CrowdDiffuse: Information Diffusion over Crowds with Social Network

Cheng-Te Li * Hsun-Ping Hsieh Shou-De Lin
Graduate Institute of Networking and Multimedia, National Taiwan University, Taipei, Taiwan

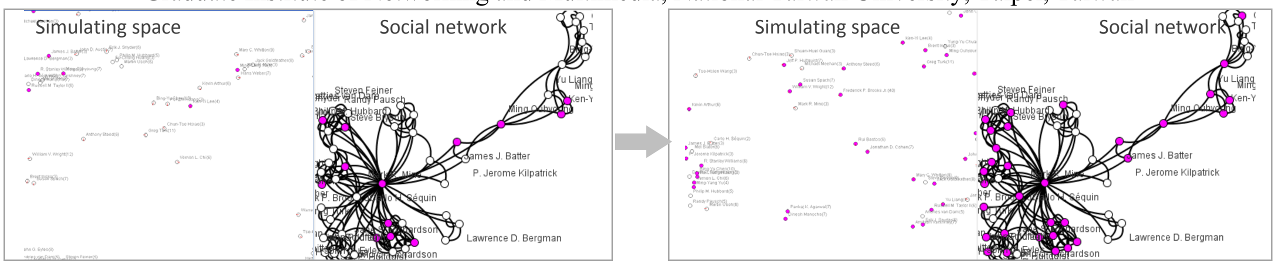

Figure 1: Spatio-social information diffusion with spatial and link-based social condition (purple indicates being informed).

1. Introduction

Crowd simulation is to produce collective behaviors through simulating the movement process of a number of agents. Information diffusion is to investigate how information propagates among people, especially on social network. This work considers both the spatial and social perspectives to simulate the dynamics of information diffusion over crowds with underlying social relationships. Existing works [Chao 2009; Neumann 2000] model the communications between agents so that elicited emotions can be spread in the space while others (e.g. [Kempe 2003]) study how information diffuses over a social network. However, the former considers only spatial aspects and the latter utilizes only the social relationships. To combine both spatial and social aspects for diffusion, we propose a framework, *CrowDiffuse*, which simulates the crowd with an underlying social network and allows information to propagate over agents. In *CrowDiffuse*, agents can affect each other if they satisfy both spatial and social condition. The spatial condition is satisfied if two agents come close enough in the space while the social condition is met if two agents possess a relationship in the social network. In addition, we further investigate the *targeted diffusion* problem, which will be described in the following.

2. Spatio-Social Information Diffusion

We use Reynold's flocking model [1987] for crowd simulation. The DBLP[1] database is used to construct the social network, in which each node (i.e., agent) stands for an author and each link is co-authorship.

Diffusion Method. The diffusion process starts from specifying a set of agents as initial informed ones. As the simulation proceeds, more and more agents will be informed. Finally all agents in the environment are informed. The diffusion from one agent to another is determined by the spatial and social conditions. One agent x can successfully affect y if and only if both conditions are satisfied. For the spatial condition, when an agent x is exposed to the local visible range of y who has been informed, the spatial condition is satisfied. The social condition is divided into link-based and community-based cases. We consider *Independent Cascade* model [Kempe 2003] as the link-based case: an informed agent v is given a single chance to affect each of its neighbor u who is not informed yet in the network. And it succeeds with a probability $p_{v,u}$ independently of the spreading history so far. If v succeeds, u will be informed. That says, when v and u must have a relationship in the social network and meet the probability $p_{v,u}$, the social condition will be satisfied. Figure 1 shows the diffusion example with one initial informed agent and $p_{v,u}$=0.5. We can find the information is propagated from few to most agents. For community-based case, it is looser: if u and v meet and belong to the same community [Newman 2004] and satisfy $p_{u,v}$, it meets the social condition.

Targeted Diffusion. We use our spatio-social diffusion model to answer "*if someone want to distribute an idea or certain information, how to select initial informed agents to diffuse the idea efficiently?*" That is, how to pick the initial informed agents to finally affect all agents with as less time as possible? To study the targeted diffusion, we consider spatial and social factors to devise six strategies of selecting initial informed agents: (1) *Random Selection*, (2) select agents located in *Center Location*, (3) select agents with *Densest Neighborhood* in the space. The following consider centrality measures in social network area [Wasserman 1994]: select agents with highest (4) *Degree* (more relationships to others), (5) *Closeness* (averagely closer to others based on graph distances), and (6) *Betweenness* (higher chances to be bridges between communities).

Experimental Results. Under some rounds, if a strategy can informs more agents, comparing to other strategies, it is regarded as a more effective strategy. We set the number of initially-informed agents to 4, the probability of independent cascade $p_{v,u}$=0.5, and #community=5. For each strategy, we simulate the diffusion 100 times and compute the average number of informed agents. Figure 2 shows the results. We can find the social strategies (i.e., degree, closeness, betweenness) generally have better performance. We think it is because agents are harder to satisfy the social condition than the spatial condition. The closeness is best since agents with higher closeness tend to reach others with fewer steps in the network. For the other three, the dense-neighbor is better since agents with higher density in the neighborhood have higher potential to meet other agents spatially to trigger the social condition.

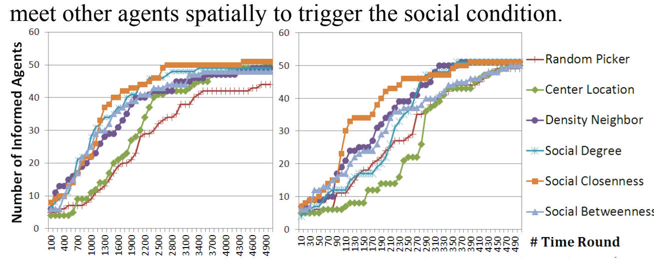

Figure 2: Experimental results of targeted diffusion under the link-based (left) and the community-based (right) social conditions.

References

Chao, W. M. and Li, T.Y. Simulating Crowd Behaviors with a Communication Model. *CASA* 2008.

Kempe, D., Kleinberg , J., and Tardos, E. Maximizing the Spread of Influence through a Social Network. *ACM SIGKDD* 2003.

Neumann, R. and Strack, F. Mood Contagion: The Automatic Transfer of Mood between Persons. Journal of Personality and Social Psychology, 2000.

Newman, M. E. J. Fast Algorithm for Detecting Community Structure in Networks. Physical Review, E 69, 066133, 2004.

Reynolds, C.W. 1987. Flocks, Herds and Schools: A Distributed Behavior Model. *ACM SIGGRAPH* 1987.

Wasserman, S. and Faust, K. Social Network Analysis. 1994.

SIGGRAPH 2012, Los Angeles, California, August 5 – 9, 2012.
ISBN 978-1-4503-1435-0/12/0008

* Email: d98944005@csie.ntu.edu.tw

[1] http://www.informatik.uni-trier.de/~ley/db/