Diss-l-ECT: Dissecting Graph Data with Local Euler Characteristic Transforms

Bastian Rieck

JMM 2025 · Seattle

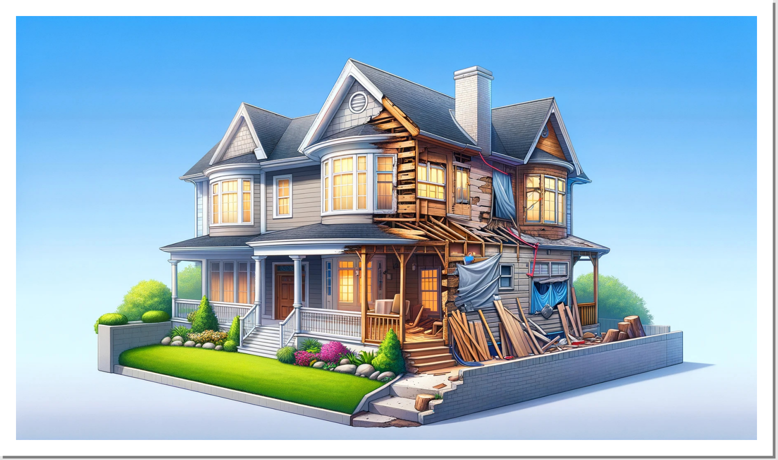
UNI
FR

Machine learning

Two perspectives

Machine learning

Two perspectives



A Shallow Introduction to Deep (Graph) Learning

How to represent graphs?

- High-dimensional vectorial representations are convenient.

How to represent graphs?

- High-dimensional vectorial representations are convenient.
- But two graphs typically have a *different* number of vertices.

How to represent graphs?

- High-dimensional vectorial representations are convenient.
- But two graphs typically have a *different* number of vertices.
- Hence, we require a *vectorised representation* $f\colon \mathcal{G} \to \mathbb{R}^d$ of graphs.

How to represent graphs?

- High-dimensional vectorial representations are convenient.
- But two graphs typically have a *different* number of vertices.
- Hence, we require a *vectorised representation* $f\colon \mathcal{G} \to \mathbb{R}^d$ of graphs.
- Such a representation f needs to be *permutation-invariant*.

Now and then

Shallow approaches

- `node2vec` (encoder–decoder)
- Graph kernels (RKHS feature maps)
- Laplacian-based embeddings

Deep approaches

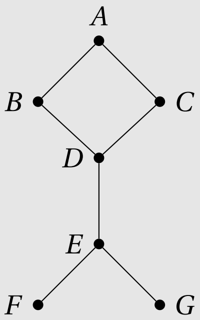
- Graph convolutional networks
- Graph isomorphism networks
- Graph attention networks

Message passing

The foundational paradigm of geometric deep learning

Idea

Neighbouring nodes in a graph may exchange *messages*, i.e. vectors $x, y, z \in \mathbb{R}^d$, which are subsequently *aggregated* to obtain a single representation.

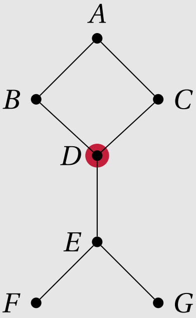


Message passing

The foundational paradigm of geometric deep learning

Idea

Neighbouring nodes in a graph may exchange *messages*, i.e. vectors $x, y, z \in \mathbb{R}^d$, which are subsequently *aggregated* to obtain a single representation.

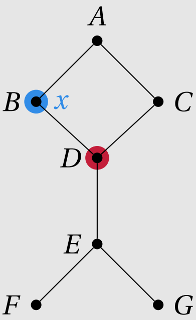


Message passing

The foundational paradigm of geometric deep learning

Idea

Neighbouring nodes in a graph may exchange *messages*, i.e. vectors $x, y, z \in \mathbb{R}^d$, which are subsequently *aggregated* to obtain a single representation.

