Chapter 2.1
Design of a Data Model for Social Network Applications

Susanta Mitra
International Institute of Information Technology, India

Aditya Bagchi
Indian Statistical Institute, India

A.K.Bandyopadhyay
Jadavpur University, India

ABSTRACT

A social network defines the structure of a social community like an organization or institution, covering its members and their inter-relationships. Social relationships among the members of a community can be of different types like friendship, kinship, professional, academic, and so forth. Traditionally, a social network is represented by a directed graph. Analysis of graph structure representing a social network is done by the sociologists to study a community. Hardly any effort has been made to design a data model to store and retrieve social-network-related data. In this paper, an object-relational graph data model has been proposed for modeling a social network. The objective is to illustrate the power of this generic model to represent the common structural and node-based properties of different social network applications. A novel, multi-paradigm architecture has been proposed to efficiently manage the system. New structural operators have been defined in the paper and the application of these operators has been illustrated through query examples. The completeness and the minimality of the operators have also been shown.

INTRODUCTION

A social network is a social structure between actors (individuals, organizations, or other social entities) and indicates the ways in which they are connected through various social relationships like friendships, kinships, professional, academic, and so forth. Usually, a social network may represent a network of acquaintance between people; a club and its members; a city or village

community; a research group communicating over the Internet; or a group of people communicating with each other through e-mail messages. Recently, the World Wide Web or just Web, as it is popularly known, has played a major role in the formation of communities (*cyber communities or Web communities*) where the members or people from different parts of the globe can join a community with common interest. For example, members of an IEEE society communicating with each other through e-mail may form a Web community. Social network applications include the traditional social network applications as studied by the social scientists Hanneman (2001), Holland and Leinhardt (1979), and Leinhardt (1977); network of acquaintances or referral system as proposed in Yu and Singh (2003) and Kuatz, Selman, and Shah (1997); and finally the Web community (Hanneman, 2001; Newman, 2003). Incidentally, in a referral system, each actor in the social community provides a set of links to its acquaintances that in turn become members of the community. In the same way, these new actors bring their acquaintances to the community again. Thus, the social network keeps on growing. This view of social network has given rise to different commercial applications like LinkedIn.com (http://www.LinkedIn.com), Ryze.com (http://www.Ryze.com), Tribe.net (http://www.Tribe.com), and so forth. For example, a commercial referral network on the Web may offer employment services, where actors provide information like *qualification, experience,* and so forth. Similarly, another referral network may offer matrimonial services, where actors provide information like *age, marital-status, sex, monthly earnings,* and so forth.

Social networks can have a few or many actors, and one or more kinds of relations between pairs of actors. For example, two houses of a village community may be connected to each other because of a family relationship yielding a *kinship* relation or they may communicate for lending or borrowing money generating an *economic* relationship. Two actors of a social network may even be connected by more than one relation. For example, an actor i may refer to another actor j, since they belong to the same professional area (e.g., computer scientist), and at the same time they may also be connected by another relation like the same hobby (e.g., playing baseball).

To build a useful understanding of a social network, a complete and rigorous description of a pattern of social relationships is a necessary starting point for analysis. This pattern of relationships between the actors can be better understood through mathematical or formal representation, like graphs. Therefore, a social network is represented as a directed graph or digraph. In this graph, each member of a social community (people or other entities embedded in a social context) is considered as a node, and communication (collaboration, interaction, or influence) from one member of the community to another member is represented by a directed edge. In order to understand the social properties and behavior of a community, social scientists analyze the corresponding digraph. The number of nodes in social network applications can be very few representing a small circle of friends or very large representing a Web community. This graphical representation is useful for the study and analysis of a social network. In addition, each social network will also have some node-related information depending on the application area or the type of social community the network is representing. For example, in a village community, each node may represent a household in the village with data relevant to such houses.

In the 1970s Leinhardt (1977) first proposed the idea of representing a social community by a digraph. Later, this idea became popular among other research workers like, network designers, Web-service application developers, and e-learning modelers. It gave rise to a rapid proliferation of research work in the area of social network analysis. A graph representing a social network has certain basic structural properties, which