2015

Supporting Financial Market Surveillance: An IT Artifact Evaluation

Irina Alić
University of Göttingen, irina.alic@wiwi.uni-goettingen.de

28th Bled eConference

#eWellBeing

June 7 - 10, 2015; Bled, Slovenia

Supporting Financial Market Surveillance: *An IT Artifact Evaluation*

Irina Alić

University of Göttingen, Germany

irina.alic@wiwi.uni-goettingen.de

Abstract

In this paper, an IT artifact instantiation (i.e. software prototype) to support decision making in the field of financial market surveillance, is presented and evaluated. This artifact utilizes a qualitative multi-attribute model to identify situations in which prices of single stocks are affected by fraudsters who aggressively advertise the stock. A quantitative evaluation of the instantiated IT artifact, based on voluminous and heterogeneous data including data from social media, is provided. The empirical results indicate that the developed IT artifact instantiation can provide support for identifying such malicious situations. Given this evidence, it can be shown that the developed solution is able to utilize massive and heterogeneous data, including user-generated content from financial blogs and news platforms, to provide practical decision support in the field of market surveillance.

Keywords: Big Data, User-Generated Content, Decision Support Systems

1 Introduction

Financial market manipulation has received much attention from regulatory authorities, resulting in trading suspensions of those companies that may have been hijacked by fraudsters. Recent suspensions included two large groups of penny stock companies suspended by the U.S. regulatory authority on a single day (SEC, 2012) before they could harm investors. All companies have been traded on over-the-counter (OTC) markets with low regulatory standards. For the assessment of potentially affected companies, the supervisory authority used information technology to recognize thinly traded, low-capitalization stocks that could be affected by fraudsters. The system is calibrated to detect companies with low-priced penny stocks that are traded in low volumes or not at all; such companies must also be delinquent in their public disclosures (SEC, 2012). Unfortunately, the system reveals its weak point: namely, insufficient real-time surveillance on voluminous and heterogeneous data.

The pump-and-dump manipulation scheme is the most common form of information-based market manipulation in financial markets (Dunham, 2007; Zaki, 2013). It first appears as the spreading of false positive information to market participants by fraudsters; this information, typically user-generated content (UGC) on financial blogs and news platforms, is enthusiastically spread by fraudsters attempting to pump up a stock price to an artificial level. Usually, the affected stocks are penny stocks that trade below $5. Before starting the manipulation, fraudsters buy a significant quantity of the stocks over a longer period of time. Motivated by spread positive information, uninformed market participants relying on the bogus story buy the stocks, effectively forcing abnormal price increases (pump) (Bouraoui, 2009; Huang & Cheng, 2013). Finally, fraudsters make a profit by selling their stocks at the increased price level, which accordingly causes the stock price to drop (dump).

This paper addresses the question of whether the assessment of user-generated content has the potential to help regulatory authorities and financial institutions detect such situations. Therefore, the focus of this study is to present a quantitative evaluation of the final IT artifact that has been developed within a major EU project in order to demonstrate the feasibility of the approach. Accordingly, an evaluation based on a voluminous dataset is performed; for this purpose, OTC trading data and corresponding UGC related to 1700 companies were collected during 2012 and 2013. The collected data is used by the developed and instantiated IT artifact to identify situations of abusive behavior, and to assess whether those identified situations do actually demand more attention compared to those that were not identified by be suspicious.

At this point, it is important to note, that the EU project involved several activities in order to provide services including technologies development as well as the development of data mining algorithms suitable to handle big data complexity and provide metadata for further usage for the different use cases. The corresponding development processes and associated findings resulted in more than 20 research articles and project deliverables ("Project FIRST," 2013). Thus, in this present paper, the voluminous data utilized for this evaluation is provided by the mentioned services.

In the next section, related background research is introduced. Next, I provide an overview of the developed IT artifact instantiation. Section 4 describes the quantitative evaluation based on real data, and the final section concludes this paper.

2 Background

Relevant areas of study relating to this research in terms of stock fraud detection include literature on (1) stock touting impacts based on spam mails; (2) the impact of user-generated news on stock activity; and (3) financial markets services architecture. Each area will be briefly discussed in the following paragraphs.

Several studies examine of the impact of stock touting using spam emails. To generate profit, scammers spread misleading information. Authors (Frieder & Zittrain, 2008) investigate market activity prior to and following a stock touting email campaign, making use of a dataset containing about 75,000 spam emails. The research reveals high market activity beginning one day before email spamming commences and continuing until the day with the most considerable number of touting mails. The authors find that volume and return respond positively to touting, whereas the returns dip significantly following the conclusion of the campaign. Hence, the study suggests two main indicators for market manipulation: abnormal price changes and trading volumes, which will utilized by our IT artifact. Research conducted by (Bouraoui, 2009) demonstrates similar findings. The author provides an evolution model of volatility to assess the impact of stock spam emails based on a sample of 110 penny stock companies. The research shows abnormal returns three days after the commencement of spam emails. In both studies, the outcomes have been explained as the behavioral effect of market participants who have responded positively to the touting. Thus, these studies consider statistical approaches to explicate the influence of email stock promotion. Other studies introduce data mining techniques that help identify stock touting spam emails. Research by (Zaki, Theodoulidis, & Solis, 2011) observes spam massages to detect highly fraudulent stock activities by utilizing data mining techniques to identify stock touting spam emails. The accuracy detection of these experiments ranges from 58% to 71%. A considerable amount of research on the predictive power of UGC (such as tweets, financial forums, and blogs) on stock prices has been documented in the literature. In one instance, the research (Delort, Arunasalam, Leung, & Milosavljevic, 2011) introduce evidence of manipulation and examine the effect of such misuse in online financial forums. The authors show that manipulative user-generated content regarding companies with lower prices and market capitalization positively correlates to stock returns, volatility, and volume. One recent study (Smailović, Grčar, Lavrač, & Žnidaršič, 2013) presents a support vector, machine-based, sentiment classifier; here, a set of about 150,000 tweets thematizing eight companies (e.g., Apple, Google, and Microsoft) served as the data basis. The authors find that positive sentiment predicts positive movement in the closing price. A study proposed by (Zhang & Skiena, 2010) examines the ways in which blog and news data is reflected in trading volumes and returns. The authors demonstrate a significant positive correlation between media content and trading volume as well as stock returns. Therefore, in this study the UGC data will be incorporated.

Several commercial stock market systems for monitoring and detecting abuse in structured and unstructured data exist; some, such as the Securities Observation, News Analysis and Regulation Systems (SONAR), are presented in scientific research (Goldberg, Kirkland, Lee, Shyr, & Thakker, 2003). SONAR, which aims to monitor the stock market, applies data mining, text mining, statistical regression, and rule-based detection to recognize both abuse patterns in structured data and unusual trading following publication of the news. A study by (Mangkorntong & Rabhi, 2008) compares two different surveillance systems as event-processing systems in such areas as memory usage, scalability, and flexibility. The authors reveal the strengths and weaknesses of the two systems and suggest a generic approach that uses numerous different event-processing systems to support the detection process (Mangkorntong & Rabhi, 2007).

The next section introduces the Financial Market Surveillance Decision Support System (FMS-DSS), which was developed within an EU project between 2010 and 2013 ("Project FIRST," 2013). As an IT artifact instantiation it demonstrates how regulatory authority can be supported in detecting malicious pump-and-dump market manipulations utilizing voluminous and heterogeneous data streams.

3 Instantiated IT Artifact

During the research project, the opportunity to develop an instantiated IT artifact (FMS-DSS) was provided. Within the research consortium consisting of research institutions, industry partners, and financial regulatory authority, software components have been developed in close researcher-practitioner collaboration. The problem owner (i.e., domain experts and regulatory authority) intervened according to the project needs and aligned the design principles with their surveillance issues (Alić, Siering & Bohanec 2013).

Generally, decision support system configurations are built on the basis of three basic technology components related to: (1) data, (2) models, and (3) user interface (UI) (Turban, Sharda, & Delen, 2010). The following subsections will present these components. The data management component preprocesses and stores the needed data. The model component assesses whether or not given stock are suspicious of being currently affected by pump-and-dump manipulation schemes. The user interface component allows meaningful representation of suspicious situations. During operation, FMS-DSS continuously searches for the appearance of user-generated content related to monitored companies. For this purpose, unstructured input data is continuously retrieved, preprocessed, and stored in a database.

Based on rules and models that were developed in close collaboration with the domain experts that were involved in the project, the UI shows alerts, indicating assessments of how likely a stock is affected by manipulation (ranging from very high, high, medium, low, to very low).

