ETH zürich

MACIEJ BESTA, TORSTEN HOEFLER, ET AL. Motif Prediction with Graph Neural Networks





























9





9











Case 2: Find Missing Data







Case 2: Find Missing Data



10













The seal









C. T. C.











Station and





$$s_{u,v}^{CN} = |\Gamma(u) \cap \Gamma(v)|$$

$$s_{u,v}^{Jaccard} = \frac{|\Gamma(u) \cap \Gamma(v)|}{|\Gamma(u) \cup \Gamma(v)|}$$



and and





$$s_{u,v}^{CN} = |\Gamma(u) \cap \Gamma(v)|$$

$$s_{u,v}^{Jaccard} = \frac{|\Gamma(u) \cap \Gamma(v)|}{|\Gamma(u) \cup \Gamma(v)|}$$



State -



all the section of the

Warmup: Link Prediction



$$s_{u,v}^{CN} = |\Gamma(u) \cap \Gamma(v)|$$

$$s_{u,v}^{Jaccard} = \frac{|\Gamma(u) \cap \Gamma(v)|}{|\Gamma(u) \cup \Gamma(v)|}$$





$$s_{u,v}^{CN} = |\Gamma(u) \cap \Gamma(v)|$$

$$s_{u,v}^{Jaccard} = \frac{|\Gamma(u) \cap \Gamma(v)|}{|\Gamma(u) \cup \Gamma(v)|}$$

$$s_{u,v}^{Salton} = \frac{|\Gamma(u) \cap \Gamma(v)|}{\sqrt{d_u d_v}}$$

State States





States and the second



How to

assess?

One obtains a "score" s(e) = s(u,v) for each (missing) link in a graph

State - ---





How to

assess?

One obtains a "score" s(e) = s(u,v) for each (missing) link in a graph

The second

The higher the score, the more probable a given link e is to appear in the graph







Motifs: small, recurring subgraphs (e.g. modelling molecules, gene interactions, groups of people)



Motifs: small, recurring subgraphs (e.g. modelling molecules, gene interactions, groups of people)









Maller and the second



Is this important?













Is this important?

A lot of work into motifs exists, <u>recent</u> renewed interest under the theme "Higher-order network organization"

=

Seeing a graph through the perspective of motifs instead of edges ("higher order")

network motifs	Q
Scholar About 899'000 results (0.08 sec) YEAR -	Ŧ

(0)

🌙 19°C \land 🦔 📴 🖻 👛 🌈 Φ) POL

Network motifs: simple building blocks of complex networks [PDF] sciencemag.org

Complex networks are studied across many fields of science. To uncover their structural design principles, we defined "**network motifs**," patterns of interconnections occurring in complex networks at numbers that are significantly higher than those in randomized …

☆ ワワ Cited by 7268 Related articles All 78 versions

[HTML] Network motifs: theory and experimental approaches [HTML] nature.com

Transcription regulation networks control the expression of genes. The transcription networks of wellstudied microorganisms appear to be made up of a small set of recurring regulation patterns, called **network motifs**. The same **network motifs** have recently been ...

 $\cancel{2}$ 99 Cited by 3112 Related articles All 45 versions





Any time

Is this important? 🚵 W 🔟 W 🖳 F 💆 S 💆 🛩 🖉 SP 🔇 🛆 S 🛥 M 🔇 S G G S S ■ scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=higher+order+organization Google Scholar higher order organization Articles About 4'660'000 results (0.11 sec) Higher-order organization of complex networks Since 2021 AR Benson, DF Gleich, J Leskovec - Science, 2016 - science.sciencemag.org \equiv Networks are a fundamental tool for understanding and modeling complex systems in Since 2020 physics, biology, neuroscience, engineering, and social science. Many networks are know Since 2017 to exhibit rich, lower-order connectivity patterns that can be captured at the level of indivi Custom range... $\cancel{2}$ $\cancel{9}$ Cited by 563 Related articles All 26 versions Sort by relevance Higher-order organization by mesoscale self-assembly and transfe Sort by date hybrid nanostructures H Cölfen, S Mann - Angewandte Chemie International Edition, 2003 - Wiley Online Libra The organization of nanostructures across extended length scales is a key challenge in include patents design of integrated materials with advanced functions. Current approaches tend to be include citations based on physical methods, such as patterning, rather than the spontaneous chemical $\cancel{2}$ $\cancel{9}$ Cited by 1821 Related articles All 11 versions Create alert Higher-order genome organization in human disease Î T Misteli - Cold Spring Harbor perspectives in biology, 2010 - cshperspectives.cshlp.org Genomes are organized into complex higher-order structures by folding of the DNA into chromatin fibers, chromosome domains, and ultimately chromosomes. The higher-order organization of genomes is functionally important for gene regulation and control of gen ☆ ワワ Cited by 190 Related articles All 7 versions [HTML] In vivo mapping of eukaryotic RNA interactomes reveals princi higher-order organization and regulation JGA Aw, Y Shen, A Wilm, M Sun, XN Lim, KL Boon... - Molecular cell, 2016 - Elsevier Identifying pairwise RNA-RNA interactions is key to understanding how RNAs fold and interact with other RNAs inside the cell. We present a high-throughput approach, sequencing of psoralen crosslinked, ligated, and selected hybrids (SPLASH), that maps . $\cancel{2}$ $\cancel{99}$ Cited by 172 Related articles All 13 versions

0

A lot of work into motifs exists, recent interest under the theme "Higher-order network organization"

Seeing a graph through the perspective of motifs instead of edges ("higher order")

network motifs			
Scholar	About 899'000 results (0.08 sec)		

YEAR

Network motifs: simple building blocks of complex networks [PDF] sciencemag.org R Milo, S Shen-Orr, S Itzkovitz, N Kashtan... - ..., 2002 - science.sciencemag.org

Complex networks are studied across many fields of science. To uncover their structural design principles, we defined "network motifs," patterns of interconnections occurring in complex networks at numbers that are significantly higher than those in randomized ...

