

Computing Lens for Exploring the Historical People's Social Network

Junjie Huang

University of Chinese Academy of Sciences
Beijing, China
huangjunjie17@mails.ucas.ac.cn

Tiejian Luo

University of Chinese Academy of Sciences
Beijing, China
tjluo@ucas.ac.cn

Abstract—A typical social research topic is to figure out the influential people's relationship and its weights. It is very tedious for social scientists to solve those problems by studying massive literature. Digital humanities bring a new way to a social subject. In this paper, we propose a framework for social scientists to find out ancient figures' power and their camp. The core of our framework consists of signed graph model and novel group partition algorithm. We validate and verify our solution by China Biographical Database Project (CBDB) dataset. The analytic results on a case study demonstrate the effectiveness of our framework, which gets information that consists with the literature's facts and social scientists' viewpoints.

Keywords—Digital Humanities, Information System, China Biographical Database Project (CBDB), CLHPSoNet, Social Network Analysis, Group Partition

I. INTRODUCTION

As we all know, microscopes can help bioscientists better observe the internal structure of cells. Our framework hopes to help researchers not only focus on the relationships of specific research characters but recognize the role of these characters in the entire large social network, which we hope it works like computing lens for some social problems.

The traditional method to study historical persons' relationship is mainly through reading literature, accessing a large number of historical documents and demonstrating the relationships. And it is mostly qualitative analysis. However, this method requires too much professional knowledge. For example, for someone who does not understand the ancient literature, it is impossible to dig out information in historical documents. In addition, the traditional method used to discuss the relationship between two people was limited to the direct relationship between historical people but did not consider the influence of these historical people's friends and even the entire social network on him. Therefore, this paper proposes a new social network research framework (CLHPSoNet) to help people understand the influence of social networks of historical people on the social relationships of historical people. This method can be studied quantitatively without the need for rich background knowledge. The comparison between our framework and traditional methods are shown in Table I. Our method can get some answers about the problems like who is the central figures? or what's the relationship between

some people?¹. These questions are not only concerned by researchers, but history enthusiasts also have great interests.

TABLE I
COMPARISON OF OUR FRAMEWORK AND TRADITIONAL METHODS

	Traditional methods	Our framework
Data source	Historical Documents	Database
Research methods	Qualitative	Quantitative
Relationship between research objects	Direct Relationship	Network
Expert knowledge requirements	Professional	Normal
Results display	Papers, Tables	Web UI

The major contributions of this paper are as follows:

- A novel research framework for exploring the social relationship of historical people is presented.
- We propose graph partitioning algorithms to solve relevant domain problems.
- Based on the proposed model and algorithm, we have built an application to help people analyze and understand the social relationships of the ancients.

The rest of the paper is organized as follows. We introduce related work in Section 2. We introduce our modeling in Section 3 and system designing in Section 4. An Application is built in Section 5 to show our framework work well in some domain problems. We conclude the paper and point out future directions in Section 6.

II. RELATED WORK

A. Research on Historical Figures' Relationship

The study of social relationships between historical people plays an important role in the study of history. In addition to a richer account of the life experiences of the characters, it also reflects the historical background of the historical people to some extent. It attracted a lot of research attention.

For example, Liu [1] lists historical document records to discuss the Su Shi's² (a song dynasty poet) association with Wang Anshi³ (a song dynasty poet), from life experience, political, literature, and other aspects. Guo [2] use a lot of historical documents to discuss the friendship between Ouyang

¹<https://www.zhihu.com/question/20589740>

²https://en.wikipedia.org/wiki/Su_Shi

³https://en.wikipedia.org/wiki/Wang_Anshi

Xiu and Wang Anshi. With the development of digital humanities, some researchers input the historical document records to computers to count, compute and visualize the relationship of historical people. Yu et al [3] have some statistics on the friendship between Su Shi and Wang Gong(a song dynasty poet). Zhu [4] proposed a ‘Chro-Ring’ approach to visualize the history of poets. However, most of the research in this area is qualitative, straightforward and historical-based.

B. Research on China Biographical Database Project (CBDB)

The China Biographical Database(CBDB)⁴ [5] is a freely accessible relational database with biographical information about approximately 417,000 individuals, primarily from the 7th through 19th centuries. It is developed by Harvard University, Institute of History and Philology of Academia Sinica and Peking University. Recent CBDB’s version released in April 2017, it has been continuously updated. The data is meant to be useful for statistical, social network, and spatial analysis as well as serving as a kind of biographical reference.

Based on CBDB dataset, a lot of new applications and system are proposed. Liu et al [6], [7] use the database to compare the poetry of different dynasties, it shows the accessibility of increasingly larger datasets strengthens researchers’ research potential. Guo [8] proposed a powerful tool for studying genealogical records. Sie et al [9] developed a text retrieval and mining system for Taiwanese historical people. There exist lots of valuable information waiting for researchers to mine. Our paper uses the dataset to implement our framework. Of course, other historical files or database can be added to make it more complete.

C. Signed Graphs

Signed social networks are such social networks in which social ties can have two signs: positive and negative [10]. It was first addressed by social balance theory, which originated in cognitive theory from the 1950s [11]. For all signed triad, triads with an odd number of positive edges are defined as balanced. It was proven that if a connected network is balanced, it can be split into two opposing groups and vice versa [12]. In reality, however, social networks are rarely, structural balance theory has been developed extensively, e.g. weak structural balance, k-balance, and other different criteria to quantify and evaluate balance in a signed social network [10], [13], [14]. Researchers defined clustering problems on signed graphs can be used as a criteria to measure the degree of balance in social networks [15]. Then it becomes a cluster problem and it’s NP-hard, a lot of heuristic methods are proposed to solve the clustering problems [13], [15]–[17]. Recently, with the development of network representation learning, researchers begin to use machine learning methods to learn low-dimensional vector representations for nodes of a given network. And it has been proven to be useful in many tasks of network analysis such as link prediction [18], node classification [19], and visualization [20]. However,

network embedding has various methods (e.g. DeepWalk [21], node2vec [19], t-SNE [20])been proposed for unsigned network embedding. How to do network embedding for signed graph attracted the interest of researchers. A framework like SiNE [18] is the latest attempt, it inspires the use of embedding methods for signed graph.

III. MODELING

A. Data Preparation

As we said in related work, we use the CBDB’s API⁵ to collect data. We cleaned the collected data. According to the following rules, we divided the dynasty where people belong to:

- 1) The dynasty is marked
- 2) The year of birth is within the year of the dynasty
- 3) The year of death is within the year of the dynasty

We use NetworkX⁶ to do some statistics on the networks, which the number of nodes $|V|$, the number of edges $|E|$, average clustering coefficient [22] and average path length. The results show in the Table II. Besides, Figure 1 shows the degree distributions in networks. It shows that the social network in Tang dynasty is not complete. Other networks show a power law phenomenon and small world phenomenon. For our following works, we use Song dynasty social networks.

TABLE II
NETWORK INFORMATION OF DIFFERENT DYNASTY

	$ V $	$ E $	Average Clustering Coefficient	Average Path Length
Tang(618, 907)	365	286	0.016	1.60
Song(960, 1279)	17,114	30,330	0.121	4.08
Yuan(1271, 1368)	6,424	11,864	0.150	4.00
Ming(1368, 1644)	8,350	14,609	0.070	4.65
Qing(1636, 1912)	3,128	3,059	0.021	7.71

In order to get an intuitive impression, we use Gephi⁷ to visualize the network to get Figure 2 and Figure 3. It shows that most of the important people are contained in a big connected component, which shows a small world phenomenon. And if you focus on someone e.g.Wan Anshi, you will find that even someone e.g.Hui Qin⁸ whom Wan Anshi don’t know can be reached by little steps, such as Sun Di.

B. Domain Problem

After our data preparation, we have a whole undirected weighted graph $G = (V, E, w)$. w_{ij} is the number of records between v_i and v_j . And we want to solve some problems people concerned. For example, how is the relationship between Eight Great Prose Masters of the Tang and Song⁹. Or who is the central figures in the New Policies in Song dynasty¹⁰?

⁵<https://projects.iq.harvard.edu/cbdb/cbdb-api>

⁶<https://networkx.github.io/>

⁷<https://gephi.org/>

⁸https://en.wikipedia.org/wiki/Qin_Hui

⁹https://en.wikipedia.org/wiki/Eight_Masters_of_the_Tang_and_Song

¹⁰[https://en.wikipedia.org/wiki/New_Policies_\(Song_dynasty\)](https://en.wikipedia.org/wiki/New_Policies_(Song_dynasty))

