ICEB 2012 Proceedings

International Conference on Electronic Business (ICEB)

Fall 10-12-2012

Influence Optimization in a Social Network with Stochastic Costs for Targeting Individuals

Qiaoni Shi

Yaodong Ni

Influence Optimization in a Social Network with Stochastic Costs for Targeting Individuals

Qiaoni Shi, Yaodong Ni
School of Information Technology and Management
University of International Business and Economics, Beijing 100029, China
yufuning@163.com

Abstract: In this paper, we study the problem of targeting a set of individuals to trigger a cascade of influence in a social network such that the influence diffuses to all individuals with the minimum time, given that the cost of initially influencing each individual is with randomness and the budget available is limited. We adopt the incremental chance model to characterize the diffusion of influence, and propose three stochastic programming models that correspond to three different decision criteria. Modified greedy algorithms are presented to solve the proposed models efficiently. Experiments are performed to show that the algorithms we present are effective.
Keywords: social networks; complete influence time; stochastic costs; greedy algorithms.

§1 Introduction

A social network is a graphical model of the relationships within a group of individuals, in which each node represents an individual and each edge denotes the relationship between the two individuals it connects. Social networks have attracted much attention from various areas, such as computer science [9], economics [3], sociology [6, 22], epidemiology [8].

The diffusion of influence in social networks plays an important role in many real applications [14, 19, 20]. For example, in virus marketing, a new product is promoted to only a small number of people, but can be adopted by a large proportion of the population. This is because the recommendation of the product from one friend to another is actually a type of influence spread throughout the social network. A variety of models characterizing how influence diffuses have been proposed, such as the linear threshold model [15], the independent cascade model [12], the voter model [10] and the incremental chance model [19]. Motivated by the need for solving practical problems, decision makers attempt to optimize the target set of individuals to whom the influence is initially sent. A number of optimization problems with different objectives have been investigated, including maximizing the total number of influenced individuals [15, 16], maximizing the expected lift in profit [9, 21], minimizing the size of the initial target set of individuals that guarantees complete influence [5]. A recently proposed problem is the complete influence time (CIT) minimization problem [19, 20], which is to determine the set of individuals to target initially such that the time taken to influence all the individuals in the network is minimized.

The restriction of limited resources almost always exists in real world. For example, in the context of this paper, the budget for influencing the initial target set of individuals is almost always limited. In a number of papers on influence diffusion in social networks [15, 19], the restriction of limited resources is represented by the assumption that the target set

Proceedings of the Twelfth International Conference on Electronic Business, Xi'an, China, October 12-16, 2012, 263-272.

has a fixed size. In other works [10, 20], a more general scenario is considered, where a predetermined amount of budget is given and the cost of initially influencing each individual may be different.

In real life, uncertainty always exists. In the problem we are concerned with, the cost to initially influence an individual may be uncertain because it usually depends on many dynamic or vague factors such as how the mood of the individual is when she is influenced, to what extend the individual favors the influence, as so on. It would be inappropriate to quantify the costs as constants in such an uncertain situation. However, uncertain costs of influencing the target set of individuals have not been discussed in any of the above works.

In this paper, we focus on the CIT minimization problem where a predetermined amount of budget is provided and the cost of initially influencing each individual is given as a random variable. We adopt the incremental chance model to describe the diffusion of influence, because compared to that under other models, the diffusion process is progressive under the incremental chance model, which guarantees the complete influence that is required by the CIT minimization problem.

Since random variables are involved in the optimization problem we study, we consider three different decision criteria in the theory of stochastic programming, and propose three stochastic programming models. In order to solve the stochastic programming models, we present a number of modified greedy algorithms, which flexibly trade off between solution performance and computational complexity. A number of experiments are performed, by which we show that the modified greedy algorithms are effective.

The rest of this paper is organized as follows. Firstly, we briefly introduce some preliminaries of the incremental chance model in section 2. Then in section 3, we propose three stochastic programming models. Section 4 describes the modified greedy algorithm. We perform experiments in section 5 to show the effectiveness of the proposed algorithms and draw conclusions in section 6.