99 Cited by 7268 Related articles All 78 versions

[HTML] Network motifs: theory and experimental approaches

[HTML] nature.com

🌙 19°C \land 🦔 🖗 🖬 📥 🌈 Φ) POL

(0)

P 🖻

1

12:22 AM

2/18/2021

U Alon - Nature Reviews Genetics, 2007 - nature.com

Transcription regulation networks control the expression of genes. The transcription networks of wellstudied microorganisms appear to be made up of a small set of recurring regulation patterns, called network motifs. The same network motifs have recently been ...

へ Ĝ 🔄 🍙 🌈 🕼 POL

99 Cited by 3112 Related articles All 45 versions

0)

8:58 PM 6/15/2021



Motifs: small, recurring subgraphs (e.g. modelling molecules, gene interactions, groups of people)









Motifs: small, recurring subgraphs (e.g. modelling molecules, gene interactions, groups of people)



Link prediction is well understood for, well, links (scores assigned to links, similarity based methods, etc.)





Motifs: small, recurring subgraphs (e.g. modelling molecules, gene interactions, groups of people)

Link prediction is well understood for, well, links (scores assigned to links, similarity based methods, etc.)

How to generalize to motifs?



Motifs: small, recurring subgraphs (e.g. modelling molecules, gene interactions, groups of people)

Link prediction is well understood for, well, links (scores assigned to links, similarity based methods, etc.)

How to generalize to motifs?

General vision: assign some score to motifs (make them comparable)

Motifs with higher scores are more probable



Idea: Generalize Link Pr

Motifs: small, recurring subgridinteractions, groups of people differences to link prediction!

Link prediction is well understood for, well, links (scores assigned to links, similarity based methods, etc.)

How to generalize to motifs?

General vision: assign some score to motifs (make them comparable)

Motifs with higher scores are more probable



Idea: Generalize Link Pr ...but there are so many Motifs: small, recurring subgr differences to link prediction! interactions, groups of people Let's go over them [1] ... Link prediction is well understood How to generalize to for, well, links (scores assigned to motifs? links, similarity based methods, etc.)

General vision: assign some score to motifs (make them comparable)

Motifs with higher scores are more probable

[1] M. Besta et al.: "Motif prediction with graph neural networks", KDD'22










Motif prediction:

Difference 1: There May Be Many Potential New Motifs for a Fixed Vertex Set



















15



Link prediction:

Motif prediction:





Link prediction:



A link to be predicted does not exist

Motif prediction:



Link prediction:



Motif prediction:

•--•

A link to be predicted does not exist



Link prediction:



Motif prediction:



A link to be predicted does not exist



Link prediction:



Motif prediction:



A link to be predicted does not exist

How to consider such edges in the score functions?



Link prediction:



Motif prediction:



A link to be predicted does not exist

How to consider such edges in the score functions? Example: some existing relationships in a group of people



Link prediction:

Motif prediction:





Link prediction:



Motif prediction:





Link prediction:

Motif prediction:

We don't want these links! (i.e., these links appearing would make it impossible for a motif in question to appear



Link prediction:



Motif prediction:

We don't want these links! (i.e., these links appearing would make it impossible for a motif in question to appear



Example: chemical bonds



Link prediction:

We don't want these links! (i.e., these links appearing would make it impossible for a motif in question to appear

Motif prediction:

Example: chemical bonds

How to consider such edges in the score functions?



Link prediction: Motif prediction: We don't want these links! (i.e., these links appearing would make it impossible for a motif in question to appear No such effect (a link to be predicted is never a Example: How to consider ", deal breaker") chemical bonds such edges in the score functions?



Link prediction:

Motif prediction:





Link prediction:

Motif prediction:

We want **this**:



Link prediction:

Motif prediction:

We want **this**:

















Starting Simple: Motif Scores Based on Independent Links



A motif:
$$M = (V_M, E_M)$$

Starting Simple: Motif Scores Based on Independent Links



A motif: $M = (V_M, E_M)$

The second second second

Starting Simple: Motif Scores Based on Independent Links

 $E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$



Starting Simple: Motif Scores Based on Independent Links

A motif: $M = (V_M, E_M)$ $E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$