3.1 Data

Selection of potential input data categories based on pump-and-dump manipulation scheme evidence: In meetings with experts (i.e., compliance officers and the regulatory body), the specific decision problem in revealing typical factors for pump-and-dump market manipulation was explained. The three main types of information incorporated into the decision model are: the manipulation of information concerning the company, the manipulation of the financial instrument, and news as user-generated content about the company and its financial instrument. Consequentially, the three main input data categories of market abuse suspiciousness are Company, Financial Instrument, and News:

- Company: According to the experts, there are two determiners for company suspiciousness. First, in those cases where the company is already involved in financial market manipulation, the financial authority issues litigation releases and puts the company on a blacklist, which is later refined into company, country, and industry blacklists.

The second determiner of a company's reliability is its history; the manipulator often targets newer companies and companies that have gone bankrupt and have recovered again. The History attribute is thus refined into the attributes Age and Bankruptcy.

- Financial Instrument: This category is refined into the attributes Market and Trading. If a company's financial instrument is listed in a market segment with low regulatory requirements, and the company itself has low market capitalization, then this instrument is seen as an additional indicator of suspiciousness. A change in trading volume or trading behavior is also seen as suspicious.

- News: The user-generated content spread in social communities is closely analyzed by the model. The attribute News is refined into attributes Content and Sentiment; the former analyzes whether the web publication includes suspicious phrases (e.g., increase in revenue, new product development), and the latter captures the sentiment expressed within the news.

This input data is provided by the developed services and described in (Grčar, 2012; Smailović, Žnidaršič, & Grčar, 2012).

3.2 Model Description

Based on further interviews, the attributes structure was transformed into a hierarchical tree (Alić et al., 2013), with the root node 'Pump and Dump' in the following P&D, based on a qualitative- multi-attribute-method as proposed in (Bohanec & Zupan, 2004), differentiating between the pure user-generated-content data (News) and the heterogeneous data regarding the company and related financial instruments (Comp_FinInst). Hence, the tree consists of the two sub-trees: one for 'Company' and its related financial instrument 'Comp_FinInst' and the second one for 'News'. The proposed model aggregates the attributes into assessment of pump-and-dump market manipulation:

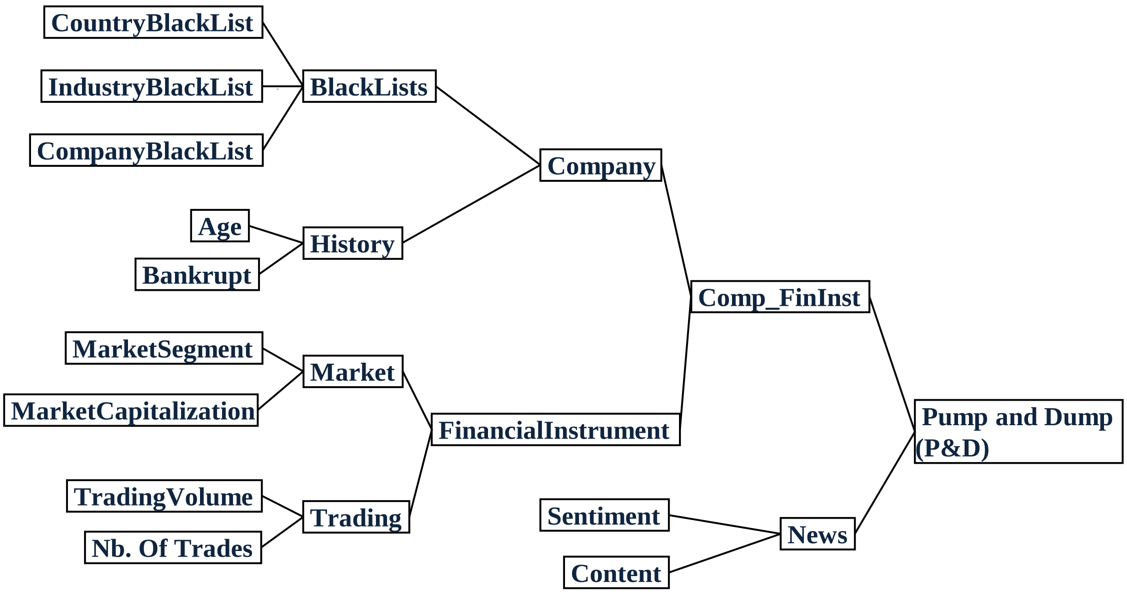

Figure 1: The hierarchical tree of attributes (Alić et al., 2013)

Attribute scales: For each attribute, the qualitative values are scaled in the range from highly suspicious (red colored) to not suspicious (green colored), where v-low is an abbreviation for very low, and v-high an abbreviation for very high (Figure 2).

