

1106-94-2011

Catherine A Bliss* (catherine.bliss@uvm.edu), Vermont Center for Complex Systems, Dept. of, Mathematics, Vermont Advanced Computing Core, 16 Colchester Ave., Burlington, VT 05404, **Christopher M Danforth** (chris.danforth@uvm.edu), Vermont Center for Complex Systems, Dept. of, Mathematics, Vermont Advanced Computing Core, 16 Colchester Ave., Burlington, VT 05405, and **Peter Sheridan Dodds** (pdodds@uvm.edu), Vermont Center for Complex Systems, Dept. of, Mathematics, Vermont Advanced Computing Core, 16 Colchester Ave., Burlington, VT 05405. *Estimation of global network statistics from incomplete data.*

Complex networks underlie an enormous variety of social, biological, physical, and virtual systems. A profound complication for the science of complex networks is that in most cases, observing all nodes and all network interactions is impossible. Previous work addressing the impacts of partial network data is surprisingly limited, focuses primarily on missing nodes, and suggests that network statistics derived from subsampled data are not suitable estimators for the same network statistics describing the overall network topology. We generate scaling methods to predict true network statistics, including the degree distribution, from only partial knowledge of nodes, links, or weights. Our methods are transparent and do not assume a known generating process for the network, thus enabling prediction of network statistics for a wide variety of applications. We validate analytical results on four simulated network classes and empirical data sets of various sizes. We perform subsampling experiments by varying proportions of sampled data and demonstrate that our scaling methods can provide very good estimates of true network statistics while acknowledging limits. Lastly, we apply our techniques to a set of rich and evolving large-scale social networks, Twitter reply networks. (Received September 15, 2014)