International Conference on Business Research
University of Moratuwa, Sri Lanka
December 04, 2023

https://doi.org/10.31705/ICBR.2023.11

DETECTING POSSIBLE OUTLIERS IN THE COLOMBO STOCK EXCHANGE

B.M.E.P. Bandaranayake, M.K.P.L Perera, W.R. Lakmini*, N. Nageswaran, S.D. Perera, A. Gunawardana, and A.A.D.S.N.A. Perera

Department of Decision Sciences, University of Moratuwa, Sri Lanka
*lakminiwr.19@uom.lk**

ABSTRACT

Insider trading poses a significant challenge for stock markets, including the Colombo Stock Exchange, as it undermines investor confidence. The purpose of this study is to develop an innovative methodology that can effectively identify and flag suspicious transactions and investors. The proposed approach considers a multitude of parameters that influence the behavior of insider traders, which have been overlooked in current detection methods. This method mainly focuses on assessing the change in the trading behavior of investors in relation to price-sensitive corporate events compared to their past behavior and the behavior of peer investors. As an initiation, we used trade data for the year 2016- 2021 and identified potential suspicious investors and transactions. By utilizing this approach, we seek to enhance the effectiveness of identifying insider trading and mitigate its adverse effects on the market.

Keywords: Colombo Stock Exchange, Insider trading, Market manipulation, Outlier detection, Suspicious transactions

1. Introduction

The stock market is a marketplace where publicly traded company stocks and derivatives are bought and sold (Chen, n.d.). It is a dynamic system that constantly changes from time to time and provides a platform for buyers and sellers of securities to meet, interact, and transact (Asalatha, 2019). The stock market operated by the Colombo Stock Exchange (CSE) is the only share market in Sri Lanka, and it is responsible for providing a transparent and regulated environment where companies and investors come together (Colombo Stock Exchange, n.d.). The CSE is operated under the Securities Exchange Commission (SEC), which is in charge of regulating the securities market (Securities and Exchange Commission of Sri Lanka, n.d.).

Market manipulation has a significant impact on stock markets. It is the artificial inflation or deflation of the price of securities for the investor's personal gain (Hayes, n.d.). Manipulation activities create artificial price movements, mislead investors, and diminish market integrity (Sahil, 2021). Insider trading is the trading of a company's stock or other securities by individuals with potential access to nonpublic information about the company (Bainbridge, 1998). This creates disadvantages for regular investors. Insider trading can be illegal or legal depending on the time of the trade (Putniņš, 2020). Due to insider trading happening in the stock market, most investors have lost their confidence in trading and refrain from investing in stocks.

In Sri Lanka, there are three main types of insider trading identified. They are insider dealing, pump-and-dump, and front-running. Insider dealing is the buying or selling of a security by someone who has access to material nonpublic information about the security (Colombo Stock Exchange, n.d.). Insider dealing can be illegal or legal depending on when the insider makes the trade. It is illegal when the material information is still nonpublic (Alexander, 2016). Pump-and-dump is a form of securities fraud that involves artificially inflating the price of an owned stock through false and misleading positive statements to sell the cheaply purchased stock at a higher price (Sahil, 2021). Front-running is when a broker enters an equity trade with foreknowledge of a block transaction that will influence the price of the equity, resulting in an economic gain for the broker (Team, n.d.).

In this study, we aim to develop a methodology that can identify and flag suspicious transactions and investors more accurately and efficiently than the methods currently used by the CSE. In literature, most studies have considered price and quantity to identify market manipulations (Zhai et al., 2018). However, as insider traders do not leave footprints when they conduct manipulations (Alexander, 2016), we cannot detect anomalies only by considering price and quantity. Hence to address these limitations, we developed a novel methodology that incorporates a comprehensive range of parameters. These parameters encompass investors' historical trading behaviors concerning specific securities, as well as their reactions to significant corporate events. By implementing this framework, we intend to enhance the accuracy and reliability of identifying insider trading activities.

2. Literature Review

In literature, most studies on detecting market outliers in stock markets were focused on the price and quantity of a particular stock and to identify such outliers, different machine learning algorithms were used.