Attribute	Scale
P&D	*v-low*; low; med; high; **v-high**
├**Comp_FinInst**	*v-low*; low; med; high; **v-high**
│ ├**Company**	*v-low*; low; med; high; **v-high**
│ │ ├**BlackLists**	*low*; med; **high**
│ │ │ ├CountryBlackList	*no*; **yes**
│ │ │ ├IndustryBlackList	*no*; **yes**
│ │ │ └CompanyBlackList	*no*; **yes**
│ │ └**History**	*low*; med; **high**
│ │ ├Age	*old*; med; **new**
│ │ └Bankrupt	*no*; was; **is**
│ └**FinancialInstrument**	*v-low*; low; med; high; **v-high**
│ ├**Market**	*low*; med; **high**
│ │ ├MarketSegment	*no*; **yes**
│ │ └MarketCapitalization	*high*; med; **low**
│ └**Trading**	*low*; med; **high**
│ ├TradingVolume	*low*; med; **high**
│ └NumberOfTrades	*low*; med; **high**
└**News**	*v-low*; low; med; high; **v-high**
├Sentiment	*v-low*; low; med; high; **v-high**
└Content	*v-low*; low; med; high; **v-high**

Figure 2: The attribute scales (Alić et al., 2013)

The scales for each attribute value are defined by the regulatory authority members and can be reconfigured when in use by the end user.

In our case, for example, the default setting for the attribute ‘Market Capitalization’ is ‘**low**’ (highly suspicious) if the company’s value is under 5 million (given currency), ‘medium’ if the value is between 5 and 30 million, and ‘**high**’ (not suspicious) if the value is greater than 30 million. In the next example the attribute ‘Market Segment’ has two values ‘**no**’ and ‘**yes**’. In this case the value ‘no’ means that the stock is not traded at the market segment with low regulations. Accordingly, the value ‘yes’ means, that the stock is traded at the market with low regulations (e.g. OTC market).

Manipulation scheme indicators: Within the predefined timespan the proposed calculation intends to identify abnormal changes which can be seen as indicators (Eren & Ozsoylev, 2006; Goldstein & Guembel, 2008; Zaki, Diaz, & Theodoulidis, 2012) for pump-and-dump abuse. Hence, the suspiciousness is assessed as follows:

Firstly, to assess recent changes in trading, long-and short-term average trading volumes are computed by taking the monthly and three-day averages of the trading volume:

$$Trading\ Volume\ Short\ Term = \frac{(\sum_{i=1}^{n} TV_i)}{n}$$

$$Trading\ Volume\ Long\ Term = \frac{(\sum_{i=1}^{m} TV_i)}{m}$$

$$Trading\ Volume\ Short\ Term = \frac{(\sum_{i=1}^{n} TV_i)}{n}$$

where TVi is the trading volume of the i-th day;^{n} n = 3 days; m = 30 days.

Secondly, to assess recent changes in number of trades, long-and short-term average trading volumes are computed by taking the monthly and three-day averages of the trading volume:

$$Number\ of\ Trades\ Short\ Term = \frac{(\sum_{i=1}^{n} NT_i)}{n}$$

$$Number\ of\ Trades\ Long\ Term = \frac{(\sum_{i=1}^{m} NT_i)}{m}$$

where NTi is the number of trades of the i-th day; n = 3 days; m = 30 days.

Additional indicators for pump-and-dump market manipulation are calculated accordingly as presented in Table 1.

Name	**Description / Definition**
Sentiment Long-Term Period	Sentiment of news based on assessment of long-term sentiment. Based on the overall picture of the mood of the news $SentiLong = (\sum_{I=1}^{m} Si)/m$
Sentiment Short-Term Period	Sentiment of news based on short-term sentiment (daily). Based on the overall picture of the mood of the news of one to three days $SentiShort = (\sum_{I=1}^{n} Si)/n$
User-Generated Content Long-Term Period	Content of news based on assessment of specified terms. Based on the overall picture of the mood of the news $ContentLong = (\sum_{I=1}^{m} Ci)/m$
User-Generated Content Short-Term Period	Content of news based on short-term specified terms (daily). Based on the overall picture of the mood of the news of one to three days $ContentShort = (\sum_{I=1}^{m} Ci)/n$

