Discrete Mathematics, Algorithms and Applications

Efficient Respondents Selection for Biased Survey using Homophily-high Social Relation Graph *†

Donghyun Kim

Department of Computer Science, Kennesaw State University, Marietta, GA 30060, USA. donghyun.kim@kennesaw.edu

Jiaofei Zhong

Department of Mathematics and Computer Science, California State University, East Bay, Hayward, CA 94542, USA. jiaofei.zhong@csueastbay.edu

Minhyuk Lee

Department of Mathematics and Physics, North Carolina Central University, Durham, NC 27707, USA. mlee28@eagles.nccu.edu

Deying Li

Key Laboratory of Data Engineering and Knowledge Engineering (Renmin University), MOE, School of Information, Renmin University of China, P.R. China. deyingli@ruc.edu.cn

Yingshu Li

Department of Computer Science, Georgia State University, Atlanta, GA 30303, USA. yli@cs.gsu.edu

Alade O. Tokuta

Department of Mathematics and Physics, North Carolina Central University, Durham, NC 27707, USA. atokuta@nccu.edu

Online social relationships which can be extracted from various online resources such as online social networks are getting much attention from the research communities since they are rich resources to learn about the members of our society as well as the relationships among them. With the advances of Internet related technologies, online surveys are established as an essential tool for a wide range of applications. One significant issue of online survey is how to select a quality respondent group so that the survey result is reliable. This paper studies the use of pair-wise online social relationships among the members of a society to form a biased survey respondent group, which might be useful for various applications. We first introduce a way to construct a homophily-high social relation graph. Then, we introduce the *minimum inverse k-core dominating set problem*

*The preliminary version of this paper was appeared in CSoNet 2014 [1].
†This work was supported in part by US National Science Foundation (NSF) CREST No. HRD-1345219. This paper was jointly supported by the Fundamental Research Funds for the Central Universities, and the Research Funds of Renmin University of China 10XNJ032.

(MIkCDSP), which aims to compute a biased respondent group using the homophily-high social relation graph. We show the problem is NP-hard and most importantly propose a greedy approximation for it. Our simulation based on a real social network shows the proposed algorithm is very effective.

Keywords: dominating set, social networks, approximation algorithm, k-core, vertex connectivity, homophily

1. Introduction

These days, online surveys are established as an essential tool for a wide range of applications such as marketing and political decision making. It is known that in 2006, around 20% of global data-collection expenditure was spent for online survey research [2]. In 2012, US spent more than $1.8 billion for all survey research spending [3]. There are a number of reasons, not to mention its low cost (than the traditional methods), that online survey becomes so popular [4]. In online survey researches, how to find a right sample group of respondents is a long lasting conundrum since this is directly related to the reliability of the survey.

Frequently, a biased survey respondent group is considered to be lack of reliability. This is because the result from a survey is mainly used to obtain an averaged statistical information about the general public by consulting with a sample group from the public, and the survey result from a sample group without representativeness is not reliable for this purpose. Due to the reason, many efforts are made to find a representative and unbiased respondent group [3]. However, the bias in the survey is not always something to avoid. Consider a product quality manager of a new smartphone, e.g. iPhone 5c, who wants to collect the feedback via an online survey from users so that he/she can improve the quality of the product. Also, suppose most of the customers using the product are happy with it. Then, while the manager is only interested in hearing complaints from the users, it is likely that the online survey result from the respondents selected by the methods whose common goal is to make the result representative and unbiased, would be mostly about their satisfaction about the new product. As a result, such survey is quite wasteful in practice to the manager only interested in complaints. Therefore, it would be helpful to form a biased respondent group so that it includes more unsatisfied users.

Recently, online social relationships which can be extracted from various online resources such as online social networks are receiving lots of attentions as they are rich sources to learn about the interest of each member as well as the relationship among them. Due to the reason, the online resources are investigated for a wide range of applications [5,6,7,8,9,10,11,12]. This paper investigates the use of pairwise online social relationship information to compute a biased but representative respondent group such that the rate of the minority opinion group (e.g. those who are not satisfied with the product) in the respondent group can be magnified. To the best of our knowledge, this is the first effort in the literature which exploits online social relationships to enhance to the quality of online minority opinion survey.

