Visualization for Changes in Relationships between Historical Figures in Chronicles

Masahiko ITOH

Institute of Industrial Science
The University of Tokyo, Japan
imash@tkl.iis.u-tokyo.ac.jp

Mina AKAISHI

Faculty of Computer and Information Sciences
Hosei University, Japan
PRESTO, Japan Science and Technology Agency
mina@hosei.ac.jp

Abstract

A huge number of historical documents that have been accumulated for a long time are currently being digitized. However, it is difficult for us to analyze and obtain insight into the past from such documents. This paper proposes an interactive visualization system to extract networks of historical figures from historical data and to show time-varying changes in their relationships. It enables users to explore changes in the structure of the network interactively. Moreover, it extracts characteristics of each relationship, such as hostile or friendly relations, and visualizes them on the network. It enables us to understand changes in past society better by exploring changes in relationships between people.

1 Introduction

A huge number of historical documents that have been accumulated for a long time are currently being digitized, enabling us to search for and access such documents easily and rapidly. However, it is difficult for us to analyze and obtain insight into the past from such documents by simply using statistical techniques. Exploring such data from various viewpoints requires an interactive environment.

In this paper, we propose an interactive visualization framework for exploring changes in relationships between historical figures that are extracted from historical documents. It enables us to support understanding of changes in past society through exploring changes in relationships between people, such as the appearance of groups of people, or whether such groups had friendly or unfriendly relations. The proposed system allows us to check whether visualization results accurately represent already-known knowledge and explore novel phenomenon or patterns that are not already well-known.

Our framework first provides a mechanism for extracting networks of historical figures from historical data and for visualizing time-varying changes in the structure of networks along with a timeline in a 3D space. It also allows users to extract networks of people with multiple aspects such as “battle” or “gift” and to compare the differences and similarities in their evolutions. Users can observe the appearances of groups of people, the important persons in these groups, and changes in the structure of the groups.

It is desirable to grasp not only structural changes but also characteristics of relationships between people, such as friendly or unfriendly relationships. We therefore provide mechanisms for extracting components of relationships and visualizing them by coloring them different colors. This enables users to observe what kind of characteristics a group of people had. It also enables users to observe temporal changes in characteristics of relationships between two people, e.g. the users can find out that two generals changed from being allies in 1561 to opponents in 1564.

The contribution of our system is that it interactively enables users to explore temporal changes in networks of historical figures and in the details of relationships between figures in the networks.

2 Framework for Visualizing Changes in Relationships between Historical Figures

This paper proposes a framework for extracting networks of historical figures and characteristics of relationships between such figures from historical documents and for visualizing and interactively exploring temporal changes in such relationships (Figure 1).

Figure 1 shows an overview of our visualization framework. A historical document database (Figure 1(a)) stores records of events that have attribute values such as names of people, keywords, and dates related to the events. Users can

select records related to a specific viewpoint in which they are interested by specifying ranges of years, lists of people, and/or keywords. The system then generates a network of historical figures from the selected records by extracting co-occurring people in the same events (Figure 1(b)). If two people have a high co-occurrence ratio, they have a connection. Users can extract networks related to specific people such as ODA Nobunaga, who was the best-known feudal lord in Japanese history, or TOKUGAWA Ieyasu, who was the founder and first *shogun* of the Tokugawa *shogunate*, by inputting their names. They can also extract networks related to a battle in the warring states period in Japanese history (called *Sengoku Jidai*) by inputting the range of years related to *Sengoku Jidai* and keywords such as "battle", "attack", and/or "rebellion". Moreover, they can extract and compare networks related to various viewpoints.

Our framework allows users to extract a network of a specific viewpoint with a specified time window, e.g. every year. It enables users to observe sequential changes in the network of people.

The historical document database stores keywords summarizing events. We can extract information about why two or more people have connections by using events and keywords related to them. However, keywords are too diverse for visualizing the reasons for connections. We therefore extract clusters of keywords and assign colors to extracted clusters (Figure 1(c)).

Our framework then visualizes a time-varying network extracted in Figure 1(b) with assigning a person to a node of the network and a connection between people to an edge (Figure 1(d)). It also assigns colors to the edge that has characteristics defined by co-occurring clusters of keywords between two people (Figure 1(d)). This enables users to explore temporal changes in relationships between people.

3 Historical Document Database

We utilize the historical database Dai-Nihon Shiryo Database[1]. The Dai-Nihon Shiryo is a collection of historical documents arranged chronologically and dating from the ninth to the seventeenth centuries. It has been published by the University of Tokyo since 1901. Currently, it is being digitized and compiled into a database. Each record in the database represents one event. It includes the date of the event, a list of people related to the event, a list of these people's official titles, a list of place names related to the event, a list of keywords representing the event, and text as attributes. The current database consists of about 230,000 records.

[1] http://wwwap.hi.u-tokyo.ac.jp/ships/db.html

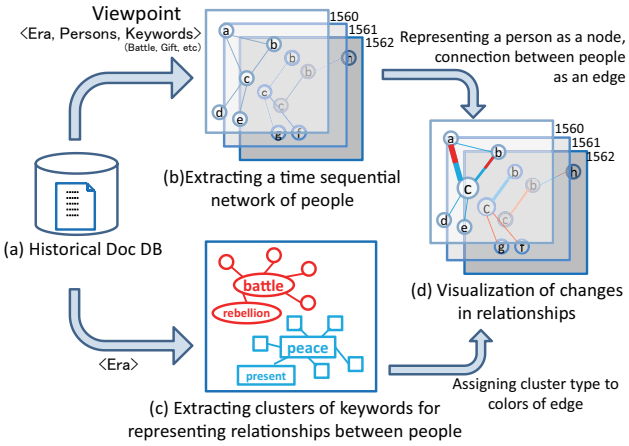


Figure 1. Overview of our visualization framework for changes in relationships between historical figures.

4 Extracting a Network of Historical Figures

Our framework enables us to extract networks of people from a specific viewpoint. We first filter records out from a database by keywords, range of date, and/or persons for specifying the viewpoint and then aggregate filtered records every year. We next extract the networks of people for every year on the basis of co-occurring people in the same event.

More specifically, we define a connection between people depending on person dependency. The following formula defines the person dependency from one person to another in a particular year y, which is an extension of term dependency described by Akaishi et al. [2, 1].

$$pd_y(p, p') = \frac{records_y(p \cap p')}{records_y(p)} \quad (1)$$

$records_y(p \cap p')$ represents the number of records representing events in which persons p and p' both appear in year y. $records_y(p)$ represents the number of records in which only person p appears. If $pd_y(p, p') \geq \alpha$ and $pd_y(p', p) \geq \alpha$, then we define a bidirectional edge between people p and p'. If $pd_y(p, p') \geq \delta$ and $pd_y(p', p) \geq \mu$, then we define a edge from person p to person p'[2].

Size and color of a node represent the importance of a person. The size of node is defined by an attraction power of person p described by $\sum_{p'' \in p} pd_y(p'', p)$, which means the summation of other people's person dependency on person p. The color of the node is defined by the ratio of the number of inlinks to outlinks of the node. We use the following colors as default colors that can be modified by a visualization system mentioned at Section 6.1.

[2] We use $\alpha = 1$ and $\delta = \mu = 0.8$ in this paper.

- $\#In\_Link > \#Out\_Link$: dark purple
- $\#Out\_Link > \#In\_Link$: light purple
- $\#In\_Link = \#Out\_Link$: gray

The length of an edge represents the strength of a relationship between two people. It is defined by the difference in person dependencies $pd_y(p, p') - pd_y(p', p)$. If person dependency is strong, the edge shortens. If one person one-sidedly depends on another, the edge lengthens.

5 Extracting Characteristics of Relationships between Historical Figures

We first extract keywords in events related to two people to represent characteristics of their relationship. We then add such keywords as labels to their relationships and attach a color to every label. This enables us to investigate characteristics of relationships between two people.

The number of keywords is 3034 in only 20 years (from 1560 to 1580). It is therefore difficult for us to distinguish colors. Moreover, it is difficult for us to understand relationships between keywords, e.g., most keywords are related to battles, and the rest are related to peace.

To solve the problems mentioned above, we first extract clusters of keywords on the basis of co-occurring keywords. We then extract common clusters of keywords related to two people for every year. Such clusters of keywords represent components of the relationship between people.

5.1 Extracting Clusters of Keywords

We first extract a network of keywords from 1560 to 1580 by replacing a person with a keyword in Formula 1 in Section 4[3].

We extract a set of nodes (keywords) that are directly connected by edges as a cluster. We categorize 3034 keywords in 70 clusters. Table 1 shows examples of clusters of keywords. We use the keyword that has highest frequency as a cluster label.

