THE WEB CONFERENCE ACM

APRIL 30-MAY, 4, 2023

Austin, TX

Everything Evolves in Personalized PageRank

Zihao Li *
(UIUC)

Dongqi Fu *
(UIUC)

Jingrui He
(UIUC)

Contact: {zihaoli5, dongqif2, jingrui}@illinois.edu

GitHub: https://github.com/DongqiFu/EvePPR

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

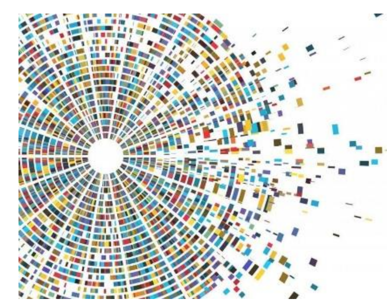
Roadmap

- Motivation
- Proposed EvePPR Method
- Experiments
- Conclusion

Personalized PageRank

- Methodology
  - Obtains broader attention within the trend of deep learning
  - E.g., PageRank-based Graph Representation Learning [1] and Graph Neural Networks [2]

- Applications [3]
  - Search Engine
  - Social Network Analysis
  - Recommender System
  - Bioinformatics
  - Many more ……



[1] Bryan Perozzi, Rami Al-Rfou, Steven Skiena. DeepWalk: online learning of social representations. KDD 2014
[2] Klicpera, et al. Predict then Propagate: Graph Neural Networks meet Personalized PageRank. ICLR 2019
[3] David F. Gleich. PageRank Beyond the Web. SIAM Rev. 2015

PageRank in the Dynamic Setting

- Static Solution [1]
  - $\boldsymbol{v} = \alpha \boldsymbol{P} \boldsymbol{v} + (1 - \alpha)\boldsymbol{h}$
    - $\boldsymbol{v}$ is personalized PageRank vector
    - $\boldsymbol{P}$ is the transition matrix
    - $\boldsymbol{h}$ is the stochastic vector (e.g., personal interest)

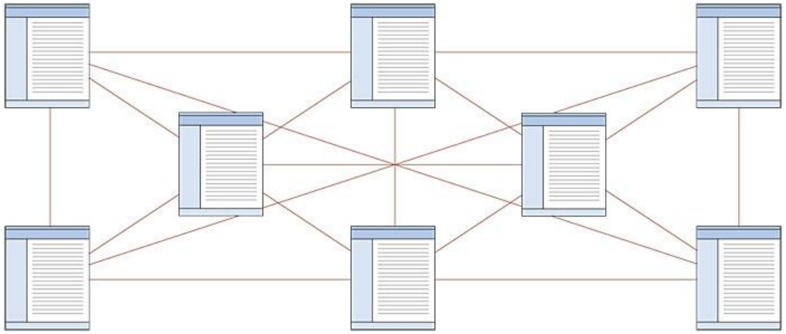
- Dynamic Solution
  - Previous dynamic PPR works focus on modeling the evolving graph structure $\boldsymbol{P}^{(t)}$, e.g.,
    - Gauss-Southwell [1]
    - Local Push [2]
    - Offset Score Propagation [3]

[1] Lawrence, et al. The pagerank citation ranking: Bring order to the web. technical report, Stanford University, 1998
[2] Ohsaka et al. Efficient PageRank Tracking in Evolving Networks. KDD 2015
[3] Zhang et al. Approximate Personalized PageRank on Dynamic Graphs. KDD 2016
[4] Yoon et al. Fast and Accurate Random Walk with Restart on Dynamic Graphs with Guarantees. WWW 2018