§2 The Incremental Chance Model

We denote a social network by a weighted undirected graph $G = (N, E, W, \boldsymbol{\varepsilon})$. Each node in G represents one individual in the group we consider, and N denotes the set of all nodes in G. Each edge in G represents a relationship between the pair of individuals corresponding to the two nodes the edge connects, and the set of all edges is denoted by E. For any two vertices i and j in N, the edge connecting them, if there exists one, can be denoted by either (i, j) or (j, i), since G is an undirected graph. In the following of this paper, we call that i and j are neighbors if there exists an edge (i, j). We denote by W the weight function which assigns a positive weight to each edge to measure the strength of the relationship the edge represents. For example, a large weight $W(i, j)$ on edge (i, j) shows that i and j are closely related.

The cost for initially influencing node i in G is assumed as a random variable ε_i, and $\boldsymbol{\varepsilon} = \{\varepsilon_i | i \in N\}$ is a random vector recording the stochastic costs for initially influencing all the individuals. The target set of individuals who are chosen to influence initially is denoted by X, which we equivalently represent by a binary vector $\boldsymbol{x} = \{x_i | i \in N\}$, where $x_i = 1$ if node i is initially targeted or $x_i = 0$ otherwise. Therefore, the cost for initially influencing the target set $\boldsymbol{x}$ is

$$C(\boldsymbol{x}, \boldsymbol{\varepsilon}) = \sum_{i \in N} \varepsilon_i x_i.$$

According to the incremental chance model [19], at any time step, each node is in either of the two states: "active" and "inactive". At any time step, a node that has been influenced is called an active node, or otherwise called an inactive node. At any time step t, an inactive node j turns to be active with the probability

$$p_t^j = \frac{\sum_{i \in N_t^a(j)} W(i,j)}{\sum_{i \in N(j)} W(i,j)},$$

where $N(j)$ denotes the set of neighbors of node j, and $N_t^a(j)$ is the set of active neighbors of j at time step t. We call p_t^j the influence chance of j at time t. Although the process of the diffusion of influence is stochastic, it is still clear to describe. At the beginning, the nodes in the target set are influenced and set active first. Then the cascade of influence is triggered, and inactive nodes get influenced according to their influence chance. Note that the process is progressive, that is, active nodes will stay active and never turn to be inactive. Thus, the diffusion of influence will always terminate as all nodes in the social network will turn to be active ultimately with probability 1 [19], which achieves the complete influence. The duration of the process of influence diffusion is called the complete influence time (CIT). We denote the CIT by $\tau(\boldsymbol{x})$ given that the target set of nodes is $\boldsymbol{x}$.

§3 The Stochastic Programming Models

In this paper, we focus on the following problem. Given a social network G and a predetermined amount of budget B, what is the target set of individuals X, or equivalently $\boldsymbol{x}$, to whom we initially send the influence, such that the cost for initially influencing $\boldsymbol{x}$ is under the budget and the CIT $\tau(\boldsymbol{x})$ is minimized. The solution to the problem is not straightforward, because for a given $\boldsymbol{x}$, both $C(\boldsymbol{x}, \boldsymbol{\varepsilon})$, the cost for initially influencing $\boldsymbol{x}$, and the CIT $\tau(\boldsymbol{x})$ are random variables, which makes the problem a stochastic programming problem. According to different decision criteria of decision makers, solutions to the problem may be quite different. In this section, we adopt three well-known decision criteria in the theory of uncertain programming [18], and provide three stochastic programming models.

