Feature-enriched hyperbolic network geometry

Roya Aliakbarisani[1,2], M. Ángeles Serrano[1,2,3], and Marián Boguñá[1,2]

[1] Departament de Física de la Matèria Condensada, Universitat de Barcelona, Martí i Franquès 1, E-08028 Barcelona, Spain ,

[2] Universitat de Barcelona Institute of Complex Systems (UBICS), Barcelona, Spain

[3] Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, E-08010 Barcelona, Spain

Graph-structured data provide a comprehensive description of complex systems, encompassing not only the interactions among nodes but also the intrinsic features that characterize these nodes. These features play a fundamental role in the formation of links within the network, making them valuable for extracting meaningful topological information. Notably, features are at the core of deep learning techniques such as Graph Convolutional Neural Networks (GCNs). GCNs aggregate information from the neighborhood of each node in a graph, allowing them to propagate information and capture the graph topology. An implicit assumption made by GCNs is that there must exist correlations between connected (or topologically close) nodes in the graph so that they are similar, and similar nodes should share common features. Only when this is the case, GCNs can detect patterns in the data. Despite their undeniable effectiveness, GCNs are criticized for their lack of explainability, a problem referred to as the black box problem. To solve the black box problem, we must first understand in detail the structure of the data that feeds GCNs.

In this paper [1], we introduce a simple yet comprehensive framework to describe real graph-structured datasets. It employs the $\mathbb{S}^1$ model, also known as the geometric soft configuration model [2, 3], to describe the network between nodes $\mathscr{G}n$, reflecting the topology. The key aspect of our proposed framework is to treat features as tangible entities and to view the set of nodes and their features as a bipartite graph $\mathscr{G}_{n,f}$ connecting nodes to features. The objective is to develop a model for this bipartite graph $\mathscr{G}_{n,f}$ that is correlated with the network of topology $\mathscr{G}_n$. To achieve this, we propose a geometric model called the bipartite-$\mathbb{S}^1$ model [4, 5] where the similarity space is shared between $\mathscr{G}_n$ and $\mathscr{G}_{n,f}$. In this model, similar to the $\mathbb{S}^1$ model, each node is assigned two hidden variables $(\kappa_n, \theta n)$, where κ_n denotes the expected degree of the node in the bipartite graph, and θ_n is its angular coordinate on a one-dimensional sphere, equal to that of the node in $\mathscr{G}_n$. Similarly, features are equipped with two hidden variables (κ_f, θ_f), indicating their expected degrees and angular positions in the common similarity space. Then, nodes are connected to features with a probability depending on their distance. As the $\mathbb{S}^1$ model is isomorphic to a purely geometric model in the hyperbolic plane, the $\mathbb{H}^2$ model, we can map the bipartite-$\mathbb{S}^1$ model to the hyperbolic plane as well.

Through this framework, we can identify correlations between nodes and features in real data and generate synthetic datasets that mimic the topological properties of their connectivity patterns. The approach provide insights into the inner workings of GCNs by revealing the intricate structure of the data. In the experimental results, we compare the topological properties of real networks, including Cora, Facebook, Citeseer, and Chameleon with their synthetic counterparts generated by the bipartite-$\mathbb{S}^1$ model and demonstrate that the model accurately replicates the topological features of these real networks.

References

1. Aliakbarisani R., Serrano, M.Á., Boguñá, M.: Feature-enriched hyperbolic network geometry. arXiv:2307.14198v2 (2023)
2. Serrano, M.Á., Krioukov, D., Boguñá, M.: Self-Similarity of Complex Networks and Hidden Metric Spaces. Phys. Rev. Lett. 100(7), 078701 (2008)
3. Krioukov, D., Papadopoulos, F., Kitsak, M., Vahdat, A., Boguñá, M.: Hyperbolic geometry of complex networks. Phys. Rev. E 82(3), 036106 (2010)
4. Serrano, M.Á., Boguñá, M., Sagués, F.: Uncovering the hidden geometry behind metabolic networks. Mol. Biosyst. 8(3), 843 (2012)
5. Kitsak, M., Papadopoulos, F., Krioukov, D.: Latent geometry of bipartite networks. Phys. Rev. E 95(3), 032309 (2017)