Detection of Collusive Behavior in Energy Markets *

Radu-Casian Mihailescu, Matteo Vasirani and Sascha Ossowski

Rey Juan Carlos University
Centre for Intelligent Information Technologies
Madrid, Spain
{raducasian.mihailescu, matteo.vasirani, sascha.ossowski} @urjc.es

Abstract. Fundamental changes in the electrical energy sector are drawing on serious implications. One key arising challenge regards current energy markets, which are undergoing a transformation towards accommodating a more decentralized and sustainable provision of energy. As the number of traders in the market is increasing steadily and the trading activities are becoming more complex, the energy markets are becoming more exposed to potential fraud. In this paper we address the problem of detecting collusive behavior, where a group of individual traders act together, inconsistently with the competitive model, to artificially manipulate the market and elicit illegal profits. We investigate collusion attacks in the energy market and propose a novel mechanism, showing the effectiveness and practical applicability of our method to real scenarios.

1 Introduction

Market surveillance represents a serious challenge and it refers broadly to the detection of abnormal market behaviors, which are known to be predominant especially in the emerging markets. In practice, an efficient way for influencing the market and gaining illegal profits is represented by the class of collusion-based malpractices. Conceptually, collusive behavior represents an attempt of a group of individual traders, that act together, to artificially manipulate the market (e.g. through price or market share) for maximizing their gains, in a manner inconsistent with a competitive model and in detriment to the other participants in the market.

While collusion has been reported in various market domains, the damages discovered were averaged at about a 25% increase in costs incurred by costumers [6]. Unfortunately, most of the collusion cases that have been discovered thus far came as a result of investigations triggered not through economic analysis, but rather due to customers' complaints or suspicious competition complaints (e.g. stainless steel industry, graphite electrodes, facsimile paper).

Though there are many ways in which collusion could be discovered, we address in this work collusion detection via the analysis of economic data, freely

* AT2012, 15-16 October 2012, Dubrovnik, Croatia. Copyright held by the author(s).

available in the market. The set of laws that regulate the market require continuous surveillance of trading activities. This can essentially be achieved either through online or offline surveillance. The former is constrained to analysing short-term data as well as being restricted to a limited time-window, thus being prone to overlooking occurrences of more complex types of fraud. Alternatively, offline techniques can encompass a wider spectrum of illegal trading strategies, while having an anticipatory or retroactive character. Obviously though, both approaches remain dependent on the amount of input data available.

In this paper, we propose an effective offline method to identify collusive groups with respect to the domain of energy markets, which are being reoriented towards replacing the traditional top-down energy supply with a decentralized, market-oriented provision [10, 7, 8].

This paper is organized as follows: in Section 2 we provide a survey of some related work. Section 3 introduces our novel mechanism and discusses how to apply it to detect collusive behavior in the energy market. In Section 4 we show the experimental results. Finally, Section 5 concludes the paper and points directions for future work.

2 Related Work and Energy Markets

Emerging financial markets are inherently exposed to malpractices, whereas a subset of traders collaborate tacitly to manipulate the market for maximizing their individual gains. In the course of the last years, new challenges in the *Electricity Market* have come to the forefront, due to the transformations occurring in the power supply infrastructure. As new distributed energy resources (DERs: e.g. wind plants, photovoltaics, combined heat and power units) are pervading the electricity grid, they provide for an increasing number of participants in the market. This trend has driven the liberalization of electricity markets and the creation of power exchanges with the emphasis on decentralized power provision.

More formally, the participants in the market, consumers and suppliers of power, can be denoted by the *market agent* set $\mathcal{A} = \{a_1, \ldots, a_n\}$, which exists in a bijective relationship with the set of devices $\mathcal{D} = \{d_1, \ldots, d_n\}$ connected to the grid. The day-ahead market is discretized over a nonempty and finite set of distinct and successive time periods $\mathcal{T} = \{t_1, \ldots, t_m\}$. Under this setting, a bid of agent a_i timestamped at tmp_i^j is represented by the function $b_i^j : \mathcal{T} \rightarrow \mathbb{R} \times \mathbb{R}$, specifying respectively, for time slot t_i, the amount of electricity requested or offered and the intended price per unit, as $b_i^j(t_i) = (v_j, p_j)$. To summarize the trading activity for agent a_i at trading day tmp_i^j we associate the time-series $X_i^j = \{b_i^j(t_1), \ldots, b_i^j(t_m)\}$, which captures the existing bids for each time-slot in $\mathcal{T}$, else being considered a null bid.