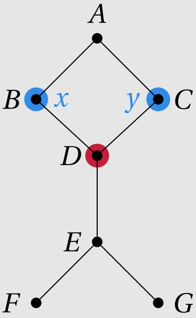


Message passing

The foundational paradigm of geometric deep learning

Idea

Neighbouring nodes in a graph may exchange *messages*, i.e. vectors $x, y, z \in \mathbb{R}^d$, which are subsequently *aggregated* to obtain a single representation.

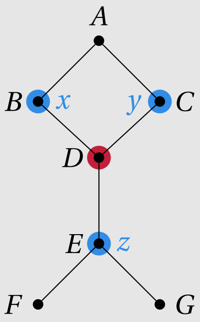


Message passing

The foundational paradigm of geometric deep learning

Idea

Neighbouring nodes in a graph may exchange *messages*, i.e. vectors $x, y, z \in \mathbb{R}^d$, which are subsequently *aggregated* to obtain a single representation.

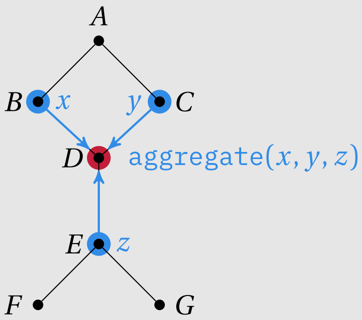


Message passing

The foundational paradigm of geometric deep learning

Idea

Neighbouring nodes in a graph may exchange *messages*, i.e. vectors $x, y, z \in \mathbb{R}^d$, which are subsequently *aggregated* to obtain a single representation.



Message passing

Example

Setup

- $(G, \mathbf{X})$: graph with n vertices and node attributes in $\mathbb{R}^d$
- $h_i^{(t)}$: hidden attributes of vertex i at step t

Message passing

Example

Setup

- $(G, \mathbf{X})$: graph with n vertices and node attributes in $\mathbb{R}^d$
- $h_i^{(t)}$: hidden attributes of vertex i at step t

A simple message-passing neural network

$$h_i^{(t)} = \sigma\left(\mathbf{W_1^{(t)}} h_i^{(t-1)} + \mathbf{W_2^{(t)}} \sum_{j \sim i} h_j^{(t-1)}\right)$$

Message passing

Example

Setup

- $(G, \mathbf{X})$: graph with n vertices and node attributes in $\mathbb{R}^d$
- $h_i^{(t)}$: hidden attributes of vertex i at step t

A simple message-passing neural network

$$h_i^{(t)} = \sigma\left(\mathbf{W_1^{(t)}} h_i^{(t-1)} + \mathbf{W_2^{(t)}} \sum_{j \sim i} h_j^{(t-1)}\right)$$

(Message passing can also be extended to other topological domains like cell complexes, CW complexes, or simplicial complexes.)

Hence: Graphs are ubiquitous!

> *In many ways, graphs are the main modality of data we receive from nature. This is due to the fact that most of the patterns we see, both in natural and artificial systems, are elegantly representable using the language of graph structures. Prominent examples include molecules (represented as graphs of atoms and bonds), social networks and transportation networks.*

P. Veličković, *Everything is connected: Graph neural networks*, *Current Opinion in Structural Biology* 79, 2023, p. 102538

But message passing is not always up to the task!

> *While GNNs have the ability to ignore the graph-structure in such cases, it is not clear that they will. In this work, we show that GNNs actually tend to overfit the graph-structure in the sense that they use it even when a better solution can be obtained by ignoring it.*

M. Bechler-Speicher, I. Amos, R. Gilad-Bachrach and A. Globerson, *Graph Neural Networks Use Graphs When They Shouldn't*, Proceedings of the 41st International Conference on Machine Learning, ed. by R. Salakhutdinov, Z. Kolter, K. Heller, A. Weller, N. Oliver, J. Scarlett and F. Berkenkamp, vol. 235, Proceedings of Machine Learning Research, PMLR, 2024, pp. 3284–3304

This talk

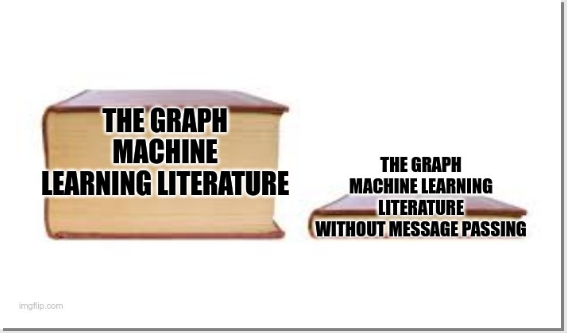
Slogan

Use concepts from *geometry* and *topology* to harness more information from graphs.

