# Dethroning Aristocracy in Graphs via Adversarial Perturbations – (Student Abstract)

**Adarsh Jamadandi and Uma Mudenagudi**
KLE Technological University, Hubli, India.
{adarsh.jamadandi, uma}@kletech.ac.in

## Abstract

Learning low-dimensional embeddings of graph data in curved Riemannian manifolds has gained traction due to their desirable property of acting as effective geometrical inductive biases. More specifically, models of Hyperbolic geometry such as Poincaré Ball and Lorentz/Hyperboloid Model have found applications for learning data with hierarchical anatomy. Gromov's hyperbolicity measures whether a graph can be isometrically embedded in hyperbolic space. This paper shows that adversarial attacks that perturb the network structure also affect the hyperbolicity of graphs rendering hyperbolic space less effective for learning low-dimensional node embeddings of the graph. To circumvent this problem, we introduce learning embeddings in pseudo-Riemannian manifolds such as Lorentzian manifolds and show empirically that they are robust to adversarial perturbations. Despite the recent proliferation of adversarial robustness methods in the graph data, this is the first work exploring the relationship between adversarial attacks and hyperbolicity while also providing resolution to navigate such vulnerabilities.

## Introduction

Graph nodes embedded in Euclidean space incur large distortions and are unsuitable for modeling real-world graphs such as Social, Internet or Biological Networks, this has prompted for introduction of curved Riemannian manifolds such as hyperbolic geometry to model hierarchical data (Nickel and Kiela 2017). Gromov's $\delta$-hyperbolicity measures the *Tree-likeliness* of the graphs, which helps decide if a given graph is suitable for embedding in hyperbolic space. The lower the $\delta$-hyperbolicity, the more hyperbolic the graph is ($\delta = 0$ for Trees) and more suitable hyperbolic space is, as an embedding space.

Authors (Borassi, Chessa, and Caldarelli 2015) have coined the term 'aristocratic' to graphs with smaller hyperbolicity, indicating only few vertices controlling a larger aspect of the network. In this paper, we show that introducing adversarial perturbations that disrupt the network structure, also destroys the hyperbolicity of graphs, effectively making hyperbolic geometry less effective as an embedding space and dethroning *aristocracy*.



Adversarial attacks are deliberate data perturbations that degrade Deep learning models' performance drastically. The idea has been recently extended to explore the robustness of graph models as well (Bojchevski and Günnemann 2019). However, this is the first work that explores the relationship between adversarial attacks on unsupervised node embeddings that use models of hyperbolic geometry as embedding space and hyperbolicity. We hypothesize and empirically show that embedding graphs in Lorentzian manifolds, is more robust to vulnerabilities.

## Gromov's Hyperbolicity

Gromov's ($\delta$) hyperbolicity introduced by (Gromov 1987) helps ascertain if the graph is inherently hyperbolic. Mathematically, we define hyperbolicity as - Let $\{a, b, c, d\}$ be the vertices of the graph $G(V, E)$ and let ($\mathbb{S} = \{S_1 = d(a, b) + d(d, c)\}$, $\{S_2 = d(a, c), d(b, d)\}$, and $\{S_3 = d(a, d) + d(b, c)\}$). The $\delta(a, b, c, d)$ is given by

$$\delta(a, b, c, d) = \frac{1}{2} \max_{\{a,b,c,d\} \in V(G)} \text{hyp}(a, b, c, d) \quad (1)$$

where, hyp$(a, b, c, d)$ = Difference of two largest values in $\mathbb{S}$. The graph $G(V, E)$ can be viewed as a metric space with $d(.)$ giving distance (geodesic) between vertices.

## Attack Model

We use the attack model introduced by authors (Bojchevski and Günnemann 2019) -

$$\hat{A}^* = \arg \max_{\hat{A} \in \{0,1\}^{N \times N}} \mathcal{L}(\hat{A}, Z^*) \quad (2)$$

where, $Z^* = \min_Z \mathcal{L}(\hat{A}, Z)$, subjected to $||\hat{A} - A||_0 = 2f$. We assume the attacker is restricted to modifying only few entries $f = ||\hat{A} - A||_0 = 2f$ of the adjacency matrix A resulting in $\hat{A}$. The perturbations introduced strive to bring down the quality of the embeddings, this is measured by the loss function $\mathcal{L}(\hat{A}, Z)$ of the model under attack. In practice, the final outcome is either addition/flipping of random edges, resulting in *poisoning* of the graph under consideration.

## Results and Discussions

We advocate for using Lorentzian manifolds as embedding space for graph data. We consider two types of Lorentzian

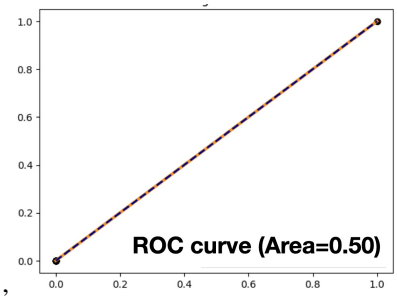



Figure 1: We try to embed PolBlogs in the Poincaré disk model. It is evident from the graph, that hyperbolic space fails to embed graphs with high hyperbolicity.

