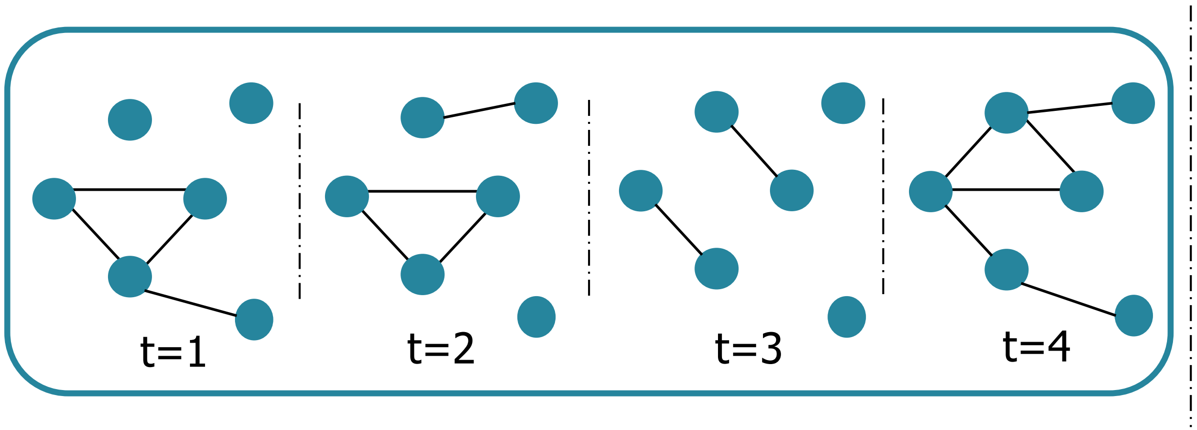
Higher-Order Temporal Network Prediction and Interpretation

Bart Peters (TUD)
In collaboration with Huijuan Wang (TUD) & Alberto Ceria (LEI)

Graphs & Data Seminar
13/02/2025

Temporal networks

- Pairwise temporal network G



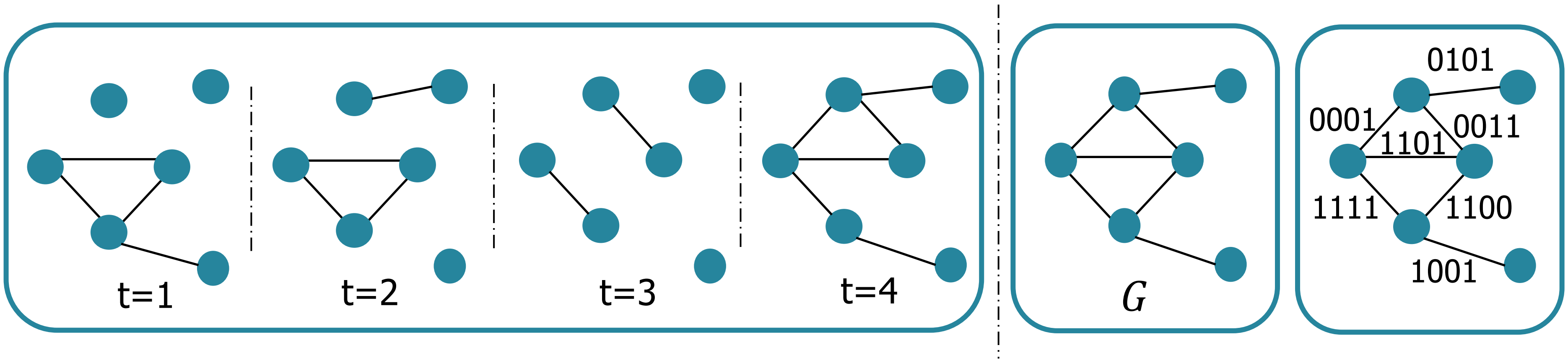
Sequence of network snapshots:
$G = \{G_1, G_2, G_3, G_4\}$ with $G_t = (V, E_t)$

V = set of nodes

E_t = set of links (interactions) at t

Temporal networks

- <u>Pairwise</u> temporal network G
- Two representations



Sequence of network snapshots:
$G = \{G_1, G_2, G_3, G_4\}$ with $G_t = (V, E_t)$