Motif edges that do **not** yet exist

The second second



Starting Simple: Motif Scores Based on Independent Links

motif:
$$M = (V_M, E_M)$$

 $E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$

Α

Motif edges thatMotif edges thatdo not yet existalready exist



Starting Simple: Motif Scores Based on Independent Links

A motif:
$$M = (V_M, E_M)$$

 $E_M = E_M, \mathcal{N} \cup E_M, \mathcal{E}$
Motif edges that
do not yet exist
Motif edges that
already exist



spcl.inf.ethz.ch

Motif Prediction Score Functions

Starting Simple: Motif Scores Based on Independent Links

A motif:
$$M = (V_M, E_M)$$

$$E_M = E_{M,N} \cup E_{M,\mathcal{E}}$$

$$Motif edges that do not yet exist$$

$$Motif edges that already exist$$

$$\int S(e)$$

$$s_{\perp}(M) = \prod_{e \in E_{M,\mathcal{N}}} s(e)$$

The second second second



A motif: $M = (V_M, E_M)$ **Motif Prediction Score Functions** Starting Simple: Motif Scores Based on Independent Links $E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$ Motif edges that Motif edges that already exist do **not** yet exist $s_{\perp}(M) = \prod_{e \in E_M \ N} \stackrel{\checkmark}{} s(e)$



Motif Prediction Score FunctionsA motif:
$$M = (V_M, E_M)$$
Starting Simple: Motif Scores Based on Independent Links $E_M = E_M, \mathcal{N} \cup E_M, \mathcal{E}_M$ Motif edges that do not yet existMotif edges that do not yet exist $s_{\perp}(M) = \prod_{e \in E_M, \mathcal{N}} s(e)$ s(e) is any link prediction score which outputs into [0, 1] (e.g., Jaccard)

A LOT A LOCATION STORE








spcl.inf.ethz.ch









Example heuristic for the Jaccard score s(e) for links:



What if the arrival of some (motif) links impacts the chances for other motif links to appear?



Example heuristic for the Jaccard score s(e) for links:



What if the arrival of some (motif) links impacts the chances for other motif links to appear?

Adding correlation between links



Example heuristic for the Jaccard score s(e) for links:



What if the arrival of some (motif) links impacts the chances for other motif links to appear?

Adding correlation between links



12 martin the the

Example heuristic for the Jaccard score s(e) for links:



What if the arrival of some (motif) links impacts the chances for other motif links to appear?

Adding correlation between links

$$s(M) = f(\mathbf{s}(\mathbf{e})) = \langle \mathbf{w}, \mathbf{s}(\mathbf{e}) \rangle$$

s(M) is a convex combination of the vector of link prediction scores s(e)

Example heuristic for the Jaccard score s(e) for links:

A motif:
$$M = (V_M, E_M)$$

 $E_M = E_M, \mathcal{N} \cup E_M, \mathcal{E}$
Motif edges that
do **not** yet exist
Motif edges that
already exist





What if the arrival of some (motif) links impacts the chances for other motif links to appear?

The weight vector (incorporates do no user's domain knowledge)

$$s(M) = f(\mathbf{s}(\mathbf{e})) = \langle \mathbf{w}, \mathbf{s}(\mathbf{e}) \rangle$$

Motif edges thatMotif edges thatdo not yet existalready exist

 $E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$



s(M) is a convex combination of the vector of link prediction scores s(e)

Adding correlation between links

Example heuristic for the Jaccard score s(e) for links:



A motif: $M = (V_M, E_M)$

 $E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$

What if the arrival of some (motif) links impacts the chances for other motif links to appear?

Motif edges that do **not** yet exist

Motif edges that already exist



The weight vector (incorporates user's domain knowledge)

$$s(M) = f(\mathbf{s}(\mathbf{e})) = \langle \mathbf{w}, \mathbf{s}(\mathbf{e}) \rangle$$



s(M) is a convex combination of the vector of link prediction scores s(e)

Example heuristic for the Jaccard score s(e) for links:

$$s(M)^J = \frac{1}{|E_M|} \left(\sum_{e_{u,v} \in E_{M,\mathcal{N}}} \frac{|N_u \cap N_v|}{|N_u \cup N_v|} + |E_{M,\mathcal{E}}| \right)$$



What if the arrival of some (motif) links reduces, or even prevents, motif's appearance?

A motif: $M = (V_M, E_M)$

12 manutation to

 $E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$





What if the arrival of some (motif) links reduces, or even prevents, motif's appearance?

Incorporating deal-breaker links

A motif: $M = (V_M, E_M)$

the section of the section of the

 $E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$





What if the arrival of some (motif) links reduces, or even prevents, motif's appearance?

Incorporating deal-breaker links

A motif: $M = (V_M, E_M)$

$$E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$$
$$E_M^* = E_M$$





What if the arrival of some (motif) links reduces, or even prevents, motif's appearance?

Incorporating deal-breaker links

A motif: $M = (V_M, E_M)$ $E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$ $E_M^* = E_M \cup \overline{E}_{M,\mathcal{D}}$

The second second second





What if the arrival of some (motif) links reduces, or even prevents, motif's appearance?

Incorporating deal-breaker links



$$E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$$

$$E_M^* = E_M \cup \overline{E}_{M,\mathcal{D}}$$

edges





What if the arrival of some (motif) links reduces, or even prevents, motif's appearance?

Incorporating deal-breaker links



 $E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$

 $E_M^* = E_M \cup \overline{E}_{M,\mathcal{D}}$ Motif Deal-breaker

edges

al-breaker edges





What if the arrival of some (motif) links reduces, or even prevents, motif's appearance?

