When studying large transactional networks such as telephone call detail data, credit card transactions, or web clickstream data, graphs are a convenient and informative way to represent data. In these graphs, nodes represent the transactors, and edges the transactions between them. When these edges have a time stamp, we have a "dynamic graph" where the edges are born and die through time. I will present a framework for representing and analyzing dynamic graphs, with a focus on the massive graphs found in telecommunications and Internet data. The graph is parameterized with three parameters, defining an approximation to the massive graph which allows us to prune noise from the graph. When compared to using the entire data set, the approximation actually performs better for certain predictive loss functions. In this talk I will motivate the methodology with an application to telecommunications fraud, and will describe other applications to viral marketing and acquaintanceship networks.

Joint work with Shawndra Hill, NYU Information Systems Group, and Deepak Agarwal and Bob Bell, AT&T Research.