Big Data and Visual Analytics for Building Performance Comparison

Dimosthenis Ioannidis[1,2], Angeliki Fotiadou[1], Stelios Krinidis[1], George Stavropoulos[1], Dimitrios Tzovaras[1], and Spiridon Likothanassis[2]

[1] Information Technologies Institute
Centre for Research and Technology Hellas
Thermi-Thessaloniki, Greece
[2] Pattern Recognition Laboratory
Computer Engineering and Informatics
University of Patras, Rio, Patras, Greece

Abstract. In this paper, big data and visual analytics techniques for comparing building performance under different scenarios and designs are presented. Large data consist of building information, energy consumption, environmental measurements and occupancy information, which are combined and correlated utilizing data analytics techniques, so as to extract useful semantic information about building performance. Also, visual analytics techniques are exploited to visualize them in a compact and comprehensive way taking into account properties of human cognition, perception and sense making. They analyze and visualize the performance of the buildings under comparison in the spatio-temporal domain assisting the analyst to compare them and detect patterns, templates and crucial points that are difficult to be detected otherwise. The performance comparison of different buildings or buildings of different designs or buildings with space usage rearrangement is an important factor in engineering that leads to building renovation and construction with low energy consumption and gas emissions in conjunction with comfort, utility and durability.[1]

Keywords: Big data analytics, Visual analytics, Building Performance.

1 Introduction

Building behavior simulation is a fundamental task for energy management, building's spaces refinement, new building designing, *etc.* Building behavior simulation can lead to design buildings with higher energy efficiency and performance. It is a multi-parametric task, where each parameter affects different building's aspects and in different level. To this end, comparison tools able to correlate, combine and analyze all buildings information in conjunction with advanced visualization mechanisms are necessary to evaluate buildings' design

[1] This work was supported by the Greek funded research project ECOBUILDING (GSRT-national)

R. Chbeir et al. (Eds.): AIAI 2015, IFIP AICT 458, pp. 421–430, 2015.
DOI: 10.1007/978-3-319-23868-5_30

performance and to compare them, so as to select and decide the optimal building's design. Building's behavior is comprised by a large number of heterogeneous information (*e.g.* geometric information, occupancy, energy consumption, *etc.*), which is very difficult to be handled, correlated and visualized. Statistical data analysis is useful if the relationship between variables is well defined. However, if the analyst does not know what to expect from the data then it often becomes necessary to visually explore the data in order to identify an appropriate statistical analysis technique. On the other hand, visual analytics techniques can correlate, analyze, extract semantic information and visualize the building information, assisting the analyst to make a decision about its performance.

There is a large number of visualization techniques in the literature, which can be divided into the following categories [11,15,17]:

a) **1D to 3D graphics:** This category includes x-y-z plots, bar charts, line graphs, scatter plots, *etc.* [11]. They are widely used to visualize an estimation of certainty about a state or hypothesis or the distribution of an attribute or even to visualize a data model [15].

b) **Geometric techniques:** Data with more than three dimensions are very difficult to be comprehensible by humans, so they are projected onto a 2D or 3D space providing an overview of their attributes. Matrix of Scatter Plots, Parallel Coordinates [19], BallotMaps [18], Prosection Views, Radial coordinates, Star coordinates, Kiviat diagrams, *etc.* [17], are some representative examples. Most of them represent data as a single 2D or 3D point cloud or object, and as a consequence, important information is missing.

c) **Pixel oriented techniques:** They visually map each multivariate data item to a pixel or block with visual attributes such as color, size, position [13]. Thus, the data are visualized as a set of pixels scattered around.

d) **Iconographic techniques:** The attributes of the data are mapped as properties of an icon or glyph, which vary depending on the values of attributes (*e.g.* icons in format of faces of the Chernoff Faces, icons as stars of the Star Glyphs and icons in stick shape of the stick figures) [2].

e) **Graphs layout techniques:** Graphs can be visually represented by matrix visualization [8], node-link diagrams [3], hybrid views of node-link diagrams and matrix visualization [9], *etc.*

f) **Clutter reduction methods:** Visual clutter is a commonly-found problem in information visualization [4]. With ever increasing sizes of networks, reducing visual clutter has become even more important for visual analysis of large networks. Edge bundling [10], geometry-based technique [3], skeleton-based technique [5] and Parallel edge splatting [1], are some representative techniques belonging to this category.

g) **Combinations:** Mixed techniques that incorporate the advantages of each above mentioned visualization methodologies, providing visualizations and visual analytics techniques focused on the building performance [6].