The rest of this paper is organized as follows. Section 2 describes how to construct a desirable input of the proposed algorithm, a homophily-high social relation graph, in which two nodes are adjacent only if they share similar opinion on a subject of interest from readily available pair-wise online social relationships. In Section 3, we introduce several important notations and definitions. Especially, we introduce the formal definition of our problem of interest, the *minimum inverse k-core dominating set problem (MIkCDSP)*, corresponding justification, and its NP-hardness result. Section 4 proposes a new greedy approximation algorithm for MIkCDSP by exploiting the homophily-high graph. Some case study simulation result and analysis using a real social network is presented in Section 5. Finally, we conclude this paper in Section 6.

2. Construction of Homophily-high Social Relation Graph

In this section, we introduce a new approach to compute a very special social relation graph from readily available pair-wise social relations. The resulting graph has the following two interesting features. First, any two adjacent nodes in the graph are likely to have a similar opinion on the subject of interest. In other word, the homophily [13] of the graph should be high. Second, the average degree of nodes in the majority opinion group should be higher than the average degree of the whole graph.

Our proposed method construct a new social relation graph as follows. Given a subject of interest and the members of a society, an existing data mining techniques such as [14] is used to compute the pair-wise social proximity among the members on the subject. The accuracy of this step is out of the scope of this paper and we just assume their existence. Then, for each member of the society, we add a corresponding node to the graph. Next, for each pair of members, we set an edge between them only if their opinion is similar on the subject than a threshold level. Clearly, the resulting graph is with higher homophily as the threshold becomes smaller. Note that if we have a majority opinion group in the whole society, their average node degree should be higher than the average degree of the graph as they have more neighbors. Therefore, the graph constructed in this way strictly satisfies the two properties.

By the construction, it is apparent that any two nodes whose opinions are significantly different are not adjacent in the resulting graph. Now, we verify that the constructed graph satisfies the second property that we discussed earlier via a simulation. In this simulation, we have 1000 nodes, and each node has a weight between 0 to 1, but the distribution strictly follows exponential distribution, and therefore, most of them has a weight closer to 0. Then, in the graph constructed by following the method above with threshold value 0.05, average node weight is 0.083422 and average node degree is 445.37. Meanwhile, the number of nodes whose node weight in the generated graph is smaller than the average (0.083422) is $613 > (1000/2)$, and the average degree of those nodes is $561.151713 > 445.37$. This means that the nodes with weight smaller than average forms the majority and at the same time, the **majority group's average node degree is higher than the average node degree** in G. As a result, G satisfies these two requirements.

3. Notations, Definitions, and Problem Statement

In this paper, $G = (V, E)$ represents a homophily-high social relation graph with a node set $V = V(G)$ and an edge set $E = E(G)$. We assume the relationship between the members are symmetric and thus the edges in E are bidirectional. Also, we use n to denote the number of nodes in V, i.e. $n = |V|$. For any subset $D \subseteq V$, $G[D]$ is a subgraph of G induced by D. For each node $v \in V$, $N_{v,V}(G)$ is the set of nodes in V neighboring to v in G. Given a graph G, a subset $D \subseteq V$ is a *dominating set (DS)* of G if for each node $u \in V \setminus D$, $\exists v \in D$ such that $(v, u) \in E$. Given a graph G, the goal of the *minimum dominating set problem (MDSP)* is to find a minimum size DS of G. Given a graph G, a subset $D \subseteq V$, and a positive integer k which is no greater than Δ, the maximum node degree of G, D is an *inverse k-core* in G if for each $v \in D$, $|N_{v,D}(G)| \leq k$.

In this paper, we study an online survey sample (a group of survey respondents) selection problem such that the rate of minority opinions from the sample can be higher than their rate in the overall group. In G, by construction, a subset of nodes whose average degree is smaller has a better chance to include higher rate of minority opinion holders. Meanwhile, the concept of DS has been widely used to select a quality representative group for the whole society (e.g. clusterheads in wireless networks) in many existing researches. Therefore, it would be desirable to compute a DS such that the average degree of the DS nodes in G is minimized. However, there is one drawback of this approach. That is, there can be more than one minority opinion in the society. Therefore, it is more desirable to limit the number of the neighbors of each DS node in G is limited in the whole DS. Given a homophily-high social relation graph G, a subset $D \subseteq V$, and a positive integer k, D is an *inverse k-core dominating set (IkCDS)* of G if (a) D is a DS of G and (b) for each $v \in D$, $|N_{v,D}(G)| \leq k$. Certainly, IkCDS is suitable for our purpose based on the discussions so far. Meanwhile, it is noteworthy that there are a number of ways to compute IkCDS of a social relation graph. Apparently, it is more desirable to reduce the size of IkCDS since it will cost less for the actual survey. As a result, the problem of computing a biased online survey respondent group can be formulated as MIkCDSP shown below. Given a graph G and a positive integer k, a *minimum inverse k-core dominating set (MIkCDS)* is an IkCDS of G of minimum cardinality. Finally, the *minimum inverse k-core dominating set problem (MIkCDSP)* aims to find a minimum size IkCDS of G.