The first model is motivated by the natural idea that we may estimate a random variable by its expected value, which is commonly called the expected value model (EVM). In the EVM model for our problem, the objective is to minimize the expected value of the CIT $\tau(\boldsymbol{x})$, and a target set $\boldsymbol{x}$ is feasible if the expected value of the cost $C(\boldsymbol{x}, \boldsymbol{\varepsilon})$ is below the budget B. The EVM model is as follows:

$$\begin{cases} \min \ E\left[\tau(\boldsymbol{x})\right] \\ \text{subject to :} \\ \quad E[C(\boldsymbol{x}, \boldsymbol{\varepsilon})] \le B, \\ \quad \boldsymbol{x} = \{x_i = 0 \text{ or } 1 \mid i \in N\}. \end{cases} \tag{1}$$

For many cases, estimating a random variable by its expected value is not always appropriate, especially for random variables with large variances. For our problem, when the variance of $C(\boldsymbol{x}, \boldsymbol{\varepsilon})$ is large, it is quite possible that $C(\boldsymbol{x}, \boldsymbol{\varepsilon})$ exceeds B with a high probability even though $E[C(\boldsymbol{x}, \boldsymbol{\varepsilon})] \le B$ holds at the same time. Similarly, the CIT $\tau(\boldsymbol{x})$ can be much greater than its expected value $E[\tau(\boldsymbol{x})]$ with a large probability. In practice, decision makers often predetermine a confidence level α as a safety margin and solve problems in the sense that the constraints are satisfied at least α of time. This idea is usually referred to as chance-constrained programming (CCP) [4]. For our problem, the CCP model is as follows:

$$\begin{cases} \min \ \{\overline{\tau} \mid \Pr\{\tau(\boldsymbol{x}) \le \overline{\tau}\} \ge \alpha\} \\ \text{subject to :} \\ \quad \Pr\{C(\boldsymbol{x}, \boldsymbol{\varepsilon}) \le B\} \ge \beta\}, \\ \quad \boldsymbol{x} = \{x_i = 0 \text{ or } 1 \mid i \in N\}, \end{cases} \tag{2}$$

where α and β are two predetermined values between 0 and 1.

For some cases, decision makers may take a third decision criterion, under which they attempt to maximize the chance to meet some stochastic events that they expect to happen. In order to make decisions under this criterion, Liu provided the dependent-chance programming (DCP) [17]. In our problem, there may be a deadline for the diffusion of influence such that exceeding the deadline will lead to a huge penalty. In this view, we may need to maximize the probability of the event that the CIT does not exceed the deadline τ_0. The DCP model is as follows:

$$\begin{cases} \max\ \Pr\{\tau(x) \le \tau_0\} \\ \text{subject to :} \\ \quad E[C(\boldsymbol{x}, \boldsymbol{\varepsilon})] \le B, \\ \quad \boldsymbol{x} = \{x_i = 0 \text{ or } 1 \mid i \in N\}. \end{cases} \tag{3}$$

where τ_0 is a predetermined value representing the deadline.

§4 The Modified Greedy Algorithms

In order to solve the stochastic programming models that we propose above, we need to estimate the values of a number of uncertain functions. Here, we employ the stochastic simulation techniques to estimate the following three types of uncertain functions:

- $U_1 : \boldsymbol{x} \to E[\xi(\boldsymbol{x})]$,
- $U_2 : \boldsymbol{x} \to \min\{\overline{f} \mid \Pr\{\xi(\boldsymbol{x}) \le \overline{f}\} \ge \alpha\}$,
- $U_3 : \boldsymbol{x} \to \Pr\{\xi(\boldsymbol{x}) \le f_0\}$,

where $\xi(\boldsymbol{x})$ is an uncertain function. The functions $E[\tau(\boldsymbol{x})]$ and $E[C(\boldsymbol{x}, \boldsymbol{\varepsilon})]$ in models (1) and (3) are two instances of U_1; the objective function is a straightforward instance of U_2; and $\Pr\{C(\boldsymbol{x}, \boldsymbol{\varepsilon}) \le B\}$ in model (2) and $\Pr\{\tau(\boldsymbol{x}) \le \tau_0\}$ in model (3) are instances of U_3. We omit the details of the stochastic simulation techniques, and interested readers can refer to the book by Fishman [11].

