ICCEPM 2022
The 9th International Conference on Construction Engineering and Project Management
Jun. 20-23, 2022, Las Vegas, NV, USA

Detecting Red-Flag Bidding Patterns in Low-Bid Procurement for Highway Projects with Pattern Mining

Chau Le[1]*, Trang Nguyen[2], Tuyen Le[3]

[1] *Department of Civil, Construction and Environmental Engineering, North Dakota State University, 202L CIE, Fargo, ND 58102, USA,* E-mail address: chau.le@ndsu.edu
[2] *Department of Accounting and Information System, North Dakota State University, Fargo, ND 58102, USA,* E-mail address: trang.nguyen@ndsu.edu
[3] *Glenn Department of Civil Engineering, Clemson University, 316 Lowry Hall, Clemson, SC 29634, USA,* E-mail address: tuyenl@clemson.edu

Abstract: Design-bid-build (DBB) is the most common project delivery method among highway projects. State Highway Agencies (SHAs) usually apply a low-bid approach to select contractors for their DBB projects. In this approach, the Federal Highway Agency suggests SHAs heighten contractors' competition to lower bid prices. However, these attempts may become ineffective due to collusive bidding arrangements among certain contractors. One common strategy is the rotation of winning bidders of a group of contractors who bid on many of the same projects. These arrangements may also be specific to a particular region or vary in time. Despite the practices' adverse effects on bidding outcomes, an effective model to detect red-flag bidding patterns is lacking. This study fills the gap by proposing a novel framework that utilizes pattern mining techniques and statistical tests for unusual pattern detection. A case study with historical data from an SHA is conducted to illustrate the proposed framework.

Keywords: bidding pattern, pattern mining, frequent itemset mining, procurement, bid tabulation.

1. INTRODUCTION

Design-bid-build (DBB) is highway projects' most common project delivery method. A State Highway Agencies (SHA) usually apply a competitive bidding procedure and identify the lowest responsible bidder with a responsive bid among the bidders to award a DBB contract [1]. This approach assumes that project owners benefit from a competitive marketplace to obtain reasonable bid prices. Studies have also shown that a specific saturation level in the market is necessary to ensure project costs and client satisfaction [2, 3]; a lack of competition may lead to a monopoly that reduces the procurement process's efficiency [3]. Therefore, the Federal Highway Agency (FHWA) suggests that SHAs find solutions to enhance contractor competition continuously [4].

However, these attempts may become ineffective due to collusive bidding arrangements among certain contractors. In contrast to project owners' perspective, some contractors attempt to rig the bidding process to shift from competitive pricing to monopoly pricing to maximize their profits [5]. Contractors' collusive bidding behaviors are not rare in public projects and are considered unethical & illegal [6, 7]. One common strategy is the rotation of winning bidders of a group of contractors who bid on many of the same projects [4]. These arrangements may also be specific to a particular region, particularly rural areas where the local contractor community is small [4]. A contractor may also convene other contractors in its circle to give it an edge in bidding on a new project [5].

To protect the public interest and ensure the transparency and integrity of public project procurement, public project owners should be able to detect contractors' collusive biding behaviors by identifying (1) bidding patterns or (2) irregular bids [8]. However, past studies were mostly focused on detecting bidding irregularities by analyzing unit prices in historical bid tabulation data. For example, Chotibhongs and Arditi [7] analyzed the bidding data of 108 contracts of a construction owner in ten years to develop a regression model to formulate the cost structure of a bid and then identify potential collusive bidders using the residual and stability tests. Ballesteros-Pérez, et al. [6] developed a new method to detect abnormally high or low bids for construction contract auctions. Differences between contractors' bids and the owner's pre-bid estimates were analyzed in [9] to develop a probabilistic method to identify abnormal bids in infrastructure projects. In the same line of research, Benford's law was applied to more than 100,000 asphalt bid items to determine irregular unit prices [10].

However, few studies focused on developing quantitative measurements of collusion or detecting bidding patterns, e.g., the co-occurrence of some contractors in many of the same contracts to alternate the winning bidder or favor a specific contractor. This study fills the gap by proposing a novel framework that utilizes pattern mining techniques and statistical tests for unusual pattern detection. Actual historical data from an SHA was collected to illustrate the proposed framework.

