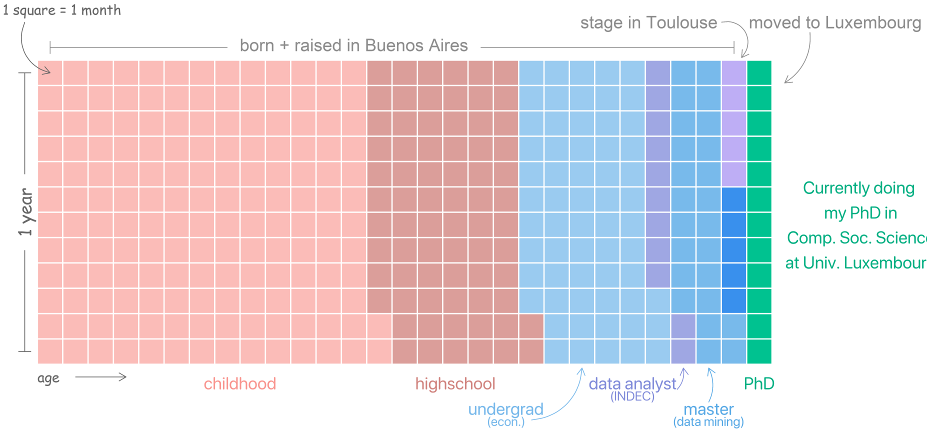
Machine Learning on Graphs

Diego Kozlowski

Faculty of Science, Technology and Medicine

November 18, 2020

about me



Currently doing my PhD in Comp. Soc. Science at Univ. Luxembourg

Networks

Deep Learning on Networks

Science of Science

Networks

Networks, or graphs, are a way of representing information where there is a **relation between elements**. The generality of this definition make it a proper tool for many different scenarios:

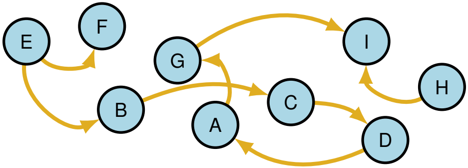
- In Social Media for example, the users interaction are know to build a Social Network, but also in real-life interactions between people in a company, school, village, etc. can be modeled with networks,
- in Chemistry the interaction between molecules can be thought of as a network
- in transportation, moving from one point to another can be represented as a network where nodes are intermediate points and edges represent the possibility of movement,
- in the economics, we can represent the banking system as a network of financial entities and money transfers, also international trade can be represented in this way, etc.

Network types

A network can either be:

- **Directed** if the edge between vertices has a one-way direction,
- **weighted** if edges are categorical or in $\mathcal{R}$,
- **attributed** if the nodes have features,
- **signed** if edges can take negative values,
- **heterogeneous** if there can be different types of edges/vertices. Heterogeneous networks can represent multiple types of relations. For example, *knowledge graphs* represent the relation between concepts. Something like *Paris is in France* could be represented as two nodes: *Paris* and *France* connected by the relation *is in*

Representation



We can represent a network using an **adjacency matrix**, where a 1 at the position *ij* represents that it exists a link between nodes *i* and *j*. **there is no natural order for rows and columns**.

$$
\begin{array}{c c}
 & \begin{array}{ccccccccc} A & B & C & D & E & F & G & H & I \end{array} \\
\begin{array}{c} A \\ B \\ C \\ D \\ E \\ F \\ G \\ H \\ I \end{array} &
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & 0 & 0 \\
0 & 0 & \boxed{1} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \boxed{1} & 0 & 0 & 0 & 0 & 0 \\
\boxed{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \boxed{1} & 0 & 0 & 0 & \boxed{1} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\end{array}
$$

Tasks on networks

In networks, given the flexibility of the representation, there exists many different tasks:

- network classification
- node classification
- link prediction
- community detection
- network generation

Tasks on networks

Besides the different possible tasks, there is also an important distinction:

- **transductive** framework: all nodes, their features and their links are visible during training, and only the labels of the test set are removed. This is *semi-supervised learning*
- **inductive** setting: Is a fully *supervised learning* problem. Nodes/Networks to classify are not seen during training. Is a more general, though more complex task.

