Mazzarisi et al. *EPJ Data Science* (2024) 13:66
https://doi.org/10.1140/epjds/s13688-024-00500-2

EPJ Data Science
a SpringerOpen Journal

RESEARCH Open Access

A machine learning approach to support decision in insider trading detection

Piero Mazzarisi[1,2*†], Adele Ravagnani[2†], Paola Deriu[3], Fabrizio Lillo[2,4], Francesca Medda[5,3] and Antonio Russo[3]

*Correspondence:
piero.mazzarisi@unisi.it
[1]Dipartimento di Economia Politica e Statistica, Università di Siena, Siena, Italy
[2]Scuola Normale Superiore, Pisa, Italy
Full list of author information is available at the end of the article
†Equal contributors

Abstract

Identifying market abuse activity from data on investors' trading activity is very challenging both for the data volume and for the low signal to noise ratio. Here we propose two complementary unsupervised machine learning methods to support market surveillance aimed at identifying potential insider trading activities. The first one uses clustering to identify, in the vicinity of a price sensitive event such as a takeover bid, discontinuities in the trading activity of an investor with respect to her own past trading history and on the present trading activity of her peers. The second unsupervised approach aims at identifying (small) groups of investors that act coherently around price sensitive events, pointing to potential insider rings, i.e. a group of synchronised traders displaying strong directional trading in rewarding position in a period before the price sensitive event. As a case study, we apply our methods to investor resolved data of Italian stocks around takeover bids.

Keywords: Machine learning; Insider trading; Market abuse; Unsupervised learning; Statistically validated networks

1 Introduction

In financial markets, market abuse refers to an intentional conduct that violates market integrity and natural demand-supply dynamics through misuse of privileged information, unlawful disclosure of inside information, unfair trading practices, price manipulation, creation of unfair market conditions, and deception of market players, to name but a few examples.

In the literature, the area of market abuse covers a number of different conducts that nonetheless could be grouped into two main categories: 1) insider dealing: the act of utilizing inside information in order to make, change, or cancel deals, or to encourage a third-party to deal using this knowledge and unlawful disclosure of inside information, by releasing information without correct permissions; 2) market manipulation, subdivided in trade base manipulation, action trade manipulation, or information based manipulation: in other terms an umbrella for a series of actions which distort market performance.

In this paper the focus is on insider trading, which is maybe the simplest market abuse conduct to conceive, but also one of the most widespread and difficult to enforce, since it is recognized as an illicit *a probatio diabolica* for the difficulties inherent in the search for a *smoking gun*.[1] Knowing in advance how the price will likely move in response to the release of the information to the market (a.k.a. price sensitive event (PSE), such as, for example, the announcement of a takeover bid), can be easily exploited to make a profit. Such a type of practice is prohibited or criminalized in most jurisdictions around the world [9]. However, rules are specific of each country and efforts in persecuting insider trading vary considerably.

The "proof" and the subsequent imposition of a sanction (either administrative or criminal) to a trader that has operated as an *insider* is however a complex process, involving many steps: (i) the detection of alerts pointing to anomalies that appear attributable to abusive behavior, (ii) the concrete assessment of the allegedly suspicious conduct with respect to possible rationale that may have supported the strategy under analysis, (iii) the investigation phase aimed at gathering evidence and clues of the abusive conduct, and (iv) the subsequent legal trial to confirm the fact that the unlawful conduct was committed.

The scope of this work concerns the first step, i.e. the analysis of the trading behavior of investors in the presence of a price sensitive event, and, in part, the second one. The goal is defining a support decision system based on an unsupervised machine learning methodology that is able to provide an indication of whether the trading behavior of an investor or a group of investors is anomalous or not, thus supporting the monitoring and surveillance processes by the competent Authority and the assessment of the conduct.

In fact, the assessment of the trading behavior of an investor, alleged to be suspicious in terms of fault of abnormality remains a crucial point. In the context of insider trading, the anomaly arises because of the exploitation of the inside information, which likely results in some discontinuity for the trading behavior, when compared with some proper benchmark (e.g. the trading behavior of the same investor in the past or the trading activities of other market participants). For example, the trading of only one stock *could* be suspicious when it happens: (i) for a specific time frame proximate to (and preceding) the dissemination of a price-sensitive news item, (ii) by assuming a rewarding position in relation to the price movement, and (iii) in the absence of other similar investments (i.e. in the same security over a longer timeframe or in securities similar in terms of capitalization, sector, risk, asset class, and for countervalues at least comparable in size to the investment under analysis).

Moreover, how much the behaviour of a trader is anomalous is key for the competent Authority. Rather than splitting all investors into two classes, i.e. anomalous or not, it seems therefore more reasonable to assign a score to each trader as a measure of how much her behavior differs from the most suspicious one. Three considerations guided the design of our methodology:

- First, anomalies need to be studied in a dynamical framework, in which any deviation from ordinary trading of an investor immediately before a price sensitive event must be characterized with respect to her activity in a past reference period when no anomalies are assumed. Such a comparison should include not only the trading in the asset under investigation but in the whole market.
- Second, the deviations from ordinary trading of an investor should be compared with the trading behavior of her peers, defined as the investors who, in the past reference

[1] As in a murder, finding the killer with the gun immediately after the fire shoot is a proof of guilt rarely available to investigators.

period, trade similarly to the investor under investigation. In other words, a modification in the trading behavior before a price sensitive event might require the assessment of the existence of a similar change by a (large) group of peers. For example, such cases might be due to some public leakage of information or some market dynamics which lead the investor and her peers to modify the trading.

- Third, when we look at small groups of traders, the synchronisation of a group in buying or selling a stock may be the signal of the spreading of confidential information within the group (*insider ring*). The study of such synchronised behavior may further reveal when the confidential information started to spread before the price sensitive event. For instance, the appearance of synchronised buyers at some date before the announcement of a takeover bid could permit to infer from market data when the confidential information has been formed, then exploited by some insiders.

The inference of such dynamic patterns can help the competent authority by providing a better understanding of the market dynamics and can help in the identification of individual and collective suspects of insider trading activity.

Our contribution[2] In this paper, we propose an unsupervised machine learning approach to the problem of insider trading detection, that leverages the availability of a rich dataset including all daily transactions of all (anonymized) investors in the Italian stocks, from the beginning of 2019 to the third quarter of 2021. The methodology is based on two standard and well-known techniques, i.e. the k-means clustering algorithm [23] and the statistically validated co-occurrence networks [43], that are generalized in a dynamic framework for the study of anomalies in the stock market. The anomaly detection approach proposed here is suited for the specific application to insider trading analysis. The methodology has been built following the analysis and the reasoning adopted by the Italian competent Authority (Consob) in the first steps of the investigation for insider trading, also taking into account the expertise acquired by the analysis of the output cases already discussed in court.

In the first instance, we propose a dynamic clustering approach based on the *k-means* algorithm [23]. The main idea is representing market data in a Euclidean space, in such a way the trading activity on one stock by an investor in a given time period is described by a point, whose coordinates are some chosen trading features: (i) the total bought or sold *turnover*, (ii) the trading concentration (also called *magnitudo*), and (iii) the financial *exposure* on the stock. The clustering of investors in such feature space at consecutive times within a given reference period allows us to accurately describe the usual trading behavior. The typical trading behavior is associated with clusters persistent in time. Such a characterization of the market provides the ideal context in which we can assess the discontinuity of the investors' behavior, if any, *both with respect to her past trading behavior and with respect to the trading behavior before the PSE of her peers*. For example, when a price sensitive event for the stock is within the reference period, some operations that place an investor in a rewarding position never assumed before, i.e. the switching towards the cluster she never belonged to, can be classified as anomalous. The goal is then to reveal all discontinuities towards a rewarding position with respect to the price sensitive event

[2] The methodology presented in the paper was conceived in 2022 for the purpose of developing a proof of concept. It is, in no way, a tool used in the analysis and investigations carried out by Consob. The methodology may possibly constitute the future one of the tools to help and support the preliminary analysis and detection activities more efficiently. Any subsequent enforcement activity will, in any case, be based on the broader set of information that is gathered in the course of investigations and other possible types of analysis.

and rank them according to some anomaly score, thus providing the competent authority with a list of suspicious behavior to be further analyzed.

The second clustering approach we employ is based on the so-called Statistically Validated Networks (hereinafter SVN). This is an unsupervised learning method that was introduced in [43] and further employed in other works such as [42]. This method aims at detecting unexpected structures in the projection of a bipartite network that represents a complex system. Similarly to [42], we start from a bipartite network of investors-trading days, which represents the trading activity for a specific asset in the Italian Stock Exchange. The network is then projected into a network of traders only. In the new projected network, a *statistically validated* link captures the synchronicity in the trading of two agents, i.e. the two nodes. In such a validated projected network, clusters of nodes represent groups of traders who are synchronized in the kind and time of trading actions. Given these clusters, our analysis focuses on a time window before the PSE and the goal is to detect groups of traders suspicious of market abuse. For instance, investors who are synchronized and trade in a rewarding position right before a PSE can be interpreted as suspicious.

We aim to highlight the extensive array of advanced methods available in the scientific literature for addressing various challenges in unsupervised clustering over the years. Notable examples include generalizations of k-means such as probabilistic mixture models, time series clustering methods like correlation-based hierarchical networks or the dynamic time warping algorithm, and Graph-Based clustering approaches, see [24] for a reference. In this study, we propose employing simple methods like k-means for the purpose of ensuring explainability[3] in detecting anomalous behavior among investors, alongside a time series clustering approach like SVN to capture complex patterns of synchronicities within data, serving as an initial step in identifying "insider ring" anomalies. In a companion paper [38], we explore more advanced techniques, such as autoencoders, to address the same anomaly detection problem of insider trading, albeit completely data-driven. Interestingly, a comparative analysis of the findings of the two papers reveals a high degree of agreement, underscoring the robustness of the proposed methodology (see below).[4]

Finally, given the different nature and goals of the k-means and SVN clustering approaches, these two methods can generally lead to different results, potentially revealing the presence of single insiders and/or insider rings. Interestingly, one method can be complementary to the other, especially in the case of a large number of suspects. Two pieces of evidence are directly signaled by the overlapping between the outputs of the two methods, obtaining the most suspicious ones among all potential insiders. Even more interestingly, the SVN-based approach can be used in a *human-in-the-loop* manner [45]: a trader z who is considered as suspicious by the k-means method can be used as seed in the statistically validated network of investors to find other possible suspects (neighbors of z in the SVN displaying synchronized trading behavior with z) or even potential insider rings (cliques of connected nodes in the SVN displaying high synchronization).

[3] Note that from the point of view of the competent Authority, the result of a possible insider dealing (in terms of anomalous behavior in investors' trading history) should be explainable and understandable even to professionals without a strong mathematical/statistical background, as it might be in court.

[4] Notice that a direct validation of results is unattainable in the unsupervised learning context.

The coupled use of the two methods needs to be intended as a support tool to the competent authority in the first steps of insider trading analysis. As such, the proposed methodology contributes to the field of *human-centered* decision support systems [7].

Paper organization. The paper is organized as it follows. Section 2 reviews the European regulation and the scientific literature. Section 3 describes the novel methodology for the identification of individual and collective potential insider trading activities. Section 4 presents the dataset we use in our empirical analysis and Sect. 5 presents the results obtained with our method, with a special focus on one PSE. More insights on the methods, robustness analysis, and results for other PSEs are presented in the Appendix Sects. A and B. Finally, Sect. 6 draws some conclusions and presents some suggestions for further work.

2 Literature review

Insider trading has been of tremendous interest in the financial literature starting from the pioneering work of [28] because of the clear connection between information and price dynamics, opening to a general proof of equilibrium in the market. The problem of detecting illegal insiders is however less explored, with few exceptions. The first work in this direction is by [33], who uses data on illegal insider trading from the US Securities and Exchange Commission (SEC) to show that almost half of the pre-announcement price run-up observed before a takeover occurs in the days of insiders' operations, thus providing a potential alert of insider activity. Extending the work of [19, 33] use a similar data set from SEC to characterize the trading behavior of an insider. They find evidence that the size of an illegal insider's trade depends mainly on the value of the private information at hand, but also on the risk to be detected and the expected penalty. By exploiting a unique hand-collected data set based on insider trading cases filed by SEC, [2] provides clear evidence that social relationships underlie the illegal spreading of inside information, thus supporting the existence of insiders operating in a coordinated way, the so-called *insider rings*. The clear limitation of available data sets has slowed down the study of insider trading detection. Based on public data, many works have proposed alerts of insider activities. [27] show that abnormal returns prior to a PSE are likely associated with insider trading. [37] devise an econometric methodology to test for such an association. Similar conclusions are drawn for the Athens Stock Exchange by [41]. [17] explores early detection of insider trading in option markets: he leveraged both supervised and unsupervised techniques using option, stock and news data. It emerges that the implied volatility has a major role in identifying suspicious cases of insider trading.[4] provide evidence of insider trading by looking at the abnormal volumes for the out-of-the-money call options. In this context, all the works point to support the intuition that insider trading is statistically associated with abnormal activity in the market, driven by inside information.

Other types of market abuses are investigated in the literature, see [44] for a review. [34] proposes a general framework for market abuse detection, which is based on modeling security's trading volumes, returns and market's concentration as diffusion processes. Detecting stock-price market manipulation is the goal of [21] and [30]. They tackle the problem via supervised learning algorithms such as Neural Networks, Support Vector Machines, Decision Trees, k-Nearest Neighbor [24]. The data set employed by [21], is made up of cases pursued by SEC and the features used for models' training include returns, volumes, volatility. On the other hand, [30] apply the methods on both daily and tick trading data of manipulated stocks according to the China Security Regulatory Commission.

Also [35] focuses on market manipulation but related to cryptocurrencies and by employing trade records. The authors analyze pump-and-dump and crowd pump manipulation schemes; in order to identify them, they propose classifiers, based on Random Forest [24].

From a methodological perspective, insider trading detection naturally fits into the framework of anomaly detection, which basically amounts to identify data instances that cannot be associated with normal behavior and that are rare in the data set. The goal is to define a region of the features' space characterizing normal observations; observations that do not lie in this region are defined as anomalies [13]. Identifying this normality region is not straightforward: the boundary between normal and anomalous behavior is not always sharp, behavior that is actually anomalous could be disguised in order not to be identified, the definition of normal behavior could be time-varying and it is strongly dependent on the application domain, it is difficult to distinguish noise from anomalous behavior [13]. From a practical point of view, there are three main aspects that determine the formulation of an anomaly detection method: the nature of the input data, the type of anomaly, the availability of data labels, and the desired output of the technique. Data instances can be of various types (binary/categorical/continuous, univariate/multivariate) and independent among them or related to each other, as is the case of time series and sequences, spatial data, and graph data, for which ad hoc methodologies have to be employed [1, 3]. Concerning the type of anomalies, the standard case is represented by *point anomalies*, which are single elements identified as anomalous; they could be *global* or *local* depending on whether the entire features' space or a specific region of it is considered [20]. Interestingly, there are cases when an element can be seen as normal, but when a given context is taken into account, it turns out to be an anomaly. We refer to this type as *contextual anomalies* [13], also referred to as *conditional anomalies* [40]. It may happen, for instance, that an investor has operated on a stock and, without context, such an operation looks similar to other operations in the market. However, when compared to the own past behavior of the investor or to the operations of other investors, some discontinuity or synchronization patterns may be revealed. As such, the operation could turn out to be identified as anomalous. Contextual anomalies problems can be tackled by algorithms for point anomaly detection once the context is included as a new feature. Finally, we could have data instances that are normal if considered individually, while they are anomalous together: they are the so-called *collective anomalies* [13], which occur when data display some dependence structure. Availability of data labels plays a crucial role in the choice of the approach: supervised when each observation is labeled as normal or anomalous, or unsupervised when no labels are provided. The latter is the case of interest in this paper. Typically, outputs of anomaly detection algorithms associate a score to each observation, thus quantifying the magnitude of the anomaly. Setting a suited threshold, the ranked list of anomalies can provide labels for each data instance.

Our insider trading detection approach is an unsupervised clustering methodology which aims at identifying *contextual anomalies* by relying on continuous and multivariate data instances. The (anonymized) transaction reporting database we are provided with is similar to the SEC data employed by [2, 19, 33]: it tracks the daily activity of all investors trading in the Italian Stock Exchange. However, [2, 19, 33] have different goals, as illustrated above, and adopt supervised approaches. We aim at identifying the identity of investors with suspicious activity related to PSEs and we are not provided with labels. The goal and the unavailability of labels come with two other peculiarities of our method: its simplicity and complete explainability.

The majority of the works present in the literature such as [4, 17, 21, 27, 30, 35, 37, 41] are based on stock-prices and volumes data, or trade records: the activity of a given investor cannot be tracked. This difference prevents us to us them as benchmarks for our work. However, our approach could be employed as a validation tool for that kind of methods. Indeed, given the presence of insiders, it is not obvious that their anomalous activity emerges at a macroscopic level, especially for highly liquid assets. Our method could shed light on the following research question: can insider trading be always detected by studying the dynamics of macroscopic quantities such as stock-price returns, volumes and volatility? This is left for further research.

2.1 European regulation

The European legislator defines *market abuse* as any unlawful conduct on the financial markets involving the commission or attempted commission of insider dealing, i.e. the unlawful disclosure of inside information or the recommendation to others to engage in insider dealing, as well as market manipulation. Such conducts, preventing full and effective market integrity and compromising public confidence, are expressly prohibited and administratively sanctioned, leaving Member States the possibility of also imposing criminal sanctions.