In real applications, a social network usually contains a huge number of individuals, thus the computation for solving CIT minimization problems is very time-consuming. An efficient approach to solving optimization problems in complex systems is the greedy algorithm, which usually avoids the brute-force search and obtains satisfactory solutions [7]. For the CIT minimization problems in social networks, by the traditional greedy algorithm, we put nodes in the target set once a time unless the constraints are violated [19]. Every time we add a new node to the target set, we select the one that brings the most improvement in the value of the objective function. Whenever a node is put in the target set, it will never be taken out of the set, and thus the target set is efficiently constructed. However, traditional greedy algorithm computes the values of the objective functions in the proposed models by stochastic simulation, which can be still time-consuming for large-scale social networks.

In this paper, we modified the traditional greedy algorithm for the sake of saving the computation in stochastic simulation. The idea is to introduce an integer parameter r to trade off between solution performance and computational complexity [19]. In the traditional greedy algorithm, when we need to add a new node to the target set, we compute, by stochastic simulation, the change of the value of objective function for every node that is possible to be selected. However, in the modified greedy algorithm, every time we add a node to the target set, we first determine r nodes, each of which is potentially a good choice for being put in the target set; and then compute by stochastic simulation the change of the value of objective function for adding each of the r nodes. We design several efficiently-computable heuristic functions to indicate which node is a potentially good alternative for being put in the target set, thus the computational complexity of stochastic simulation is alleviated.

In order to present the heuristic functions, we first define the distance graph of a social network. The distance graph of a social network is a directed graph that has exactly the same set of nodes with the original social network. For each pair of nodes that are connected by an edge in the social network, there are two directed edges in opposite directions connecting them in the distance graph. The weight on each edge (i, j) in the distance graph represents the distance between i and j, which is defined by

$$d(i, j) = \frac{\sum_{l \in N(j)} W(l, j)}{W(i, j)}.$$

For each set of nodes $\boldsymbol{x}$ and a node i, the heuristic function returns a non-negative value $h(\boldsymbol{x}, i)$ quantifying the fitness of putting i in $\boldsymbol{x}$ in terms of improving the performance of the solution. For example, if $h(\boldsymbol{x}, i) \geq h(\boldsymbol{x}, j)$, then adding i to $\boldsymbol{x}$ is potentially better than putting j in $\boldsymbol{x}$. Here, we present three heuristic functions.

We define by $SP(i, j)$ the length of the shortest path from i to j in the distance graph and by $SP(\boldsymbol{x}, i) = min_{l \in \boldsymbol{x}} SP(l, i)$ the shortest path length from set of nodes $\boldsymbol{x}$ to i in the distance graph. If $\boldsymbol{x}$ is an empty set, we define $SP(\boldsymbol{x}, i)$ as a sufficiently large number, for example, the sum of the weights on all the edges in the distance graph. The first heuristic function is motivated by the intuition that it usually takes much time to spread the influence to a node that is distant from the influenced nodes. We call the heuristic function "Shortest Path Length" (SPL) and define it by

$$h_1(\boldsymbol{x}, i) = SP(\boldsymbol{x}, i).$$

We define the maximin path length of a set of nodes $\boldsymbol{x}$ by $\max_{j \in N} SP(\boldsymbol{x}, j)$, which describes how far from the set to the farthest node in the network. The idea of the second heuristic function is that it gives a high score to a node if adding it to the target set will highly reduce the maximin path length. We call this heuristic function "Maximin Path Length Reduction" (MPLR) and define it by

$$h_2(\boldsymbol{x}, i) = \max_{j \in N} SP(\boldsymbol{x}, j) - \max_{j \in N} SP(\boldsymbol{x} \cup \{i\}, j).$$

Given the distance graph of a social network, it is not difficult to generate a spanning forest with the set of roots being $\boldsymbol{x}$, such that each node in the forest is connected to its closest root. Such a spanning forest is called the shortest path forest and denoted by $F_{SP}(\boldsymbol{x})$. Let $Edge(F_{SP}(\boldsymbol{x}))$ be the set of edges in $F_{SP}(\boldsymbol{x})$. We present the third heuristic function named "Shortest Path Forest Weight Reduction" (SPFWR) as follows.

$$h_3(\boldsymbol{x}, i) = \sum_{(l,j) \in Edge(F_{SP}(\boldsymbol{x}))} d(l, j) - \sum_{(l,j) \in Edge(F_{SP}(\boldsymbol{x} \cup \{i\}))} d(l, j).$$

Finally, we present the modified greedy algorithms for solving the EVM model (1), the CCP model (2) and the DCP model (3) in Table 1, 2 and 3, respectively.