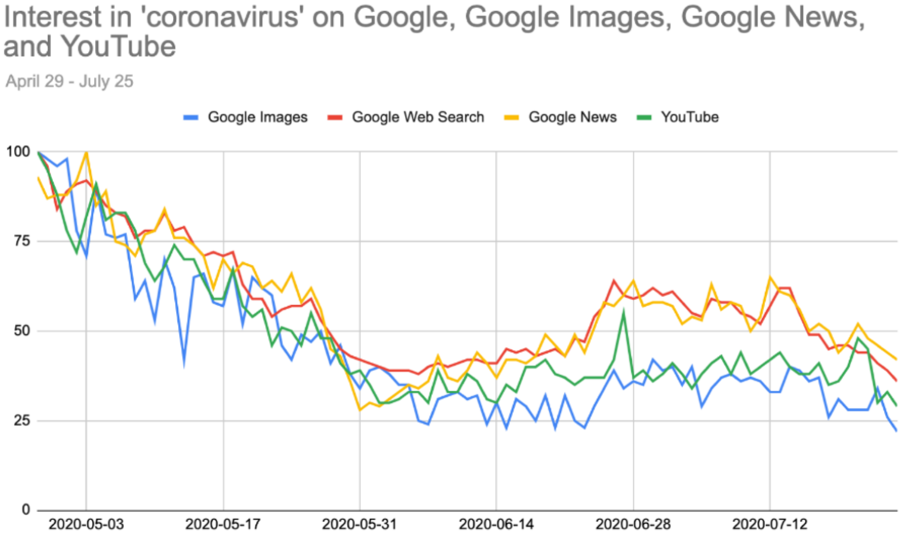
PageRank in the Fully Dynamic Setting

- What if $\boldsymbol{h}$ is also evolving with time, like $\boldsymbol{h}^{(t)}$?

- E.g., web structure $\boldsymbol{P}^{(t)}$ can evolve, as well as user interest $\boldsymbol{h}^{(t)}$



Link expiration or breaking news



- Then, how to solve $\boldsymbol{v}^{(t)} = \alpha \boldsymbol{P}^{(t)} \boldsymbol{v}^{(t)} + (1-\alpha)\boldsymbol{h}^{(t)}$?

Roadmap

- **Motivation**
- **Proposed EvePPR Method**
- **Experiments**
- **Conclusion**

Problem Setting and Theoretical Contribution

- **Input**: Evolving graph structures $\{\boldsymbol{P}^{(1)}, \boldsymbol{P}^{(2)}, \ldots, \boldsymbol{P}^{(t)}\}$ and evolving stochastic vectors $\{\boldsymbol{h}^{(1)}, \boldsymbol{h}^{(2)}, \ldots, \boldsymbol{h}^{(t)}\}$

- **Output**: We aim to solve $\{\boldsymbol{v}^{(1)}, \boldsymbol{v}^{(2)}, \ldots, \boldsymbol{v}^{(t)}\}$ effectively and efficiently.

- Targeting this setting, we provide the solution **EvePPR** with theoretical time complexity and error bound.

- Also, we provide the fast and accuracy-comparable version, **EvePPR-APP** with theoretical analysis.

EvePPR

- Core Idea
  - When interests vary, i.e., $\Delta\boldsymbol{h} = \boldsymbol{h}^{(t)} - \boldsymbol{h}^{(t-1)}$, decompose $\Delta\boldsymbol{h}$ into multiple single-source interests.
  - For each $\Delta h(\mathrm{i}) \neq 0$, execute a single-source tracking $\boldsymbol{v}_{mid}$
  - Then, combine multiple $\boldsymbol{v}_{mid}$ to get $\boldsymbol{v}^{(t)}$

- Theoretical Analysis
  - At each timestamp t, EvePPR can get exact $v^{(t)}$ satisfying $\boldsymbol{v}^{(t)} = \alpha\boldsymbol{P}^{(t)}\boldsymbol{v}^{(t)} + (1-\alpha)\boldsymbol{h}^{(t)}$

  - The time complexity is $O(m(l+1)log_{\alpha}\varepsilon)$
    - m and l is num. of non-zero entries of $\boldsymbol{P}^{(t)}$ and $\boldsymbol{h}^{(t)}$, resp.
    - ε is the tolerance to terminate the tracking

EvePPR-APP

- Core Idea
  - Trade off the tracking accuracy to tracking efficiency
  - Instead of decomposing $\Delta\boldsymbol{h}$ in EvePPR, use OSP [1] to get $\boldsymbol{v}_{mid}$ and refine $\boldsymbol{v}_{mid}$ through Gauss-Southwell [2] to get $\boldsymbol{v}^{(t)}$

- Theoretical Analysis
  - The time complexity is $O\left(mlog_{\alpha}\varepsilon + n\left(\frac{(1+\alpha)\alpha}{(1+\alpha)^2}\right) + \frac{||\Delta h||_1}{\varepsilon}\right)$
    - m is num. of non-zero entries of $\boldsymbol{P}^{(t)}$
    - ε is the tolerance to terminate the tracking

  - The tracking error (i.e., L_1 norm) is bounded by $\frac{n\varepsilon}{1-\alpha}$