5.2 Extracting Clusters for each Connection between Historical Figures

We next extract clusters and their frequencies every year related to each person dependency (edge) that is extracted in Section 4.

More specifically, we collect records, in which two people related to the edge appear in the same time. We then obtain keywords appearing in collected records and obtain clusters and frequencies for the edge by using the result of keyword clustering described in Section 5.1. Table 2 shows an example of extracted clusters and frequencies for the edge between ODA Nobunaga and TAKEDA Katsuyori[4] every year.

Cluster label	Keywords
Battle	Ikko, Shin Buddhism, Follower, Battle, Rebellion, Crucifixion, Naval battle, Navy,...
Gift	Whale meat, Gift, Persimmon, Wild goose, Garden lantern, Present,...
Allowance	Loyalty, Allowance, Supply,...
Grant	Incense, Grant,...
Conferment	Commission, Conferment
Succession	Successor, Succession, Head of family
...	...

Table 1. Example of keyword clusters.

Year	Cluster name : frequency
1574	Battle: 2, Allowance: 1, Unknown: 1
1575	Battle: 5, Unknown: 1
1576	Peace: 1

Table 2. Clusters and frequencies for the edge between ODA Nobunaga and TAKEDA Katsuyori every year.

6 Visualization of Changes in Relationships between Historical Figures

We propose a visualization system that interactively enables us to explore time varying changes in the structure of person dependency networks and changes in characteristics of relationships among people.

6.1 Visualizing Changes in Structure

We visualize networks of people, which are extracted in Section 4, by using TimeSlice [6] as a platform for visualizing temporal changes in the structure of a network in a 3D space.

A TimeSlice is a 2D plane for visualizing a snapshot of a network with a specified timing (Figure 2). The TimeSlice only displays the network belonging to a selected year. It can be dragged along the timeline, which is represented by

[3] We use $\alpha = \delta = \mu = 0.8$ in this case.

[4] TAKEDA Katsuyori was a famous rival lord of Nobunaga

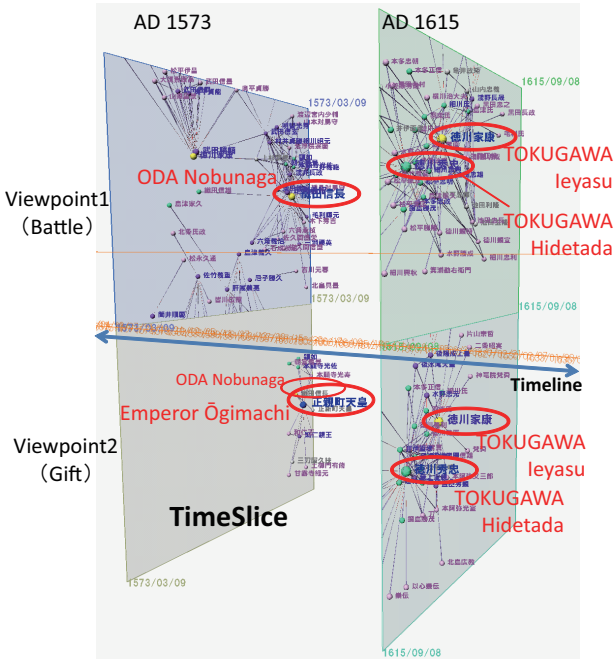


Figure 2. Comparing networks of people related to "Battle" and "Gift" in 1573 and 1615.

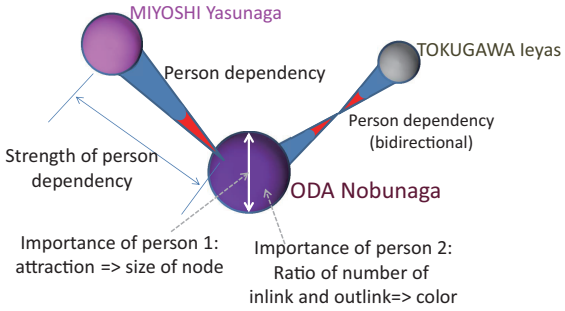


Figure 3. Visualization of connection between people.

the third dimension. Users can seamlessly change the positions of the TimeSlice along the timeline to change the year being visualized. Such manipulation generates an animated temporal change in the structure of a network.

