Supplemental Material of “Fine-grained Feature-based Social Influence Evaluation in Online Social Networks”

Guojun Wang, *Member, IEEE,* Wenjun Jiang, Jie Wu, *Fellow, IEEE,* and Zhengli Xiong

Appendix

We analyze and prove the time complexity of our two proposed algorithms.

Proof of Theorem 1.

Proof: Suppose the total number of nodes is m, the size of the feature set is n, the number of friends of u is $|N_u|$, the number of common friends between u and v is $|C_{uv}|$, and the average degree is a.

For calculating the similarity S_{uv} (line 4), each feature should be considered. It takes the complexity of $O(n)$. For calculating the direct affinity A^d_{uv} (line 5), the tie strength of $t_{u,v}$ should be found. It will be very time-consuming if we search $e(u,v)$ from the whole edge list. To avoid this, we can keep an index list according to the neighbor list for each node. By doing this, we only need to search the index of $e(u,v)$, and then find $t_{u,v}$; it takes $O(|N_u|)$. Similarly, for indirect affinity s^{id}_{uv} (line 6), the time complexity is $O(|C_{uv}|)\cdot(O(|N_u|)+O(|N_v|))$. It takes constant complexity for line 7 and line 8, that is, $O(1)$. In summary, the complexity of lines 3-7 is $O(|N_u|)\cdot(O(n)+O(|C_{uv}|)\cdot(O(|N_u|)+O(|N_v|)))$. It is worth noting that this is an upper bound, which may be much larger than the real complexity. The reason is that, for many users who are at the bottom of the node list, the similarity, affinity, or even impact, may have been calculated before. In this case, it is unnecessary to recompute. Searching those values only takes a time cost of $O(|N_u|)$.

The number of average common neighbors is usually much smaller than the average degree; the feature set is usually very small compared to the size of the network. Thus, $O(n)$ and $O(|C_{uv}|)$ can be neglected without loss of generality. Then, we can use a to replace $O(|N_u|)$ and $O(|N_v|)$. Therefore, it takes the time complexity of $O(a^2m)$ for Algorithm 1. □

In an online social network, the degree distribution fits with the power-law distribution. Therefore, the average degree a can be taken as a constant. We can deem the total complexity as $O(m)$ if the network is sparse, as it usually happens.

In another case of dense network, since the algorithm only needs local information (i.e., the information of neighbors within at most 2-hops), we can easily distribute the computation into several parts to improve the efficiency. Only some copies of the marginal nodes between each part are needed, so as to keep the consistency with the whole social graph.

Proof of Theorem 2.

Proof: Let us first consider the process of adjusting a user u's social influence (lines 3-6). Both line 3 and line 6 take constant time cost $O(1)$. For calculating the proportion of contribution from a neighbor v (line 5), since all the impacts have been calculated by Algorithm 1, it only needs to search in the neighbor set of v (to find the index of u), which takes $O(|N_v|)$. So, it takes a total of $O(|N_v|\cdot|N_u|)$ for lines 3-6.

Then, for each adjustment iteration (lines 2-9), it takes the time complexity of $O(|N_v|\cdot|N_u|\cdot m)$. We can use the average degree a to replace $O(|N_u|), O(|N_v|)$. Therefore, it takes the time complexity of $O(a^2m)$ for each iteration in Algorithm 2. □

Due to the same reason with Algorithm 1, we can deem the total complexity as $O(m)$ if the network is sparse. Moreover, since only local information is needed for updating the social influence, Algorithm 2 can also be distributively executed to improve the efficiency.

- *G. Wang, W. Jiang, and Z. Xiong are with the School of Information Science and Engineering, Central South University, Changsha, Hunan Province, 410083, P. R. China. E-mail: csgjwang@csu.edu.cn, wenjj8a@gmail.com. G. Wang is the corresponding author.*
- *J. Wu is with the Department of Computer and Information Sciences, Temple University, Philadelphia, PA 19122, USA. E-mail: jiewu@temple.edu.*

This work is supported by NSFC grants 61272151 and 61073037, ISTCP grant 2013DFB10070, the China Hunan Provincial Science & Technology Program under Grant Number 2012GK4106, the Ministry of Education Fund for Doctoral Disciplines in Higher Education under Grant Number 20110162110043, NSF grants ECCS 1231461, ECCS 1128209, CNS 1138963, CNS 1065444, and CCF 1028167.