2. METHODOLOGY

This section presents a novel method of detecting red-flag patterns from historical bid tabulation data using frequent itemset mining (FIM) and the Fisher Exact Test. The method includes three main steps:

1) Preparation of a database of the contractors co-occurring in the bidding of past contracts and those winning the contracts from bid tabulation data,
2) Determination of sets of contractors and their frequencies of bidding together by applying an FIM algorithm to the prepared database, and
3) Detection of cases in which the winning proportion of a contractor was improved significantly with the co-occurrence of some particular contractor(s). The Fisher Exact test is used to compare two independent population proportions (i.e., with and without the co-occurrence).

The proposed method is not aimed to conclude actual collusion cases. It is expected to be used to detect potential red-flag patterns only to assist decision-making processes. Detailed investigations from project owners are necessary to determine a group of contractors actually colluded or not, which is out of the scope of this paper.

2.1. Step 1 — Data preparation

The bid tabulation data of an SHA was collected for this study, including 1,590 contracts from 2009 to 2019. Each contract has information about its bidders and the winning contractor. A total of 248 distinct contractors participated in bidding on these contracts, and each contractor was assigned an identification number (ID) ranging from 1 to 248. Table 1 shows the data attributes used for data analysis. For example, contractors 18, 70, and 90 competed for contract 8099, and contractor 18 won that contract.

Table 1. Data attributes

No.	Contract ID	Bidder ID	Winner ID
1	8099	18, 70, 90	18
2	8038	70, 90, 182	90
3	8036	53, 83	83
4	8009	53, 83, 152	83
5	7938	4, 97, 115, 208	115
…	…	…	…
1586	2204	97, 115, 87, 200, 76, 197	87
1587	2197	97, 219	219
1588	2184	97, 115, 87, 77, 219, 76	77
1589	2156	77, 188, 117, 124	124
1590	2125	149, 146, 9	9

2.2. Step 2 — Determination of the frequencies of contractors bidding together

FIM is a type of data mining initially designed for market basket analysis to deal with transaction databases and discover frequent patterns or groups of items that frequently co-occur in a customer's transactions [11]. With a database of transactions of items as input, an FIM algorithm (e.g., Apriori, EClaT, or FP-Growth) can output all itemsets and their frequencies of co-occurrence (also known as support) [12]. Users can also set a minimum support threshold to extract only itemsets with support not less than the threshold [11]. From its inception, FIM has been applied to numerous domains, such as product recommendation, bioinformatics, e-learning, web browsing analysis, sociology, and text mining [11, 13].

An FIM (i.e., Apriori) was applied to determine the co-occurrence of contractors in the bidding of past contracts (see Figure 1). Using a different FIM algorithm (e.g., EClaT or FP-Growth) would produce the same result with only a difference in the running time. The output is groups of contractors and their frequencies of bidding together. For example, contractors 27 and 30 submitted bids for the same contract 181 times, and contractors 27, 30, and 62 tendered together 133 times. The group size of 1 is a special case; it only indicates the number of contracts a contractor submitted a bid. Take contractor 27 as an example. It bid for 382 contracts out of 1,590 contracts in the dataset.

Input

Contract ID	Bidder ID
8099	18, 70, 90
8038	70, 90, 182
8036	53, 83
8009	53, 83, 152
7938	4, 97, 115, 208
…	…
2204	97, 115, 87, 200, 76, 197
2197	97, 219
2184	97, 115, 87, 77, 219, 76
2156	77, 188, 117, 124
2125	149, 146, 9

An FIM algorithm →

Output

Itemset/ Set of Contractors	Size	Frequency
27	1	382
62	1	237
32	1	234
30	1	218
39	1	197
140	1	184
27, 30	2	181
…	…	…
27, 62, 30	3	133
…	…	…
107, 130, 163, 94, 208, …	11	1

Figure 1. Extraction of groups of contractors and their frequencies from the database

The original output had 23,612 rows/groups of contractors with the group size from 1 to 11 and the frequency from 1 to 382. Since this study was focused on detecting signs of collusion between contractors, the group size of 1 was excluded from further analysis. Also, groups of contractors that rarely bid together were included using a minimum frequency threshold of 4. After applying two filter criteria, only 1,964 contractor groups remained, with the group size from two to eight contractors.