⁴<https://projects.iq.harvard.edu/cbdb>

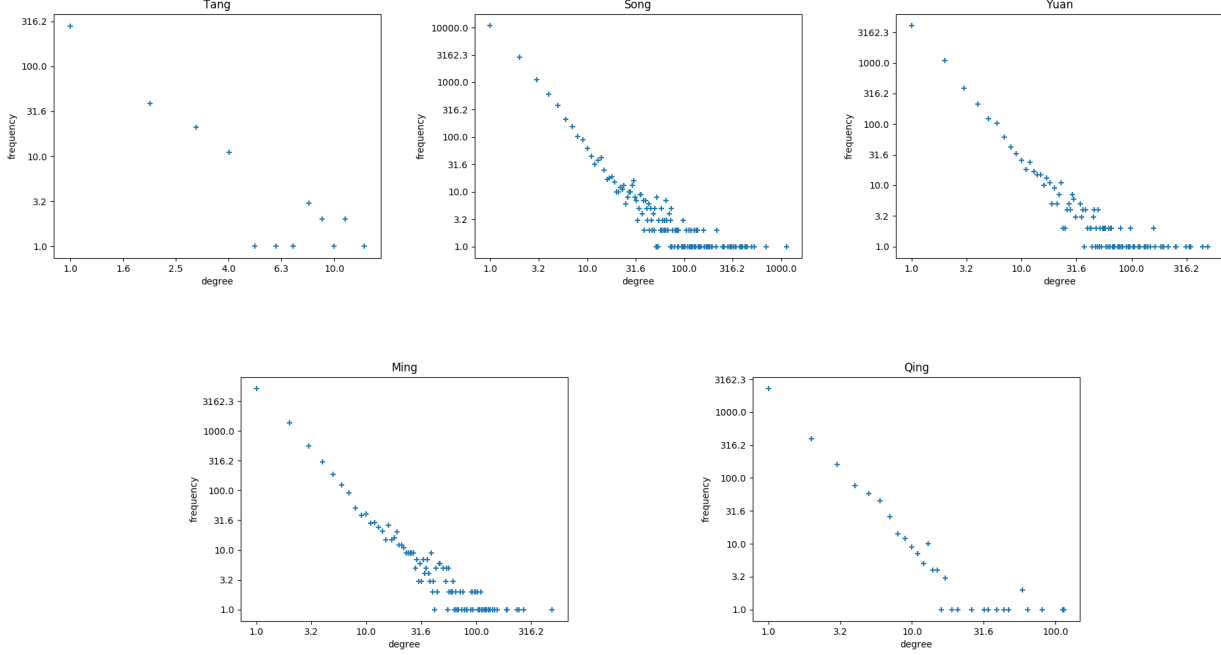


Fig. 1. Logarithmic plot of the degree distribution showing that the degree distribution in the networks follows a power law.

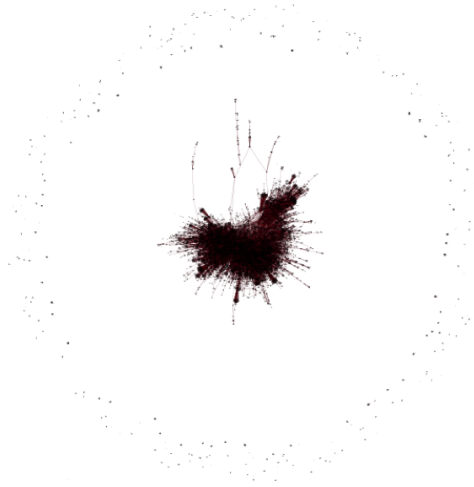


Fig. 2. The social network in Song dynasty is visualized, it shows that the network consists of a large core component and other marginal components.

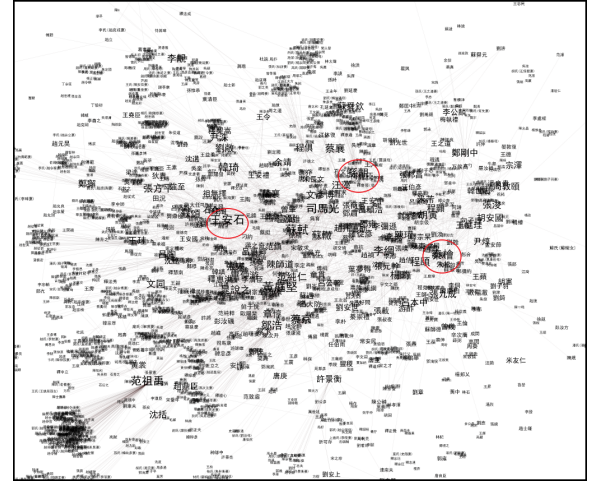


Fig. 3. Wang Anshi (1021-1086) as the focus, the relationship between people born in 1000-1100. It shows that Wan Anshi can get a relationship with somebody like Qing Hui (1091-1155)

To answer the questions about the relationship for some people, we need do some data modelings.