§5 Experiments

In this section, we perform a number of experiments on randomly-generated networks. By the experiments, we show the effectiveness of the modified greedy algorithms and how they trade off between solution performance and computational complexity. In all the following experiments, we set the number of nodes $|N| = 1000$ and the number of samples in the stochastic simulation $M = 500$.

In order to generate a social network, we define a parameter p to specify how social the individuals in the social network will be. Given the set of nodes N, for every pair of nodes in the network, we generate an edge between them with probability p. Therefore, a high value of p will result in a social network where each individual is possibly connected to a lot of other individuals. After we generate the social network in such a stochastic way, we add as few

Table 1 A modified greedy algorithm for solving the EVM model (1)

Input: G, C, B, r, h.
Output: $\boldsymbol{x}$.
  $\boldsymbol{x} = \emptyset$;
  $T = \{i \in N \mid E[C(\{i\}, \boldsymbol{\varepsilon})] > B\}$;
  while $|\boldsymbol{x} \cup T| < |N|$
    $n = 0$;
    $R = \emptyset$;
      while $n < r$ and $|\boldsymbol{x} \cup T \cup R| < |N|$
        $i_n = \arg\max_{i \in N \setminus (\boldsymbol{x} \cup T \cup R)} h(\boldsymbol{x}, i)$;
        $R = R \cup i_n$;
        $n = n + 1$;
    $j = \arg\max_{i \in R} (E[\tau(\boldsymbol{x})] - E[\tau(\boldsymbol{x} \cup \{i\})])$;
    $\boldsymbol{x} = \boldsymbol{x} \cup \{j\}$;
    $T = \{i \in N \setminus \boldsymbol{x} \mid E[C(\boldsymbol{x} \cup \{i\}, \boldsymbol{\varepsilon})] > B\}$;
  return $\boldsymbol{x}$;

Table 2 A modified greedy algorithm for solving the CCP model (2)

Input: G, C, B, α, β, r, h.
Output: $\boldsymbol{x}$.
  $\boldsymbol{x} = \emptyset$;
  $T = \{i \in N \mid \Pr\{C(\{i\}, \boldsymbol{\varepsilon}) \leq B\} < \beta\}$;
  while $|\boldsymbol{x} \cup T| < |N|$
    $n = 0$;
    $R = \emptyset$;
      while $n < r$ and $|\boldsymbol{x} \cup T \cup R| < |N|$
        $i_n = \arg\max_{i \in N \setminus (\boldsymbol{x} \cup T \cup R)} h(\boldsymbol{x}, i)$;
        $R = R \cup i_n$;
        $n = n + 1$;
    $j = \arg\max_{i \in R} (\min\{\overline{\tau} | \Pr\{\tau(\boldsymbol{x}) \leq \overline{\tau}\} \geq \alpha\} - \min\{\overline{\tau} | \Pr\{\tau(\boldsymbol{x} \cup \{i\}) \leq \overline{\tau}\} \geq \alpha\})$;
    $\boldsymbol{x} = \boldsymbol{x} \cup \{j\}$;
    $T = \{i \in N \setminus \boldsymbol{x} \mid \Pr\{C(\boldsymbol{x} \cup \{i\}, \boldsymbol{\varepsilon}) \leq B\} < \beta\}$;
  return $\boldsymbol{x}$;

Table 3 A modified greedy algorithm for solving the DCP model (3)