V = set of nodes

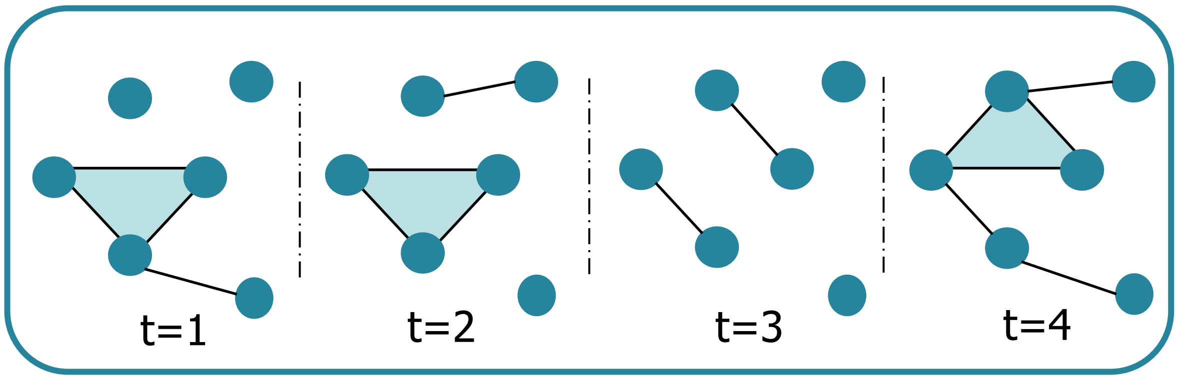
E_t = set of <u>links</u> (interactions) at t

Aggregated network G

Activity of <u>link</u> i:
$\boldsymbol{x}_i = \{x_i(1), x_i(2), x_i(3), x_i(4)\}$

Higher-order temporal networks

- Higher-order temporal network H



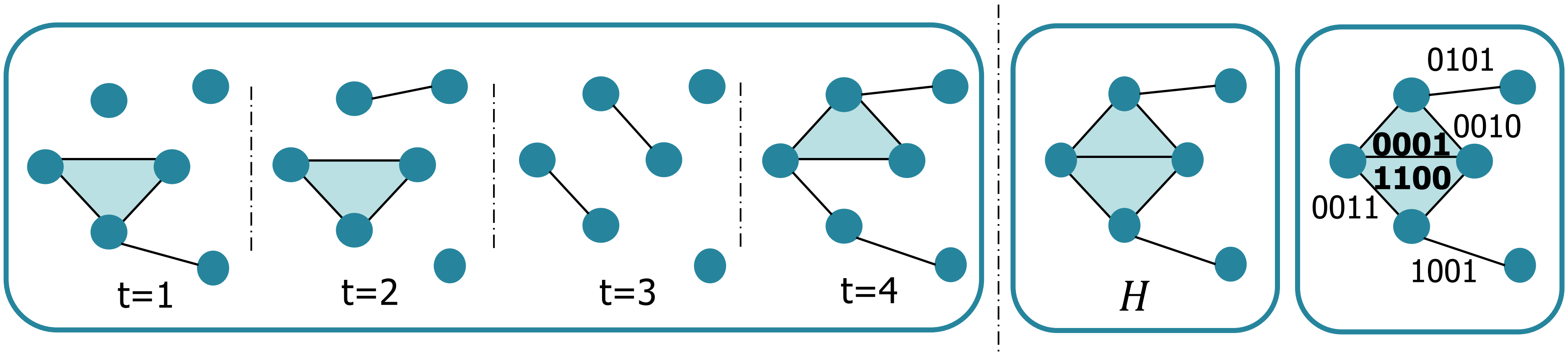
Sequence of network snapshots:
$H = \{H_1, H_2, H_3, H_4\}$ with $H_t = (V, \mathcal{E}_t)$

V = set of nodes

$\mathcal{E}_t$ = set of hyperlinks (interactions) at t

Higher-order temporal networks

- Higher-order temporal network H
- Two representations



Sequence of network snapshots:
$H = \{H_1, H_2, H_3, H_4\}$ with $H_t = (V, \mathcal{E}_t)$