One of the widely used methods for outlier detection is Isolation Forest (Xu et al., 2017). The study by Liu et al. (2008) discusses how the Isolation Forest machine learning algorithm is a more efficient and effective approach for outlier detection. This study has proposed Isolation Forest as a new model-based approach to anomaly detection which focuses on anomaly isolation rather than normal instance profiling. The authors have compared this method with other anomaly detection methods such as ORCA (Outlier detection and Robust Clustering for Attributed graphs, Eswar et al., 2021), LOF (Local Outlier Factor, Cheng et al., 2019), and Random Forest (Shi et al., 2006) in terms of detection performance and processing time. The results showed that Isolation Forest outperformed other methods. Further, the authors have shown that Isolation Forest can handle high dimensional data.

Another widely used method for outlier detection is the one-class Support Vector Machine (SVM) algorithm. A study published by Chen et al. (2013) has looked into using SVMs to detect anomalies from daily and tick trading data. The article presented a novel approach to anomaly detection using a one-class SVM algorithm for unsupervised data.

The study by Mazzarisi et al. (2022) provides a machine-learning approach to support decisions in insider trading detection. Here the author proposes a clustering approach based on the k-mean algorithm (Zhang et al., 2018) to find groups of investors with similar behavior in trading a given stock. The authors then identify the potential suspects of insider trading if those investors display a discontinuous trading behavior compared to their cluster. Another recent study which has been done by Li et al. (2022) uses multi-task deep neural networks (Ruder, 2017) in insider trading detection. The model takes into account the correlation and differences between different tasks and has been shown to identify insider trading of enterprises in different industries.

Perera et al. (2014) have shown that insider trading can provide market participants with valuable information and lead to market inefficiencies. Regression analysis is used in this study to examine the relationships between insider trading volume and market returns on the Colombo Stock Exchange.

Most studies conducted in machine learning based on insider trading detection have been considered labelled data. We have identified the need to address this limitation and develop a method that is appropriate for unlabeled data. Further, it has been observed that the previous studies have only considered the price and quantity of the transactions for their analysis. From those, we can only identify anomaly transactions This novel study aims to fill this gap by applying factors that can help identify potential individuals who can be considered as insider traders. Hence in our study, we aim to develop a method which can give effective results in flagging suspicious transactions. This method incorporates factors impacting insider trader's characteristics.

3. Methodology

3.1. Data and data preprocessing

In this study, we mainly utilize trade data provided by CSE which included transaction information over 10 years, from 2012 to 2022. The data set included the details of the transaction date, transaction number, trade time, security name, IDs of buy clients, IDs of sell client, IDs of broker firm of the sell client, IDs of broker firm of the buy client, price, and quantity. In adherence to ethical concerns, the CSE did not disclose the real names of the buy-and-sell clients. Instead, CSE provided unique IDs for each individual, ensuring the confidentiality and privacy of the market participants.

From the CSE website[1], we further obtained corporate announcements, financial reports, rules, and circulars as well as regulatory filings. We have narrowed down the selection to a subset of events that have played a significant role in facilitating insider trading based

[1] https://cse.lk/

on historical data. Incidents include announcements for annual reports (annual, interim, or quarterly), cash dividends and scrip dividends rights issues, debenture issues, capitalization of reserves, rating review dealings by directors, related party transactions, mergers acquisitions, sub-division of shares, amalgamation and significant events disclosed under corporate disclosures such as potential investments and changes of company name, and among others.

However, since unique IDs were provided for security names, sell clients, and buy clients we replaced the buy client and sell client IDs with a special character to convert the data type into integer values. This change was made to reduce computer storage requirements. Additionally, we collected data by analyzing the announcement section on the CSE website, specifically focusing on gathering information related to corporate events.

3.2. Data analysis

Different stocks have different behaviors, and buyers and sellers have different trading patterns. Thus, the behavior of the insider trader depends on different factors. Figure 1 shows the characteristics considered when identifying the behavior of an insider trader. In our approach, we considered the following three criteria:

1. Relative buying and selling behavior of an investor compared to their past pattern of buying and selling,
2. Relative buying and selling behavior of an investor compared to other investors in the market, and
3. Proximity of an investor's trade to a corporate event.

