Summary of this thesis

The Internet, social network service on the Internet, financial network all over the world, biological network of molecular interactions; we have so many kind of networks today. Information, substances, fund, goods, people and other things are traveling over these networks. These various networks often behave in ways that cannot be understood by simply observing their isolated components. This kind of unpredictability is sometimes called ``network effect''. Also, completely different types of networks may share the same properties such as degree distribution and self-similarity.

Complex networks generally have very diverse structures, so it is often difficult to get clues for the analysis. To overcome this difficulty , advanced visualization techniques were developed in this study, and they were applied to visually reveal the characteristics of large, complex networks with directed link. Furthermore, based on the results, we designed and executed a quantitative analysis of the structure to clarify the mechanism that works in the network, especially by using the structural feature known as the ``bow-tie'' structure. To find out structural characteristics of network and its implications by focusing on the bow-tie's diversity is the objective of this study.

Visualizing a complex network can provide us helpful insights to understand topological structures and temporal evolution of the network. Real networks often consist of asymmetric relations and possess flow from upstream to downstream and circulation. Because such structures have vital roles in those networks, expressing link direction in the visualization is essential. While many graph layout algorithms have been studied, there are only few tools that enable us to find detailed structure of such flow and circulation especially for a large-scale network with millions or billions of nodes, recently available for researchers.

First, I apply the visualization method to real data of social networks in order to study the so-called bow-tie structure in social networks. To perform these analyses, it is necessary to extract both the macroscopic features and details of the network. According to the conventional standard, the network visualization should satisfy the condition that the number of crossing edges must be small and that a symmetric drawing can be obtained from a symmetric network. However, these conditions are not appropriate to use the visualization as a clue for quantitative analysis. In this research, instead of conventional conditions like the edge crossing count, the following two conditions are proposed:

1. the same visualization can always be obtained from the same network,
2. and that the visualization of different networks is not the same.

Using these conditions as a guide, two-stage visualization method based on physical simulation that satisfies these standards was developed. The first stage is an initial layout of nodes using the standard method of multidimensional scaling, but with a similarly measure being consistent with directions of links in an approximate way. The second step is a physical simulation using the so-called Barnes-Hut's hierarchical algorithms to calculate electric repulsion force with an additional of magnetic field to visualize links' direction.

Neither the first stage nor the second stage can be applied naively to actual calculations of the data of the scale analyzed in this study. Therefore, in the first stage the matrix dimensions are reduced to work with networks of millions of nodes. In the second stage, optimization and parallelization are performed so that the number of interactions between vertices are reduced. As a result, the combination of these two methods provided an excellent visualization within reasonable time that well represents both the overall features and the details.

I also applied the method to visualize the temporal evolution of the production network, namely how the bow-tie structure changes along with time, to find a stable core part and ephemeral peripheral part. Combined with quantitative analysis, I found that it is the way of transformation of this structure that makes both integrity and flexibility possible within a single economic network.

By using the visualization, I discovered that a bow-tie in the WWW usually has clusters, which are locally-located mini bow-ties that are loosely connected to each other, resulting in a formation of SCC as a whole. To quantify the mutual connectivity among such local bow-tie, I define a quantity to measure how a local bow-tie connects to others in comparison with random graphs. I found that there are striking difference between the WWW and other social and artificial networks including a million firms' nationwide production network among firms in Japan and thousands of symbols' dependency in the programming language of Emacs LISP, in which a global bow-tie exits. I argue that the difference comes from a self-similar structure and development of the WWW speculated by others.

How and why such a large-scale and complex structures as discussed in this research have developed continues to be one of the central issues in complex network study. The concept of self-similarity is a very important guide to contemplate this issue. This is because the similarity of the 43structure between the parts and the whole is not only a feature of the structure itself, but also the process and mechanism that generated the structure.

Through quantitative analysis, it could be affirmed that the relationship between local bow-tie structures strongly affects the properties of the entire network system. The local bow-tie structure of the web link network means limited self-similarity and integrity. In contrast, production network of Japanese companies shows more consistent self-similarity. In addition, time series analysis showed that the central SCC part is more stable while the surround- ing part of IN and OUT is susceptible to change. Furthermore, the difference in the economic role of the IN and OUT sections is reflected in the difference in how the two sections are connected to the central part. The important thing is that these parts are tightly integrated to keep self-similarity. Thanks to its characteristic topological structure, production networks have realized two seemingly contradictory properties of flexibility and stability within single network.

It is unlikely that a huge network like the web will grow without self-similarity. The future work for web link analysis may be to find hidden link struture to complete the self-similarity. As for the production network analysis, working on the aggregated relational data will overestimate the flow and spreading of money and information through the relation. Future work on the economic network will be to make up for these shortcomings and improve the availability of risk assessment methods to bring more economic policy implication.