C. Data Model

1) *signed networks*: As we mentioned in the related work about the signed graph. In order to answer some domain problems, the network needs to be a signed graph to make the relationship more semantics, which means to make the relationship friendly or unfriendly. We assigns a sign to each edge in E , then $G = (V, E, w, s)$ is achieved. A total 445

kinds of relationship is manually annotated. Some signed rules(Top 10) showed in the Table III. It shows that the number of positive relationships is much more than the number of negative relationships.

2) *Top and Central People*: Centrality is a common way to quantify the importance of the vertices in a network. It has a lot of different declinations are used in social network analysis to identify and highlight specific nodes [23]. In order to find the important people in our network, we use

TABLE III
THE STATISTICS OF TOP10 SIGNED RULES

Positive			Negative			Neutral		
Relationship(Chinese)	Relationship	Counts	Relationship(Chinese)	Relationship	Counts	Relationship(Chinese)	Relationship	Counts
为Y作墓志铭	Make epitaph for Y	10189	弹劾	impeach	400	被致书由Y	The book was written by Y	3624
墓志铭由Y所作	Epitaph made by Y	8502	被Y弹劾	Invoked by Y	385	致书Y	mail to Y	3302
书序由Y所作	Book sequence made by Y	4774	遭到Y的反对/攻讦	Y's opposition/attack	348	同僚	A collegae	538
为Y所著书作序	Preface to Y's book	4323	反对/攻讦	Opposition/attack	334	未详	Unknown	353
书跋由Y所作	Book made by Y	4111	反对/不支持Y的政策	Oppose/Do not support Y's policy	297	(暂时保留, 待删除: 吏部供职)	(Temporarily reserved pending deletion: Appointment)	331
祭文由Y所作	The memorial was made by Y	3975	其政策被Y反对/不支持	Its policy is opposed by Y/not supported	281	上司为Y	Boss is Y	213
为Y所著书作跋	Writing for Y's book	3868	遭Y排挤	Excluded by Y	199	为Y之考官	The examiner for Y	167
为Y作祭文	Make a memorial for Y	3667	排挤	Exclude	196	奏录Y之文	Record the story of Y	20
为Y作临赠言(送诗、序)	Make a parting speech for Y (Farewell poetry)	3623	得罪Y	Offend Y	177	同场屋/同应举	Same-room housing	20
临得到Y所作赠言(送诗、序)	Make a goodbye to Y's speech (Farewell poems, prefaces)	3520	不合	Not fit	152	以宦官事Y	With official business Y	18

the **Degree Centrality, Betweenness Centrality, Closeness Centrality and Eigenvector Centrality** [24]. In Table IV, Top 15 degree centrality figures are listed, which is order by **Degree Centrality**. It's easy to find that most high degree centrality people are famous in the history.

TABLE IV
THE CENTRALITY OF TOP 15 PEOPLE ORDER BY DEGREE CENTRALITY

EngName	PersonId	Degree Centrality	Betweenness Centrality	Closeness Centrality	Eigenvector Centrality
Zhu Xi	3257	0.066	0.170	0.399	0.400
Unknown	9999	0.041	0.065	0.341	0.191
Zhou Bida	7197	0.031	0.061	0.359	0.203
Su Shi	3767	0.029	0.083	0.380	0.179
Wu Cheng	10084	0.026	0.055	0.345	0.072
Liu Kezhuang	3595	0.026	0.055	0.342	0.105
Wang Anshi	1762	0.025	0.054	0.361	0.138
Wei Liaoweng	4001	0.025	0.054	0.356	0.131
Lou Yue	3624	0.023	0.042	0.348	0.139
Huang Tingjian	7111	0.023	0.049	0.363	0.132
Ouyang Xiu	1384	0.022	0.046	0.362	0.135
Lv Zuqian	7055	0.020	0.027	0.334	0.111
Yu Ji(2)	10801	0.020	0.047	0.346	0.073
Wang Yun(5)	28617	0.020	0.022	0.321	0.048

Degree Centrality defines the number of links that have a node. For a graph, $G = (V, E)$, the degree centrality for vertex v is:

$$C_d(v) = \frac{deg(v)}{|V| - 1}$$

$deg(v)$ is the degree of v . The degree centrality values are normalized by dividing by the maximum possible degree $|V| - 1$ in a simple graph. It shows that people who know many other people are actually important and powerful. *e.g.* Anshi Wang who once served as prime minister and implemented reforms New Policies¹¹. This reform involves many aspects, such as finance, military, education an so on. As the reformer, Wang Anshi were in the political center and contacted with many people. What's more, in the network, the distribution of 'long tail' of this measure exists like many other real-world network [23], which means a lot of people have a low degree centrality.

