Analysing Consumer Purchase Patterns through the Amazon Product Co-Purchasing Network

Preksha Shah
202203004

Madhav Kanjilimadom
202203018

August 2025

Abstract

This project aims to explore the Amazon product co-purchasing network [1], a large-scale, real-world example of a complex system of interacting components. The dataset, collected in May 2003 by analysing the "Customers Who Bought This Item Also Bought" feature, provides a directed graph where if a product i is frequently co-purchased with product j, the graph contains a directed edge from i to j.

The network is fairly large, comprising 410,236 nodes (products) and 3,356,824 edges (co-purchase links). A key finding is that the entire network forms a single giant weakly connected component (WCC), indicating that all products are interconnected within a single marketplace. More significantly, a massive strongly connected component (SCC) containing 95.1% of all nodes forms the core of the network, revealing a strong, cyclical recommendation system where many product discoveries are highly reciprocal.

By applying network science principles discussed in the course, this project aims to characterise the network's structure and draw conclusions about the consumer behaviour and market dynamics it represents.

Link to Dataset

Stanford Large Network Dataset Collection

References

[1] Jure Leskovec, Lada A. Adamic, and Bernardo A. Huberman. The dynamics of viral marketing. *ACM Transactions on the Web (TWEB)*, 1(1):5, 2007.