Finding arbitrary shaped clusters with related extents in space and time

C. Pölitz[1] and G. & N. Andrienko[1]

[1]University of Bonn & Fraunhofer IAIS, Germany

Abstract

The paper deals with density-based clustering of events, i.e. objects positioned in space and time, such as occurrences of earthquakes, forest fires, mobile phone calls, or photos with geographical references. Finding concentrations of events in space and time can help to discover interesting places and time periods. The spatial and temporal properties of event clusters, in particular, their spatial and temporal extents and densities, can be related to each other. Therefore, we suggest a two-step clustering method that considers relations of the spatial and temporal dimension. We demonstrate the work of the method on several examples of real data.

1. Introduction

The focus of this work is detecting spatio-temporal clusters (concentrations) of events. By events we mean objects positioned in space (in particular, geographical space) and time. Examples of such objects are occurrences of earthquakes, forest fires, disease cases, mobile phone calls, and photos taken by Flickr or Panoramio users. Concentrations of events in space and time may indicate interesting and/or important phenomena. Places with re-appearing concentrations of events may be interesting or problematic, depending on the nature of the data.

Clusters of events may have different shapes in the space-time continuum. A cluster wide in space but narrow in time means that many events occurred in a large area during a short time interval. Such a cluster in forest fires data may correspond to spreading of forest fires due to weather conditions, in disease data - to a disease outbreak, and in Flickr photos data - to festivities taking place throughout a city or region. A cluster compact in space but extended in time means that events re-occur in the same place. Such a cluster in forest fires data may indicate arsons, in Flickr photos data - a place attractive for tourists, and in earthquakes data - a geological fault.

Our work aims at finding arbitrarily shaped clusters in space and time. The density-based clustering techniques are the most suitable for this task. We extend such a method to find spatio-temporal clusters in event data [GAA04].

Further, the extents and densities of the clusters in the spatial and temporal dimensions can be related. Our ideas are based on the first law of geography by W. Tobler saying "Everything is related to everything else, but near things are more related than distant things" [Tob70]. We apply this principle separately to the spatial and temporal components of the data. Thus, events that occurred in the same place or very closely in space can be related even in case of long time intervals between them. Similarly, events that occurred simultaneously or close in time may be related even if they are not very close in space. Therefore, we propose to relax the time constraints for clusters with small spatial extent and to consider strict time limits for clusters that cover large areas. According to this idea, we suggest a method consisting of two clustering steps. In the first step, all spatial clusters are extracted by means of DBSCAN with a user-specified spatial distance threshold. In the second step, DBSCAN is applied again to each spatial cluster with a certain temporal distance threshold, which is chosen depending on the characteristics of the spatial cluster.

To make the method applicable to large data sets, we enhance the clustering by using index structures that speeds up data access. For geographical and temporal data, an index structure like R-Tree or kd-Grid can be used. These data structures reduce the average time for finding neighbours in the space of the events from linear to logarithmic. Such nearest neighbour searches are repeated up to all points in the data set during our proposed algorithm.

2. Related Work

The clustering task is a well investigated area in computer science. There are several different approaches with different strengthes and weakenings. For this work the density based clustering method DBSCAN is of special interest. Our proposed method is based on this method by Kriegel et. al. [EKJX96]. Furthermore additional research has been done on data with certain attribute types like geographical positions or timestamps.

Birant and Kut [BK07] do spatial temporal clustering with DBSCAN using two distance functions and two distance thresholds for the maximum distance in the spatial and in the temporal space. This leads to clusters that are arbitrarily shaped in space and time.

In [AA09] Andrienko et. al. describe also a spatial temporal cluster analysis of event data with two fixed thresholds for spatial and temporal distance. Here the focus is on the visual exploration of events with geographical position and timestamp. The visual representations of the clustering supports reasoning of the gained results.

In [NM04] Neill and Moore propose a method for detecting clusters in space and time. It is assumed that certain subsets of the data have entries that follow some distribution. These subsets are regions with significantly higher numbers of entries than expected according to the assumed distribution. The found regions are hyperrectangles containing a set of entries within a certain time interval.

Iyengar [Iye04] introduces a method to discover clusters in space and time that grow or shrink. This is done by searching for pyramidical shapes in space-time containing points very likely being generated by a Poisson process.

Additionally, a space-time clustering analyses of seismic activity in a mine is outlined by Kijko and Funk [KF96]. They use a single link clustering as explained by Day et. al. [DE84]. The applied link distance to create clusters depends on spatial and temporal distances as well as on a regulation value. This regulation value makes sure that simultaneous events with large spatial distance d are in one cluster, as well as events with the same position and a temporal distance of t.

Index structures that can be applied to spatial data like R-Trees and kd-Trees are described by Guttman [Gut84] and Bentley [Ben75].

3. Method

The suggested method acts in a stepwise manner. First, spatial clustering with the DBSCAN algorithm is applied with a spatial distance threshold ε^s that is given by the user. This produces k spatial clusters $C_1, \cdots, C_k$. During the spatial clustering, a nearest neighbor query is performed for every yet unclustered entry. To speed up this process, a spatial kd-Tree index is used. This allows our method to be applied to large amounts of data within less processing time than simply doing sequential searches.

$$\varepsilon_i^t = T(C_i) = \min\{1, \frac{N}{n_i}\} \cdot \frac{D}{d_i} \cdot \varepsilon^t \tag{1}$$

$$n_i = \|\{e \in C_i\}\| \tag{2}$$

$$d_i = \sqrt{(max\{e.x\} - min\{e.x\})^2 + (max\{e.y\} - min\{e.y\})^2} \tag{3}$$

$$t_i = \sqrt{(max\{e.t\} - min\{e.t\})^2} \tag{4}$$

$$\rho_{d_i} = d_i/n_i \tag{5}$$

$$\rho_{t_i} = t_i/n_i \tag{6}$$

Figure 1: *The characteristics of the spatial clusters C_i, which are listed in expressions 2, $\cdots$,6, can be used for the definition of the temporal thresholds and also for filtering out uninteresting spatial clusters. In this way, the user can parameterize and control the further temporal clustering.*

Secondly, for each spatial cluster C_i, a further run of DB-SCAN clustering along the temporal dimension is applied using a temporal kd-Tree index for a better performance. The temporal distance threshold ε_i^t is computed according to equation 1. It depends on the number of entries n_i and the spatial extent d_i of the corresponding spatial cluster C_i. Three parameters must be specified by the user according to the desired properties of the clusters he/she wishes to detect: ε^t, N, D. ε^t is the maximum temporal distance between entries in a cluster that has at least N entries and a spatial extent of D. For instance to find tourist features within many "geotagged" photos a small spatial distance threshold of $\varepsilon^s = 5$ meters is appropriate. Further we state that within clusters of extent of $D = 20$ at least $N = 5$ photos should be made within the temporal threshold. As we assume that many photos being made there over a long time the temporal distance threshold should be large like $\varepsilon^t = 1$ hour. Besides the possibility to find the tourist features our proposed method would sharpen the temporal distance threshold for a spatially larger extended cluster C_i with more points, i.e. $\varepsilon_i^t < \varepsilon^t$. This configuration could identify a parade within the photo data having large spatial but low temporal extent.

According to equation 1, the individual temporal distance threshold of a cluster increases as the spatial extent shrinks. On the other hand, when the cluster is larger than the specified spatial extent D, the temporal distance threshold decreases. Hence, temporal clustering will detect dense temporal clusters within large spatial clusters and sparser temporal clusters within small spatial clusters. Additionally, the number of entries in the spatial cluster influences the temporal threshold in a way that the more entries there are, the smaller the threshold will be.

A different order of the clustering steps would be less effective. If we do temporal clustering first, we are very likely to retrieve very few large clusters greatly extended in time. This is due to the fact that there are less separation possibilities in the one dimensional time than in the two dimensions

Algorithm 1 Stepwise DBSCAN

1: ε^s *// set spatial density threshold*
2: Y = DBSCAN(X,ε^s) *//do spatial clustering*
3: SC = Y.getClusters *//retrieve spatial clusters*
4: **for all** $C_i \in SC$ **do** {for all found spatial clusters}
5:     ε_i^t = T(C_i) *//set temporal density threshold*
6:     Y = DBSCAN(C_i.instances, ε_i^t) *//do temporal cluster-ing*
7:     STC_i = Y.getClusters *//retrieve temporal clusters*
8:     sort(STC_i) *//sort temporal clusters w.r.t. time*
9:     F.put(i,STC_i) *//map list of temporal clusters to index i of corresponding spatial*
10: **end for**
11: **for all** $STC_i \in F$ **do** {for all temporal clusters of spatial cluster i}
12:     plot(STC_i) *//plot temporal cluster characteristics*
13:     **for all** $c \in STC_i$ **do** {for all clusters}
14:        draw(c) *//draw all spatial temporal clusters*
15:     **end for**
16: **end for**

of the geographical space.

The second step of clustering results in $0, \cdots, m$ temporal clusters in each spatial cluster C_i found during the first step of the method. These temporal clusters ($C_{i,j}$) have certain properties that may be important for analysis (analog to the properties of the spatial clusters 2): the number of entries contained $n_{i,j}$, the spatial and temporal extents $d_{i,j}, t_{i,j}$, the spatial and temporal densities $rho_{d_{i,j}}, rho_{t_{i,j}}$. Since the clusters do not overlap in time, their properties can be regarded as time-dependent functions, which can be investigated by visual or analytical methods. For example, the functions can be plotted in a time graph. The clusters can also be represented as collections of points in a map or a space-time cube. The pseudo code of the proposed method is given in Algorithm 1. The computional complexity of our method results from the first spatial clustering and the iterative temporal clustering. The spatial clustering needs to find at most for each of the n entries its ε^s-neighbourhood which can be done in $O(n \cdot \log n)$ with the kd-Tree as index structure. This results in k clusters with $n_i (i = 1, \cdots, k)$ entries that must by further clustered. k times clustering of n_i entries has complexity of $\sum_{i=1}^{k} n_i \cdot \log n_i$. Since $n_i \leq n \forall i$ and $\log n_i \leq \log n \forall i$ and $\sum_{i=1}^{k} n_i \leq n$ due to noise we have $\sum_{i=1}^{k} n_i \cdot \log n_i \leq \sum_{i=1}^{k} \log n \cdot \sum_{i=1}^{k} n_i \leq k \cdot \log n \cdot n$. This means the whole method has a complexity of $O(n \cdot \log n \cdot k)$ which is $O(n \cdot \log n)$ depending on n.

4. Types of spatio-temporal clusters

Using the proposed method, we can detect clusters in space and time with different shapes. We outline a taxonomy of the possible shapes of spatio-temporal clusters. The types

of clusters are exemplified by referring to the Flickr photos data.

- Stationary - clusters limited in space and extended in time over the whole time span of the data or from some time moment to the end of the period under study. For example, in the Flickr photos data, stationary clusters exist in the places of major tourist attractions, such as monuments.
- Reappearing - several temporal clusters occurring in the same place and separated by time intervals in which no or very few events occurred in this place. Two types of reappearing clusters can exist:
  - Regular (periodic) - temporal clusters separated by time intervals of approximately equal lengths. For example, periodic clusters of Flickr photos exist in places where certain happenings occur regularly, such as Oktoberfest in Munich.
  - Irregular - temporal clusters separated by time intervals of unequal lengths. For example, irregular clusters of Flickr photos may correspond to concerts or football games.
- Occasional - clusters limited in time such that there are no other temporal clusters that occurred in the same place. An example may be a cluster of Flickr photos portraying a traffic accident.
- Tracks - temporally dense clusters in which later events are shifted in space with respect to earlier events. An example is a sequence of photos taken by a Flickr user during a walk along the way.

Now we shall demonstrate how our method detects different types of spatio-temporal clusters in real data.

5. Application to real data and discussion

We tested our method on a subset of the all photos uploaded to Flickr [†] covering the area of Berlin and the period from March to September 2009. The subset contains about 45.000 photos. The spatial clustering in the first step of our method is done with the spatial distance threshold of 200 meters. For the temporal clustering in the second step, we defined that the temporal threshold of 60 minutes must be used for an area with the spatial extent 6000 meters containing 500 events. A cluster must consist of at least 10 events.

Out of the results, we have chosen two examples that demonstrate the behavior of our method. figure 2 presents several temporal clusters detected within one spatial cluster. The temporal threshold chosen according to the characteristics of the parent spatial cluster is 1.2 minutes. On the left, the clusters are shown on a map. The area is marked with a red circle. In the middle, the spatio-temporal clusters are represented in a space-time cube that visualizes the events as points in the tree-dimensional space (latitude, longitude
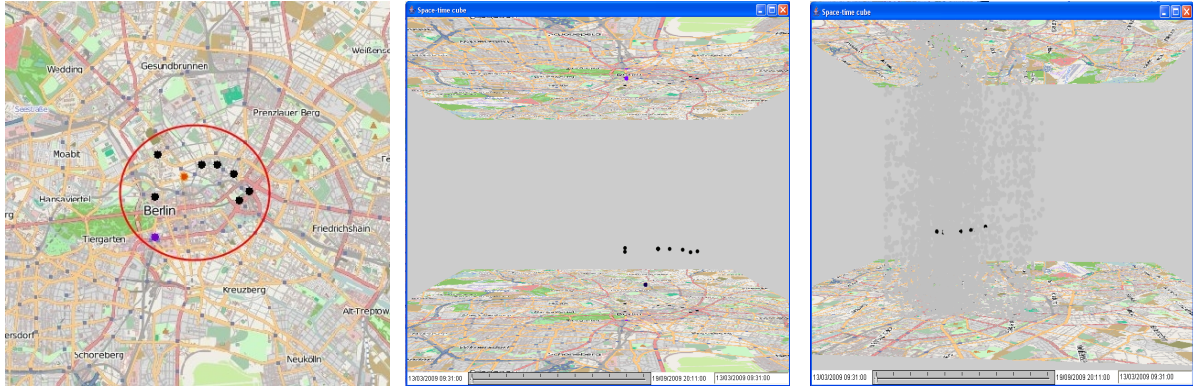

[†] www.flickr.com

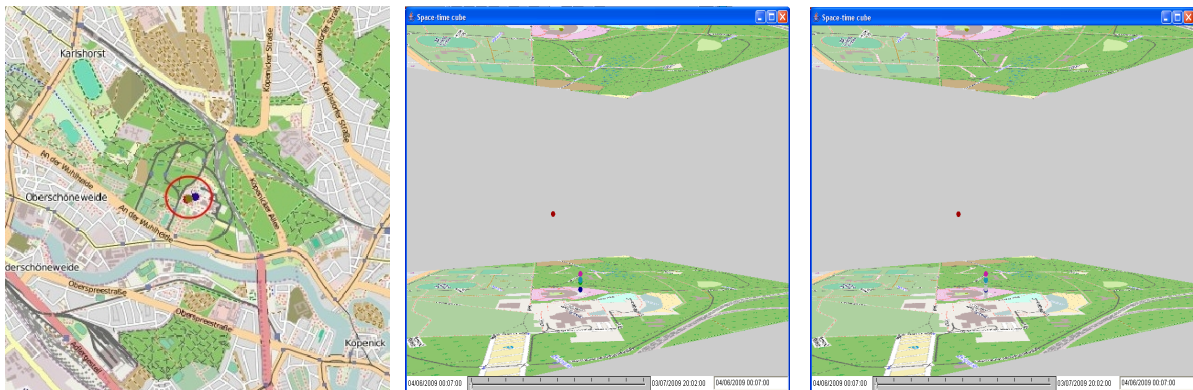**Figure 2:** *Temporal clusters in a spatially extended area*

Figure 3: *Temporal clusters in a spatially compact area*

and time). The black points correspond to a cluster extended in space but dense in time. By further investigation we found that the events of this cluster belong to a city tour of one person. Two other temporal clusters refer to a conference called "re-publica" (purple) [†] and to a labor union event (orange). These clusters are very compact both in space and in time.

On the right of figure 2, all events of the parent spatial cluster of these temporal clusters are shown in the space-time cube. The points representing most of the events are in gray, which means that the temporal clustering treated them as "noise". This is because the very small temporal distance threshold did not allow the algorithm to unite these events in clusters. Hence, in a broad spatial cluster we could find temporally compact subclusters referring to meaningful happenings.

Another example is shown in figure 3. Here six temporal clusters are found in a spatially small area using the temporal threshold of 5 minutes. The clusters refer to three different happenings that took place in this area. Four

clusters are associated with the Juggling convention [‡] that takes place over four days from the 4th to the 7th of June. The first cluster contains entries related to the opening event of the convention, the second to the event itself, the third to the gala and the last to games at the convention at the last day. The other two clusters refer to two concerts at the "Kindl-Bühne Wuhlheide" [§], which is a big stadium in Berlin. The first concert was on the 12th of June, the second on the 3rd of July.

Investigations on the noise of the temporal clustering (presented on the right figure 3) show that several entries during the time of the Juggling convention were classified as noise. There are four time intervals containing only noise but referring to the Juggling convention. This means that the temporal threshold of 5 minutes was too small and the number of temporally close events was not high enough to build additional clusters. For instance between 0:07 and 0:08 on the 4th of June five temporally close events exist but the next event happens only at 10:17.

[†] http://www.re-publica.de/09/

[‡] http://www.circulum.de/convention/joomla15a/

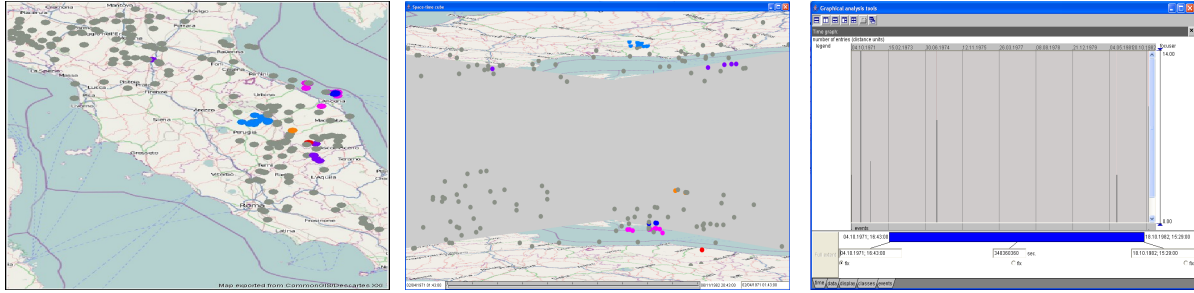[§] http://www.wuhlheide.de/en_index.html

Figure 4: *Earthquakes in central Italy. Left: the detected clusters of earthquakes on the map. Center: the clusters in a space-time cube. Right: the variation of the number of events over time.*
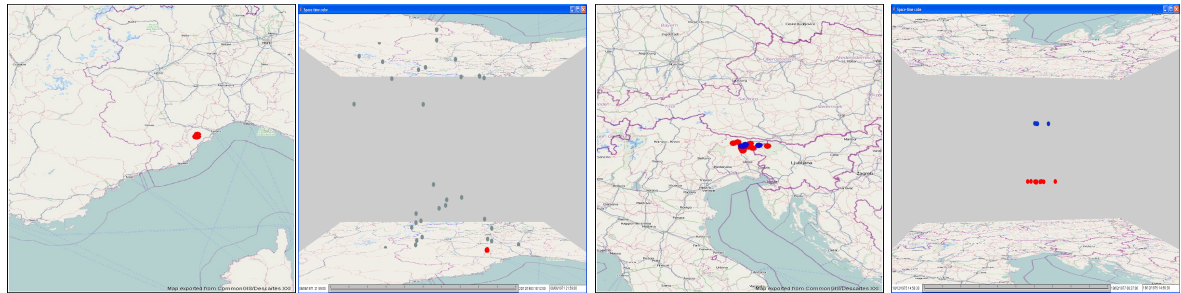

Figure 5: *Earthquakes in Italy and France are shown on the left side and in North Italy, Austria and Switzerland on the right side.*

cluster i	begin	end	n_i	d_i	x	y
noise	2009-03-13 09:31	2009-09-19 20:11	20,163	<null>	13.3906	52.51995
re-publica	2009-04-01 18:12	2009-04-01 18:12	30	0	13.38847	52.52413
re-publica	2009-04-01 18:13	2009-04-01 18:13	38	0	13.38847	52.52413
re-publica	2009-04-01 18:14	2009-04-01 18:14	32	0	13.38847	52.52413
city tour	2009-05-01 00:00	2009-05-01 00:00	40	2,768	13.39599	52.52648
labor union	2009-09-12 15:37	2009-09-12 15:37	38	0	13.37714	52.50969
noise	2009-06-04 00:07	2009-06-12 20:33	80	<null>	13.54658	52.4623
juggling	2009-06-04 19:40	2009-06-04 21:49	86	0	13.54665	52.46231
juggling	2009-06-05 15:22	2009-06-05 15:39	46	0	13.54665	52.46231
juggling	2009-06-05 20:17	2009-06-05 23:22	182	0	13.54665	52.46231
juggling	2009-06-06 13:02	2009-06-06 13:50	52	0	13.54665	52.46231
concert a	2009-06-12 21:25	2009-06-12 23:01	69	0	13.54524	52.46199
concert b	2009-07-03 20:00	2009-07-03 20:02	174	0	13.54543	52.46217

Table 1: *Temporal and numeric characteristics of the clusters of Flickr photos found by our proposed method*

The exact times and additional characteristics of the clusters presented in figures 2 and 3 can be found in table 1. The first part of the table refers to the clusters in the inner city, the second part to the concerts and the Juggling convention. These examples demonstrate that within one run of our method we could find different types of clusters having different semantics. We found an occasional cluster (labor union), several re-appearing clusters (conferences, juggling convention, concerts) and a track (city tour).

We compared the results of our approach with the traditional DBSCAN that used two distance thresholds. For the minimal spatial distance threshold we have chosen 200 meter and for the temporal distance threshold 1 hour - similar to

the parameter values used on the previous experiment. The differences are demonstrated on the spatial temporal clusters representing the "city tour" and the "juggling event". On the table on figure 6 on the last page statistics of the found clusters are shown just like the results before. Further the pictures of the investigated clusters found by DBSCAN are shown on figure 6 as well (in the middle the "'city tour"' and on the right the "juggling event" can be seen).

We clearly see that DBSCAN also finds the "juggling event" but with different temporal clusters. The "city tour" on the other hand is not extracted as accurately as in our proposed method. Further the found clusters have all very low variances in space and time compared with the clusters found by the stepwise approach. This shows that our method can find more arbitrarily shaped clusters in space and time.

Additionally, we tested our method on a data set containing information about the earthquakes in Europe in the last century. We focused on a period from the beginning of 1971 to 1999. During this time, there were about 13.000 earthquakes noted in the data set. In this study, we take into account only the geographical and temporal components of the earthquake events. In the first step, the method looks for regions in Europe having many earthquakes during the mentioned time period. Depending on the sizes of the detected regions and the density of the earthquakes in them, the method further searches for temporal intervals with high

density of earthquake events.

The spatial clustering was performed with the spatial distance threshold of 10.000 meters. For the further temporal clustering, we defined that in a region with the spatial extent of 60.000 meter containing 500 earthquakes an event must occur in at most 100 days after the previous event in order to be included in the same cluster as the previous event. Each cluster must contain at least 5 earthquakes. According to the above-given taxonomy of spatio-temporal clusters, in this scenario we are mainly interested in reappearing and occasional clusters.

Out of the results of our clustering method, we picked three examples that shall show the capability to find desired clusters in space and time. figure 4 presents one big spatial cluster in central Italy containing seven temporal clusters. The points belonging to the temporal clusters are colored in different hues and the events that were classified as noise are shown in gray. On the left we see the area of the earthquakes on a map, in the middle the events are visualized in a space-time cube, and on the right the cardinalities of the temporal clusters are represented on a time graph as a time-dependent function. These three visualization techniques are complementary. While the space-time cube is good for getting an overview of the spatio-temporal distribution of events, the map conveys more precisely the spatial positions of the events, and the time graph conveys shows conveniently the temporal positions and numeric characteristics of the clusters.

On the map we see that the temporal clusters are also compact in space except for the violet-colored cluster 7. This cluster contains one outlier located far north west from the other events. Nevertheless, although this earthquake happened 270 km away, this was only 2 hours before the next event in the cluster. Hence, this spatially isolated earthquake may be related to the other earthquakes in the cluster.

Concerning the temporal extent, the earthquakes in all clusters occurred in intervals from one to three days. The exact temporal and numeric characteristics of the clusters can be seen in table 2. Here the separation of the clusters from the noise can be easily seen. The table is ordered by the time of the earthquakes. After each cluster, the earthquakes classified as noise appeared after a longer time than the earthquakes in the cluster. For instance, cluster 9 contains earthquakes from 04.02.1972 02:42 to 07.02.1972 10:48, each of them happening from 1 hour to 16 hours after the previous one. After the last earthquake of the cluster, the next event happened six days later and the following one two additional days later. Both earthquakes were classified as noise.

Two further results of our method are shown in the left and right parts of figure 5, respectively. Each result is visualized on a map and in a space-time cube. On the left there is an area near the border between Italy and France. The area on the right includes the north of Italy and parts of Austria and Switzerland. However, all clusters of earthquakes found lie in Italy.

The cluster shown on the left contains six earthquakes that occurred from 17.01.1972 12:08 to 28.01.1972 15:29. This cluster is temporally well separated from the noise before and after it (shown in table 2). The map and space-time cube show that this cluster is very dense and small in space and time with respect to the surrounding noise.

On the right, two temporal clusters that occurred from 06.05.1976 20 to 07.05.1976 07:18 and from 15.09.1976 03:15 to 15.09.1976 19:18 can be seen. Here the noise is not visualized for a better visibility on the clusters. From the map it is evident that the earthquakes happened densely in space while the spatial extent of the area is larger than in the previously considered example. According to our typology, on the left of figure 5 we have an occasional cluster and on the right a re-appearing cluster.

The examples presented in this section demonstrate that our method is indeed able to detect spatio-temporal clusters of various shapes.

6. Conclusion and Future Work

Our experiments show that the proposed method is able to find interesting places and times using event data. We showed that our assumption of a relation between the spatial and temporal properties of the clusters enables our method to find differently shaped clusters of events that are likely to be semantically related. In the future we plan to investigate several possibilities to visualize the discovered spatio-temporal clusters in a more comprehensible way. The goal is to integrate the two-step clustering method into a Visual Analytic environment where the user can find, examine, and interpret various types of spatio-temporal clusters.

Another important aspect that must be considered is scalability. The amounts of event data are often so large that they do not fit into main memory. In addition to the use of index structures, possibilities for incremental clustering must be investigated.

Further research will be done towards automated classification of discovered spatio-temporal clusters according to the definitions given in section 4. We shall also investigate the possibilities for combining the method with techniques for time series analysis.

7. Acknowledgments

Spatial Cluster	Temporal Cluster i	Begin	End	n_i	d_i	Mean X	Mean Y
7	noise	06.06.1971 21:59	09.07.1971 21:07	2		7,4842857142	44,474
7	6	17.01.1972 12:08	28.01.1972 15:29	6	3697,09273941669	8,20666666	44,205
7	noise	23.03.1972 18:23	22.12.1983 18:12	33		7,4842857142	44,474
3	noise	03.05.1971 21:31	29.04.1976 20:31	140		13,384293785	46,946073446
3	4	06.05.1976 20	07.05.1976 07:18	13	64677,4667099047	13,180769230	46,254615384
3	noise	09.05.1976 00:53	14.09.1976 02:30	21		13,384293785	46,946073446
3	5	15.09.1976 03:15	15.09.1976 19:18	5	30917,9161385909	13,22	46,28
3	noise	17.09.1976 17:07	26.11.1983 21:34	175		13,384293785	46,946073446
8	noise	02.04.1971 01:43	12.09.1971 00:30	28		11,90265625	43,558645833
8	10	04.10.1971 16:43	05.10.1971 00:07	5	0	13,2	42,8
8	noise	05.12.1971 07:12	25.01.1972 21:17	3		11,90265625	43,558645833
8	9	04.02.1972 02:42	07.02.1972 10:48	14	53133,7911486772	13,435	43,632142857
8	noise	13.02.1972 08:54	14.06.1972 19:21	13		11,90265625	43,558645833
8	11	16.06.1972 18:09	17.06.1972 10:03	6	0	13,42	43,63
8	12	18.06.1972 07:44	19.06.1972 06:48	5	10984,324745169	13,494	43,616
8	noise	25.06.1972 17:10	01.12.1974 01:21	40		11,90265625	43,558645833
8	13	02.12.1974 01:14	04.12.1974 17:59	9	21608,5916155424	12,9111111	42,84
8	noise	05.12.1974 05:22	09.08.1981 16:04	34		11,90265625	43,558645833
8	7	10.08.1981 23:02	12.08.1981 02:49	5	253279,562163937	12,882	42,914
8	noise	20.09.1981 20:40	16.10.1982 01:44	42		11,90265625	43,558645833
8	8	17.10.1982 04:43	18.10.1982 15:29	10	32223,4031652449	12,55	43,158
8	noise	19.10.1982 01:22	06.11.1982 20:43	4		11,90265625	43,558645833

