Analyzing Co-Purchase Network of Books in Japanese Online Store

Yuki Kitayama,[*1*2] *Mariko Yoshida,*[*3*2] *Shinya Takami,*[*4] *and Takashi Iba*[*3*5]

**1 Faculty of Environment and Information Studies, Keio University*
**2 Visiting Assistant, Rakuten Institute of Technology Lab.*
**3 Faculty of Policy Management, Keio University*
**4 Rakuten Institute of Technology Lab.*
**5 Visiting Researcher, Rakuten Institute of Technology Lab.*

Introduction

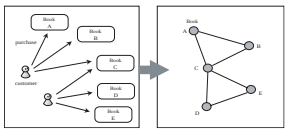
We visualize and analyze the co-purchase network of books in order to clarify the hidden laws in the book sales market. In the co-purchase network, a node represents a book, and the nodes are linked each other if they are purchased by same person. The source data of the network is collected by "Rakuten Books" (http://books.rakuten. co.jp/), which is one of the biggest online book stores in Japan.

*This research was done as an analysis by Rakuten Institute of Technology Lab., and the data do not include any personal

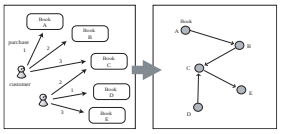
Analyzing the co-purchase network

We analyzed two types of network. One is the full connection, and the other is the sequential connection. In the former type, all the nodes which the customer purchases connect each other. In the latter method, nodes connect as the sequential order of customer's purchase. It means that an undirected graph is generated in the former type and a direct graph is generated by the latter.

Full Connection

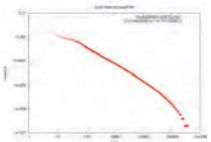


Sequential Connection

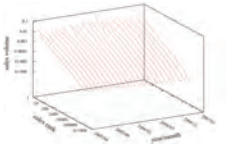


Analyzing the book sales market

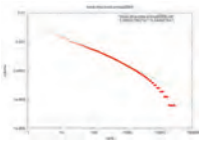
As an analysis of book sales market, we have already found that the relation between sales volume and sales rank follows power law. In addition, the relation between purchase volume and purchase rank of each customers also follows power law. Although the titles of books replaced everyday, these laws can be observed both of monthly and annually.



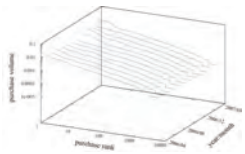
The relation between the sales volume and sales rank (Apr. 2006 - Mar. 2007)



The transition of monthly power laws which represents the relation between the sales volume and sales rank (Feb. 2006 - Mar. 2007)

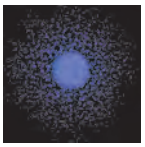


The relation between the purchase volume and purchase rank (Apr. 2006 - Mar. 2007)



The transition of monthly power laws which represents the relation between the purchase volume and purchase rank (Apr. 2006 - Mar. 2007)

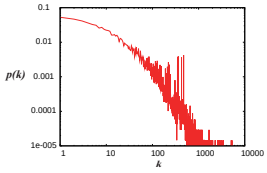
Co-purchase network of Books (Full Connection)



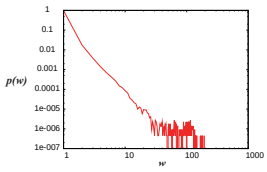
Target Customers: 30,000
Network Type: Full connection
Weight Threshold: weight > 0
The number of nodes: 14,038
The number of edges: 215,260

Weight > 0

Degree distribution
(30,000 customers, Full connection)

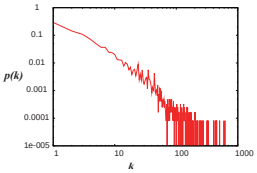


Weight distribution
(30,000 customers, Full connection)



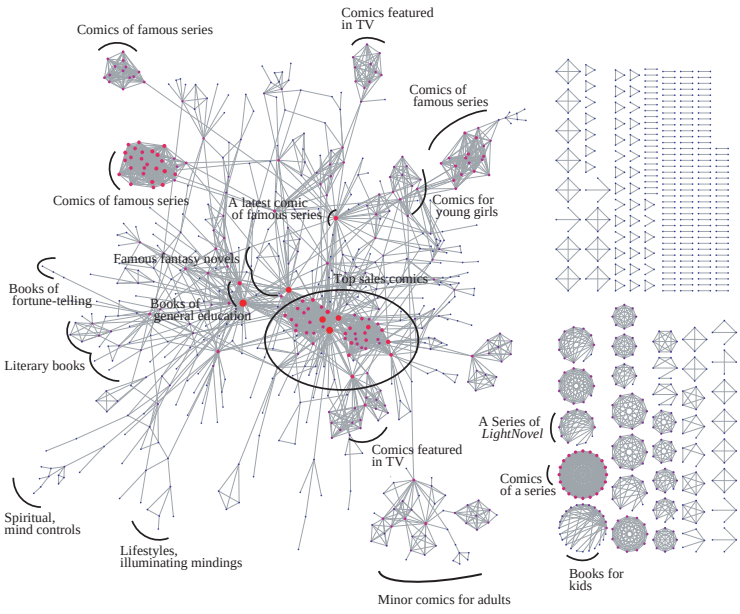
Weight > 1

Degree distribution
(30,000 customers, Full connection)

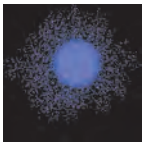


Analyzing Genre in co-purchase network of books

The following figure shows the characteristics of some genres in co-purchase network of books.
(5,000 customers, full connection, weight > 1)



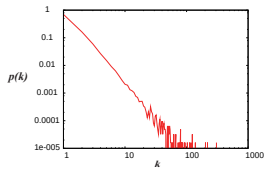
Co-purchase network of Books (Sequential Connection)



Target Customers: 30,000
Network Type: Sequential connection
Weight Threshold: weight > 0
The number of nodes: 68,701
The number of edges: 2,215,260

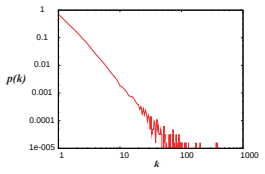
Weight > 0

In-degree distribution
(30,000 customers, Sequential connection)



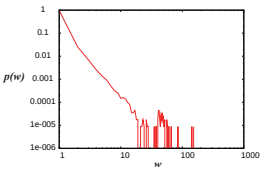
The distribution follows power law, therefore the network can be considered as scale-free network.

Out-degree distribution
(30,000 customers, Sequential connection)



The distribution follows power law, therefore the network can be considered as scale-free network.

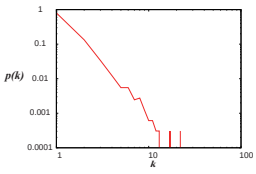
Weight distribution
(30,000 customers, Sequential connection)



The combination of co-purchases seems to be self-organized at the market level.

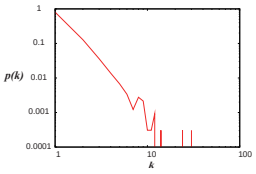
Weight > 1

In-degree distribution
(30,000 customers, Sequential connection)



The distribution follows power law, therefore the network can be considered as scale-free network.

Out-degree distribution
(30,000 customers, Sequential connection)



The distribution follows power law, therefore the network can be considered as scale-free network.