2.3. Step 3 — Detection of red-flag patterns

The high co-occurrence level of a group of contractors alone is not an indicator of collusion among the contractors because they simply bid on many projects and accidentally appeared together. In this study, an additional step was conducted to see whether there was a significant improvement in the winning proportion of a contractor by the co-occurrence of certain contractors compared without those contractors' participation. For each of the 1,964 groups of contractors, the winning proportion of each contractor in the group was calculated in two cases: (1) the contractor participated with all other contractors in the group and (2) the contractor participated but without the co-occurrence of all other contractors in the group (see Figure 2). The two proportions were compared using the Fisher Exact Test, which applies to both small and large samples. More details about the test can be found in [14].

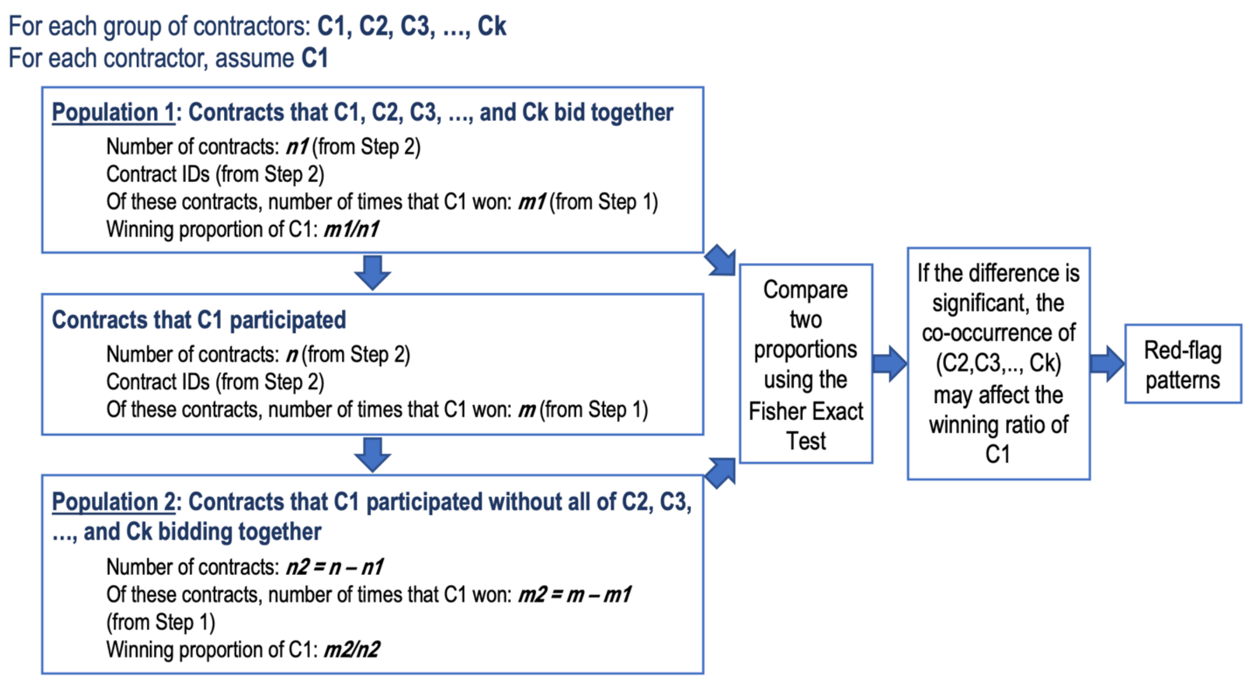

Figure 2. Detection of red-flag patterns

3. RESULTS AND DISCUSSION

Tables 2, 3, and 4 show some representative red-flag patterns obtained by applying the proposed method to the input bid tabulation data. The tables correspond with the group sizes of 2, 3, and 4, respectively. An example of groups of two contractors is case 3 in Table 2. When contractor 121 and contractor 177 bid together, contractor 121 was the winning bidder of 16 out of 33 contracts, achieving a winning proportion of 48.5%. Conversely, contractor 121 bid on 67 contracts without contractor 177, and it won only three times, with a winning proportion of only 4.5%. Since 4.5% is dramatically smaller than 48.5% (the difference is also statistically significant according to the

Fisher Exact Test at the significant level of 0.05), one may suspect the influence of contractor 177 on the winning chance of contractor 121, indicating a red-flag sign to be further investigated by the project owner. One of the probable reasons was the collusion between contractor 121 and contractor 177 to obtain a favorable result for contractor 121.