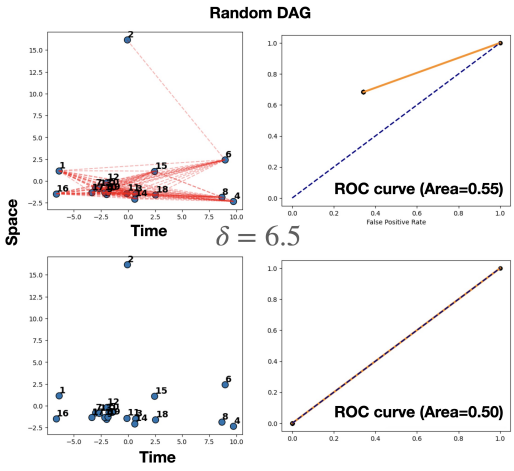
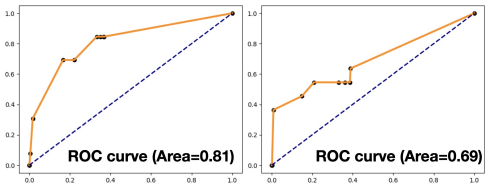
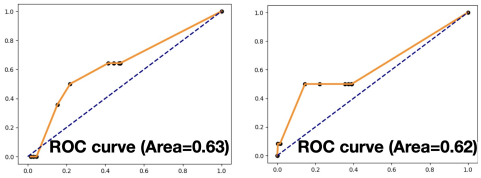



Figure 2: For a random Directed Acyclic Graph with $\delta = 6.5$, we try to embed in both de Sitter space (Top Row) and hyperbolic space (Bottom Row). From the figure, its evident that, graphs, especially DAGs are more naturally embedded in Lorentzian manifolds.




(a) (L)Cora before attack. (R)Cora after attack.




(b) (L)Citeseer before attack. (R)Citeseer after attack.

Figure 3: The first column shows the accuracy of graph embeddings quantified by the AUC curve before introducing the adversarial perturbations. The second column shows the accuracy after poisoning the graph.

| Dataset | EdgeFlips | $\delta_{old}$ | $\delta_{new}$ |
|---|---|---|---|
| Cora | 1000 | 2.0 | 2.5 |
| Citeseer | 1000 | 3.0 | 3.5 |
| PolBlogs | 1000 | 1.0 | 1.5 |

Table 1: We compute -hyperbolicity of standard graph data sets before and after introducing adversarial perturbations. Its evident from the Table below that hyperbolicity increases for random adversarial attacks rendering embedding in hyperbolic space ineffective.

manifolds and employ Multi-dimensional scaling algorithm (Clough and Evans 2017) to embed datasets - Cora, Citeseer and Polblogs.

1. Minkowski space - A 4-dimensional pseudo-Euclidean space with three spatial dimensions and one time dimension. $\{x_i^0, x_j^0\}$ represent the time co-ordinates and $\{x_i^k, x_j^k\}$ represent spatial co-ordinates and $c$ is the speed of light, which indicates the speed of flow of information in this case. The metric is given by,

$$d_{\mathbb{M}_{i,j}} = -c^2(x_i^0 - x_j^0)^2 + \sum_{k=1}^{d}(x_i^k - x_j^k)^2 \tag{3}$$

2. de Sitter space - A maximally symmetric Lorentzian manifold. Its distance metric is given by,

$$d_{deS_2}(\mathbf{x}, \mathbf{y}) = \lambda \operatorname{arcosh}\left(\frac{-\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{L}}}{\lambda^2}\right) \tag{4}$$

Table 1 shows the increase in hyperbolicity for 1000 random edge flips. Figure 3a and b, shows the accuracy of graph embedding in the Minkowski space before and after poisoning the graph structure. We compare the performance with graph embedding in Poincaré disk model given by

$$\begin{aligned}\cosh(\zeta d_{i,j}) = \cosh(\zeta r_i)\cosh(\zeta r_j) - \\ \sinh(\zeta r_i)\sinh(\zeta r_j)\cos(\pi - |\pi - |\theta_i - \theta_j|)\end{aligned} \tag{5}$$

From Figure 1, it is clear that, hyperbolic space fails to accurately embed the graph. To further test the capacity of Lorentzian manifolds, we generate a random DAG with $\delta = 6.5$ and embed the graph in both de Sitter space and Poincaré disk. from Figure 2 we can see that de Sitter space fairs better in handling graphs with high hyperbolicity.

## References

Bojchevski, A.; and Günnemann, S. 2019. Adversarial Attacks on Node Embeddings via Graph Poisoning. In *ICML*.

Borassi, M.; Chessa, A.; and Caldarelli, G. 2015. Hyperbolicity measures democracy in real-world networks. *Phys. Rev. E* .

Clough, J. R.; and Evans, T. S. 2017. Embedding graphs in Lorentzian spacetime. *PLOS ONE* .

Gromov, M. 1987. *Hyperbolic Groups*. Springer New York.

Nickel, M.; and Kiela, D. 2017. Poincaré Embeddings for Learning Hierarchical Representations. In *NeurIPS*.