THE GEOMETRY OF SOCIAL NETWORKS

ANTHONY BONATO

Abstract. We propose a broad geometric view of real-world complex networks, inspired by seminal work on Blau space for social networks. We describe how the examination of mathematical models for on-line social networks led to a hypothesis regarding their dimension. A thesis is presented on how all complex networks possess an underlying geometry related to but distinct from their underlying graph distance metric space.

1. Introduction

Online social networks (or OSNs) such as Facebook, Twitter, and Instagram are highly prevalent in our daily lives and commerce. The availability of social networking data has witnessed an active multi-disciplinary trend towards the modelling and mining of these big data sets.

Not unlike earlier studied complex networks such as the web graph (see [2]), OSNs exhibit several common evolutionary properties. OSNs possess the *small world property*, which demands that there is a short path joining any two nodes, and if two nodes share a common neighbor, they are more likely to be adjacent. OSNs have power law degree distributions. In a graph G, let N_k denote the number of nodes of degree k. The degree distribution of G follows a *power law* in some range of k if N_k is proportional to k^{-b}, for a fixed *exponent* $b > 2$. Other properties include densification power laws (where the network becomes more dense over time), and constant or even shrinking distances over time.

Several stochastic models have emerged simulating the evolution of OSNs. Such models serve to simulate known properties of social networks, and suggest new ones. We focus in this note on *geometric models* of OSNs; that is, models which posit nodes as points in fixed metric space, and edges are determined by various probabilistic rules along with their relative proximity in the space. We present the view that the geometry is indeed a fundamental construct of social networks, and is in alignment with other central role of geometric models in the physical or natural sciences.

2. Blau Space and models for OSNs

Sociologists considered a geometric view of social networks long before the advent of OSNs. In *Blau space* [7], nodes correspond to points in a multi-dimensional metric space, and the link structure is governed by the principle of *homophily*: nodes with similar socio-demographic attributes are more likely joined.

Blau space suggests a view towards random geometric graphs as natural objects for the modelling of OSNs. Random geometric graphs are well-studied within discrete mathematics and graph theory. In such stochastic models, nodes are points in a fixed metric space $\mathcal{S}$, and are each assigned a *ball of influence* in $\mathcal{S}$. Nodes are joined with some prescribed probability

The author gratefully acknowledge support from NSERC.

if they arise in each other's ball of influence. Several geometric models for OSNs and complex networks have independently and recently emerged; see [1, 5, 6, 10].

We focus here on the MGEO-P model which uses both geometry and ranking first presented in [3]. The MGEO-P model has five parameters: n the total number of nodes, m the dimension of the metric space, $0 < \alpha < 1$ the attachment strength parameter, $0 < \beta < 1 - \alpha$ the density parameter, $0 < p \leq 1$ the connection probability. We identify the agents of an OSN with points in $\mathbb{R}^m$ with the L_∞-metric, each chosen uniformly at random in the unit hypercube (with the torus metric to avoid boundary effects). Nodes are ranked by their popularity from 1 to n, where n is the number of nodes. Here, 1 is the highest ranked node with n the lowest. Each node has a ball of influence that is a function of their rank (along with α and β); the higher the rank, the larger the volume of the ball of influence. A node is joined to another with probability p if it is in its ball of influence. See Figure 1.

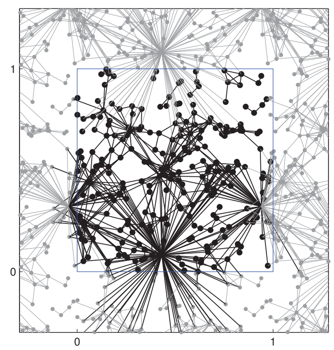

FIGURE 1. A simulation of the MGEO-P model resulting in graph with 250 nodes in the unit square with torus metric. Each figure shows the graph replicated in grey on all sides in order to illustrate the torus metric. The blue square indicates the unit square.

As shown in [3], MGEO-P model provably generates graphs satisfying the following properties with probability tending to 1 as n tends to ∞.

(1) The degree distribution follows a power law with exponent $1 + \frac{1}{\alpha}$.
(2) The average degree of node is
$$\frac{p}{1-\alpha} n^{1-\alpha-\beta}(1 + o(1)).$$
(3) The diameter (that is, length of a longest shortest path connecting vertices) is $n^{\Theta(\frac{1}{m})}$.