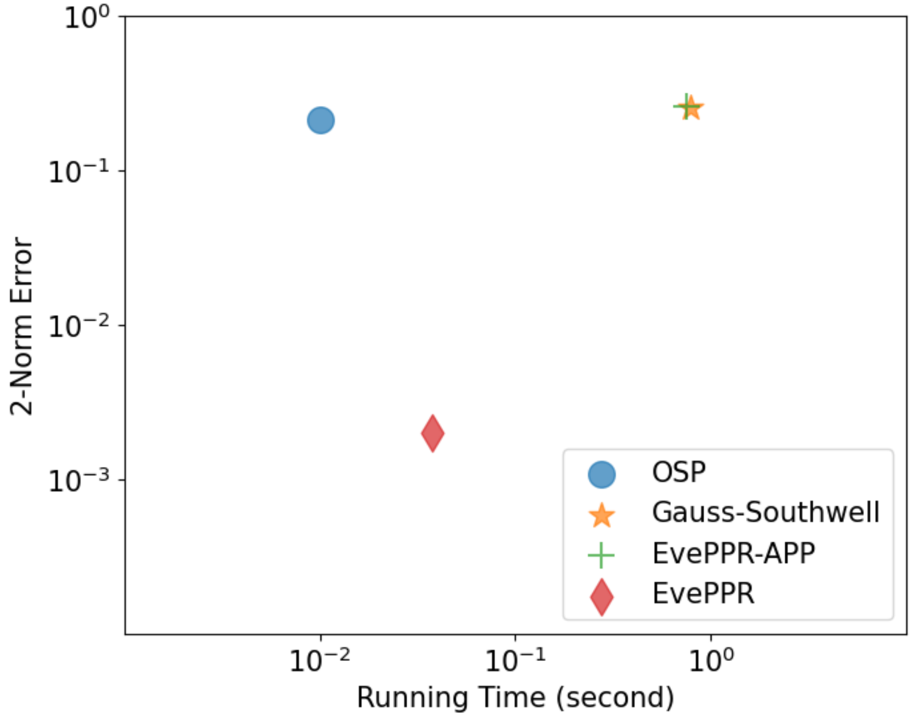
[1] Naoto Ohsaka, Takanori Maehara, Ken-ichi Kawarabayashi. Efficient PageRank Tracking in Evolving Networks. KDD 2015
[2] Minji Yoon, Woojeong Jin, U Kang. Fast and Accurate Random Walk with Restart on Dynamic Graphs with Guarantees. WWW 2018

Roadmap

- **Motivation**
- **Proposed EvePPR Method**
- **Experiments**
- **Conclusion**

PageRank Tracking Experiment

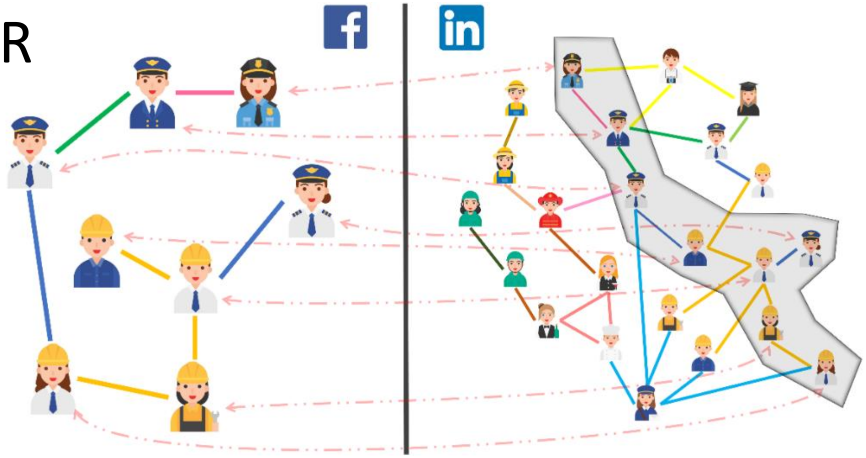
- Our method is fast and has less tracking errors



Tracking Error and Running Time of Different PageRank Algorithms in MathOverflow Network (24,818 nodes and 506,550 edges)

Application Experiment – Graph Alignment

- Node similarity retrieval across graphs [1]
- Can be rewritten in the form of PPR
  $$\boldsymbol{s} = \alpha \widetilde{\boldsymbol{W}} \boldsymbol{s} + (1-\alpha)\boldsymbol{h}$$
  - $\widetilde{\boldsymbol{W}}$ encodes graph topology, node feature, edge feature of two graphs, and $\boldsymbol{h}$ encodes the prior aligning knowledge
  - $\boldsymbol{s} \in \mathbb{R}^{n_1 \times n_2}$ encodes node-pair similarity

