Alternative Symmetrizations of Hitting Times in Graphs

Timothy Chu
Carnegie Mellon University
tzchu@andrew.cmu.edu

Gary Miller
Carnegie Mellon University
glmiller@cs.cmu.edu

November 20, 2018

Abstract

In this paper, we list out a number of proposals for metrics on graphs. We also sketch out impossibility results. By default, edge weights are conductances and not resistances. The properties we would like to have are:

1. A metric ball is connected. This is of use in machine learning, when the underlying graph is local (or at least semi-local).
2. If our graph is a line: $A - x$, $x - B$, then A and B should have some distance between them. If x is connected to a large graph, but A is connected only to x and B only to x, then either A and B get closer, or A and B stay the same distance.
3. Rayleigh monotonicity property.
4. A model that has some kind of physical or simple mathematical meaning. We speculate such a model has more interesting mathematics underneath.

We examine the use of htiting times, commute times, voltages, p-voltages, maximum flow, low conductance cuts, shortest paths, and more. The number of ways we can approach this problem is manifold:

1. Determine when a vertex is central.
2. Determine when a point x is closer to A than to B.
3. Build balls to fixed point A, and then create a symmetrized measure from there. (Dijkstra, doing all pairs shortest path at once).
4. Build a measure that is some function of its neighbors.
5. Make a differential equation that dissipates things faster than heat (like if there were positively charged protons at every vertex or something).

It seems like I would need an alternative strategy. Level sets and low conductance cuts don't quite fit into this mold.

It seems as if I would need to be very creative, or have inspiration from some other sources, (crystallography, glass-making, gravitation, strong and weak forces, Maxwell,

fluid mechanics, terrestrial mechanics, microbial action, neurons, humanities, or more) because many machine learners have undoubtedly looked at this problem. Thus far, I have also been – unnecessarily – negative on things like spectral clusering or diffusion maps, and it seems worthwhile to look into that line of work and their limitations. Here's a few approaches that Tim Chu has tried, doesn't believe them to work, and will list out why they don't work:

1. Heat half lives
2. Voltage level sets
3. Time weighted heat flow.
4. New symmetrization of heats.
5. Laplacian powers for Effective Resistance. (?)
6. Allowing the net flow at the start and end point to be distributed, at some cost.
7. ????

Some questions:

1. It would seem as if commute times should fail the $A-x$ $x-B$ test but they don't.

Heat half lives have the property that they are connected. In this note, we quickly determine whether they satisfy Rayleigh.

Graph: Take the line times 100, and take a clique and a single edge from the middle of the line to that clique. Picture attached. The dotted line is the edge that we will attach.

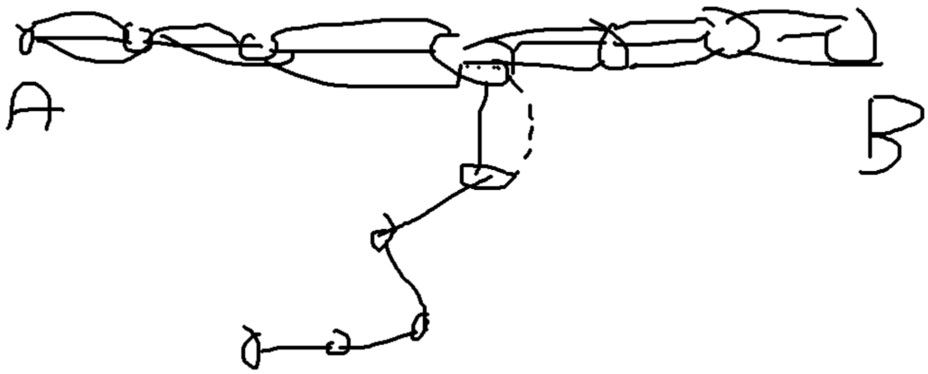

Figure 1: