An Ensemble Framework for Detecting Community Changes in Dynamic Networks

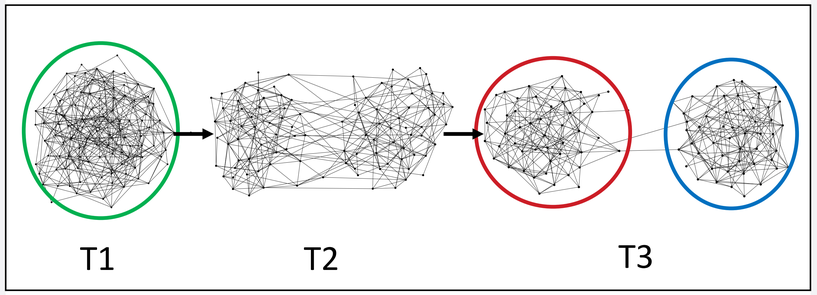
Timothy La Fond, Geoff Sanders, Christine Klymko, Van Emden Henson

DATA SCIENCE INSTITUTE

Lawrence Livermore National Laboratory

UNIVERSITY OF CALIFORNIA National Laboratories

GRAPH COMMUNITIES CHANGE OVER TIME IN NETWORKS

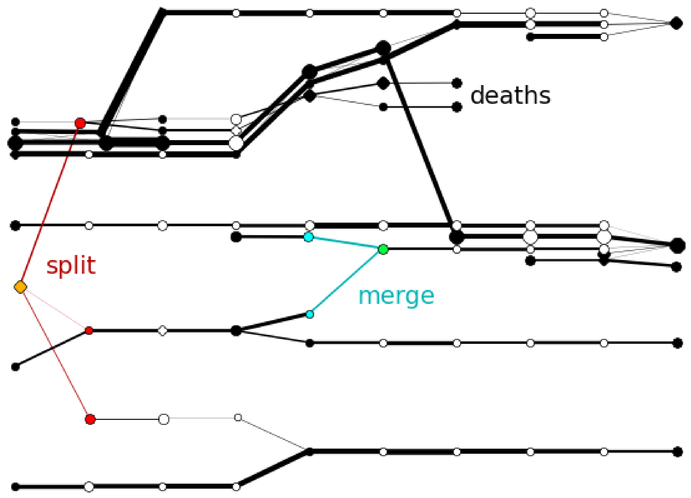


Nodes: individuals; Edges: messages

Nodes organized into Communities

Dynamic Underlying Model: Communities split, merge, appear, disappear, etc. over time

CREATE TIMELINE OF COMMUNITY CHANGES



Quickly generated summary of large network shows community evolution succinctly

Need to find community structure at every time slice: clustering algorithm must be efficient

APPLY STATIC, STOCHASTIC CLUSTERING ALGORITHM

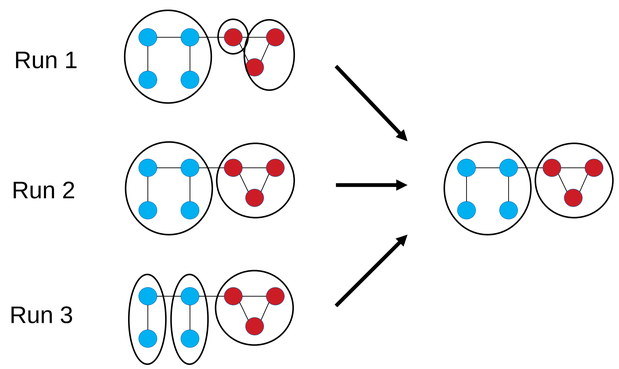
Find clusters at each time using static clustering algorithm

Examples are MCMC [1] and Louvain [2] agglomerative clustering algorithms

Similar clusters in adjacent times are linked to show evolution

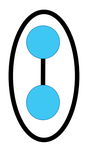
Algorithms are stochastic, random community output

RESOLVE VARIANCE VIA ENSEMBLE APPROACH

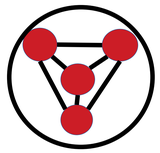


Cluster with multiple runs and pick representative communities as final [3]

HIDE INSIGNIFICANT CLUSTERS



Positive Objective Function but Likely Random

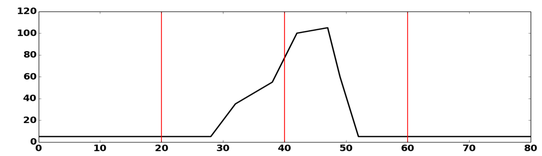


Unlikely to be Random

Hiding insignificant cluster objects improves clarity, use probability of appearing by chance to determine significance

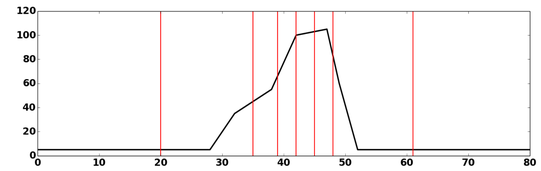
TIME ADAPTIVE ALGORITHM

P: 0.20
R: 0.66



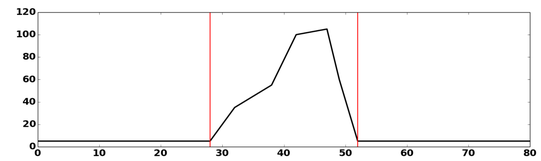
Equal Width

P: 0.82
R: 0.58



Equal Quantity

P: 0.86
R: 0.75



Event Start/End

Data should be divided **temporally** in a manner appropriate to rate of evolution in the network

CONCLUSIONS

Dynamic Clustering aids in the analysis of real-world networks

Due to the variance of practical clustering algorithms, an ensemble approach is desirable

Eliminating unimportant micro-clusters from visualization leads to more interpretable results

A Time-Adaptive clustering algorithm will allow for better representation of events as they occur in the network

COLLABORATORS

CASC
Center for Applied Scientific Computing

1. Tiego Peixoto, “Efficient Monte Carlo and Greedy Heuristics for the Inference of Stochastic Blockmodels,” Phys. Rev. E 89, 012804 (2014). 2. Vincent Blondel et al. "Fast unfolding of communities in large networks." *Journal of statistical mechanics: theory and experiment* 2008.10 (2008): P10008.
3. Timothy La Fond, Geoff Sanders, Christine Klymko, and Van Emden Henson. “An ensemble framework for detecting community changes in dynamic networks." *High Performance Extreme Computing Conference (HPEC), 2017 IEEE*. IEEE, 2017.

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

LLNL-POST-752913

Colors to use in charts and graphs

Charts and Tables

Since there will be a lot of information on your poster, keeping tables and charts simple and easy to understand is much better than an in-depth and complicated graphic. There are a few simple tricks to simplify charts and tables:

1. Remove unnecessary borders, lines, drop shadows, and backgrounds
2. Remove unnecessary labels and markers
3. Use flat styles and solid colors
4. Highlight important values