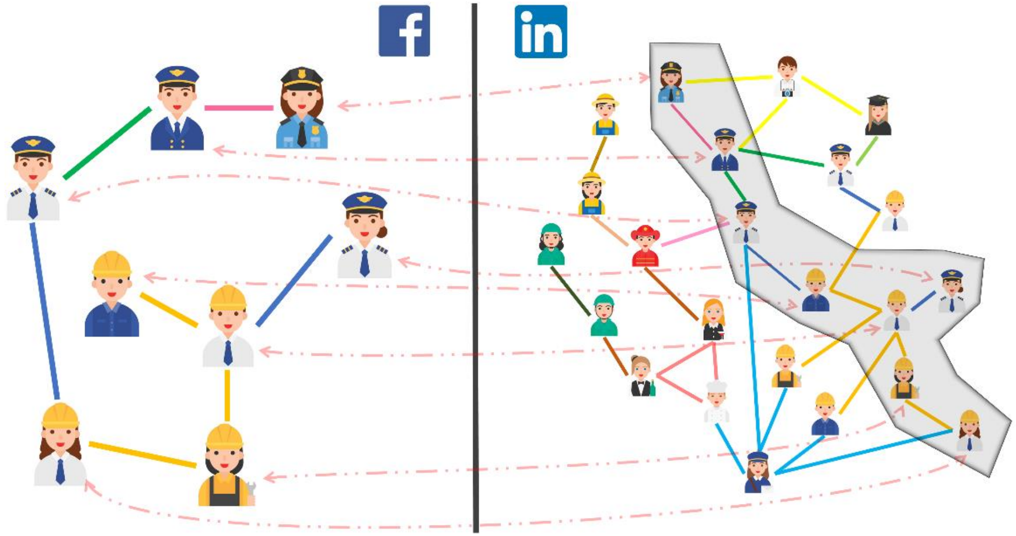


- Our EvePPR in the fully dynamic setting allows graph topology, node feature, edge feature, and prior aligning knowledge from two graphs co-evolve, i.e., $\widetilde{\boldsymbol{W}}^{(t)}$ and $\boldsymbol{h}^{(t)}$