Incorporating deal-breaker links





What if the arrival of some (motif) links reduces, or even prevents, motif's appearance?

Incorporating deal-breaker links

$$s(M) = f(\mathbf{s}(\mathbf{e})) = \langle \mathbf{w}, \mathbf{s}(\mathbf{e}) \rangle$$

A motif:
$$M = (V_M, E_M)$$

 $E_M = E_M, \mathcal{N} \cup E_M, \mathcal{E}$
 $E_M^* = E_M \cup \overline{E}_M, \mathcal{D}$
Motif edges edges



What if the arrival of some (motif) links reduces, or even prevents, motif's appearance?

Incorporating deal-breaker links

A motif: $M = (V_M, E_M)$ $E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$ $E_M^* = E_M \cup \overline{E}_{M,\mathcal{D}}$ **Deal-breaker** Motif edges edges $s^*(M) = f(\mathbf{s}^*(\mathbf{e})) = \max(0, \langle \mathbf{w}, \mathbf{s}^*(\mathbf{e}) \rangle)$



Incorporating deal-breaker links

A motif: $M = (V_M, E_M)$

 $E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$

 $E_M^* = E_M \cup \overline{E}_{M,\mathcal{D}}$

Motif

edges

What if the arrival of some (motif) links reduces, or even prevents, motif's appearance?

The weight vector (incorporates user's domain knowledge)

$$s^*(M) = f(\mathbf{s}^*(\mathbf{e})) = \max(0, \langle \mathbf{w}, \mathbf{s}^*(\mathbf{e}) \rangle)$$





A motif: $M = (V_M, E_M)$

What if the arrival of some (motif) links reduces, or even prevents, motif's appearance?

Incorporating deal-breaker links

The weight vector (incorporates user's domain knowledge)

$$s^*(M) = f(\mathbf{s}^*(\mathbf{e})) = \max(0, \langle \mathbf{w}, \mathbf{s}^*(\mathbf{e}) \rangle)$$

 $E_M^* = E_M \cup \overline{E}_{M,\mathcal{D}}$ **Deal-breaker** edges

Motif

edges

 $E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$



A deal-breaker edge that **does not** exist (i.e., it might arrive later):

$$s_i^*(e) = -s_i(e)$$



A motif: $M = (V_M, E_M)$ **Motif Prediction Score Functions** What if the arrival of some (motif) links $E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$ reduces, or even prevents, motif's appearance? $E_M^* = E_M \cup \overline{E}_{M,\mathcal{D}}$ Incorporating deal-breaker links The weight vector (incorporates **Deal-breaker** Motif user's domain knowledge) edges edges $s^*(M) = f(\mathbf{s}^*(\mathbf{e})) = \max(0, \langle \mathbf{w}, \mathbf{s}^*(\mathbf{e}) \rangle)$ A deal-breaker edge that **does not** A deal-breaker edge that **does already** exist: exist (i.e., it might arrive later): $s_{i}^{*}(e) = -s_{i}(e)$ $s^{*}(e) = 0$



A motif: $M = (V_M, E_M)$ **Motif Prediction Score Functions** What if the arrival of some (motif) links $E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$ reduces, or even prevents, motif's appearance? $E_M^* = E_M \cup \overline{E}_{M,\mathcal{D}}$ Incorporating deal-breaker links The weight vector (incorporates Deal-breaker Motif user's domain knowledge) edges edges $s^*(M) = f(\mathbf{s}^*(\mathbf{e})) = \max(0, \langle \mathbf{w}, \mathbf{s}^*(\mathbf{e}) \rangle)$ A deal-breaker edge that **does not** A deal-breaker edge that **does already** exist: exist (i.e., it might arrive later): $s_i^*(e) = -s_i(e)$ $s^{*}(e) = 0$

Another formulation (perhaps more intuitive):



Motif Prediction Score FunctionsA motif:
$$M = (V_M, E_M)$$
What if the arrival of some (motif) links
reduces, or even prevents, motif's appearance? $E_M = E_{M,N} \cup E_{M,\mathcal{E}}$ Incorporating deal-breaker linksThe weight vector (incorporates
user's domain knowledge) $E_M = E_M \cup \overline{E}_{M,\mathcal{D}}$ $s^*(M) = f(s^*(e)) = max(0, \langle \mathbf{w}, s^*(e) \rangle)$ A deal-breaker edge that does not
exist (i.e., it might arrive later):
 $s^*_i(e) = -s_i(e)$ A deal-breaker edge
that does already exist:
 $s^*(e) = 0$

Another formulation (perhaps more intuitive):

$$s_{\perp}^{*}(M) = \prod_{e \in E_{M}} s(e) \cdot \prod_{e \in \overline{E}_{M,\mathcal{D}}} (1 - s(e))$$

and and and and



Check the paper [1] for details about heuristics (based on parwise Jaccard, Common Neighbors, and Adamic-Adar scores)

 $E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$ $E_M^* = E_M \cup E_{M.\mathcal{D}}$

<u>A m</u>otif: $M = (V_M, E_M)$

Deal-breaker Motif

edges

edges

$$\mathbf{s}^*(M) = f(\mathbf{s}^*(\mathbf{e})) = \max(\mathbf{0}, \langle \mathbf{w}, \mathbf{s}^*(\mathbf{e}) \rangle)$$

user s domain kn

A deal-breaker edge that **does not** exist (i.e., it might arrive later):

$$s_i^*(e) = -s_i(e)$$

A deal-breaker edge that **does already** exist:

$$s^*(e) = 0$$

$$s_{\perp}^{*}(M) = \prod_{e \in E_{M}} s(e) \cdot \prod_{e \in \overline{E}_{M,\mathcal{D}}} (1 - s(e))$$



<u>A m</u>otif: $M = (V_M, E_M)$ **Motif Prediction Score Functions** Check the paper [1] for details about heuristics $E_M = E_{M,\mathcal{N}} \cup E_{M,\mathcal{E}}$ (based on parwise Jaccard, Common $E_M^* = E_M \cup E_{M,\mathcal{D}}$ Neighbors, and Adamic-Adar scores) **Deal-breaker** edges edges $s^*(M) = f(\mathbf{s}^*(\mathbf{e})) = \max(0, \langle \mathbf{w}, \mathbf{s}^*(\mathbf{e}) \rangle)$ A deal-breaker edge that **does not** A deal-breaker edge exist (i.e., it migh These are all heuristics... but recent results for $s_{i}^{*}(e) =$ learning-enhanced link prediction [2] post a question: can we use learning for motif prediction as well?

$$s_{\perp}^{*}(M) = \prod_{e \in E_{M}} s(e) \cdot \prod_{e \in \overline{E}_{M,\mathcal{D}}} (1 - s(e))$$

[1] M. Besta et al.: "Motif prediction with graph neural networks", KDD'22 [2] M. Zhang et al.: "Link Prediction Based on Graph Neural Networks", NeurIPS'18





The graph structure may be arbitrary, maybe one could arrive at better heuristics by **learning**?





The graph structure may be arbitrary, maybe one could arrive at better heuristics by **learning**?







The graph structure may be arbitrary, maybe one could arrive at better heuristics by **learning**?







The graph structure may be arbitrary, maybe one could arrive at better heuristics by **learning**?







The graph structure may be arbitrary, maybe one could arrive at better heuristics by **learning**?







The graph structure may be arbitrary, maybe one could arrive at better heuristics by **learning**?







The graph structure may be arbitrary, maybe one could arrive at better heuristics by **learning**?





The graph structure may be arbitrary, maybe one could arrive at better heuristics by **learning**?





The graph structure may be arbitrary, maybe one could arrive at better heuristics by **learning**?





The graph structure may be arbitrary, maybe one could arrive at better heuristics by **learning**?




The animation borrowed from T. Hoefler

The graph structure may be arbitrary, maybe one could arrive at better heuristics by **learning**?

How does deep learning work?





A motif: $M = (V_M, E_M)$

The second of the



A motif: $M = (V_M, E_M)$

The second and the





A motif:
$$M = (V_M, E_M)$$













































layer-wise weight update







The sector of the

Motif Prediction: Deep Learning Formulation



Will a k-clique appear or not?

[1] M. Zhan et al. 2018. An end-to-end deep learning architecture for graph classification.AAAI Conference on Artificial Intelligence

Will a k-clique

appear or not?

The samples are used to train a neural network model called Deep Graph Convolutional Neural Network (DGCNN) [1]