The main goal of visual analytics techniques is to assist humans, who posses different visual perceptual abilities and skills, such as memorizing visual information (visual memory) and matching identical forms (visual closure) [14]. Even for one individual, visual perception and visual illusion could be an issue [16].

The main motivation of the proposed work is the limited research in deploying visual analytic technologies in building performance simulation, and comparing analysis tools [7]. Comparing alternative building designs before their implementation is a crucial issue in building construction industry. However, the output of the literature simulation tools comprised of complicated information, which could be very difficult assessed by the analyst. Also, the data mining results (coming from the analytics component) are often complex and unruly, so the visualization component allows for the optimal representation of different data structures assisting the user to efficiently and effectively perform data exploration and focus on the problematic or critical areas of interest.

In this paper, taking into account the limited research in deploying visual analytic technologies in building performance comparison tools [7], new visual analytics schemes are proposed that offer new opportunities on evaluation and comparison of building information assisting to selection of the optimal building alternative before their implementation/construction. Also, the building alternatives comparison is enhanced by the extraction of meaningful semantic information for the building's operation and performance is proposed.

The rest of the paper is organized as follows. Section 2 presents the knowledge mining and visual analytics techniques utilized for the building alternatives comparison, while Section 3 presents some indicative use cases demonstrating the performance and the efficiency of the proposed comparison tool. Finally, conclusions are drawn in Section 4.

2 Data and Visual Analytics

Nowadays technology has enabled and has set as a priority the evaluation of the building performance towards energy savings and occupants' comfort, with the use of the necessary and corresponding software, during the design process. These Building Performance Software can simulate with high accuracy the energy consumption of every designed building, even more so if higher level of details are given for the materials, the construction itself, the occupants, the heating and cooling systems as well as the orientation and the place of location. Whether more detailed or less the description of the construction and therefore more realistic the simulation, the program generates a large amount of data regarding the temperature, humidity, energy consumption, CO_2 *etc.* for the whole measurement period set by the user, which need to be thoroughly analyzed in order to retrieve the correctness or not of the design choices. Handling however and analyzing this big amount of data can be a painful, time consuming process with uncertain results if some parameters are overlooked.

2.1 Knowledge Mining

The big amount of data could provide very useful information about the building's operation and performance, but also their combination, correlation and analysis could provide enhanced semantic information as well.

Table 1. Events produced for a three-month period of time.

Type of event	Number of events/data
Environmental/Energy consumption	52.852
Space occupancy events	47.823
Occupancy trajectories	188.325
Occupancy trajectory points	747.059.925
Total	**747.348.925**

An example of the amount of data acquired and processed by the system during a three-month period of time is presented in Table 1, where a summary of different events is presented.

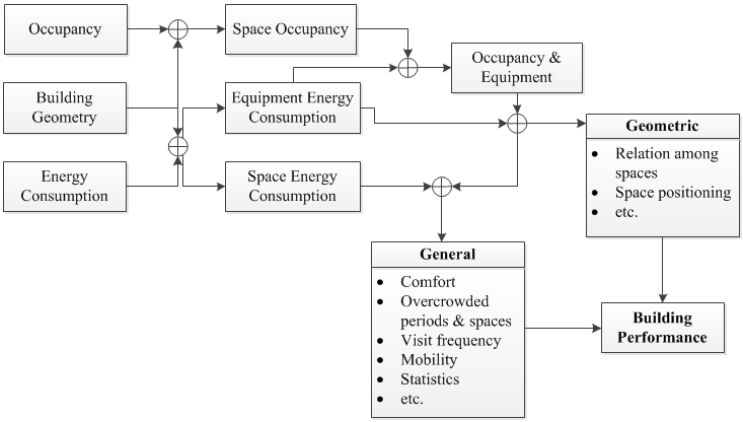

Fig. 1. Data mining procedure for the building performance evaluation

Combining occupancy and building geometry information, one can derive useful information about space occupancy and information for the transitions among spaces, which can lead to meaningful information of the preferences of the occupants, the logical transition sequence among spaces in a building, *etc.* Furthermore, space occupancy and transitions, combined with equipment and space energy consumption, can lead to semantic information about the usage of

each equipment, the needs for extras, or the redeployment specific equipment (*e.g.* the printer or the photocopier is better to moved to a more central space, so as to reduce the employees transitions, *etc.*).

