Title of PhD thesis: *Analysis and Visualization of Complex Heterogeneous Networks*

Complex systems in real-life are affected by the presence of heterogeneity at different levels of their composition. Heterogeneous complex networks differ from ordinary homogeneous networks in a number of respects. At the most basic level, the links in the networks have numerical weights or other descriptive properties associated with them, and so instead of just recording the presence or absence of an interaction or relationship, they record the intensity, strength, or some other characteristic of that interaction or relationship. Additionally, the nodes in the network could denote multiple types of entities, each well-defined and distinct from the other. The same could be true for links in the networks, expressing several different types of relationships or interactions. This variation of node and link types in the network gives rise to semantics of relationships in a network---a description of the types of nodes and links together with rules that dictate which node types can be connected by which link types. An added layer of heterogeneity could result from the association of any number of descriptive properties with both nodes and links.

This PhD thesis presents the results of four studies carried out to study heterogeneous complex networks, and to develop analytic and visual methods to help answer some of the questions involved in their analysis. The results are presented in the form of a collection of papers, each of which addresses a separate study within the context and theme of this thesis. The first study deals with the identification of key players in a weighted covert network by using generalizations of the three classical node centrality measures that allow for a combination of link counts and link weights. It is found that for a covert network where fewer strong ties are more important than numerous weak ties, centrality rankings of important nodes over different observations become less sensitive to minor changes when the influence of link weights is favoured over that of link counts. The second study describes a method for visually exploring a heterogeneous network based on the concept of network semantics. It allows the user to construct semantically meaningful structural patterns and query the network against those patterns, and interactively explore the resulting network. Semantically-guided transformations of a heterogeneous network thus obtained can then be further queried or analysed by means of standard network analysis methods and measures. The third study proposes a topologically-driven method of link prediction for multi-relational networks based on semantically meaningful definitions of compound relationships composed of the distinct relationships given in the network. This prediction method retains the variation in node and link type information, and successfully predicts links of typed relationships, thus preventing loss of information. And finally, the fourth study extends a result from the first study, and investigates the relationship between the stability of the three classical node centrality rankings under different levels of influence of link weights in the presence of random error in the network. It adopts an experimental methodology simulating different observations of a ground-truth network perturbed with different types and sizes of error, and measures the extent of agreement between observed and true node centrality rankings under three different comparison metrics. It is found that there is a significant connection between the stability of a node centrality measure and the extent to which it (the centrality measure) is influenced by link weights.

The main premise of the thesis is that network heterogeneity is a useful and necessary feature of complex networks, and thus methods for analysis and visualization of complex networks must account for that heterogeneity. Disregarding link weights, or network semantics in terms of the variation in the meaning of different entities and relationships in a network lead to loss of important network information, and this must be avoided. This thesis finds that in some situations this additional information can be measured, accounted for and is beneficial to the analysis of complex networks.