Betweenness Centrality is a centrality measure of a vertex within a graph. Vertices that occur on many shortest paths

between other vertices have higher betweenness than those that do not. The betweenness computed as follows:

$$C_b(v) = \sum_{s,t \in V} \frac{\sigma(s, t|v)}{\sigma(s, t)}$$

where $\sigma(s, t)$ the number of shortest (s, t) -paths, $\sigma(s, t|v)$ is the number of those paths passing through some node v other than s, t . It measures the person's ability to act as a bridge between other people. In our networks, some people with high degree centrality may not have high betweenness centrality like Lv Zuqian.

Closeness Centrality is the reciprocal of the average shortest path distance to over all other reachable nodes. The result is "a ratio of the fraction of actors in the group who are reachable, to the average distance" from the reachable actors. It defines as follows:

$$C_c v = \frac{n - 1}{|V| - 1} \frac{1}{\sum_{u=1}^{n-1} d(v, u)}$$

where $d(v, u)$ is the shortest-path distance between v and u , and n is the number of nodes that can reach u . It can measure someone's independence, influence and the ability of information transmission.

Eigenvector Centrality computes the centrality for a node based on the centrality of its neighbors.

$$Ax = \lambda x$$

where A is the adjacency matrix of the graph G with eigenvalue λ . The power iteration method is used to compute the eigenvector, the default is 100 iterations. Eigenvector centrality assigned to the vertices according to the score their neighbors received. It takes the number of people and the power of the people someone knows into considerations.

3) **Subgraph Extraction**: In order to measure the social balance, we proposed an Algorithm 1. If $d = 0$, the graph will be the direct relationship between the people we care about. Algorithm 1 can extract the subgraph centered on the seed node, which can better reflect the relationship between

¹¹[https://en.wikipedia.org/wiki/New_Policies_\(Song_dynasty\)](https://en.wikipedia.org/wiki/New_Policies_(Song_dynasty))

Algorithm 1 Extract Subgraph

Input: seed people $V_{seed} = \{v_1, v_2, \dots, v_n\}$, depth d , graph $G = \{V, E, w, s\}$
Output: subgraph G_{sub}
Initialisation : $V_{sub} = V_{seed}, k = d, currentSet = V_{seed}$
1: **while** $k > 0$ **do**
2: $nextNodeSet = \emptyset, k = k - 1$
3: **for** n in $currentSet$ **do**
4: $neighborSet = getNeighbors(n)$
5: **for** n_2 in $neighborSet$ **do**
6: **if** $n_2 \notin V_{sub}$ **then**
7: $nextNodeSet = nextNodeSet \cup \{n_2\}$
8: **end if**
9: **end for**
10: **end for**
11: $V_{sub} = V_{sub} \cup nextNodeSet$
12: $currentSet = nextNodeSet$
13: **end while**
14: $G_{sub} = getSubGraph(V_{sub})$
15: **return** G_{sub}

seed nodes. Because of small world phenomenon, d should less than 4 most cases and the number of nodes $|V_{sub}|$ in subgraph grow very fast.

IV. SYSTEM DESIGN

A. Problem Definition

In the section, we proposed our framework workflow as Figure 4. It includes signed graph modeling, subgraph extraction, computing and visualize. The final output is consists of three parts: Top and Central People, Direct Relationship and Group Partition.

After modeling, we have a signed subgraph and top central people. Then we want to know how to parted into different groups based on the social relationship? It can be formal defined *CCproblem* as follows [13]:

Problem 1: (CC problem). Let $G = (V, E, s)$ be a signed graph and we be a nonnegative edge weight associated with each edge $e \in E$. The correlation clustering problem is the problem of finding a partition P of V such that the imbalance $I(P)$ is minimized. $I(P)$ is defined as for a partition $P = \{S_1, S_2, \dots, S_l\}$:

$$I(P) = \sum_{1 \leq i \leq l} \omega^-(S_i, S_i) + \sum_{1 \leq i < j \leq l} \omega^+(S_i, S_j)$$

where $\omega^+(S_i, S_j) = \sum_{e \in E^+ \cap E[S_i, S_j]} w_e$ and $\omega^-(S_i, S_j) = \sum_{e \in E^- \cap E[S_i, S_j]} w_e$

As we said in the related work, the CC problem is NP-hard [15] and there are some algorithms can give the solutions.

