Visualizing and Analyzing Networks of Co-Purchased Books, CDs and DVDs

T. Iba¹² and M. Mori³

¹Faculty of Policy Management, Keio University, Endo 5322, Fujisawa, Kanagawa, Japan

³Research Fellow, Rakuten Institute of Technology Lab., Shinagawa Seaside Rakuten Tower, 4-12-3 Higashi-shinagawa, Shinagawa-ku, Tokyo, Japan

³Rakuten Institute of Technology Lab., Shinagawa Seaside Rakuten Tower, 4-12-3 Higashi-shinagawa, Shinagawa-ku, Tokyo, Japan

Theme: Economy, Communications and Management

Introduction

Every day, many customers buy many kinds of products like books, CD, and DVD etc, at online stores. Then, the companies often have giant data stock of the transactions. In academic viewpoint, the data is important because it is able to become a clue to understand the complexity of the market. The main issue is what kind of order is emerged as a result of compiling the customers' actions? For the purpose, we propose the method to investigate the co-purchase network of the market, and also visualize and analyze the network with using the real market data of Books, CDs, and DVDs of the online store "Rakuten Books" (http://books.rakuten. co.jp/), which is one of the biggest online stores in Japan. Note that this research was done as an analysis by Rakuten Institute of Technology, and the data do not include any personal information.

Method

The co-purchase network is compiled by the following way. We describe a node A if there is the product A is purchased by the target customers. Then we describe an edge to connect node A and node B if the product A and the product B is purchased by a customer. For describing the edge, we try two types of connection method: "full connection" and "sequential connection" (Figure 1). In the former method, all the nodes which user bought connect each other. In the latter method, nodes connect as the sequential order of user bought. It means that an undirected graph is generated by the former method and a direct graph is generated by the latter.



Results

We visualize the map as a network of the relation among products based on choices by customers. In the case of "full connection" with threshold to visualize the link, we can understand that there are los of components stands for the hidden relationship of the products (Figure 2). We also found the rank distribution of link weight follows power-law in both case of the full-connection and sequential-connection method (Figure 2 and 3).



Figure 2: Co-purchase Networks of Books, CDs, DVDs and Rank distributions of Linkweight (Full-Connection, visualizing links more than weight 2)

Figure 3: Co-purchase Networks of Books, CDs, DVDs and Rank distributions of Link-weight (Sequential-Connection)



We thank the project members: R. Nishida, S. Itoh, Y. Kitayama, and M. Yoshida for the discussion, visualization and analysis. We also thank the members of Rakuten Institute of Technology.



References

- Y. Kitayama, M. Yoshida, S. Takami and T. Iba (2008): Analyzing Co-Purchase Network of Books in Japanese Online Store, poster, International Conference of Network Science '08 (submitted)
- R. Nishida, M. Mori and T. Iba (2008): Analyzing Co-Purchase Network of CDs in Japanese Online Store, poster, International Conference of Network Science '08 (submitted)
- S. Itoh, S. Takami and T. Iba (2008): Analyzing Co-Purchase Network of DVDs in Japanese Online Store, poster, International Conference of Network Science '08 (submitted)

International Workshop and Conference on Network Science 2008