Atlantis Highlights in Engineering, volume 5

International Conference on Precision Machining, Non-Traditional Machining and Intelligent Manufacturing (PNTIM 2019)

Research on Surrounding String Recognition Technology Based on Big Data and Semantic Analysis

Xie Rongwei
Guangzhou Power Supply Bureau Co., Ltd.
Guangdong, 510000, China
E-mail: jackxrw@163.com

Zheng Nannan
Guangzhou Women and Childrens Medical Center
Guangdong, 510000, China

***Abstract*—In bidding process, even if the bidder knows it in advance, electronic bid evaluation algorithm based on price factor deviation rate and published in the bidding documents, where the deviation rate of the tender is calculated by directly weighting the average deviation rate of the price component and its weight is the proportion of each price component to the total quotation, has a linear structure. Therefore, semantic analysis ability based on the price factor deviation rate in big data is used as the surrounding string recognition technology in this paper, and bidding security identification analysis ability is established with the help of linear indexing algorithm of linear mechanism to provide service for the fairness of bidding.**

Keywords-Big Data; Semantic Analysis; Surrounding String Recognition Technology

I. INTRODUCTION

In the bidding of engineering quantity list, the quantity of project is given, which means that the bidding price of each tender is linear function with comprehensive unit price. If bidder knows it in advance that the electronic bid evaluation algorithm published in the bidding document also has a linear structure, then he may combine some bidders to collude the bidding quotation privately, and use the potential defect of electronic bidding algorithm itself to influence or even control the lowest bidding price, which will damage legitimate rights and interests of tenderer or other bidders. In other words, collusion on bidding quotations violates *The People's Republic of China Bidding Law*, but if the electronic bidding algorithm itself is flawed, it will be impossible to prevent such violations. Moreover, linear weighted structure of tender deviation rate means tender deviation rate is only related to arithmetic mean of price constituents, and has nothing to do with its value distribution, which offers bidders opportunity to collude bidding quotes and adjust the value of each price component in illegal bid so that these illegal bids can be made legal in form. Therefore, safety of a class of electronic bid evaluation algorithms based on price factor deviation rate is analyzed in this paper, where the linear structure will be used to design a kind of beacon algorithm and theoretically analyze it to demonstrate its potential safety hazards[1-2].

II. SURROUNDING ALGORITHM

Electronic bid evaluation algorithm based on deviation rate is designed to win the bid as much as possible at reasonable low price, but the algorithm itself has serious defects, which mainly exist in following two designs.

1) Benchmark value of price component (labor fee, material fee, machinery fee, and comprehensive fee) of all the items in bidding documents is equal to the arithmetic mean of remaining sample values after discrete sample values are removed from all bids. However, according to the law of large numbers, when the number of bids is $m \geq 30$, arithmetic mean should be quite close to market price of factor so that it can be easily estimated by market price. What's more, according to the law of large numbers, arithmetic mean of all legal tenders labor costs should be very close to its market price of 978 yuan. If all bids are legal, in bid evaluation algorithm based on deviation rate, labor cost reference value of list item 001 should satisfy $\overline{\mathrm{RG}_1} \approx 978$.

2) The deviation rate of tender is calculated by direct weighted average of deviation rate in price component (labor fee, material fee, machinery fee, comprehensive fee, measure fee and so on), and its weight is the proportion of each price component in total quotation, where the security of bid evaluation algorithm is extremely fragile and easy to be manipulated since linear weighted structure means that as long as the deviation rate of all price components in tender i reaches the minimum at the same time, deviation rate PBSi of tender will be naturally minimized. Then according to the law of large numbers, bid price of the bid will be the lowest bid price with a high probability.

Next, how bid evaluation algorithm based on bias rate is manually manipulated due to these two design flaws will be analyze, and it will be proved that only a small number of bidders collude with the bid price can the bid price of specified bid be brought to theoretical minimum with great probability[3].

If the number of bids is *m* in which bids *1, 2, ..., n* have been colluded with the bid price, and it is intended to make bid 1 the winner with the lowest bid price, deviation rate of bidding book 1 will be controlled to the minimum by the following steps with the help of law of large numbers and linear structure of bid evaluation algorithm so that purpose of winning the bid can be achieved. In addition, it should be

paid much attention that at this time only the bids *n+1, n+2,...,m* are legal.

The first step is to study the bidding documents and get familiar with the requirements in bidder's instructions and various parameters in evaluation algorithm. For example, *Bidding Documents (2019 Edition)* states as follows.

1) The bidder's quotation lower than bidding control price of 80% and higher than bidding control price, the bid will be directly rejected without entering next review.

2) When bidder's bid price is lower than 85% of bidding control price and lower than 95% of arithmetic mean of bid price from all bidders (excluding the bid price from bidder whose quote is directly scraped as it is mentioned in (1)), the bidder's bid price will be deemed to be lower than the cost quote, and the bid will be treated as scrap.

