HI!

All You Need Is PageRank.

Metapopulation model is done, and now we dive into the topology of commuter flows between geographies. A brief overview of problems addressed in this journey.

We have an aggregated month of intercity commutes between Russian cities. Flights, trains, buses. It makes a graph of nearly 1K nodes connected by 20K weighted edges. First thing to do was an implementation of that Metapopulation model. Each city is treated separately plus we consider traffic flows.

Next question is funnier. Which city has the biggest impact on the dynamics? In case it is the only city - PageRank is All You Need. As we want to find several cities that maximize spread, the question is trickier. This problem is NP-complete. Good news is there are good approximations. We use this recent Stop-and-Stare algorithm and find sub-optimal places to control (or block).

Next, we look at higher order structures - motifs. We know that in nature there are recurring patterns of interaction. We also know these patterns are not random. For example, there are 13 possible directed structures of 3 nodes. The test for non-random nature is simple - count motifs, rewire the original network so that nodes preserve their degrees, and count motifs. Randomize many times and get expected frequencies of structures to appear at random.

An important property of networks is structural roles - nodes are embedded differently. So we take counts of motifs around every node and use them as feature vectors.

For every node we look at neighbors. And we combine their features. First we do so linearly - that makes a variant of RolX algorithm.

In order to check the compactness hypothesis we first project these multi dimensional feature vectors into 2D space with t-SNE algorithm. And we also run k means in order to identify the number of structural roles. Those clusters are reasonable - Moscow falls into a cluster of one, two major cities - St. Petersburg and Yekaterinburg go together and so on.

Does it make sense? It seems so - these violet-colored nodes go along the trans-siberian railway.

The disadvantage of RolX is that every aggregation doubles the feature vector size. Yep, the same story of exponential growth. So we go for Graph Neural Networks that make better features of smaller size. At least we know that in graph-related machine learning problems these outperform everything.

Our ultimate goal is to be able to compare topology between counties and regions. Ironically, the super-novel Graph Neural Networks implement a 50-something year old russian algorithm of a space exploration era. Weisfeiler-Lehman test of isomorphism is able to distinguish between most of the real-world graphs.

That is how we address the core issue: Compare topology of population and commuter flows between geographies. So that authorities of alike communities can exchange relevant knowledge. Just because tactics for a village vary from those for a city.

This project has a side quest.

1. Let's get it squared.

2. Medics are today's rock stars. Working the front line.

3. Michael Jordan said: "Talent wins games, but teamwork and intelligence wins championships".

4. I can hardly find examples of more regulated jobs.

5. These days they experience major disruption

of everything

6. Normal organization structure and processes

7. Were meant for normal times.

8. Things have changed though.

9. Organizations restructure in order to respond to increased demand. It is pure luck that many medical workers have several specializations and can jump-start into action.

We tested similarity measures for unsupervised settings. Dialectic method states these can measure structural differences.

10. Change of structure makes people suffer. Who is going to break down?

Definitely, not them. I put these two issues together as we have reached a point when realities collide. Let’s move to the questions.