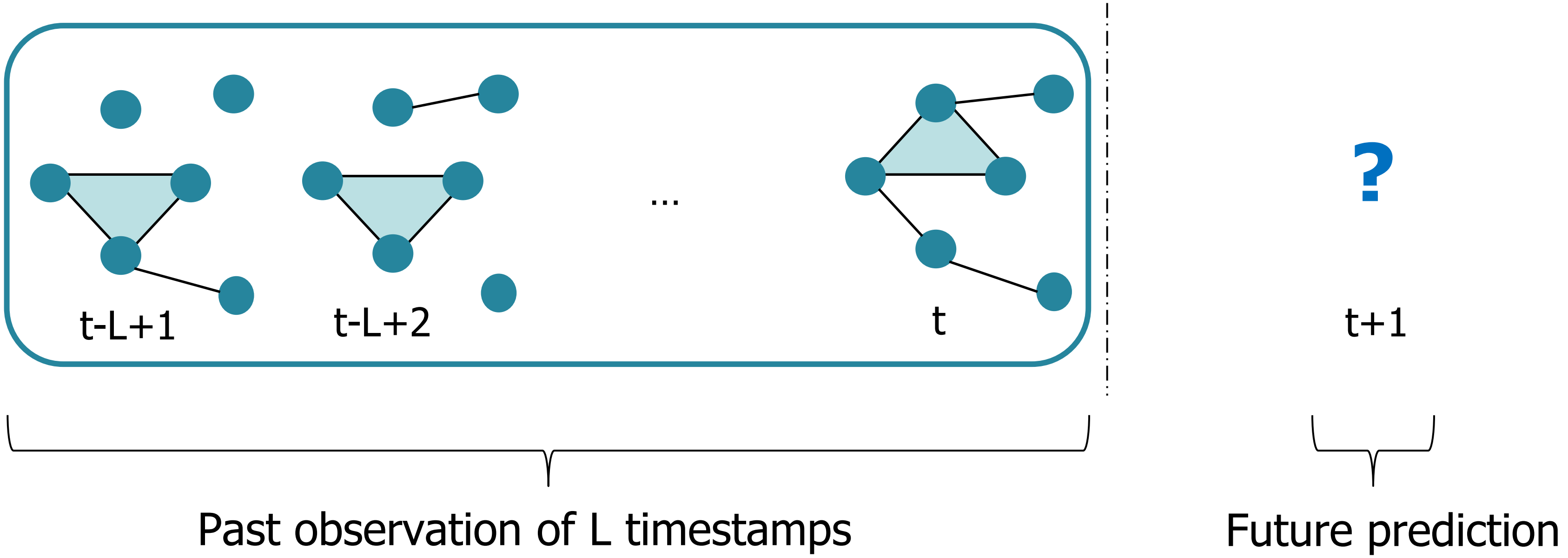
V = set of nodes
$\mathcal{E}_t$ = set of hyperlinks (interactions) at t