B. Algorithms

1) *Graph Partition:* Doreiana [16] proposed a heuristic approach proposed is a simple greedy neighborhood search procedure that assumes a prior knowledge of the number of clusters in the solution. This heuristic is implemented

in software Pajek¹². It can be a solution algorithm for our problems.

2) *Community Detection:* Community Detection is developed by the concept known as *modularity* [25]. While origin modularity approaches take for granted that links are positively valued, Traag et al [17] extend an existing model to negative links. We use it as one solution of our group partition problems.

3) *Network Embedding:* Network Embedding is a new method to model the signed graph. SiNE [18] is a deep learning framework for Signed Network Embedding. It defines a new objective function guided by the extended structural balance theory. The framework has two hidden layers and it will embed the node to d -dimension (In his paper, $d = 20$ achieved the best performs). After the node embedding, it can be used K-means clustering methods to part the nodes to different groups.

Overall, the problem is an open issue to a certain extent. Different datasets, research questions, and algorithms will lead to different results. There is no particularly good silver bullet.

V. APPLICATION

A. Development tools

Based on our framework, we use Python3.5 and JavaScript to build a WebApp for Song Dynasty. The web backend framework is Flask1.0.1 and the web UI framework is Reactjs16.3.2, antd3.4.4 and D3jsv4. Both our source code and demo system are publicly available online¹³.

B. A Case Study for Eight Great Prose Masters of Song

In order to study the relationship between in Eight Great Prose Masters of the Song (Su Shi, Wang Anshi, Ouyang Xiu, Zeng Gong, Su Zhe, Su Xun), we input these people and set $d = 0$, which means just take their relationship into consideration. And we choose Community Detection as our algorithm. We will get a three-part report by our visualize system.

1) *Top Central People:* First of all, we can get the central people. Figure 5 shows that Su Shi, Wang Anshi, Ouyang Xiu all have high central values, indicating that they are important nodes in the network. In the social network, many people know each other through them. This result meets with common sense. These three people's popularity in the understanding of ordinary people is relatively high than other three people.

2) *Direct Relationship:* The second part reports the direct relationship between given people. We describe in detail the positive relationship, negative relationship and final signed relationship between them. Figure 6 shows that most of the relationships between them are positive, and the source of the negative relationship is mainly due to political opinion differences. It is worth pointing out that the relationship between Ouyang Xiu and Wang Anshi is more complicated.

In the early days, Ouyang Xiu appreciated Wang Anshi's talent very much. He had recommended him to the court

¹²<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>

¹³<https://github.com/huangjunjie95/CLHPSoNet>

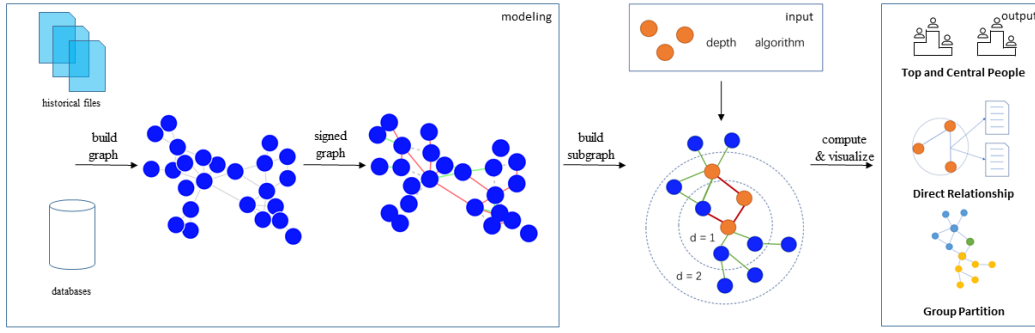


Fig. 4. CLHPSoNet Workflow is showed in the Figure. It includes signed graph modeling, subgraph extraction, computing and visualize. The final output is consists of three parts: Top and Central People, Direct Relationship and Group Partition.