Deep Learning on Networks

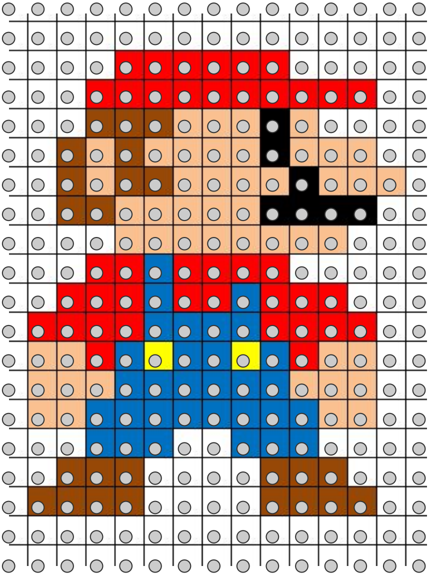
Problem

Traditional Deep Learning models for images or text depend on the regularities that not all networks have. The two main difficulties are:

- Nodes have a variable number of neighbors,
- neighbors cannot be ordered.

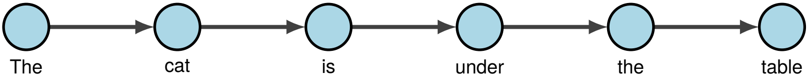
Images as networks

Images can be thought of as regular networks, where each node is connected to the adjacent pixels following a regular spatial pattern.



Sequences as networks

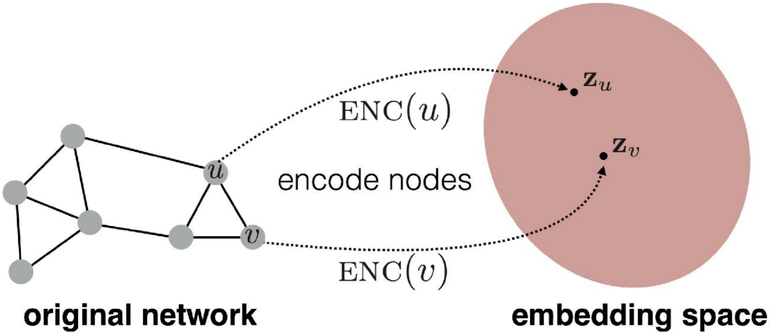
Also a text can be seen as a sequential network



The same can be said with any sequential data, like time series, sensor data, etc.

Embeddings

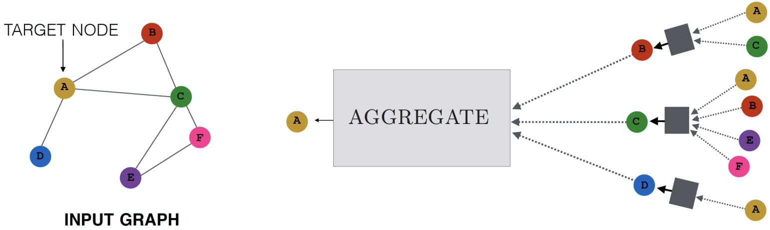
- For any of the tasks on networks mentioned before, the first step is to build an embedding representation of nodes.
- We train an encoder that maps the nodes to a low-dimensional embedding space
- The goal is that the distances in the embedding preserve the distances in the network.



Hamilton 2020

Spatial/Message Passing I

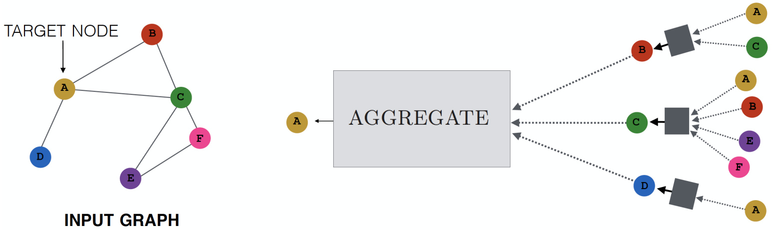
- In the spatial framework, we are going to build the representation of each node as an *AGGREGATE* of its neighbors,
- at each iteration (layer) embedding of a node encodes information from more distant neighbors,
- We can initialize the embeddings the feature-vector of each node



Hamilton 2020

Spatial/Message Passing II

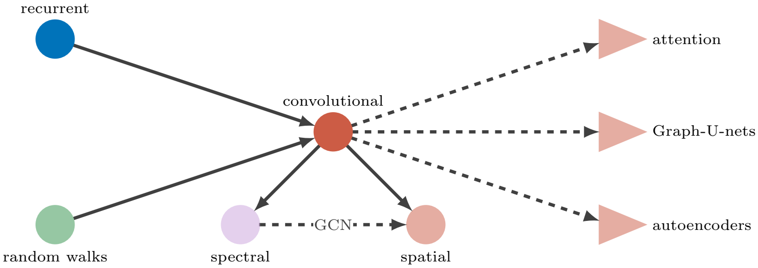
- how do we *AGGREGATE* and *UPDATE* the representation from a previous layer with the neighbors information, defines the different types of models.
- Kipf and Welling 2017 use the average of the previous layer with the neighbors, while *Hamilton2017* use a concatenation as *update* and the average as *aggregate*
- Xu et al. 2019 use the *sum* of the neighbors.



Hamilton 2020

Further reading I

There are other frameworks for deep learning on graphs. For a timeline of different models and future directions:



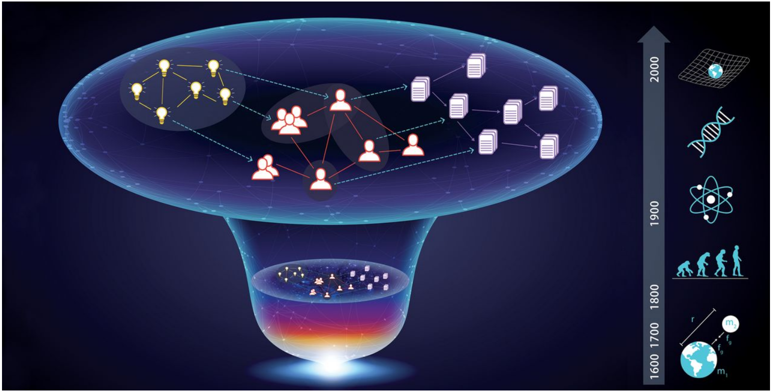
timeline

Further reading II

- **recurrent** (Scarselli et al. 2009)
- **random walks** (Perozzi and Skiena 2014; Grover and Leskovec 2016)
- **spectral methods** (Bruna et al. 2013)
- **GCN** (Kipf and Welling 2017)
- **Spatial** (message passing) (Hamilton, Ying, and Leskovec 2017; Xu et al. 2019)
- **attention** (Veličković et al. 2018; Thekumparampil et al. 2018)
- **autoencoders** (Kipf and Welling 2016)
- **pooling layers** (Gao and Ji 2019)

Science of Science

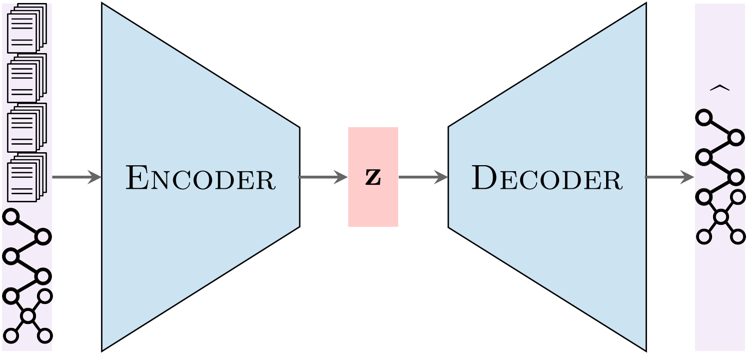
"The Science of Science is based on a transdisciplinary approach that uses large data sets to study the mechanisms underlying the doing of science" (Fortunato et al. 2018)



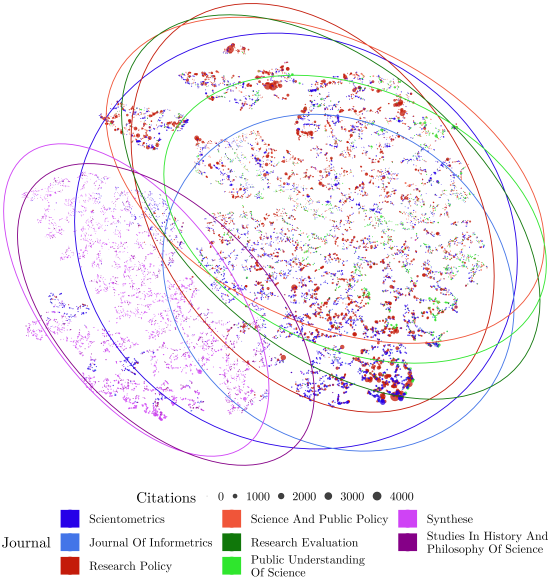
(Fortunato et al. 2018)