Bidders do not risk direct taxation, that is, they can reasonably assume that the bidder's total offer is no less than 85% of bidding control price. For the sake of convenience, *d=0.95* in bidder' s notice can be seen as price control factor.

The second step is to conduct a market survey to determine the market price of price component (labor fee, material fee, machinery fee, and comprehensive fee) in each list item so that the total price of bid book 1 can be determined[4].

Taking list item 001 (build wall) in Table 1 as an example and assuming that the maximum likelihood estimate on market price of labor fee is $\hat{RG}_1 = 978$ through market survey, when the number of bids is m ≥ *30*, arithmetic mean of labor costs in all legal tender should be very close to $\hat{RG}_1$ according to the law of large numbers. In order not to become a scrap, all bidders' quotations will be adjusted proportionally according to the bidding control price. For the sake of simplicity, it is assumed that the bidding control price is higher than market cost , when quotation according to market price will not become a scrap, so labor fee for list item 001 of tender 1 is $RG_{11} = \hat{RG}_1$. Additionally, it needed to be paid more attention that at this time quotation of tender 1 is completely legal in form.

The third step is to make an illegal bid where the list deviation rate of control bid 1 is the smallest according to the law of large numbers.

Deviation rate of labor costs in list item is first considered. Taking list item 001 (build wall) as an example, bidder 1 colludes with bidding bidders 2,3,…,n to make *n-1* illegal bids. In theory, arithmetic mean *RG1~* of labor costs from item 001 in *n-1* illegal bids can be controlled so that deviation rate of labor fee in bid book 1 can theoretically reach the minimum.

In order to keep tender 1 from being scrap, bid price of bidder 1 cannot be lower than the value obtained by multiplying the arithmetic mean of all bidders by *d*. In other words, labor average arithmetic mean RG1 of *m* (including illegal ones) list item 001 must at least satisfy the following conditions.

$$d \times \overline{RG_1} \le RG_{11} \quad (1)$$

It is easy to make it according to the law of large numbers. In fact, according to the law of large numbers, when the number of legal tenders is m ≥ 30, the arithmetic mean of *m-n* legal tenders should be very close so that equation (2) can be obtained as follows.

$$d \times \overline{RG_1} \times d \times \frac{RG_{11} + (n-1)\overline{RG_1} + (m-n)\hat{RG}1}{m} \le RG_{11} \quad (2)$$

Conditions arithmetic mean $\overline{RG_1}$ of labor fee in *n-1* illegal bid list item 001 should satisfy can be worked out as follows with the help of $RG_{11} = \hat{RG}_1$.

$$\overline{RG_1} \le \frac{(\frac{1}{d}-1)m + n - 1}{n-1} \hat{RG}_1 \quad (3)$$

If the price control factor d=0.95, then

$$\overline{RG_1} \le \frac{0.05m + n - 1}{n-1} \hat{RG}_1 \quad (4)$$

where arithmetic mean $\overline{RG_1}$ of labor fee in *n-1* illegal bid list item 001 is the only one that is needed to be controlled, and its distribution is not required, which allows bidders *1, 2, ..., n* to collude with bid price, and adjust the value of labor cost in the item 001 from bids *2, ..., n*, so that these illegal bids will be legal in form.

Deviation rate of labor fee in other list items from tender 1 is controlled to be the minimum, through which illegal tender is made and deviation rate of all items in tender 1 such as material fee, mechanical fee and comprehensive fee is controlled to be minimized at the same time so that list deviation rate of tender 1 can theoretically reach the minimum value[5].

The fourth step is to make illegal bid with the help of linear weighting property of bidding rate algorithm, where deviation rate PBS1 of tender 1 is controlled to be the minimum.

In electronic bid evaluation algorithm based on deviation rate, according to cost deviation rates such as inventory deviation rate and measure deviation rate, which are directly weighted, deviation rate of tender is calculated through average conversion, and its weight is the proportion of each price component in total quotation, where the security of bid evaluation algorithm is extremely fragile and easy to be manipulated since linear weighted structure means that as long as the deviation rate of all price components (labor fee, material fee, machinery fee, comprehensive fee, measure fee and so on) in tender *1* reaches the minimum at the same time, deviation rate P_{BSi} of tender will be naturally minimized. Then bid price of the bid will be the lowest bid price with a high probability.

Meanwhile, linear weighted structure of tender deviation rate means tender deviation rate is only related to arithmetic mean of price constituents (labor fee, material fee, machinery fee, comprehensive fee), and has nothing to do with its value distribution, which offers bidders *1,2,...,n* opportunity to collude bidding quotes and adjust the value of each price component in illegal tender *2,...,n* so that these illegal tenders can be made legal in form.