Table 1: Calculation of average values of the input variables

Thirdly, in order to calculate jumps in e.g. price (Frieder & Zittrain, 2008), the deviation of the short-term as related to the long-term average is calculated by dividing the short-term average by the long-term average and multiplying by 100, as presented in Figure 3. Three cases are assessed: when the short-term value is smaller than, greater than, or equal to the long-term value. Suspiciousness is assessed using aggregated numerical input values, which are then mapped according to qualitative scales, as defined by the problem owner as high, med and low or as v-high, high, med, low, and v-low.

Accordingly, structured data (such as e.g. trading volume) and unstructured data (such as e.g. user-generated content) are thereby taken into account in order to identify abnormal changes that may be indicators of pump-and-dump market manipulations. The recalibration of the indicator values or even the deployment of predefined default values can be adjusted by the end user.

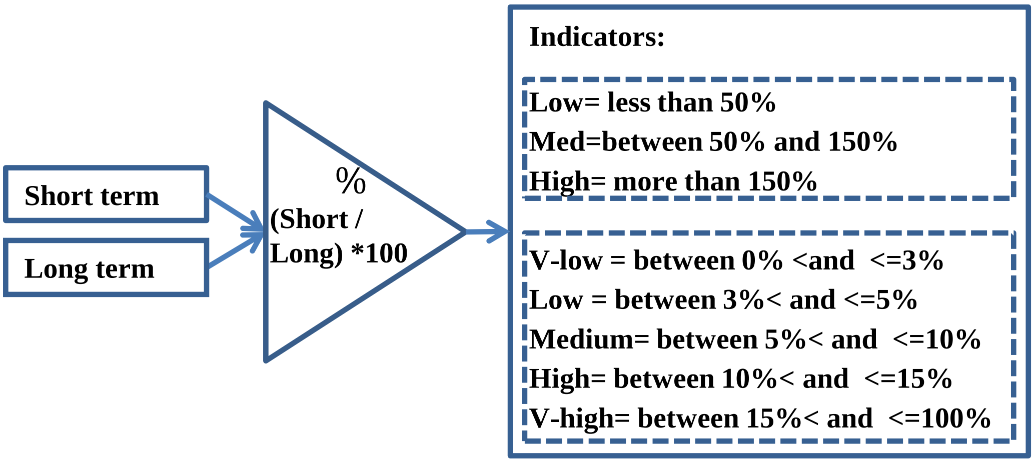

Figure 3: Calculation of artificial jumps as decision rules calibration

Output Calculation: The final pump-and-dump alert output value (P&D), is an aggregation of the lower level attributes Comp_FinInst and News, whereas the Comp_FinInst aggregates Financial Instrument and the issuing Company. News aggregates Content and Sentiment of the News regarding the Financial Instrument and the issuing company. P&D aggregates News and Comp_FinInst and presents the final indicator which indicates whether a suspicious market situation prevails. The scales consist of five values representing the decision rules as depicted in Figure 4.

	Comp_FinInst	News	P&D
1	v-low	v-low	**v-low**
2	v-low	low	**low**
3	v-low	med	**low**
4	v-low	high	**med**
5	v-low	v-high	**high**
6	low	v-low	**low**
7	low	low	**low**
8	low	med	**med**
9	low	high	**med**
10	low	v-high	**high**
11	med	v-low	**low**
12	med	low	**med**
13	med	med	**med**
14	med	high	**high**
15	med	v-high	**high**
16	high	v-low	**med**
17	high	low	**med**
18	high	med	**high**
19	high	high	**high**
20	high	v-high	**v-high**
21	v-high	v-low	**low**
22	v-high	low	**med**
23	v-high	med	**high**
24	v-high	high	**v-high**
25	v-high	v-high	**v-high**

Figure 4: P&D alert output aggregation

For the v-high alert, three possible combinations exist; the alerts ‘high’ and ‘med’ are respectively defined by seven and eight possible combinations. The alert ‘low’ has six possible permutations, and the alert ‘v-low’ has one combination.

3.3 User Interface

Initialized by a routine at the end of the day (midnight), the Surveillance System begins the search for the appearance of user-generated content regarding the monitored company every day for a predefined time period (e.g., within the last 30 days). For this purpose, the unstructured data is continuously retrieved, preprocessed by the developed services, and stored in the database where it can be assessed by the FMS-DSS. The FMS-DSS subsequently collects other related input data.