Several approaches have attempted to automate the process of collusion detection based on economic data. A common method would be to apply supervised learning techniques, given that proven fraudulent activities, previously detected, could be obtained. Of course such datasets for training are not usually available in significant amount. Thus, some work has looked into unsupervised learning, namely using graph clustering algorithms. In [4], the collusive marker that the authors base their detection mechanism on, addresses circular trading, where a group of market agents are trading heavily among themselves aiming to raise the price of their shares. A Markov clustering algorithm is introduced and applied to the stock flow graph, which summarizes a trading database. In [9], the authors have adopted cross trading as a collusive marker and compared the performance of different off-the-shelf clustering algorithms. A similar approach is introduced in [2], by means of employing a spectral clustering method.

In this paper we address the phenomenon of collusion in the emerging energy markets where the collusive markers identified above do not hold, as novel domain driven markers need to be derived.

3 Collusion Detection Methodology

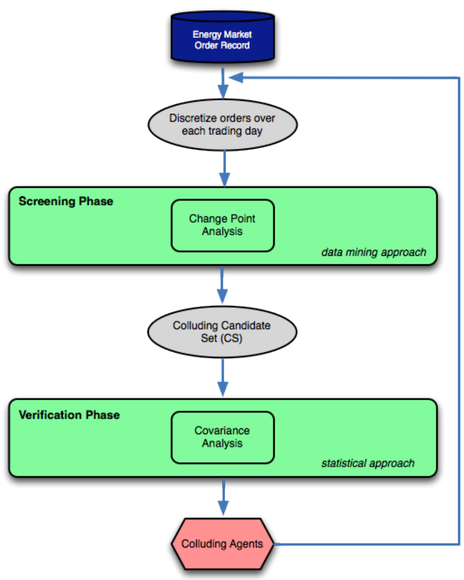

Fig. 1. Mechanism flow-chart

To start with, for discovering collusion in a market we need to be specific as to what behavioral patterns, that might be indicative of collusion, we will be looking for.

Our goal is to detect the presence of colluders that are coordinating their behavior at the expense of the rest of the market participants, by adopting a behavior inconsistent with what a competitive environment might entail. In this work, the *collusive marker* that we investigate in order to provide evidence of economic collusion consists of detecting colluders in the energy market based on the similar trading behaviors of agents. Colluders can generally be differentiated by similar trading patterns (which also depart from competition), as opposed to those outside their coalition. Thus, within a coalition of colluders their trading behavior should exhibit correlations when they should normally be independent.

Consequently, we design our collusion detection method as a two-stage process: a *screening phase* and a *verification phase*, as depicted in the diagram of Fig 1. The *screening phase* has the role of performing a triage through the *order record* of the market and identifying a set of market agents that are worthy of closer scrutiny. Due to the combinatorial explosion of possible colluder subsets and the data-intensive analysis, this phase is ought to output a set of candidates that will be further addressed during the *verification phase*.

Specifically, the *screening phase* is testing each market agent, looking for structural breaks in its behavior. Such behavioral breakpoint could be associated with the formation of a coalition of colluders (or with its dissolution). We approach this by running a *change-point analysis (CPA)* over the discretized order record of each market agent. If any behavioral breakpoints, which may be conductive to colluding coalitions, have been identified, they will represent the input for the second phase. Obviously such behavioral changes could be expected even when no collusion takes place. During *verification* we are essentially looking at two aspects: *i)* whether there is a price or volume correlation between the candidates' breakpoint and *ii)* if there appears inconsistencies with the competitive model.

While this methodology preserves a broad outlook into the realm of collusion detection in financial markets in general, further approaches can be directly derived for different contexts based upon the available data of the specific markets and agents. Additional domain-specific insights (see case study in Section 4.2), such as inferring estimates of costs, may ease the distinction between collusion and competition.