Each node in the network on the TimeSlice represents a person, and each edge represents a person dependency. Figure 3 shows the directed edge from MIYOSHI Yasunaga[5] to ODA Nobunaga, and the bidirectional edge between TOKUGAWA Ieyasu and ODA Nobunaga. The size of nodes represents attraction power of people, and the colors of nodes represent the ratio of the number of inlinks to outlinks of the nodes described in Section 4. Suitable length of edges is defined by the difference in person dependencies as described in Section 4. However, the layout algorithm defines the final length of edges.

We adopt automatic and dynamic graph layout algorithms based on a force-directed model [4] to visualize the structures of networks. Users can select and drag nodes to change their positions for interactively checking details on linkage between nodes.

Moreover, a user can incrementally add new TimeSlices on the timeline side by side to compare different years (in Figure 2, 1573 and 1615) by clicking the timeline. They can also add TimeSlices to represent different viewpoints at the top and bottom (in Figure 2, "battle" and "gift") to compare multiple viewpoints. Common people in different viewpoints are represented as light green.

Multiple TimeSlices visualize networks of people related to different viewpoints and with different timings. They arrange the same nodes in different TimeSlices in the same positions. The positions of nodes on different TimeSlices can be completely synchronized with one another even if users drag a node. Interactions such as selecting and highlighting nodes are propagated to other TimeSlices. Panning and zooming operations are also propagated to other TimeSlices. These mechanisms help users to compare multiple situations easily.

To compare people and clusters of people in different TimeSlices in detail, our framework utilizes mechanisms for a parallel view [6] in a 3D environment as shown in Figure 5. This enables us to change the normal 3D view to a parallel view seamlessly and explore information interactively through different aspects using multiple views. We can see an overview by using the normal 3D view to find areas of interest and then change the view mode to a parallel view to explore the area in detail.

6.2 Visualizing Components of Relationships

We visualize components of relationships extracted in Section 5 by assigning color to a cluster such as "battle" or "gift", width of an edge to frequency of the cluster appearance, and length of the colored edge to constitution ratio of clusters. Figure 4 shows how to represent the components of relationships between ODA Nobunaga and TAKEDA Katsuyori every year described in Table 2.

In the example in Figure 4 and Figure 6, we use shades of red for clusters related to "adversarial relationship" or "battle", shades of blue for clusters related to friendship, shades of green for other clusters, and gray for keywords that are not categorized in any clusters. Users can specify

[5]MIYOSHI Yasunaga was a military commander who fought against Nobunaga.

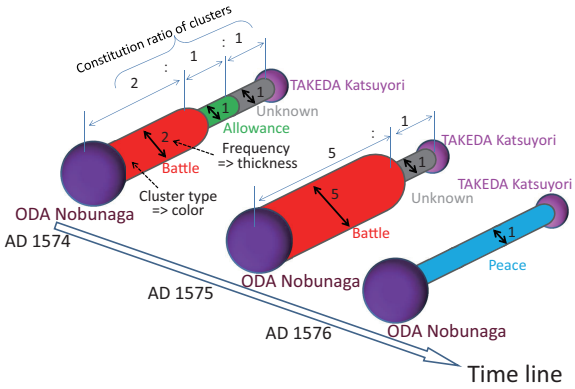


Figure 4. Visualization of components of clusters relationship between two people.

arbitrary groups and clusters in groups and assign arbitrary colors to such groups in the current system. Figure 3 shows an example of grouping.

We can also visualize a set of edges arranged on the timeline to overview changes in relationships between two people every year as shown in Figure 4 and Figure 7 by selecting an edge on the network to specify a person dependency.

group of clusters for adversarial relationship or battle (shades of red)	group of clusters for friendship (shades of blue)
battle, death, Ikko-ikki, hostage, plunder	peace, gift, prayer, grant, celebration, conferment

Table 3. Example of grouping

7 Examples

7.1 Changes in Structure of "Battle" Network

Figure 5 shows a visualization of temporal changes in structure of network of people related to the viewpoint of "battle" from 1568 to 1574.

We adapt keywords for "battle" such as "war", "attack", "surrender", "reconciliation", "formation", "reinforcement", "invasion", "subjugation", and "retreat" for selecting records from the database, and extract the network every year from 1550 to 1650.