This talk

Slogan

Use concepts from *geometry* and *topology* to harness more information from graphs.



Geometry and Topology in Graph Learning

Does adding knowledge about the *graph topology* help?

M. Horn*, E. De Brouwer*, M. Moor, Y. Moreau, B. Rieck† and K. Borgwardt†, *Topological Graph Neural Networks*, International Conference on Learning Representations, 2022, arXiv: `2102.07835 [cs.LG]`

- Providing a new layer for use in graph neural networks.

Does adding knowledge about the *graph topology* help?

M. Horn*, E. De Brouwer*, M. Moor, Y. Moreau, B. Rieck† and K. Borgwardt†, *Topological Graph Neural Networks*, International Conference on Learning Representations, 2022, arXiv: `2102.07835 [cs.LG]`

- Providing a new layer for use in graph neural networks.
- Create new synthetic data sets with pronounced topological features.

Does adding knowledge about the *graph topology* help?

M. Horn[*], E. De Brouwer[*], M. Moor, Y. Moreau, B. Rieck[†] and K. Borgwardt[†], *Topological Graph Neural Networks*, International Conference on Learning Representations, 2022, arXiv: `2102.07835 [cs.LG]`

- Providing a new layer for use in graph neural networks.
- Create new synthetic data sets with pronounced topological features.
- Getting high predictive performance on these data sets. 🥳

Does adding knowledge about the *graph topology* help?

M. Horn*, E. De Brouwer*, M. Moor, Y. Moreau, B. Rieck† and K. Borgwardt†, *Topological Graph Neural Networks*, International Conference on Learning Representations, 2022, arXiv: `2102.07835 [cs.LG]`

- Providing a new layer for use in graph neural networks.
- Create new synthetic data sets with pronounced topological features.
- Getting high predictive performance on these data sets. 🥳
- Getting mediocre performance on other benchmark data sets. 😢

Does adding knowledge about the *graph topology* help?

M. Horn[*], E. De Brouwer[*], M. Moor, Y. Moreau, B. Rieck[†] and K. Borgwardt[†], *Topological Graph Neural Networks*, International Conference on Learning Representations, 2022, arXiv: `2102.07835 [cs.LG]`

- Providing a new layer for use in graph neural networks.
- Create new synthetic data sets with pronounced topological features.
- Getting high predictive performance on these data sets. 🥳
- Getting mediocre performance on other benchmark data sets. 😢
- Getting better performance if node features are randomised. 😱

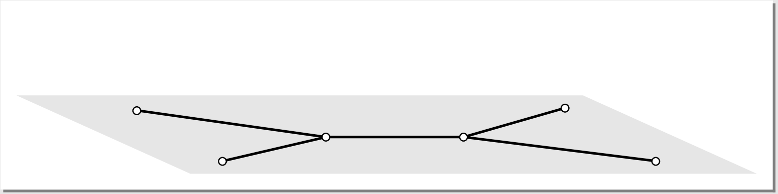
Does adding knowledge about the *graph topology* help?

M. Horn*, E. De Brouwer*, M. Moor, Y. Moreau, B. Rieck† and K. Borgwardt†, *Topological Graph Neural Networks*, International Conference on Learning Representations, 2022, arXiv: `2102.07835 [cs.LG]`

- Providing a new layer for use in graph neural networks.
- Create new synthetic data sets with pronounced topological features.
- Getting high predictive performance on these data sets. 🥳
- Getting mediocre performance on other benchmark data sets. 😢
- Getting better performance if node features are randomised. 😱

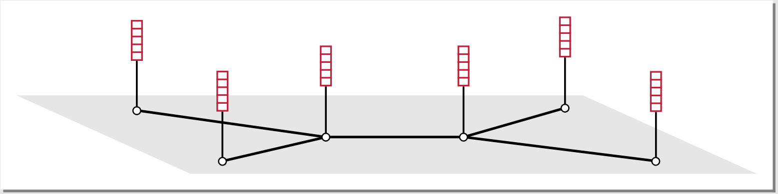
On data sets with pronounced topological structures, we found that our method helps GNNs obtain substantial gains in predictive performance.