3.1 CPA based Behavioral Screening Phase

This section demonstrates the potential usefulness of change-point analysis techniques for detection of colluding behavior in Energy markets. The approach undertaken in this paper falls under the class of nonparametric change point detection methods [1], that do not rely on pre-specified parametric models and thus avoid strong model assumptions.

Change-point detection is the problem of discovering time points at which properties of time-series data change. Suppose that $X = \{x_1, \ldots, x_n\}$ is a sequence of independent random variables, such that the first r observations $X^A = \{x_1^A, \ldots, x_r^A\}$ are distributed as F_A and the remaining observations $X^B = \{x_1^B, \ldots, x_{n-r}^B\}$ come from another unknown distribution F_B, where $F_A \neq F_B$. Hence, integer r is called change point. Representing the trading activity of each market agent $a_i \in \mathcal{A}$ in terms of time-series enables us to perform a change-point analysis (Algorithm 1).

So, when a behavioral breakpoint occurs this corresponds to a step change of the mean value of X at r from α to $\alpha + \delta$, where δ represents the minimum increase of the mean value of X. Then, the deviation from the average may indicate that a collusion attack is being launched, if the cumulative deviation is noticeably higher than the random fluctuations, lower-bounded by δ.

A popular method based on a recursive nonparametric change point detection scheme uses a combination of cumulative sum charts (CUSUM) and bootstrapping to detect the changes [5]. The analysis begins with the construction of the CUSUM chart by calculating and plotting a cumulative sum, based on the timeseries data X.

The cumulative sums can be recursively defined using a new sequence $\{S_n\}$:

$$\begin{cases} S_n = S_{n-1} + (x_j - \bar{X}) \\ S_0 = 0 \end{cases} \tag{1}$$

where by $\bar{X}$ we denote the mean of the sample:

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{2}$$

Thus, the cumulative sum series can be obtained iteratively by adding to the previous sum the difference between a current value and the sample mean. This means that the sequence $\{S_n\}$ always ends at zero ($S_n = 0$), as the differences computed at each iteration sum to zero. Based on these remarks a CUSUM chart can be interpreted as follows. An upward slope of the chart indicates that the corresponding values tend to be above the overall mean of the sample, while a downward slope indicates a period of time where the values tend to be below. When a sudden turn occurs this indicates that around this time, the mean has shifted, which represents a potential changepoint. Likewise, a relatively straight CUSUM represents a period where the average did not change.

In order to associate a confidence level with a changepoint occurrence, a bootstrap analysis can be performed. To start with, a number of bootstrap samples are generated by sampling without replacement, which essentially is a random reordering of the original sample values. Thus, considering the initial sample X of size n, a bootstrapping sample at iteration i will be obtained by permuting without replacing k elements, generating $X_k^i = \{x_1^i, x_2^i, \ldots, x_n^i\}$. For each of these samples, the bootstrap CUSUM is computed similarly. Moreover, for each sample we need to determine the maximum, minimum and difference of the bootstrap CUSUM denoted respectively as $S_{max}^i = \max_{j=0,1,\ldots,n} S_j$, $S_{min}^i \min_{j=0,1,\ldots,n} S_j$ and $S_{diff}^i = S_{max}^i - S_{min}^i$. A bootstrap analysis consists of performing a large number of bootstraps and counting the number of bootstraps for which S_{diff}^i is less than S_{diff} of the initial sample. Let N be the number of bootstrap samples performed. Then, the probability of a changepoint occurrence for a fixed point r is given by:

$$\mathcal{P}_r = \frac{\#\{j : S_{diff}^i \leq S_{diff}\}}{N} 100[\%] \tag{3}$$

The bootstrapping technique basically compares the S_{diff} value of the original data with the S^i_{diff} values from a number of bootstrap samples, which estimate how much S_{diff} would vary if no change took place and checks whether these results are consistent. It is clear that a better estimate can be obtained by increasing the number of bootstrap samples, however statistically significant result can typically be obtained for a reasonable number of generations.

Now, if a changepoint has been detected, in order to determine when the change has occurred, different estimators can be employed. A straightforward approach would be to determine the changepoint r as the point furthest from zero in the CUSUM chart:

$$\mid S_r \mid = \underset{i=0,1,\ldots,n}{\operatorname{argmax}} \mid S_i \mid \tag{4}$$

The point r estimates last point before the change occurred, while the point $r+1$ estimates the first point after the change.

Alternatively, the changepoint occurrence can be estimated by applying the mean square error estimator (MSE). The idea behind this estimator is that of partitioning the data in two sequences $X_1 = \{x_1, \ldots, x_r\}$ and $X_2 = \{x_{r+1}, \ldots, x_n\}$ and estimating the mean of each sequence and compare this against the initial data:

$$MSE(r) = \sum_{i=1}^{r}(x_i - \bar{X}_1)^2 + \sum_{i=r+1}^{n}(x_i - \bar{X}_2)^2 \tag{5}$$

where $\bar{X}_1 = \frac{\sum_{i=1}^{r} x_i}{r}$ and $\bar{X}_2 = \frac{\sum_{i=r+1}^{n} x_i}{n-r}$.

Algorithm 1: Change point analysis

Data: X_n, N, k, δ
Result: changepoint r (if any)
$S_0 \leftarrow 0$;
for $i \leftarrow 1$ *to* n **do**
    $S_i \leftarrow S_{i-1} + (x_i - \bar{X})$;
end
for $j \leftarrow 1$ *to* N **do**
    generate bootstrap sample X_k^j, $S_0^j \leftarrow 0$;
    for $i \leftarrow 1$ *to* n **do**
        $S_i^j \leftarrow S_{i-1}^j + (x_i^j - \bar{X}^j)$;
    end
end
for $j \leftarrow 1$ *to* N **do**
    $S_{diff}^j = S_{max}^j - S_{min}^j$, $cnt \leftarrow 0$;
    if $S_{diff}^j \leq S_{diff}$ **then**
        $cnt \leftarrow cnt + 1$;
    else
    end
end
$\mathcal{P} = \frac{cnt}{N}$
if $\mathcal{P} \geq 1 - \delta$ **then**
    apply estimator to determine changepoint r
else
end

3.2 Verification Phase for Collusive Coalitions

Now, having identified a candidate set of potential colluders, which allows to narrow our search space, we proceed to the second phase. Here, the focus is on detecting correlations between any members of the candidate set, resulting in a coalition structure[1]. Considering two timeseries $X = \{x_1, x_2, \ldots x_n\}$ and $Y = \{y_1, y_2, \ldots y_n\}$ representing the trading activities of agents a_x and a_y respectively, we are interested to capture linear dependencies between the two variables: X and Y. As a measure of similarity, we evaluate the covariance, which determines how X and Y vary together:

$$Cov(X, Y) = \frac{1}{n} \sum_{i=1}^{n} (x(i) - \bar{x})(y(i) - \bar{y}) \tag{6}$$

where $\bar{x}$ is the sample mean of the X values and $\bar{y}$ is the sample mean of the Y values. The covariance measure ranges from 1 for perfectly correlated results,

[1] If all members of the candidate set show correlations we refer to this as the grand colluding coalition, while no correlations corresponds to the empty set.

through 0 when there is no relation between X and Y, to -1 when the results are perfectly correlated negatively.

More generally, if we consider k variables we can construct the covariance matrix $(k \times k)$, where an element (i, j) represents the covariance between the ith and jth variables. Removing the dependence of the covariance on the ranges of the variables can be done by standardization, dividing the result by the standard deviations of X and Y. The result is the correlation coefficient between X and Y:

$$\rho(X,Y) = \frac{\sum_{i=1}^{n}(x(i)-\bar{x})(y(i)-\bar{y})}{\left(\sum_{i=1}^{n}(x(i)-\bar{x})^2 \sum_{i=1}^{n}(x(i)-\bar{x})^2\right)^{1/2}} \tag{7}$$

Algorithm 2 summarizes the verification phase according to the computations previously detailed.

Algorithm 2: Collusion detection

Data: order record set O from one trading day discretize over a fixed time slot sequence $\tau = \{t_1, \ldots, t_k\}$; set of market agents $A = \{a_1, \ldots, a_n\}$

Result: collusion candidate set $\mathcal{C} = \{S_1, \ldots, S_l\}$

```
C ⟵ ∅;
for each market agent a_i do
    extract order record X_i associated to agents a_i from O
    run change point analysis for given X_i
    if probability of change point occurrence ≥ 95% then
        C ← C ∪ a_i;
        determine change point r with estimator ;
    else
end
generating covariance matrix for elements a_i ∈ C , l = 1;
for i ← 1 to sizeof(C) do
    for j ← i to sizeof(C) do
        M(i, j) ← Cov(X_i, Y_i);
        if M(i, j) ≥ 1 − δ then
            S_l ← {a_i, a_j};
            if a_i ⊂ S_k or a_j ⊂ S_k then
                append S_l to S_k;
            else
                l ← l + 1;
                append S_l to C;
            end
        else
        end
    end
end
```

4 Experimental results

4.1 Data preparation

Prior to running our collusion detection mechanism, the dataset needs to undergo a pre-processing phase. As discussed in Section 2, we retain from the order record of the Energy Market, for every agent, a time-series for each day consisting of their bids, with respect to the predefined time-slots $\mathcal{T}$ of the day-ahead market. Therefore, agent a_i is characterized at day j by $X_i^j = \{b_i^j(t_1), \ldots, b_i^j(t_m)\}$.

Now, in order to run a meaningful change-point analysis over this data, detecting relevant behavioral breakpoints, we need to span the investigation over a time-window of several days. Moreover, we need to relate the agents' trading patterns to the temporal organization of the day-ahead market, $\mathcal{T}$. Specifically, let's assume a time-window of length l days and a fixed discretization of the day-ahead market $\mathcal{T} = \{t_1, \ldots, t_m\}$. This requires constructing for each agent a_i the set of time-series $X_k^i = \{b_1(t_k), \ldots, b_l(t_k)\}$, $k = \overline{1, m}$. Next, recall that a bid, $b_i^j(t_i) = (v_j, p_j)$, from an energy supplier consists of the amount of electricity offered and the intended price per unit. This further implies that for each X_k^i there corresponds a time-series denoting price $P_k^i = \{p_1(t_k), \ldots, p_l(t_k)\}$ and another for the intended trading volume $V_k^i = \{v_1(t_k), \ldots, v_l(t_k)\}$. This representation of the data is used during the *screening phase* for generating the collusion candidate set, as well as during the *verification phase* for detecting price or volume correlations.

4.2 Case study

In this section we report on results[2] obtained from applying our model to real datasets, collected from the Philippine Wholesale Electricity Spot Market (WESM) [3]. We ran the analysis on the Market Bids submitted by the trading participants over a one month time-window (January 2012), for the Luzon region. The list of registered WESM market participants consists of 60 members; prices are listed in Pesos per MWh; the nominated energy quantity is given in MW.

We proceed with the *screening phase* by conducting an exhaustive change-point analysis over the bid records of each market participant. Table 1 summarizes the results obtained at this stage, outlining the *colluding candidate set* as input for the following phase. For the given scenario we have generated 1000 bootstrap samples for each run of the algorithm. The results indicate that out of the total number of market agents, behavioral breakpoints have been detected for 10 agents, some of which exhibiting multiple ones. An illustration of this process is given in Figure 2 for market agent a_6. Figure 2a is a representation

[2] We remind the reader that the analysis of economic data only, has the role of discovering suspicious behavior and is not meant to provide conclusive evidence of collusion, nor substitute antitrust authorities, but rather to provide supporting evidence and triggers for deciding whether antitrust authorities should actively engage and further pursue such an investigation.

