ABSTRACT

Dissemination of information is a significant aspect of any large scale distributed system. This thesis analyzes in details two distinct classes of dissemination processes. One class comprises of the processes which are meant for large scale technological communication systems and are controlled by the underlying stochastic algorithms, while the processes in the other class happen in a more self-organized fashion through social interactions among the humans.

One of the most desirable features of the dissemination services, implemented on large scale peer-to-peer communication systems is the maximization of the coverage, i.e., the number of distinct nodes that receive the message under the constraints of available network resources (e.g., energy in the nodes) as well as time. However, redundant delivery of the information to the nodes by the underlying algorithms for these services, causes wastage of both resource and time. Through a detailed stochastic analysis of the multiple random walker based dissemination algorithms we identify that redundancy quickly increases with increase in the concentration of the walkers. Based on this postulate, we design a very simple distributed algorithm which dynamically estimates the concentration and thereby carefully proliferates walkers in sparse regions. We test our algorithm on various kinds of network topologies whereby we find it to be performing well specially in networks that are highly clustered as well as sparse.

On the other hand, dynamics of information dissemination through social interactions among the humans are crucially affected by their social behaviors such as the tendency of selective visits to different places, joining different social groups etc. A significant part of the thesis is devoted to a systematic study of the effect of

these issues on the dissemination processes. We model the system as a special kind of bipartite network where one partition corresponds to the set of the active entities which are the humans and the other partition stands for passive entities such as popular places, social groups. The active partition grows with time and is highly dynamic in nature while the other partition is fixed in size and is static in nature. Such bipartite networks are known as alphabetic bipartite network. Fundamentally, this bipartite construction represents a form of indirect spreading of information in the system through the interactions among the two classes of entities, e.g., information flows from one active entity to another active entity through some passive entity and similarly from one passive entity to another passive entity through some active entity.

In this thesis, we study this indirect spreading phenomena in two ways. The fixed sized passive partition plays significant role in the indirect flow of information in the system. This can be best understood through a hypothetical information-flow network constructed among these passive entities. Hence, we first do a detailed structural analysis of this network. Mainly, we exploit a classical graph theoretic concept - '*Random threshold graph*' and employ statistical techniques to find out mathematical formulas for the largest connected component size, degree distribution, edge density and clustering coefficient of this network. Next, in order to understand the dynamical aspects of the indirect spreading, we carry out a detailed functional analysis of the bipartite spreading processes among the active and the passive entities. Furthermore, we derive mathematical formulas for a few important epidemiological metrics, e.g., prevalence, survival probability and epidemic thresholds.

Many instances of technological communication systems are highly associated with the behavior of the target system. Hence, we believe that the insights gained from the detailed studies in this thesis done on two different classes of spreading processes will definitely help design engineers to develop sophisticated and efficient algorithms for various information dissemination services.

Keywords: Information dissemination, Intergroup network, Random threshold graph, Alphabetic bipartite network, α-BiN, Indirect spreading, Direct spreading, Prevalence, Survival probability, Epidemic threshold.