Based on end-user defined rules, the output appears as a signal which can be either a v-high, high, medium, low, or v-low alert. The following figure depicts a detailed view of alerts as displayed on the end-user screen. This view allows the user to drill down to a more detailed level for each alert. The related P&D output values are then stored in the database, enabling the user to post-analyze the suspect alerts.

Figure 5: FMS-DSS prototype

The FMS-DSS user interface depicted in figure 5 consists of a screen showing the listed ‘Alerts’ and ‘Details’ with specific information on each alert.

The warnings listed in the ‘Alerts’ screen provide general information to the end user. Here, the company name in the list is replaced with an International Securities Identification Number (ISIN), and the exchange where the security is traded. The screen also displays the alert number under which the specific alert is stored in the database.

The Date specifies the exact alert date. The user can select an alert from the list to view more detailed information.

The 'Details' screen depicts the alert selected above. Here, the display divides the information into two parts; the upper part displays unstructured data, and the lower part displays corresponding structured data.

In the unstructured data area, the graph indicates the appearance of positive (green) and/or negative (red) news. The yellow line represents the difference between positive and negative news. In the alert shown in this screen shot, the yellow line appears mostly in a positive sentiment area and the light blue bars represent the sentiment amount; in this example, we note the high appearance of sentiment on the 28th and 29th March 2012.

In the structured data, we can see three colored horizontal lines showing the highest (green), lowest (red), and average (black) of daily stock prices. The dark blue vertical bars represent trading volume. The alerts, appearing as yellow dots, are evident on the 29th and 30th of March and the 3rd of April 2012.

On the right side of the screen comprises some of the phrases collected from social-net news that have been used in the calculation of sentiment values; this feature allows the end user to observe the corresponding text and follow it to its source and author. Also of the right side of the screen, the Calculation tab contains further information on the calculation of the P&D output alert.

For the user's convenience, a history of alerts and the corresponding database ID of the specific company are also listed. The end user is thus provided with a range of information to use in the task of financial market monitoring.

4 IT Artifact Evaluation

This section presents an empirical evaluation of the presented FMS-DSS. In particular, the P&D output data is used for the performance assessment.

This section is organized as follows: first, the data required for the evaluation will be described; then, after presenting descriptive statistics of this data, the relationship between the P&D output values and actual observed stock price changes is analyzed.

There are two different sources of unstructured data considered in this dataset: user-generated content gathered from a variety of financial blogs, and news platforms. The unstructured input data is preprocessed and stored in a database, which is then assessed by the other system components as described previously; the data published on regulatory authority pages is likewise stored as input data in the database.

Structured financial data is automatically downloaded by the system from a data vendor (which was involved as one industry partner in the research project), while regulatory data regarding the blacklist is stored as a list in a database and can be modified by the end user.

Overall, three different data sources are considered: user-generated content, regulatory authority data, and stock price data for the period from 01.01.2012 to 03.09.2013. As a result, the P&D output sample generated by the FMS-DSS consists of 1700 OTC stock trades, with the corresponding structured and unstructured data containing 118096 entries.

4.1 Descriptive Analysis

The P&D output data is clustered within the five alert groups: very high, high, medium, low, and very low suspiciousness. The evaluations of suspiciousness in the given dataset of 118,096 entries are presented in Table 2.

Alert	Count	Percentage
v-high	982	0.83%
High	22215	18.81%
Med	92062	77.96%
Low	2780	2.35%
v-low	57	0.05%

Table 2: Examination of the highest and lowest daily price changes

Statistical analysis reveals prevalent medium- and high-value alerts. This may be explained by the sensitivity of the model. Certainly the model is developed as a qualitative model where pre-defined default values might be too receptive. Allowing the end user to configure their own aggregation rules and set their own thresholds would therefore allow for improvement in alert sensitivity, to the effect that false positive alerts can be reduced. The output data presented above already incorporates abnormal volume changes and other market abuse indicators as defined in previous sections.

4.2 Manipulation Examination

According to the domain experts involved in the research project, v-high and high alerts are the most relevant situations that demand further investigation. On that account, the evaluation proceeds by grouping v-high and high alerts into a suspicious class (v-h, h), and the other alerts into a non-suspiciousness class (m, l, v-l).

