Analyzing Illegal Psychostimulant Trafficking Networks Using Noisy and Sparse Open Source Data

This paper applies analytical approaches to map illegal psychostimulant (cocaine and methamphetamine) trafficking networks in the U.S. using purity-adjusted price data from the System to Retrieve Information from Drug Evidence (STRIDE) dataset. We make two assumptions for any two connected nodes: (i) the purity-adjusted price is lower at the origin than at the destination and (ii) price perturbations are trans- mitted from origin to destination, resulting in more price co-movement between the two nodes than between two unconnected nodes. We then adopt a two-step analytical approach: we formulate the data aggregation problem as an optimization problem to temporally match data in different states to analyze co-movement, then construct an inferred network of connected states and examine its properties.

We find, first, that the inferred cocaine network created from the optimally ag- gregated dataset explains 46% of the anecdotal evidence, compared to 28.4% for an over-aggregated and 14.5% for an under-aggregated dataset. Second, our network re- veals a number of phenomena, some aligning with what is known and some previously unobserved. To demonstrate the applicability of our method, we compare our cocaine data analysis results with parallel analysis of methamphetamine data. These results likewise align with prior knowledge but also present new insights about locations that are not highlighted in the anecdotal evidence. Our findings show that an optimally aggregated dataset can provide a more accurate picture of an illicit drug network than can suboptimally aggregated data.