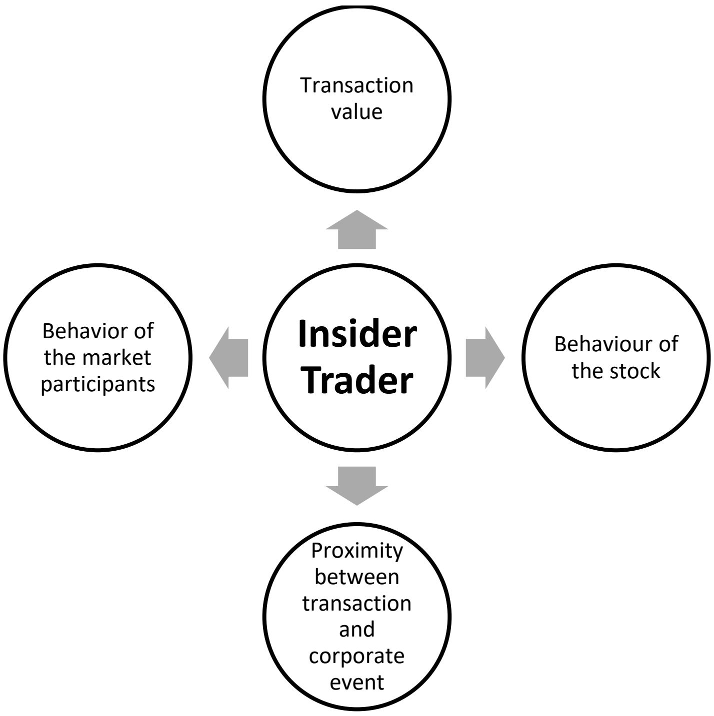

Figure 1. Characteristics considered when identifying an insider trader's behavior.

When developing the model, several assumptions were considered:

- According to Section 135 of SEC Act No. 19 of 2021, the transactions should be material to be considered as insider trade (SEC ACT, 2021). To align with that statement, we considered transactions above the value of LKR 150,000 in our analysis.
- A time duration of ±7 days from the date on which the selected transaction occurred was considered for determining the number of trades done by an investor for a selected stock.
- A time duration of ±7 days from the date on which the selected transaction occurred was considered for calculating the number of trades done by all the investors for a selected stock.
- All transactions with a trade ratio [Trade Ratio = Number of trades by an investor/ Number of trades by all investors] greater than 1 between a specific investor's transactions and the total trades by all investors were considered.
- A time duration of ±1 day from the date on which the selected transaction occurred was considered for calculating the relative deviation of a trade’s price from the average price of trades for a selected stock.
- Only the following corporate events were considered in the analysis: right issues, rating reviews, cash dividends, scrip dividends, debenture issues, corporate disclosure, credit rating, capitalization of reserves, annual reports, and interim reports.

We systematically analyze every transaction within the CSE, applying our parameters to each transaction by taking into account both the buyer and the corresponding security involved. To illustrate this process, let's examine a specific transaction as an example which is illustrated by Table 1.

Table 1. Example of a transaction made on a security by an investor.

Trade date	Security	Buy client	Price	Qty ('000)	Value ('000)	A	A/B	PROX
2021-01-29	ABCD	2972	160	5	800.0	67	15.96	0 days 00:04:06

As shown by Table 1, we determine the buyer's activity by calculating the total number of transactions by the buyer 2972 conducted for the security ABCD. Additionally, we consider the average transaction count for the security ABCD within the specific date range under examination, as further elaborated below. Moreover, we assess the proximity of this transaction to the nearest corporate event, as explained in detail later. Our approach entails the computation of these parameters for all transactions, and the values may vary based on the transaction date. Subsequently, we make informed decisions for each transaction to assess its potential for being an insider trade.

Moreover, we have defined the following four parameters in our analysis:

1. Number of trades by an investor(A), for a particular stock
2. Number of trades by all investors (B), for a particular stock
3. Trade ratio (= A/B), and
4. Proximity to a corporate event (PROX).

It should be mentioned that these criteria also can vary depending on economic conditions and the knowledge of experts. For ease of understanding, each of the parameters considered in the analysis is briefly described as follows:

1. Number of trades by an investor (A): The number of transactions made on a specific security by a specific investor was examined. Our focus is on transactions that occur within a 7-day period before and after each transaction. The objective here is to determine how frequently a particular buyer engages in transactions with a specific stock during the specified 14-day timeframe. Table 2 illustrates that the total number of transactions made on security "COMB" by an investor within a 14-day time period is 29727.

Table 2. Example of a total number of transactions made on a security by an investor within 14 days.

Trade date	Security	Buy client
2021-01-22 (Lower bound)		
2021-01-29 (Transaction date considered)	COMB	29727
2021-02-05 (Upper bound)		

2. Number of trades by all investors (B): All the investors and their interactions with specific stocks were examined. We account for an average of all the transactions done for a certain security within the specified 14-day timeframe. Table 3 shows the average number of transactions made by all investors for the security "COMB" within 14 days.

Table 3. Example of an average number of transactions made by all investors for a security within 14 days.

Trade date	Security	Buy client
2021-01-22 (Lower bound)		
2021-01-29 (Transaction date considered)	COMB	All the investors of COMB
2021-02-05 (Upper bound)		

3. Trade ratio = A/B: The main goal of evaluating this ratio is to determine the frequency of transactions by a particular buyer involving a specific security, in comparison to the frequency of transactions by other investors in the same security. Here, we aim to provide insights into the trading behavior of individual buyers and their level of involvement in a particular security as compared to their peers.

4. Proximity to corporate event (PROX): We are trying to identify the proximity of the most recent upcoming corporate events related to the specific stock for the specific transactions. This enables us to gauge the degree to which market participants are reacting to or preparing for such upcoming events.

4. Results/Analysis and Discussion

The novel approach is being developed for demonstration purposes and the primary goal is to showcase its effectiveness. Hence as an initiation to our study, we selected trade data from the years 2012 and 2022. Initially, we fed our data set into the model, performed calculations according to the methodology, and identified suspicious transactions. To identify potentially suspicious investors, we have examined transactions with a trade ratio exceeding 1. Additionally, we considered proximity to a corporate event, but the criteria for determining suspicious activity can vary based on the specific characteristics of the corporate event and the stock in question. Therefore, different thresholds are applied to different scenarios and cases.

Given the limitations of the available data, we face challenges in assessing the significance of the factors we have employed, as we lack a means to validate our results. To address this gap, we conducted a comparative analysis by referencing insider trading publications provided by the Securities and Exchange Commission. In this study, our primary emphasis was on identifying variables that have the potential to detect insider traders, our approach enables us to flag potential insider trading activities, and it should be noted that further investigations are essential to validate these suspicions.

With our model, we were able to identify transactions that could be flagged as suspicious. Table 4 shows some of the transactions which were flagged as suspicious according to the parameters related to security "XXXX" (unique code given by us due to ethical concerns) These transactions have been done before the announcement of a cash dividend.

Table 4. Suspicious transactions for security XXXX on January 29, 2021.

Trade date	Security	Buy client	Price	Qty ('000)	Value ('000)	A	A/B	PROX
2021-01-29	XXXX	10309	160	5	800.0	67	15.96	0 days 00:04:06
2021-01-29	XXXX	10309	160	2	320.0	67	15.96	0 days 00:04:06
2021-01-29	XXXX	10309	159	2	317.5	67	15.96	0 days 00:04:11

It can be seen that for an investor with the buy client ID 10309 who has completed three transactions worth LKR 1,437,500, the trade ratio (A/B) is high and the proximity to the transaction is close to the disclosure of the cash dividend. This shows how the behavior of buy client 10309 changed respectively to his past behavior and respective to other buyers closely to the announcement of price sensitive corporate event. As a result, we can determine that this buy client and the transactions appear suspicious.

Table 5 illustrates some of the transactions related to "YYYY" security and transactions done before the disclosure of the rating review. It can be seen from Table 5 that the transactions were done by the buy client ID 6395 within an hour before the disclosure of corporate review that the position of YYYY has been upgraded. The trade ratio is high and the proximity to the event is also close. Hence, we can label this series of transactions and the investor as suspicious.

Table 5. Suspicious transactions for security YYYY on January 29, 2021.

Trade date	Security	Buy client	Price	Qty ('000)	Value ('000)	A	A/B	PROX
2021-01-29	YYYY	6395	160	3.600	576	56	14.9	0 days 00:13:23
2021-01-29	YYYY	6395	160	3.775	604	56	14.9	0 days 00:13:26
2021-01-29	YYYY	6395	160	4.125	660	56	14.9	0 days 00:16:47
2021-01-29	YYYY	6395	160	5.875	940	56	14.9	0 days 00:17:03
2021-01-29	YYYY	6395	160	4.125	660	56	14.9	0 days 00:17:03
2021-01-29	YYYY	6395	160	5.875	940	56	14.9	0 days 00:17:23
2021-01-29	YYYY	6395	160	4.125	660	56	14.9	0 days 00:17:23
2021-01-29	YYYY	6395	165	2.000	330	56	14.9	0 days 00:52:04
2021-01-29	YYYY	6395	165	5.567	918.555	56	14.9	0 days 00:52:04
2021-01-29	YYYY	6395	165	13.00	2145	56	14.9	0 days 00:52:04
2021-01-29	YYYY	6395	165	12.00	1980	56	14.9	0 days 00:52:04
2021-01-29	YYYY	6395	165	10.00	1650	56	14.9	0 days 00:52:04
2021-01-29	YYYY	6395	165	2.000	330	56	14.9	0 days 00:52:04

To assess the effectiveness of our approach, we further examined insider trading activities reported by the SEC and determined if our findings align with the disclosures made by the exchange. As part of the evaluation, we focused on an incident reported in February 2016 concerning a specific listed company with a unique security code assigned to LLL.

The SEC has stated this as “a violation of Section 32 (1) of the Securities and Exchange Commission of Sri Lanka Act No. 36 of 1987, by purchasing 344, 500 LLL shares on February 11, 2016, prior to the release of the Interim Financials of LLL for the nine months ended on December 31, 2015, that was released to the market on February 15, 2016” (SEC SL, n.d.). Through our methodology, we have captured that investor as a suspicious investor. Table 6 illustrates the transactions we flagged as suspicious related to LLL. Therefore, with these results, we can confirm that the model we developed has the potential to flag suspicious transactions and investors.

Hence, we hope by implementing this framework we can take a substantial step towards mitigating the limitations of the generic outlier detection and work towards building a more robust method capable of flagging suspicious activities.

Table 6. Suspicious transactions for security LLL on February 11, 2016.

Trade date	Security	Buy client	Price	Qty ('000)	Value ('000)	A	A/B	PROX
2016-02-11	LLL	7555	6.2	52.5	325.5	8	3.4	3 days 11:27:57
2016-02-11	LLL	7555	6.2	39.5	244.9	8	3.4	3 days 11:27:57
2016-02-11	LLL	7555	6.2	250	1550	8	3.4	3 days 11:27:57

5. Conclusion and Implications

In recent years, insider trading activities have become increasingly sophisticated. As a result, detecting such activities is a challenging task that may not yield a comprehensive and precise outcome. Hence our key focus is combating insider trading effectively.

We have implemented a novel approach focusing on trading data from a specific day and examining transactions within a 14-day time period. We took into account the investors' historical trading patterns and their overall activity in the securities market. Furthermore, we paid close attention to significant corporate events related to securities. Consequently, we have successfully pinpointed certain transactions that exhibit suspicious characteristics and have identified the potential individuals who can be involved.

This novel method offers distinct advantages when compared to the previously discussed methods. While Isolation Forest and One-Class Support Vector Machines are powerful for outlier detection models using price and volume. Moreover, clustering methods also provide an efficient method of grouping investors with similar behaviors and identifying their patterns. But our approach provides a special focus on different factors such as behavior of the stock, transaction value, behavior of market participants, and proximity to corporate events.

This approach appears to be an effective way of flagging suspicious investors and their activities. We believe that implementing this method could effectively aid the CSE in identifying individuals associated with insider trading activities. With the current economic and trading patterns, the thresholds can be changed according to their requirements. Though the accuracy of identifying suspicious individuals may not be able to quantify, this approach can indeed serve as a highly effective way to significantly narrow down the pool of individuals warranting investigation for insider trading activities.

Acknowledgement

The authors would like to thank the Colombo Stock Exchange for providing the necessary data for this study as well as sharing their expertise for the study to be a success.

References

Alexander, R. C. H. (2016). *Insider dealing and money laundering in the EU: law and regulation*. Routledge.

Asalatha, R. (2019). Stock Markets: An Overview and A Literature Review. *MPRA, Paper*, (101855).

Bainbridge, S. M. (1998). Insider trading: an overview. *Available at SSRN 132529*.

Chen, J. (n.d.). What Is the Stock Market, What Does It Do, and How Does It Work? Retrieved on May 22, 2023, from https://www.investopedia.com/terms/s/stockmarket.asp

Chen, Y., Qian, J., & Saligrama, V. (2013, May). A new one-class SVM for anomaly detection. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 3567-3571). IEEE.

Cheng, Z., Zou, C., & Dong, J. (2019, September). Outlier detection using isolation forest and local outlier factor. In *Proceedings of the conference on research in adaptive and convergent systems* (pp. 161-168).

Colombo Stock Exchange. (n.d.). Retrieved on May 30, 2023, from https://www.cse.lk/

Eswar, S., Kannan, R., Vuduc, R., & Park, H. (2021). ORCA: Outlier detection and Robust Clustering for Attributed graphs. *Journal of Global Optimization*, *81*(4), 967-989.

Hayes, A. (n.d.). *Investopedia*. Retrieved on May 30, 2023, from https://www.investopedia.com/terms/m/manipulation.asp

Mazzarisi, P., Ravagnani, A., Deriu, P., Lillo, F., Medda, F., & Russo, A. (2022). A machine learning approach to support decision in insider trading detection. *arXiv preprint arXiv:2212.05912*.

Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation Forest. *2008 Eighth IEEE International Conference on Data Mining*, 413–422. https://doi.org/10.1109/ICDM.2008.17

Li, G., Li, Z., Wang, Z., & Zhang, K. (2022). Identification of Insider Trading in the Securities Market Based on Multi-task Deep Neural Network. *Computational Intelligence and Neuroscience*, *2022*.

Perera, K. L. U. D., & Nimal, P. D. (2014). Information Content of Insider Trading Volumes in Colombo Stock Exchange.

Putniņš, T. J. (2020). An overview of market manipulation. *Corruption and Fraud in Financial Markets: Malpractice, Misconduct and Manipulation*, *13*.

Ruder, S. (2017). An overview of multi-task learning in deep neural networks. *arXiv preprint arXiv:1706.05098*.

Sahil, M. (2021). Stock Market Manipulation. *Available at SSRN 3903191*.

Securities and Exchange Commission of Sri Lanka. (n.d.). Retrieved on May 30, 2023, from https://www.sec.gov.lk/

SEC SL (n.d.). *Securities and Exchange Commission of Sri Lanka*. Retrieved on May 30, 2023, from https://www.sec.gov.lk/insider-dealing/

Securities and Exchange Commission of Sri Lanka Act, No 19 of 2021. (2021). *Published as a Supplement to Part II of the Gazette of the Democratic Socialist Republic of Sri Lanka on September 24, 2021.*

Shi, T., & Horvath, S. (2006). Unsupervised learning with random forest predictors. Journal of Computational and Graphical Statistics, 15(1), 118-138.

Team, C. (n.d.). *The illegal practice of purchasing a security based on advanced non-public information*. Retrieved on May 30, 2023, from https://corporatefinanceinstitute.com/resources/capital-markets/front-running/

Xu, D., Wang, Y., Meng, Y., & Zhang, Z. (2017). An Improved Data Anomaly Detection Method Based on Isolation Forest. *2017 10th International Symposium on Computational Intelligence and Design (ISCID)*, 287–291. https://doi.org/10.1109/ISCID.2017.202

Zhai, J., Cao, Y., & Ding, X. (2018). Data analytic approach for manipulation detection in the stock market. *Review of Quantitative Finance and Accounting*, *50*, 897-932.

Zhang, Z., Zhang, J., & Xue, H. (2008, May). Improved K-means clustering algorithm. In *2008 Congress on Image and Signal Processing* (Vol. 5, pp. 169-172). IEEE.