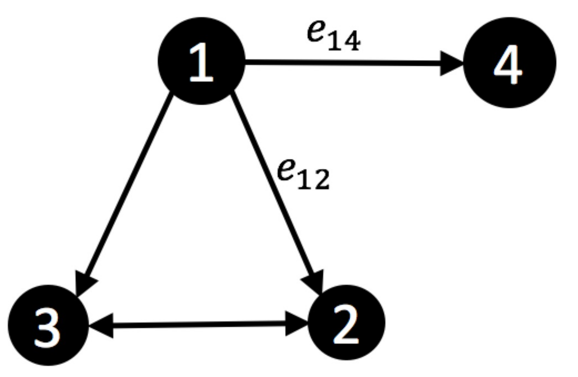
Ranking Users in Social Networks with Higher-Order Structures

Huan Zhao, Xiaogang Xu, Yangqiu Song, Dik Lun Lee, Zhao Chen, Han Gao

Motivation

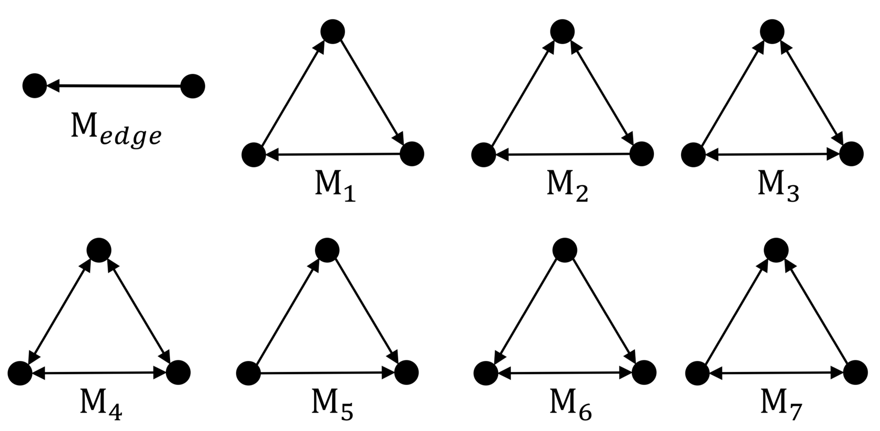
PageRank have been used to measure the authority or the influence of a user in social networks. However, conventional PageRank only makes use of edge-based relations, ignoring higher-order structures captured by motifs.



The weight of e_{12} should be larger than e_{14} due to the existence of the triangle.

Network Motifs

Network motifs are subgraphs or graphlets whose frequency are significant in a graph, which is called higher-order structure. We show in this paper that motif-based and traditional edge-based relations are complementary to each other in computing the authority of a node in complex networks.



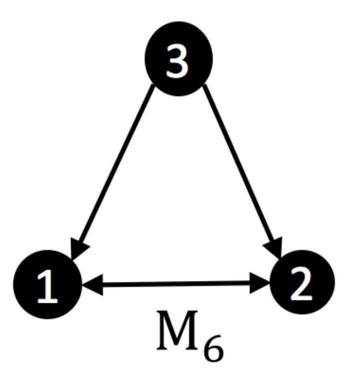
Motif Definition

Definition 1 ***Network Motif.*** *A motif M is defined on k nodes by a tuple $(\mathbf{B}, \mathcal{A})$, where $\mathbf{B}$ is a $k \times k$ binary matrix and $\mathcal{A} \subset \{1, 2, ..., k\}$ is a set of anchor nodes.*

Definition 2 ***Motif Set.*** *The motif set in an unweighted directed graph $\mathcal{G}$ with an adjacency matrix $\mathbf{W}$, denoted as $\mathcal{M}(\mathbf{B}, \mathcal{A})$, is defined by*

$$\mathcal{M}(\mathbf{B}, \mathcal{A}) = \{(set(\mathbf{v}), set(\chi_{\mathcal{A}}(\mathbf{v})))|\mathbf{v} \in V^k, \\ v_1, ..., v_k, distinct, \mathbf{W_v} = \mathbf{B}\}.$$

where $\chi_{\mathcal{A}}$ is a selection function that takes the subset of a k-tuple indexed by $\mathcal{A}$, and $set(\cdot)$ is an operator that transforms an ordered tuple to an unordered set, $set((v_1, v_2, ..., v_k)) = \{v_1, v_2, ..., v_k\}$. $\mathbf{v}$ is an ordered vector representing the k nodes, and $\mathbf{W_v}$ is the $k \times k$ adjacency matrix of the subgraph induced by $\mathbf{v}$.



M_6

$\mathcal{A} = \{1, 2, 3\}$

$$\begin{array}{cccc} & 1 & 2 & 3 \\ 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 1 & 1 & 0 \end{array}$$

B =

Given a motif $\mathcal{M}$, the definition of the **motif-based adjacency matrix** is defined by:

$$(\mathbf{W}_M)_{ij} = \sum_{(\mathbf{v}, \chi_{\mathcal{A}}(\mathbf{v})) \in \mathcal{M}} \mathbf{1}(\{i, j\} \subset \chi_{\mathcal{A}}(\mathbf{v})),$$

where $i \neq j$, and $\mathbf{1}(s)$ is the truth-value indicator function, i.e., $\mathbf{1}(s) = 1$ if the statement s is true and 0 otherwise. Note that the weight is added to $(\mathbf{W}_M)_{ij}$ only if i and j appear in the anchor set.

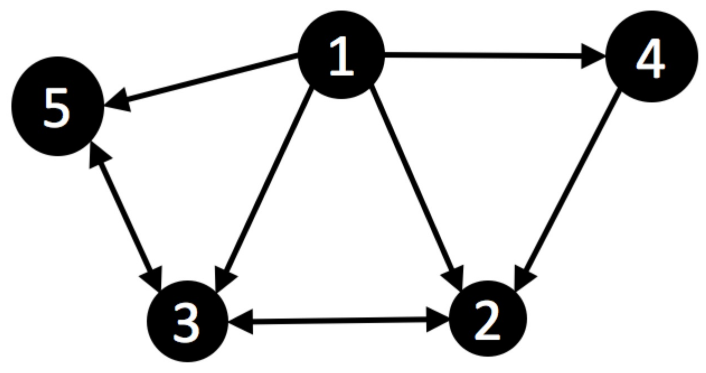
Motif-based Adjacency Matrix Computation

For 3-node motifs, the following equations can be used to compute the motif-based adjacency matrix. Let $\mathbf{W}$ be the adjacency matrix, then $\mathbf{B} = \mathbf{W} \odot \mathbf{W}^T$ and $\mathbf{U} = \mathbf{W} - \mathbf{B}$, representing **bidirectional** and **unidirectional** relations. $\odot$ means elementwise product.

Table 1: Computation of motif-based adjacency matrices for M_1 to M_7.

Motif	Matrix Computation	$\mathbf{W}_{M_i} =$
M_1	$\mathbf{C} = (\mathbf{U} \cdot \mathbf{U}) \odot \mathbf{U}^T$	$\mathbf{C} + \mathbf{C}^T$
M_2	$\mathbf{C} = (\mathbf{B} \cdot \mathbf{U}) \odot \mathbf{U}^T + (\mathbf{U} \cdot \mathbf{B}) \odot \mathbf{U}^T + (\mathbf{U} \cdot \mathbf{U}) \odot \mathbf{B}$	$\mathbf{C} + \mathbf{C}^T$
M_3	$\mathbf{C} = (\mathbf{B} \cdot \mathbf{B}) \odot \mathbf{U} + (\mathbf{B} \cdot \mathbf{U}) \odot \mathbf{B} + (\mathbf{U} \cdot \mathbf{B}) \odot \mathbf{B}$	$\mathbf{C} + \mathbf{C}^T$
M_4	$\mathbf{C} = (\mathbf{B} \cdot \mathbf{B}) \odot \mathbf{B}$	$\mathbf{C}$
M_5	$\mathbf{C} = (\mathbf{U} \cdot \mathbf{U}) \odot \mathbf{U} + (\mathbf{U} \cdot \mathbf{U}^T) \odot \mathbf{U} + (\mathbf{U}^T \cdot \mathbf{U}) \odot \mathbf{U}$	$\mathbf{C} + \mathbf{C}^T$
M_6	$\mathbf{C} = (\mathbf{U} \cdot \mathbf{B}) \odot \mathbf{U} + (\mathbf{B} \cdot \mathbf{U}^T) \odot \mathbf{U}^T + (\mathbf{U}^T \cdot \mathbf{U}) \odot \mathbf{B}$	$\mathbf{C}$
M_7	$\mathbf{C} = (\mathbf{U}^T \cdot \mathbf{B}) \odot \mathbf{U}^T + (\mathbf{B} \cdot \mathbf{U}) \odot \mathbf{U} + (\mathbf{U} \cdot \mathbf{U}^T) \odot \mathbf{B}$	$\mathbf{C}$

An example is given to show the motif-based (M_6) adjacency matrix for the graph in the leftside. Each entry means the frequency of two nodes occur in M_6.