Table 2. Red-flag patterns from groups of two contractors

No.	Case	Number of contracts		Winning proportion	Fisher Exact Test p-value
		Bidding	Winning		
1	177 with 212	59	25	0.424	0.004
	177 without 212	91	19	0.209	
2	165 with 115	46	19	0.413	0.004
	165 without 115	80	14	0.175	
3	121 with 177	33	16	0.485	0.000
	121 without 177	67	3	0.045	
4	121 with 212	25	12	0.480	0.000
	121 without 212	75	7	0.093	
5	18 with 137	15	7	0.467	0.001
	18 without 137	161	15	0.093	

An example of groups of three contractors is case 5 in Table 3. Contractor 30 bid together with both contractor 16 and contract 201 six times, and it won in five out of those six times, with a winning proportion of 83.3%. Meanwhile, out of 212 contracts that contractor 30 participated in but both contractor 16 and contractor 201 did not, contractor 30 only won in 23 contracts, at the winning proportion of only 10.8%. The vast difference in the winning proportions in the two scenarios indicates a possible influence of the co-occurrence of contract 16 and contractor 201 on the winning chance of contractor 30. For example, contractor 30 might negotiate with contractor 16 and contractor 201 in advance of bidding to help it win.

Table 3. Red-flag patterns from groups of three contractors

No.	Case	Number of contracts		Winning proportion	Fisher Exact Test p-value
		Bidding	Winning		
1	165 with both 97 & 115	26	13	0.500	0.003
	165 without both 97 & 115	100	20	0.200	
2	121 with both 177 & 212	23	11	0.478	0.000
	121 without both 177 & 212	77	8	0.104	
3	222 with both 19 & 27	9	8	0.889	0.000
	222 without both 19 & 27	79	21	0.266	
4	16 with both 30 & 120	7	7	1.000	0.002
	16 without both 30 & 120	47	17	0.362	
5	30 with both 16 & 201	6	5	0.833	0.000
	30 without both 16 & 201	212	23	0.108	

Regarding groups of four contractors, case 4 in Table 4 is interesting. Contractor 16 bid together with all of the contractors 27, 30, and 120 in six contracts, and it won in all of them, with a perfect winning proportion of 100%. Without the co-occurrence of all contractor 27, contractor 30, and contractor 120, the winning proportion of contractor 16 was much smaller, at 37.5%. Again, the project owner may suspect collusion among the four contractors.

Table 4. Red-flag patterns from groups of four contractors

No.	Case	Number of contracts		Winning proportion	Fisher Exact Test p-value
		Bidding	Winning		
1	121 with all 165, 177, & 212	14	7	0.500	0.005
	121 without all 165, 177, & 212	86	12	0.140	
2	222 with all 18, 19, & 27	8	7	0.875	0.001
	222 without all 18, 19, & 27	80	22	0.275	
3	30 with all 27, 62, & 201	7	5	0.714	0.000
	30 without all 27, 62, & 201	211	23	0.109	
4	16 with all 27, 30, & 120	6	6	1.000	0.005
	16 without all 27, 30, & 120	48	18	0.375	
5	30 with all 16, 27, & 201	6	5	0.833	0.000
	30 without all 16, 27, & 201	212	23	0.108	

4. CONCLUSIONS

The primary contribution to the body of knowledge of this study is a novel method of detecting red-flag bidding patterns from historical bid tabulation data using FIM and the Fisher Exact Test. The method provides the identity of the contractor with potential collusive behaviors along with other contractors possibly involved in the collusive arrangement. The contractor's winning proportions with and without the arrangement are also available to project owners to investigate the contractor further if necessary. SHAs can quickly apply the proposed method to their pre-existing bid tabulation data to enhance their bid evaluation process without the need to collect any additional data. Whereas the method was specifically developed for highway projects, it can be used by any major project owners that maintain bid tabulation data of their past projects.