Table 2: *Characteristics of the clusters of earthquakes*

| cluster i	begin	end	n_i	d_i	x	y
city tour	2009-05-01 00:00	2009-05-01 00:00	34	0	13.39559	52.527
juggling	2009-06-04 12:27	2009-06-04 13:52	10	0	13.54665	52.46231
juggling	2009-06-04 19:11	2009-06-04 23:04	96	0	13.54665	52.46231
juggling	2009-06-05 11:37	2009-06-05 11:49	15	0	13.54665	52.46231
juggling	2009-06-05 15:22	2009-06-05 15:39	46	0	13.54665	52.46231
juggling	2009-06-05 16:47	2009-06-05 16:58	12	0	13.54665	52.46231
juggling	2009-06-05 20:17	2009-06-05 23:22	182	0	13.54665	52.46231
juggling	2009-06-06 13:02	2009-06-06 13:50	52	0	13.54665	52.46231
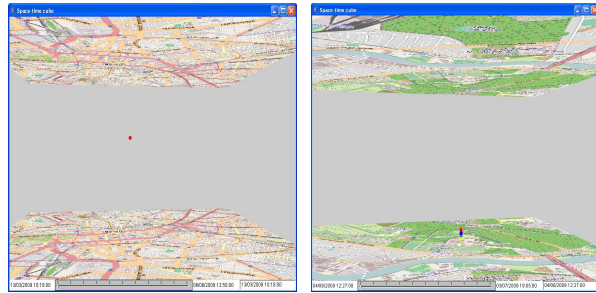
Figure 6: *Spatial temporal clusters found by DBSCAN. On the left temporal and numeric characteristics and on the right spatial and temporal extent of the clusters are shown.*

References

[AA09] ANDRIENKO G., ANDRIENKO N.: Interactive spatio-temporal cluster analysis of vast challenge 2008 datasets. In *VAKD '09: Proceedings of the ACM SIGKDD Workshop on Visual Analytics and Knowledge Discovery* (New York, NY, USA, 2009), ACM, pp. 5–11. 2

[Ben75] BENTLEY J. L.: Multidimensional binary search trees used for associative searching. *Commun. ACM 18*, 9 (September 1975), 509–517. 2

[BK07] BIRANT D., KUT A.: St-dbscan: An algorithm for clustering spatial-temporal data. *Data Knowl. Eng. 60*, 1 (2007), 208–221. 2

[DE84] DAY W. H., EDELSBRUNNER H.: Efficient algorithms for agglomerative hierarchical clustering methods. *Journal of Classification 1*, 1 (December 1984), 7–24. 2

[EKJX96] ESTER M., KRIEGEL H.-P., JÖRG S., XU X.: A density-based algorithm for discovering clusters in large spatial databases with noise, 1996. 2

[GAA04] GATALSKY P., ANDRIENKO N., ANDRIENKO G.: Interactive analysis of event data using space-time cube. In *Information Visualisation, 2004. IV 2004. Proceedings. Eighth International Conference on* (July 2004), pp. 145–152. 1

[Gut84] GUTTMAN A.: R-trees: A dynamic index structure for spatial searching, 1984. 2

[Iye04] IYENGAR V. S.: On detecting space-time clusters. In *KDD '04: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (New York, NY, USA, 2004), ACM, pp. 587–592. 2

[KF96] KIJKO A., FUNK C. W.: Space-time interaction amongst clusters of mining induced seismicity. *Pure and Applied Geophysics 147*, 2 (1996), 277–288. 2

[NM04] NEILL D. B., MOORE A. W.: *Detecting space-time clusters: prior work and new directions*. Tech. rep., Carnegie Mellon University, 2004. 2

[Tob70] TOBLER W. R.: A computer movie simulating urban growth in the detroit region. *Economic Geography 46* (1970), 234–240. 1