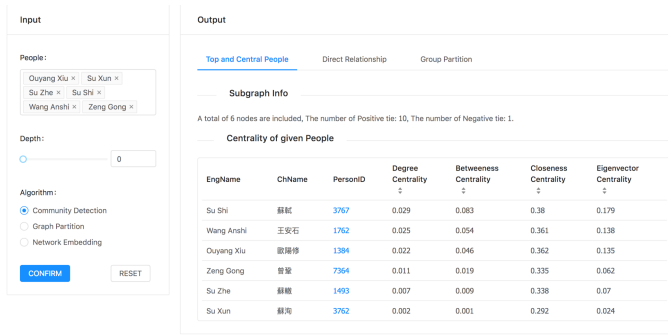


Fig. 5. Input: People(Su Shi , Wang Anshi, Ouyang Xiu, Zeng Gong, Su Zhe , Su Xun), Depth(0), Algorithm(Community Detection), Output: People's Centrality Metrics

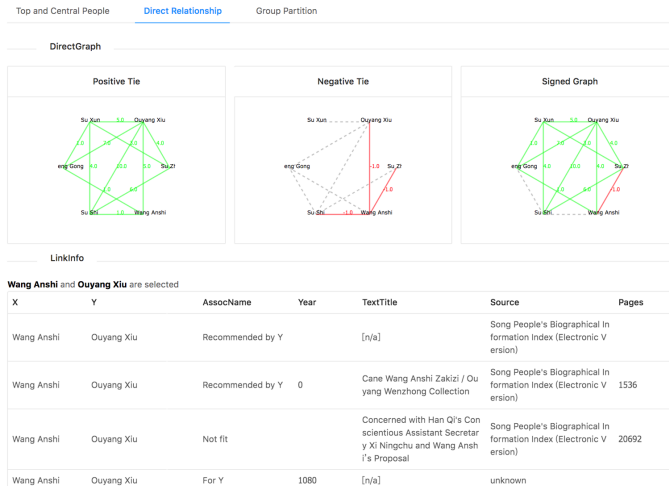


Fig. 6. Focus on the Relationship between Wang Anshi and Ouyang Xiu

several times, and Wang Anshi was also very grateful for the grace of understanding. Later, Wang Anshi presided over the reform process. Ouyang Xiu's political views are different

from Wang. The relevant studies have different opinions on the relationship between the two in the later period [2]. However, according to our analysis, we believe that the relationship between the two is still positive. After the death of Ouyang Xiu, Wang Anshi also made a memorial service for him. With regard to the results of this part, we have found relevant research to support our findings.

3) *Group Partition*: Finally, we obtained the results of the graph partition results in Figure 7. The system divide Su Shi, Su Zhe, Su Xun into a group, and divide Wang Anshi, Ouyang Xiu, Zeng Gong into another group.

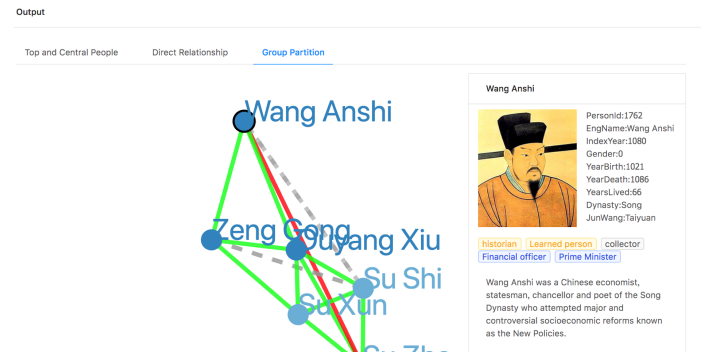


Fig. 7. The Partition between Eight Great Prose Masters of Song

In fact, Su Shi, Su Zhe, Su Xun are a family like Alexandre Dumas family and the Brontë family in history. Su Xun is the father of Su Shi and Su Zhe. Although they are a family, their personality, works, and experience are totally different. But if we focus on the six figures networks, it will be found that Su family all opposed Wang's opinions, which matches the historical facts. Compare to Su family, Ouyang Xiu, Zeng Gong is more friendly to Wang. And when you make depth $d \geq 1$, it will give different results. The partition of this six

people in a large network is not stable, which are not the same like $d = 1$.

VI. CONCLUSION AND FUTURE WORK

In this paper, we present a new research framework for exploring the social relationship of historical people. Based on the proposed the signed social networks model and group partition algorithm, we have built an application to help people analyze and understand the social relationships of the ancients. Via our framework, the social research questions can be transformed into a computing task. People not only can easily use our application to visualize ancient figures' social network but also more clearly understand kinds of literature opinion. In the case study, our system produced some information that meets the facts and social experts' judgment. There are still several works to be done in future. For example, CBDB dataset need to be enriched and improved. How to apply machine learning methods to label social network edges is a research problem. Further study is required for more reasonable and stable for graph partition and structural balance.