Future studies on determining collusion patterns are needed. For example, future research might explore bidding patterns under different project conditions, such as project work types, contract sizes, and geographic areas. Spatial and temporal analyses may be incorporated into the current method to examine bidding patterns in different regions and time frames. Stronger indicators of collusive bidding behaviors may be necessary apart from the significant increase in winning ratios from without to with the co-occurrence of the contractors in a pattern. For example, analyzing both bidding patterns and irregular bids simultaneously may better support project owners in detecting collusion among bidders.

REFERENCES

[1] C. Le, T. Le, H. D. Jeong, and E.-B. Lee, "Geographic Information System–Based Framework for Estimating and Visualizing Unit Prices of Highway Work Items," *Journal of Construction Engineering and Management*, vol. 145, no. 8, p. 04019044, 2019/08/01 2019, doi: 10.1061/(ASCE)CO.1943-7862.0001672.

[2] P. G. Carr, "Investigation of Bid Price Competition Measured through Prebid Project Estimates, Actual Bid Prices, and Number of Bidders," *Journal of Construction Engineering and Management*, vol. 131, no. 11, pp. 1165-1172, 2005, doi: doi:10.1061/(ASCE)0733-9364(2005)131:11(1165).
[3] S. O. Cheung and L. Shen, "Concentration Analysis to Measure Competition in Megaprojects," *Journal of Management in Engineering,* vol. 33, no. 1, 2017, doi: 10.1061/(asce)me.1943-5479.0000464.
[4] FHWA. "Suggested Guidelines for Strengthening Bidding and Contract Procedures." https://www.fhwa.dot.gov/programadmin/contracts/ashtobid.cfm (accessed.
[5] X. Wang, K. Ye, M. Chen, and Z. Yao, "A Conceptual Framework for the Inclusion of Exogenous Factors into Collusive Bidding Price Decisions," *Journal of Management in Engineering,* vol. 37, no. 6, 2021, doi: 10.1061/(asce)me.1943-5479.0000981.
[6] P. Ballesteros-Pérez, M. Skitmore, R. Das, and M. L. d. Campo-Hitschfeld, "Quick Abnormal-Bid-Detection Method for Construction Contract Auctions," *Journal of Construction Engineering and Management,* vol. 141, no. 7, p. 04015010, 2015, doi: doi:10.1061/(ASCE)CO.1943-7862.0000978.
[7] R. Chotibhongs and D. Arditi, "Detection of Collusive Behavior," *Journal of Construction Engineering and Management,* vol. 138, no. 11, pp. 1251-1258, 2012, doi: 10.1061/(asce)co.1943-7862.0000542.
[8] FHWA, "4D Modeling: A State Transportation Agency's Approach to Risk Management," ed: Federal Highway Administration (FHWA), 2017.
[9] R. Signor, P. E. D. Love, A. T. N. Belarmino, and O. Alfred Olatunji, "Detection of Collusive Tenders in Infrastructure Projects: Learning from Operation Car Wash," *Journal of Construction Engineering and Management,* vol. 146, no. 1, 2020, doi: 10.1061/(asce)co.1943-7862.0001737.
[10] A. Erfani, K. Zhang, and Q. Cui, "TAB Bid Irregularity: Data-Driven Model and Its Application," *Journal of Management in Engineering,* vol. 37, no. 5, p. 04021055, 2021, doi: doi:10.1061/(ASCE)ME.1943-5479.0000958.
[11] P. Fournier-Viger, J. C.-W. Lin, B. Vo, T. T. Chi, J. Zhang, and H. B. Le, "A survey of itemset mining," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,* vol. 7, no. 4, p. e1207, 2017/07/01 2017, doi: 10.1002/widm.1207.
[12] C.-H. Chee, J. Jaafar, I. A. Aziz, M. H. Hasan, and W. Yeoh, "Algorithms for frequent itemset mining: a literature review," *Artificial Intelligence Review,* vol. 52, no. 4, pp. 2603-2621, 2019/12/01 2019, doi: 10.1007/s10462-018-9629-z.
[13] J. M. Luna, P. Fournier‐Viger, and S. Ventura, "Frequent itemset mining: A 25 years review," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,* vol. 9, no. 6, p. e1329, 2019.
[14] R. L. Ott and M. Longnecker, *An Introduction to Statistical Methods and Data Analysis*, Seventh ed. Cengage Learning, 2015.