Collusion Set Detection using Graph Clustering

Team: Puneet Mangla (CS17BTECH11029)
Yash Khasbarge (CS17BTECH11044)
Rushikesh Tammewar (CS17BTECH11041)

Objective: Implement Collusion clustering algorithm proposed in “Collusion Set Detection using Graph Clustering” Palshikar et. al. for detecting traders engaged in illegal market practices called as “modus operandi”. The algorithm takes the trading transactions as input and returns the candidate set of traders who are involved in modus operandi with high chance.

Algorithm:

Input: StockFlowGraph G, k, m, h

Procedure:

1. Use only k nearest neighbors based on the trading amount for G.
2. Create a separate cluster for every trader.
3. for every possible pair of clusters, compute the collusion index and order every pair according to collusion index.
4. Merge the pair of clusters having the highest collusion index and being khm-compatible with each other.
5. Repeat step 3 and 4 until no further change in the set of clusters.
6. Gather the clusters with large size and compute their collusion index. Clusters with high collusion index can be declared as possible candidates.

Output: Set of set of traders involved in modus operandi

Dataset Summary:

- total transactions: 23432
- Columns: SellerID, BuyerID, Amount in Lakhs
- Unique SellerID: 3939
- Unique BuyerID: 6224

Example:

	A	B	C
1	6941	707	84
2	17371	707	27
3	18216	707	5
4	76646	707	29
5	78095	707	56
6	117584	707	569
7	136859	707	7
8	138668	707	24
9	83910	1014	836
10	99088	1014	430
11	3375	1304	1216

Experimental Results:

- As mentioned in paper, we select
    - k = 4
    - m = 1
    - h = 0.6
- Since the traders present in both buyer and seller lists can only form a cycle, we remove the rest of the traders from the dataset. Finally, 2969 traders were only the possible candidates.
- After running algorithm, we note down the following size of clusters:
-

Size of cluster	Number of clusters
1	2467
2	237
3	5
4	2
5	1

- We note that, from this clustering, the clusters with size equal to or more than 3 should be the candidate sets.
- We also report the collusion index of clusters with size >= 3
    - cluster {136850, 239051, 246655} collusion index 0.3008356545961003
    - cluster {11166, 16310, 10942} collusion index 0.3338008415147265
    - cluster {20378, 25315, 134414, 11599} collusion index 1.2740606111839046
    - cluster {13728, 27700, 30429, 22959} collusion index 1.8742378518199174
    - cluster {87536, 113914, 218343} collusion index 8.557377049180328
    - cluster {40115, 22824, 33449, 63691, 9310} collusion index 1.9138367967562089
    - cluster {231171, 129313, 40411} collusion index 0.8173713703515028
    - cluster {43162, 16259, 19886} collusion index 0.4364691081912217
- For the sake of completeness, we report all the clusters with size >= 2,
    - {49189, 20311} {188458, 257006} {6784, 8245} {126443, 57413} {49240, 93492} {40123, 8284} {139400, 52399} {131272, 111245} {183984, 24777} {57562, 57586} {170322, 49372} {139505, 47601} {179636, 82167} {123128, 237230} {8441, 20433} {1368, 8444} {100105, 164099} {82180, 198574} {157648, 33073} {57656, 112546} {164173, 75806} {16721, 122499} {90461, 94702} {115073, 175505} {18146, 188811} {41360, 124403} {34152, 213409} {188834, 195086} {164267, 180923} {105092, 115172} {70680, 16889} {115246, 215863} {33329, 28866} {10491, 82486} {58560, 594} {14033, 8830} {246960, 164501} {17057, 158061} {123590, 93903} {102125, 107223} {90848, 146693} {140030, 129231} {149004, 254727} {64289, 58149} {210110, 8998} {78121, 9010} {13529, 9058} {38423, 50031} {2948, 9077} {189332, 70702} {25524, 85006} {194674, 164804} {164850, 94198} {129008, 156701} {91195, 67879} {25674, 95220} {42072,

