

# Visual Analytics for Building Management

Petr Glos and Lubomír Popelínský



Fig. 1. 3D model of the new Masaryk University Campus Bohunice

**Abstract**—We first shortly describe Masaryk University building management system and visual analytic tools that are used in building management at Masaryk university. Then we introduce a novel method that combine visualization with automatic classification for finding unexpected trends in temperature.

**Index Terms**— Building management, visualization, data mining, classification, decision tree, outliers

---

## 1 INTRODUCTION

Masaryk University maintains a digital version of building passport that currently consists of approximately 200 buildings and 17,000 rooms. For processing this digital data and its visualization, the ESRI ArcGIS software has been successfully applied. However, for deep understanding of running processes in a room, namely for anomaly detection, we need an analytic tool that can predict (or at least detect) such rare events. A rare event is a pattern that does not occur very often but is important for precaution, e.g. a fire (or an increase of a temperature), fast repeated switching on/off of a device, or a water pipe disruption.

In [8] we presented an ongoing project on mining spatio-temporal frequent patterns from building management data. We focused on two tasks that are important for facility management - mining frequent patterns [7] and mining rare events in spatio-temporal data [3].

In this paper we focus on another part of the analytic tool: the anomaly detection in temperature trends. We first describe shortly the building management system and then we summarize visualization tools that are currently used. In the second part of this text we introduce an analytic method that serve for detection of

anomalies – unexpected behavior in temperature – in a room. Since the number of temperature measurements varies (from less than 90 to more than 80,000) depending on a way of measurement and the actual temperature changes in the particular place, various aggregated values are computed first. This method combines automatic classification and outlier detection.

## 2 GEODATABASE OF BUILDINGS AND TECHNOLOGIES

Several distinct constructions (building primitives – engineering constructions, holes in constructions and panes of wholes) were defined for building and room representation in a geodatabase and every building is made up of these objects. The building primitives have polygon geometry representations with 2D coordinates and two attributes of height (upper and bottom part) for 3D representation. Each room, floor and building has unique identification called position code. For example, BMA01 is the position code of Masaryk university headquarters building, BMA01N02 is the position code of the second floor of this building and BMA01N02040 is the position code of one room on this floor (see Fig 2).

The resulting passport data is available to university employees, students and even for public via the internet/intranet as well as it is used by other university's information systems. Additionally, the building passport is used to generate 2D maps and 3D models of the buildings (see Fig 3).

- 
- Petr Glos with Masaryk University, Institute of Computer Science, E-Mail: glos@ics.muni.cz.
  - Luboš Popelínský, with Masaryk University, Faculty of Informatics, E-Mail: popel@fi.muni.cz.

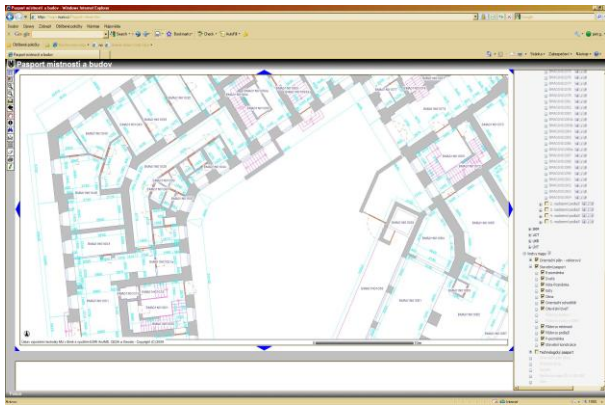


Fig. 2. Internet floor map with positions codes. High resolution image can be found at [http://gis.ics.muni.cz/bms/demo/SP\\_RM.U.PNG](http://gis.ics.muni.cz/bms/demo/SP_RM.U.PNG).

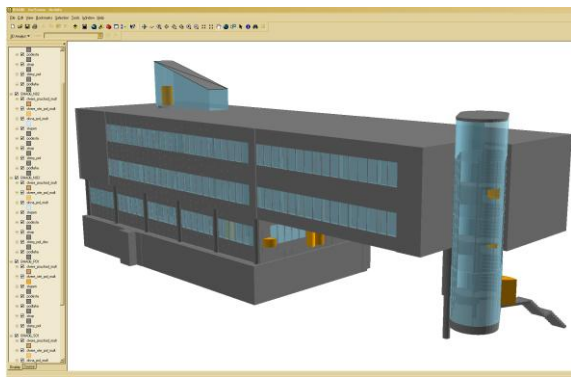


Fig. 3. 3D model of a new campus building. High resolution image can be found at [http://gis.ics.muni.cz/bms/demo/SP\\_A7.PNG](http://gis.ics.muni.cz/bms/demo/SP_A7.PNG).

### 3 BUILDING MANAGEMENT SYSTEM

Masaryk University implements a Building Management System (BMS) based on BACNet open protocol. BACnet is a data communication protocol for building automation and control network (<http://www.bacnet.org>). The BMS applications are used for monitoring and controlling technologies of the new University Campus buildings such as HVAC (heating, ventilation, air conditioning and cooling), lighting, fire alarm system, intrusion system, access system, CCTV, etc. The data from HVAC are very interesting because room environment is very important for building users and also operation costs depends on HVAC behaviour. BMS provides means of storing a historical building operation data in the relational database.

A typical record of the database contains an identification of the measurement, time, and a value. The identification of measurement consists of the position code (unique code of room, floor and building) and type of value such as temperature or humidity. Tuple [BMA01N02040, TK14, 4.9.2007 15:00:00, 20.511] represents one measurement of room temperature (TK14) in room with position code BMA01N02040. The room temperature was 20.511 °C on September 4th, 2007 at 15:00.

In this way, we can obtain a spatiotemporal data of the room environment and HVAC operation and we are able to use the data for analyses and visualizations.

### 4 GIS VISUALIZATIONS

The objects of BMS database and the geodatabase of buildings are linked through the position code and we can apply GIS tools for analysing the BMS spatiotemporal data and visualization. We use

ESRI ArcGIS for creating 2D and 3D thematic maps to show either the actual operation values or the aggregating values over a given time interval. We utilize 2D and 3D temporal animations for visualization of operation values for given time period. Especially temporal animations are very interesting because they make it possible to see the history of long-term values quickly and detect the trends and anomalies of a building operation.

#### 4.1 2D Thematic Maps

They are two typical scenarios for preparation of 2D thematic maps. In the first scenario we create floor maps or maps of buildings for a given value (e.g. temperature, humidity, differential pressure, electricity consumption) in a given instant of time. For this purpose we prepare data view of a required feature class (e.g. geometries for a ground plan of rooms) and required values from the BMS database (e.g. temperatures of rooms for given instant of time). The records from these views are joined through the position code.

In the second scenario we create floor maps or maps of buildings for aggregated values (e.g. minimum or maximum of temperatures, electricity consumption, etc.) for a given time interval. For this purpose we prepare a query that gathers the required data and then we join the query results (e.g. electricity consumption of a given day for all campus buildings) and the required feature class from the geodatabase again using the position code.

After that, we prepare the desired symbology and legend for the requested map and do the GIS analyses on the BMS data. The resulting map can be exported in the desired data format (e.g. PDF, JPG or PNG). The map (Fig. 4) represents temperatures measured by ceiling sensors and sensors positioned on top of computer racks in a computer room. Blue colour represents the lowest temperature and red colour is used for the highest temperature.

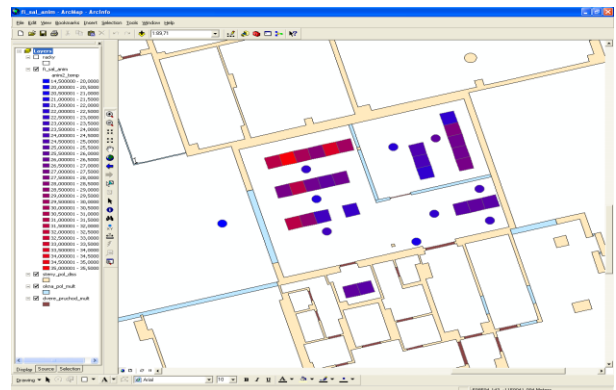


Fig.4. Computer room temperatures. High res. image can be found at [http://gis.ics.muni.cz/bms/demo/VIZ\\_2D\\_UVT\\_sal\\_teploty.PNG](http://gis.ics.muni.cz/bms/demo/VIZ_2D_UVT_sal_teploty.PNG).

#### 4.2 3D Thematic Maps

Similarly to the 2D, there are also two typical scenarios for preparation of 3D thematic maps. The difference is in the method of visualization. We need to prepare 3D objects that represents the buildings and rooms. For this purpose we can use the height attributes (upper and bottom heights of the building primitives) for the extrusion method of ArcScene software. The extrusion method allows to create a 3D model of rooms and buildings from building primitives.

Using the retrieved 3D representation we can create the required 3D map for given values as in the case of 2D maps and we can also export the result – there are the data formats available: VRML, PDF, JPG or PNG). The map (Fig. 5) illustrates the values of selected room temperatures in a given building. Dark blue colour is used for the lowest temperature while light ping colour represents the highest temperature.

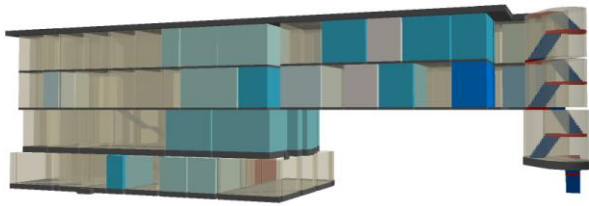


Fig. 5. 3D model of a new campus building. High resolution image can be found at [http://gis.ics.muni.cz/bms/demo/VIZ\\_3D\\_A5\\_teploty.jpg](http://gis.ics.muni.cz/bms/demo/VIZ_3D_A5_teploty.jpg).

### 4.3 2D Temporal Animations

Again, there are two typical scenarios. We can create animations of 2D maps using the BMS data for either values measured in a given time frame (e.g. history of room temperatures of one floor of given building) or aggregated values in a given time interval (e.g. history of daily electricity consumption for all building of campus). For this purposes we prepare a data view of the requested feature class and the required BMS data for the given time interval (e.g. room temperatures of required floor of building for one week or month) and again join them using the position code.

We can use the same symbology as for the thematic maps and export the resulting animation to a video format AVI. The animation ([http://gis.ics.muni.cz/bms/demo/fi\\_sal\\_video.avi](http://gis.ics.muni.cz/bms/demo/fi_sal_video.avi) - 28 MB) represents a two-day history of temperatures in a computer room. You can see that the temperatures are relatively steady, but the temperatures in north-western racks were increased between 20:00 and 24:00 on August 23rd.

### 4.4 3D Temporal Animations

We use 3D representations of rooms and buildings and prepare temporal animations with same scenarios as in the case of 2D temporal animations. The 3D animation give us better imagination of the chosen characteristics of building operation because we can see and analyze all floors of the building together. It is also simple to rotate and pan the model and switch the visibility of floors on or off to get a complex view of the monitored situation.

The animation ([http://gis.ics.muni.cz/bms/demo/UKB\\_spotreby.avi](http://gis.ics.muni.cz/bms/demo/UKB_spotreby.avi) - 4MB) illustrates a three-month history of daily electricity consumptions of selected buildings of the campus. The dark brown colour represents the highest consumption and the light brown colour is used for the lowest values of consumption. We can produce the video from these animations in the AVI format too. While the geographic visualizations are applicable to obtain overview information about buildings, for deep analyses we use data mining techniques.

## 5 MINING AND VISUALIZING EXCEPTIONS

### 5.1 Exceptions

For efficient building management it is very important to find such parts of a building that strongly differs – e.g. in terms of consumption of resources, like heating – from other parts that are supposed to be similar. Such events are actually very rare. The main idea is following. To obtain information about rare events that may happen in building management we first have to define similarity on

parts of building (rooms). This similarity measure generates classes of similar objects. Then, using learning techniques, we learn a classifier that with high accuracy classifies all rooms to those classes. We use the C4.5 algorithm (J48 implementation). Rooms that are not classified correctly are candidate places where a rare event has occurred. Those rooms are visualized together with an evidence for a rare event – a logical formula (e.g. a branch of a decision tree) for which that room is as an exception. It helps not only to find the place but also to explain the reason for a possible anomaly.

### 5.2 Data and Data Aggregation

Attributes of the data about rooms extracted from the building management system can be divided into four groups. The first group concerns room identification like the room position code and room description (text). The second group describes a position of a room in a building – its orientation (east, west), the area, etc. The third group provides information about the way how a temperature is measured. Since the number of temperature measurements can vary from tens to tens of thousands, the fourth group contains computed attributes that are actually aggregates – a minimum and maximum temperature, a difference between the max. and min. temperature and also a number of temperature measurements.

### 5.3 Algorithm

Below is an algorithm for finding exceptions in building management data. The algorithms are influenced by the following parameters, *MinAcc* - minimum accuracy of the decision tree on the learning data, *MaxNoice* - ratio of incorrectly classified examples =  $\frac{\text{incorrectly classified}}{\text{majority class in the leaf}}$

- 
0. Define similarity classes
  1. Learn classifiers
    - For each class,
    - find a model (decision tree) that reaches accuracy  $\text{MinAcc} : 100\% > \text{accuracy} \geq \text{MinAcc}$
  3. Find exceptions:
    - 3a) find all branches in all trees such that  $\text{MaxNoice} \geq \text{noice} > 0$
    - 3b) sort them by a number of appearance (and than by noice) in ascending order
    - 3c) take first *MaxExceptions* examples and display them together with the corresponding branch and layers
- 

### 5.4 User Settings

It is clear that the process of finding anomalies can be hardly automatic. A user has to define at least a similarity or neighborhood measure [1] that serves for dividing data into similarity classes. It is necessary that the class is nominal. In the case that the similarity measure is not discrete, the class is automatically discretized into 2 to 5 classes, either equidistant or with equal frequency of examples, and the discretization that maximize the accuracy is chosen. A user can also set the parameters mentioned above, especially a level for outlier detection, which is highly data-dependent. Additionally, setting *MaxExceptions*, the maximal number of exceptions to display, prevents the algorithm from finding and displaying too many candidates for outliers. The generation of classifiers is fully automatic. The result is then displayed with all layers where an outlier has appeared. Places that have been detected most frequently are displayed first.