In Europe the current legal framework of reference is represented by Directive 2014/57/EU (so-called MAD II), European Regulation 596/2014 (so-called MAR) as well as a bunch of Delegated Regulations supplementing the MAR Regulation with regard to regulatory technical standards on several aspects, see for example [14, 18]. The European legislator envisaged equipping the competent authorities of each Member State with adequate tools, powers and resources to ensure the effectiveness of supervision. In addition, the European provisions concerning market abuse require criminalisation of the most serious misconduct leaving to national legislators the power to criminalise certain misbehaviours. The Italian case is noteworthy because yet the implementation of the first Market Abuse Directive (MAD), in 2005, was a hook to introduce severe administrative sanctions in addition to the pre-existing criminal penalties [15]. Therefore, since long ago, in Italy Consob's supervisory activities in this area may give rise to both administrative and criminal sanctioning proceedings, the latter by reporting the detected conducts to the Judicial Authority.

3 Methodology

In this Section, we present the methodology based on two different clustering approaches, which allow us to find groups of investors with similar behavior in trading a given stock. We apply it to the problem of contextual anomaly detection for insider trading. In the first method, groups are identified by partitioning investors depending on some trading features (turnover, magnitudo, and exposure). In the second method, investors are identified as similar when they trade in a coordinated way (e.g. they buy or sell on the same days), displaying some significant correlation in the trading activity. Finally, we combine the two methods together to devise a human-in-the-loop procedure for detecting contextual anomalies as a result of both the discontinuity and the coordination of the trading behavior of investors in the presence of a price sensitive event. The application is then shown for a particular case study in the next section.

3.1 Method based on k-means clustering

The *k-means* clustering algorithm [23] is an unsupervised learning method for finding clusters of points in a set of unlabeled data that lie in a Euclidean space. Each data point x is a *n-tuple* of real numbers characterizing the trading features of each investor. The method aims to partition N data points into K clusters, each point belonging to the cluster with the nearest mean, which is named the *centroid* of the cluster. Such a particular point in the feature space summarizes the average characteristics of all points in the cluster. The output of the clustering algorithm is a partition of the feature space into *Voronoi cells*. More specifically, given a set of data points $\{x_i\}_{i=1,\dots,N}$ with $x_i \in \mathbb{R}^n$ and the number K of clusters in which we aim to partition the feature space, the $k-means$ algorithm finds K sets $\mathcal{S} = \{S_k\}_{k=1,\dots,K}$ of points corresponding to K centroids, in such a way the squared distance (variance) of points from the centroid within each cluster is minimized, i.e.

$$\text{argmin}_{\mathcal{S}} \sum_{k=1}^{K} \sum_{i \in S_k} ||x_i - c_k||^2 = \text{argmin}_{\mathcal{S}} \sum_{k=1}^{K} |S_k| \text{Var}(S_k) \tag{1}$$

where $c_k = \frac{1}{|S_k|} \sum_{i \in S_k} x_i$ is the mean of points belonging to S_k, whose cardinality is $|S_k|$. The vector c_k defines the position of the centroid in the feature space.

The problem in Eq. (1) is in general computationally difficult (NP-hard), however, there exist several heuristic methods converging quickly to a local minimum. Here, we use a two-phase iterative algorithm, see e.g. [31], that minimizes the within-cluster variances by alternating *batch updates* (reassigning points to their nearest cluster centroid, all at once, followed by recalculation of cluster centroids) and *online updates* (reassigning points one by one only if it reduces the within-cluster variances).

3.1.1 Trading features

Let N be the number of investors, M the number of stocks, and S the number of trading days. For a given stock $j \in \{1, \dots, M\}$ and a given time window $\Delta = (t, t + D]$ with $D > 0$, e.g. a week ($D = 5$) or a month ($D = 20$), each investor i is associated with a triple of features (i.e. $K = 3$), summarizing her trading activity during that period:

1. the *signed turnover*, namely the aggregated Euro turnover of operations within the period, with positive (negative) sign for a net buying (selling) volume,

   $$A_i^{(j)} = \sum_{t \in \Delta} A_b(i,j,t) - \sum_{t \in \Delta} A_s(i,j,t),$$

   where $A_b(i,j,t)$ and $A_s(i,j,t)$ are the total Euro turnover bought and sold by investor i for stock j in day t;
2. the *magnitudo* (or portfolio concentration), namely the relative Euro turnover (without sign) traded in the stock with respect to the total amount traded in any stock,

   $$a_i^{(j)} = \frac{\tilde{A}_i^{(j)}}{\sum_{j=1}^{M} \tilde{A}_i^{(j)}}, \quad \text{with} \quad \tilde{A}_i^{(j)} = \sum_{t \in \Delta} A_b(i,j,t) + \sum_{t \in \Delta} A_s(i,j,t);$$

3. the *maximum exposure* (in Euro) to the stock in the time window,

$$E_i^{(j)} = \left(\max_{t \in \Delta} |\alpha_t| \right) \mathrm{sign}(\alpha_{\tilde{t}})$$

where α_t is the position[5] (in Euro turnover) of investor i in stock j at day t and $\tilde{t} = \mathrm{argmax}_{t \in \Delta} |\alpha_t|$.

For a stock j, a data point having the three features as coordinates in an Euclidean space $\mathbb{R}^3$ describes the trading behavior of an investor i, aggregated over a time window Δ.

While the magnitudo (i.e. relative turnover) $a_i^{(j)}$ is by definition between 0 and 1, both the aggregated turnover and the exposure depend strongly on the size of the trader's portfolio. In absence of any normalization, the clustering algorithm would tend to group together investors of similar sizes, independently from their idiosyncratic trading behavior. Hence, in order to avoid such a spurious effect, it is convenient to re-scale each value within the interval $[-1, 1]$, as follows. First, let us consider the whole period $[t_1, t_S]$, containing $m > 1$ time windows of length D, possibly overlapping.[6] Then we compute the three features for each investor i within each time window, i.e. $x_i \equiv \{A_{i,s}^{(j)}, a_{i,s}^{(j)}, E_{i,s}^{(j)}\}_{s=1,\dots,m}$. Finally, we define the re-scaled signed turnover and the re-scaled maximum exposure as

$$A_{i,s}^{(j)} \hookleftarrow \frac{A_{i,s}^{(j)}}{\max_s |A_{i,s}^{(j)}|},$$

and

$$E_{i,s}^{(j)} \hookleftarrow \frac{E_{i,s}^{(j)}}{\max_s |E_{i,s}^{(j)}|}.$$

In this way, we are able to map the trading activity of an investor i on a stock j in a given time window to three normalized features defined in the space $[-1, 1] \times [0, 1] \times [-1, -1]$. Finally, the clustering analysis for the time window Δ includes only those investors that are active for at least one day in Δ.

A pictorial example is shown in Fig. 1, where a generic investor, i.e. the black dot, belonging to some cluster is identified by three coordinates in the feature space, namely the signed turnover, the magnitudo, and the maximum exposure as defined above.

Such an approach is particularly useful when we aim to characterize the trading behavior of an investor within a given time window, with respect to the typical operations done in the whole reference period. In particular, when t_S is the date of a price sensitive event for the stock j, it is possible to compare the trading behavior of an investor in the previous D days with the past. From the point of view of the regulator, this serves to verify whether the current behavior of an investor is coherent or not with its own past. A graphical explanation of representative trading behaviors as captured by the clustering analysis, together

[5]In general, information on the precise composition of the portfolio of each investor is not available, but only the daily aggregated information on her operations. As a proxy of asset positions we assume that they are zero on Jan. 1, 2019, then each position at day t is obtained by aggregating bought and sold turnovers from Jan. 1, 2019 up to t.

[6]An example is a reference period of one year, with twenty days long windows (i.e. one business month), which are rolled week by week starting from January.

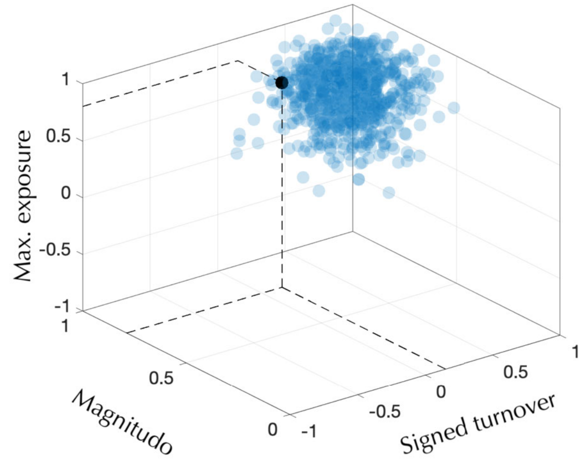


Figure 1 A pictorial representation of the trading activity (black dot) of an investor by using the three features, i.e. signed turnover, magnitudo, and maximum exposure, as defined in the main text

with an intuition on such discontinuity pattern, are shown in the Appendix Sect. A.1. Finally, notice that the trading features are devised to capture various types of discontinuities associated with turnovers (volumes in Euro), with a similar intuition by [4].

It is worth noticing that the choice of three features has been motivated by two main reasons: (i) the modeling choices trace the rationale of insider trading investigation (by the competent Authority), and (ii) outputs are directly explainable. Moreover, explainability is even easier since we can obtain a graphical representation of the trading behaviors of investors (in the case of three features only).

3.1.2 Dynamic clustering

In the case of time-resolved data, a natural question is how to generalize the clustering method to capture the temporal evolution. Applying the *k-means* algorithm to consecutive time windows in a row is straightforward. The challenge is preserving the stability of clusters to achieve a coherent description of the system under investigation. As such, the solution depends on the particular problem. For example, [29] proposes a dynamic *k-means* clustering method for anomaly detection of various Internet threats. The authors assume that the position of centroids is fixed over time but subject to some systemic rearrangement (signaling an anomaly) when the variance of data within a cluster changes above a given threshold value. This solution paves the way for the anomaly detection of Internet attacks. Here we devise a more flexible approach, allowing the centroids to change positions smoothly over time but preserving the patterns of stability, if any.

The *k-means* clustering algorithm is applied to the set of features computed within each time period Δ, by rolling the window over the whole period $[t_1, t_S]$. Two consecutive windows are overlapping in order to smooth the evolution of clusters. In the empirical application below, we consider a one-month-long window, shifted week by week. For each window, the output of the method is the positions of centroids, together with the vector of labels, indicating which investors belong to each cluster. Unfortunately, the convergence to a global minimum is not ensured, possibly resulting in very different centroids' positions moving from one time period to the next one. Moreover, any permutation of labels does not change the loss function in Eq. (1). However, when we roll the time window over the whole period, it is convenient to associate two clusters in a row with the same label when both of them are formed in large part by the same investors.

In order to recover a pattern of stability for clusters, the centroids' positions that solve the minimization problem at a given time are then used as starting point of the clustering method when the time window is shifted one step forward.[7] Once the new centroids' positions are found in the new time window, we consider all possible permutations of labels over the set $\{1, \dots, K\}$ and compute the Jaccard similarity coefficient [24] between the current clusters and the previous ones. This is a metric measuring the overlap between the two sets of elements, which is equal to one when each element is also in the other set and vice versa, zero if not. Any other value between zero and one suggests a partial overlapping between the two clusters, such that the larger is the overlapping, the larger is the coefficient. The final assigned labels are the ones maximizing the Jaccard similarity. In this way, cluster stability tends to be preserved over time.

3.1.3 Identification and classification of potential insiders

The method identifies the potential insider investors once the dynamic clusters have been obtained. To a PSE, one can associate a rewarding position (buy or sell), which tells us which trading direction would have produced a profit. In the empirical application below, the PSE is a takeover bid, thus the rewarding position is to buy before the PSE. An insider i that aims at maximizing the profit by exploiting the sensitive information on the takeover bid would purchase during some previous period the largest possible volume ($A_i \to 1$), in particular concentrating her investment on the stock ($a_i \to 1$), likely reaching her maximal historical exposure ($E_i \to 1$). As such, the best rewarding position in the feature space will be represented by the unit vector $\mathbf{1} \equiv (1, 1, 1)$.[8] Such considerations allow us to identify the cluster with a rewarding position (w.r.t. the PSE).

Following the principles exposed in the Introduction, an investor suspected of insider trading has discontinuous behavior with respect to her own past trading and with respect to the trading behavior of her peers. The proposed clustering method permits the identification of the set of suspects with discontinuous trading behavior in a data-driven manner. An investor is defined as discontinuous if she has never belonged to the cluster in the past (non-overlapping) time windows.[9] If this is the case, there are two mutually exclusive possibilities: (i) the investor has never traded the stock in the considered period; (ii) the investor has traded the stock in the past but taken a different position, e.g., lower exposure ($E_i \ll 1$) and larger portfolio diversification ($a_i \ll 1$). We name the first type of investors as *hard discontinuous* traders, while the second one as *soft discontinuous*.

Finally, given the identified set of discontinuous traders, it is interesting to obtain a ranking of the suspects depending on some score function measuring how anomalous their trading behavior is. To this end, we can consider any decreasing function of the Euclidean distance $d(x_i, \mathbf{1}) \equiv \sqrt{||x_i - \mathbf{1}||^2}$ of the vector of features x_i characterizing the investor i

[7] At initial time, random positions are used. The algorithm is then repeated many times for different input seeds, in order to find the global minimum over the number of initializations.

[8] Notice here that such considerations are independent of the size of the investor. Two investors with different capital but trading similarly in a given period are considered similar in the feature space. This is coherent with the current regulation about insider trading, which persecutes illegal behavior independently from the volume of the investment or the profit, if any.

[9] Before such an analysis, we need to ensure the stability in time of the cluster under investigation. To this end, we verify that the position of the centroid is almost constant for all time windows.

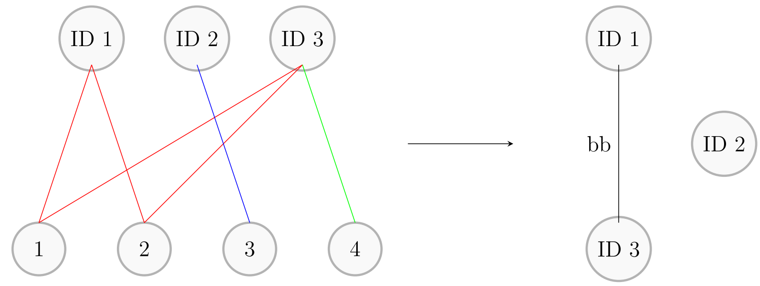


Figure 2 Example of a bipartite network (left) and its projected network (right). In the bipartite network, the first layer of nodes (on the top) is made up of 3 investors and the second layer (on the bottom) of 4 trading days. Three types of edge are allowed: red = *buying* state, green = *selling* state, blue = *buying-selling* state. The edge *bb* in the projected network of investors represents a co-occurrence of *buying* states

from the best rewarding position $\mathbf{1}$. In the empirical analyses below, the score is defined as $s_i = \exp(-d(x_i, \mathbf{1}))$, but other functions can be used.

3.2 Method based on statistically validated networks

The second clustering approach we employ is based on the so-called Statistically Validated Networks, which is an unsupervised learning method introduced in [43] and further employed in other works such as [5, 6, 36, 42].

This method aims at detecting structures in the projection of bipartite networks which represent complex systems. Analogously to [42], the complex system under our investigation is the activity of traders in the Italian Stock Exchange. Our goal is to identify clusters of investors who are synchronized in the kind and time of trading actions. This clustering is the starting point for a subsequent analysis which allows us to detect groups of traders who are likely to be suspicious of insider trading before a price sensitive event.

3.2.1 The SVN

Given the stock under investigation, the first step of the method is constructing a bipartite network, with two sets of nodes: traders on one side and trading days on the other. Only links between a trader and a trading day can occur, whose specific value represents a particular trading activity state. Figure 2 provides an example of this bipartite network. Each edge colors stands for a different trading state.

More specifically, let us consider $i = 1 \ldots, N$ traders and $t = 1, \ldots, T$ trading days. In order to characterize investors' trading activity, the following metric, which we denominate *directionality*, is associated to investor i in day t:

$$r(i,t) = \frac{V_b(i,t) - V_s(i,t)}{V_b(i,t) + V_s(i,t)} \qquad (2)$$

where $V_b(i,t)$ and $V_s(i,t)$ are the total volumes (in shares) bought and sold by trader i in day t. The metric $r(i,t)$ is by definition between -1 and 1. This choice permits a fair comparison between heterogeneous investors. In this way, we account also for trading behavior that is heterogeneous over time (in the size of the traded volumes).

Depending on the value of $r(i,t)$ compared to a fixed threshold θ (set to 0.01), three different states are defined:

- $r(i,t) > \theta$, *buying* state i.e. $s(i,t) = b$;
- $r(i,t) < -\theta$, *selling* state i.e. $s(i,t) = s$;
- $-\theta \leq r(i,t) \leq \theta$, *buying-selling* state i.e. $s(i,t) = bs$.

Similarly to [42], we verify the sensitivity of the results from the parameter θ, showing that it plays no important role.

Given the states $\{s(i,t)\ i = 1,\dots,N,\ t = 1,\dots,T\}$, we build a bipartite network where

- one layer is made up of traders: $A = \{1,\dots,N\}$;
- the other layer is made up of trading days: $B = \{1,\dots,T\}$;
- only links of the type (i,t) $i \in A$, $t \in B$ are admitted;
- each link can be b, s, bs, depending on $s(i,t)$.

The goal of the analysis is to perform a clustering of traders in relation to significant co-occurrences of the trading states over days. To this end, we consider the *projected* network of traders composed of N nodes and with weighted links that count the number of co-occurrences of states. In particular, a link between node i and node j exists if in the bipartite network, the traders i and j share at least one trading day of activity. This means that, being E and E_P the sets of edges in the bipartite and in the projected network respectively,

$$(i,j) \in E_P \iff \text{for } i,\ j \in A\ \exists t \in B \text{ s.t. } (i,t) \in E \text{ and } (j,t) \in E\,.$$

Since edges in the bipartite system can be of three types according with the states b, s and bs, in the projected network of traders we can have 9 types of links: bb, ss, $bsbs$, bs, sb, bbs, bsb, sbs, bss. For instance, if between i and j there exists a link of the type bb, this means that there is at least one trading day on which i is in state b and so is j. If we consider the bipartite network represented in the left part of Fig. 2, this situation corresponds to investors $i = 1$ and $j = 3$ since on day $t = 1$ and $t = 2$, i and j are both in state b. In the projected network of investors, represented in the right part of Fig. 2, there is a link bb between investors 1 and 3. On the other hand, investor 2 is isolated given she does not share any day of activity with other investors.

The projected network of traders is a weighted network where weights count the number of days for which a given co-occurrence of states between two traders occurs. Let us consider the example in Fig. 2: in the projected network, the edge $(i,j) = (1,3)$ of the type bb has weight equal to 2 since there are two days on which investors 1 and 2 are both in state b. The weight of this link is given by the number of days for which trader i and j are in state b i.e.

$$w_{ij} = |H| \text{ where } H \subset B \text{ s.t. } s(i,t) = b \text{ and } s(j,t) = b\ \forall t \in H.$$

If two traders (i,j) do not share trading days, $w_{ij} = 0$.

Thus, we obtain a multilink weighted graph with 9 types of link which can be formalized as $G_P = (V_P, E_P, L)$ where $V_P = A$, $L = \{bb, ss, bsbs, bs, sb, bbs, bsb, sbs, bss\}$,

$$E_P = \{(w_{ij}, l_{ij})\ i,j \in V_P,\ l_{ij} \in L\}$$

and w_{ij} as defined above.

In general, this graph is almost complete in many empirical examples, because of random activities creating non-zero "small" weights. As such, before performing the clustering, the first step is to identify statistically significant weights against the null hypothesis of randomness, the so-called *statistically validated* links. These links should highlight the structure and the organization of the system since their presence cannot be explained by a random assignation of co-occurrences. To this end, the Statistically Validated Networks (SVN) method [42, 43] is defined as follows.

Let us consider the link $(i,j) \in E_P$. First, we define as N_i^Q the number of days in which trader i is in state Q (b, s or bs). Analogously, N_j^R is the number of days in which trader j is in state R (b, s or bs). Then, N_{ij}^{QR} is the number of days in which i is in state Q and j is in state R.

The null hypothesis is defined by assuming the random co-occurrence of state Q for investor i and state R for investor j. More specifically, the probability of observing X days out of T in which the two traders are in the given states, is described by the hypergeometric distribution $H(X|T,N_i^Q,N_j^R)$ and its p-value is defined as

$$p(N_{ij}^{QR}) = \mathbb{P}(X \geq N_{ij}^{QR}) = 1 - \sum_{X=0}^{N_{ij}^{QR}-1} H(X|T,N_i^Q,N_j^R) \,.$$

If $p(N_{ij}^{QR})$ is lower than some chosen significance level p, the link is validated.

The procedure requires the validation of any possible link in the projected network. As such, it is a multiple hypothesis testing procedure and the significance level p must be corrected for the increased probability of false positives when performing more than one test at once [8]. Many solutions can be proposed depending on how we control for both false and positive discovery rates. The Bonferroni correction sets the significance level as equal to the usual single-test level divided by the number of tests performed, i.e. $p = \alpha/N_{test} = 2\alpha/(9N(N-1))$. Below, we set $\alpha = 0.01$. A less stringent approach is the so-called False Discovery Rate (FDR), see [8]. Once computed all the p-values, they are sorted in increasing order i.e. $p_1 \leq p_2 \leq \cdots \leq p_{N_{test}}$; then, the FDR significance level is $p_{\bar{k}}$ where $\bar{k} = \arg\max_{k=1,\dots,N_{test}} k$ such that $p_{\bar{k}} \leq \bar{k}\alpha/N_{test}$. As a result, the FDR approach is more prone to validate false positives but it has the advantage of a large statistical power by increasing the true positive rate. In the Appendix Sect. B.3, we show how the results are affected by the specific choice of correction for multiple hypothesis, comparing Bonferroni, FDR, and other possible solutions.

Once all tests are performed and all links in the projected network are validated,[10] the clustering analysis of traders can be carried out. We follow the approach proposed in [42]: communities of traders are detected on the network which only considers diagonal links i.e. bb, ss, $bsbs$. In the empirical analysis below, we find in fact that the number of diagonal links is predominant compared to non-diagonal links. Moreover, our goal is to find clusters of investors with similar trading activities around PSEs, i.e. the ones captured by diagonal links.

3.2.2 *Identifying clusters with Infomap*

Similarly to [43], the Infomap method for community detection is employed to find clusters in the validated network. Infomap is a community detection algorithm that minimizes

---

[10] An implementation of the SVN method can be found at the following link: https://github.com/cbongiorno/Bipartite-Tools [10, 12, 43].

the map equation over possible network partitions [11]. It is part of the flow models for community detection indeed, the map equation for a given network partition represents the information cost of a random walker which moves on the partition. Infomap amounts to find the network partition for which the information cost is minimum. The map equation allows us to obtain clusters that are less affected by the resolution limit (i.e. the tendency of many community detection algorithms of incorporating small clusters inside the larger ones) and, as such, it is preferred to other methods. From a practical point of view, we relied on its implementation in the Python package *infomap*.

3.2.3 Identification and classification of potential insider rings

After the identification of the groups of investors, the method is used to identify potential insider rings, i.e. small groups of investors which trade in a synchronized way and in a rewarding position before the PSE. If, for example, the rewarding position is to buy, looking for coordinated suspicious clusters consists in finding groups of investors who are in the b state in the proximity of the PSE.

In order to characterize clusters' suspicious behavior, some aggregated metrics are considered. The focus is on a time window $\bar{\Delta}$ before the PSE; this could be considered as an average reference period observed by the competent Authorities while investigating insider trading related to the particular PSE. Then, metrics are computed on $\bar{\Delta}$ as averages over each cluster, in particular the mean directionality R_C and the mean expected profit π_C. R_C is the average over the cluster C of the metric $r(i,t)$ in Eq. (2). The mean expected profit π_C is defined as the average expected profit of the traders in cluster C computed with respect to the takeover bid share price p_{TB} announced at the time of the PSE:

$$\begin{aligned}
&\pi_C = \frac{1}{N_C}\sum_{i\in C}\pi_i \quad \text{with} \\
&\pi_i = p_{TB}\left(\sum_{t\in\bar{\Delta}}[V_b(i,t) - V_s(i,t)]\right) - \sum_{t\in\bar{\Delta}}[A_b(i,t) - A_s(i,t)],
\end{aligned} \tag{3}$$

where N_C is the number of traders in cluster C, $V_{b/s}(i,t)$ are the volumes (in shares) bought/sold by trader i in day t, $A_{b/s}(i,t)$ are the amounts (in Euro) bought/sold by trader i in day t.

Once these two metrics are computed for each cluster, a list of clusters with suspicious behavior related to the PSE under investigation can be obtained. If for instance the PSE is the announcement of a takeover bid and so, the rewarding position is to buy, clusters synchronized in buying in $\bar{\Delta}$ have suspicious coordinated behavior. They correspond to clusters with mean directionality close to +1. A kind of ranking of these suspicious clusters can be provided by referring to the mean directionality, such that $R_C = +1$ corresponds to the most suspicious case. If R_C is equal for two clusters, the ranking is based on the mean expected profit, such that the higher the profit the more a cluster is suspicious.

3.3 Human-in-the-loop approach

The above methods capture two different discontinuities for trading behavior, signaling potential insider trading. The k-means approach allows us to assess the discontinuity of

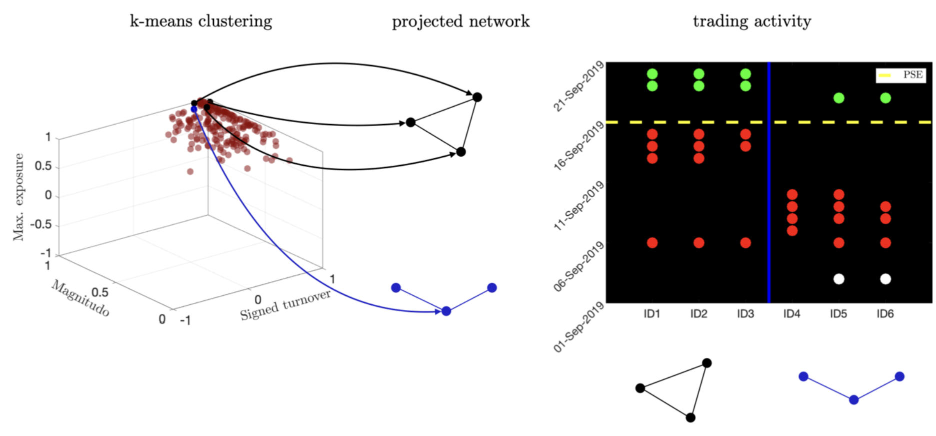


Figure 3 Graphical representation of the human-in-the-loop approach

the investors' behavior, if any, both with respect to her past trading behavior and with respect to the trading behavior before the PSE of her peers (i.e. change of cluster she belonged). The SVN approach captures instead the synchronicity in the trading for clusters of investors, thus signaling some coordination based on the exchange of information. In the case of inside information, a specific timeline of operations can emerge from data, e.g. buying before a takeover bid and selling hereafter (in some synchronized way). As such, the two methods lead in general to different results but can be used in a complementary way, in particular in a *human-in-the-loop* manner: a trader z who belongs to the cluster of (hard or soft) discontinuous investors for the k-means analysis, can be used as seed in the projected network of statistically validated links amongst investors if z has at least one connection in such a network. There are two possibilities: (i) z belongs to a clique of connected nodes in the SVN displaying high synchronization; (ii) neighbors of z are not connected to each other. In the first case, z is a potential insider according to the k-means method. Moreover, it is likely to belong to an insider ring, as suggested by SVN. In the second case, the SVN method provides other possible suspects since the trading behavior of the z's neighbors is synchronized with z, which is suspicious according to the k-means approach.

A pictorial representation of the proposed *human-in-the-loop* approach is in Fig. 3: from left to right, one selects some suspects in the k-means cluster, and the SVN result helps to find the new clusters of synchronized investors, if any, in the projected network of statistically validated links; finally, the trading activity of such investors, in particular, the trading states as defined above, is investigated. Notice that a beneficial way to visualize such activities is by constructing the right plot: the x-axis represents (anonymized) investors and the y-axis trading days; black points correspond to no-activity, red to b state, green to s state, and white to bs state. In this way, each vertical sequence of points represents the activity over time of a given trader. Vertical light blue lines separate the clusters of the projected network, and the dotted yellow horizontal line marks the date of the PSE. The example shows two potential insider trading behaviors of investors who buy before the announcement of a takeover bid and sell hereafter, displaying some synchronicity for the corresponding trading states.

4 Data

4.1 Transaction reporting database

The analysis is based on transaction reports collected by Consob for the Italian stocks, according to the directive 2014/65 by European Union, also called MiFID II[11] The relevant dataset was built aggregating the daily transactions of all investors operating in any of the Italian stocks, in the period from January 1, 2019 to September 30, 2021. In details, the dataset was built according to the following rules: i) all the information related to the identity of individual investors have been anonymized; ii) with reference to each stock (identified by its ISIN code), each data point keeps a record of:

1. anonymous identifier of the investor;
2. type of investors (household: H, investment firm: IF, legal entity: L);
3. trading venue of the operation (Borsa Italiana - MTA, London Stock Exchange - LSE, off-exchange, etc.) out of a total of 224 venues;
4. day of the operation;
5. buy and sell volumes (in shares);
6. buy and sell Euro volumes;
7. number of buy and sell contracts;
8. price of both the first and the last contracts (if there are more than one contract, otherwise they coincide);
9. minimum and max prices of contracts (if there are more than one contract, otherwise they coincide);
10. average price of buy (sell) contracts.

In the period covered by the dataset, 2,253,707 investors were observed, operating in 286 Italian stocks. This dataset was recently used in the investigation of the trading behavior of Italian investors during the Covid pandemic [16].

4.2 Price sensitive events database

In addition to the transaction reporting database, a data set containing several price sensitive events (PSEs) was built; such events, obviously public, had all been analyzed by the competent Authority with the aim of market abuse detection, by way of standard analytics methodologies. PSEs are events or a set of circumstances relating to listed companies that, when made public, had a significant impact on the price of the company's shares.

Our focus is on insider dealing in the Italian Stock Exchange. Investors who know in advance when a PSE will occur can trade in a rewarding manner before the information spreads, thus closing their position after the PSE and making a profit. For instance, if a trader knows a few days before its public announcement that a takeover bid will occur for a given stock, she could exploit such information by buying shares of the stock involved in the takeover. When the information is released in the market, the shares' price aligns with the offer price, typically increasing because of the premium paid by the company purchasing the stock. Then, the insider can sell to make a risk-free profit or hold the position until the takeover to earn the premium.

PSEs dataset contains a list of takeover bids for a number of stocks. As known, a takeover bid is a public offer made by a physical person or a legal entity that is willing to buy other

[11] In a nutshell, the MiFIDII/MiFIR regime has introduced new regulations for European financial markets and, among them, the transaction reporting obligation that requires investment firms or intermediaries executing transactions in financial instruments to communicate "complete and accurate details of such trans- actions to the competent authority as quickly as possible, and no later than the close of the following working day".

Table 1 Price sensitive events database. For each case, the stock name, the type of the event, its date, and the time window for the analysis are reported

Stock	PSE date	Reference period
IMA	July 28, 2020	June 29, 2020 - July 28, 2020
UBI	Feb 17, 2020	Jan 16, 2020 - Feb 17, 2020
PANARIAGROUP	Mar 31, 2021	Mar 1, 2021 - Mar 31, 2021
CARRARO	Mar 28, 2021	Jan 4, 2021 - Mar 28, 2021
MOLMED	Mar 17, 2020	Dec 2, 2019 - Mar 17, 2020

shareholders' shares at a price higher than the stock market value. As we saw, takeover bids can be exploited by an informed trader by buying before the event.

Our data report for each PSE the stock, the type of the event, its date, and the time window for insider trading analysis, here named as *reference period*. This is the period for investigation: it begins when some inside information is formed (e.g. the agreement between two companies for a takeover) and it ends when the information becomes public (i.e. PSE). As such, there is no general operative definition for the reference period of insider trading analysis but it depends on the inference (by the competent Authority) of the time at which some information starts to be considered price sensitive. In Table 1, the PSEs database is displayed.

5 Results

We tested our methods on a set of specific PSEs, namely takeover bids, which are listed in Table 1 together with the considered reference period. For space reason, in the following we present in detail a case study related to the takeover bid for the Italian stock "Industria Macchine Automatiche" (IMA) announced on July 28, 2020. In Tables 3-7 and in the Appendix Sects. A and B we present also results obtained by investigating the other PSEs.

The whole data set related to IMA covers the period January 2, 2019 - February 15, 2021 i.e. $T = 541$ trading days. The starting number of active investors[12] is 26,911 and the total number of records is 214,122. Summary statistics of the data set is reported in the top panel of Table 2 by referring to the investors' grouping of our database, as explained in Sect. 4.1. We observe 95.5% of investors are represented by households and their corresponding exchanged volume amounts at 6.7% of the total. On the other hand, investment firms which are the 0.5% of total investors, trade 59.0% of the total volume.

By looking at the number of days with trading activity for each trader in the investigated period, we observe that this number is small for a wide range of investors. This is shown in Fig. 4 where the probability density function of investors' trading activity days is displayed.

In the SVN analysis below, we will consider a threshold on the number of days in which an investor has traded, then obtaining the restricted set of the most active investors in the market. Summary statistics of this restricted data set are reported also in the top panel of Table 2. Finally, in the bottom panel of Table 2, the same statistics, but aggregated over all types of investors, are shown for the other assets involved in a takeover bid analysis.

5.1 K-means results

The *k-means* clustering method described in Sect. 3.1 is applied by defining: (i) t_S as the date of the PSE, namely July 28, 2020, (ii) the length of the time window D as the duration

[12] An investor is active if she executed at least one transaction in the observed time frame. In the text the words investor and trader are used indifferently.