Figure 5 (a) shows networks around ODA Nobunaga in 1568, 1569, and 1570 from left to right. We can see that the node of ODA Nobunaga first appears in 1568 and that it is bigger than others. That is because Nobunaga entered Kyoto with ASHIKAGA Yoshiaki to help Yoshiaki became the 15th *Shogun* of the Ashikaga *shogunate*. We can also find out the node of Nobunaga becomes the center of clusters of various nodes, and there are two clusters of people around Nobunaga in 1570. That is because Nobunaga was involved in two battles in 1570. The first was the "Battle of Anegawa". We can find out some big, dark purple nodes (key people) in the upper cluster around Nobunaga because they were commanders or *daimyo*[6] who directly fought for Nobunaga. The second is "Ishiyama Hongan-ji War" which was the war against Nobunaga caused by a network of commanders around Osaka, temples, and communities belonging to the *Ikko*, which was a powerful faction of religious zealots. We can find a complicated cluster of people including some huge, dark purple nodes of people at the bottom in 1570 in Figure 5 (a). Since this war continued for over 10 years, this cluster and key people appear over again.

Figure 5 (b) shows networks of people around ODA Nobunaga in 1572, 1573, and 1574 from left to right. We can see some huge, dark purple nodes of people around Nobunaga in 1572. We assume that is because some famous *daimyo* and commanders created an anti-Nobunaga encircling net to counter Nobunaga around 1572. We can recognize most of these nodes as key people in the anti-Nobunaga encircling net. However, we can observe some other big nodes of people such as UESUGI Kenshin and MOURI Terumoto[7] who did not belong to the anti-Nobunaga net. It is difficult for us to conceive why these nodes are near Nobunaga, e.g. because they formed an alliance with or fought against him. We can see the node of ASHIKAGA Yoshiaki shrinks and becomes gray from 1573 to 1574. We assume that is because ASHIKAGA Yoshiaki was deposed as *Shogun* by Nobunaga, who brought the Ashikaga *shogunate* to an end in 1573. We can see a huge cluster of people in 1573, and two big nodes ("AZAI" and "ASAKURA") in the 1573 cluster disappear in 1574. That is because there was huge battle (Nobunaga vs. AZAI Nagamasa and ASAKURA Yoshikage[8]), after which the AZAI and ASAKURA clans fell. Their nodes never appear after 1574.

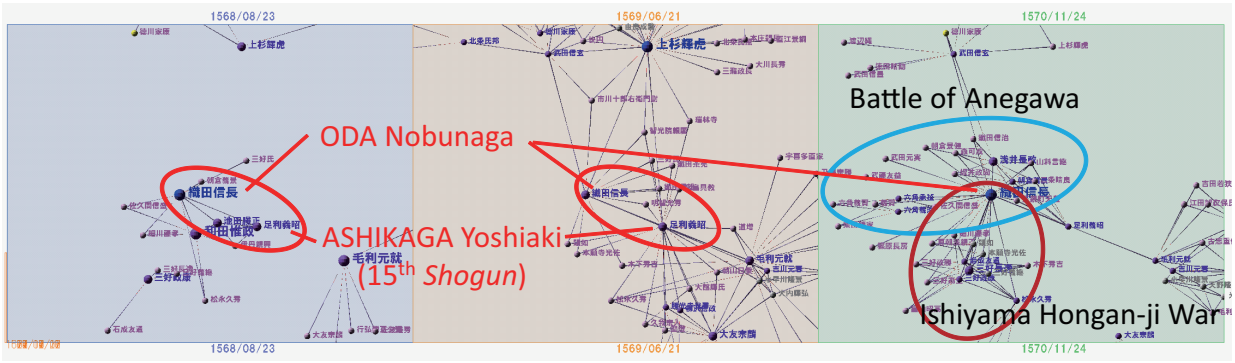
7.2 Comparing Structures of "Battle" and "Gift" Networks

Figure 2 is an example of simultaneously visualizing two networks related to different viewpoints, "battle" and "gift", from 1550 to 1650. The "gift" network is extracted by inputting keywords related to "gift". TimeSlices at the top visualize networks related to the viewpoint "battle", and the bottom ones visualize networks related to the viewpoint

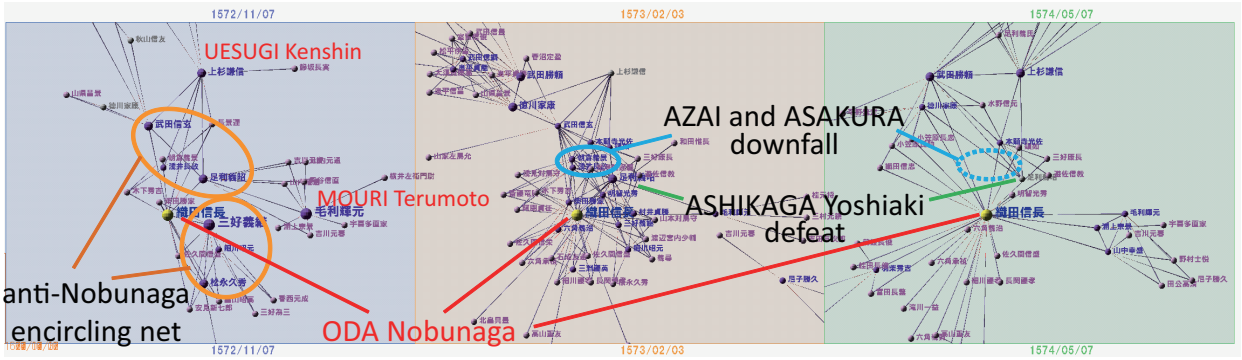
[6]The powerful territorial lords in pre-modern Japan

[7]UESUGI Kenshin and MOURI Terumoto were two of the most powerful *daimyo* of the *Sengoku Jidai*

[8]AZAI Nagamasa and ASAKURA Yoshikage were Japanese *daimyo* of the *Sengoku Jidai*



(a) Network of people around "ODA Nobunaga" from 1568 to 1570



(b) Network of people around "ODA Nobunaga" from 1572 to 1574

Figure 5. Changes in structure of network of people related to "battle" from 1568 to 1574.

"gift". Figure 2 also visualizes TimeSlices with different timings. We place TimeSlices at the left in the position for 1573 and TimeSlices at the right in the position for 1615.

In 1573, we can observe big communities around ODA Nobunaga in the "battle" TimeSlice, but a small cluster with few nodes around Emperor Ogimachi in the "gift" TimeSlice. In contrast, TimeSlices (both at the top and bottom) in 1615 have big clusters around TOKUGAWA Ieyasu and have many common nodes (light green nodes).

We can assume from these results that only a few *daimyo* such as Nobunaga connected with the Emperor through exchanging presents, and most of them connected by fighting around 1573. However, most *daimyo* connected by exchanging gifts with the Tokugawa family in addition to battles around 1615.

These visualizations enable us to show changes in clusters, key people, and differences and similarities among different viewpoints. However, it is difficult for us to recognize the detailed information about relationships between people and characteristics of clusters of people from these results.

7.3 Visualizing Characteristics of Relationships between People Related to a Battle

In Figure 6, we add colors to represent characteristics of relationships mentioned in Section 5 to the network related to "battle" described in Section 7.1 and then visualize their changes.

Figure 6 (a) shows there are many edges with a high ratio of blue around ODA Nobunaga and ASHIKAGA Yoshiaki. We can imagine they tried to build friendly relationships with other people. People around Ishiyama Hongan-ji are connected by blue edges in 1573 in Figure 6 (c). In this case, these people are known as an anti-Nobunaga group. We can then see many red edges appear in 1576 in Figure 6 (d) and can recognize some famous commanders under Nobunaga in the same cluster of people (around Hongan-ji). These represent Nobunaga attacking Hongan-ji with many commanders. Figure 6 (c) shows that there are many red and thick edges around AZAI and ASAKURA. These represent the battle (Nobunaga vs. AZAI and ASAKURA) described in Section 7.1.

In Figure 6 (b) and (d), we can see people belonging to the anti-Nobunaga encircling net as huge nodes. We ex-

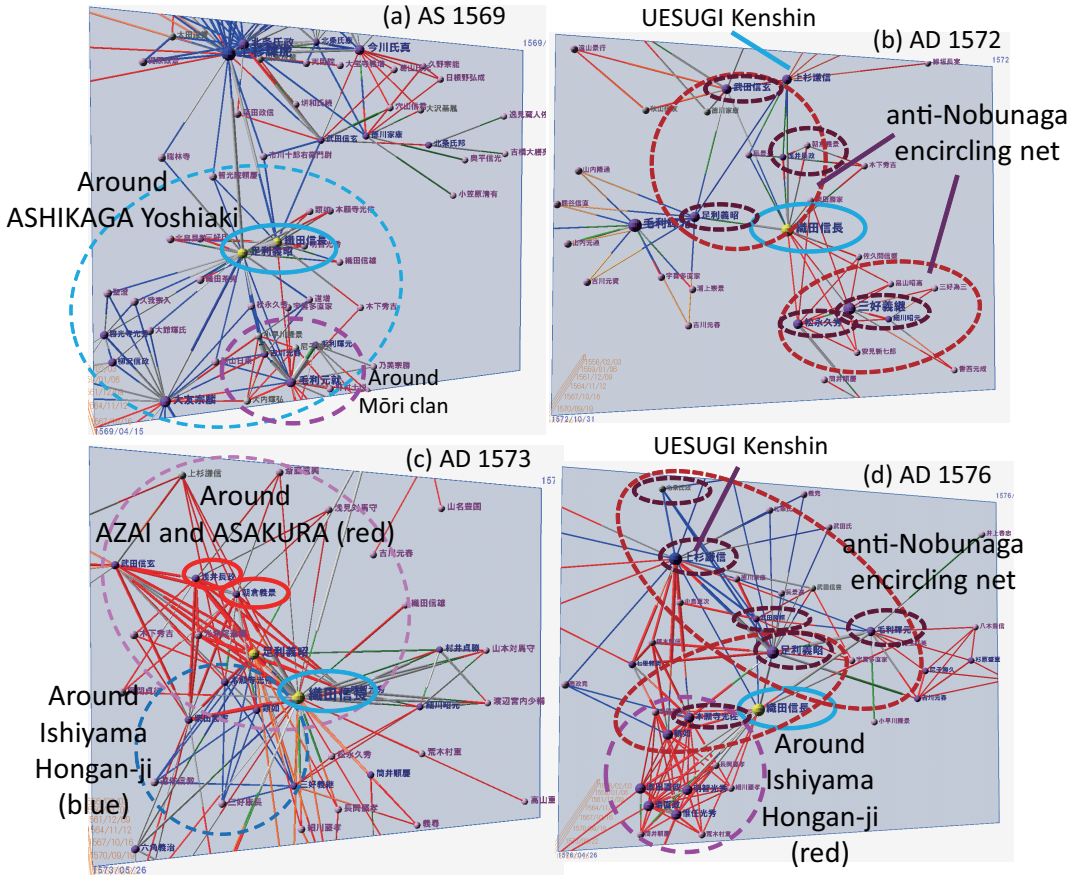


Figure 6. Changes in relationships between people around ODA Nobunaga (blue circle) from 1569 to 1576.

pected they would have red edges to Nobunaga, but some famous *daimyo* do not. Moreover, although UESUGI Kenshin changed from being an ally of Nobunaga to an enemy, the edge between Kenshin and Nobunaga does not change from blue to red. We assume that the historical documents we use in this paper do not include enough descriptions about hostile relationships between Nobunaga and these *daimyo*.

7.4 Example of Changes in Relationships between two People

Figure 7 visualizes details of temporal changes in characteristics of selected relationships between people. Figure 7 (b) shows red edges between TOKUGAWA Ieyasu and TAKEDA Katsuyori. These represent their relationships in continual battles. Figure 7 (c) shows changes in the relationship between ODA Nobunaga and UESUGI Kenshin for each year. Figure 7 (d) visualizes complex relationships between Nobunaga and ASHIKAGA Yoshiaki.

7.5 Related Work

There have been many studies for extracting networks of people from the Web [7, 8] or literature [3] on the basis of co-occurrence of extracted named entities. Itsubo et al. extracted networks of historical people from historical documents using co-occurring place names among people [10]. They also added labels to people in the network by using clustering results of people. Akaishi et al. provided a system for visualizing networks of terms from chronological data called *WordColony* [1]. However, these researchers did not focus on extracting temporal changes in structures of networks of people or characteristics of each relationship between people. Moreover, most of them did not provide an interactive environment for exploring extracted results.

Many systems and/or much researches have attempted visualizing relationships between people in social networks [5] and co-authorship networks [11, 9, 12]. BiblioViz [11] visualizes research topics using SOM to explore research topics related to co-authorship relation-

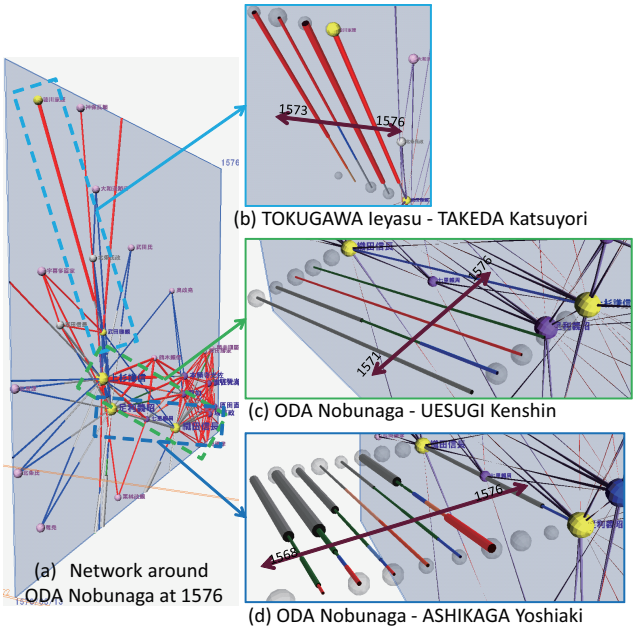


Figure 7. Details of temporal changes in relationships between people.

ships. NeL^2 [9] enables us to interactively explore temporal changes in co-authorship relationships. In contrast, our system enables us to explore both temporal changes in networks and characteristics of relationships.

Conclusion

The objective of this research is to provide tools for visualizing a large number of historical and chronological documents to support searches for new knowledge. In this paper, we have proposed the framework for visualizing time varying structural changes in networks of historical figures and for visualizing characteristics of relationships between them extracted from records about their activities.

This paper has only treated historical events from 1560 to 1580 for extracting clusters representing characteristics of relationships. We plan to expand data for extracting clusters from the *Sengoku Jidai* (from around 1450) to the early stage of Edo period (around 1650).

In future work, we plan to evaluate the effectiveness of our system with expert historians. We will try to check that our system can visualize already-known knowledge and can extract new knowledge such as different activities from well-known knowledge about particular people or unexpected relationships between people.

References

[1] Mina Akaishi, Yoshikiyo Kato, Ken Satoh, and Koichi Hori. Narrative based Topic Visualization for Chronological Data. In *Proc. IV 2007*, pages 139–144, 2007.

[2] Mina Akaishi, Ken Satoh, and Yuzuru Tanaka. An Associative Information Retrieval based on the Dependency of Term Cooccurrence. *Lecture Notes in Artificial Intelligence*, 3245:295–206, 2004.

[3] David K. Elson, Nicholas Dames, and Kathleen R. McKeown. Extracting social networks from literary fiction. In *Proc. ACL 2010*, pages 138–147, 2010.

[4] Thomas M. J. Fruchterman and Edward M. Reingold. Graph Drawing by Force-Directed Placement. *Software Practice and Experience*, 21(11):1129–1164, 1991.

[5] Jeffrey Heer and Danah Boyd. Vizster: Visualizing Online Social Networks. In *Proc. INFOVIS 2005*, pages 32–39, 2005.

[6] Masahiko Itoh, Masashi Toyoda, and Masaru Kitsuregawa. An Interactive Visualization Framework for Time-Series of Web Graphs in a 3D Environment. In *Proc. IV 2010*, pages 54–60, 2010.

[7] Yutaka Matsuo, Junichiro Mori, Masahiro Hamasaki, Keisuke Ishida, Takuichi Nishimura, Hideaki Takeda, Koiti Hasida, and Mitsuru Ishizuka. POLYPHONET: An Advanced Social Network Extraction System. *Journal of Web Semantics*, 5(4):262–278, 2007.

[8] Peter Mika. Flink: Semantic Web Technology for the Extraction and Analysis of Social Networks. *Journal of Web Semantics*, 3:211–223, 2005.

[9] Nagayoshi Nakazono, Kazuo Misue, and Jiro Tanaka. NeL2: Network Drawing Tool for Handling Layered Structured Network Diagram. In *Proc. APVis 2006*, pages 109–115, 2006.

[10] Takahiko Osaki, Sho Itsubo, Fuminori Kimura, Taro Tezuka, and Akira Maeda. Visualization of Relationships among Historical Persons Using Locational Information. In *Proc. W2GIS 2011*, pages 230–239, 2011.

[11] Zeqian Shen, Michael Ogawa, Soon Tee Teoh, and Kwan-Liu Ma. Biblioviz: a system for visualizing bibliography information. In *Proc. APVis 2006*, pages 93–102, 2006.

[12] Lei Shi, Chen Wang, and Zhen Wen. Dynamic network visualization in 1.5D. In *Proc. PVIS 2011*, pages 179–186, 2011.