Statistical analysis of the Metropolitan Seoul Subway System: Network structure and passenger flows

Woo-Sung Jung[a], **Keumsook Lee**[b], **Jong Soo Park**[c] and **M. Y. Choi**[d]

[a]Department of Physics, Pohang University of Science and Technology
Pohang 790-784, Republic of Korea
wsjung@postech.ac.kr

[b]Department of Geography, Sungshin Women's University
Seoul 136-742, Republic of Korea

[c]School of Computer Science and Engineering, Sungshin Women's University
Seoul 136-742, Republic of Korea

[c]Department of Physics and Astronomy, Seoul National University
Seoul 151-747, Republic of Korea

The Metropolitan Seoul Subway system, consisting of 380 stations, provides the major transportation mode in the metropolitan Seoul area. Focusing on the network structure, we analyze statistical properties and topological consequences of the subway system. We obtain various network measurements including the path length, clustering coefficient, diameter, and radius as well as the efficiency of the network. The path length, diameter, and radius are also computed in terms of the physical distance between stations.

We further study the passenger flows on the system, and construct the maximum spanning tree of the flows. The Metropolitan Seoul Subway system operates a smart card system which keeps track of the travel information of every passenger. Here we analyze the passenger flows on a single day, based on the transaction data of the smart card on 24 June, 2005. The total number of transactions or passenger flows is as many as 4,909,316 on that specific day.

It is found that the weight distribution displays a power-law behavior whereas the strength distribution follows a log-normal one with a peak at $s \approx 4 \times 10^4$. Note that in the subway network the weight of a link connecting two stations represents the passenger flow between them and the strength of a station corresponds to the number of passengers arriving at and departing from that station. In a metropolis, most facilities are located near stations, so that each station is naturally abundant in passengers, the number of which reflects the capacity of facilities located near the station. The fact that a majority of stations are used by a similar number of passengers, corresponding to the peak of the strength distribution, thus indicates that with residential and commercial facilities taken into account, places near most stations are already developed fully to accommodate dense and compact location of facilities. The peak of the strength distribution thus gives a measure for the characteristic capacity of the facilities near a station in the fully urbanized Seoul. Also revealed is the power-law behavior of the degree distribution of the spanning tree.

Keywords
Passenger flow, Transportation, Subway, Power law