# Show Me Your Friends and I'll Tell You Who You Are. Finding Anomalous Time Series by Conspicuous Cluster Transitions

Martha Tatusch[(✉)] , Gerhard Klassen , Marcus Bravidor ,
and Stefan Conrad

Heinrich Heine University, Universitätsstr. 1, 40225 Düsseldorf, Germany
{tatusch,klassen,bravidor,conrad}@hhu.de

**Abstract.** The analysis of time series is an important field of research in data mining. This includes different sub areas like trend analysis, outlier detection, forecasting or simply the comparison of multiple time series. Clustering is also an equally important and vast field in time series analysis. Different clustering algorithms provide different analysis aspects like the detection of classes or outliers. There are various approaches how to apply cluster algorithms to time series. Previous work either extracted subsequences or feature sets as an input for cluster algorithms. A rarely used but important approach in clustering of time series is the grouping of data points per point in time. Based on this technique we present a method which analyses the transitions of time series between clusters over time. We evaluate our approach on multiple multivariate time series of different data sets. We discover conspicuous behaviors in relation to groups of sequences and provide a robust outlier detection algorithm.

**Keywords:** Outlier detection · Time series analysis · Clustering

## 1 Introduction

Time series data is collected in various domains. Not only the behavior of users on different platforms, but also the tracking of vehicles and objects or the recording of financial or weather data can be displayed as time series. For further analysis, the various data types can be converted into numerical (mostly discrete) values so that sequences of numerical vectors are derived. These can then be processed in a variety of ways. Information can be obtained through analyses such as clustering, prediction or comparison of time series and different outlier detection methods.

Depending on the context, different aspects can be relevant for the user. For example, not all clustering algorithms consider the same types of clusters, and outlier detection techniques do not always address the same types of outliers. In some cases, very special solutions have to be found for specific problems, whereby there are many algorithms that can be applied to a wide range of application areas.
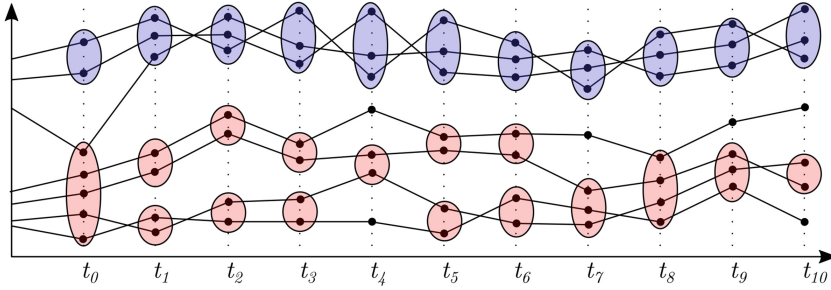
**Fig. 1.** Example for a time series over-time clustering. The blue color indicates stable clusters while red stands for instability. (Color figure online)

In this paper we focus on databases of multivariate time series with discrete values, same length and equivalent time steps. We detect anomalous subsequences with regard to groups of time series of the given database. Therefore we cluster the multivariate data of all time series per timestamp and analyze the stability of all subsequences over time. Thereby we call the resulting clustering *over-time clustering.* In Fig. 1 an example for such a clustering is displayed. For the sake of simplicity, only univariate time series are plotted. Since the data is clustered independently at each point in time, there is at first no time-related connection between the clusterings.

There are several proposals for clustering time series depending on the application. Some methods cluster the time series of a database as a whole [10,12,19], extract feature sets first [22], or consider subsequences of a single time series only [3]. However, these are not suitable when it comes to detecting irregularities or gathering information per time point.

Outlier detection in time series is in most cases not based on clustering. Because of various underlying data such as single or multiple time series with uni- or multivariate data points and different definitions of what an outlier is, there are several approaches to their identification. Some papers consider data points [1] or subsequences [15] that are anomalous with regard to a single time series [5,17], such as peaks. Others look for so called *change points* [6,16], that imply that the course of the considered time series significantly changes from that point on. Yet others analyse data from several time series that are very similar, such as sensor data, and detect irregularities in relation to the entire data set [1,11,13]. Finding these abnormalities usually presupposes that either the course of a single time series follows consistent patterns or that the courses of several time series are highly correlated.

In this paper we assume that the exact course of the individual time series is not important, but the trend which groups of sequences follow. By anomalies we denote subsequences that deviate from one trend and therefore cannot be assigned steadily to a group of sequences. In that case, we say that the sequence possesses a weak stability. We present an algorithm that identifies such unstable

sequences in a database of multivariate time series and is robust against missing data points.

## 2   Related Work

Anomaly detection in time series is a wide field of research. It can be distinguished in the detection of outliers within a single time series and the detection of outliers in multiple time series. Outliers in single time series are usually categorized in two classes:

*Additive outliers*, which represent surprisingly large or small values in a short period. In case additive outliers occur consecutively they are often summarized as additive outlier patches.

*Innovational outliers* are characterized by their impact on subsequent observations. Additionally the influence of innovational outliers can grow with time.

There are also several different categories of outliers, which can be described as a mix of both main classes. For example, additive outliers which cause a move of following observations to a new level are called *level shift outliers* and have a permanent impact on the ongoing time series. In case the influence of the level shift outlier is decreasing over time, it is called a *transient change outlier*. Additive outliers, which occur periodically are named *seasonal additive outliers*.

Additive and innovational outliers are often identified with extensions of autoregressive-moving-average (ARMA) models [2,18]. Other techniques include the use of decomposition methods such as STL, a seasonal-trend decomposition procedure based on LOESS [7]. Yet other methods evaluate derivatives of the dynamic time warping (DTW) [20] similarity in order to detect anomalies.

The detection of outliers in multiple time series is handled differently. Methods of this kind are often using the peers of a time series to determine whether it is anomalous or not. Beside other techniques, recent approaches use Probabilistic Suffix Trees (PST) [21] and Random Block Coordinate Descents (RBCD) [23] in order to detect outliers. However, while these approaches focus on the deviation of one time series to the others, we focus on the behaviour of a time series compared to its peers. More concretely, we assume that a time series which has a similar development to a group of other time series over a subsequence is expected to move on with the same group. Therefore we first cluster per point in time and then analyse the transition of time series regarding these clusters. This is realized by the analysis of cluster transitions of time series over time. Transitions of this kind are also analysed in cluster evolution methods. Landauer et al. [14] makes use of such a method in order to calculate an anomaly score for a single time series in a sliding window. In contrary to Landauer et al. we relate to multiple time series. The analysis of the time series behavior not only reveals new kinds of outliers but also detects different types of additive and innovational outliers.

This approach is very different from clustering whole time series or their subsequences, since the outlier detection would rely on the single fact whether a sequence is assigned to a cluster or not. Such an approach would not take

the cluster transitions of the time series into account, which can be an expressive feature on its own. Hence, our approach detects anomalous subsequences, although they would be assigned to a cluster in a subsequence clustering.

## 3    Fundamentals

In order to create a good basis of knowledge to avoid later misunderstandings, we will provide some definitions which our work is based on. As these terms are used in many different areas, it is useful to explain which interpretations are considered in this paper.

**Definition 1 (Time Series).**  *A multivariate time series $T = o_{t_1}, \ldots, o_{t_n}$ is an ordered set of $n$ real valued data points of arbitrary dimension. The data points are chronologically ordered by their time of recording, with $t_1$ and $t_n$ indicating the first and the last timestamp, respectively.*

**Definition 2 (Data Set).**  *A data set $D = T_1, \ldots, T_m$ is a set of $m$ time series of same length and equal points in time. The set of data points of all time series at a timestamp $t_i$ is denoted as $O_{t_i}$.*

**Definition 3 (Subsequence).**  *A subsequence $T_{t_i,t_j,l} = o_{t_i,l}, \ldots, o_{t_j,l}$ with $j > i$ is an ordered set of successive real valued data points beginning at time $t_i$ and ending at $t_j$ from time series $T_l$.*

**Definition 4 (Cluster).**  *A cluster $C_{t_i,j} \subseteq O_{t_i}$ at time $t_i$, with $j \in \{1, \ldots, q\}$ being a unique identifier (e.g. counter), is a set of similar data points, identified by a cluster algorithm or human. This means that all clusters have distinct labels regardless of time.*

**Definition 5 (Cluster Member).**  *A data point $o_{t_i,l}$ at time $t_i$, that is assigned to a cluster $C_{t_i,j}$ is called a member of cluster $C_{t_i,j}$.*

**Definition 6 (Noise).**  *A data point $o_{t_i,l}$ at time $t_i$ is considered as noise, if it is not assigned to any cluster.*

**Definition 7 (Clustering).**  *A clustering is the overall result of a clustering algorithm or the set of all clusters annotated by a human for all timestamps. In concrete it is the set $\zeta = \{C_{t_1,1}, \ldots, C_{t_n,q}\}$ of all $q$ clusters.*

In Fig. 2 an example for the above definitions can be seen. The data points of a data set containing five time series ($T_a$, $T_b$, $T_c$, $T_d$, $T_e$) are clustered for the timestamps $t_i, t_j$ and $t_k$. For simplicity, all data points of a time series $T_l$ are denoted by the identifier $l$.

In $t_i$ the data points $o_{t_i,a}, o_{t_i,b}$ of time series $T_a$ and $T_b$ are cluster members of cluster $C_{t_i,l}$. The data point $o_{t_i,e}$ is marked as noise, as it is not assigned to any cluster in $t_i$. In total, the shown clustering consists of five clusters. It can be described by the set $\zeta = \{C_{t_i,l}, C_{t_i,u}, C_{t_j,v}, C_{t_j,f}, C_{t_k,g}\}$.
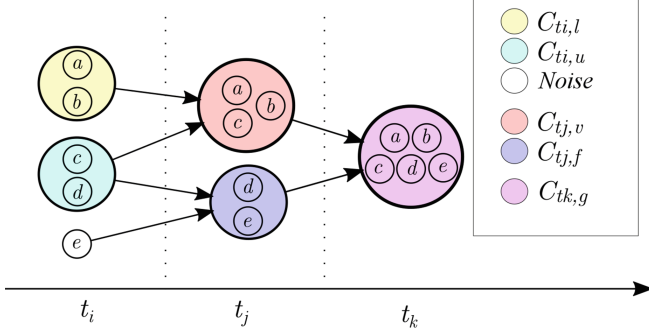
**Fig. 2.** Example for the transitions of time series $T_a, \ldots, T_e$ between clusters over time.

## 4   Method

After the clarification of important foundations, the basic idea of the algorithm is described. Therefore further terms have to be explained before.

Let $C_{t_i,a}$ and $C_{t_j,b}$ be two clusters, with $t_i, t_j \in \{t_1, \ldots t_n\}$. First, we introduce the term *temporal cluster intersection* for the purpose of measuring the stability of a time series:

$$\cap_t \{C_{t_i,a}, C_{t_j,b}\} = \{T_l \mid o_{t_i,l} \in C_{t_i,a} \wedge o_{t_j,l} \in C_{t_j,b}\}$$

with $l \in \{1, \ldots, m\}$. The result is the set of time series that are assigned to both of the clusters under consideration. This means all sequences that were grouped together at time $t_i$ <u>and</u> $t_j$. The transition of a time series from $t_i$ to $t_j$ can now be described by the proportion of cluster members from the corresponding cluster in $t_i$ who migrated together into the cluster in $t_j$:

$$p(C_{t_i,a}, C_{t_j,b}) = \begin{cases} \emptyset & \text{if } C_{t_i,a} = \emptyset \\ \frac{|C_{t_i,a} \cap_t C_{t_j,b}|}{|C_{t_i,a}|} & \text{else} \end{cases}$$

with $t_i < t_j$. In Fig. 2 an example for transitions of time series between clusters is sketched. There, the proportion for $C_{t_i,l}$ and $C_{t_j,v}$ would be

$$p(C_{t_i,l}, C_{t_j,v}) = \frac{|\{T_a, T_b\}|}{|\{o_{t_i,a}, o_{t_i,b}\}|} = \frac{2}{2} = 1.0.$$

This proportion can be used to measure the stability of a sequence with a *subsequence score*. It is defined as

$$subsequence\_score(T_{t_i,t_j,l}) = \frac{1}{k} \cdot \sum_{v=i}^{j-1} p(cid(o_{t_v,l}), cid(o_{t_j,l}))$$

with $l \in \{1, \ldots, m\}$, $k \in [1, j-i]$ being the number of timestamps between $t_i$ and $t_j$ where the data point exists and *cid*, the cluster-identity function

$$cid(o_{t_i,l}) = \begin{cases} \emptyset & \text{if the data point is not assigned to any cluster} \\ C_{t_i,a} & \text{else} \end{cases}$$

returning the cluster which the data point has been assigned to in $t_i$. The function returns an empty set, either if the object is classified as noise or if it does not exist at the considered time. Note, that the subsequence score is normalized to $[0,1]$ by $k$, as the proportion $p$ is a percentage between 0 and 1, as well.

In the example of Fig. 2, the score of time series $T_a$ between time points $t_i$ and $t_k$ would be:

$$subsequence\_score(T_{t_i,t_k,a}) = \frac{1}{2} \cdot (1.0 + 1.0) = 1.0.$$

A notable characteristic is, that the score is always 0, if the last data point of the considered subsequence is marked as noise. However, this circumstance does not lead to any handicaps in most cases as all partial sequences of these subsequences are treated normally. Nevertheless, the handling of sequences with an endpoint that is labeled as noise will be analyzed in more detail later on.

For now describing the concrete procedure of detecting conspicuous sequences, we first provide a vague definition of them:

**Definition 8 (Anomalous Subsequence).**  *A subsequence $T_{t_i,t_j,l}$ is called anomalous, if it is significantly more unstable than its cluster members at time $t_j$.*

With the help of the subsequence score which measures the stability of a subsequence, anomalous ones can now be distinguished by comparing the stability of grouped subsequences at a given time point. Every possible subsequence gets an outlier score indicating the probability of being anomalous, by calculating the deviation of its stability from the best subsequence score of its cluster. A formal description of the best subsequence score can be given by:

$$best\_score(t_i, C_{t_j,a}) = max(\{subsequence\_score(T_{t_i,t_j,l}) \mid cid(o_{t_j,l}) = C_{t_j,a}\})$$

The outlier score of a subsequence is then calculated as follows:

$$outlier\_score(T_{t_i,t_j,l}) = best\_score(t_i, cid(o_{t_j,l})) - subsequence\_score(T_{t_i,t_j,l})$$

As the best score lies between 0 and 1, an outlier score of 100% can only be achieved in completely stable clusters. The smaller the best score of the considered cluster is, the smaller is the greatest possible outlier score.

Regarding the example in Fig. 2, the time series $T_d$ would get the following *outlier_score* between time points $t_i$ and $t_k$:

$$outlier\_score(T_{t_i,t_k,d}) = 1.0 - (0.5 \cdot (0.5 + 1.0)) = 0.25$$

With the outlier score, now a more precise definition of an outlier can be given.

**Definition 9 (Outlier).**  *Given a threshold $\tau \in [0,1]$, a subsequence $T_{t_i,t_j,l}$ is called an outlier, if its probability of being an outlier is greater than or equal $\tau$. That means, if*

$$outlier\_score(T_{t_i,t_j,l}) \geq \tau.$$

Although $\tau$ is a constant, it can be interpreted as a dynamic threshold. That is, because the greatest possible deviation from the best subsequence score – and thus the greatest outlier score – depends on the best score of the considered cluster. Clusters with low stability have a lower probability of containing an outlier than stable ones, since all their cluster members show irregularities and that represents a pattern of instability. In this context, the small subsequence score is thus not conspicuous.

Intuitive outliers from the over-time clustering that were marked as noise get a special treatment. Subsequences that consist entirely of noise data points are automatically identified as outliers. Since subsequences whose last data point is labeled as noise are not assigned to a cluster from which the best score can be calculated, no outlier score can be determined for them. Therefore, they are not included in the regular outlier calculation. In the following we will differentiate between *anomalous subsequences*, *intuitive outliers* and *noise*.

Take another look at the case where the last element of an examined subsequence $T_{t_i,t_j,l}$ is marked as noise. Suppose the subsequence $T_{t_i,t_{j-1},l}$ gets a high outlier score and is detected as outlier. Then one would expect that the subsequence under consideration $T_{t_i,t_j,l}$ would be identified as an outlier as well. This will only be the case, if the previous data point was categorized as noise as well and the sequence was therefore recognized as an intuitive outlier. However, for the sequence $T_{t_i,t_k,l}$ with $k > j$, which at the last time point $t_k$ is assigned to a cluster again for the first time this would also be the case. Thus in the end $T_{t_i,t_j,l}$ would be covered.

Yet a marginal case is when a data point is labeled as noise at the last time of the entire time series. In this scenario, a sequence with end time $t_m$ would never be detected as an outlier if it is not marked as noise in $t_{m-1}$.

*Remark 1 (Stability).* The stability is not only influenced significantly by a small sample size when considering constant data points [4]. When examining the over-time stability, a small sample size leads to high sensitivity to cluster transitions, as well. As more data points are considered, the simpler it is to draw meaningful conclusions about the stability.

## 5   Experiments

In order to evaluate the presented method, we performed several experiments on different real world data. We also present two artificially generated data sets which are used to illustrate the handling of some marginal cases. In order to cluster the data per point in time, we used DBSCAN [9] with adapted parameters.
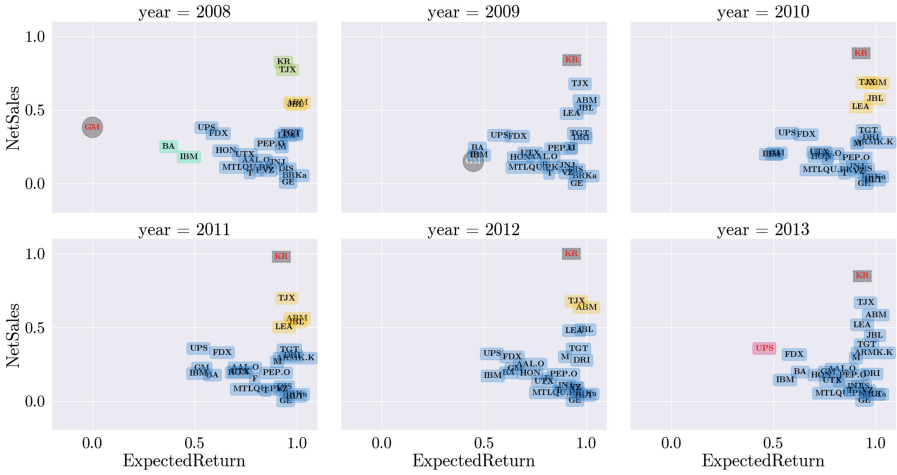
**Fig. 3.** Two dimensional experiment on the EIKON Financial Data Set with $\tau = 0.6$, $minPts = 2$ and $\epsilon = 0.15$. The colors indicate cluster belongings, whereby grey objects represent outliers. Circles are outliers by distance and boxes are intuitive outliers, as well. Red color or font indicates noise. (Color figure online)

## 5.1   EIKON Financial Data Set

Eikon is a set of software products released by Refinitiv (formerly Thomson Reuters Financial & Risk). It contains a database with financial data of thousands of companies for the past decades. For illustration reasons we randomly selected thirty companies and two features. The selected features are a figures which were taken from the balance sheet of the company. In economics it is common to normalize these figures by the companies' total assets in order to make it comparable to other companies. Beside this, we normalized the features with a min-max normalization. The clustering was done with DBSCAN and $\epsilon = 0.15$, $minPts = 2$ as parameters. The outlier detection parameter was chosen to be $\tau = 0.6$. In Fig. 3 one can see the illustrated results. The presented technique found two outlier subsequences. The first, which is labeled as $GM$ is detected from the year 2008 until 2009. This is because $GM$ is noise in the year 2008, which leads to a subsequence score of 0. In 2009 $GM$ merges with a cluster, which has a high reference score. The second outlier detected is the subsequence $T_{t_{2009}, t_{2013}, KR}$. It is detected as an intuitive outlier.

## 5.2   Airline On-Time Performance Data Set

The Airline on-time performance data set [8] was originally collected by the U.S. Department of Transportation's Bureau of Transportation Statistics. It contains records of 3.5 million flights. Every flight has a set of 29 features, such as the departure delay, the delay reason, the arrival delay and the airline which processed the flight. In order to detect anomalies in this data set, we constructed
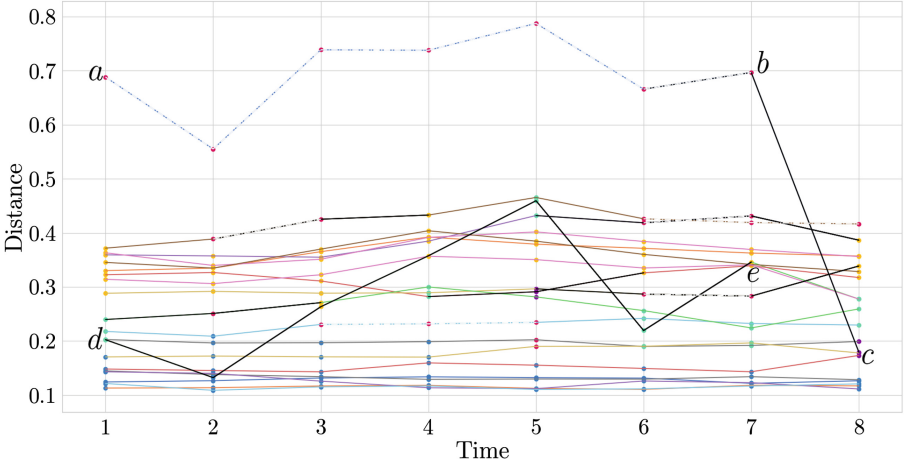
**Fig. 4.** One dimensional experiment on the Airline On-Time Performance Data Set with $\tau = 0.4$, $minPts = 3$ and $\epsilon = 0.03$. Black sequences represent anomalies, while white dashed ones stand for intuitive outliers. The color of the dots emphasize which cluster the data points are assigned to. Red dots represent noise. (Color figure online)

a time series for every airline by calculating the average of their features for every day. Before applying our technique, we normalized the data with the min-max normalization and clustered it with DBSCAN. Every observation represents a flight of an airline. In order to illustrate the results we executed our algorithm to one feature, namely the flight distance. We applied DBSCAN for eight time points with the following parameters: $minPts = 3$ and $\epsilon = 0.03$. Additionally we chose $\tau = 0.4$. The result can be seen in Fig. 4.

The figure shows two kinds of outliers: Intuitive outliers and outliers which were identified by their distance to a reference time series. Since the time series which is labeled with the points $a$, $b$ and $c$ has a large distance to other time series it is detected as an intuitive outlier from $a$ to $b$. Due to this, the time series' accumulated subsequence score is zero and thus it is also detected as an outlier at the last time stamp $c$. From point $a$ to $b$ it is not detected as an outlier by it's distance to the reference subsequence score, since the neighborhood of the sequence at time point 8 have also a low stability score. Regarding the time points 1 to 8 and the objects in the neighborhood, there are at most two peers which remained together. The subsequence labeled with $d$ and $e$ is a good example for the presented method. It illustrates the detection of outliers by the change of cluster neighbors of the subsequence.

### 5.3   Simulated Data

In order to test our method in a targeted manner, two experiments were performed on simulated data. Both a univariate and a multivariate data set with two features are considered. In both cases, a time span of 8 time points is examined.
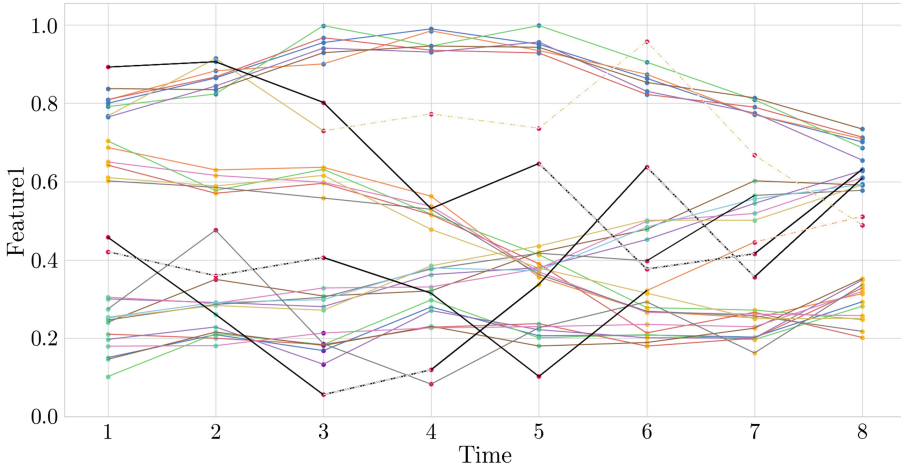
**Fig. 5.** Illustration of the detected outliers on the simulated one-dimensional data set with $\tau = 0.55$, $minPts = 3$ and $\epsilon = 0.05$. Black sequences represent anomalous subsequences, while white dashed ones stand for intuitive outliers. The color of the dots emphasize which cluster the data points are assigned to. Red dots represent noise. (Color figure online)

The one-dimensional data set was generated so that initially four starting points (for four groups) were selected. In addition, the maximum deviation from the centroid and the number of members were chosen for each group. The centroids were then calculated randomly for each time point, whereby the distance of the centroids of a cluster of two successive time points could not exceed 0.06. After generating the normal data points, 5 outlier sequences were randomly inserted. The starting points were chosen randomly and the distance between two consecutive points could not be greater than 0.3. For all points, care was taken to ensure that they were between 0 and 1.

As shown in Fig. 5, anomalous sequences from five time series have been found. Regarding the first time stamp the first and second black line show time series that are entirely recognized as conspicuous ones. Since their data points often switch between being noise (red dots) and different cluster members, this result is meaningful. Between time point 6 and 7 one additional black line in added. This can be explained by the stability of the sequence's cluster at time 7. All its cluster members migrate together from time point 6 to 7, so that an outlier is very conspicuous.

Looking at the completely randomly generated time series with the uppermost noise point at time 2, it is noticeable that it was not recognized by our algorithm. This is due to the fact that the purple cluster at time 3 and the turquoise cluster at time 5 do not have a high stability and the deviation of the sequence from the best possible score is therefore not very large. In the last time
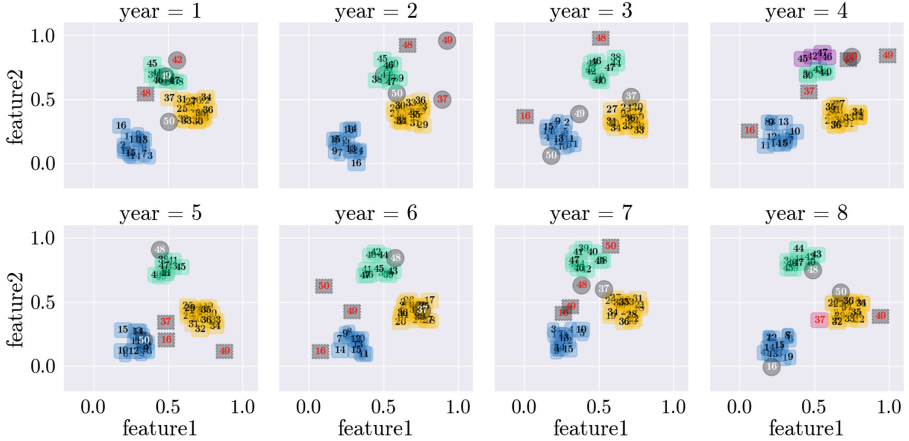
**Fig. 6.** Illustration of the detected outliers with $\tau = 0.5$, $minPts = 4$ and $\epsilon = 0.11$ on the artificially generated data two-dimensional set. The colors indicate cluster belongings, whereby grey objects represent outliers. Circles are outliers by distance and boxes are intuitive outliers, as well. Red color or font indicates noise. (Color figure online)

points, the time series migrates stably with the yellow cluster, so that it does not behave uncommonly.

If the data points of a time series change from one point in time to another from a cluster to noise, they are not initially interpreted as conspicuous. This is a problem if the time series remains as noise as the time at which it split from the cluster is not recognized as an intuitive outlier. This behavior can for example be seen in the striped line regarding the first time stamp. Between the times 6 and 7, the sequence was not detected as an outlier.

The second data set was created as follows: First, three starting points as centroids and the number of members of the three clusters were chosen. The maximum deviation of two consecutive centroids was set to 0.05 and that of the member data points to the centroid was set to 0.1. One time series was assigned to each group, which was allowed to deviate from the centroid by up to 0.25. Finally, two time series with completely random data points were added, so that a total of 5 outlier sequences should be noticeable. Here, too, we made sure that all data points are between 0 and 1 for each feature.

In Fig. 6 the results for an over-time clustering made by DBSCAN with $minPts = 4$ and $\epsilon = 0.11$ and an outlier threshold of $\tau = 0.5$ are shown. The time series 16, 37, 48 are generated with higher deviation and 49 and 50 completely random. It can be seen that all these time series were found by our algorithm as outliers (grey). Since the data points of these time series often are outliers as well as change their cluster members, this is a correct result. However, the first two time points are assumed to be normal for time series 16. This is desired too, as it moves stable with its cluster members at this time.

Although the data points of 42, 45, 46 and 47 split from their cluster members in time point 4, they are not identified as outliers. Since they migrate together and even merge back to their former cluster members in time point 5, their behavior is not conspicuous. The sequence 42 is identified as anomalous between time points 1 and 2 (turquoise cluster), since all its cluster members migrated completely stable from time point 1 to 2.

In total, the following outlier sequences can be read from Fig. 6: $T_{3,8,16}$, $T_{1,2,42}$, $T_{1,7,37}$, $T_{1,8,48}$, $T_{1,8,49}$, $T_{1,8,50}$. All are justified and correspond to the desired result. There is one striking observation, though: Although 37 is conspicuous over the entire period, it is only found as outlier between time 1 and 7. The reason for this is that the marginal case mentioned in Sect. 4 has occurred. Since the data point of the time series was classified as noise at the very last point in time, but not at the time before, the sequence is not found by our algorithm.

## 6    Conclusion and Future Work

In this work we presented a robust outlier detection algorithm for multiple multivariate time series. By analyzing the cluster transitions of time series over time, we are able to identify anomalous sequences. Instead of using sliding windows, our method performs an analysis of all possible subsequences. The shown results are sound and enable a new field of research. However, there are still some interesting aspects which may be examined in future work. The most important issue is the determination of the outlier detection parameter $\tau$. We assume an interdependence of $\tau$ and hyperparameters that are used for the clustering algorithm. Further not all intuitive outlier sequences have to be conspicuous in regard to the time series database. Considering the deviation of time series can lead to an enhanced analysis of those. Finally, it could be useful to identify whole outlier clusters. Therefore a cluster score could be computed and evaluated.

## References

1. Ahmad, S., Lavin, A., Purdy, S., Agha, Z.: Unsupervised real-time anomaly detection for streaming data. Neurocomputing **262**, 134–147 (2017)
2. Ahmar, A.S., et al.: Modeling data containing outliers using ARIMA additive outlier (ARIMA-AO). J. Phys: Conf. Ser. **954**, 012010 (2018)
3. Banerjee, A., Ghosh, J.: Clickstream clustering using weighted longest common subsequences. In: Proceedings of the Web Mining Workshop at the 1st SIAM Conference on Data Mining, pp. 33–40 (2001)
4. Ben-David, S., Von Luxburg, U.: Relating clustering stability to properties of cluster boundaries. In: 21st Annual Conference on Learning Theory (COLT 2008), pp. 379–390 (2008)
5. Cheng, H., Tan, P.N., Potter, C., Klooster, S.: Detection and characterization of anomalies in multivariate time series. In: Proceedings of the 2009 SIAM International Conference on Data Mining, pp. 413–424 (2009)

6. Cho, H., Fryzlewicz, P.: Multiple change-point detection for high-dimensional time series via sparsified binary segmentation (2014)

7. Cleveland, R.B., Cleveland, W.S., McRae, J.E., Terpenning, I.: STL: a seasonal-trend decomposition procedure based on loess (with discussion). J. Official Stat. **6**, 3–73 (1990)

8. ASA Statistics Computing and Graphics: Airline on-time performance. http://stat-computing.org/dataexpo/2009/the-data.html. Accessed 15 July 2019

9. Ester, M., Kriegel, H.P., Sander, J., Xu, X.: A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, pp. 226–231 (1996)

10. Ferreira, L.N., Zhao, L.: Time series clustering via community detection in networks. Inf. Sci. **326**, 227–242 (2016)

11. Hill, D.J., Minsker, B.S.: Anomaly detection in streaming environmental sensor data: a data-driven modeling approach. Environ. Model Softw. **25**(9), 1014–1022 (2010)

12. Huang, X., Ye, Y., Xiong, L., Lau, R.Y., Jiang, N., Wang, S.: Time series k-means: a new k-means type smooth subspace clustering for time series data. Inf. Sci. **367–368**, 1–13 (2016)

13. Keogh, E., Lonardi, S., Chiu, B.Y.C.: Finding surprising patterns in a time series database in linear time and space. In: Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2002, pp. 550–556 (2002)

14. Landauer, M., Wurzenberger, M., Skopik, F., Settanni, G., Filzmoser, P.: Time series analysis: unsupervised anomaly detection beyond outlier detection. In: Su, C., Kikuchi, H. (eds.) ISPEC 2018. LNCS, vol. 11125, pp. 19–36. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-99807-7_2

15. Lin, J., Keogh, E., Fu, A., Van Herle, H.: Approximations to magic: finding unusual medical time series. In: 18th IEEE Symposium on Computer-Based Medical Systems (CBMS 2005), pp. 329–334 (2005)

16. Liu, S., Yamada, M., Collier, N., Sugiyama, M.: Change-point detection in time-series data by relative density-ratio estimation. Neural Netw. **43**, 72–83 (2013)

17. Malhotra, P., Vig, L., Shroff, G.M., Agarwal, P.: Long short term memory networks for anomaly detection in time series. In: ESANN (2015)

18. Munir, M., Siddiqui, S.A., Chattha, M.A., Dengel, A., Ahmed, S.: FuseAD: unsupervised anomaly detection in streaming sensors data by fusing statistical and deep learning models. Sensors **19**(11), 2451 (2019)

19. Paparrizos, J., Gravano, L.: k-shape: efficient and accurate clustering of time series. In: Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, pp. 1855–1870 (2015)

20. Salvador, S., Chan, P.: Toward accurate dynamic time warping in linear time and space. Intell. Data Anal. **11**(5), 561–580 (2007)

21. Sun, P., Chawla, S., Arunasalam, B.: Mining for outliers in sequential databases. In: ICDM, pp. 94–106 (2006)

22. Truong, C.D., Anh, D.T.: A novel clustering-based method for time series motif discovery under time warping measure. Int. J. Data Sci. Anal. **4**(2), 113–126 (2017)

23. Zhou, Y., Zou, H., Arghandeh, R., Gu, W., Spanos, C.J.: Non-parametric outliers detection in multiple time series a case study: power grid data analysis. In: AAAI (2018)