A Maximum Entropy Model for the Network of Commercial Transactions between Cities based on Data from Electronic Invoices

Cesar I. N. Sampaio Filho [(a)], Rilder S. Pires [(b)], Humberto A. Carmona [(a)], José S. Andrade Jr. [(a)]
[(a)] Departamento de Física, Universidade Federal do Ceará
[(b)] Laboratório de Ciência de Dados e Inteligência Artificial, Universidade de Fortaleza

We analyze the network of transactions among cities based on the electronic invoice database for the municipalities in the Ceará state, Brazil. This database consists of approximately 6 billion records, each containing 43 fields of information, registered during the period between the years 2016 to 2019. All the transactions are grouped in a unique dataset and represented as an asymmetrical adjacency matrix corresponding to a directed graph with connections weighted by the number of transactions among cities. Due to the large size of Ceará state (≈149,000 km2), its unequal distribution of wealth, and spatially heterogeneous population density, we initially determine communities of cities based on their mutual intensity of trades and then verify to which extent their economic interests somehow reflect what we define as a "community cohesiveness". For the first task, we use the Infomap algorithm to detect the partition M that provides the shortest description length and captures the optimal community structure of the network in terms of its associated flow dynamics. Surprisingly, the partition identified has five modules, whose two-dimensional geographical projections are all simply-connected domains, i.e. , consisting of single pieces without holes. Having described the topological properties of the transaction network, we proceed with the analysis of our database from the perspective of traded products by building bipartite structures described in term of adjacency matrices between municipalities and products, considering both the contexts of selling and purchasing. We then make use of the revealed comparative advantage (RCA) concept, widely used in foreign trade analyses, to define a non-monetary and binary activity index that is capable to distinguish the relative advantage of a city in a class of goods or services as evidenced by trade flows. Finally, through the Pairwise Maximum-Entropy method, we can associate to each of the five communities previously characterized, their corresponding binary Ising-like Hamiltonian models. The local fields and couplings computed for a given community are those that best reproduce the average product activities of its cities as well as the statistical correlations between the product activities of all pairs of its cities. In an analogy with critical phenomena, our results reveal that each community operates at a "temperature" that is close to the corresponding "critical point", suggesting a high degree of "economic cohesiveness" in its trade network of cities.