What does typical graph data look like?



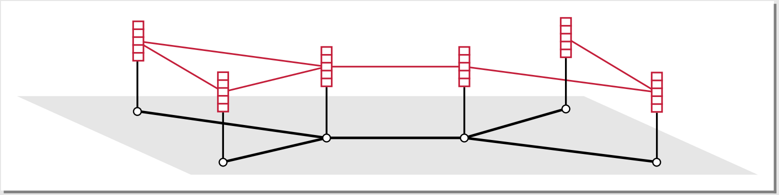
Only topology

What does typical graph data look like?



Node features (signals)

What does typical graph data look like?



Topology plus feature-induced geometry

Going back to the data

C. Coupette, J. Vreeken and B. Rieck, *All the World's a (Hyper)Graph: A Data Drama*, *Digital Scholarship in the Humanities* 39.1, 2024, pp. 74–96, arXiv: `2206.08225 [cs.LG]`, URL: `https://hyperbard.net`

We introduce a novel dataset,
With full documentation as Appendix.
Raw data stem from all of Shakespeare's plays,
We model them as graphs in many ways,
And demonstrate representations matter.
The data readily accessible,
All code is publicly available.
What follows, to avoid redundancy,
Conveys our main ideas, as you will see
A tragedy in the Community.

(Probably the only machine-learning paper imitating a dramatic Shakespearean style that was *not* co-authored using large language models.)

Going back to the data, redux

R. Ballester*, E. Röell*, D. B. Schmid*, M. Alain*, S. Escalera, C. Casacuberta and B. Rieck, *MANTRA: The Manifold Triangulations Assemblage*, Preprint, 2024, arXiv: `2410.02392 [cs.LG]`

> *We present MANTRA, the* ***man****ifold* ***tr****iangulations* ***a****ssemblage, […] comprising triangulations of combinatorial 2-manifolds and 3-manifolds. A noteworthy aspect of MANTRA is the* ***conspicuous absence of any intrinsic vertex or edge features such as coordinates or signals****. We argue that this absence renders tasks more topological, as models can only rely on topology, instead of non-topological information contained in features.*

Going back to the data, conclusion

- Message passing cannot easily leverage topological information.

Going back to the data, conclusion

- Message passing cannot easily leverage topological information.
- Topological methods cannot easily make use of node signals.

Going back to the data, conclusion

- Message passing cannot easily leverage topological information.
- Topological methods cannot easily make use of node signals.
- We need something that *combines* a geometrical and a topological perspective.

The Euler Characteristic Transform

The Euler Characteristic Transform

Ingredients

The Euler Characteristic of a simplicial complex K

Let K_i denote the i-skeleton of K. Then the *Euler Characteristic* of K is defined as:

$$\chi(K) := \sum_{i=0}^{p} (-1)^i |K_i|$$

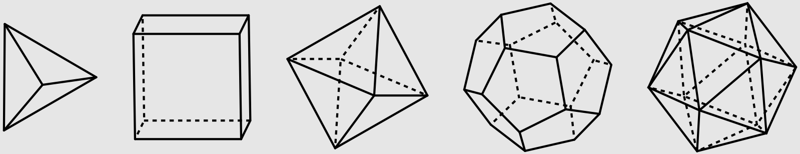
The Euler Characteristic Transform

Ingredients

The Euler Characteristic of a simplicial complex K

Let K_i denote the i-skeleton of K. Then the *Euler Characteristic* of K is defined as:

$$\chi(K) := \sum_{i=0}^{p} (-1)^i |K_i|$$



Platonic solids (with $\chi = 2$)