## 6 RESULTS

The algorithm has been tested on the data collected at Masaryk University Campus Bohunice in 2009 for all 267 rooms. Information about temperature has been first aggregated as described above. The data contained the following attributes: RoomID - room identifier

Corner - yes/no  
 Orientation - east/west/other  
 Floor position - inner/outer  
 Description - text  
 Floor - -1,1,2,3  
 Area - in square meters  
 Vzorkovani\_typ - COV for sampling data by change of value (temperature)  
 - POL for sampling data by change of time  
 Vzorkovani\_hodnota - in °C for COV, in minutes for POL  
 NumTempMeas - number of temperature measurement  
 Temperature\_min - minimal temperature  
 Temperature\_max - maximal temperature  
 Difference\_min\_max - difference between min. and max. temperature

We defined four similarity measures - floor, floor position (inner/outer), number of temperature measurements for POL sampling (discretized into 3 bins equidistant, NumTemp-POL), number of temperature measurements for COV sampling (discretized into 5 bins with equal frequency in a bin, NumTemp-COV). No limit has been set to *MaxExceptions*.

Class	MinAcc	MaxNoise	num. of outliers
floor	0.95	0.20	6
floor position	0.90	0.40	9
NumTemp-POL	0.75	0.20	5
NumTemp-COV	0.75	0.20	9

The most interesting patterns are below.

### floor position

Difference\_min\_max <= 9.0082 AND  
 Difference\_min\_max > 7.0286 AND  
 not to west =>  
 outer floor (2 outliers)

Difference\_min\_max > 9.0082: =>  
 outer floor (7 outliers)

### POL data, class: number of measurement

Difference\_min\_max <= 17.606 =>  
 num of measurements < 45681.333333 (3 outliers)

### COV data, class: number of measurement

outer AND  
 area > 10.43 AND  
 Difference\_min\_max <= 9.1335  
 => -1st floor (1 outlier)

outer AND  
 area > 10.43 AND  
 Difference\_min\_max > 9.1335  
 => 3<sup>rd</sup> floor (1 outlier)

floor = -1 AND not to east  
 => num of measurements < 1614 (1 outlier)

## 7 CONCLUSION

We described visual analytics tools that are used in the building management system of Masaryk University Brno. We introduced a novel algorithm for finding anomalies in temperature measurements. The detected anomaly may be caused by a defect of a measurement tool or the network that transmits data. Actually, most of the anomalies discovered in our experiments were of that kind. Nevertheless, even in that case the information about anomaly is useful. Moreover, the discovered rules have been evaluated as reasonable by an expert.

## ACKNOWLEDGMENTS

We thank Michal Batko for his collaboration. This work has been partially supported by Faculty of Informatics, Masaryk University and also by the Grant Agency of the Czech Republic under the Grant No. MSM 0021622418 Dynamic Geovisualization in Crisis Management and by Faculty of Informatics, Masaryk University Brno.

## REFERENCES

- [1] Ester M. et al. Spatial Data Mining: A Database Approach. *Advances in spatial databases: 5th international symposium, SSD '97*, Berlin, 1997.
- [2] Hodge V., Austin J.: A survey of outlier detection methodologies. *Artificial Intelligence Review Volume 22, Number 2, October 2004*, Springer 2004
- [3] Huang Y., Pei J., Xiong H. Mining C-Location Patterns with Rare Events from Spatial Data Sets. *Geoinformatika 10, 2006, pp. 239-260*.
- [4] Glos, P. Building and Technology Passport of Masaryk University. In *Proceedings of 2008 ESRI International User Conference*. San Diego, California : ESRI Press, 2008
- [5] Glos, P. Using ArcGIS for Visualizing Historical Data from BMS. In *Proceedings of 2009 ESRI International User Conference*. San Diego California : ESRI Press, 2009
- [6] John G.H.: Robust Decision Trees: Removing Outliers from Databases. In *Proceedings of KDD-95 Conference*, AAAI 1995.
- [7] Popelínský L., J. Blaťák: Toward mining of spatiotemporal maximal frequent patterns. In *Proceedings of ECML/PKDD Workshop on Mining Spatio-Temporal Data (MSTD)*, Porto 2005.
- [8] Popelínský L., P. Glos: Facility management and mining spatio-temporal data. *Proceedings of Znalosti 2010 Czech-Slovak AI conference*, Jindřichův Hradec, 2010.
- [9] Simoff S. J., Böhlen, M. H., Mazeika A. (eds.) (2008) *Visual Data Mining*. LNCS 4404 Springer Verlag 2008.