Multiscale Decomposition of Collective Organization in Nature and Society

Zach Gelbaum[1], Mat Titus[1], James R. Watson[1]

January 7, 2019

[1] The Prediction Lab LLC, Corvalllis, OR, USA, 97330

1 Introduction

Graphs and networks provide a natural framework for studying the dynamics of complex systems of many interacting agents. Examples include biological groups such as fish schools or insect colonies, sports teams, or networks of financial institutions such as banks and other lenders. These systems can be described using a dynamically evolving network whose edges represent the interaction between agents, each agent has its own dynamic for its individual evolution, and additional information can be given as dynamic signals associated to each agent.

The importance of scale is implicit in any study of collective behavior, and the most natural approach to multiscale analysis is through wavelets. In the setting of graphs, one can use the framework of spectral graph wavelets to decompose a dynamic network as above into the dominant scales of variation locally at each vertex, thus revealing multi-scale geometric structure essential for understanding the collective evolution of the system. Moreover, by using Hermitian graph wavelets, this multiscale analysis can be directly related to heat diffusion on the network and thus, and as is now well known, directly related to the geometric structure of the network. Such geometric formulations of dynamics are information rich and provide highly efficient representations of the key structural features of a network through time.

Many such systems have a hierarchical structure of influence, where some agents exert greater influence on the dynamics of the system than others. This could be a subset of animals that act as leaders for the group, highly connected banks that strongly influence the behavior of a financial system, or a captain on a sports team, or the hierarchy may be relatively "flat," with agents acting autonomously and exerting a close to equal influence on the overall group, such as with a swarm of autonomous robots or drones. Moreover, these hierarchies can evolve in response to changes in the environment in which the system is operating, or according to an internal dynamic of the group.

A problem of central importance in the study of all of these systems is the quantification of this collective organization using a general, data driven methodology. Often, measures of graph centrality are used. In (REF), the authors prove that for a system of agents in a stochastic environment attempting to measure and respond to an ambient signal, the optimal agent for measuring the ambient signal and communicating that information to the rest of the network (the "leader") is determined by maximizing so called "information centrality" (IC). This result strongly suggests that IC should be utilized in analyses of hierarchical leadership structure. However, IC is not easily interpretable in a quantitative way and it is not clear how a given value of IC for an agent relates to the rest of the network structure, or can be compared to values of IC at different points in time as the network evolves. Moreover, once one has identified a leader agent, one would like a way to quantify how strong an influence that leader is exerting on the rest of the team on an agent by agent basis; who is this agent influencing and how much, through time?

To address these questions, we have developed a new a wavelet based measure of centrality, "scale centrality". We have shown that minimizing scale centrality is equivalent to maximizing information centrality, so scale centrality identifies the same leaders. However, because our measure is derived from Hermitian graph wavelets and related to the most important scales and geometric structure of the network, its use yields much more information. Using scale centrality, not only can we identify key agents exerting outsize influence on the system, we can associate to each agent its primary scale of influence (the leader having the largest), and from this scale directly quantify the influence of each agent on the rest of the system through time, thus quantifying the hierarchical structure of influence between agents in the network as it changes through time.

2 Hermitian Graph Wavelets, Geometric Learning and Scale Centrality

Networks can be naturally viewed as discrete geometric objects, and many of the classical tools for studying the geometric structure of surfaces and manifolds can be extended and applied to networks. Such tools form the basis for much of manifold learning, whereby information rich geometric structure is identified in complex data sets. The most powerful techniques in geometric learning are spectral in nature, based on various Laplacian operators and their spectral decompositions. Moreover, many of these techniques are multi scale in nature.

Spectral graph wavelets are part of this spectral geometric toolkit and provide a localization of the above global multiscale geometric learning techniques. However, one critical difference is that other methods based on Laplacian eigenfunctions do not allow one to identify which scales matter most; how does one choose which scales to look at? Once one has chosen to use wavelets, another key question is what wavelet kernel to use, as this is a choice left to the user. We have developed a particular wavelet kernel closely related to the heat kernel that we call Hermitian Graph Wavelets. Hermitian Wavelets not only describe the local, multiscale geometric structure of networks, but also characterize the dominant scales of importance of each agent in a network, as well as any signals produced by the agents in a network, and can do all of this in a dynamically varying system. Hermitian wavelets also provide a clear meaning to the scale parameter in terms of natural measures of distance on the network, and in determining the effective scale of an agent in the network can quantify precisely the influence of an agent on every other agent in the network and track it through time.