In the literature, significant price changes have been described as an indicator (Hanke & Hauser, 2008; Zaki et al., 2011) of stock manipulation. In the following, this measure is applied in order to determine if the developed IT artifact can provide support to detect potentially suspicious pump-and-dump abuse cases.

In order to evaluate the system's capability to detect such relevant cases, actual price changes in situations that were assessed by the system as highly suspicious (v-h, h) are compared with those that were assessed as less suspicious (m, l, v-l). In doing so, for each v-high, high, medium, low and v-low P&D output value, the changes between the highest (pH) (p_H) and the lowest (pL) daily stock price at the alert date are calculated. To assess the changed market price value for the lowest and highest daily prices, the following measure is calculated:

$$f = \frac{p_H - p_L}{p_H} * 100$$

A hypothesis is thus formulated to verify whether actually observed price changes related to the (v-h, h) class are significantly higher than the actual price changes related to the (m, l, v-l) class:

$$H_0: \Delta p_{(v-h,h)} \leq \Delta p_{(m,l,v-l)}$$
$$H_1: \Delta p_{(v-h,h)} > \Delta p_{(m,l,v-l)}$$

In doing so, and given suitably sized samples (Nv-h, h =23197; Nm,l,v.l=94899) a t-test is applied to test the formulated hypothesis. The null hypothesis can be rejected at the 1% level. The results are presented in Table 3.

	N=118096	**Mean**	**Std. dev.**	**Median**
(v-h, h)	N=23197	1.24	3.65	0.41
(m, l, v-l)	N=94899	0.4	1.46	0
p-value	<0.01			
t-value	34.31			

Table 3: Evaluation results

Undoubtedly, it appears that focusing on the group of v-high and high alerts is advantageous, as the t-value is highly significant. Hence, it can be stated that suspicious abuse cases where fraudsters first buy shares and then attempt to manipulate their price by spreading extremely euphoric content on social media, can be detected by the implemented FMS-DSS.

This statistical evaluation encourages in such a way that the default values as defined by the practitioners and regulatory authority might be seen as the best default values. A potential shortcoming from the end-user (e.g. compliance officer in a bank) perspective might be the huge number of potentially suspicious daily alerts (55 in averages) the system provides.

In order to reduce the overwhelming number of positive alerts to capture only the most probable suspect patterns, further refinements might be undertaken by the end user, such as a threshold routine that calculates the changes between the highest and the lowest daily stock price. For every price change greater than or equal to a specified percentage, the routine will trigger an alert.

An examination of such price changes results in a reduced list of 268 alerts, in other words, about one alert every two days. Thus, the procedure might reduce the noise of less relevant alerts and pinpoint the most suspicious manipulation cases. It appears that filtering based on price changes significantly increases the appearance of v-high values from 0.83% to 23.88% and of high values from 18.81% to 34.33%. The appearance of a medium alert, however, falls significantly from 77.96% to 41.04%. Hence, the final procedure will filter the most promising suspect alerts and help reduce their daily appearance to the lowest reasonable number.

5 Conclusion

This work has documented the effectiveness of the FMS-DSS and the importance of a market surveillance solution in improving the detection of information-based market manipulation. It has shown how the regulatory authority can be supported by the use of a market surveillance service. The evaluation results show that user-generated content can be utilized in near-real time using the developed FMS-DSS. The P&D alert results show that the FMS-DSS is able to handle voluminous and heterogeneous data and provide timely daily alerts by distinguishing suspect and non-suspect market situations. Most notably, this is the first study that investigates the effectiveness of market surveillance decision support that considers three different data sources: user-generated content, data provided by regulatory, and time series data from a data vendor. Our results provide convincing evidence for a long-term analysis (of approximately two years) of real data and suggest that the developed artifact may be effective in detecting real abusive cases of pump-and-dump market manipulation. However, some limitations are worth noting; even if the research is supported statistically, the artifact has not yet been evaluated in the real-world conditions of a compliance office. Future research should therefore include a subsequent effort in order to evaluate the acceptance and use of the running artifact.

Acknowledgments

The presented research program received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) within the context of the Project FIRST, large scale information extraction and integration infrastructure for supporting financial decision making, under grant agreement no. 257928. The author thanks all of the members of the FIRST project consortium for their contributions to the IT artifact developments described in this paper.