all the second and

















25







Evaluation: considered motifs & scenarios



and and and



Evaluation: considered motifs & scenarios



and the state



Providence of















Evaluation

CN	(Mul) -	49.99 ± 0.45	49.78 ± 0.39	50.19 ± 0.47	50.13 ± 0.65	50.37 ± 0.79	51.55 ± 0.32	52.93 ± 0.61	55.42 ± 0.58	54.81 ± 0.58
CN	(Min) -	49.98 ± 0.33	49.72 ± 0.49	50.26 ± 0.59	50.35 ± 0.27	50.48 ± 0.32	51.77 ± 0.40	52.99 ± 0.75	54.75 ± 0.48	54.60 ± 0.85
CN	(Avg) -	49.76 ± 0.32	49.50 ± 0.64	50.18 ± 0.64	50.28 ± 0.51	50.91 ± 0.35	51.70 ± 0.59	53.32 ± 0.35	54.93 ± 0.77	54.20 ± 0.68
AA	(Mul) -	63.05 ± 0.71	62.09 ± 0.57	60.67 ± 0.95	54.95 ± 0.92	51.25 ± 0.63	51.40 ± 0.68	53.92 ± 0.52	55.15 ± 0.77	54.93 ± 0.61
AA	(Min) -	63.34 ± 0.68	62.81 ± 0.80	61.59 ± 0.94	54.81 ± 0.77	51.26 ± 0.38	51.60 ± 0.68	54.15 ± 0.89	54.59 ± 0.65	54.94 ± 0.27
AA	(Avg) -	63.96 ± 0.68	63.66 ± 0.48	62.71 ± 0.52	55.78 ± 0.74	51.28 ± 0.55	51.79 ± 0.62	54.52 ± 0.57	55.20 ± 0.53	54.55 ± 0.39
Jaccard	(Mul) -	67.17 ± 0.92	62.01 ± 0.72	59.71 ± 0.93	69.62 ± 1.09	57.60 ± 0.73	52.75 ± 0.97	51.75 ± 1.10	51.68 ± 0.85	50.93 ± 0.64
Jaccard	(Min) -	69.20 ± 0.80	67.11 ± 0.46	65.24 ± 0.80	73.88 ± 0.88	63.36 ± 1.17	56.50 ± 0.94	52.30 ± 0.54	51.86 ± 0.77	50.87 ± 0.53
Jaccard	(Avg) -	70.12 ± 0.78	68.59 ± 0.71	68.69 ± 0.77	75.35 ± 0.60	67.93 ± 0.87	61.22 ± 1.11	51.76 ± 0.91	49.74 ± 0.75	47.66 ± 0.58
SEAL	(Mul) -	76.68 ± 0.61	74.00 ± 0.50	71.80 ± 0.95	76.25 ± 1.90	63.66 ± 4.01	59.48 ± 4.87	68.53 ± 0.88	67.49 ± 1.27	67.88 ± 1.43
SEAL	(Min) -	77.15 ± 0.43	74.62 ± 0.55	73.11 ± 0.99	78.00 ± 1.49	69.70 ± 3.56	64.49 ± 5.47	66.40 ± 1.44	62.94 ± 1.98	62.88 ± 3.57
SEAL	(Avg) -	77.91 ± 0.91	75.98 ± 0.99	75.71 ± 0.66	77.50 ± 2.35	72.68 ± 3.21	66.95 ± 6.79	66.05 ± 0.78	65.14 ± 0.89	66.99 ± 1.32
SEAN	1, no	86.24 ± 0.99	85.57 ± 0.94	88.61 ± 0.71	91.20 ± 1.03	96.16 ± 0.55	98.40 ± 0.22	83.39 ± 0.94	86.12 ± 0.66	87.86 ± 1.06
embe	aaing	90.78 ± 1.30	90.00 ± 1.84	91.53 ± 1.53	93.06 ± 0.61	97.26 ± 0.23	98.90 ± 0.18	83.81 ± 0.53	87.56 ± 0.79	88.59 ± 1.51
SE	AM	3-star	5-star	7-star	3-clique	5-clique	7-clique	3-db-star	5-db-star	7-db-star

1 1 1 1 1 1 1 1

SEAM: learning from Subgraphs, Embeddings and Attributes for Motif prediction



Evaluation

		5 500	5 500	, 500	5 chque	5 chque	, enque			
SEA	ΔM´	3-star	5-star	7-star	3-clique	5-clique	7-clique	3-db-star	5-db-star	7-db-star
empe	aaing	90.78 ± 1.30	90.00 ± 1.84	91.53 ± 1.53	93.06 ± 0.61	97.26 ± 0.23	98.90 ± 0.18	83.81 ± 0.53	87.56 ± 0.79	88.59 ± 1.51
SEAM	l, no	86.24 ± 0.99	85.57 ± 0.94	88.61 ± 0.71	91.20 ± 1.03	96.16 ± 0.55	98.40 ± 0.22	83.39 ± 0.94	86.12 ± 0.66	87.86 ± 1.06
SEAL ((Avg) -	77.91 ± 0.91	75.98 ± 0.99	75.71 ± 0.66	77.50 ± 2.35	72.68 ± 3.21	66.95 ± 6.79	66.05 ± 0.78	65.14 ± 0.89	66.99 ± 1.32
SEAL	(Min) -	77.15 ± 0.43	74.62 ± 0.55	73.11 ± 0.99	78.00 ± 1.49	69.70 ± 3.56	64.49 ± 5.47	66.40 ± 1.44	62.94 ± 1.98	62.88 ± 3.57
SEAL	(Mul) -	76.68 ± 0.61	74.00 ± 0.50	71.80 ± 0.95	76.25 ± 1.90	63.66 ± 4.01	59.48 ± 4.87	68.53 ± 0.88	67.49 ± 1.27	67.88 ± 1.43
Jaccard ((Avg) -	70.12 ± 0.78	68.59 ± 0.71	68.69 ± 0.77	75.35 ± 0.60	67.93 ± 0.87	61.22 ± 1.11	51.76 ± 0.91	49.74 ± 0.75	47.66 ± 0.58
Jaccard	(Min) -	69.20 ± 0.80	67.11 ± 0.46	65.24 ± 0.80	73.88 ± 0.88	63.36 ± 1.17	56.50 ± 0.94	52.30 ± 0.54	51.86 ± 0.77	50.87 ± 0.53
Jaccard	(Mul) -	67.17 ± 0.92	62.01 ± 0.72	59.71 ± 0.93	69.62 ± 1.09	57.60 ± 0.73	52.75 ± 0.97	51.75 ± 1.10	51.68 ± 0.85	50.93 ± 0.64
AA ((Avg) -	63.96 ± 0.68	63.66 ± 0.48	62.71 ± 0.52	55.78 ± 0.74	51.28 ± 0.55	51.79 ± 0.62	54.52 ± 0.57	55.20 ± 0.53	54.55 ± 0.39
AA	(Min) -	63.34 ± 0.68	62.81 ± 0.80	61.59 ± 0.94	54.81 ± 0.77	51.26 ± 0.38	51.60 ± 0.68	54.15 ± 0.89	54.59 ± 0.65	54.94 ± 0.27
AA	(Mul) -	63.05 ± 0.71	62.09 ± 0.57	60.67 ± 0.95	54.95 ± 0.92	51.25 ± 0.63	51.40 ± 0.68	53.92 ± 0.52	55.15 ± 0.77	54.93 ± 0.61
CN ((Avg) -	49.76 ± 0.32	49.50 ± 0.64	50.18 ± 0.64	50.28 ± 0.51	50.91 ± 0.35	51.70 ± 0.59	53.32 ± 0.35	54.93 ± 0.77	54.20 ± 0.68
CN	(Min) -	49.98 ± 0.33	49.72 ± 0.49	50.26 ± 0.59	50.35 ± 0.27	50.48 ± 0.32	51.77 ± 0.40	52.99 ± 0.75	54.75 ± 0.48	54.60 ± 0.85
CN	(Mul) -	49.99 ± 0.45	49.78 ± 0.39	50.19 ± 0.47	50.13 ± 0.65	50.37 ± 0.79	51.55 ± 0.32	52.93 ± 0.61	55.42 ± 0.58	54.81 ± 0.58