2.1 Our Wavelets

Given a network represented by a graph with adjacency A, often one also has additional data in the form of one or more functions $f : V \to \mathbb{R}^n$, which assigns to each vertex a value. For example, f could be a a component of a vector describing the internal state of each agent, e.g., vital signs etc. Such a function is often referred to as a *signal* on the network in analogy with classical signal processing. We can consider the set of all signals on G as a vector space $L^2(V)$ with inner product given by

$$\langle f, g\rangle = \sum_{x \in V} f(x)g(x),$$

and where as usual we define $\|f\|^2 = \langle f, f\rangle$. The graph Laplacian is defined as an operator on this space of signals as follows: Let D be the row (or column) sum of A, meaning that $D_{i,j} = 0$ for $i \neq j$ and $D_{i,i} = \sum_j A_{i,j}$. The Laplacian is given by

$$\Delta = D - A.$$

In the general framework of (REF) one constructs spectral graph wavelets as kernels of operators given by $g(s\Delta)$ where $s \in (0, \infty)$ is a continuous scale parameter and $g : [0, \infty) \to \mathbb{R}$ satisfies certain conditions. For fixed a fixed scale s and $x \in V$ the wavelet kernel takes the form

$$\psi_{s,x}(y) = \sum_{k=0}^{N} g(s\lambda_k)\phi_k(x)\phi_k(y).$$

. For a given signal f, we can decompose it into its wavelet coefficients $W_f(x, s) = \langle f, \psi_{x,s}\rangle$,

$$f = \sum_{s,x} W_f(x, s)\psi_{x,s}.$$

While this decomposition is not orthogonal, it does satisfy frame bounds, giving an approximate Parseval relation: there exist constants $0 < A < B$ such that

$$A\|f\|^2 \leq \sum_{s,x} |W_f(x, s)|^2 \leq B\|f\|^2.$$

Moreover, with enough properly chosen scales B can be brought arbitrarily close to A.

2.1.1 Heat Kernel

In order to define a wavelet with a clear relationship of the scale parameter to the geometric structure of the network, one must choose an appropriate g. For this we turn to the heat kernel, defined as the integral kernel of the heat semigroup, $e^{-t\Delta}$ (REF):

$$H_t(x,y) = \sum_{k=0}^{N} e^{-t\lambda_k}\phi_k(x)\phi_k(y).$$

The term "heat" refers to the fact that the heat semigroup when applied to any function $f \in L^2(G)$ solves the heat equation:

$$\frac{d}{dt}e^{-t\Delta}(f) = \Delta e^{-t\Delta}(f).$$

When f is taken to be a point mass δ_x, one obtains the heat kernel. One of the most important features of the heat kernel is the well known fact that its decay reflects the geometry of G in terms of so called intrinsic metrics. This relationship between the heat kernel and geometry is made precise in the following result due to (REF): If we let $\rho(x,y)$ denote an intrinsic metric on our network, then

$$\zeta_s(t,r) = \frac{1}{s^2}\left(rs \cdot arcsinh\frac{rs}{t} - \sqrt{t^2 + r^2s^2} + t\right),$$

then we have

$$H_t(x,y) \leq e^{-\zeta_s(t,\rho(x,y))}$$

where s is the jump size of the intrinsic metric ρ. Note the bounds dependence on the choice of intrinsic metric. Such intrinsic metrics can encode various different geometries. For example, given a set of interacting agents, one could consider a number of different interaction networks among them and for each such network define an intrinsic metric that captures a certain notion of distance, perhaps focused on a particular aspect of the interaction. Thus the appearance of the intrinsic metric above as a sort of free parameter serves to greatly increase the possibility for comparing different geometric structures with rich meaning and interpretation.

2.1.2 wavelet localization

Using the above bounds we can obtain analogous bounds for our wavelet kernel by following the method of proof in (REF), in particular one can obtain the following bounds (with the same notation as above and letting $r = \rho(x,y)$):

$$|\psi_{t,x}(y)| \leq \left[\frac{r^2}{t}\left(1 + \frac{s}{\sqrt{t^2+s^2r^2}}\right)\left(\frac{1}{sr + \sqrt{t^2+s^2r^2}}\right) - \left(\frac{t}{\sqrt{t^2+s^2r^2}} + 1\right)\right] e^{-\xi_s(t,r)}$$

This bound, which to our knowledge is new, gives precise quantitative information on the decay of $|\psi_{t,x}(y)|$ as $\rho(x, y) \to \infty$ and ensures sharp localization of wavelets. Note however the proof depends essentially upon our choice of g and its relation to $H_t(x, y)$, which was a primary motivation for our choice.

Another benefit of our choice of wavelet is the following interpretation of scale it provides: As mentioned above the heat kernel provides local, multiscale structural information on the network, but does not distinguish between scales in a manner allowing for choosing the most important. With our choice of wavelet, for each vertex x we define the dominant scale $DS(x)$ as the value of t that maximizes $\|\psi_{t,x}\|^2$. This is possible as one can show, e.g. using the above bounds or from reasoning directly from the heat equation, that $\lim_{t\to^+0} |\psi_{0,x}(y)| = 0$ for all x, y and $\lim_{t\to\infty} |\psi_{t,x}(y)| = 0$, which together with the fact that ψ is continuous on $[0, \infty) \times G \times G$ tells us it must achieve a maximum, $DS(x)$. That this maximum is unique follows from the maximum principle of the heat equation. We can interpret this scale in the following way: $\|\psi_{t,x}\|^2$ measures the total energy of the derivative of $H_t(x, \cdot)$ over the network. When this is maximum is when heat is maximally flowing from vertex x to the rest of the graph, i.e., is the scale at which it is most strongly influencing the rest of the network.

2.2 Scale Centrality

Given the above notion of dominant scale, it is natural to attempt to determine leaders based on their scale of influence. We define scale centrality as follows: Given an agent x and Hermitian wavelet $\psi_{s,x}(y)$, we let

$$SC(x) = \int_0^\infty \|\psi_{s,x}\|^2 ds.$$

. From this definition one can show that the agent with maximum IC is the agent with minimum SC,

$$\arg\min_{x\in G} SC(x) = \arg\max_{x\in G} IC(x)$$

Note that $SC(x)$ does not define a specific scale. However here again the special nature of Hermitian graph wavelets are essential: for large classes of graphs, including those arising in complex systems of agents, the function $s \mapsto \|\psi_{s,x}\|$ takes a specific shape for each x. In particular, one can show that for large classes of graphs $\arg\min_{x\in G} SC(x) = \arg\min_{x\in G} DS(x)$. This result matches our intuition: if an agent is highly influential, is effective scale will be small, because what this means is that it takes a very short time for its influence to diffuse to the rest of the network. By contrast, an agent that is isolated and not interacting strongly with the rest of the network with have a large effective scale, indicating that it takes a longer time for its influence to diffuse out to the rest of the network.

3 Application to Hierarchical Leadership Structure in Biological Collectives and Competitive Sports Teams

From the above discussion, we see that small effective scale $DS(x)$ means an agent is highly influential, thus providing an intuitive understanding of leadership in terms of geometric scale. Given $DS(x)$ for a leader x, one can then consider e.g. the values of $H_{DS(x)}(x, y)$ to measure the influence of a leader agent x on agent y. In this way not only can we sort agents according to their effective scale of influence, **we can also measure and plot each agents influence on every other agent, and we can do all of this through time as the network evolves**.

As a first application, we will study two datasets of schooling fish. In the first, a subset of fish have been trained to be leaders, acting more independently of the school and moving towards food sites on their own. In this way these fish tend to "lead" the rest of the school. Using the above analysis we will first identify these fish, and then we will track their influence through time, thus obtaining a dynamic measure of hierarchy of the fish school through time. As a second test, we will apply the same methods to a dataset of untrained fish, where the school evolves more naturally and alternates between various regimes of behavior. We will study the transitions between these different regimes and identify the fish driving the transition. Preliminary studies of these transitions suggest a few fish drive drive each transition. Identifying these will be a first step towards looking for any early indications of which fish will be the leaders for a given transitions. For another application, we plan to apply these methods to data collected from competitive sports. Specifically, we have the location of players and the ball from a soccer game (the Premier League in the UK) and will analyze patterns of leadership (and who are the followers) over the period of the game.

In addition to the application of Scale Centrality to these empirical datasets, we will explore its utility in situations where nodes in the graph come and go, and the size of the graph changes over time, and when there is uncertainty over edge weights.