	1	2	3	4	5
1	0	1	2	0	1
2	1	0	1	0	0
3	2	1	0	0	1
4	0	0	0	0	0
5	1	0	1	0	0

Higher-order PageRank

Traditionally, the PageRank over the graph is defined as follows:

$$\mathbf{x} = d\mathbf{H}^T\mathbf{x} + \frac{1-d}{N}\mathbf{e},$$

where $\mathbf{x} \in \mathbb{R}^N$ and $\mathbf{x}_i$ is the PageRank value of the i-th node in $\mathcal{G}$, $\mathbf{e} \in \mathbb{R}^N$ is a vector with every entry equal to 1, and $d \in (0, 1)$ is a damping factor.By integrating edge-based and motif-based higher relations, we replace $\mathbf{P}$ with the following $\mathbf{H}_{M_k}$.

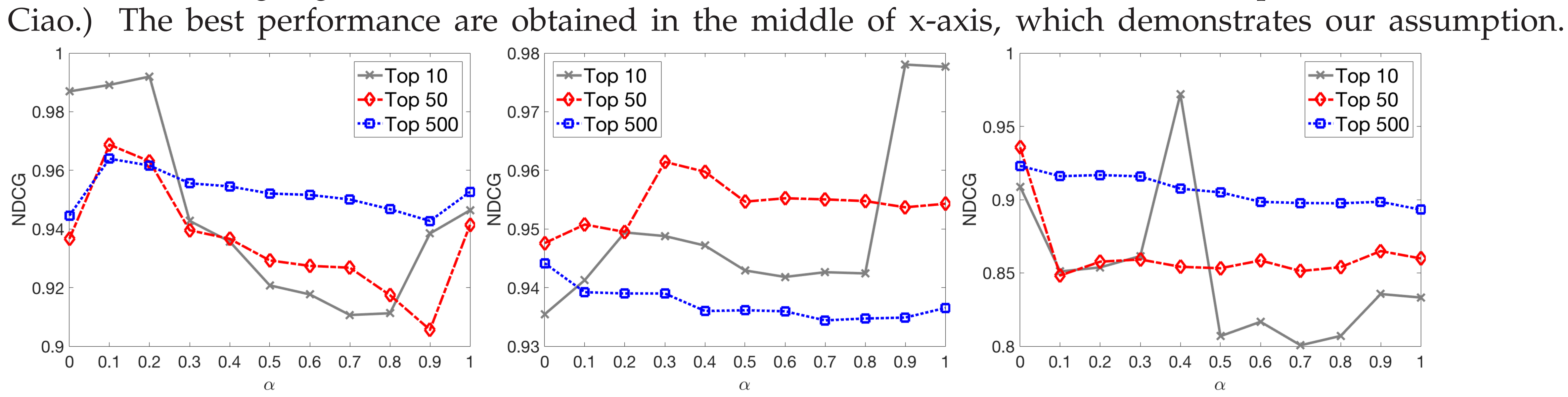
$$\mathbf{H}_{M_k} = \alpha \cdot \mathbf{W} + (1 - \alpha) \cdot \mathbf{W}_{M_k}.$$

Experimental Results

We test our framework on three datasets, and **NDCG** is used as evaluation metric.

	DBLP			Epinions			Ciao		
TopK	10	50	500	10	50	500	10	50	500
IND	0.9879	0.9639	0.9400	0.9476	0.9563	0.9343	0.9218	0.8651	0.9120
BET	0.9796	0.9710	0.9559	0.9566	0.9559	0.9403	0.9421	0.8961	0.8911
CLO	0.9875	0.9614	0.9285	0.9308	0.9346	0.9382	0.9021	0.9225	0.9251
BPR	0.9464	0.9414	0.9527	0.9777	0.9543	0.9365	0.8332	0.8599	0.8932
WPR	0.9154	0.8871	0.9350	0.9777	0.9543	0.9365	0.8332	0.8599	0.8932
M_1	0.9753	0.9590	0.9623	0.9777	**0.9656**	0.9406	0.9802	0.9347	0.9392
M_2	0.9890	0.9424	0.9585	0.9777	0.9581	0.9417	**0.9905**	0.9453	0.9401
M_3	0.9895	0.9508	0.9586	0.9788	0.9568	0.9378	0.9768	0.9576	0.9441
M_4	0.9809	0.9477	0.9528	0.9827	0.9557	0.9395	0.9719	0.9357	0.9401
M_5	0.9877	0.9513	0.9574	0.9777	0.9551	0.9454	0.9792	**0.9792**	0.9401
M_6	0.9634	0.9525	0.9588	**0.9957**	0.9596	0.9382	0.9459	0.9459	**0.9427**
M_7	**0.9920**	**0.9766**	**0.9640**	0.9780	0.9614	**0.9442**	0.9514	0.9500	0.9418

The following figures show the influence of α. (M_4 on DBLP, M_7 on Epinion, and M_7 on Ciao.) The best performance are obtained in the middle of x-axis, which demonstrates our assumption.



Collaboration work of HKUST, ZJU and Wechat, Tencent

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

ZHEJIANG UNIVERSITY 1897

Tencent 腾讯