Remark 3.1. It is noteworthy that as k decreases, the DS will be more diversified and biased and the rate of minority opinion in the survey will increase. At the same time, the size of the IkCDS will decrease. This means that with very small k value, the survey respondent set can be very small and less practical given that the usual degree of social relation graphs is not small. On the other hand, with very high k

Algorithm 1 Greedy-MIkCDSA ($G = (V, E), k$)

Prepare an empty set D, i.e. $D \leftarrow \emptyset$.
For each $v_i \in V$, prepare a counter n_i which is initialized to 0, i.e. $n_i \leftarrow 0$.
Suppose $X_j = \{v_i | v_i \in V \text{ and } n_i = j\}$.
while $X_0 \neq \emptyset$ **do**
  Find $v_i \in V \setminus \left((\bigcup_{j \geq k} X_j) \bigcup D \bigcup Q\right)$ so that $|N_{v_i, X_0}(G)|$ is maximized, where $Q = \{w_1, \ldots, w_q\}$ such that $w_l \in Q$ has at least one neighbor in $(\bigcup_{j \geq k} X_j)$ and $w_l \in D$ is true. A tie can be broken arbitrarily.
  Set $D \leftarrow D \cup \{v_i\}$.
  for each node $v_j \in N_{v_i, V}(G)$ **do**
    $n_j \leftarrow n_j + 1$.
  end for
end while
Output D.

value, the survey respondent group can be negligibly biased, which also may not be desirable for our application. While selecting proper k value is very significant, it is also application dependent. Since this question is the out of this paper, we assume that k value is given as a part of the inputs of the problem.

Theorem 3.2. *MIkCDSP is NP-hard.*

Proof. A special case of MIkCDSP with $k = \Delta$ is equivalent to the minimum dominating set problem, the problem of computing a minimum size dominating set of G, which is proven to be NP-hard [15]. As a result, MIkCDSP is NP-hard. □

Remark 3.3. Given any graph G and a non-negative integer k, there exists a feasible solution of MIkCDSP in G. This claim is true since (a) a feasible solution of MIkCDSP with $k = 0$ is clearly a feasible solution of MIkCDSP with any $k \geq 1$ (as the condition that each DS node cannot have any neighboring DS node is stricter than the condition that each DS node can have at most $k \geq 1$ DS node(s)), and (b) the following coloring strategy can be used to compute an independent set of G, the subset of nodes in G which are pairwise disjoint with each other, which is a feasible solution of MIkCDSP with $k = 0$: (i) initially color all nodes white, (ii) pick each white node turn into black and its neighbors in gray until there is not white node left, and (iii) return the set of black nodes. Clearly, the set of black nodes is a dominating set and each pair of black nodes are not neighboring from each other.

4. Greedy-MIkCDSA: A Greedy Approximation for MIkCDSP

In this section, we introduce Greedy-MIkCDSA, a simple greedy strategy for MIkCDSP and show that its performance ratio is Δ, where Δ is the maximum degree of the input social relation graph. The formal description of Greedy-MIkCDSA is Algorithm 1. Given an MIkCDSP instance (G, k), Greedy-MIkCDSA first prepares an empty set D (Line 1), which will eventually include the output, an inverse k-core dominating set (IkCDS) of G. For each node $v_i \in V$, we create a counter n_i which is initialized to 0 (Line 2). The counter will be used to track the number of neighbors of v_i in D. Depending on the counter, we create a partition of the nodes in V, $X_0, X_1, \ldots$, where X_j is the subset of nodes in V whose counter is j (Line 3). This means that initially X_0 is equal to V and each of the rest is empty. Clearly, the number of the subsets is bounded by n. From Lines 4-10, we iteratively pick a node v_i from $\left((\bigcup_{j\geq k} X_j)\bigcup D \bigcup Q\right)$, i.e. v_i is a node which is (a) with a counter n_i whose value is less than k (i.e. has less than k neighbors in DS), (b) not selected as a DS node yet, and (c) without any neighboring node w_l which is in D and, at the same time, in X_j for some $j \geq k$, such that the number of neighbors of v_i in X_0 is the maximum. Any tie can be broken arbitrarily. This loop is repeated until all nodes in V is either in D or dominated by some node in D while maintaining $G[D]$, the subgraph of G induced by D, as an inverse k-core.

Clearly, Algorithm 1 produces a feasible solution of MIkCDSP since the algorithm repeatedly constructs D until X_0 becomes empty (which means D is a DS of G) and by Line 5, the degree of $G[D]$, the graph induced by D in G, will be bounded by k (which means D is an inverse k-core). One may wonder if there is a situation in which some node x, which has to be included in D to dominate some other node y, cannot be included in D since it has already k neighbors in D. However, this never becomes a problem since if x cannot be selected, then y itself will be included in D by our algorithm, which means that D is always a valid output. Now, we show Algorithm 1 is a Δ-approximation algorithm for MIkCDSP.

Lemma 4.1. *Given a graph $G = (V, E)$, let OPT_{MDSP} and OPT_{MIkCDS} be an optimal solution of MDSP and an optimal solution of an MIkCDSP instance (G, k) for some $k \geq 1$, respectively. Then, $|OPT_{MDSP}| \leq |OPT_{MIkCDS}|$.*

Proof. By definitions, the goal of MDSP is to find a DS of G with minimum cardinality and the goal of MIkCDSP is to find a DS of G with minimum cardinality such that for each node in the DS, the node is allowed to be adjacent with at most k other nodes in the DS. Therefore, in any given G, an IkCDS of G is also a DS of G, but our choice of IkCDS is more limited than that of DS. As a result, this lemma is true. □

Lemma 4.2. *Given a graph $G = (V, E)$, suppose we have an α-approximation algorithm of MDSP such that the output O of the algorithm is also a feasible solution of MIkCDSP. Then, we have $|O| \leq \alpha |OPT_{MIkCDS}|$.*

Proof. As discussed in the proof of the previous lemma, clearly, an output O of Algorithm 1 is a dominating set and therefore a feasible solution of MDSP. By the definition, if Algorithm 1 is an α-approximation algorithm of MDSP, then we

Fig. 1. All nodes in P_i is adjacent to o_i. P_i may contain several nodes in D, which is an output of Algorithm 1.

have $|O| \leq \alpha|OPT_{MDSP}|$. By combining this with Lemma 4.1, we have $|O| \leq \alpha|OPT_{MDSP}| \leq \alpha|OPT_{MIkCDS}|$, and thus this lemma is true. □

Now, we show that this α is Δ (Lemma 4.3) and therefore, Algorithm 1 is an Δ-approximation algorithm for MIkCDSP (Theorem 4.4).

Lemma 4.3. *The performance ratio of Algorithm 1 for MDSP is Δ, where Δ is the maximum degree of G.*

Proof. Given $G = (V, E)$ and k, consider $OPT_{MDSP} = \{o_1, o_2, \ \dots\ , o_l\}$ be a minimum DS of G. Then, for each $o_i \in OPT_{MDSP}$ in the increasing order of i, we compute

$$P_1 = \{o_1\} \bigcup N_{o_1, V \setminus OPT_{MDSP}}(G),$$

and

$$P_i = \Big(\{o_i\} \bigcup N_{o_i, V \setminus OPT_{MDSP}}(G)\Big) \setminus \Big(\bigcup_{1 \leq j \leq i-1} P_j \Big) \text{ for } i \neq 1.$$

That is, P_1 is the union of the first node in the optimal solution OPT_{MDSP} and its neighbors except the other nodes inside OPT_{MDSP} in G. Likewise, P_i is the union of the i-th node in the optimal solution OPT_{MDSP} and its neighbors except the other nodes inside OPT_{MDSP} as well as except those in $P_1 \cup P_2 \cup \ \cdots\ \cup P_{i-1}$ in G. Clearly, V is partitioned into $\mathcal{P} = \{P_1, P_2, \ \dots\ , P_l\}$ such that each $P_i \in \mathcal{P}$ exactly includes one $o_i \in OPT_{MDSP}$.

Suppose Algorithm 1 is applied to (G, k) and outputs $D = \{z_1, z_2, \ \dots\ , z_p\}$ (see Fig. 1). Then, we have the following two important observations. First, all nodes in P_i is adjacent to o_i. Second, each P_i can include some nodes in D (e.g. $\{z_{q_1}, z_{q_2}, z_{q_3}\}$ in Fig. 1). Therefore, the size of $P_i \bigcap D$ is bounded by Δ (the node degree of o_i in

(a) Minimum dominating set.

(b) Minimum inverse k-core dominating set.

Fig. 2. Selected dominating nodes by each strategy.

$G[P_i]$), and we have $|D| \leq \max_{1 \leq i \leq l} |P_i \cap D| \cdot |OPT_{MDSP}| = \Delta \cdot |OPT_{MDSP}|$, and this lemma is true. □

Theorem 4.4. *The performance ratio of Algorithm 1 for MIkCDSP is Δ, where Δ is the maximum degree of G.*

Proof. The proof naturally follows from Lemma 4.2 and Lemma 4.3. □

5. Case Study

In this section, we evaluate the performance of Algorithm 1 using one real social network example, the jazz musician social network [16] under the assumption that this social network graph satisfies the two properties discussed in Section 2. This social network with undirected edges represents the collaboration network of jazz musicians. In this simulation, Algorithm 1 is compared with an existing approximation algorithm for the minimum dominating set problem by Guha and Khuller [17] since it can be one trivial way to solve MIkCDSP and there is no existing algorithm designed directly to compute MIkCDSP. Fig. 2 shows the output of each algorithm with the jazz musician social network graph as an input. Note that the subset of dark blue nodes constitutes the dominating set (output). We acknowledge that a social network visualization tool called Cytospace [18] is used for this visualization. Following our discussion in the earlier part of this paper, we assume a node with a higher degree is more likely to be a member of the majority group. In Fig. 2, it is clear that Algorithm 1 outperforms the minimum connected dominating set algorithm in terms of average node degree. Meanwhile the size difference of the dominating set is not significant (14 vs. 16 out of hundreds nodes).

In Fig. 3, we evaluate the performance of Algorithm 1 by comparing the average degree of the dominating nodes generated by Algorithm 1 with that of the dominating nodes by the minimum dominating set (approximation) algorithm by Guha and

Fig. 3. Cumulative degree distribution of (a) InputGraph: the all nodes in the input graph, (b) MIkCDS: the nodes in the output of our Algorithm 1 with $k = 6$, and (c) MCDS: the nodes in the output of an existing approximation algorithm for the minimum dominating set problem by Guha and Khuller.

Khuller [17] as well as the average node degree of whole graph. In this figure, x-axis represents the degree of nodes and y represents the cumulative distribution of the nodes (up to the specified degree) in the dominating set (or the whole graph). In this simulation, note that, we fix k to 6. From this figure, we can clearly learn that Algorithm 1 successfully select a biased (minority) representative subgroup (dominating set). In detail, the percentage of the nodes whose degree is no greater than 10 in the output of MIkCDSA is almost 50% while that of the nodes in the output of the minimum dominating set algorithm is around 20% which is similar to the graph average. As a result, in case of the jazz musician social network, Algorithm 1 effectively constructs a biased (minority) representative group.

6. Concluding Remarks

This paper introduces a new approach to use the pair-wise social relationships to enhance the result of minority opinion survey. To perform this task efficiently, we introduce a new NP-hard optimization problem, propose a new greedy heuristic algorithm for it, and show the algorithm in fact has a theoretical performance guarantee. We also introduced a way to compute a desirable input graph for the proposed algorithm. To the best of our knowledge, this is the first attempt in the literature to use online social relation information to improve the result of biased online survey. Our approach requires the expected similarity of the opinions between each pair of users from readily available online social relation to construct the biased respondent group. Therefore, it is rather localized and consumes less resources than analyzing the opinion of each user for a subject and find the minorities, and thus is more practical in big data environment. As a future work, we plan to further investigate the use of online social relationships to improve the reliability of online voting systems as well as other uses of our homophily-high social relation graph.