Input: G, C, B, τ_0, r, h.
Output: $\boldsymbol{x}$.
$\boldsymbol{x} = \emptyset$;
$T = \{i \in N \mid E[C(\{i\}, \boldsymbol{\varepsilon})] > B\}$;
while $|\boldsymbol{x} \cup T| < |N|$
  $n = 0$;
  $R = \emptyset$;
    while $n < r$ and $|\boldsymbol{x} \cup T \cup R| < |N|$
      $i_n = \underset{i \in N \setminus (\boldsymbol{x} \cup T \cup R)}{\arg\max} \; h(\boldsymbol{x}, i)$;
      $R = R \cup i_n$;
      $n = n + 1$;
  $j = \underset{i \in R}{\arg\max} \, (\Pr\{\tau(\boldsymbol{x} \cup \{i\}) \le \tau_0\} - \Pr\{\tau(\boldsymbol{x}) \le \tau_0\})$;
  $\boldsymbol{x} = \boldsymbol{x} \cup \{j\}$;
  $T = \{i \in N \setminus \boldsymbol{x} \mid E[C(\boldsymbol{x} \cup \{i\}, \boldsymbol{\varepsilon})] > B\}$;
return $\boldsymbol{x}$;

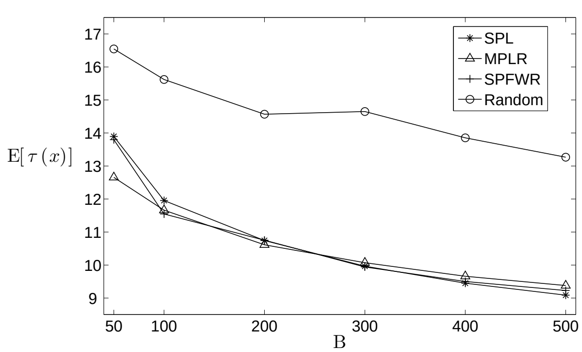


Figure 1 Results of modified greedy algorithms with different B for the EVM model.

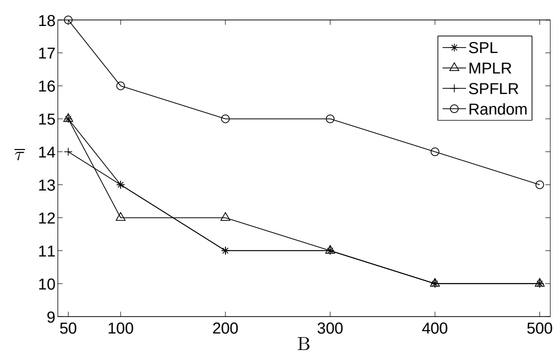


Figure 2 Results of modified greedy algorithms with different B for the CCP model.

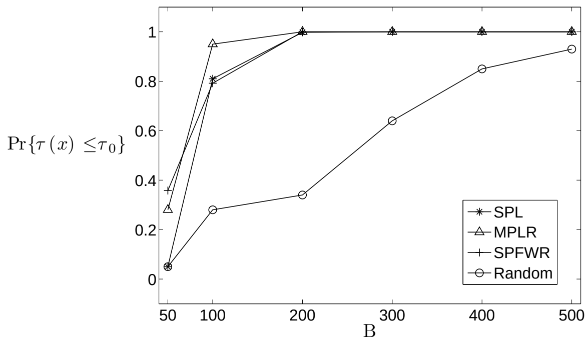
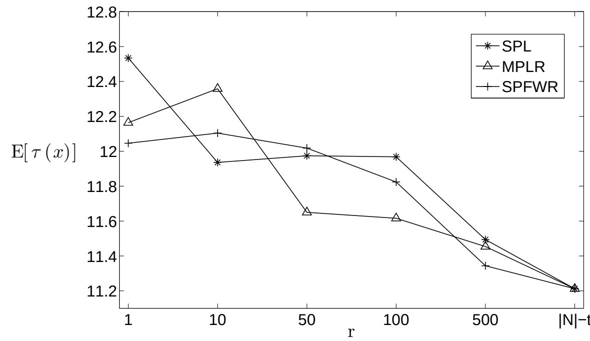


