nisan, Noam, ed. Algorithmic game theory. Cambridge University Press, 2007.

7- Leyton-Brown, Kevin, and Yoav Shoham. "Essentials of game theory: A concise multidisciplinary introduction." Synthesis Lectures on Artificial Intelligence and Machine Learning 2.1 (2008): 1-88.

8- M. E. J. Newman, Communities, modules and large-scale structure in networks, Nature Physics: 25-31, 22 Dec 2011

9- P. Esfandiar, Application of Evolutionary Game theory to Social Networks: A Survey ,2009