ACKNOWLEDGMENT

This paper is inspired by Peking University Digital Humanities Forum, and Prof Luo's Network Science Course.

REFERENCES

- [1] L. Naichang, "Su shi's association with wang anshi," *Journal of North-east Normal University: a Philosophy and Social Science Edition*, no. 3, pp. 45–51, 1981.
- [2] G. Yongxin, "The friendship between ouyang xiu and wang anshi," *Literary heritage*, vol. 5, p. 016, 2001.
- [3] Y. Shihua and Z. Guangyu, "Sharing weal and woe and supporting each other - on friendship between su shi and wang gong," *Journal of Jiangsu University of Science and Technology (Social Science Edition)*, vol. 13, no. 2, pp. 50–58, 2013.
- [4] Y. Zhu, J. Yu, and J. Wu, "Chro-ring: a time-oriented visual approach to represent writer's history," *The Visual Computer*, vol. 32, no. 9, pp. 1133–1149, 2016.
- [5] M. A. Fuller, "The china biographical database user's guide," 2015.
- [6] C.-L. Liu and K.-F. Luo, "Tracking words in chinese poetry of tang and song dynasties with the china biographical database," *arXiv preprint arXiv:1611.06320*, 2016.
- [7] C.-L. Liu, "Flexible computing services for comparisons and analyses of classical chinese poetry," *arXiv preprint arXiv:1709.05729*, 2017.
- [8] X. Guo, "An investigation of the digitization of chinese genealogical records," *Taiwan University of Information Network and Multimedia Institute Dissertation*, pp. 1–40, 2016.
- [9] S.-H. Sie, H.-R. Ke, and S.-B. Chang, "Development of a text retrieval and mining system for taiwanese historical people," in *Pacific Neighborhood Consortium Annual Conference and Joint Meetings (PNC)*, 2017. IEEE, 2017, pp. 56–62.
- [10] J. Leskovec, D. Huttenlocher, and J. Kleinberg, "Signed networks in social media," in *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 2010, pp. 1361–1370.
- [11] F. Heider, "Attitudes and cognitive organization," *The Journal of psychology*, vol. 21, no. 1, pp. 107–112, 1946.
- [12] D. Easley and J. Kleinberg, *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press, 2010.
- [13] M. Levorato, L. Drummond, Y. Frota, and R. Figueiredo, "An ils algorithm to evaluate structural balance in signed social networks," in *Proceedings of the 30th Annual ACM Symposium on Applied Computing*. ACM, 2015, pp. 1117–1122.
- [14] A. Srinivasan, "Local balancing influences global structure in social networks," *Proceedings of the National Academy of Sciences*, vol. 108, no. 5, pp. 1751–1752, 2011.
- [15] P. Doreian and A. Mrvar, "A partitioning approach to structural balance," *Social networks*, vol. 18, no. 2, pp. 149–168, 1996.
- [16] —, "Partitioning signed social networks," *Social Networks*, vol. 31, no. 1, pp. 1–11, 2009.
- [17] V. A. Traag and J. Bruggeman, "Community detection in networks with positive and negative links," *Physical Review E*, vol. 80, no. 3, p. 036115, 2009.
- [18] S. Wang, J. Tang, C. Aggarwal, Y. Chang, and H. Liu, "Signed network embedding in social media," in *Proceedings of the 2017 SIAM International Conference on Data Mining*. SIAM, 2017, pp. 327–335.
- [19] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2016, pp. 855–864.
- [20] L. v. d. Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of machine learning research*, vol. 9, no. Nov, pp. 2579–2605, 2008.
- [21] B. Perozzi, R. Al-Rfou, and S. Skiena, "Deepwalk: Online learning of social representations," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2014, pp. 701–710.
- [22] J. Saramäki, M. Kivelä, J.-P. Onnela, K. Kaski, and J. Kertesz, "Generalizations of the clustering coefficient to weighted complex networks," *Physical Review E*, vol. 75, no. 2, p. 027105, 2007.
- [23] M. Grandjean, "A social network analysis of twitter: Mapping the digital humanities community," *Cogent Arts & Humanities*, vol. 3, no. 1, p. 1171458, 2016.
- [24] S. Wasserman and K. Faust, *Social network analysis: Methods and applications*. Cambridge university press, 1994, vol. 8.
- [25] M. Girvan and M. E. Newman, "Community structure in social and biological networks," *Proceedings of the national academy of sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.