How I use it

I use the GNN on research publications, where each article is a node with metadata and text, and they are linked by their citations. I train the emebdding for the *link prediction task*, i.e., I try to predict if to articles are related through a citation.



autoencoder (Kozlowski et al. 2020).



articles embedding (Kozlowski et al. 2020).

Computational Mechanics

- As a future project, we would like to explore the field of computational mechanics, and build useful artifacts for exploring the topics within the field,
- for this, we want to share with you a drive sheet for suggestions on the most relevant journals from the field,
- you are welcome to join us!

Acknowledgement

The Doctoral Training Unit **Data-driven computational modelling and applications** (DRIVEN) is funded by the Luxembourg National Research Fund under the PRIDE programme (PRIDE17/12252781).

https://driven.uni.lu

DTU DRIVEN

Fonds National de la Recherche Luxembourg

Bibliography I

[1] William L. Hamilton. “Graph Representation Learning”. In: *Synthesis Lectures on Artificial Intelligence and Machine Learning* 14.3 (2020), pp. 1–159.

[2] Thomas N. Kipf and Max Welling. “Semi-supervised classification with graph convolutional networks”. In: *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings*. 2017, pp. 1–14. arXiv: 1609.02907.

[3] Keyulu Xu et al. “How powerful are graph neural networks?” In: *7th International Conference on Learning Representations, ICLR 2019* (2019), pp. 1–17. arXiv: arXiv:1810.00826v3.

[4] F Scarselli et al. “The Graph Neural Network Model”. In: *IEEE Transactions on Neural Networks* 20.1 (2009), pp. 61–80.

[5] Bryan Perozzi and Steven Skiena. “DeepWalk: Online Learning of Social Representations”. In: (2014).

Bibliography II

[6] Aditya Grover and Jure Leskovec. “node2vec”. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD ’16*. New York, New York, USA: ACM Press, 2016, pp. 855–864. ISBN: 9781450342322. DOI: 10.1145/2939672.2939754. URL: http://dl.acm.org/citation.cfm?doid=2939672.2939754.

[7] Joan Bruna et al. *Spectral Networks and Locally Connected Networks on Graphs*. 2013. arXiv: 1312.6203 [cs.LG].

[8] William L. Hamilton, Rex Ying, and Jure Leskovec. “Inductive representation learning on large graphs”. In: *Advances in Neural Information Processing Systems* 2017-Decem.Nips (2017), pp. 1025–1035. ISSN: 10495258. arXiv: 1706.02216.

[9] Petar Veličković et al. “Graph attention networks”. In: *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings* (2018), pp. 1–12. arXiv: 1710.10903.

[10] Kiran K. Thekumparampil et al. “Attention-based Graph Neural Network for Semi-supervised Learning”. In: (2018), pp. 1–15. arXiv: 1803.03735. URL: http://arxiv.org/abs/1803.03735.

Bibliography III

[11] Thomas N. Kipf and Max Welling. “Variational Graph Auto-Encoders”. In: 2 (2016), pp. 1–3. arXiv: 1611.07308. URL: http://arxiv.org/abs/1611.07308.

[12] Hongyang Gao and Shuiwang Ji. “Graph U-Nets”. In: *Proceedings of the 36th International Conference on Machine Learning*. Long Beach, California, 2019.

[13] Santo Fortunato et al. “Science of science”. In: *Science* 359.6379 (2018). ISSN: 10959203. DOI: 10.1126/science.aao0185.

[14] Diego Kozlowski et al. “Semantic and Relational Spaces in Science of Science: Deep Learning Models for Article Vectorisation”. In: *arXiv preprint arXiv:2011.02887* (2020).

Other resources

Implementation of models:

https://pytorch-geometric.readthedocs.io/