[1] Si Zhang, Hanghang Tong. FINAL: Fast Attributed Network Alignment. KDD 2016

Application Experiment – Graph Alignment

- Real-World Datasets

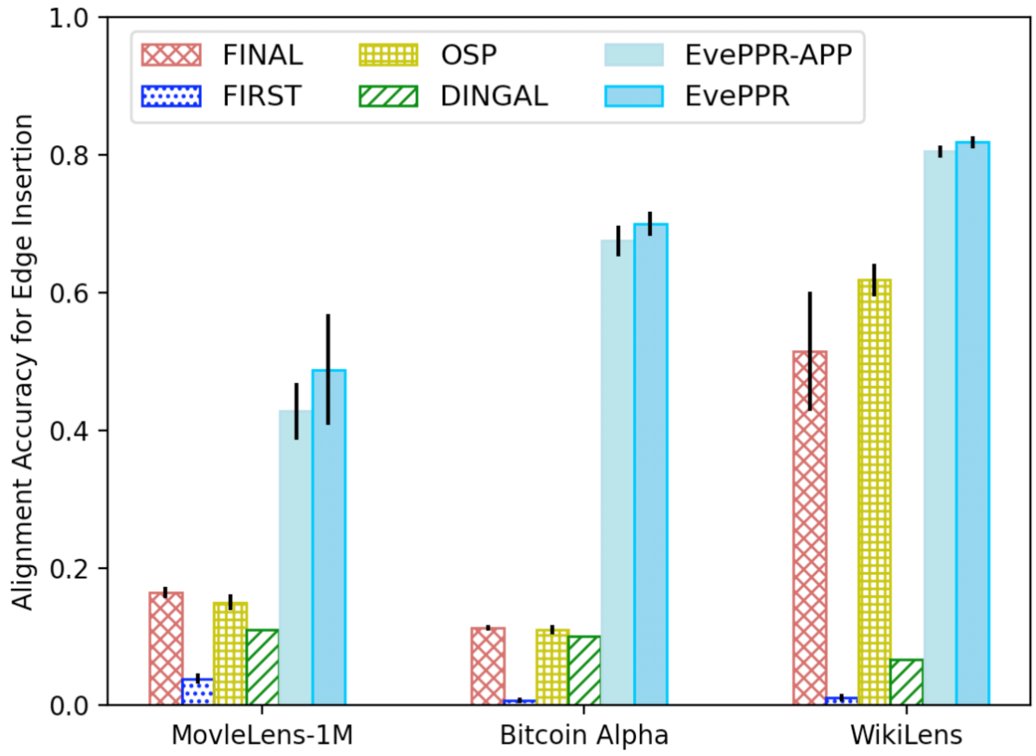


Extracted Subgraphs	Format	|V|	|E|
MovieLens-1M	Bipartite	90	375
Bitcoin Alpha	Unipartite	100	423
WikiLens	Bipartite	150	553

Graphs	Format	|V|	|E|	Time Span
MovieLens-1M	Bipartite	9,746	1,000,209	35 months
Bitcoin Alpha	Unipartite	3,783	24,186	64 months
WikiLens	Bipartite	5,437	26,937	46 months

Application Experiment – Graph Alignment

- Our method achieves highest alignment accuracy



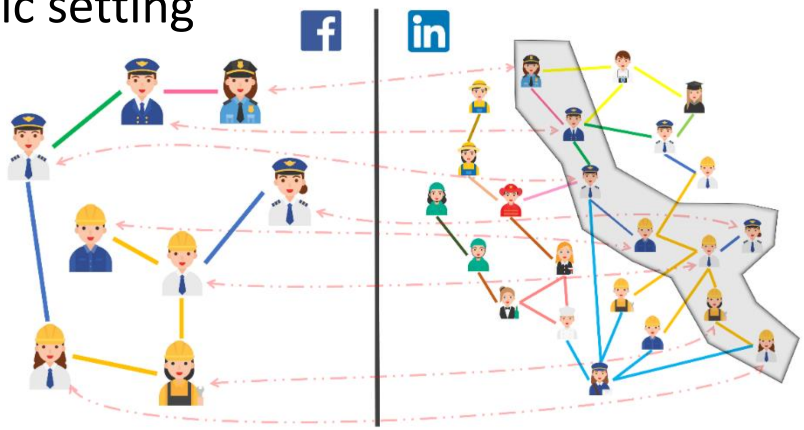
Alignment Accuracy of Different Graph Alignment Algorithms in Different Networks

Roadmap

- Motivation
- Proposed EvePPR Method
- Experiments
- Conclusion

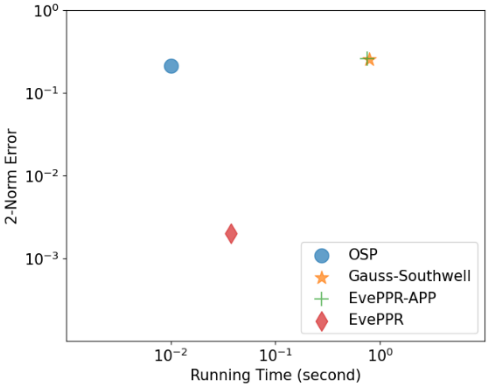
Conclusion

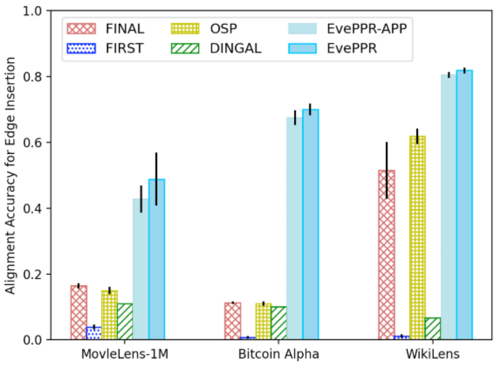
- Problem
  - Tracking PPR solution in the fully dynamic setting



- Algorithm: EvePPR
  - Bounded accuracy
  - Bounded time complexity
  - Approximation solution

- Evaluation
  - Effectiveness in PPR tracking
  - Efficiency in PPR tracking
  - Knowledge graph alignment
  - Ablation studies





Thanks!

Zihao Li * (UIUC)

Dongqi Fu * (UIUC)

Jingrui He (UIUC)

Contact: {zihaoli5, dongqif2, jingrui}@illinois.edu

GitHub: https://github.com/DongqiFu/EvePPR

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN