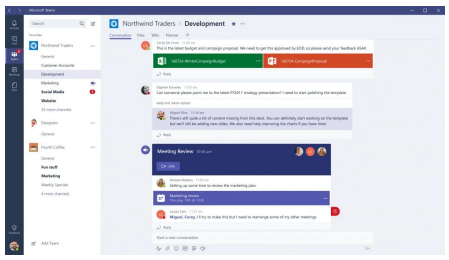
Workplace Recommendation with Temporal Network Objectives

Kiran Tomlinson,[1] Jennifer Neville,[2] Longqi Yang,[2] Mengting Wan,[2] Cao Lu[2]
[1]Cornell University, [2]Microsoft

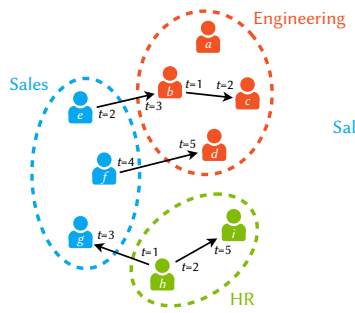
Background

Workplace communication platform (Microsoft Teams)

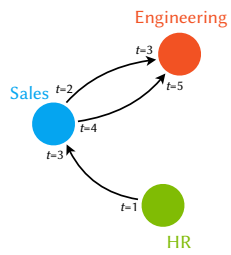


Remote work makes communication more siloed [1]

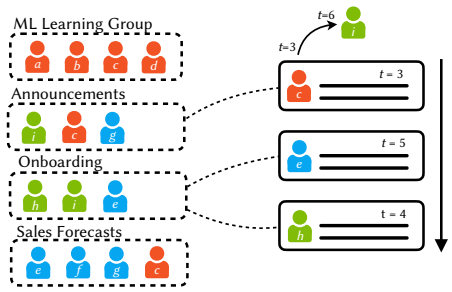
Temporal workplace communication network



Cross-cluster network

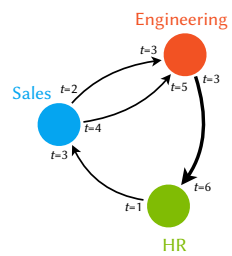


Teams channels



Post recommendation

New edge because of recommendation



Goal: Recommend posts to improve cross-team information flow in a company's communication network.

Methods

Information flow metrics

Graph G with temporal edges (u, v, d, a, w);
source u, target v, departure time d, arrival time a, weight w

$\mathscr{P}_t(u, v)$: set of all temporal paths from u to v arriving no later than t

Information Latency [2] at time t:
"How fresh could your information about another node be?"

$$\mathrm{IL}(G, t) = \sum_{u,v \in V} t - \max_{P \in \mathscr{P}_t(u,v)} d(P)$$

Total Information (new) at time t:
"What fraction of another node's state do you know?"

$$\mathrm{TI}(G, t) = \sum_{u,v \in V} \mathrm{TI}(u, v, t)$$

$$\mathrm{TI}(u, v, t) = \min\{1, \underbrace{\lambda \mathrm{TI}(u, v, t-1)}_{\text{discounted prior knowledge}} + \sum_{(z,v,d,a,w) \in E_t} \underbrace{w \lambda^{t-d} \mathrm{TI}(u, z, t-1, d)}_{\text{information arriving at time } t \text{ via } z}\}$$

$$\mathrm{TI}(u, u, t) = 1 \quad \text{(self-information always 1)}$$

Theory: edge addition problems

Find k edges in/out of u whose addition optimizes IL/TI now (myopic) or given future edge stream

Hardness Results (for both IL & TI)

	Myopic	*With future edges*
In-edges	NP-hard	NP-hard
Out-edges	P	NP-hard

Theorem. Greedy algorithm is a $1 - 1/e$ approximation for the hard problems. (Submodularity!)

Our recommendations take a greedy approach to optimizing information flow, inspired by this result.

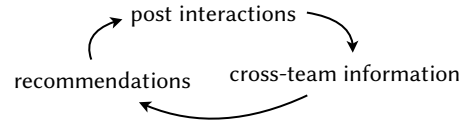
Our method: TIER

Temporal Information and Engagement Recommender

1. Track matrix of cross-cluster information metrics (IL/TI) as communication occurs
2. When recommending post t, compute **network score** n_p and **relevance score** r_p
   - n_p: normalized improvement in information matrix if new edge is added
   - r_p: normalized score from any traditional engagement-based rec sys
3. Rank posts by $n_p + \alpha r_p$

Offline evaluation

- How to evaluate before A/B test?
- Challenge: feedback cycle
- *Simulate user actions!*
- Given historical post interaction, replace with recommended post w.p. ρ