The Euler Characteristic Transform

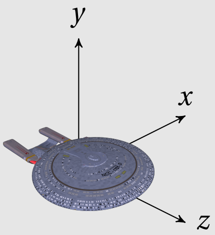
Ingredients, continued

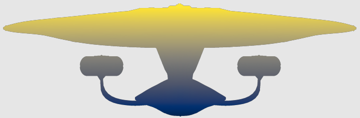
A filtration of K

Given a simplicial complex K with *vertex features* $x_v \in \mathbb{R}^d$, and a *direction* $w \in S^{d-1}$, we obtain a real-valued function on the simplices of K via:

$$f(\sigma) = \max_{v \in \sigma} \langle x_v, w \rangle$$

Given $t \in \mathbb{R}$, this lets us filtrate K via $K_t := \{\sigma \in K \mid f(\sigma) \leq t\}$.





Filtration in y-direction (colours correspond to values of f)

The Euler Characteristic Transform

Putting it all together

Definition

The Euler Characteristic Transform maps a direction to a *function* from $\mathbb{R}$ to $\mathbb{Z}$:

$$\begin{aligned}\mathrm{ECT}\colon S^{d-1} &\to \mathbb{Z}^{\mathbb{R}}\\ w &\mapsto ((w, t) \mapsto \chi(K_t))\end{aligned}$$

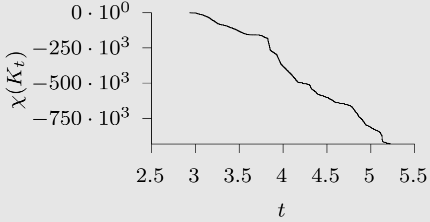
The Euler Characteristic Transform

Putting it all together

Definition

The Euler Characteristic Transform maps a direction to a *function* from $\mathbb{R}$ to $\mathbb{Z}$:

$$\begin{aligned} \mathrm{ECT}\colon S^{d-1} &\to \mathbb{Z}^{\mathbb{R}} \\ w &\mapsto ((w,t) \mapsto \chi(K_t)) \end{aligned}$$



The ECT restricted to the filtration in y-direction

The Euler Characteristic Transform

Representation and properties

- The ECT is a sufficient shape statistic and *injective*.

The Euler Characteristic Transform

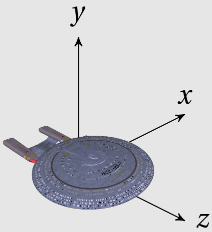
Representation and properties

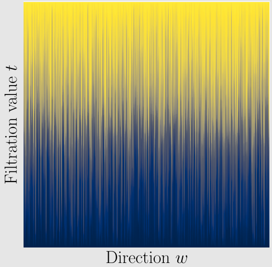
- The ECT is a sufficient shape statistic and *injective*.
- Discretised versions can be represented as *images*, with rows indexing the directions, and columns indexing the filtration values.

The Euler Characteristic Transform

Representation and properties

- The ECT is a sufficient shape statistic and *injective*.
- Discretised versions can be represented as *images*, with rows indexing the directions, and columns indexing the filtration values.

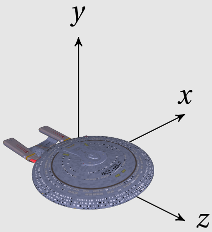


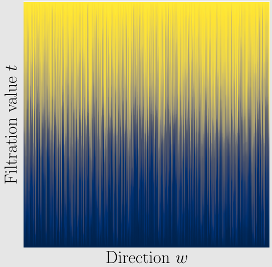


The Euler Characteristic Transform

Representation and properties

- The ECT is a sufficient shape statistic and *injective*.
- Discretised versions can be represented as *images*, with rows indexing the directions, and columns indexing the filtration values.