References

[1] D. Kim, J. Zhong, M. Lee, D. Li, and A.O. Tokuta, "Efficient Respondents Selection for Biased Survey using Online Social Networks," *Proc. of the 2nd Workshop on Computational Social Networks (CSoNet 2014)*, August 4-5, 2014, Atlanta, GA, USA, Co-located with COCOON 2014, Computing and Combinatorics, LNCS 8591, pp. 608-615, 2014.

[2] V. Vehovar and K.L. Manfreda, "Overview: Online Surveys," *The SAGE Handbook of Online Research Methods*, London: SAGE (edited by N.G. Fielding, R.M. Lee, and G. Blank), pp. 177-194, 2008.

[3] G. Terhanian and J. Bremer, "A Smarter Way to Select Respondents for Surveys?," *International Journal of Market Research*, vol. 54, no. 6, pp. 751-780, 2012.

[4] B. Duffy, K. Smith, G. Terhanian, and J. Bremer, J, "Comparing Data from Online and Face-to-face Surveys," *International Journal of Market Research*, vol. 47, no. 6, pp. 615-639, 2005.

[5] F. Wang, K. Xu, and H. Wang, "Discovering Shared Interests in Online Social Networks," *Proc. of The 4th International Workshop on Hot Topics in Peer-to-peer computing and Online Social Networking (HotPOST 2012)*, June 21, 2012, Macau, China.

[6] W. Zhang, W. Wu, F. Wang, and Kuai Xu, "Positive Influence Dominating Sets in Power-Law Graphs," *Social Network Analysis and Mining*, vol. 2, no. 1, pp. 31-37, Mar. 2012.

[7] H. Zhang, T. N. Dinh, and M. T. Thai, "Maximizing the Spread of Positive Influence in Online Social Networks," *Proc. of the IEEE International Conference on Distributed Computing Systems (ICDCS 2013)*, July 8-11, 2013, Philadelphia, PA, USA.

[8] C. Kahl, S. Crane, M. Tschersich, and K. Rannenberg, "Privacy Respecting Targeted Advertising for Social Networks," *Proc. of the 5th IFIP WG 11.2 International Conference on Information Security Theory and Practice: Security and Privacy of Mobile Devices in Wireless Communication (WISTP 2011)*, pp. 361-370, June 1-3, 2011, Heraklion, Crete, Greece.

[9] D. Kim, D. Li, O. Asgari, Y. Li, and A.O. Tokuta, "A Dominating Set Based Approach to Identify Effective Leader Group of Social Network," *Proc. of the 19th Annual International Computing and Combinatorics Conference (COCOON 2013)*, June 21-23, 2013, Hangzhou, China.

[10] D. Kim, H. Guo, Y. Li, W. Wang, S.-S. Kwon, and A.O. Tokuta, "Social Relation Based Long-term Vaccine Distribution Planning to Suppress Pandemic," *Proceeding of the 4th International Conference on Computational Social Networks (CSoNet 2015)*, August 4-6, 2015, Beijing, China.

[11] D. Kim, W. Wang, M. Tetteh, J. Liang, S. Park, and W. Lee, "Biased Respondent Selection under Limited Budget for Minority Opinion Survey," *Proceeding of the 4th International Conference on Computational Social Networks (CSoNet 2015)*, August 4-6, 2015, Beijing, China.

[12] H. Ma, Y. Zhu, D. Li, D. Kim, and J. Liang, "Improving the Influence under IC-N Model in Social Networks," *Discrete Mathematics, Algorithms and Applications (DMAA)*, vol. 7, no. 3, 2015.

[13] H. Bisgin, N. Agarwal, and X. Xu, "A Study of Homophily on Social Media," *World Wide Web*, vol. 15, no. 2, pp. 213-232, 2012.

[14] B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis," *Foundations and Trends in Information Retrieval*, vol. 2, no. 1-2, pp. 1-135, 2008.

[15] M.R. Garey and D.S. Johnson, "Computers and Intractability: A Guide to the Theory of NP-completeness," *Freeman*, San Francisco, 1978.

[16] P.Gleiser and L. Danon, "Community Structure in Jazz," *Advances in Complex Systems*, vol. 6, no. 4, pp. 565-573, 2003.
[17] S. Guha and S. Khuller, "Approximation Algorithms for Connected Dominating Sets", *Algorithmica*, vol. 20, pp. 374-387, 1998.
[18] Download website, *http://www.cytoscape.org/download.html*