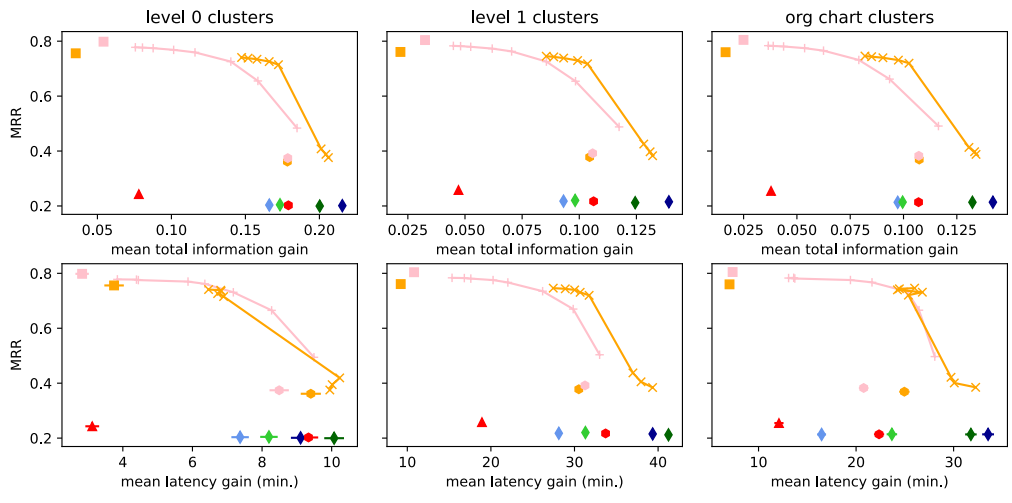


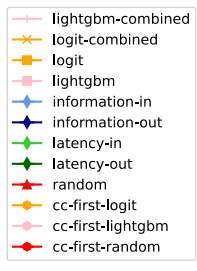
Data

- One month of internal Microsoft communication
- ~180k nodes, ~100m edges
- Teams posts, Teams chats, Outlook emails, SharePoint file sharing
- Clusters from org chart or clustering on email network (level 0, level 1)

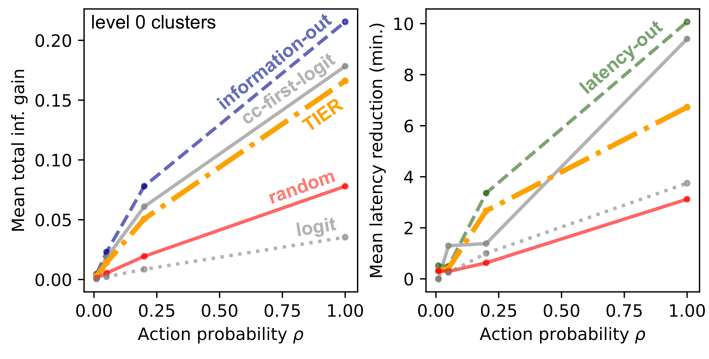
Results

TIER can improve cross-team information flow with a very small relevance drop.

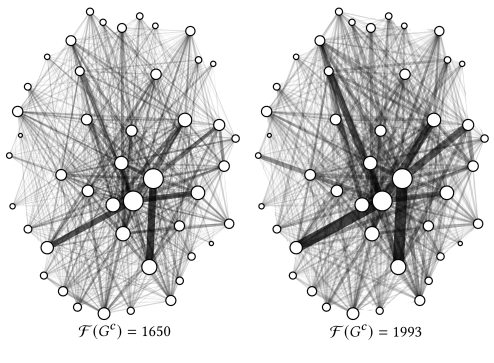


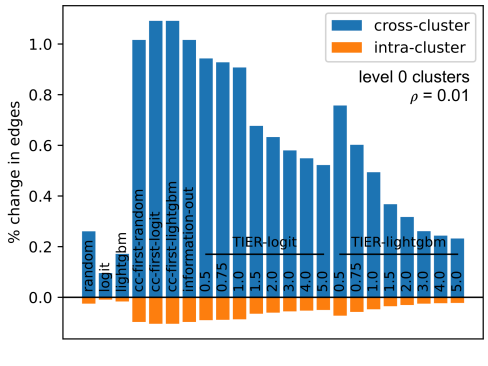


Relevance matters: the more likely users are to engage with posts, the more impact we can have (but diminishing returns).

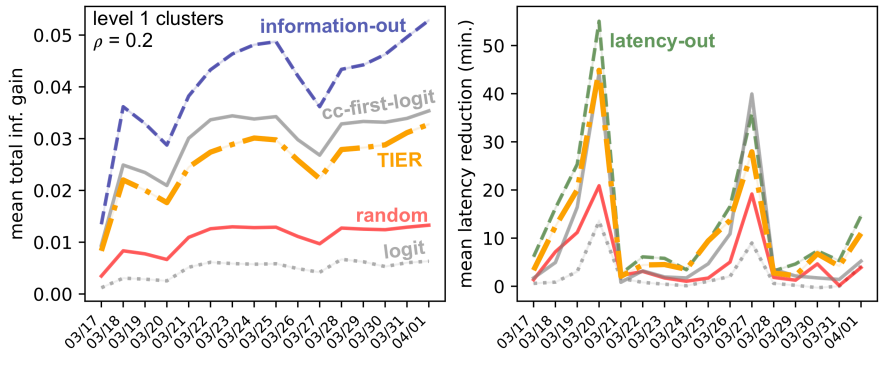


A relatively small loss in intra-cluster edges can create a large cross-cluster gain.





Information gains follow periodic workweek cycles.



References

[1] Longqi Yang, David Holtz, Sonia Jaffe, Siddharth Suri, Shilpi Sinha, Jeffrey Weston, Connor Joyce, Neha Shah, Kevin Sherman, Brent Hecht, et al. 2022. The effects of remote work on collaboration among information workers. *Nature Human Behaviour* 6, 1 (2022), 43–54.

[2] Gueorgi Kossinets, Jon Kleinberg, and Duncan Watts. 2008. The structure of information pathways in a social communication network. In *KDD*. 435–443.

Contact info
Kiran Tomlinson
kt@cs.cornell.edu

Full paper →