References

Alić, I., Siering, M., & Bohanec, M. (2013). Hot Stock or Not? A Qualitative Multi-Attribute Model to Detect Financial Market Manipulation. In *26th Bled eConference eInnovations: Challenges and Impacts for Individuals, Organizations and Society* (pp. 1–15).

Bohanec, M., & Zupan, B. (2004). A function-decomposition method for development of hierarchical multi-attribute decision models. *Decision Support Systems*, *36*(3), 215–233. doi:10.1016/S0167-9236(02)00148-3

Bouraoui, T. (2009). Stock spams: An empirical study on penny stock market. *International Review of Business Research Papers*, 1–14.

Delort, J.-Y., Arunasalam, B., Leung, H., & Milosavljevic, M. (2011). The impact of manipulation in internet stock message boards. *International Journal of Banking and Finance*, *8*(4).

Dunham, K. (2007). Pump and Dump Scams. *Information Systems Security*, *16*(1), 65–71. doi:10.1080/10658980601051755

Eren, N., & Ozsoylev, H. N. (2006). Hype and dump manipulation. In *Eropean Finance Association 2007 Ljubljana*.

Frieder, L., & Zittrain, J. (2008). Spam works: Evidence from stock touts and corresponding market activity. *Hastings Communications and Entertainment Law Journal*, *479*.

Goldberg, H., Kirkland, J., Lee, D., Shyr, P., & Thakker, D. (2003). The NASD Securities Observation, New Analysis and Regulation System (SONAR). In *Innovative Applications of Artificial Intelligence* (pp. 11–18). Acapulco, Mexico.

Goldstein, I., & Guembel, A. (2008). Manipulation and the allocational role of prices. *Review of Economic Studies*, *75*(1), 133–164.

Grčar, M. (2012). FIRST Consortium *D6.4 Ontology reuse and evolution*; http://www.project-first.eu/public_deliverables.

Hanke, M., & Hauser, F. (2008). On the effects of stock spam e-mails. *Journal of Financial Markets*, *11*(1), 57–83. doi:10.1016/j.finmar.2007.10.001

Huang, Y. C., & Cheng, Y. J. (2013). Stock manipulation and its effects: pump and dump versus stabilization. *Review of Quantitative Finance and Accounting*. doi:10.1007/s11156-013-0419-z

Mangkorntong, P., & Rabhi, F. (2007). A High-Level Approach for Defining & Composing Event Patterns and Its Application to E-Markets. In *The Second International Workshop on Event-driven Architecture, Processing and Systems (EDA-PS 2007) at the 33rd International Conference on Very Large Data Bases (VLDB 2007)* (pp. 1–4).

Mangkorntong, P., & Rabhi, F. a. (2008). A Domain-Driven Approach for Detecting Event Patterns in E-Markets. *World Wide Web*, *12*(1), 69–86. doi:10.1007/s11280-008-0053-1

Project FIRST. (2013). *Dissemination Material: http://www.project-first.eu/public_deliverables*.

SEC, G. (2012). *SEC Suspends Trading in Common Stock of Three Hundred Seventy-Nine Compnies Quoted on OTC May 14 , 2012* (No. No . 66980). *SEC* (Vol. May).

Smailović, J., Grčar, M., Lavrač, N., & Žnidaršič, M. (2013). Predictive Sentiment Analysis of Tweets: A Stock Market Application. In *Human-Computer Interaction and Knowledge Discovery in Complex, Unstructured, Big Data Lecture Notes in Computer Science* (pp. 77–88). Springer Berlin Heidelberg.

Smailović, J., Žnidaršič, M., & Grčar, M. (2012). *Web-based experimental platform for sentiment analysis*. *streammining.ijs.si*.

Zaki, M. (2013). *An Ontological Approach for Monitoring and Surveillance Systems in Unregulated Markets*. Manchester.

Zaki, M., Diaz, D., & Theodoulidis, B. (2012). Financial Market Service Architectures: A “Pump and Dump” Case Study. *2012 Annual SRII Global Conference*, 554–563. doi:10.1109/SRII.2012.68

Zaki, M., Theodoulidis, B., & Solis, D. D. (2011). A Data Mining Approach for the Analysis of “Stock-Touting” Spam Emails. *Journal of Manufacturing Technology Management*, *22*(6), 70–79.

Zhang, W., & Skiena, S. (2010). Trading Strategies to Exploit Blog and News Sentiment. In *4th Int’l AAAI Conference on Weblogs and Social Media (ICWSM)*.