Aggregated network H

Activity of hyperlink i:
$\boldsymbol{x}_i = \{x_i(1), x_i(2), x_i(3), x_i(4)\}$

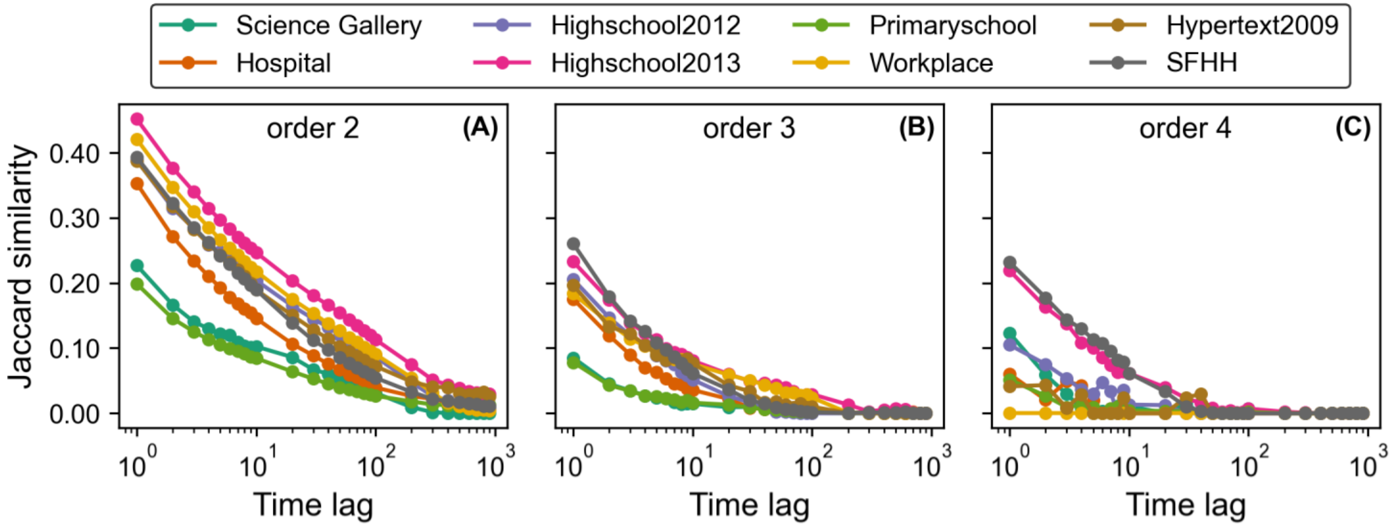
Goal

- Predict future hyperlink activity
- Understand prediction mechanism



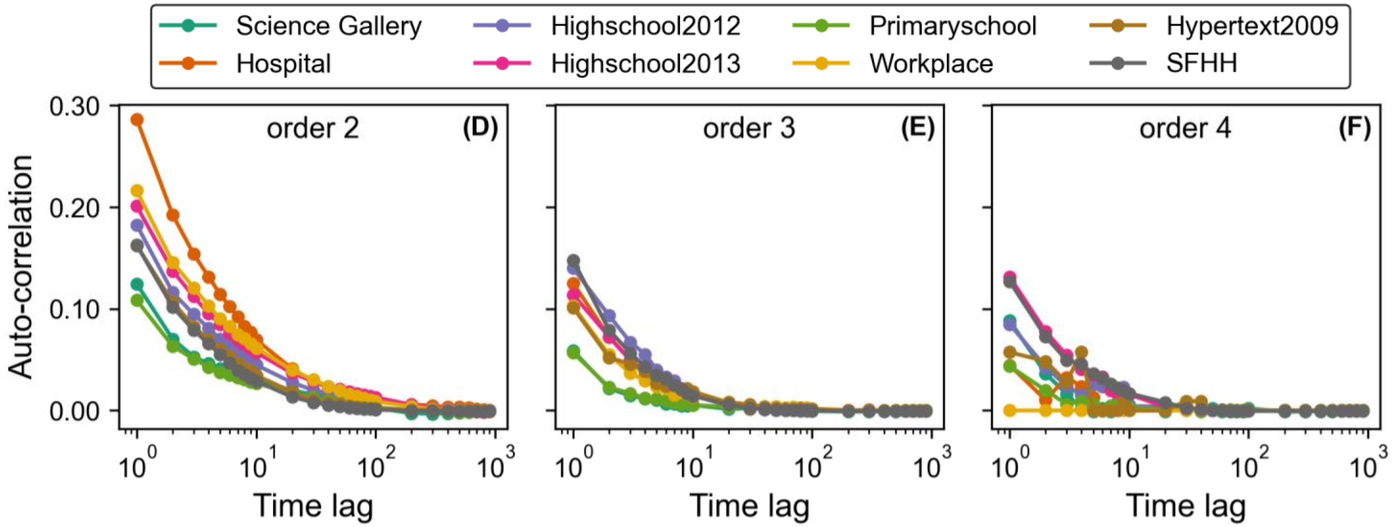
Memory in HOTNs: Jaccard similarity

- Similarity of network topology over time
  - $$J_d(\Delta) = \frac{\left|\mathcal{E}_t^d \cap \mathcal{E}_{t+\Delta}^d\right|}{\left|\mathcal{E}_t^d \cup \mathcal{E}_{t+\Delta}^d\right|}$$

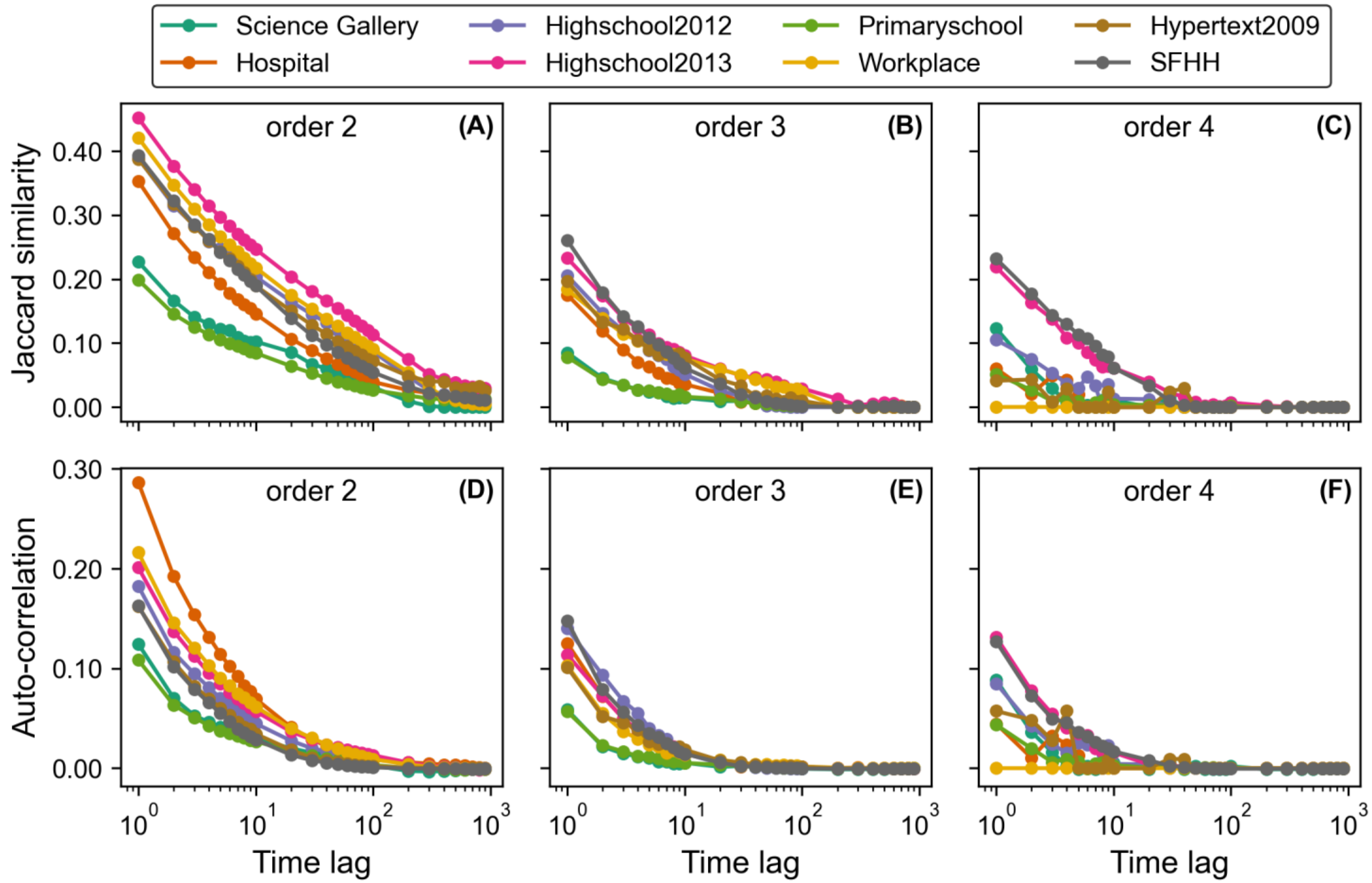


Memory in HOTNs: auto-correlation

- Similarity of hyperlink activity over time
  - Pearson correlation coefficient between $\{x_i(t)\}_{t=1,2,\ldots,T-\Delta}$ and $\{x_i(t)\}_{t=\Delta+1,\Delta+2,\ldots,T}$



Memory in HOTNs

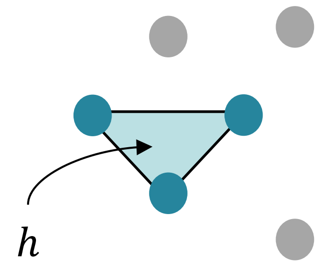


Prediction: self-driven model

- Recent events (interactions) have more influence than past events
- $w_i(t+1) = \sum_{k=t-L+1}^{k=t} x_i(k) e^{-\tau(t-k)}$
- Only uses past activity of target hyperlink i

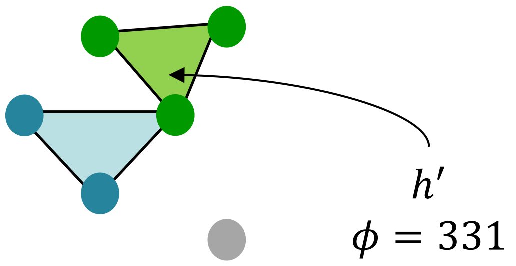
‘Neighboring’ hyperlinks

- Definition: For a target hyperlink h, a neighboring hyperlink h' is called a type ϕ-neighbor of h, with $\phi = (dd'o)$, where:
  - $d =$ order of h
  - $d' =$ order of h'
  - $o =$ #overlapping nodes



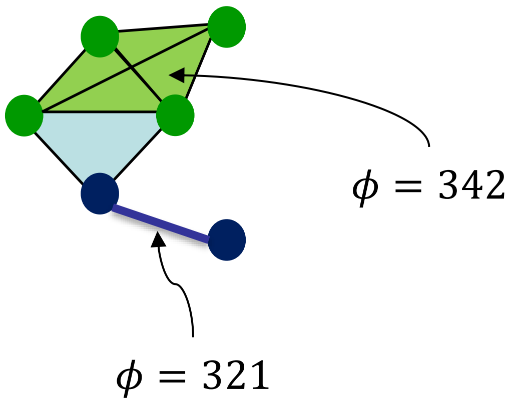
‘Neighboring’ hyperlinks

- Definition: For a target hyperlink h, a neighboring hyperlink h' is called a type ϕ-neighbor of h, with $\phi = (dd'o)$, where:
  - d = order of h
  - d' = order of h'
  - o = #overlapping nodes



‘Neighboring’ hyperlinks

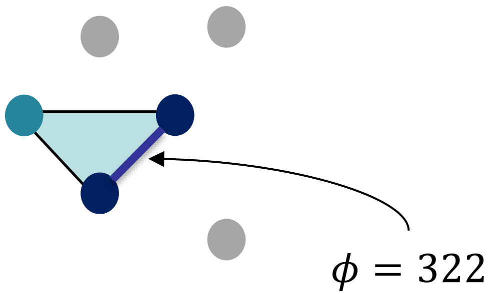
- Definition: For a target hyperlink h, a neighboring hyperlink h' is called a type ϕ-neighbor of h, with $\phi = (dd'o)$, where:
  - d = order of h
  - d' = order of h'
  - o = #overlapping nodes



‘Neighboring’ hyperlinks

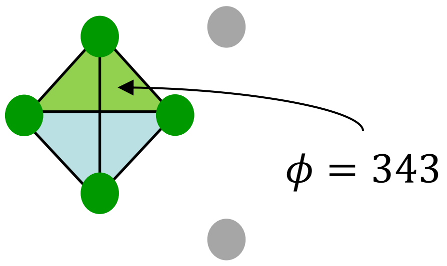
- Definition: For a target hyperlink h, a neighboring hyperlink h' is called a type ϕ-neighbor of h, with $\phi = (dd'o)$, where:
  - d = order of h
  - d' = order of h'
  - o = #overlapping nodes

- Sub-hyperlinks: $h' \subset h$



'Neighboring' hyperlinks

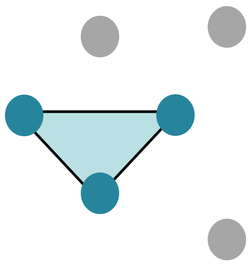
- Definition: For a target hyperlink h, a neighboring hyperlink h' is called a type ϕ-neighbor of h, with $\phi = (dd'o)$, where:
  - $d =$ order of h
  - $d' =$ order of h'
  - $o =$ #overlapping nodes



- Sub-hyperlinks: $h' \subset h$
- Super-hyperlinks: $h' \supset h$

‘Neighboring’ hyperlinks

- Definition: For a target hyperlink h, a neighboring hyperlink h' is called a type ϕ-neighbor of h, with $\phi = (dd'o)$, where:
  - d = order of h
  - d' = order of h'
  - o = #overlapping nodes



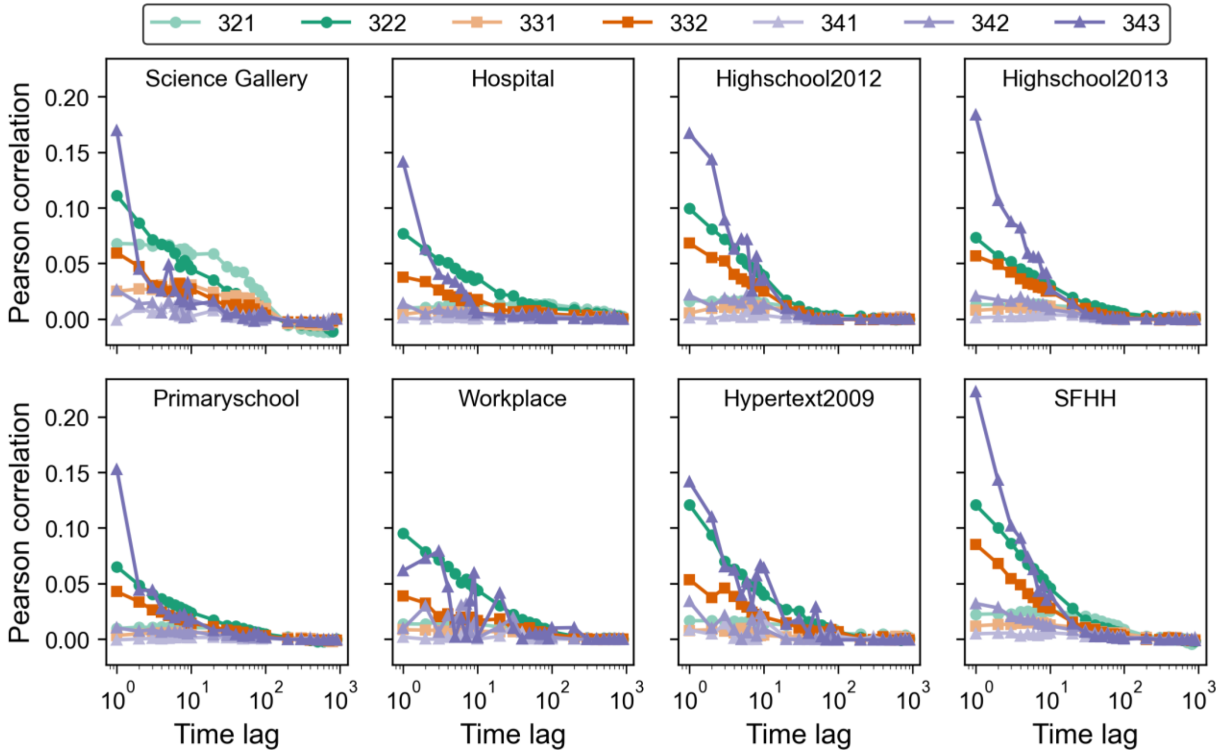
- Sub-hyperlinks: $h' \subset h$
- Super-hyperlinks: $h' \supset h$

- All possible types ϕ for hyperlink order $d \in [2,3,4]$:

order d	ϕ			
	target	**sub-hyperlinks**	**sup-hyperlinks**	**other**
2	222	-	232, 242	221, 231, 241
3	333	322	343	321, 331, 332, 341, 342
4	444	422, 433	-	421, 431, 432, 441, 442, 443

Memory in HOTNs: target and neighbors

- Similarity of target and neighbor activity over time
  - Pearson correlation coefficient between target's activity and average 'lagged' activity of neighbors



Prediction: general model

- Sum over past activations of target and neighbors

- $$w_i(t+1) = \sum_{\phi \in \Phi^d} c_\phi \, y_i^\phi(t) + c_d$$
  $$y_i^\phi(t) = \sum_{k=t-L+1}^{k=t} \sum_{j \in S_i^\phi} x_j(k) e^{-\tau(t-k)}$$

- Coefficients learned using Lasso regression

Prediction: general model

- Sum over past activations of target and neighbors
- $w_i(t+1) = \sum_{\phi \in \Phi^d} c_\phi \, y_i^\phi(t) + c_d$

  $y_i^\phi(t) = \sum_{k=t-L+1}^{k=t} \sum_{j \in S_i^\phi} x_j(k) e^{-\tau(t-k)}$
- Coefficients learned using Lasso regression

- Coefficients + memory analysis: target is most important, followed by sub- & super-hyperlinks

Prediction: refined model

- Sum over past activations of target and sub- & super-hyperlinks
- $w_i(t+1) = \sum_{\phi \in \Phi^d} c_\phi \, y_i^\phi(t) + c_d$

  $y_i^\phi(t) = \sum_{k=t-L+1}^{k=t} \sum_{j \in S_i^\phi} x_j(k) e^{-\tau(t-k)}$

General model

order d	ϕ			
	target	**sub-hyperlinks**	**sup-hyperlinks**	**other**
2	222	-	232, 242	221, 231, 241
3	333	322	343	321, 331, 332, 341, 342
4	444	422, 433	-	421, 431, 432, 441, 442, 443

Refined model

Prediction: results

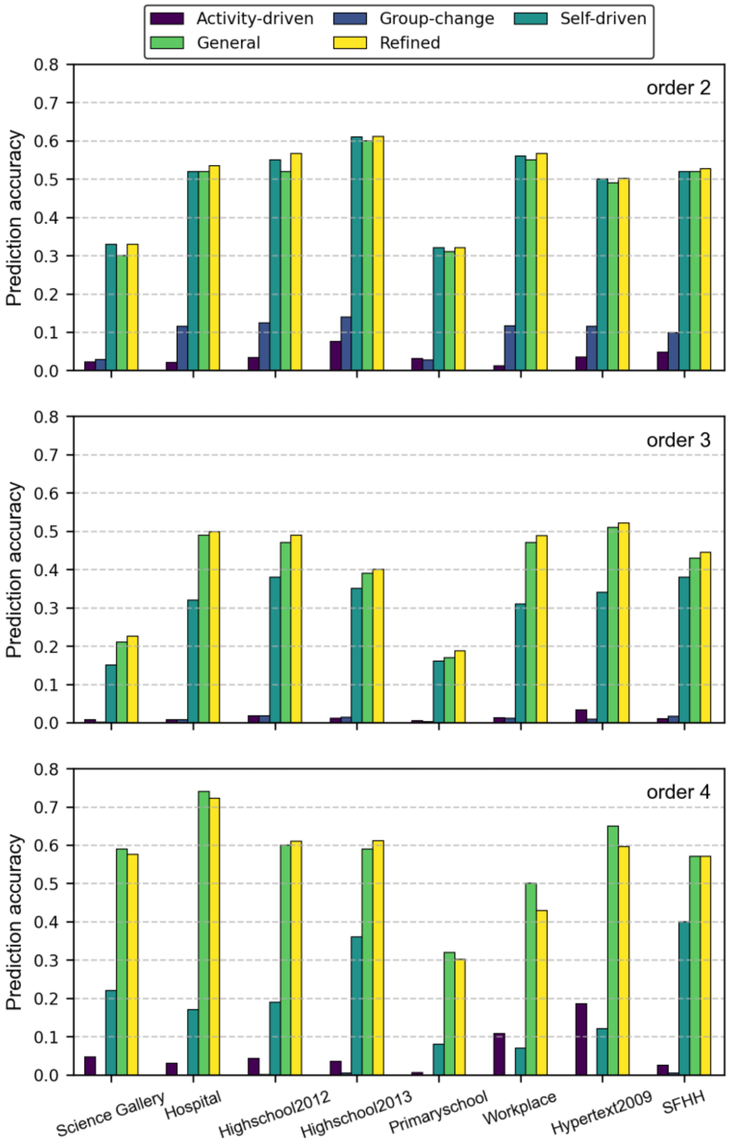
- Pairwise self-driven model as baseline:

$$w_i(t+1) = \sum_{k=t-L+1}^{k=t} x_i(k) e^{-\tau(t-k)}$$

- General and refined model:

$$w_i(t+1) = \sum_{\phi \in \Phi^d} c_\phi \, y_i^\phi(t) + c_d$$

$$y_i^\phi(t) = \sum_{k=t-L+1}^{k=t} \sum_{j \in S_i^\phi} x_j(k) e^{-\tau(t-k)}$$



Conclusions

- General and refined model outperform baseline models

- Past activity of the target itself is the most important factor in forecasting its activity, followed by the past activity of its sub- & super-hyperlinks