BEYOND METAMATRIX: TOWARDS A ROBUST SEMANTICS FOR REASONING ABOUT NETWORKS

MAKSIM TSVETOVAT, KATHLEEN M. CARLEY

1. Critique of Traditional SNA models

Traditional social network analysis operates on a simple set of concepts: nodes of a social network are people or groups of people - generally a homogenous set, and links between nodes represent existence of a connection in case of binary networks or frequency of communications or closeness of relationship in real-valued networks. Thus, computing simple graph-theoretic measures upon the resultant graph produces interpretable results that allow detection of powerful or important nodes, communication gatekeepers, etc.

With the PCANS (and later MetaMatrix) models, there is a trend towards expansion of network analysis to operate on a heterogeneous set of concepts. PCANS was originally designed to operate on concepts of People, Knowledge and Tasks. Later, this model was expanded to include Resources and Organizations, and its expansion is continuing.

However, with the expansion of the number of node types that take part in the analysis, the concept of a node or an edge increasingly becomes overloaded with a plethora of meanings. For example: An edge between nodes $A_1 \in Agent$ and $A_2 \in Agent$ preserves the original SNA meaning of "a connection exists". Even at this point, a real-valued edge weight carries an ambiguous meaning (is it "strength of connection", "distance", or "frequency of communication"?). An edge between nodes $A_3 \in Agent$ and $O_1 \in Organization$ has an even more ambiguous meaning of "agent is connected to organization" or "agent is a part of the organization".

Of even more importance is a semantic ambiguity of edge direction, i.e. - does it mean the same thing if an *Agent* is connected to an *Organization* or *Organization* is connected to *Agent*? Since the MetaMatrix contains many sub-matrices with heterogeneous node types, and traditionally only included an Upper Triangular portion of the matrix, the directionality of heterogeneous edges is either lost completely (as the lower triangular went ignored) or potentially misinterpreted (does reversal of an edge's direction change the meaning of an edge?).

We also cannot overlook the fact that some edges have different properties then others. For example, subordination edges ("isSuperiorTo") for people or inclusion edges ("isPartOf") for organizations (as well as a number of others) are transitive - i.e. the boss of my boss is also my boss. A number of other special properties can be defined as well. Graph models to date do not offer sufficient reasoning capabilities to resolve transitivity of edges, especially in a context where nodes and edges of multiple types coexist.

acknowledgements go here.

A further complication comes with the introduction of larger datasets, especially these of machine-collected data. Such data often consists of "statements" of the form:

$$subject \longrightarrow verb \longrightarrow object$$

Person1 isConnectedTo Person2

In text source written for human understanding, the statements are mainly multi-lateral and multi-mode - combining information about two or more actors with information about organizations, knowledge, resources, or other entities. To enter into an SNA dataset, such statements need to be converted into sets of bi-lateral edges. Then, re-assembling the original statement from the SNA data becomes a daunting if not impossible task.

The purpose of this paper is to offer an alternative to graph-based models of social networks that would offer formidable expressive power over heterogeneous network models, enable automated reasoning and inference of network properties, yet being at least partially backward-compatible with existing SNA models (thus allowing cross-validation with well-researched datasets).

The proposed social structure semantics is rooted in object-oriented knowledge representation, as well as frame-based reasoning a symbolic inference (chaining) algorithms. We proceed to define a regular language for expression of complex statements related to social structure, and define a set of operations upon the social structure semantics that facilitate reasoning about social structure.

Carnegie Mellon University
E-mail address: maksim@cs.cmu.edu, kcarley@ece.cmu.edu