K. Turner, S. Mukherjee and D. M. Boyer, *Persistent Homology Transform for Modeling Shapes and Surfaces*, *Information and Inference* 3.4, 2014, pp. 310–344

The Euler Characteristic Transform and machine learning

Dramatis personæ

Julius (von Rohrscheidt)

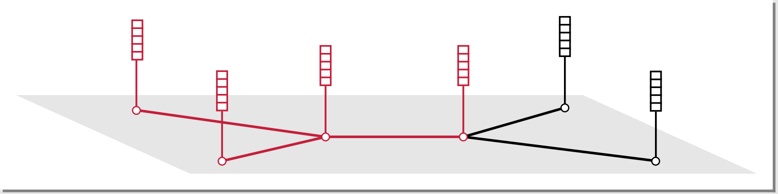
Ernst (Röell)

- Using the ECT in a differentiable setting and *learning* relevant directions for classification: E. Röell and B. Rieck, *Differentiable Euler Characteristic Transforms for Shape Classification*, International Conference on Learning Representations, 2024, arXiv: `2310.07630 [cs.LG]`
- Using *local* ECTs for node classification: J. von Rohrscheidt and B. Rieck, *Diss-l-ECT: Dissecting Graph Data with local Euler Characteristic Transforms*, Preprint, 2024, arXiv: `2410.02622 [cs.LG]`

Local Euler Characteristic Transforms

Main idea

Use the ECT of a k-hop neighbourhoods of an attributed graph as a descriptor of that node's role.



Local Euler Characteristic Transforms

Model	Cornell	Wisconsin	Texas	Roman-Empire	Amazon
GCN	45.0 ± 2.2	44.2 ± 2.6	47.3 ± 1.5	73.3 ± 0.8	42.3 ± 0.7
GAT	44.7 ± 2.9	48.2 ± 2.0	51.7 ± 3.2	76.4 ± 1.2	44.6 ± 0.9
GIN	46.5 ± 3.1	49.7 ± 2.5	54.2 ± 2.9	56.8 ± 1.0	44.1 ± 0.8
H2GCN	66.2 ± 3.5	70.2 ± 2.3	72.3 ± 3.0	64.2 ± 0.9	40.1 ± 0.7
ℓ-ECT	67.1 ± 4.1	78.5 ± 2.6	74.8 ± 3.1	81.1 ± 0.4	49.8 ± 0.3

Conclusion

- The (local) Euler Characteristic Transform provides a new geometrical-topological perspective to machine-learning research.

Conclusion

- The (local) Euler Characteristic Transform provides a new geometrical-topological perspective to machine-learning research.
- It is particularly interesting due to its *versatility* and *stability* properties, often outperforming more complex methods.

Conclusion

- The (local) Euler Characteristic Transform provides a new geometrical-topological perspective to machine-learning research.
- It is particularly interesting due to its *versatility* and *stability* properties, often outperforming more complex methods.
- Numerous open questions remain:

Conclusion

- The (local) Euler Characteristic Transform provides a new geometrical-topological perspective to machine-learning research.
- It is particularly interesting due to its *versatility* and *stability* properties, often outperforming more complex methods.
- Numerous open questions remain:
  - How to efficiently represent ECTs?

Conclusion

- The (local) Euler Characteristic Transform provides a new geometrical-topological perspective to machine-learning research.
- It is particularly interesting due to its *versatility* and *stability* properties, often outperforming more complex methods.
- Numerous open questions remain:
  - How to efficiently represent ECTs?
  - How to compare ECTs while being *invariant* to certain transformations?

Conclusion

- The (local) Euler Characteristic Transform provides a new geometrical-topological perspective to machine-learning research.
- It is particularly interesting due to its *versatility* and *stability* properties, often outperforming more complex methods.
- Numerous open questions remain:
  - How to efficiently represent ECTs?
  - How to compare ECTs while being *invariant* to certain transformations?
  - How many directions are required *in practice*?

Conclusion

- The (local) Euler Characteristic Transform provides a new geometrical-topological perspective to machine-learning research.
- It is particularly interesting due to its *versatility* and *stability* properties, often outperforming more complex methods.
- Numerous open questions remain:
  - How to efficiently represent ECTs?
  - How to compare ECTs while being *invariant* to certain transformations?
  - How many directions are required *in practice*?

Our research