



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

APEROL from Networks: Analyzing Pipeline and Embedding Representations for Optimized Learning (from Networks)

Annunziata Marco, Mondin Silvia, Turola Sveva

Computer Engineering MEng, 2025

Contents

1	Motivation	1
2	Datasets	2
3	Methods	3
4	Experiments	4
5	References	5

1 Motivation

Graphs have emerged as a natural model for the representation and analysis of data and complex systems in various domains. For instance, link prediction in social networks can help to understand the spreading process of rumors or epidemics [3], and in biological networks, it is commonly used for predicting novel interactions between proteins [2].

However, most traditional machine learning algorithms are not designed to work directly with graph-structured data, as they require numeric vectors or matrices as inputs. To address this challenge, graph embedding techniques have been developed to learn low-dimensional vector representations of nodes, edges, or entire graphs while preserving their topological properties.

Consequently, in order to design an effective graph embedding technique, it is necessary to consider several criteria [1]. Two of these are adaptability and scalability. An embedding is considered adaptable if it can be utilized for various data and tasks without the need for retraining. The scalability of an embedding is determined by its capacity to process large-scale networks within a reasonable amount of time.

The objective of this study is to evaluate known embeddings from the perspective of these two criteria.

From the perspective of adaptability, the objective is to determine the potential impact of the graph domain on the precision of an embedding. In other words, the objective is to ascertain whether there exist approaches that are better suited for a particular field, whether there is one embedding that consistently outperforms the others, or if the embeddings are comparable. Eventually, we would like to assess the adaptability of these embeddings when they are applied to different tasks.

In the context of scalability, the objective is to identify the embeddings that are better suited to process large graphs. The objective of this study is to evaluate the tradeoff between the accuracy of the predictions and the time required for training and inference.

2 Datasets

3 Methods

4 Experiments

Bibliography

- [1] Anthony Baptista et al. “Zoo guide to network embedding”. In: *Journal of Physics: Complexity* 4.4 (2023), p. 042001.
- [2] Björn H Junker and Falk Schreiber. *Analysis of biological networks*. Vol. 2. Wiley Online Library, 2008.
- [3] Romualdo Pastor-Satorras et al. “Epidemic processes in complex networks”. In: *Reviews of modern physics* 87.3 (2015), pp. 925–979.

Contribution of Authors and AI Usage