Table 2 Top panel. A summary of IMA's data set, before (entire set, used for the k-mean analysis) and after (restricted set, used for the SVN analysis) setting the threshold on trading activity days, is reported. *N* is the number of traders, *C* is the contracts' number, *V* is the sum of bought and sold volume (in millions of shares). Bottom panel. The same (total) quantities are obtained for the other assets involved in a takeover bid analysis. In this second table, the different types of investors are aggregated

IMA	Entire set			Restricted set		
Investor type	N	C	V	N	C	V
Households	25,705	248,069	21	4405	151,575	13
Inv. firms	145	1,259,860	184	81	1,255,272	181
Legal entities	1061	607,650	107	358	581,933	98
Total	26,911	2,115,579	312	4844	1,988,780	292

	Entire set			Restricted set		
Asset/Total	N	C	V	N	C	V
UBI	154,080	7,582,854	18,793	11,977	7,084,384	17,446
PANARIAGROUP	6015	125,416	142	663	91,096	90
CARRARO	9019	179,521	202	608	123,643	125
MOLMED	17,877	342,703	4725	2472	271,364	2587

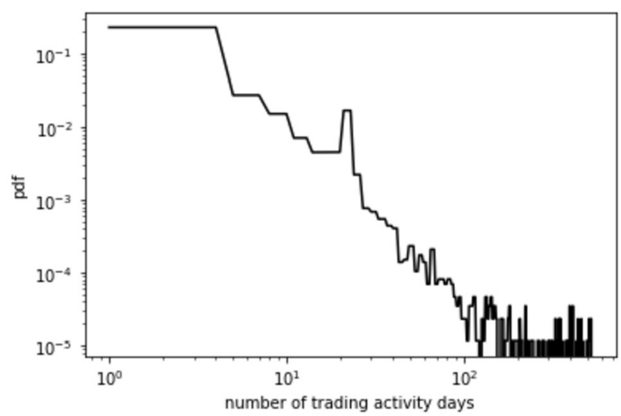


Figure 4 Probability density function (pdf) of the trading activity days for investors in IMA data set. The plot is with log scaling on both axes

of the reference period shown in Table 1, i.e. one business month of 20 days, and (iii) January 2, 2020 as initial time t_1. We consider a rolling window Δ starting from January 2020, then shifted week by week until the date of the PSE, for a total of 27 time windows. Within each time window, the features characterizing the trading behavior of an investor are computed as described above.

Before running the algorithm we find the optimal number of clusters K in each time window, as described in the Appendix Sect. A.2. The distribution of optimal Ks is shown in the left panel of Fig. 5. We select a single hyperparameter K for each time window as the rounded mean of the distribution, obtaining $K = 7$. We finally consider the *dynamic clustering* approach, achieving a stable description of the clusters in which we partition the feature space. This is confirmed by the high value of the Jaccard similarity coefficient, as shown in the middle panel of Fig. 5. In fact, the positions of the centroids in the feature space are almost stable in the whole period, except for the blue and purple clusters, which

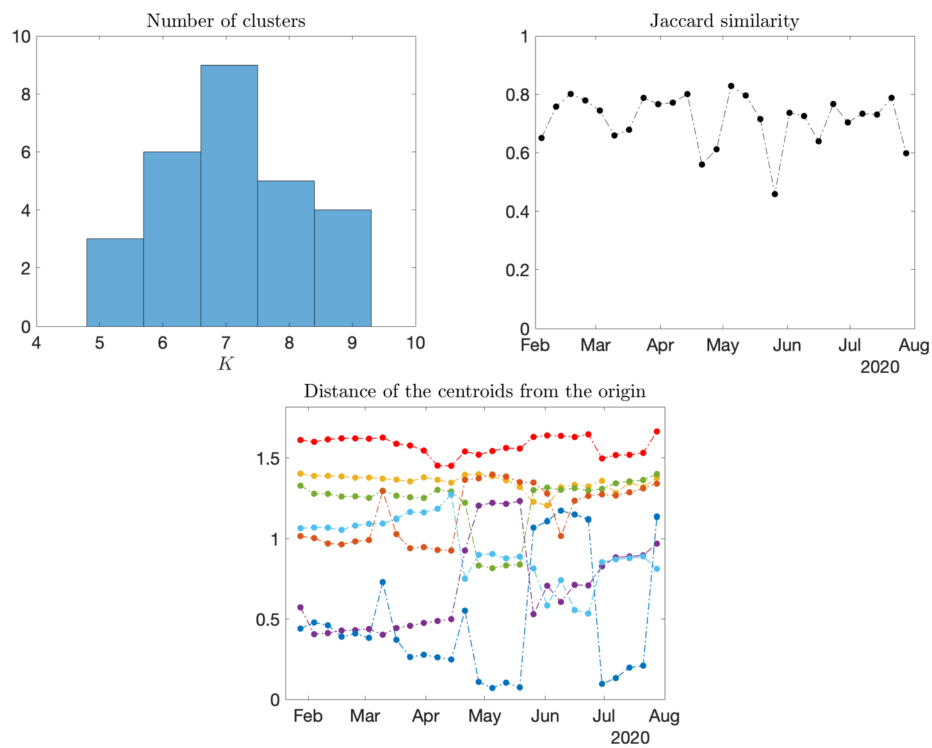


Figure 5 Distribution of optimal number of clusters K (left panel). Jaccard similarity between two clusters in a row for IMA (middle panel). Evolution of Euclidean distances of each cluster's centroid from the origin (right panel)

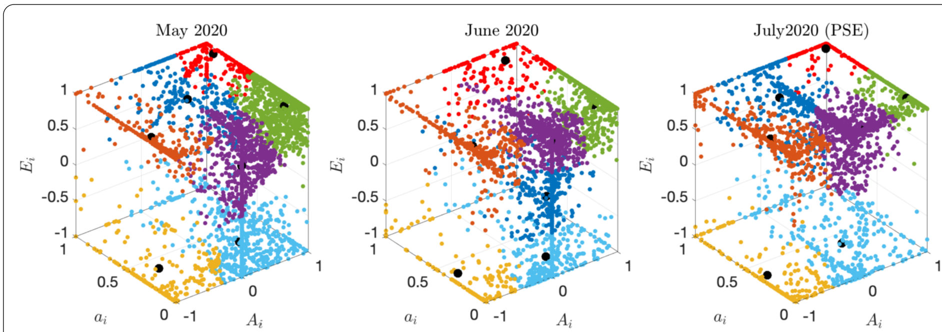


Figure 6 Inferred clusters for IMA ($K = 7$), for three different months

display some co-movements of the investors forming the clusters, see the right panel of Fig. 5.

A pictorial representation of the clusters in the Euclidean space at three different months (May, June, July) is shown in Fig. 6, where each cluster is identified by a different color. We can notice that almost all clusters are stable from one month to the next one, with the exception of the blue and the purple ones. Moving from May to June and then from June to July, the optimal partition shows some switching behavior depending on the distribution of blue and purple features at different months.

Table 3 Number of soft and hard discontinuous traders in the (red) suspect cluster for IMA (and other takeover bids). Percentage of discontinuous (soft + hard) traders in the suspect cluster, which is statistically significant according to the Chi-squared test with level of significance 5% (*) or 1% (**) when compared to other clusters, as described in the main text)

Stock	K	Soft disc.	Hard disc.	% disc. (red)	avg. % disc. (others)
IMA	7	66	237	0.86*	0.50
CARRARO	6	43	388	0.98**	0.68
MOLMED	7	25	259	0.89**	0.63
PANARIAGROUP	8	5	27	0.82*	0.60
UBI	7	204	1304	0.85**	0.67

Following the discussion in Sect. 3.1, the cluster more likely to be involved in insider activity is the red one since its centroid is the closest to the best rewarding position represented by **1**. However, belonging to the red cluster is not necessarily an indication of insider activity. The discontinuity of the trading behavior, if any, must be investigated, i.e. whether an investor belongs to the red cluster only in the reference period or she displays some coherence with the past, belonging to the red cluster also before.

For IMA there are 66 soft discontinuous and 237 hard discontinuous traders in the red cluster, summing to 303 discontinuous traders over a total of 379 investors in the red cluster, see Table 3. The question is now whether the fraction of discontinuous traders is statistically significant when compared with other clusters or not. In other words, the crucial point is to know if the cluster with rewarding position is somehow anomalous when looking at the discontinuity of trading behavior of investors. A pairwise comparison can be performed in a statistical fashion by using the χ^2-test. Given two clusters, whose investors are labeled as continuous or discontinuous traders, we can test the null hypothesis that the labels have equal probability in the two clusters (e.g. if the fraction of discontinuous traders in the red cluster is consistent with the fraction of discontinuous traders in another cluster) by considering the following statistic

$$\chi^2 = \sum_{\ell \in \{C,D\}} \frac{(n_\ell^{(1)} - n_\ell^{(2)})^2}{n_\ell^{(1)} + n_\ell^{(2)}} \tag{4}$$

where ℓ indicates the label, i.e. continuous C or discontinuous D, and $n_\ell^{(1)}$ ($n_\ell^{(2)}$) is the number of investors with label ℓ in the first (second) cluster. Under the null hypothesis, the test statistic (4) is χ^2-distributed with one degree of freedom. When the p-value associated with the test statistic is below a given confidence level, e.g. 5% or 1%, the null hypothesis is rejected and we can conclude that the two clusters are statistically different. In the case under investigation, the comparison between the red cluster and any other one is done by performing $K - 1$ χ^2-test, i.e. red vs. any other color. The results lead to the conclusion that the red cluster is statistically different from the others if the null hypothesis is always rejected. For IMA and PANARIAGROUP this is true at 5% confidence level while for the others at 1% level. In particular, the fraction of discontinuous traders in the red cluster is always larger than others. This result is summarized in Table 3, where we show the percentage of discontinuous traders in the red cluster and the average percentage for the other clusters.

Once the set of discontinuous investors has been obtained, we can sort them according to the score s_i defined in Sect. 3.1, which is inversely related to the distance from the **1**

Table 4 Ranking of the discontinuous traders in the red cluster for IMA according to the score function defined in the main text. The number of shares bought in the reference period, together with the directionality of trading and the expected profit of the trading in the reference period (by closing the position marked to the takeover bid share price) are shown

Ranking	Anonymous ID	Type	Score	Shares	Directionality	Exp. Profit (€)
1	257,853	L	1	52,000	1	443,947
2	783,031	L	1	23,535	0.49	320,457
3	1,664,331	L	1	19,500	1	277,625
4	1,139,263	L	1	17,250	0.81	241,132
5	9,280,051	L	1	16,700	1	138,132
6	9,276,483	L	1	10,500	1	94,727
7	9,249,321	L	1	9000	1	101,782
8	2,193,864	H	1	5700	1	94,006
9	1,564,905	H	1	4153	1	62,729
10	9,249,741	L	1	4001	1	51,084
11	9,253,185	L	1	4000	1	44,778
12	208,123	L	1	3559	0.31	44,961
13	9,253,505	H	1	3500	1	32,200
14	9,253,442	H	1	3245	1	35,786
15	9,239,385	H	1	3000	1	45,498
⋮	⋮	⋮	⋮	⋮	⋮	⋮
225	948,008	H	0.99	100	1	1671
226	7,882,312	H	0.98	940	1	14,202
⋮	⋮	⋮	⋮	⋮	⋮	⋮
303	135,723	L	0.55	3000	0.22	53,367

point in the cube. The ranking for IMA is shown in Table 4. Notice that there are 224 discontinuous traders in the highest rewarding position (score equal to one) over a total of 303 investors. This subset is then ranked according to the number of shares bought within the reference period. Table 4 shows also the directionality of the operations, as defined in Eq. (2), and the expected profit, as defined in Eq. (3) by using the takeover bid share price $p_{TB} = 68.0$ for IMA.

This kind of ranking is the final output of the methodology introduced here and needs to be interpreted as support to the investigation of insider trading by the competent Authority.

Finally, it should be noted that a rigorous validation procedure cannot be performed in such an unsupervised learning approach since the "labels" are not provided due to the severe privacy policy related to the data collected within the MiFIDII/MiFIR regime. However, to support both the effectiveness and robustness of the proposed k-means approach, we compare its output with (i) an anomaly detection benchmark based on baseline investor-specific trading metrics to corroborate the results and (ii) another unsupervised method proposed in a companion paper [38] for robustness analysis. The results are shown in the Appendix Sect. A.4. The analysis shows that, while simple baseline models are unable to capture the complex pattern of the trading profile that could be associated with insider trading, the more sophisticated method based on Principal Component Analysis and Autoencoders provides results comparable to those presented here. This is due to the fact that our proposed methods (and the one in [38]) are able to consider both the discontinuous temporal pattern associated with insider trading and the dynamical pattern of the investor clusters in the proximity of the PSE. Investor-specific metrics are unable in general to capture them.

Table 5 Numbers of the different types of links in the IMA's SVN obtained with the Bonferroni and FDR corrections. Numbers in parentheses are the corresponding percentage values

Edge type	Bonferroni	FDR
bb	1,320,213 (42.21)	1,468,510 (41.00)
ss	1,807,170 (57.78)	2,042,470 (57.02)
bsbs	41 (< 0.01)	287 (< 0.01)
bs	30 (< 0.01)	33,204 (0.93)
sb	49 (< 0.01)	28,541 (0.80)
bbs	31 (< 0.01)	4314 (0.12)
bsb	43 (< 0.01)	2562 (0.07)
sbs	26 (< 0.01)	1159 (0.03)
bss	53 (< 0.01)	939 (0.02)
Total	3,127,656	3,581,986

5.2 SVN results

We now describe the application of the SVN method to the IMA takeover bid to identify small groups of potential insiders. We restrict the analysis to the investors who traded at least 8 days in the investigated time period. In fact, setting this threshold allows us to reduce the statistical uncertainty that is typical of rare events. After that, the number of traders under investigation drops to $N = 4844$ as shown in Table 2, together with a summary of this restricted data set. From this table, it is evident that the investors in the restricted data set i.e. the ones who are active in at least 8 days, trade 93.6% of the total exchanged volume.

We proceed as explained in Sect. 3.2.1: states and the projected network of traders are obtained, then links are statistically validated with both Bonferroni and FDR corrections, and finally, clustering is performed via Infomap.

The Bonferroni threshold is equal to about $9.47 \cdot 10^{-11}$ while the FDR threshold is about $3.39 \cdot 10^{-4}$. Table 5 shows the number of validated edges we obtained with the two corrections. In both cases, the most validated links are of the *bb* and *ss* types and are about 3 million. This represents a consistent reduction in the number of possible edges in the projected network of traders that amounts to $N(N-1)/2 * 9 \simeq 10^8$. However, as we will see in the following, this number is larger than the one obtained for stocks with a similar number of records. A possible explanation is that in the IMA case, there are stronger synchronization signals.

The clustering is then performed on the network restricted to the diagonal links i.e. *bb*, *ss*, and *bsbs*. With the Bonferroni correction, the SVN comprises 2434 non-isolated nodes and 3,127,424 edges. With FDR instead, the SVN comprises 4673 non-isolated nodes and 3,511,267 edges. In Appendix Sect. B.2, further details about non-isolated traders' composition in terms of their type (household, investment firm, legal entity) are provided. After Infomap is applied to both the Bonferroni and FDR SVN, we obtain 69 and 13 clusters, respectively. A description of each cluster composition in terms of co-occurrence and investors types, can be obtained by relying again on hypergeometric tests as explained in Appendix Sect. B.1. The corresponding results related to the most populated clusters for the IMA case are reported in Appendix Sect. B.2. In this Appendix, an analysis of the relation between the FDR and the Bonferroni network is also available.

Once the clusters have been obtained, we display the activity of the investors, which do not correspond to isolated nodes, as described above. Figures 7 and 8 show the trading

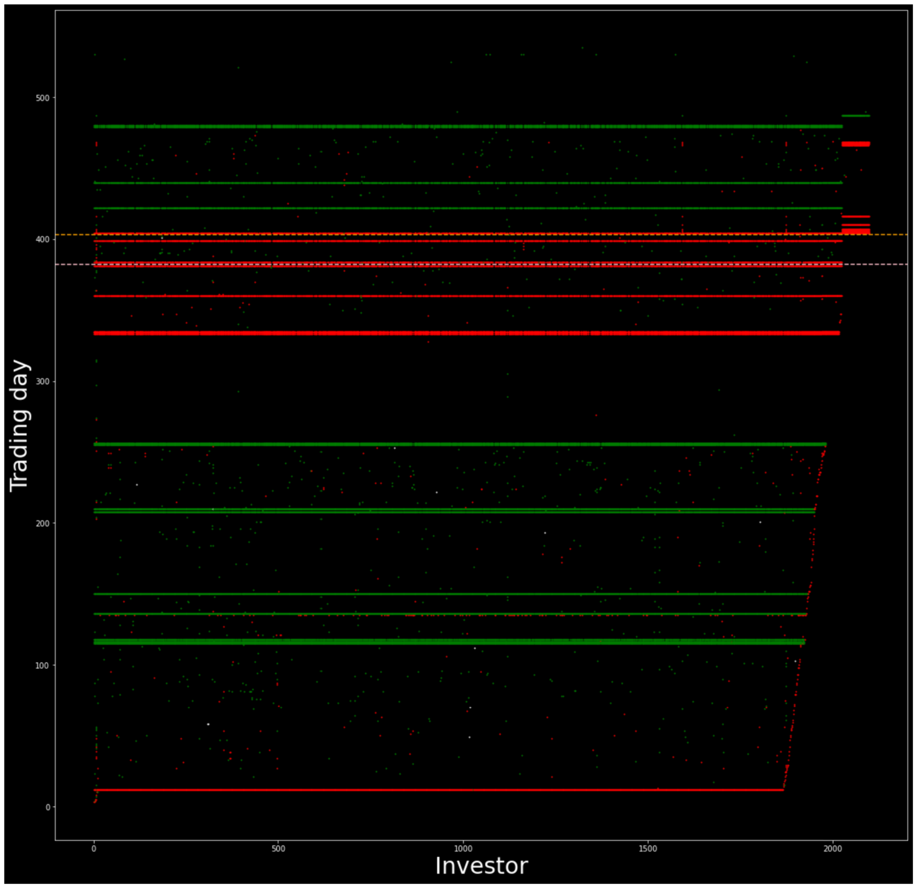


Figure 7 Graphical representation of traders' activity in the cluster 1 obtained by running Infomap on the Bonferroni SVN *bb-ss-bsbs*. The *x*-axis represents investors, the *y*-axis trading days. Black points correspond to no-activity, red to *b* state, green to *s* state and white to *bs* state. The dotted orange horizontal line marks the PSE and the dotted pink horizontal line marks the beginning of the time window $\bar{\Delta}$

activity of non-isolated nodes in the Bonferroni[13] SVN for IMA in the case of a takeover bid announced on July 28, 2020 (yellow dotted line in the figures, while the white dotted line marks the beginning of the reference period).

Figure 7 shows that the largest identified cluster is composed of more than 2000 investors, most of them being households, and displays an extremely synchronized behavior. Such behavior is likely explained by investors with portfolios managed by the same entity. In Fig. 8, we observe all the other clusters. It is clear that the degree of trading synchronization for each cluster is very high, displaying, also for this database, the capability of SVN to identify groups of a few synchronized traders.

We now turn to the use of SVN for the identification of potential insider rings. For the IMA case, the reference period $\bar{\Delta}$ corresponds to the period from June 29, 2020 to the PSE and the takeover bid share price, used to compute the mean expected profit π_C, is p_{TB} = 68.0 €. To identify suspect clusters, we compute the mean directionality R_C and π_C for the clusters of both the Bonferroni and the FDR SVN. Table 6, reports the results for the Bonferroni clusters with mean directionality greater than 0.5. Most of these clusters are

[13] In Appendix Sect. B.2, the corresponding activity plots for the FDR SVN are reported.

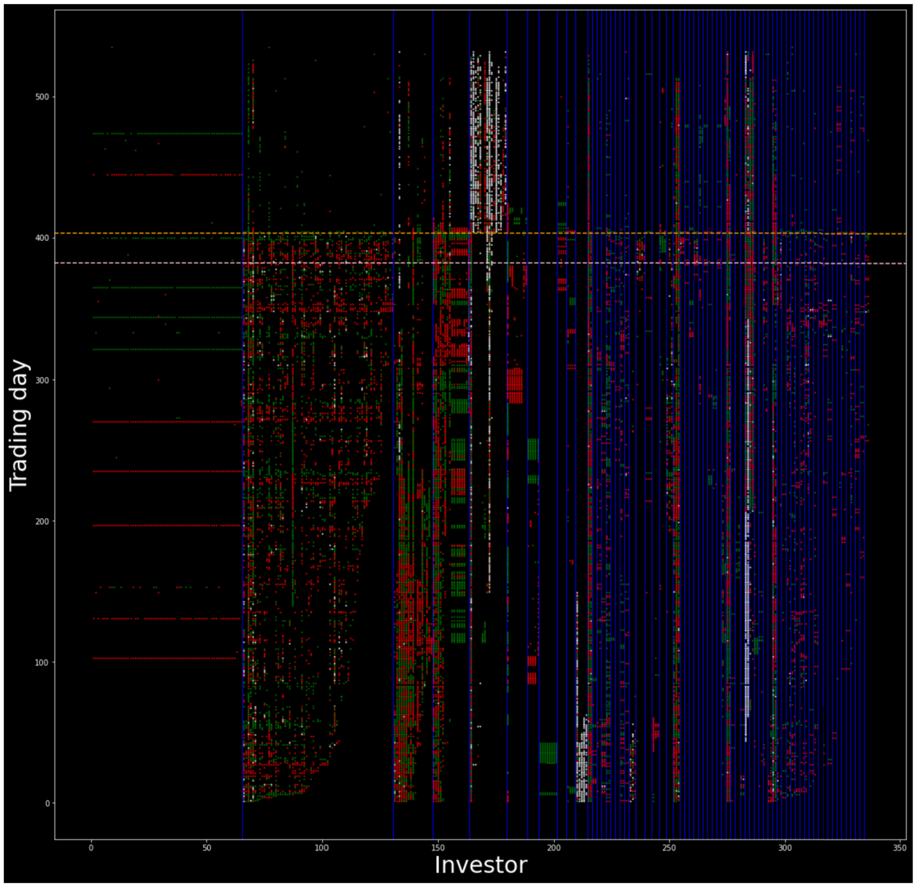


Figure 8 Graphical representation of traders' activity in the cluster 2 – 63 obtained by running Infomap on the Bonferroni SVN *bb-ss-bsbs*. The *x*-axis represents investors, the *y*-axis trading days. Black points correspond to no-activity, red to *b* state, green to *s* state and white to *bs* state. Vertical light blue lines separate the clusters, the dotted orange horizontal line marks the PSE, the dotted pink horizontal line marks the beginning of the time window $\bar{\Delta}$

made up of a couple of traders and in three cases, only one of them is active in the reference period $\bar{\Delta}$. Notice that also the large cluster of Fig. 7 is present in the table. Figure 9 shows the activity of the small clusters of Table 6. The buying synchronized behavior of these traders in the time window $\bar{\Delta}$ is evident. Several clusters in the figure were essentially inactive in the months before the PSE and started to trade in a very synchronized way right before the PSE. In particular, they assume a rewarding position (long for the IMA case) before the PSE, to close the position right after it. Such evidence is clearly suspect, thus requiring further investigations by the competent Authority.

When we use FDR instead of Bonferroni correction, the obtained clusters have typically a lower directionality. This might be related to the larger number of false positives detected by FDR. As Fig. 18 of the Appendix Sect. B.2 shows, the FDR cluster number 1 contains the majority of non-isolated nodes in the Bonferroni SVN. This issue is not specific to the IMA case but is observed also for the other four stocks. Table 7 shows the number of traders in clusters with mean directionality greater than or equal to 0.9 for the 5 takeover bids under investigation. It is observed that employing the FDR correction in the IMA case leads to a catastrophic reduction in the number of traders in clusters with high directionality. This

Table 6 Results related to clusters of the Bonferroni SVN with mean directionality $R_C > 0.5$ are reported. For each cluster, the number of its total traders and the ones active in the time window $\bar{\Delta}$ are specified. For the traders active in $\bar{\Delta}$, their type (H = household, L = legal entity, IF = investment firm) is summarised and then, mean quantities about the cluster - that are the mean expected profit π_C in euro and the mean directionality R_C - are shown

Cluster	Traders	Traders active in $\bar{\Delta}$	Traders type	π_c (€)	R_C
33	2	1	L	893,456	1.00
10	4	4	L	547,921	1.00
31	2	1	L	155,805	1.00
62	2	2	L	32,202	1.00
63	2	2	IF, L	2526	1.00
55	2	1	H	860	1.00
67	2	2	H	725	1.00
51	2	2	H	692	1.00
59	2	2	H	78	1.00
30	2	2	H	66	1.00
37	2	2	H	53	1.00
34	2	2	H	40	1.00
17	2	2	L	1,584,498	0.99
1	2098	1635	H (99.33 %)	145	0.92
29	2	2	L	259,379	0.91
64	2	2	L	92,214	0.88
18	2	2	H	1287	0.80

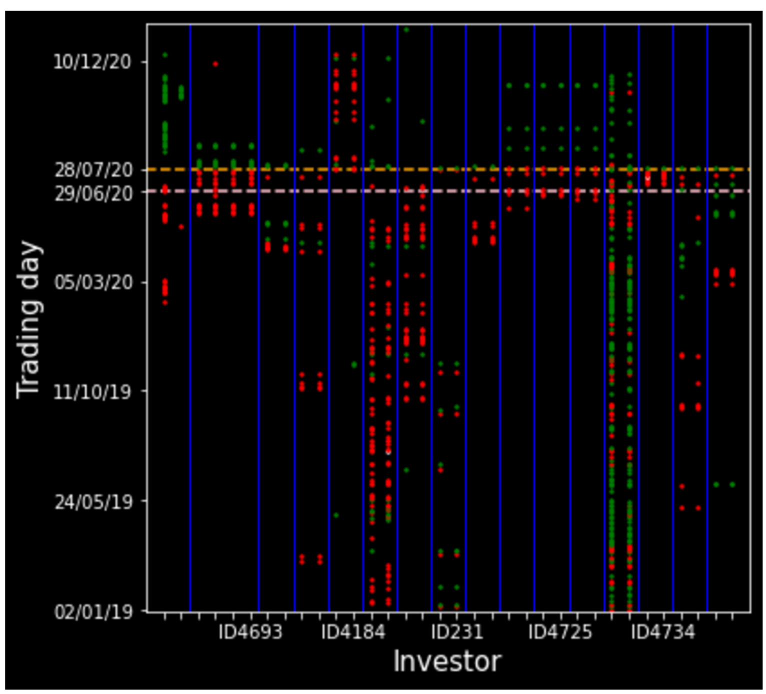


Figure 9 Graphical representation of traders' activity in clusters with mean directionality $R_C > 0.5$ (except cluster 1) obtained by running Infomap on the Bonferroni SVN *bb-ss-bsbs*. The *x*-axis represents investors, the *y*-axis trading days. Black points correspond to no-activity, red to *b* state, green to *s* state and white to *bs* state. Vertical light blue lines separate the clusters, the dotted orange horizontal line marks the PSE and the dotted pink horizontal line marks the beginning of the time window $\bar{\Delta}$

reduction is about 27 % for UBI. Also for MOLMED the FDR captures fewer traders in clusters with high directionality, even if the difference is just a couple of units. On the

Table 7 Summary table about the number of traders in clusters with mean directionality $R_C \geq 0.9$. The results related to the takeover bids under investigation are reported for both the Bonferroni and the FDR correction

Stock	Number of traders in clusters with $R_C \geq 0.9$	
	Bonferroni	FDR
IMA	1662	0
UBI	102	28
PANARIAGROUP	1	2
CARRARO	0	0
MOLMED	3	1

other hand, results are unchanged for CARRARO while, for PANARIAGROUP, the FDR correction is more effective at detecting traders in suspicious clusters but, similarly to MOLMED, the difference is just one unit.

As the above results clearly show, for the SVN method, the choice of multiple hypothesis testing corrections is key. In Appendix Sect. B.3, we show that more solutions can be considered and propose a method to optimize the significance level of the multiple hypothesis tests used for the validation of links. Finally, notice that the SVN clustering can be performed also by using an entropy-based approach, as introduced in [39]. In Appendix Sect. B.4 we prove the robustness of the results by comparing the two clustering approaches and showing there exists a very good agreement between the final outputs.

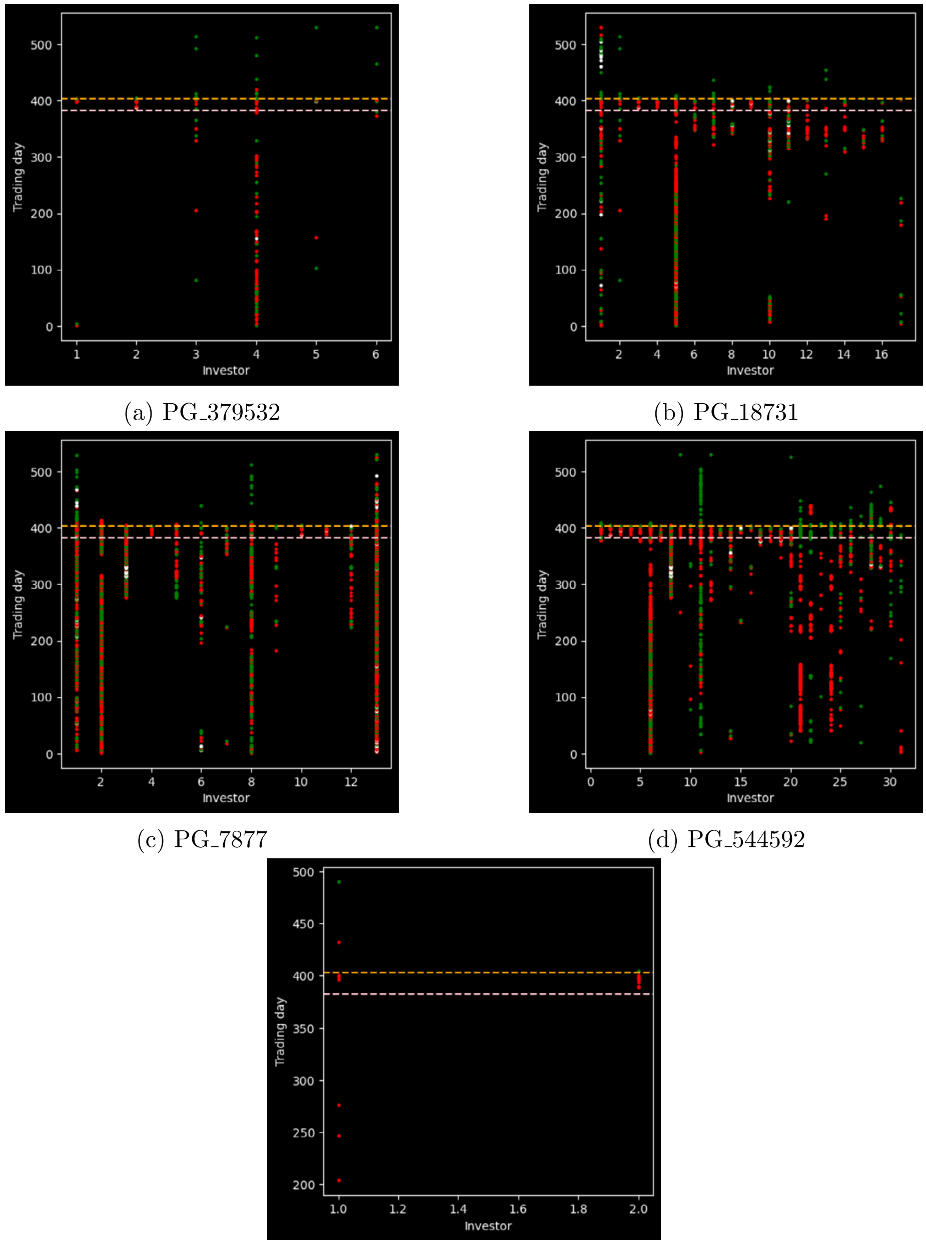
5.3 Human-in-the-loop approach and combined use of the two methods

The richer nature of the FDR SVN can be determinant when we would like to use the SVN method in a kind of *human-in-the-loop* manner. Let us suppose that another method for market abuse detection determines that an investor has suspicious behavior. Then, it is possible to identify investors who are synchronized with the suspicious one in their trading actions, by focusing on the first neighbors in the validated projected network of investors. For example, we consider the suspicious investors identified according to the *k-means*-based methodology in the IMA case. All of them are isolated nodes in the Bonferroni SVN but a few have some connections in the FDR SVN. In Fig. 10, each sub-figure shows the trading activity of a suspected trader according to k-means (the one most on the left) and of its first neighbors in the FDR SVN. It is clear that several first neighbors exhibit suspicious trading behavior around the takeover bid, but they were not identified by the k-means method. Indeed, by focusing on Fig. 11, the histogram of the first neighbors shows the largest peak for high values of directionality. This is in contrast to the histogram related to random traders, which has instead its largest peak for null directionality.

5.4 Discussion on methods

The *k-means* and the SVN based approaches for market abuse detection are two methods with different purposes: the former aims at capturing discontinuities in the trading operations of single investors (although taking into account the trading behavior of her peers). Instead the latter focuses on the identification of coordinated suspicious behavior of groups of investors. Consequently, they detect different anomalies.

Let us focus on IMA again and compare the traders identified as suspicious by the two methods. The *k-means* identified a total of 303 discontinuous investors, divided in 237 *hard* and 66 *soft*. On the other hand, the SVN Bonferroni approach detected 1662 traders in clusters with mean directionality greater than or equal to 0.9. The overlap is very small,



(a) PG_379532 (b) PG_18731

(c) PG_7877 (d) PG_544592

(e) PG_77838

Figure 10 Each sub-figure is a graphical representation of the trading activity of a suspected node (the first from the left) and its first neighbors in the FDR SVN *bb-ss-bsbs*. The *x*-axis represents investors and the *y*-axis trading days. Black points correspond to no-activity, red to *b* state, green to *s* state, and white to *bs* state. The dotted orange horizontal line marks the PSE and the dotted pink horizontal line marks the beginning of the time window $\bar{\Delta}$. The anonymous IDs of the suspected traders are reported below each sub-figure

resulting in 4 traders, 2 households and 2 legal entities. The two households are *soft* discontinuous traders and they are part of the cluster number 1 in the SVN Bonferroni. They are both characterized by a strong directionality (1 and 0.99) but different levels of expected profit (1380 € and 95,619 €). The two legal entities are instead *hard* discontinuous traders with very high expected profits (277,625 € and 241,132 €) and directionality

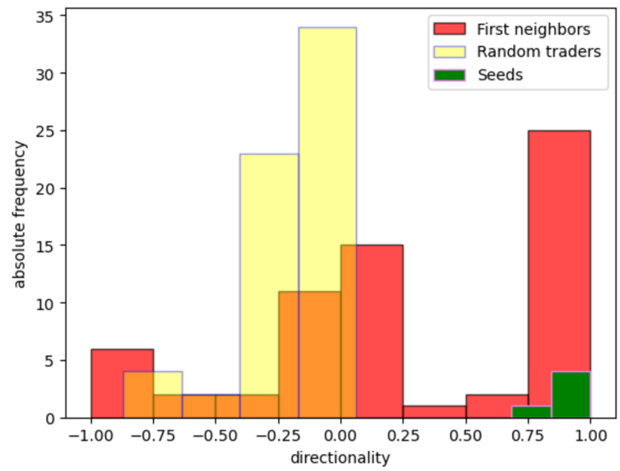


Figure 11 Histograms of the directionality of three groups of traders related to the *human-in-the-loop* approach. *Seeds* are the traders identified by the *k-means* method and that are used as input by the *human-in-the-loop* approach. *First neighbors* are the seeds' first neighbors in the FDR SVN. *Random traders* are randomly sampled traders

(1 and 0.81); they are traders number 3 and 4 in Table 4. Interestingly, they form a micro cluster of 2 highly synchronized traders, that is the cluster number 29 in the SVN Bonferroni.

This comparison highlights how the two methods provide different but complementary results. Given the complexity of the problem, it has to be tackled with an approach that captures several aspects at the same time. Therefore, the lacking of overlapping between the results of the two methods needs to be considered a strength of the proposed approach. There exists further a synergy that strengthens the connection between the two approaches, as shown when we combine them in a *human-in-the-loop* manner, and makes the proposed methodology for supporting decisions in insider trading a flexible and precise tool in the hands of the competent Authority.

Finally, let us point out that the computational times of our methods are moderate. The greatest contribution to the computational complexity of the method based on k-means comes from the trading features' computation and it is $O(NM)$ where N is the number of investors active on the asset under investigation and M is the total number of assets in our database. On the other hand, the greatest cost of the method based on SVN comes from the statistically validation of the links in the projected network of investors: it is $O\big(\bar{N}(\bar{N}-1)/2\big)$ where $\bar{N}$ is the number of investors considered for the SVN analysis. Both the implementations of these two steps have been parallelized. Overall, the computational time of each method does not exceed an hour for the largest assets in our dataset (as such, "real-time" daily indicators based on the proposed approach could be also devised as early-warnings). The methods have been run on a computer with an Intel® Core$^{\mathrm{TM}}$ i5 processor running at 1.60 GHz and using 16 GB RAM.

6 Conclusions

This paper proposes the use of two unsupervised clustering methods for the identification of potential suspects of insider trading in the vicinity of a PSE, such as a takeover bid. The first method clusters the investors in a space of three features and identifies as potential suspects those investors who display a discontinuous trading behavior with respect to their past activity and in a rewarding direction with respect to the PSE. The second method aims at detecting small groups of investors which trade in a synchronized and rewarding way in the vicinity of a PSE, pointing at possible insider rings and collusion in the insider trading activity. The two methods are complementary and indeed the overlap between the potential suspects found by them is small. In our opinion this is an advantage, since they focus on two different aspects of insider trading activity. Moreover, as shown at the end of Sect. 5.2, they can be used jointly to identify investors which are not identified by each method individually. This approach based on the identification of neighbors of suspect investors in the FDR network can also be used by considering suspects obtained with other methods (for example with the traditional supervising approach) rather than with the suspects from k-means.

There are several extensions of this work which we can foresee. First, it would be interesting to extend our analysis to other PSE, such as, for example, Accelerated Bookbuilds or corporate news releases. Second, in the k-means approach we have considered three specific features, which have been chosen both because they are financially relevant for the problem under investigation (insider trading) and because they allow to represent the investors in a three dimensional space. However, it is clear that other features could be added to the clustering analyses obtaining a richer characterization of the investors' population and thus a more precise identification of discontinuous behavior. Third, concerning the SVN approach, we used a trading day to define synchronous trading and for this reason we defined the trading state on a daily level. However, synchronous behavior can occur on longer (or shorter) time scales and one can easily extend our methodology in this direction. Finally the SVN method could shed light on the reference period to be considered for monitoring suspicious trading activity around a PSE.

Identification of insider trading is a complicated activity that requires the analysis of large and complex datasets, especially if one wants to consider the activity of potential insiders, not on an individual basis, but through a comparison with the behavior of the whole population of investors. Our proposed methodologies provide two contributions in this direction and we believe they might be very useful in the monitoring activity of supervising authorities. As clearly reported, such methodologies do not constitute any official process of the Consob but represent an in-depth analysis tool to be used when certain investigative conditions are met.

Appendix A: K-means

A.1 Features and representative trading behaviors

In this section, we present a visual illustration of certain representative trading behaviors identified through the k-means features, once they have been normalized and clustered, as depicted in Fig. 12. We specifically examine two trading features: *signed volume* and *maximum exposure*. The *magnitudo* feature directly reflects the level of investment concentration in the particular stock under consideration. Assuming a clustering number of $K = 5$ and investors' features grouped accordingly, as shown in the left plots of the Figure, we can draw insights.

For instance, consider the trading behavior of the green cluster (top left panel), which can be associated with the typical operations of a daily trader (top right panel). A net

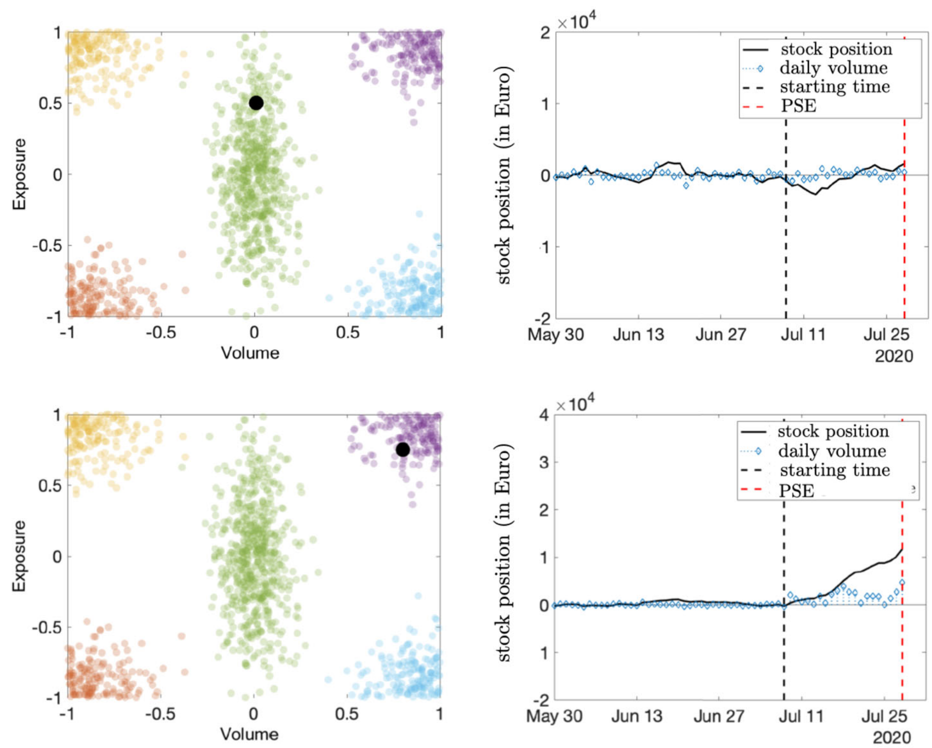


Figure 12 Pictorial representation of the clustering analysis via k-means restricted to two features: signed volume and maximum exposure (left panels). Pictorial examples of trading histories (stock position, traded shares) associated with the corresponding investor (black data point) to the left (right panels)

signed volume close to zero during a reference period (e.g., from the starting time to PSE) indicates a relatively small position held compared to historical data, despite potential fluctuations of the exposure feature between -1 and 1. This phenomenon arises from the normalization process applied to historical minor fluctuations around a net zero position, which aligns with the characteristics of a daily trader.

As another example, observe the purple cluster in the bottom left panel, characterized by signed volumes and maximum exposures close to one. This pattern is associated with an investor who accumulates a substantial number of shares during the reference period, resulting in a notably higher trading volume compared to historical data. Consequently, this investor acquired a positive stock position, reflected in the historical maximum for the exposure, as illustrated in the bottom right panel of the Figure. Representative trading behaviors, as captured by other clusters, can be drawn with similar reasoning.

We aim to highlight that the illustrative example in Fig. 12 is in line with the empirical results. In particular, the "purple" cluster emerges naturally from data for all the cases investigated in the paper. To our scope, this is particularly important because it represents the group of investors who (i) traded more in a rewarding position right before the price-sensitive event, (ii) acquired the historically largest exposure, and (iii) concentrated all the investment in the stock involved in the takeover bid. As such, it is the cluster that needs to be analyzed in more detail for the detection of potential insiders.

A.2 Implementation details: optimal number of clusters

The number of clusters K in which we partition the feature space is the only input for the clustering algorithm. In the absence of any prior information, K is a hyperparameter to optimize. This is usually done by following some heuristics [24]. Here, we adopt a standard approach in clustering analysis known as the *elbow* method. It is based on the computation of the mean variance of features $\{x_i\}_{i=1,...,N}$ with respect to the centroids' position $\{c_k\}_{k=1,...,K}$ as a function of the number of clusters K, namely

$$\mathcal{E}_K = \frac{1}{K}\sum_{k=1}^{K}\sum_{i\in S_k}||x_i - c_k||^2. \tag{5}$$

In fact, $\mathcal{E}_K$ is a non-increasing function in K, since a larger number of clusters will naturally improve the fit, thus reducing the variance. However, increasing K leads to over-fitting, eventually. To solve such an issue, the intuition is finding a cutoff point at which the diminishing of the variance in Eq. (5) becomes negligible. This is obtained by finding the number of clusters K such that the relative (negative) increments of $\mathcal{E}_k$ are smaller than 5% when more than K clusters are used. The procedure is in general repeated within each time window of length D, by covering the whole period $[t_1, t_S]$. We select the optimal number of clusters K as the (rounded) mean value over the m time windows.

A.3 k-means clustering analysis: outputs for other PSEs

The analysis of cluster stability, similarly to the one presented in Sect. 5.1, is performed also for other takeover bid events involving other stocks. The obtained results are in line with the IMA case presented above. They are summarized in Figs. 13 and 14.

A.4 Benchmarking and robustness analysis

The methodology presented in the paper is unsupervised and therefore cannot be directed validated with a ground truth. The typical approach in these cases is to make a comparison with other methods to corroborate the findings of the proposed methodology. To this end, we first compare our results with a basic anomaly detection approach, which ranks investors based on simple investor-specific metrics that serve as a baseline for validation. Second, we conduct a comparative analysis with the methodology introduced in our companion paper [38], which uses autoencoders, to further demonstrate the robustness of the proposed approach.

A.4.1 Benchmarking

We first consider the investor-specific metric π_i introduced in the second line of Eq. (3) that captures the *expected profit* of investor i from the position built with the trades performed in the period $\bar{\Delta}$ before the PSE, here set equal to one month. This benchmark metric is then used as an anomaly score, with investors ranked from the most to the least suspicious, according to their expected profit. We also consider a second investor-specific metric that captures the *trading volume fluctuation* of investor i, defined as the difference between the time average of the trading volume (without distinguishing between buying or selling) in $\bar{\Delta}$ and the time average of trading volume in the period before $\bar{\Delta}$. Similarly to the previous case, this benchmark metric is used as an anomaly score, ranking investors from most to least suspicious based on the magnitude of fluctuation, from the largest to

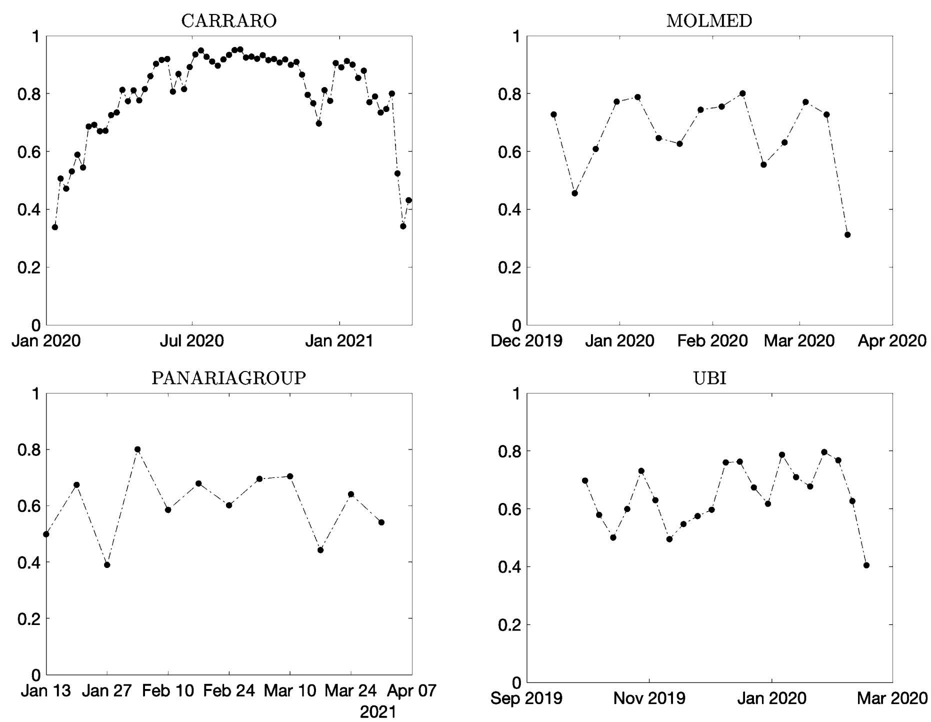


Figure 13 Jaccard similarity between two clusters in a row, by considering other takeover bid events listed in Table 1

the smallest. For both metrics, the aim is to identify investors who either profit significantly from trades made just before the PSE or show a sharp increase in trading volume.

In the case of IMA,[14] 5143 investors traded at least oncein the month before the PSE. Using the k-means based methodology proposed in the paper, we identify 303 suspects, and these investors can be ranked according to the anomaly score, as described in Sect. 5.1. For comparison, we computed the two baseline metrics, expected profit and trading volume fluctuation, for all investors and selected the n with the largest anomaly scores, i.e. the most suspicious ones. We use the Jaccard similarity to compare the set of investors detected with the k-means approach with the set of the top n suspects according to baseline methods. The results, shown in the left panel of Fig. 15, reveal a quite small overlap, with investors trading in the month before the PSE, ranked by expected profit, showing a slightly higher similarity with those detected with k-means. Thus, we conclude that our proposed approach identifies a set of suspicious investors, that are very different from those detected with the baseline methods. To understand which method more likely identifies potential insiders, we show in the right panel of Fig. 15 a scatter plot of the normalized signed turnover (A_i) and maximum exposure (E_i), both computed as described in Sect. 3.1, for the 250 most suspicious investors identified using the expected profit metric.[15] We compare these results with the 250 most suspicious investors identified using the k-means approach. The overall pattern is clear. The proposed baseline highlights sub-

[14] The results for UBI are similar and are not included for space reasons. However, they are available on request.

[15] The results for the trading volume fluctuation metric are similar and are not included for space reasons. They are, however, available upon request.

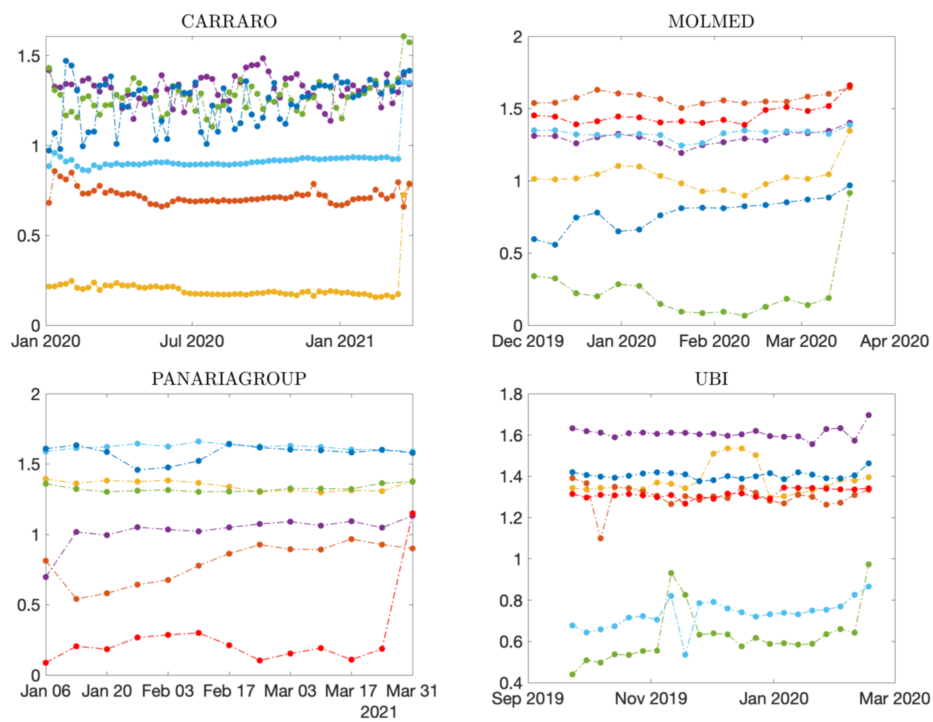


Figure 14 The evolution of the Euclidean distances of centroids from the origin by considering other takeover bid events listed in Table 1

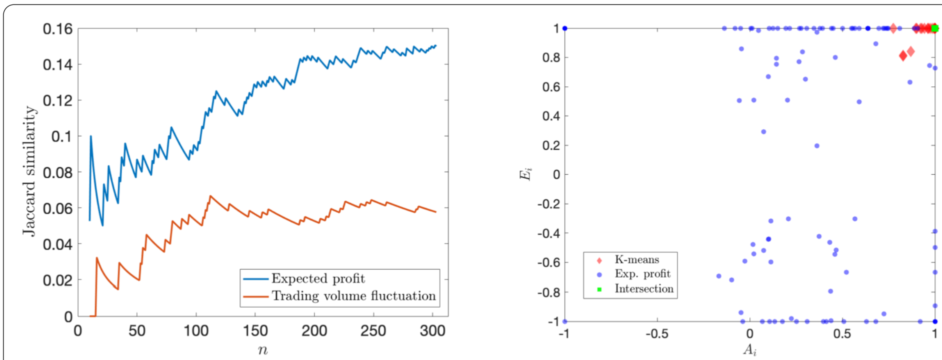


Figure 15 Jaccard similarity coefficients computed for the set of the n most suspicious investors selected with the k-means methodology and the proposed baselines (left). Normalized signed turnover A_i and maximum exposure E_i for the $n = 250$ most suspicious investors, selected with the expected profit metric (blue dots) and the k-means approach (red diamonds). Green squares represent the intersection, namely the common suspects for k-means and expected profit metric (right)

sets of investors with high turnovers and significant exposures relative to their past trading behavior, represented by the blue dots (A_i, E_i) near the point $(1, 1)$ on the Cartesian plane and by the green squares. However, many other investors flagged as suspicious by the baseline metrics exhibit no significant (positive) turnovers or exposures, indicating trading patterns consistent with their previous operations. This is not the case with the k-means approach that selects by construction suspicious investors with discontinuous trading behaviors and in a rewarding position, resulting in trading features (A_i, E_i) very close to $(1, 1)$ in the Cartesian plane or, in other words, displaying profiles consistent with insider trading.

These results clearly demonstrate that insider trading, as an anomaly detection problem, cannot be effectively addressed using baseline methods, such as ranking investors based on standard trading metrics. Such methods produce a large number of suspects, most of whom do not exhibit patterns consistent with typical insider trading activities. Instead, insider trading should be viewed as a contextual anomaly - where the context refers to the specific timeline of trading activities - and requires the development of specialized approaches, such as the k-means method, designed to address this unique challenge. The findings validate the proposed methodology.

A.4.2 Robustness analysis

In the companion paper [38], we propose an anomaly detection method based on a dimensionality reduction approach which avoids the choice of the three trading features. Specifically, the only input is the trading position (as an illustrative example, see Fig. 12) of each investor active on the asset for which a PSE occurs and the model learns a low dimensional representation of the time series. This new approach lies in the reconstruction-based paradigm of contextual anomaly detection [32], and it involves several steps. First, a dimensionality reduction step is carried out by means of Principal Component Analysis (PCA) [26] or autoencoders [22] and this allows us to obtain the reconstruction errors related to the trading profiles of each active trader. Then, the largest reconstruction errors are identified and, by verifying some trading conditions, anomalous investors are detected. Anomaly detection is thus performed in terms of reconstruction errors: under the assumption of few insiders operating in the market, the trading history of insiders is expected to be reconstructed worse than "normal" trading behaviors (associated with the majority of investors). In statistical words, either PCA or autoencoders (i.e., a nonlinear PCA) explain some large part of the variance of data (depending on the bottleneck size), which is likely not associated with the anomalies.

Here we compare the potential insiders detected with the k-means and with the autoencoder approach of [38]. In particular, for the latter, the dimensionality reduction step is performed by means of a vanilla autoencoder with one hidden layer and 16 neurons. Table 8 shows the number of potential insiders obtained with the two methods (A_{KM} and A_{AE}, respectively). by the method based on k-means (A_{KM}) and the method of [38] (A_{AE}). The big difference is mainly due to the fact that k-means uses also the magnitudo while the autoencoder uses only the trading activity in the asset under investigation. Thus k-means considers the activity of each investor on all other assets traded in the Italian Stock Exchange. To perform a fairer comparison, we consider a subset $\bar{A}_{AE} \subset A_{AE}$ obtained by eliminating the investors with magnitudo smaller than a threshold equal to the smallest magnitudo of the potential insiders in the k-means method.

Table 8 Comparison between the method based on k-means and the method of [38]. A_{KM} is the set of potential insiders according to the method based on k-means; A_{AE} is the set of potential insiders according to the method of [38] by employing an autoencoder for the dimensionality reduction step; $\bar{A}_{AE}$ is A_{AE} after a condition on the magnitudo of the investors is applied

Asset	$\|A_{KM}\|$	$\|A_{AE}\|$	$\|\bar{A}_{AE}\|$	$\|A_{KM} \cap \bar{A}_{AE}\|$
IMA	303	1502	389	293
UBI	327	2653	426	322

From Table 8, we observe that the number of potential insiders becomes comparable across the two methods. More importantly, when considering the intersection of the two subsets, we find than 96% of potential insiders according to the method based on k-means are also detected by the method based on auteoncoders. In statistical words, using the anomalies selected with k-means as targets, the true positive rate of the autoencoder methodology for IMA (UBI) is 96.7% (98.5%), while the false positive rate is 15.7% (3.1%).

Notice that we perform the comparison only for IMA and UBI because the autoencoder methodology requires many input data points (i.e. the number of investors' trading histories). These are available for IMA and UBI which were mid-cap stocks. On the contrary, the others stocks (PANARIAGROUP, CARRARO, MOLMED) were small-cap and, as such, were traded by fewer investors. Interestingly, the k-means approach does not require large datasets and can be adopted also for small cap stocks.

In conclusion, the findings support the robustness of the proposed k-means method. When compared to a competing approach, designed to detect similar contextual anomalies by an automatic identification of the relevant trading features through dimensionality reduction—the results show a high level of agreement. This indicates that the trading features used in the k-means approach are well-defined, leading to the detection of anomalies that align with insider trading patterns.

Appendix B: SVN

B.1 SVN: characterising clusters

Given traders' clusters, their composition can be characterized by relying on the method introduced in [43], analogously to what is done also in [6, 42]. In particular, we investigate the over-expression and under-expression of the investors' categories and co-occurrences of states.

The method is again based on the hypergeometric distribution as a benchmark for randomness. We have a total of N_V elements divided in N_C clusters. In order to test whether the attribute Q is over-expressed in the cluster C, we compute the p-value

$$p_o(N_{CQ}) = 1 - \sum_{X=0}^{N_{CQ}-1} H(X|N_V, N_C, N_Q)$$

where N_Q is the total number of elements in the system with attribute Q. If $p_o(N_{CQ})$ is lower than a statistical threshold corrected with Bonferroni or FDR, then the null hypothesis is rejected and we can conclude the attribute Q is over-expressed in cluster C. In a similar way, we test whether an attribute Q is under-expressed in cluster C by comparing the p-value

$$p_u(N_{CQ}) = \sum_{X=0}^{N_{CQ}} H(X|N_V, N_C, N_Q)$$

with a given statistical threshold.

In our analysis, this procedure allows to gain a macroscopic characterization of each cluster that is obtained.

Table 9 Number of traders in the Bonferroni and FDR SVN *bb-ss-bsbs* divided by type

Type	Bonferroni	FDR
Households	2287	4277
Inv. firms	24	72
Legal entities	123	324
Total	2434	4673

Table 10 Summary statistics of the 10 most populated clusters obtained by running Infomap on the Bonferroni SVN *bb-ss-bsbs*. OI/UI = over/under-expressed investor type, OC/UC = over/under-expressed co-occurrence. H = household, L = legal entity, IF = inv. firm

Cluster	Traders	OI	UI	OC	UC
1	2098	H	IF, L	*ss*	*bsbs*
2	65				*ss*
3	65				*ss*
4	17	IF, L	H		
5	16	L	H		
6	16	IF	H	*bsbs*	*bb*, *ss*
7	9	L	H		*ss*
8	5	L	H		
9	8	L	H		*bb*
12	5	IF	H	*bsbs*	*ss*

Table 11 Summary statistics of the 3 most populated clusters obtained by running Infomap on the FDR SVN *bb-ss-bsbs*. OI/UI = over/under-expressed investor type, OC/UC = over/under-expressed co-occurrence. H = household, L = legal entity, IF = inv. firm

Cluster	Traders	OI	UI	OC	UC
1	4417	H	IF, L	*ss*	*bsbs*
2	119	IF, L	H	*bb*	*ss*
3	105	IF	H	*bsbs*	*bb*, *ss*

B.2 SVN: more details about the IMA case

In Table 9, the number of different types of non-isolated traders in the Bonferroni and FDR SVN *bb-ss-bsbs* is reported.

In Tables 10 - 11, summary statistics about the most populated clusters obtained via Infomap are reported. In particular, the method illustrated in Appendix Sect. B.1 is employed to obtain the over/under-expression of traders' attributes in clusters.

In Figs. 16 - 17, plots which represent trading activity of non-isolated nodes in the FDR SVN *bb-ss-bsbs* are displayed.

Being the FDR correction less restrictive than Bonferroni, the SVN obtained with the latter is contained in the SVN achieved with the former. Figure 18 shows an heat map representing how the clusters obtained via Infomap from the Bonferroni SVN are contained in the clusters coming from the analogous FDR network. The element (i, j) in the plot is the fraction of nodes of Bonferroni cluster j that is contained in FDR cluster i. This means each column sums up to 1 and represents how the nodes in the corresponding Bonferroni cluster are rearranged in the FDR clusters. The FDR cluster number 1 is made up of 4417 elements and it turns out to contain most of the non-isolated nodes detected by the Bonferroni SVN.

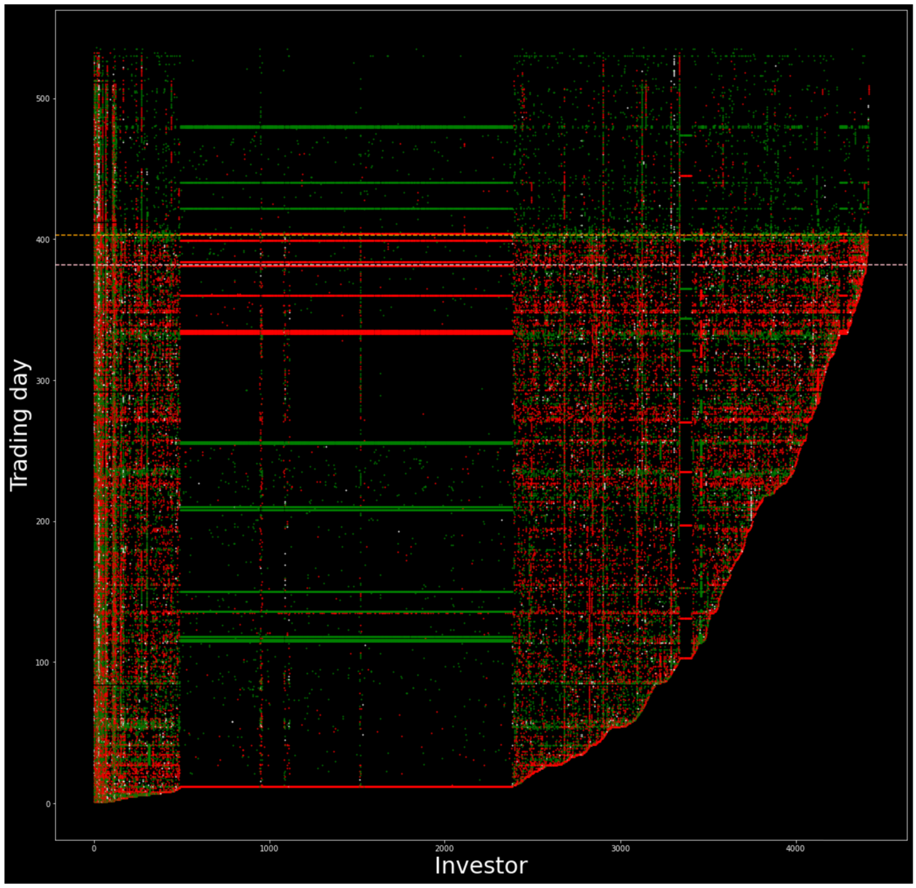


Figure 16 Graphical representation of traders' activity in the cluster 1 obtained by running Infomap on the FDR SVN *bb-ss-bsbs*. The *x*-axis represents investors, the *y*-axis trading days. Black points correspond to no-activity, red to *b* state, green to *s* state and white to *bs* state. The dotted orange horizontal line marks the PSE and the dotted pink horizontal line marks the beginning of the time window $\bar{\Delta}$

B.3 SVN and multiple hypothesis testing corrections: a comparison

The choice of the correction used for multiple hypothesis tests is fundamental and can lead to different conclusions about candidate traders as insiders. An option could be treating the statistical threshold for edge validation in our projected network of traders, as a parameter to be optimized. The optimal value of this threshold corresponds to the maximum number of traders in clusters with mean directionality greater than or equal to 0.9. In Fig. 19, this analysis is carried out for IMA, UBI, PANARIAGROUP, and MOLMED. The trend observed for IMA and UBI is similar and the maximum is achieved for a statistical threshold equal to 10^{-5} and 10^{-7}, respectively. This choice would amount to obtain for IMA, 27 more traders in suspected clusters than the Bonferroni correction and 36 for UBI. For PANARIAGROUP, 10^{-6} leads to the maximum value 4. MOLMED, instead, has an optimal threshold equal to 10^{-5}, which leads to 7 traders in suspected clusters. It is worth noting that MOLMED and PANARIAGROUP are assets with fewer records than IMA and UBI, as shown by Table 2, and so, this could also explain their fluctuating behavior observed in Fig. 19. These results confirm how the choice of an optimal statistical threshold for validation should be carried out case-by-case.

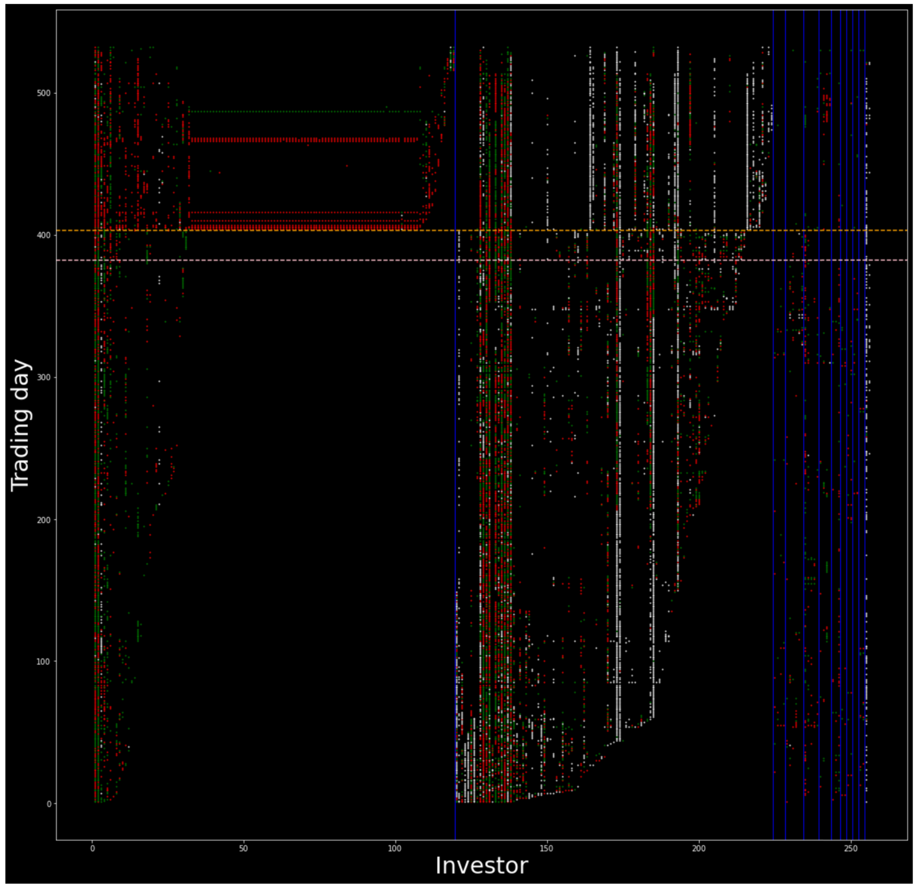


Figure 17 Graphical representation of traders' activity in the cluster 2 – 13 obtained by running Infomap on the FDR SVN *bb-ss-bsbs*. The *x*-axis represents investors, the *y*-axis trading days. Black points correspond to no-activity, red to *b* state, green to *s* state and white to *bs* state. Vertical light blue lines separate the clusters, the dotted orange horizontal line marks the PSE, the dotted pink horizontal line marks the beginning of the time window $\bar{\Delta}$

B.4 Robustness analysis of SVN: comparison with maximum entropy methods

In order to check the robustness of the SVN method for market abuse detection, the clustering is performed also with another method introduced in [39]. It is an entropy-based approach which employs the Bipartite Configuration Model (BiCM) as statistical benchmark.

B.4.1 Method

As for the SVN method, the starting point is to compute the states $\{s(i,t)\ i = 1,\dots,N,\ t = 1,\dots,T\}$ and then, to organize our data in a bipartite network where

- one layer is made up of traders: $A = \{1,\dots,N\}$;
- the other layer is made up of trading days: $B = \{1,\dots,T\}$;
- only links of the type $(i,t)\ i \in A,\ t \in B$ are admitted;
- each link can be b, s, bs, depending on $s(i,t)$.

Given the bipartite network, traders similarity is computed and its statistical significance is measured by performing multiple hypothesis tests with the BiCM as benchmark.

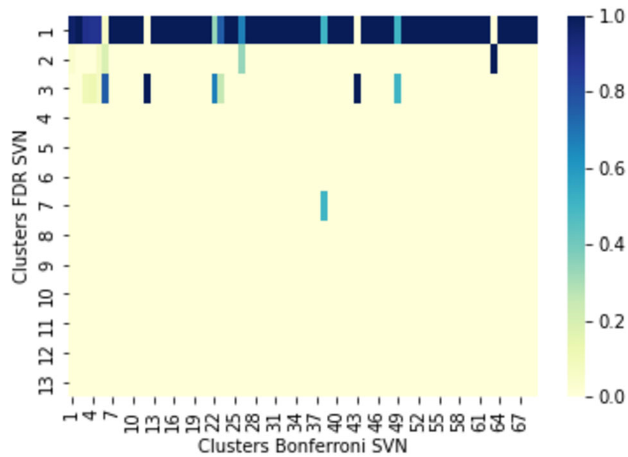


Figure 18 The element (*i*,*j*) is the fraction of nodes of Bonferroni cluster *j* that is contained in FDR cluster *i*

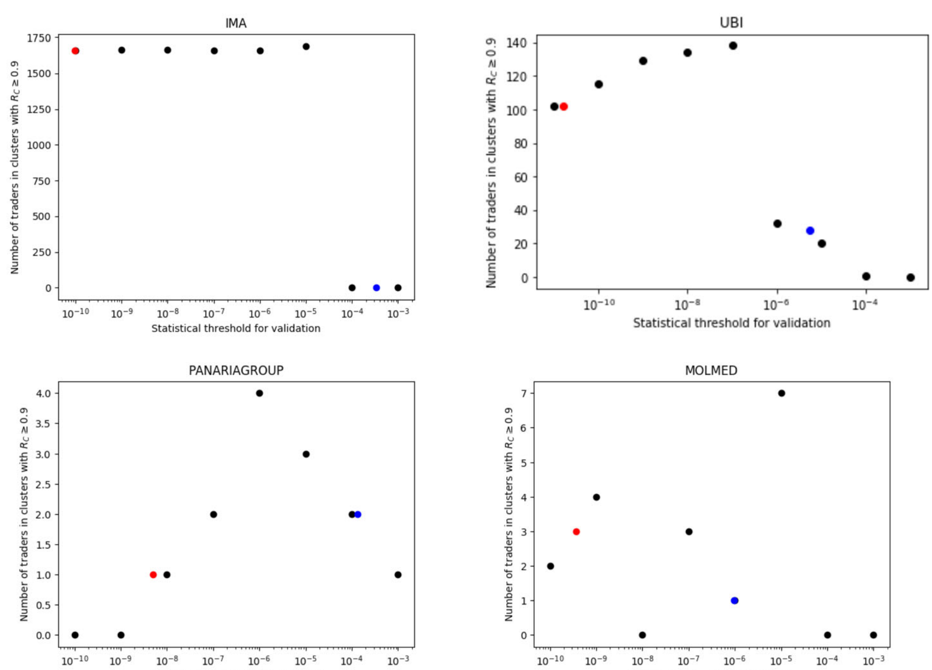


Figure 19 Each sub figure represents the number of traders in clusters with mean directionality $R_C \geq 0.9$ as a function of the statistical threshold employed for edges' validation. The red and the blue point corresponds to the Bonferroni and FDR correction respectively

For simplicity, let us consider our network as it just had a single type of diagonal link e.g. *bb* and thus, let us focus on the states of type *b*. Traders similarity is defined as the number of trading days in which *i* is in state *b* and so does *j*:

$$N_{ij} = \sum_{t=1}^{T} \sigma_{it}\sigma_{jt} = \sum_{t=1}^{T} N_{ij}^{t}$$

where $\sigma_{it} = \mathbb{I}[(i,t) \in E]$ and E is the edge set of the bipartite graph. This measure of similarity represents the number of the so-called V-motifs.

In order to quantify the statistical significance of traders' similarity, the Exponential Random Graph (ERG) class of null-models is considered. These models assign to a bipartite graph M a probability

$$P(M) = \frac{e^{-H(\theta, C(M))}}{Z(\theta)}$$

where θ is a vector of unknown parameters, $C(M)$ is a vector of constraints, $H(\theta, C(M))$ is the system's Hamiltonian and $Z(\theta)$ is the partition function.

In [39], several ERG models are considered; here, we choose to focus on the Bipartite Configuration Model (BiCM). The BiCM Hamiltonian imposes constraints on the degree sequences of both layers indeed,

$$H(\theta, C(M)) = \sum_{i=1}^{N} \alpha_i k_i + \sum_{t=1}^{T} \beta_t h_t$$

where $k_i,\ i = 1, \dots, N$ and $h_t,\ i = t, \dots, T$ are the degrees of traders and trading days respectively. More precisely, we have

$$k_i = \sum_{t=1}^{T} \sigma_{it}$$
$$h_t = \sum_{i=1}^{N} \sigma_{it} \ .$$

The parameters $\alpha_i,\ i = 1, \dots, N$ and $\beta_t,\ t = 1, \dots, T$ are Lagrange multipliers which are determined by Maximum Likelihood Estimation (MLE) starting from the biadjacency matrix M^* of an observed network.

The linear constraints of the system allow us to rewrite $P(M)$ in a factorized form:

$$P(M) = \prod_{i=1}^{N} \prod_{t=1}^{T} p_{it}^{\sigma_{it}} (1 - p_{it})^{1-\sigma_{it}}$$

where

$$p_{it} = \frac{e^{-(\alpha_i + \beta_t)}}{1 + e^{-(\alpha_i + \beta_t)}} \ .$$

The presence of linear constraints in the Hamiltonian amounts at treating links as independent random variables. This means N_{ij} is the sum of T independent Bernoulli random variables with

$$\mathbb{P}(N_{ij}^t = 1) = p_{it} p_{jt}$$
$$\mathbb{P}(N_{ij}^t = 0) = 1 - p_{it} p_{jt} \ .$$

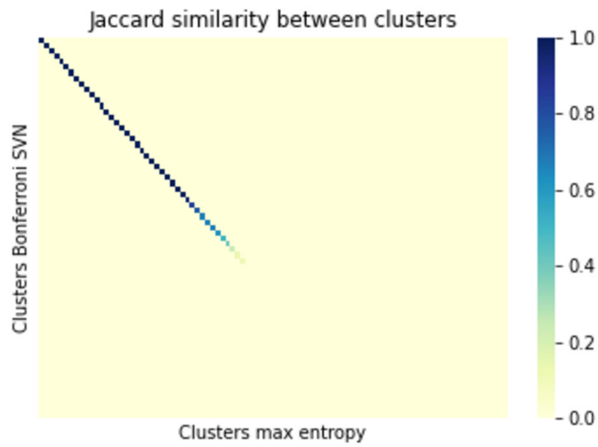


Figure 20 Heat map of the Jaccard similarity between clusters obtained with the Bonferroni SVN and the max entropy method. Clusters in the x and y axis are not in ordinal order: they are rearranged such that their correspondence is more evident

Thus, N_{ij} is a Poisson-Binomial random variable and

$$\mathbb{P}(N_{ij} = n) = \sum_{C \subset C_n} \left[\prod_{t \in C} p_{it} p_{jt} \prod_{t' \notin C} (1 - p_{it'} p_{jt'}) \right]$$

where C_n is the set of all subsets made up of n integers that can be selected from $\{1, 2, \dots, T\}$.

Given this distribution, the computation of the p-value follows:

$$p(N^*_{ij}) = \mathbb{P}(N_{ij} \geq N^*_{ij})$$

where N^*_{ij} is the value of the V-motifs in the observed network.

As for the SVN method, the link (i, j) is validated whether $p(N^*_{ij})$ is lower than a statistical threshold which is corrected with the FDR procedure.

In order to implement this method, we used the Python package *bicm*, which relies on the algorithm introduced in [25] to compute the Poisson-Binomial distribution. Once the validated projected network of traders is obtained, we performed the clustering with Infomap.

B.4.2 Results

As for the SVN method, first we obtain activity states and the bipartite network of traders and trading days. Given this network, V-motifs are validated as it is described in the previous paragraph. Our ultimate goal is to perform clustering on the validated projected network of traders with only diagonal links therefore, the maximum entropy method can be run on a reorganized version of the bipartite network: it is made up of three disjoint bipartite graphs, each one characterized by a different edge type (b, s and bs).

So, the maximum number of edges in the projected network of traders we obtain with this approach, is $3N(3N - 1)/2$. This number is greater than the corresponding one we have in the SVN method i.e. $9N(N - 1)/2$; indeed, in that case the validated network was obtained allowing for all links and not-diagonal edges were removed secondly. However, we observe $3N(3N-1)/2$ and $9N(N-1)/2$ share their dominating terms and given N = 4844, the corrections in multiple tests for the two methods have a difference which is negligible.

The validated projected network of traders that is obtained with the maximum entropy method has 3,279,920 edges and 2751 non-isolated nodes. After running Infomap, the clusters are 93 and analyses similar to the ones carried out in the SVN method can be done. However, we would like to focus on a comparison of the clusters obtained by the Bonferroni SVN and the maximum entropy method. In Fig. 20, an heat map representing the Jaccard similarity [24] between couples of clusters obtained with the two procedures is reported. The line which can be identified, shows that basically, 41 clusters have a one-to-one correspondence and among them, 31 have values of Jaccard similarity greater than 0.8.

Abbreviations
PSE, Price Sensitive Event; SVN, Statistically Validated Networks; MAD, Market Abuse Directive; *b*, *buying*; *s*, *selling*; *bs*, *buying-selling*; FDR, False Discovery Rate (correction); IMA, Industria Macchine Automatiche; UBI, Ubi Banca; PANARIAGROUP, Panariagroup Industria Ceramiche S.p.A.; CARRARO,Carraro S.p.A.; MOLMED, MolMed; ID, Identification.

Acknowledgements
This paper represents the personal opinions of the authors and do not bind the membership organization in any way. A very special thanks to Pietro Rizzo for his patience, experience and irreplaceable support. We wish to thank also Sandro Leocata, Francesco Gigante, Carlo Martinoli, Stefania La Civita, Alessio Sanfilippo, Vincenzo Vicari, Giancarlo Carotenuto and Mario Formato - working at the Consob, Market Division, Cash and Derivatives Department - for their useful comments.

Author contributions
All authors contributed to the conception and design of the work and the interpretation of the results. PM and AR developed the methodology, analyzed the data, and were major contributors to writing the manuscript. All authors read and approved the final manuscript.

Funding
This paper is funded by the project "Percorso di formazione su data analysis, network science, artificial intelligence e big data per gli abusi di mercato", a research collaboration with the Commissione Nazionale per le Società e la Borsa (CONSOB).

Data availability
The data that support the findings of this study are from the Commissione Nazionale per le Società e la Borsa (CONSOB), the government authority of Italy responsible for regulating the Italian securities market. Restrictions apply to the availability of these data, because of the severe privacy policy related to the data collected within the MiFIDII/MiFIR regime and are not publicly available.

Declarations

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Dipartimento di Economia Politica e Statistica, Università di Siena, Siena, Italy. [2]Scuola Normale Superiore, Pisa, Italy. [3]Consob, Rome, Italy. [4]Dipartimento di Matematica, Università di Bologna, Bologna, Italy. [5]University College London, London, UK.

Received: 28 July 2023 Accepted: 11 October 2024 Published online: 28 October 2024

References
1. Aggarwal CC (2013) Outlier analysis. Springer, New York
2. Ahern KR (2017) Information networks: evidence from illegal insider trading tips. J Financ Econ 125(1):26–47
3. Akoglu L, Tong H, Koutra D (2015) Graph based anomaly detection and description: a survey. Data Min Knowl Discov 29:626–688
4. Augustin P, Brenner M, Subrahmanyam MG (2019) Informed options trading prior to takeover announcements: insider trading? Manag Sci 65(12):5697–5720
5. Baltakiene M, Baltakys K, Kanniainen J (2022) Trade synchronization and social ties in stock markets. EPJ Data Sci 11(1):54

6. Baltakiene M, Baltakys K, Kanniainen J, Pedreschi D, Lillo F (2019) Clusters of investors around initial public offering. Palgrave Commun 5(1):1–14
7. Barthélemy JP, Bisdorff R, Coppin G (2002) Human centered processes and decision support systems. Eur J Oper Res 136(2):233–252
8. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc, Ser B, Methodol 57(1):289–300
9. Bhattacharya U, Daouk H (2002) The world price of insider trading. J Finance 57(1):75–108
10. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. J Stat Mech Theory Exp 2008(10):P10008
11. Bohlin L, Edler D, Lancichinetti A, Rosvall M (2014) Community detection and visualization of networks with the map equation framework. In: Measuring scholarly impact. Springer, Cham, pp 3–34
12. Bongiorno C, London A, Miccichè S, Mantegna RN (2017) Core of communities in bipartite networks. Phys Rev E 96(2):022321
13. Chandola V, Banerjee A, Kumar V (2009) Anomaly detection: a survey. ACM Comput Surv 41(3):1–58
14. De Jong W (2013) Tackling financial market abuse in the EU. Library of the European Parliament 22. 2013
15. Decreto legislativo 24 febbraio 1998, n. 58. https://www.consob.it/documents/art184d.lgs.n.58/98
16. Deriu P, Lillo F, Mazzarisi P, Medda F, Ravagnani A (2022). Preprint at. https://arxiv.org/abs/2208.00181
17. Donoho S (2004) Early detection of insider trading in option markets. In: Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining, pp 420–429
18. EU Legislation Summaries: Market abuse. https://eur-lex.europa.eu/summaries
19. Frino A, Satchell S, Wong B, Zheng H (2013) How much does an illegal insider trade? Int Rev Finance 13(2):241–263
20. Goldstein M, Uchida S (2016) A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. PLoS ONE 11(4):e0152173
21. Golmohammadi K, Zaiane OR, Díaz D (2014) Detecting stock market manipulation using supervised learning algorithms. In: Proc. Int. conf. Data Sci. Adv. Anal. (DSAA), pp 435–441
22. Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT Press, Cambridge. deeplearningbook.org
23. Hartigan JA (1975) Clustering algorithms. Wiley, New York
24. Hastie T, Tibshirani R, Friedman JH, Friedman JH (2009) The elements of statistical learning: data mining, inference, and prediction, vol 2. Springer, New York, pp 1–758
25. Hong Y (2013) On computing the distribution function for the Poisson binomial distribution. Comput Stat Data Anal 59:41–51
26. Jolliffe IT (2002) Principal component analysis. Springer, New York
27. Keown AJ, Pinkerton JM (1981) Merger announcements and insider trading activity: an empirical investigation. J Finance 36(4):855–869
28. Kyle AS (1985) Continuous auctions and insider trading. Econometrica, 1315–1335
29. Lee S, Kim G, Kim S (2011) Self-adaptive and dynamic clustering for online anomaly detection. Expert Syst Appl 38(12):14891–14898
30. Li A, Wu J, Liu Z (2017) Market manipulation detection based on classification methods. Proc Comput Sci 122:788–795
31. Lloyd S (1982) Least squares quantization in PCM. IEEE Trans Inf Theory 28(2):129–137
32. Mejri, et al (2022) Unsupervised Anomaly Detection in Time-series: an Extensive Evaluation and Analysis of State-of-the-art Methods. arXiv:2212.03637
33. Meulbroek LK (1992) An empirical analysis of illegal insider trading. J Finance 47(5):1661–1699
34. Minenna M (2003) The detection of market abuse on financial markets: a quantitative approach. Quad Finanza 54
35. Morgia ML, Mei A, Sassi F, Stefa J (2021) The doge of wall street: analysis and detection of pump and dump cryptocurrency manipulations. ACM Transactions on Internet Technology (TOIT)
36. Musciotto F, Piilo J, Mantegna RN (2021) High-frequency trading and networked markets. Proc Natl Acad Sci 118(26):e2015573118
37. Park YS, Lee J (2010) Detecting insider trading: the theory and validation in Korea exchange. J Bank Finance 34(9):2110–2120
38. Ravagnani A, Lillo F, Deriu P, Mazzarisi P, Medda F, Russo A Dimensionality reduction techniques to support insider trading detection. https://arxiv.org/abs/2403.00707
39. Saracco F, Straka MJ, Di Clemente R, Gabrielli A, Caldarelli G, Squartini T (2017) Inferring monopartite projections of bipartite networks: an entropy-based approach. New J Phys 19(5):053022
40. Song X, Wu M, Jermaine C, Ranka S (2007) Conditional anomaly detection. IEEE Trans Knowl Data Eng 19(5):631–645
41. Thalassinos E, Maditinos D, Paschalidis A (2012) Observing evidence of insider trading in the Athens stock exchange. J Economic Structures 1(1):1–26
42. Tumminello M, Lillo F, Piilo J, Mantegna RN (2012) Identification of clusters of investors from their real trading activity in a financial market. New J Phys 14(1):013041
43. Tumminello M, Micciche S, Lillo F, Piilo J, Mantegna RN (2011) Statistically validated networks in bipartite complex systems. PLoS ONE 6(3):e17994
44. West J, Bhattacharya M (2016) Intelligent financial fraud detection: a comprehensive review. Comput Secur 57:47–66
45. Wu X, Xiao L, Sun Y, Zhang J, Ma T, He L (2022) A survey of human-in-the-loop for machine learning. Future Gener Comput Syst

Publisher's Note