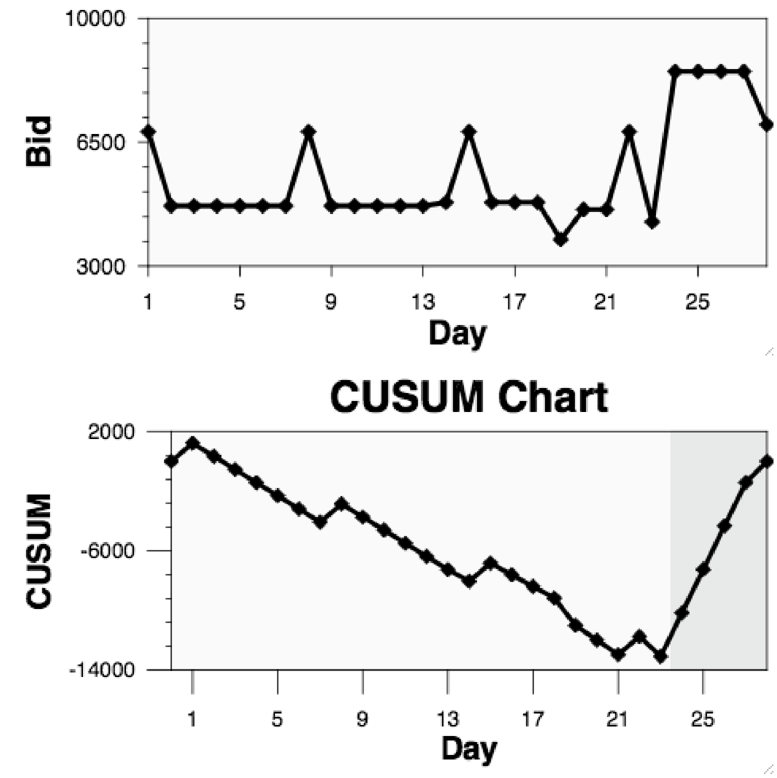

Fig. 2. a) Plot of the trading activity P_1^6 corresponding to market agent a_6 for January 2012. b) The associated CUSUM chart for the time-series of market agent a_6.

of a_6's trading activity during the specified time-window. The result of performing the screening analysis over this data, using the MSE estimator, detects the occurrence of one change-point timestamped at day 24. The confidence level associated, indicates a 100% accuracy. In Figure 2b we plot the corresponding CUSUM chart, highlighting a sudden shift in the average, associated with the change-point detection. Generally, it is not the case that change-points can be readily detected visually from the time series plot. A CUSUM chart however, can facilitate pinpointing shifts in the mean of the data, by identifying slope changes at the points where a change has occurred.

Next, the mechanism proceeds with the *verification phase*. As previously detailed, for the designated candidate set, generated during screening, we investigate further correlations between the market agents' trading patterns. This phase yields the covariance matrix, a representation of which is given in Figure 3. Here, we perform an exhaustive pairwise comparison of the candidate set. White squares denote a perfect correlation between the respective market agents, while black stands for no similarities. The color shading inbetween is indicative of the

Table 1. Results of the Change-Point Analysis with MSE Estimates for market agents with confidence level above or equal to 95%, representing the collusion candidate set.

Market Agent	Change-Point	Confidence Level
a_1	3	95%
a_2	7	97%
a_3	6	98%
a_3	20	100%
a_4	9	96%
a_4	12	99%
a_5	2	99%
a_5	22	100%
a_6	24	100%
a_7	17	100%
a_7	26	95%
a_8	12	100%
a_9	10	100%
a_9	15	100%
a_9	28	98%
a_{10}	24	100%

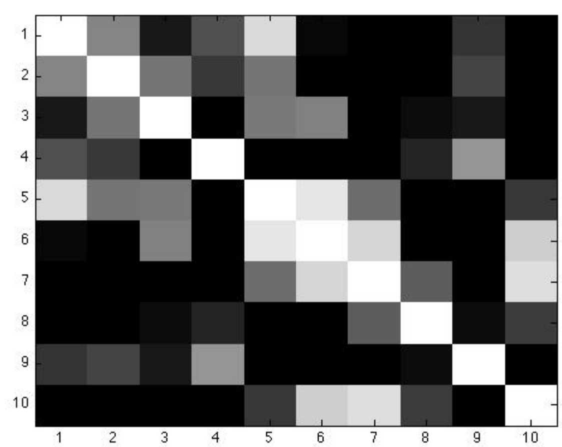

Fig. 3. Plot of the covariance matrix for the collusion candidate set of market agents using a grayscale range of colors. Correlations of 1.0 are plotted in white, while no correlation are plotted in black. The diagonal elements of the matrix represent self-correlations and are thus all 1.