III. Analysis on Surrounding Algorithm

The best empirical method to verify analysis conclusions mentioned above is to make some illegal bids based on the results of theoretical analysis, and then conduct the test through evaluation results in electronic bid evaluation system. Meanwhile, it must be pointed out that this kind of behavior is illegal, and any illegal behaviour that attempts to use the research results proposed in this paper for bidding will bear corresponding legal consequences.

If empirical method is not suitable for verification, then the next step is to perform numerical analysis on parameters in theoretical model. What's more, the key question is whether illegal bids can be guaranteed to be formally legal and not be directly rejected. In theory, linear weighted structure of tender deviation rate means tender deviation rate is only related to arithmetic mean of price constituents (labor fee, materials fee, machinery fee, general fee, measures fee, and other fees), and has nothing to do with its value distribution. However, these values may be greatly affected by the parameters of evaluation algorithm.

A. Price Control Factor

Bidding control price directly affects bidder's profit level and bidding strategy, and so does the illegal cost of illegal bidders. For example, "Tender Document (2019 Edition)" stipulates that bid price cannot be lower than the value obtained by multiplying the arithmetic mean of all bidders by d=0.95. If price control factor d=0.9 is re-specified, it can be concluded that arithmetic mean $\hat{RG}_1$ of labor costs from list item 001 in illegal tender shall satisfy:

$$\hat{RG}_1 \le \frac{(\frac{1}{d}-1)m+n-1}{n-1}\hat{RG}_1 = \frac{0.11m+n-1}{n-1}\hat{RG}_1 \qquad (5)$$

With the reducing of price control factor, the risk that illegal tender 1 becomes a scrap is increasing, and illegal cost of bidding is accordingly increasing as well.

B. Number of Tenders Participating in Bidding

The number *n* of participation in bid directly determines the illegal cost. Obviously, if a small number of bidders can control the bid evaluation result by colluding with bid price, then the bid evaluation algorithm has major safety hazards.

For example, "Bidding Documents (2010 Edition)" stipulates that if bidder's quotation is lower than bidding control price of 80% or higher than bidding control price, the bid will be directly rejected. Under this stipulation, if the quotation is $\frac{1}{0.8}=1.25$ times higher than market price, it will be very likely to exceed bidding control price. Therefore, to be formally legal, the maximum value of arithmetic mean $\hat{RG}_1$ of the labor costs from list item 001 in illegal tender shall satisfy (6).

$$\hat{RG}_1 = \frac{(\frac{1}{d}-1)m+n-1}{n-1}\hat{RG}_1 \le 1.25\hat{RG}_1 \qquad (6)$$

The solution is as follows.

$$n \ge 4\left(\frac{1}{d}-1\right)m+1 \qquad (7)$$

If price control factor d=0.95,then

$$n \ge 4(\frac{1}{d}-1)m+1 \approx 0.2m+1 \qquad (8)$$

According to the law of large numbers, it can be concluded that only 20% of bidders colluding with each other can ensure that they win the bid with great probability.

Only the situation in which *n* bidders jointly ensure that one bid is a short-listed candidate is discussed in this paper. In fact, if price component of tender 1 is added to white noise and "copied" multiple copies, the probability of achieving the bid will be greater. Therefore, electronic bid evaluation algorithm based on deviation rate has major safety hazards.

IV. Conclusion

Electronic bid evaluation algorithm whose security performance is directly related to the fairness of bid evaluation results is the core in entire electronic bidding system. Therefore, security analysis on a class of electronic bid evaluation algorithms based on weighted average of the price component factors is performed in this paper, which theoretically shows that under the condition where the algorithm parameters are known, bidder can collude quotes privately, which affects the benchmark value of price components of all the items in bidding documents so that the deviation rate of specified bids will be controlled. According to the law of large numbers, a small number of bidders can make the bid price of specified tender reach the theoretical minimum with great probability with the help of the bidding strategy collusive bidding. Further numerical analysis shows that under the conditions given in the *Bidding Documents (2019 edition),* if 20% the least required number of the bidders collude with each other for the bid, which will ensure a high probability of winning the designated bid. Moreover, results from this study will help ensure the legitimate rights and interests of bidding parties.

References

[1] Analysis of how to resolve the phenomenon of bidding and labeling in the bidding process and the post-standard supervision measures and suggestions [J]. Li Yanfeng. Tendering and bidding. 2016(09).

[2] Discussion on the phenomenon of string bidding and solution measures in project bidding [J]. Jiang Yan. Real Estate Guide. 2016(07).

[3] On the stringing behavior and prevention countermeasures in the bidding of water conservancy projects [J]. Sun Weifang. Heilongjiang Science and Technology Information. 2016(28).

[4] Analysis and identification of stringing behavior in architectural bidding [J]. Yan Jie, Lu Yuxi. Home. 2019(13).

[5] Exploration of measures to avoid the phenomenon of bidding and stringing [J]. Wang Dabin. Chinese Government Procurement. 2018(01).