3. The Logarithmic Dimension Hypothesis and the Feature Space Thesis

Property (3) of MGEO-P on its diameter suggests that, ignoring constants, for a network with n nodes and diameter D, the expected dimension based on the MGEO-P model is
$$m \approx \frac{\log n}{\log D}.$$
Hence, the model predicts a logarithmically scaled dimension of Blau space. Interestingly, this was proposed independently in at least two other models for OSNs including those in [**?**, 10]. The logarithmic relationship between the dimension of the metric space and the number of nodes has been called the *Logarithmic Dimension Hypothesis* (or *LDH*) [4].

The study [3] set out to experimentally verify LDH, using data sets from Facebook (the so-called "Facebook 100" data set [9]) and LinkedIn. Employing machine learning algorithms using motif counts (that is, counts of small subgraphs) and analysis of eigenvalue distributions, dimensions of real OSN data sets were predicted. While the results of [3] do not conclusively prove the LDH, they are suggestive of a logarithmic or sub-logarithmic dimension for Blau space. For future work, it would be useful to further validate the LDH with additional data sets. The challenge is to find representative samples of OSN data that scales with time (as found in the Facebook 100 data sets).

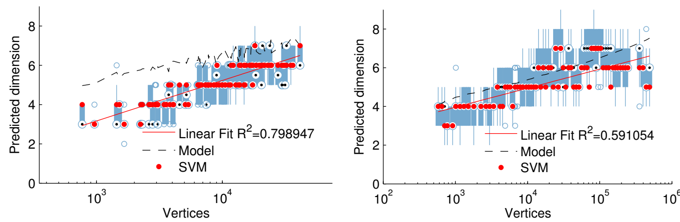

FIGURE 2. Facebook dimension at left, LinkedIn dimension at right. Each red dot is the predicted dimension computed via motif counts and a support vector machine classifier. MGEO-P predicts a dimension of $\frac{\log n}{\log D}$, which is plotted as the dashed line.

The study of Blau space inspires a geometric view not only to OSNs, but also other complex networks. We propose here the *Feature Space Thesis* which states that every complex network has an underlying metric (or *feature*) space, where nodes are identified with points in the feature space, and edges are influenced by node similarity and proximity in the space. As with Blau space, the metric of the feature space is not arbitrary, but is rather a hidden property of the network. For example, in the web graph there is an underlying *topic space*, with web pages more closely related in topics more likely to be linked. In protein-protein interaction networks, we may view these as embedded in a *biochemical space*; see [8] for more on the geometry of protein networks.

The Feature Space Thesis is not provable, per se. However, it should be used as a guide towards a deeper understanding of complex networks. For instance, while agents may be close via graph distance in an OSN, they may be far apart in Blau space; the latter metric better reflects their different sociological profile. How do we infer the underlying metric of Blau space? How are the graph structure and feature space geometry related? We propose these questions for future work.

REFERENCES

[1] W. Aiello, A. Bonato, C. Cooper, J. Janssen, P. Prałat, A spatial web graph model with local influence regions, *Internet Mathematics* (2009) **5** 175–196.

[2] A. Bonato, *A Course on the Web Graph*, American Mathematical Society Graduate Studies Series in Mathematics, Providence, Rhode Island, 2008.

[3] A. Bonato, D.F. Gleich, M. Kim, D. Mitsche, P. Prałat, A. Tian, S.J. Young. Dimensionality matching of social networks using motifs and eigenvalues, *PLOS ONE* **9**(9):e106052, 2014.

[4] A. Bonato, J. Janssen, P. Prałat, Geometric protean graphs, *Internet Mathematics* **8** (2012) 2–28.

[5] M. Kim, J. Leskovec, Multiplicative attribute graph model of real-world networks, *Internet Mathematics* **8** (2012) 113–160.

[6] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, M. Boguñá, Hyperbolic geometry of complex networks, *Phys. Rev. E*, **82**:036106, 2010.
[7] J.M. McPherson, J.R. Ranger-Moore, Evolution on a dancing landscape: Organizations and networks in dynamic blau space, *Social Forces* **70** (1991) 19–42.
[8] V. Memišević, T. Milenković, N. Pržulj, An integrative approach to modeling biological networks, *Journal of Integrative Bioinformatics*, **7**:120, 2010.
[9] A.L. Traud, P.J. Mucha, M.A. Porter, Social structure of facebook networks, *arXiv*, cs.SI:1102.2166, February 2011.
[10] Z. Zhang, F. Comellas, G. Fertin, L. Rong, Highdimensional apollonian networks, *Journal of Physics A: Mathematical and General*, **39**:1811, 2006.

Ryerson University
E-mail address: abonato@ryerson.ca