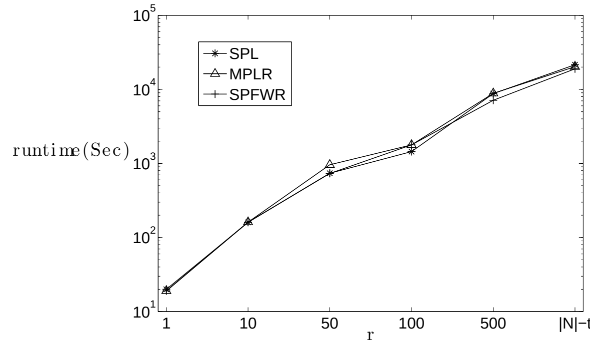
Figure 3 Results of modified greedy algorithms with different B for the DCP model.

edges as possible to make the network a connected graph. In the following experiments, without loss of generality, we randomly assign an integer weight from $\{1, 2, 3\}$ to each edge. The stochastic cost for initially influencing each node i is given as follows: we first generate an integer u_i in the interval $[0, 20]$ at uniformly random, and then set the stochastic cost ε_i as the random variable with uniform distribution $\mathcal{U}(20, 20 + u_i)$.

In order to show the effectiveness of the proposed algorithms, in the first series of experiments, we solve the three stochastic programming models with different budget B by the modified greedy algorithms. In addition, we compare the solution performance of the modified greedy algorithms with that of a trivial heuristic method called "randomly choosing nodes". The heuristic "randomly choosing nodes" randomly selects the nodes in the target set without exploiting any information of the network. In the first series of experiments, without loss of generality, we set $p = 0.02$ to generate the social network and the trade-off parameter $r = 50$ in the modified greedy algorithms. The solutions for solving EVM model, CCP model and DCP model are shown in Figure 1, 2 and 3, respectively. In the three figures, a result of a modified greedy algorithm is labelled with the heuristic function the algorithm adopts, and a result of the "randomly choosing nodes" heuristic is labelled "Random". For the CCP model, we set α and β both as 0.9; while for the DCP model, the value of τ_0 is set to be 12.



(a)



(b)

Figure 4 The solutions and runtime of the modified greedy algorithms.

From Figure 1, 2 and 3, we have some clear observations. No matter what value the budget is, the modified greedy algorithms generate better solutions than the "randomly choosing nodes" heuristic. The only expectation is that when solving the DCP model with $B = 50$, the modified greedy algorithm with heuristic function "SPL" has the same solution performance with the "randomly choosing nodes" heuristic. This is because there are at most two nodes in the target set when $B = 50$, and the power of the modified greedy algorithm may be weaken in searching such a small set. The second observation is that, no matter which algorithm is employed, increasing the budget B will improve the value of the objective for every model, which coincides with our intuition. In addition, we find that none of the three modified greedy algorithms is better than the others for all the problems.

In the next experiment, we show how the modified greedy algorithms trade off between solution performance and computational complexity. We let $p = 0.02$ and $B = 100$ and consider the EVM model. We change the value of the trade-off parameter r and obtain a number of results that are shown in Figure 4. It is not difficult to find that the stochastic simulation is entirely avoided when $r = 1$. Another extreme is that when $r(t) = |N| - t$ depends on the volume of the current target set t, the modified greedy algorithms degenerate to the traditional greedy algorithm. We find that, in general, greater r will make the modified greedy algorithms generate better solutions, however at the expense of spending more time. On the other hand, smaller r helps save computation, however the performance of solutions is poorer. Therefore, Figure 4 illustrates the fact that the modified greedy algorithms trade off between optimality and complexity.

§6 Conclusions

In this paper, we have focused on the complete influence time minimization problem given that the available budget for initially influencing nodes is limited and that the cost for targeting each node is stochastic. The incremental chance model has been adopted to characterize the mechanism of the influence diffusion, and three different stochastic programming models have been proposed. In order to solve the proposed models, we have modified the traditional greedy algorithm so that the complexity and optimality can be better traded off. Experiments have been performed on random graphs, which have shown that the algorithms we propose are effective and help decision makers get a good balance between solution performance and computational complexity.