102355} {118193, 115823} {13548, 83103} {13722, 83108} {6729, 165060} {93699, 74959} {59435, 17622} {25818, 40602} {74971, 9468} {26444, 99559} {50462, 19471} {17698, 118755} {52857, 75125} {48445, 181621} {78960, 189839} {59584, 1442} {50609, 48492} {44489, 50626} {136850, 239051, 246655} {121375, 34255} {91752, 148953} {124393, 124394} {230890, 4926} {50739, 50780} {50778, 29991} {166209, 247413} {178930, 124547} {42642, 7671} {13540, 34463} {137504, 149159} {29168, 198324} {51632, 59106} {34538, 16095} {18211, 44933} {26409, 130199} {37826, 116524} {125849, 165678} {124754, 143646} {1916, 187174} {56425, 18306} {179953, 42882} {78266, 42907} {206754, 184123} {146035, 51110} {239544, 7164} {200986, 182205} {239573, 240013} {108579, 170615} {32098, 2102} {27209, 84025} {2132, 46741} {2134, 7199} {223328, 144628} {143987, 256103} {29033, 26750} {116874, 145731} {190624, 154238} {2277, 5046} {166123, 219741} {2290, 143586} {47233, 10501} {60515, 125247} {71529, 133462} {125840, 51561} {55939, 2415} {232164, 76165} {226392, 10687} {55707, 117205} {2898, 2591} {77099, 84525} {85940, 43596} {73705, 92757} {144321, 10855} {11166, 16310, 10942} {111224, 109261} {163409, 232161} {77116, 84709} {240744, 117484} {43761, 89561} {211714, 76543} {29189, 109414} {109416, 85291} {119804, 117606} {191353, 13618} {38998, 207759} {101305, 101306} {76730, 135138} {27636, 73085} {142357, 72455} {3401, 142359} {248914, 154871} {60530, 38914} {44216, 151552} {184331, 3259} {68808, 71364} {114474, 60629} {109822, 87191} {85283, 137868} {62216, 175415} {20378, 25315, 134414, 11599} {12999, 159191} {20507, 52725} {15869, 36342} {134648, 15246} {167416, 145871} {220194, 224763} {20485, 44566} {15553, 3725} {118448, 136930} {80056, 11957} {179254, 77494} {176137, 159421} {134857, 13958} {169301, 93917} {152984, 151267} {128722, 110341} {89362, 192324} {234658, 126795} {28498, 20718} {24447, 118687} {15200, 102385} {176116, 122607} {130857, 86038} {77869, 97287} {217981, 86078} {154936, 127039} {20560, 37460} {209020, 218702} {188199, 151679} {89044, 200839} {69794, 64372} {45304, 6941} {12552, 152327} {45322, 169748} {63810, 86318} {219712, 78146} {12640, 153826} {23917, 45429} {23000, 45443} {87082, 61828} {87769, 184714} {6722, 12684} {80888, 94613} {15649, 20899} {143888, 145138} {103040, 184954} {152209, 209675} {54024, 21999} {169992, 21335} {86872, 22131} {64818, 13187} {185344, 226527} {95289, 80157} {177250, 218938} {15202, 193655} {5748, 177285} {112989, 152823} {39137, 136476} {145980, 62781} {251256, 194665} {13728, 27700, 30429, 22959} {87536, 113914, 218343} {54786, 89195} {145033, 96665} {211345, 63141} {14048, 23593} {63323, 145726} {96596, 161653} {38794, 39899} {96146, 73698} {137230, 47599} {15589, 14495} {30929, 188251} {153810, 32564} {55536, 106114} {227570, 162427} {40115, 22824, 33449, 63691, 9310} {122622, 145710} {89976, 178573} {14806, 39543} {32357, 31197} {96928, 186850} {15058, 15102} {97999, 56031} {187208, 138701} {228962, 171061} {231171, 129313, 40411} {43162, 16259, 19886} {98256, 98258}