The second second

SEAM: learning from Subgraphs, Embeddings and Attributes for Motif prediction

"db" – with dealbreaker edges





Evaluation

CN: Common Neighbors, **AA**: Adamic-Adar, **SEAL [2]**: Link Prediction using GNNs **Mul, Min, Avg**: different variants (i.e., how link scores are aggregated to form motif scores)

Participante Part

[2] M. Zhang and Y. Chen. 2018. Link prediction based on graph neural networks. NeurIPS'18

CN (Mul) -	49.99 ± 0.45	49.78 ± 0.39	50.19 ± 0.47	50.13 ± 0.65	50.37 ± 0.79	51.55 ± 0.32	52.93 ± 0.61	55.42 ± 0.58	54.81 ± 0.58
CN (Min) -	49.98 ± 0.33	49.72 ± 0.49	50.26 ± 0.59	50.35 ± 0.27	50.48 ± 0.32	51.77 ± 0.40	52.99 ± 0.75	54.75 ± 0.48	54.60 ± 0.85
CN (Avg) -	49.76 ± 0.32	49.50 ± 0.64	50.18 ± 0.64	50.28 ± 0.51	50.91 ± 0.35	51.70 ± 0.59	53.32 ± 0.35	54.93 ± 0.77	54.20 ± 0.68
AA (Mul) -	63.05 ± 0.71	62.09 ± 0.57	60.67 ± 0.95	54.95 ± 0.92	51.25 ± 0.63	51.40 ± 0.68	53.92 ± 0.52	55.15 ± 0.77	54.93 ± 0.61
AA (Min) -	63.34 ± 0.68	62.81 ± 0.80	61.59 ± 0.94	54.81 ± 0.77	51.26 ± 0.38	51.60 ± 0.68	54.15 ± 0.89	54.59 ± 0.65	54.94 ± 0.27
AA (Avg) -	63.96 ± 0.68	63.66 ± 0.48	62.71 ± 0.52	55.78 ± 0.74	51.28 ± 0.55	51.79 ± 0.62	54.52 ± 0.57	55.20 ± 0.53	54.55 ± 0.39
Jaccard (Mul) -	67.17 ± 0.92	62.01 ± 0.72	59.71 ± 0.93	69.62 ± 1.09	57.60 ± 0.73	52.75 ± 0.97	51.75 ± 1.10	51.68 ± 0.85	50.93 ± 0.64
Jaccard (Min) -	69.20 ± 0.80	67.11 ± 0.46	65.24 ± 0.80	73.88 ± 0.88	63.36 ± 1.17	56.50 ± 0.94	52.30 ± 0.54	51.86 ± 0.77	50.87 ± 0.53
Jaccard (Avg) -	70.12 ± 0.78	68.59 ± 0.71	68.69 ± 0.77	75.35 ± 0.60	67.93 ± 0.87	61.22 ± 1.11	51.76 ± 0.91	49.74 ± 0.75	47.66 ± 0.58
SEAL (Mul) -	76.68 ± 0.61	74.00 ± 0.50	71.80 ± 0.95	76.25 ± 1.90	63.66 ± 4.01	59.48 ± 4.87	68.53 ± 0.88	67.49 ± 1.27	67.88 ± 1.43
SEAL (Min) -	77.15 ± 0.43	74.62 ± 0.55	73.11 ± 0.99	78.00 ± 1.49	69.70 ± 3.56	64.49 ± 5.47	66.40 ± 1.44	62.94 ± 1.98	62.88 ± 3.57
SEAL (Avg) -	77.91 ± 0.91	75.98 ± 0.99	75.71 ± 0.66	77.50 ± 2.35	72.68 ± 3.21	66.95 ± 6.79	66.05 ± 0.78	65.14 ± 0.89	66.99 ± 1.32
SEAM, no	86.24 ± 0.99	85.57 ± 0.94	88.61 ± 0.71	91.20 ± 1.03	96.16 ± 0.55	98.40 ± 0.22	83.39 ± 0.94	86.12 ± 0.66	87.86 ± 1.06
embedding	90.78 ± 1.30	90.00 ± 1.84	91.53 ± 1.53	93.06 ± 0.61	97.26 ± 0.23	98.90 ± 0.18	83.81 ± 0.53	87.56 ± 0.79	88.59 ± 1.51
SEAM	3-star	5-star	7-star	3-clique	5-clique	7-clique	3-db-star	5-db-star	7-db-star

SEAM: learning from Subgraphs, Embeddings and Attributes for Motif prediction





spcl.inf.ethz.ch

Evaluatio	With DN	correlation	CN: (Common Ne	ighbors, AA :	Adamic-Ada	ar, SEAL [2]:	Link Predicti	on using GNNs	
		Mu	I, Min, Avg:	different vai	riants (i.e., h	ow link score	es are aggre	gated to forr	n motif scores)	
	No	correlation		[2] M. Zhan	g and Y. Chen. 2	018. Link predie	ction based on g	graph neural net	works. NeurlPS'18	
CN (Mul) -	49.99 ± 0.45	49.78 ± 0.39	50.19 ± 0.47	50.13 ± 0.65	50.37 ± 0.79	51.55 ± 0.32	52.93 ± 0.61	55.42 ± 0.58	54.81 ± 0.58	
CN (Min) -	49.98 ± 0.33	49.72 ± 0.49	50.26 ± 0.59	50.35 ± 0.27	50.48 ± 0.32	51.77 ± 0.40	52.99 ± 0.75	54.75 ± 0.48	54.60 ± 0.85	
CN (Avg) -	49.76 ± 0.32	49.50 ± 0.64	50.18 ± 0.64	50.28 ± 0.51	50.91 ± 0.35	51.70 ± 0.59	53.32 ± 0.35	54.93 ± 0.77	54.20 ± 0.68	
AA (Mul) -	63.05 ± 0.71	62.09 ± 0.57	60.67 ± 0.95	54.95 ± 0.92	51.25 ± 0.63	51.40 ± 0.68	53.92 ± 0.52	55.15 ± 0.77	54.93 ± 0.61	
AA (Min) -	63.34 ± 0.68	62.81 ± 0.80	61.59 ± 0.94	54.81 ± 0.77	51.26 ± 0.38	51.60 ± 0.68	54.15 ± 0.89	54.59 ± 0.65	54.94 ± 0.27	
AA (Avg) -	63.96 ± 0.68	63.66 ± 0.48	62.71 ± 0.52	55.78 ± 0.74	51.28 ± 0.55	51.79 ± 0.62	54.52 ± 0.57	55.20 ± 0.53	54.55 ± 0.39	
Jaccard (Mul) -	67.17 ± 0.92	62.01 ± 0.72	59.71 ± 0.93	69.62 ± 1.09	57.60 ± 0.73	52.75 ± 0.97	51.75 ± 1.10	51.68 ± 0.85	50.93 ± 0.64	
Jaccard (Min) -	69.20 ± 0.80	67.11 ± 0.46	65.24 ± 0.80	73.88 ± 0.88	63.36 ± 1.17	56.50 ± 0.94	52.30 ± 0.54	51.86 ± 0.77	50.87 ± 0.53	
Jaccard (Avg) -	70.12 ± 0.78	68.59 ± 0.71	68.69 ± 0.77	75.35 ± 0.60	67.93 ± 0.87	61.22 ± 1.11	51.76 ± 0.91	49.74 ± 0.75	47.66 ± 0.58	
SEAL (Mul) -	76.68 ± 0.61	74.00 ± 0.50	71.80 ± 0.95	76.25 ± 1.90	63.66 ± 4.01	59.48 ± 4.87	68.53 ± 0.88	67.49 ± 1.27	67.88 ± 1.43	
SEAL (Min) -	77.15 ± 0.43	74.62 ± 0.55	73.11 ± 0.99	78.00 ± 1.49	69.70 ± 3.56	64.49 ± 5.47	66.40 ± 1.44	62.94 ± 1.98	62.88 ± 3.57	
SEAL (Avg) -	77.91 ± 0.91	75.98 ± 0.99	75.71 ± 0.66	77.50 ± 2.35	72.68 ± 3.21	66.95 ± 6.79	66.05 ± 0.78	65.14 ± 0.89	66.99 ± 1.32	
SEAM, no embedding	86.24 ± 0.99	85.57 ± 0.94	88.61 ± 0.71	91.20 ± 1.03	96.16 ± 0.55	98.40 ± 0.22	83.39 ± 0.94	86.12 ± 0.66	87.86 ± 1.06	
SEAM	90.78 ± 1.30	90.00 ± 1.84	91.53 ± 1.53	93.06 ± 0.61	97.26 ± 0.23	98.90 ± 0.18	83.81 ± 0.53	87.56 ± 0.79	88.59 ± 1.51	
SEAM	3-star	5-star	7-star	3-clique	5-clique	7-clique	3-db-star	5-db-star	7-db-star	
SEAM: learning from Subgraphs, Embeddings										

Caller Charles The

"db" – with dealbreaker edges


spcl.inf.ethz.ch



breaker edges



spcl.inf.ethz.ch



The second

EHzürich

MACIEJ BESTA, TORSTEN HOEFLER, ET AL. Motif Prediction with Graph Neural Networks

Thank you for your attention