Acknowledgements

This work was supported by National Natural Science Foundation of China (No. 71101027), Program for Young Excellent Talents, UIBE (No. 12YQ08), and Research Funding of School of Information Technology and Management, UIBE (No. SIT2011001).

References

[1] R. Albert, A.-L. Barabási, Statistical mechanics of complex networks, *Reviews of Modern Phiysics*, vol. 74(1), pp 47-97, 2002.

[2] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, *Science*, vol. 286, pp 509-512, 1999.

[3] J. Brown, P. Reinegen, Social ties and word-of-mouth referral behavior, *Journal of Consumer Research*, vol. 14(3), pp 350-362, 2007.

[4] A. Charnes, W.W. Cooper, Chance-constrained pragramming, *Management Scince*, vol. 6, pp 73-79, 1959.

[5] N. Chen, On the Approximability of influence in social networks, in: *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, California, pp 1029-1037, 2008.

[6] J. Coleman, H. Menzel, E. Katz, *Medical Innovations: A Diffusion Study*, Bobbs Merrill, 1966.

[7] T.H. Cormen, C.E. Leiserson, R.L. Rivest, C. Stein, *Introduction to Algorithms*, MIT press, Cambridge, MA, 2001.

[8] Z. Dezso, A.-L. Barabási, Halting viruses in scale-free networks, *Physical Review E*, vol. 65, pp 055103, 2002.

[9] P. Domingos, M. Richardson, Mining the network value of customers, in: *Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, California, pp 57-66, 2001.

[10] E. Even-Dar, A. Shapira, A note on maximizing the spread of influence in social networks, in: *Proceedings of the 3rd International Workshop on Internet and Network Economics*, pp 281-286, 2007.

[11] G.S. Fishman, *Monte Carlo: Concepts, Algorithms, and Applications*, Springer-Verlag, New York, 1995.

[12] J. Goldenberg, B. Libai, E. Muller, Talk of the network: a complex systems look at the underlying process of word-of-mouth, *Marketing Letters*, vol. 12(3), pp 211-223, 2001.

[13] A. Grabowski, N. Kruszewska, R.A. Kosiński, Properties of on-line social systems, *The European Physical Journal*, vol. 66, pp 107-113, 2008.

[14] M. Granovetter, Threshold models of collective behavior, *American Journal of Sociology*, vol. 83, pp 1420-1443, 1978.

[15] D. Kempe, J. Kleinberg, É. Tardos, Maximizing the spread of influence through a social network, in: *Proceedings of 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Washington D.C., pp 137-146, 2003.

[16] D. Kempe, J. Kleinberg, É. Tardos, Influential nodes in a diffusion model for social networks, in: *Proceedings of the 32nd International Colloquium on Automata, Languages and Programming*, Lisbon, Portugal, pp 1127-1138, 2005.

[17] B. Liu, Dependent-chance programming: a class of stochastic programming, *Computer and Mathematics with Aplications*, vol. 34, pp 89-104, 1997.

[18] B. Liu, *Theory and Practice of Uncertain Programming*, Physica-Verlag, Heidelberg, 2002.

[19] Y. Ni, L. Xie, Z.-Q. Liu, Minimizing the expected complete influence time of a social network, *Information Sciences*, vol. 180(13), pp 2514-2527, 2010.

[20] Y. Ni, Z.-Q. Liu, Heuristic search for optimizing diffusion of influence in a social network under the resource constraint, *Soft Computing*, vol. 15(2), pp 335-344, 2011.

[21] M. Richardson, P. Domingos, Mining knowledge-sharing sites for viral marketing, in: *Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Edmonton, Alberta, pp 61-70, 2002.

[22] D.J. Watts, S.H. Strogatz, Collective dynamics of “small-world” networks, *Nature*, vol. 393, pp 440-442, 1998.