All these information and data could also provide information about the comfort of the spaces, to determine overcrowded periods and spaces and analyze their circumstances (occupancy, time, *etc.*), visit frequency, mobility and transitions among spaces, and a number of other statistics related to the building. Based on these meaningful information, one can extract Key Performance Indicators (KPIs) related to the building's performance, so as to accurately determine its operational performance, even from the design phase. Figure 1 depicts an abstract block diagram of the knowledge mining process.

2.2 Visual Analytics

In this section, an easy and more direct evaluation of the energy performance results is going to be presented. Taking into consideration all the aforementioned information, a user friendly tool has been developed which, through a set of visualization techniques, enables user to extract the useful information from a large data set. This data visualization tool uses a coarse-to-fine approach to visualize information and though its design it directs step by step the user from a more general view of the performance indicators to a more detailed one and in focus of a desired parameter (temperature, energy, *etc.*) selected by the user.

More specifically, the tool starts by exhibiting in a kiviat diagram all the main Key Performance Indicators (KPIs) referring to a specific building setup/ design that give the general description of the building performance. The KPIs are grouped into three categories based on the aspect of building performance that they address: **a)** energy performance, **b)** business performance and **c)** occupancy comfort which can be also colored identified (energy performance- green, business performance -red and occupancy comfort- blue). The green and blue colored KPIs have a positive meaning (the larger the better), while, on the contrary,

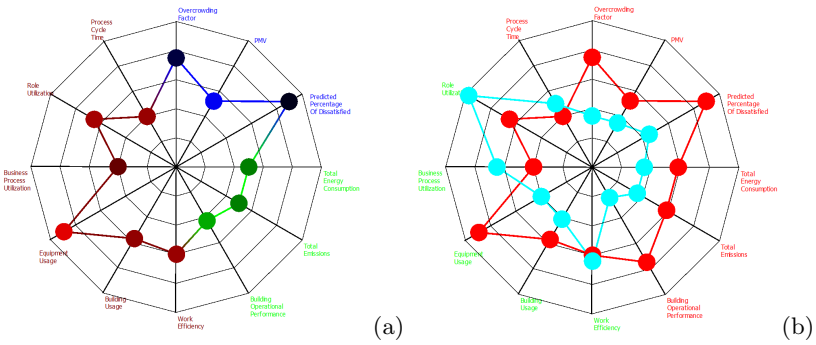

Fig. 2. (a) Kiviat Diagram: Visualization of the building's KPIs, **(b)** Building performance comparison using Kiviat diagrams

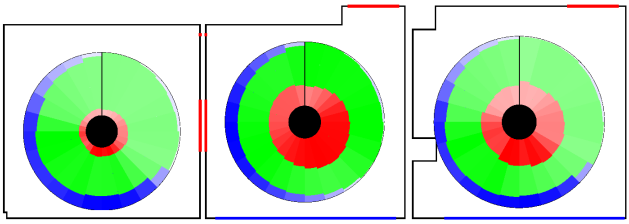

Fig. 3. Energy consumption for a single day. On each space a clock map with the respected energy consumption for the different category devices per day is displayed. The radius of each category denotes the portion of the energy consumption due to the respective category for the specific slice, where the intensity of the color denotes the relation of the consumption of the specific category with respect to the rest of the same category slices.

red colored KPIs have a negative meaning. Furthermore, the density of the color of each KPI indicates whether it has a good or bad value; darker colors of a KPI illustrate better values, while light colors indicates bad values [Figure 2(a)]. From there on, by clicking in each one of the KPIs circles, new diagrams appear with more detailed information regarding the selected performance indicator.

Every indicator, *e.g.* energy consumption, is the sum of spatio-temporal data, stemming from the various energy consuming elements placed into different spaces or even from those elements that are not space-specific, like the centralized air-conditioning device, but whose consumption can also be calculated. With the help of the described tool and by visually analyzing the energy consumption per space, the user can easily identify the most energy consuming elements of a building and focus any consumption optimization efforts on those elements. The spatio-temporal data are graphically represented in every floor plan of building simulated and for every space of the floor, by displaying the cumulative energy consumption of three different energy consuming element categories: **(1)** lightning, **(2)** heating and cooling systems and **(3)** equipment devices (computerss, printers, white appliances *etc.*). Depending on the desired time resolution, the data is displayed per month, week or day. For the latter, an extension of the Clock Map proposed in [12] was utilized, with the addition of a 3^{rd} dimension encoded in the radius using concentric cycles (Figure 3).

The data inside the clock map is organized starting from the top (black line) and moving in a clockwise direction, where each slice represents a time slot that is user configurable. The different elements are represented by different colors: blue color denotes lightning energy consumptions, green denotes HVAC consumption and red other equipment devices such as computers, kitchen equipment *etc.* Again the differentiation in color intensity depicts different values. More specifically the radius of each component in the clock map denotes the ratio of its energy consumption for the specific period of time (slice), where the intensity of the color compares the value of the ratio with the rest of the same component's values over the total visualized period. More detailed analysis over the selected

KPI of a space can be shown in new graphic representations with all the parameters that affect the indicator, which can allow a targeted comparison such as *e.g.* the occupancy versus lighting. This option can prove to be very helpful in terms of design and energy control decisions, as it can stress out important resolutions, for example, if the lights were left open during the night or even when no occupant was present at the room.

3 Experimental Results

Therefore, it can be seen that beyond the graphically representation of all important KPIs, there is another feature that can probably be considered as the most important asset of the tool; the comparison that it allows in every level of its analysis. Indeed through the general representation of the kiviat diagram until the most detailed analysis of the parameters of a certain space, the user has the ability to use this software as helpful tool for the decision that he has to make, either in design stage or in real building energy performance control.

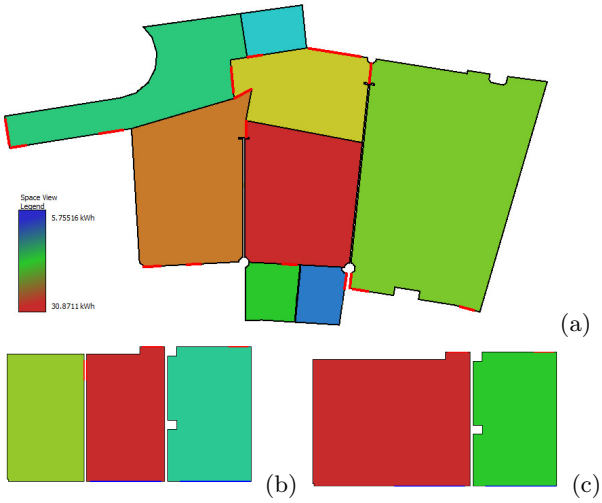

Fig. 4. Spaces energy consumption visual comparison.

Through the general view and the kiviat diagram, the user can compare the performance between two or more buildings or even between different time periods for the same building [Figure 2(b)]. By selecting one of the KPIS and enabling the option "Space View", the user can also have a very fast first impression of the energy consumption of every space and in comparison to the others; the spaces are colored in variations *e.g.* according to their energy consumption, in relation to the color scale described by the legend located at the left side of the floor graphic representations (Figure 4). Finally, in a more detailed analysis of

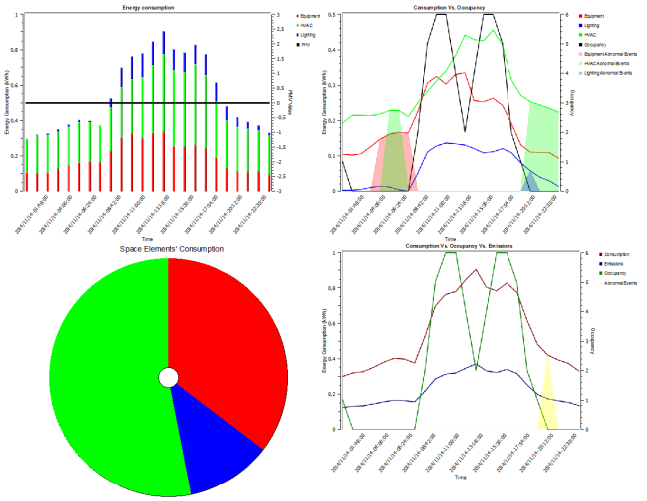

Fig. 5. Space energy consumption. **Top left:** Bar chart displaying the energy consumption per category over time. **Top right:** Energy consumption versus occupancy. **Bottom left:** Individual energy consuming per element. **Bottom right:** Total energy consumption versus CO_2 emissions and occupancy.

the parameters affecting KPIs in every space, it is achieved a thorough, graphically represented comparison in four graphs: a) the space's energy consumption analyzed in every parameter, b) the element's consumption, c) consumption VS occupancy and d) consumption VS occupancy VS emissions (Figure 5).

In order to become more clear the procedure and profit of the comparison, an attempt will be made in this paper to compare two different designs of the same building. For demonstration reasons the primary design of the building consists of two spaces with all the necessary information regarding usage, materials, location, equipment *etc.* already introduced into the design and the simulation. By preserving all the aforementioned parameters, the proportions and dimensions of the building, we divide one of the spaces into two new equal ones, creating in this mater a building that consists of 3 spaces.

The kiviat diagram that arises from the comparison of the two new buildings can be seen in Figure 6. With the help of the diagram one can easily identify that the building with 3 spaces has higher Total Energy Consumption and correspondingly higher level of Emissions than the building with the 2 spaces. This change in consumptions leads to a lower Building Operational Performance and a higher predicted Dissatisfaction of the building users, a logical conclusion considering that after the division, the new space has new heating and cooling demands and therefore there is higher energy consumption which lowers the overall building performance. The percentage however of Work Efficiency seems to be higher in this case, possibly because the Overcrowding factor is quite lower than the 2 space building, meaning that the working environment can be more

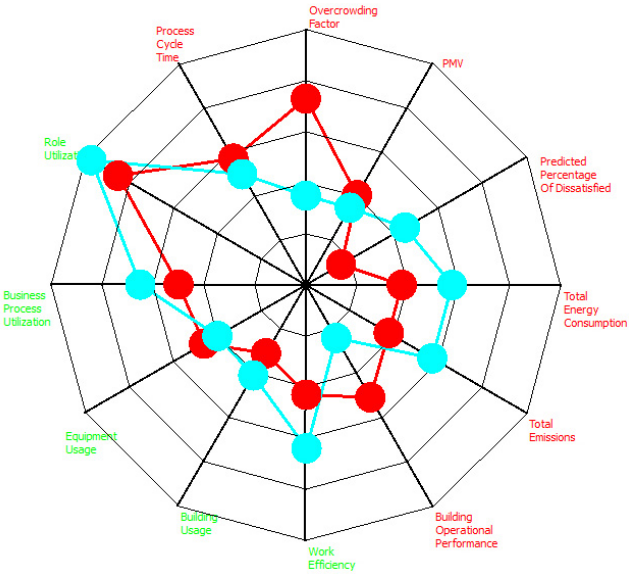

Fig. 6. Building performance comparison between alternative designs.

pleasant without the commotion that an overcrowded space has to face. Certainly this comparison doesn't conclude here as there are many other KPIs to be considered nor provides with a clear answer to which solution is better than the other; in every case selected there are always advantages and disadvantages that only the architect can decide which he/she can overlook or prioritize in order to achieve the best solution according to his/hers design scenario.

4 Conclusions

In this paper a building performance comparison application has been presented. It provides a large number of visualizations, allowing the user to evaluate the performance of a building from the energy consumption point of view using an intuitive graphical user interface. It has been designed in such a way that can deal with the large data volume produced by the simulation engines, providing to the user a set of tools that allow him/her to refine the analysis to a desired level of detail or focus on specific building elements.

Finally it must be pointed out that the proposed visual analytics techniques could provide comparative information about alternative building designs, before their construction, even at their design phase. The analyst/engineer could be able to determine the optimal solution regarding a number of well-defined building KPIs. All, these characteristics render the proposed application to a promising tool to the construction industry, as well as to the big data analytics domain.

References

1. Burch, M., Weiskopf, C.V.F.B.S.D.D.: Parallel edge splatting for scalable dynamic graph visualization. IEEE Transactions on Visualization and Computer Graphics 17(12), 2344–2353 (2011)
2. Chan, W.: A survey on multivariate data visualization. Tech. rep. Depart. of Computer Science and Engin. Hong Kong Univ. of Science and Technology (2006)
3. Cui, W., Zhou, H., Qu, H., Wong, P.C., Li, X.: Geometry-based edge clustering for graph visualization. IEEE Transactions on Visualization and Computer Graphics 14(6), 1277–1284 (2008)
4. Ellis, G., Dix, A.: A taxonomy of clutter reduction for information visualisation. IEEE Transactions on Visualization and Computer Graphics 13, 1216–1223 (2007)
5. Ersoy, O., Paulovich, C.H.F.V., Telea, G.C.A.: Skeleton-based edge bundling for graph visualization. IEEE Transactions on Visualization and Computer Graphics 17(12), 2364–2373 (2011)
6. Gehani, H.: Data visualisation for building performance analysis. Ph.D. Thesis, University College Dublin, UCD (2013)
7. Hailemariam, E., Glueck, M., Attar, R., Tessier, A., McCrae, J., Khan, A.: Toward a unified representation system of performance-related data. In: 6th IBPSA Canada Conference, pp. 117–124 (2010)
8. Henry, N., Fekete, J.D.: Matrixexplorer: A dual-representation system to explore social networks. IEEE Transactions on Visualization and Computer Graphics 12(5), 677–684 (2006)
9. Henry, N., Fekete, J., McGuffin, M.: Nodetrix: A hybrid visualization of social networks. IEEE Transactions on Visualization and Computer Graphics 13(6), 1302–1309 (2007)
10. Holten, D., van Wijk, J.J.: Force-directed edge bundling for graph visualization. Computer Graphics Forum 28(3), 983–990 (2009)
11. Keim, D., Sips, M., Ankerst, M.: Visual data-mining techniques. Bibliothek der Universitat Konstanz (2004)
12. Kintzel, C., Fuchs, J., Mansmann, F.: Monitoring large IP spaces with ClockView. In: 8th International Symposium on Visualization for Cyber Security (VizSec 2011). ACM, New York (2011)
13. Oelke, D., Spretke, D., Stoffel, A., Keim, D.: Visual readability analysis: How to make your writings easier to read. IEEE Transactions on Visualization and Computer Graphics 18(5), 662–674 (2012)
14. Pylyshyn, Z.: Things and places how the mind connects with the world. MIT Press (2007)
15. Rzezniczak, T.: Evaluation of multidimensional visualization techniques for medical patterns representation. Journal of Theoretical and Applied Computer Science 7(4), 70–85 (2013)
16. Stacey, M., Salvatore, J., Jorgensen, A.: Visual Intelligence: Microsoft Tools and Techniques for Visualizing Data. John Wiley & Sons, Inc., Indianapolis (2013)
17. Sun, G., Wu, Y., Liang, R., Liu, S.: A survey of visual analytics techniques and applications: State-of-the-art research and future challenges. Journal of Computer Science and Technology 28(5), 852–867 (2013)
18. Wood, J., Badawood, D., Dykes, J., Slingsby, A.: Ballotmaps: Detecting name bias in alphabetically ordered ballot papers. IEEE Transactions on Visualization and Computer Graphics 17(12), 2387–2391 (2011)
19. Wu, Y., Wei, F.: et al: Opinionseer: Interactive visualization of hotel customer feedback. IEEE Transactions on Visualization and Computer Graphics 16(6) (2010)