VAP: A Visual Analysis Tool for Energy Consumption Spatio-temporal Pattern Discovery

Xiufeng Liu
Technical University of Denmark
Lyngby, Denmark
xiuli@dtu.dk

Zhibin Niu
Technical University of Denmark
Lyngby, Denmark
zhini@dtu.dk

Yanyan Yang
School of Computing, University of Portsmouth
Portsmouth PO1 3HE, UK
linda.Yang@port.ac.uk

Junqi Wu
Tianjin University
Tianjin, China
wujunqi@tju.edu.cn

Dawei Cheng
Shanghai Jiaotong University
Shanghai, China
dawei.cheng@sjtu.edu.cn

Xin Wang
Southwest Jiaotong University
Chengdu, China
xinwang@swjtu.cn

ABSTRACT

In the context of urbanization and the rapid growth of energy demand, understanding the spatial and temporal dynamics of urban energy use is crucial for identifying energy-saving potentials. In this demo, we present a visual analysis tool, VAP, that allows users to explore the dynamics of urban energy use at different spatial and temporal scales. In contrast to traditional statistical and machine learning methods, the visual analysis based tool focuses on analytical thinking, user interactions and answering business questions by examining different visual analysis views. In the demonstration, conference attendees will interact with VAP and learn its capabilities in discovering typical consumption patterns and spatio-temporal shift patterns from a real-world case study of electricity.

1 INTRODUCTION

In recent years, the availability of high-resolution energy consumption data has exploded at an unprecedented rate, along with the diffusion of smart metering technology. The energy sector is increasingly in need of advanced tools and methods to gain insights from big smart meter data sets for decision-making purposes [1]. However, traditional statistical and machine learning methods fail or are too complex to answer some central business questions such as *"What is the consumption trend or pattern over time?"* and *"Does mass mobility affect energy demand?"*. On the other hand, these questions can be much easier to answer through visual analysis supported by human cognitive capabilities. The visual analysis focuses on analytical reasoning, facilitated by interactive user interfaces, and exploring the views that most effectively answer these questions [2]. The visual analysis encompasses several disciplines, including geographic information systems, information visualization, and data computing. Visual analysis has been employed in bioinformatics, physics, astronomy, and climate, but until now rarely in energy.

The most relevant work we have found is the tool implemented in [3] to study energy consumption in the Chicago area in connection with census data. This tool supports the disaggregation analysis on several spatial levels, but without further functionalities, e.g. for the analysis of spatio-temporal patterns and demand shifts caused by mass behaviors. However, much research has been attempted on spatio-temporal data analysis, which involves the identification of object spatial positions at specific moments, e.g., [2], and the detection of anomalies such as traffic congestion [4], cyber attacks [5], medical diagnoses [6] and more. We believe that there is great potential for visual analysis in the applications of energy management systems, because especially many areas in this sector can achieve better results efficiently and effectively with the help of visual analysis. Among others, these include the study of consumer behaviors and living habits, the planning of energy supply, the development of energy strategies and the design of personalized services.

In this demo, we introduce a visual analysis tool, *VAP*, to support the study of urban energy consumption patterns and dynamics on different spatial and temporal scales. This tool is unique in at least two ways:

- Unlike traditional segmentation analysis, which uses clustering algorithms to find typical patterns, VAP supports typical pattern recognition through visual mining. High-dimensional time series are first reduced to low-dimensional data points, and closely placed together on a view according to their similarity, then users can identify patterns by interactively selecting the points on the view. Therefore, pattern recognition is an interactive process embedded in human cognitive recognition. The identified patterns represent customers with similar consumption behaviors or habits, which can be used to develop targeting demand-response programs, forecast energy consumption, and provide personalized services.
- VAP supports the analysis of energy shift patterns in different spatial and temporal scales. The variation in energy demand over time has been studied intensively, while the variation in demand across different spatial spaces has rarely been studied. The demand shift patterns can be identified and visualized with VAP, and the shifts in high energy demand over time can help utilities plan energy distribution and improve energy flexibility.

In Section 2, we will describe the visual analysis framework, the tool, and the pattern discovery methods. In Section 3, we will outline the demonstration scenarios that illustrate how VAP can be used to discover typical patterns and shift patterns of energy demand through visual analysis in a real-world case study. The conference attendees can interact with VAP by first asking business questions, then probing the answers through interactions with the tool, and finally gaining knowledge from the visualized outcomes. This demo presents the elements that drive research for visual analysis in the energy sector, and provides an outlook

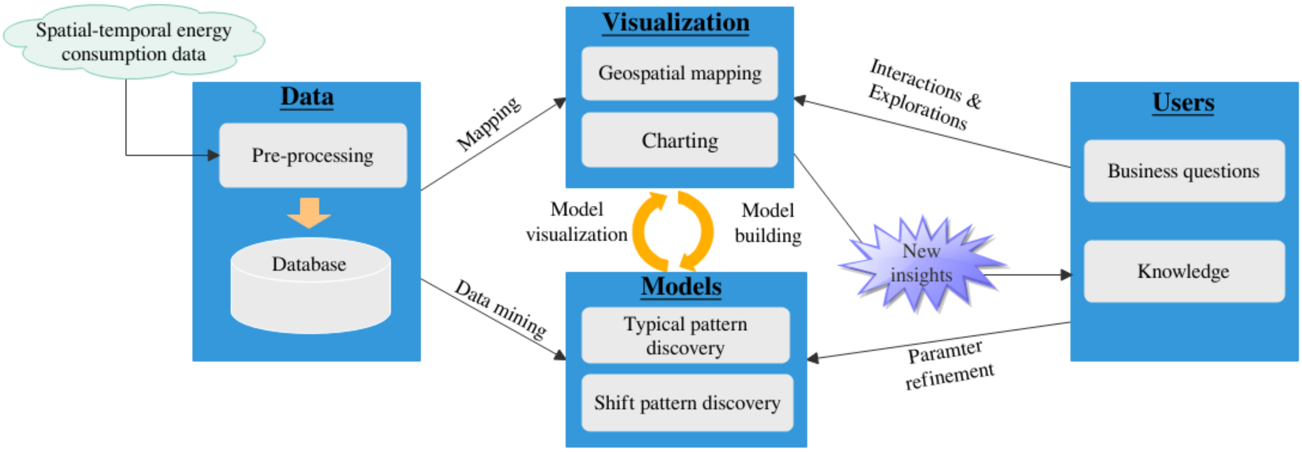


Figure 1: Overview of the visual analysis framework for spatio-temporal pattern discovery

on the use potentials on a higher spatial scale as well as on other urban energy uses.

Through the demonstration of VAP, we hope to increase the awareness of the emerging technologies for the domain-specific applications within the data management community. We also hope that this demonstration will stimulate the research of applying new database technologies in the applications for smart energy systems, including data acquisition, processing, storage, analysis and visualization. We plan to make the code open-source and create an online demonstration version for others to use, experience and extend.

2 APPROACH

Visual analysis aims to help users gain insights into the data by interacting with the visual diagrams that interpret the data. The overview of the visual analysis framework is illustrated in Figure 1, which integrates the components including Data, Models, Visualization and Users. The data for this demo is spatio-temporal energy consumption data from smart meters. The data were pre-processed, including removal of anomalies and correction of missing values. To derive knowledge from the data, the models including typical pattern recognition and spatio-temporal shift pattern recognition were implemented. The visualization aids in the presentation of the analysis results, knowledge generation and communication with users. Users typically gain knowledge by first asking business questions and then exploring analytical views to answer their questions. This is an iterative process of discovering knowledge from the data and refining parameters of the models.

2.1 Pattern recognition models

VAP supports typical consumption pattern discovery and spatial-temporal shift pattern discovery of energy demand using the visual analysis method. The two pattern recognition models are described as follows.

Typical pattern discovery model. Typical consumption patterns are often used to segment customer groups for offering personalized services in the energy sector. In order to visually analyze typical consumption patterns, high-dimensional time series must first be reduced to a lower dimension so that it can be displayed in a low-dimensional space. The proposed model supports the t-distributed Stochastic Neighbor Embedding (t-SNE) [7] and the Multi-Dimensional Scaling (MDS) [8] for high-dimensional time series data reduction. The positions of the resulting low-dimensional data points on the view are determined by minimizing the Kullback-Leibler divergence defined as follows.

$$KL(P\|P') = \sum_{i \neq j} P_{ij} \log \frac{P_{ij}}{P'_{ij}} \tag{1}$$

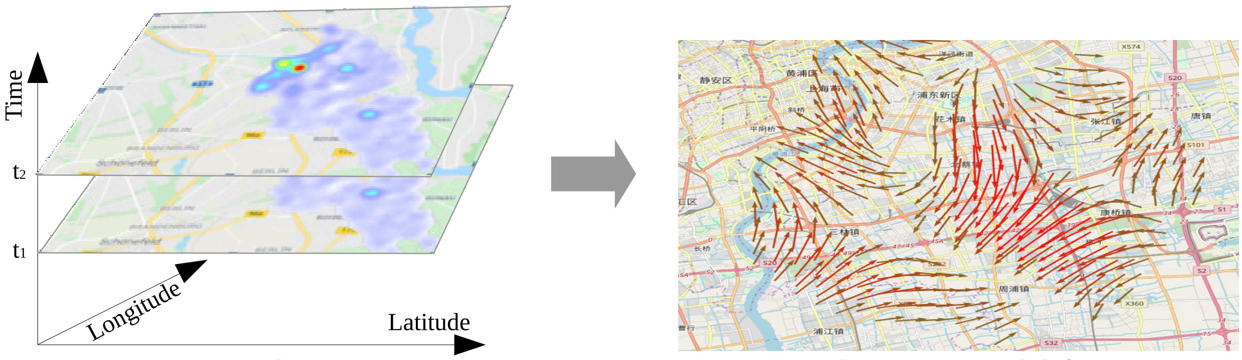
where P_{ij} is the similarity probability distribution between the high-dimensional objects, o_i and o_j, while P'_{ij} is the similarity probability distribution between the reduced low-dimensional objects, o'_i and o'_j. Here, the Pearson correlation coefficient is used as the distance metric for calculating the similarity as it can be better to reflect the correlation of the trend between two time series [9]. The similarity probability distribution of the low-dimensional objects can be obtained by:

$$P'_{ij} = \frac{\left(1 + \left\|o'_i - o'_j\right\|^2\right)^{-1}}{\sum_{k \neq l}\left(1 + \left\|o'_k - o'_l\right\|^2\right)^{-1}} \tag{2}$$

where k and l are the indices of the objects; and $\|\cdot\|$ represents the distance. The similarity probability determines whether the two objects should be placed closely or far away on the view, i.e. the more similar, the closer.

With the reduced data points, typical patterns of energy consumption can be interactively identified by selecting the closely placed points on the view in a low-dimensional space.

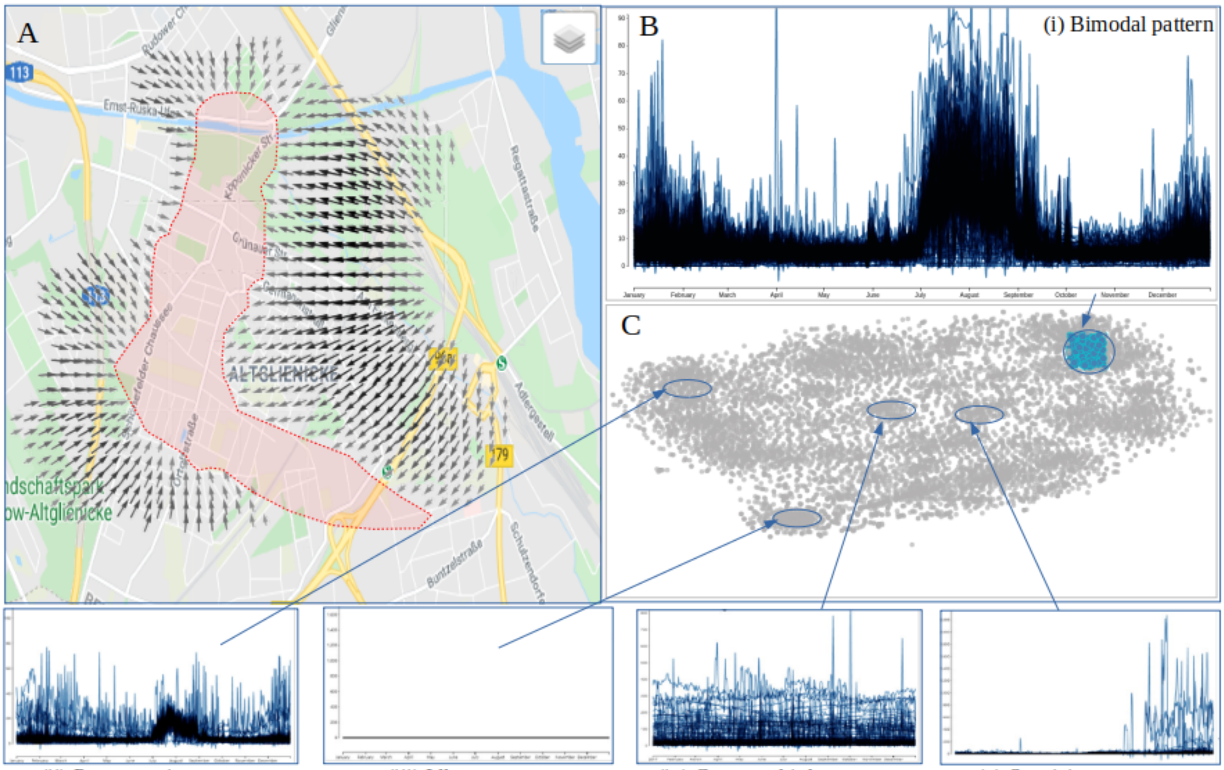
Shift pattern discovery model. This model is used to capture the changes of high energy demand locations over time. The shift pattern of energy demand is essential for planning the energy supply between different geographical areas. For example, the high-demand area may shift from commercial to residential when people go home after work. This can be happening within the time interval of 1 - 2 hours, but the detection of demand shifts is helpful for energy supply planning. Here, we use the flow map method [10] to visualize the spatial migration of high-energy demand flow, which implies the flow between spatially different areas. Figure 2 illustrates the flow map method with a schematic diagram. First, the spatio-temporal distributions of the discrete energy demand are expressed as two different density-strength maps over the time from t_1 to t_2 (see Figure 2a). The density strength map can be obtained using a kernel density estimation method (KDE). Then, the flows are obtained by the difference between the strength maps in two time steps. The resulting flow



a) Density strength map over time

b) Spatio-temporal shift patterns

Figure 2: Flow map method for shift pattern discovery



(ii) Power-saving pattern (iii) Idle pattern (iv) Constant high pattern (v) Suspicious pattern

Figure 3: Spatio-temporal pattern discovery using VAP

map represents the spatio-temporal shifts of high energy demand (see Figure 2b).

The method is formalized in the following. Let $x_1, x_2, ..., x_n$ be the discrete geographical locations of the customers, x_i denoted as a vector $(lon_i, lat_i)^T$ of longitude and latitude. The density of a position of x in a 2D space is defined as follows:

$$\widehat{f_{2D}}(x) = \frac{1}{n}\sum_{i=1}^{n} c_i K_h (x - x_i) = \frac{1}{nh}\sum_{i=1}^{n} c_i K\left(\frac{x - x_i}{h}\right) \quad (3)$$

where n is the sample size, c_i is a normalized value of the average energy consumption used to re-weight the demand strength with respect to the geographical distribution, and h is the bandwidth of the kernel K_h. Gaussian kernel is used in the implementation, but note that other kernels can also be used. Since the Gaussian kernel can cover a larger spatial area for the changes, and it has a lower computation complexity compared with other kernels with exponential functions. The flow patterns representing the energy demand shift between t_1 and t_2 can be expressed as the density difference as follows:

$$Shift(x)|_{(t_1,t_2)} = \widehat{f_{2D}}(x)|_{t_2} - \widehat{f_{2D}}(x)|_{t_1} \quad (4)$$

2.2 Visual analysis tool

The visual analysis tool VAP is implemented as a web application with an architecture of three layers *data layer, logic layer* and *presentation layer*. In the data layer, PostgreSQL is used as the database management system, with PostGIS added to support spatial data processing. In the logic layer, all algorithms are implemented in Python, including the typical pattern discovery and spatio-temporal RESTful APIs are implemented to exchange JSON-formatted data between client and server. In the presentation layer, HTML5, CSS and JavaScript are used to implement the user interface (see Figure 3). Especially the JavaScript library, Leaflet.js [11], is used for the visualization of Scalable Vector Graphics (SVG) and the mapping. The flow patterns are displayed as colored arrows on the map, and the color depth represents the rate of change of the flow patterns; the darker the color, the higher the rate. The JavaScript library d3.js [12] is used for time series visualization.

Figure 3 shows the main user interface of VAP, which consists of the three views: A, B, and C. View A displays the spatial information of the clustered customers. This view supports users in selecting different map types, displaying the geographical positions of customers with markers, and visualizing the spatial distribution density with a heat map and spatio-temporal shift patterns with flow vectors. View B shows the time series for the customers selected in view C. This view visualizes the typical consumption pattern for all selected customers. View C is an interactive navigator that allows users to explore different energy consumption patterns by selecting the points by clicking and dragging. The closer the points are to each other, the more similar the patterns will be.

VAP supports visual analysis for any type of energy consumption with spatial information. The design of the database model can be interpolated with specific energy data sets in real use cases. Figure 3 illustrates an example for the visual analysis of an electricity consumption data set (note: the coordinates are offset for anonymization), where five typical patterns were discovered, including (i) bimodal pattern, (ii) energy-saving pattern, (iii) idle pattern, (iv) constant high pattern and (v) suspicious pattern. View A and B shows the flow map and the aggregated consumption pattern for the customers selected in the reduced 2D space in view C, respectively. The pattern is a bimodal pattern with a peak in winter and summer, respectively, which may be caused by the use of electrical heating and cooling appliances. The other four typical patterns can be interpreted in a similar way by examining the consumer behaviors or habits of customers, which will be explained in the demonstration. In view A, the area covered by the arrows is a commercial area, while the area pointed to by the arrows is a residential area (i.e. the light red area). This flow map indicates that the area with high energy demand is shifted to the residential area when people go home after work.

3 DEMONSTRATION

We will use a real-world electricity consumption data set for the demonstration. We will introduce the system architecture, the visual analysis process and how to use the tool to solve practical problems in energy planning. Conference attendees will interact with the tool to perform visual analysis on the data, i.e., ask business questions, answer questions through the exploration of the views and acquire knowledge. In particular, the participants will interactively discover typical and shift patterns of electricity consumption with VAP. We aim to help participants learn more about visual analysis for energy decision makings, explore the approaches for energy demand-side management, and raise their awareness of energy savings. Two demo scenarios will be presented during the conference.

S1: Typical patterns discovery. In this demo scenario, conference attendees will interact with VAP by investigating typical energy consumption patterns and identifying the spatial distribution of customers in the study area. We will explain to the participants the meaning of each identified pattern and the reason behind it. First, an attendee can start by asking questions, e.g., *who are the early birds with a morning peak between 5:00–7:00?* They investigate the customers of interest by selecting the points at different places on the view. Second, the attendees will study the transition of consumption patterns based on point similarity (or point spacing) in 2D space. They select the closely placed points continuously, and observe the pattern transition over the spatial space. Third, they can select the scatter plots generated by different dimensional reduction methods, including t-SNE and MDS, observe difference and compare capabilities in typical pattern discovery. Fourth, we will run the k-mean algorithm on the sampled data to discover typical patterns, compare the results, and explain the advantages of using the visual analysis method.

S2: Spatio-temporal shift pattern discovery. In this demo scenario, conference attendees will examine the patterns of the energy demand interactively, and we will explain the discovered shift patterns accordingly. First, an attendee examines the shift patterns by varying the temporal granular intervals, including hourly, every four hours, daily, weekly, monthly, quarterly, and yearly. They can then learn the sensitivity of the shift pattern changes against different time granularity. Second, they select different customer groups according to the consumption intensity in a quartile value ranging from 30% to 90%. They can then learn the sensitivity against different energy consumption intensities. Third, we can further demonstrate the shift pattern dynamics through a simulation. If, for example, the data are fed to the system in a short time interval, e.g., every 10 seconds, we can observe the changes of patterns in near real time.

4 CONCLUSION

We presented a visual analysis framework and a tool that supports both spatial and temporal pattern analysis for smart energy systems. We described the technique of dimension reduction and discussed how to reduce and visualize high dimensional data to a low dimensional space. We described and demonstrated the discovery of typical consumption patterns and spatial-temporal shift patterns of energy demand using a real case study. The demonstration validated the plausibility of the proposed visual analysis framework and the effectiveness of the tool in knowledge discovery through user interactions.

Acknowledgements. This research was supported by the EU H2020 Research and Innovation Program under the Marie Sklodowska-Curie Grant Agreement (754462), the Flexible Energy Denmark project funded by Innovationsfonden (8090-00069B) and the project FlexSUS: Flexibility for Smart Urban Energy Systems (91352) funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 775970.

REFERENCES

[1] N. Iftikhar, X. Liu, F. E. Nordbjerg, and S. Danalachi. 2016. A Prediction-based Smart Meter Data Generator. In Proc. of NBIS, 2016, 173–180.
[2] N. Andrienko, and G. Andrienko. 2013. A visual analytics framework for spatio-temporal analysis and modelling. Data Mining and Knowledge Discovery 27, 1 (2013), 55–83.
[3] J. T. Trabucco, D. Lee, S. Derrible, and G. E. Marai. 2019. Visual Analysis of a Smart City's Energy Consumption. Multimodal Technologies and Interaction 3, 2 (2019), 30.
[4] F. Zhou, W. Huang, Y. Zhao, Y. Shi, X. Liang, and X. Fan. 2015. Entvis: a visual analytic tool for entropy-based network traffic detection. IEEE graphics and applications 35, 6 (2015), 42–50.
[5] L. Layman, N. Zazworka. 2014. InViz: instant visualization of security attacks. In HotSoS, (2014), 15.
[6] F. Ritter, T. Boskamp, A. Homeyer, H. Laue, M. Schwier, F. Link, and HO. Peitgen. 2011. Medical image analysis. IEEE pulse 2, 6 (2011), 60–70.
[7] LVD. Maaten, and G. Hinton. 2008. Visualizing data using t-SNE. Journal of machine learning research, 9 (2008), 2579–2605.
[8] J. B. Kruskal. 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis Psychometrics, 29 (1964), 1–27.
[9] X. Liu, N. Iftikhar, H. Huo, R. Li, and P.S. Nielsen. 2019. Two Approaches for Synthesising Scalable Residential Energy Consumption Data. Future Computer Systems 95, (2019), 586–600.
[10] D. Guo, and X. Zhu. 2014. Origin-destination flow data smoothing and mapping. IEEE Transactions on Visualization and Computer Graphics 20, 12 (2014), 2043–2052.
[11] Crickard III P. 2014. Leaflet. js essentials. Packt Publishing Ltd, 2014.
[12] F. Bao, and J. Chen. 2014. Visual framework for big data in d3. js. In Proc. of IEEE Workshop on Electronics, Computer and Applications, 2014, 47–50.