correlation strength. Finally, according to Algorithm 2 the coalition structure of colluders is determined. For the considered scenario, the coalition structure of potential colluders consists of solely one group: $CS = \{\{a_1, a_5, a_6, a_7, a_{10}\}\}$. For privacy concerns, we omit other direct reference to the suspected colluders. Note that the particular choice of the value of the correlation coefficient threshold δ is ought to impact the number of resulting colluders. One one hand, higher thresholds imply a higher level of confidence for the collusion detection process, but on the other hand it may as well reduce the number of possible suspects, disregarding certain abnormal trading behaviors. Alternatively, lower thresholds may result in including false colluders to the coalition structure and therefore reducing the accuracy of the mechanism. In this context, selecting a reliable correlation coefficient threshold is an important issue for the overall performance of the mechanism, which we plan to address in future work. Specifically, we intend to calibrate the system based on already proven cases of collusion and use such scenarios as training data. In addition to this, we plan to extend the model to integrate details regarding the devices $\mathcal{D}$, which are controlled by the market agents $\mathcal{A}$, such as DER type and geographical location. Adding this domain-dependent dimension is ought to provide further insight into differentiating between correlations that may come as a result of external conditions (e.g. weather conditions) and those that are irrespective to this regard.

5 Conclusions

In this paper we have addressed the challenge of detecting collusive traders that collaborate illegally to increase their benefits at the expense of the other market participants. We have pose this question in the domain of the emerging energy markets, that are adapting to the integration of a diversity of distributed energy generators. Such contexts are especially susceptible to various trading malpractices.

The proposed method for discovering colluders consists of two phases. Firstly we apply a *screening phase* that performs a change-point analysis in order to detect behavioral breakpoint in the traders' activities, proposing a reduced candidate set of possible colluders. Secondly, for the denominated group we run a *verification phase* aimed at revealing behavioral correlations. The procedure determines a potential coalition structure of colluders. We evaluate our mechanism on real datasets and show the effectiveness and practical applicability of our method, even for scenarios that are exploiting a minimal amount of data, that is freely available on the market. Continuing along these lines, future work will further investigate and exploit other collusive markers that may expose potential vulnerabilities in the energy market.

References

1. Boris E. Brodsky and Boris S. Darkhovsky. *Nonparametric Methods in Change-Point Problems*. Kluwer Academic Publishers, 1993.
2. Markus Franke, Bettina Hoser, and Jan Schrder. On the analysis of irregular stock market trading behavior. In Christine Preisach, Hans Burkhardt, Lars Schmidt-Thieme, and Reinhold Decker, editors, *Data Analysis, Machine Learning and Applications*, Studies in Classification, Data Analysis, and Knowledge Organization, pages 355–362. Springer Berlin Heidelberg, 2008.
3. The Philippine Wholesale Electricity Spot Market (WESM): http://www.wesm.ph/.
4. M.N. Islam, S.M.R. Haque, K.M. Alam, and M. Tarikuzzaman. An approach to improve collusion set detection using mcl algorithm. In *Computers and Information Technology, 2009. ICCIT '09. 12th International Conference on*, pages 237 –242, dec. 2009.
5. Antoch Jaromir, Hukov Marie, and Veraverbeke Noel. Change-point problem and bootstrap. *Journal of Nonparametric Statistics*, 5(2):123–144, 1995.
6. Robert H. Lande and John M. Connor. How High Do Cartels Raise Prices? Implications for Reform of the Antitrust Sentencing Guidelines. *SSRN eLibrary*.
7. R.-C. Mihailescu, M. Vasirani, and S. Ossowski. Towards agent-based virtual power stations via multi-level coalition formation. *In Proceedings of the International Workshop on Agent Technologies for Energy Systems*, pages 107–108, 2010.
8. Radu-Casian Mihailescu, Matteo Vasirani, and Sascha Ossowski. Dynamic coalition adaptation for efficient agent-based virtual power plants. In *Proceedings of the 9th German ConferenceMultiagent System Technologies, MATES*.
9. Girish Keshav Palshikar and Manoj M. Apte. Collusion set detection using graph clustering. *Data Min. Knowl. Discov.*, 16:135–164, April 2008.
10. Sarvapali Ramchurn, Perukrishnen Vytelingum, Alex Rogers, and Nick Jennings. Putting the "smarts" into the smart grid: A grand challenge for artificial intelligence. *Communications of the ACM*, 2012.