International Workshop on Visual Analytics (2012)
K. Matkovic and G. Santucci (Editors)

AMPLIO VQA – A Web Based Visual Query Analysis System for Micro Grid Energy Mix Planning

A. Stoffel[1] and L. Zhang[1] and S. H. Weber[2] and D. A. Keim[1]

[1]University of Konstanz, Germany
[2]Siemens AG

Abstract

Micro grid technology brings the opportunity to integrate renewable energies with a traditional energy mix on a regional level and to achieve specific local goals. To obtain an optimal energy mix plan, different scenarios need to be simulated and analyzed. However, effective tools for analyzing such simulation results are lacking. Here we present an interactive visual query analysis tool designed for that purpose. The tool integrates effective visualization techniques and advanced pattern detection methods.

1. Introduction

With the adoption of renewable energies such as wind, solar, biomass, and hydro power in recent years, there is a rising trend towards localized electricity systems, which generate, store, and distribute electricity to consumers independently from other regions. Such systems are called micro grids (see Figure 1). Micro grids are ideal for integrating renewable energies into the existing energy mix on a community level, because they are designed to achieve specific local goals such as reliability, carbon emission reduction, diversification of energy sources, and cost reduction, established by the community being served [Mic].

The shift from a traditional energy mix to renewable energies is complicated. A large number of parameters have to be taken into account to enable a sustainable and competitive expansion of renewable energies. The availability, import/export price, and carbon emission cost of every energy type, the power demand over time, and the planned energy mix need to satisfy different sets of local goals. For example, we may want to minimize production costs and carbon emissions. Hence Siemens AG developed a web based framework called AMPLIO as a planning tool for designing new energy mix. The tool allows the analyst to simulate different scenarios based on available energy data. However, given the complex parameter settings, the simulation result is difficult to interpret and the interrelationships between parameters are still hidden. An effective tool needs to be designed to help providers analyze the results until a satisfactory plan is achieved. Since the energy data are time related, the tool has to handle multivariate time series data.

Although a number of software systems exist for time series data analysis, (e.g. TimeSearcher [HS04], Tableau [Tab], and Spotfire [Spo]), we found it difficult to perform all the analysis tasks that we require by any single tool. For example, we want to be able to define a target pattern by selecting a time frame and a subset of dimensions that are relevant to the analysis task and search for similar patterns in the data. Also it should be easy to integrate the tool into the existing micro grid planning framework.

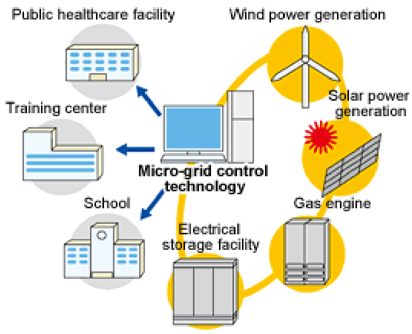

Figure 1: *Micro Grid Electricity System*

Towards this end we designed and developed AMPLIO VQA — a web application for supporting the strategic planning of a micro grid energy mix based on simulation results. The tool integrates effective visualization techniques and advanced data mining methods to help the user find the optimal plan for a micro grid. The system is designed for analyzing energy-related data; however, it can be easily extended to be used for other applications that have data of a similar nature.

The main contributions of this paper comprise: 1) an integrated system for micro grid energy mix planning, 2) the

DOI: 10.2312/PE/EuroVAST/EuroVA12/073-077

integration of line chart and recursive pattern visualization for visual pattern detection in time series data, and 3) the advanced pattern search approach for time series data mining.

2. Related Work

Data analysis and visualization is a well-studied topic, with many software systems having been developed in the past years. Among all the systems, TimeSearcher [HS04] is one of the most well-known open source applications for query based time series analysis. The tool allows the user to define one or more time boxes on the river plot visualization as a target pattern, and searches for similar patterns in the remaining part of the data. Given a selected time box, the tool can also predict the trend by looking for a data fragment which contains the most similar subset of curves in historical data and making inference based on it. The system is highly interactive and implements a simple user interface that makes search and comparison easy. However, the pattern match searching algorithm only takes the values of the start and end point of a time box into consideration; there is a risk of missing important information that lies in between. In addition, the user lacks the freedom of selecting subset of dimensions for further analysis.

A large proportion of time series data analysis software systems is designed for specific application domains. For example, Thomson Reuters' [Reu] data analysis toolkit for analyzing financial time series and CDAT [CDA] for climate time series analysis. Matlab [Mat] and R [R] also support time series data analysis and provide some basic visualization functions; however, they both require a certain level of programming experience. There are also a good number of commercial products, for example, Tableau [Tab], Spotfire [Spo], and Qlikview [Qli] that are designed for general data analysis and visualization purposes. Each of them has its own strengths, but most of them implement a large number of functionalities, some of which are not relevant to our application, and on the other hand do not include all the functionalities that are required by our application. For example, we need search for patterns based on a subset of dimensions, display level of similarity between target and found patterns, visualization which enables easy visual detection of temporal patterns.

3. System Design

3.1. System Architecture

Since the main objective of our system is to analyze simulation results generated by AMPLIO, one major request from the user is to be able to integrate the analysis functionalities into the existing web-based AMPLIO application framework and allow linking and brushing between different views. To fulfill such requirements, we implement the user interface based on the canvas element of HTML 5 and Java Script running in a browser. We develop our system as a remote Java Server Page [DLRmC09] application to provide data sources and perform computation intensive analysis tasks.

A critical issue for a web based solution is the responsiveness of the system. Transferring large amounts of data between the client and the server could result in a slow response of the client interface while it is waiting for new data. This could cause serious delays when the user links and brushes between several views. Hence we decided to put all the rendering load of the visualization onto the browser. In combination with caching of the data set in the browser, linking and brushing only requires the transfer of a small amount of meta information between the browser and the server. Thus the system updates the client interface very quickly while linking-and-brushing different views.

3.2. Visualization and Analysis Functions

Line based visualization (e.g. line chart, area chart) is one of the most widely used techniques to display time series data. A line based visualization provides an intuitive way to see the trend in data over time. However, with multivariate time series data, there is always a problem of getting an overcluttered display when the data values in different series are similar. One possible solution is the pixel based visualization technique called recursive pattern [AKK95]. A recursive pattern visualization maps the time series into colored pixels. The color represents data values, the order of pixels is defined in a recursive manner. This day the time series can be investigated using different scale, for example, hourly, daily, month, etc. In our design, each pixel represents a hour by default, and the 24 hours of a day form a row. 7 days of the week are staked up such that the top row represents Monday. And the weekly blocks are put from left to right and top to bottom. Such technique can provide a compact visualization for a large number of data values without overlapping, and the colored grid units make it easy to visually detect patterns. AMPLIO VQA combines both techniques as complementary graphical representations of the data. Linking-and-brushing is implemented over all visualization panels to support easy comparison.

For visual query analysis, we take an approach similar to TimeSearcher, that is, the user can define a time frame in the series and search for similar patterns. However our approach improves on the time box approach in two ways: 1) The user can select a subset of the dimensions on display when defining the target pattern. The search will then be based on the data values in selected dimensions. 2) Instead of checking only the start and end value, the pattern matching algorithm is based on similarity between two time frames over all the data values in the selected dimensions. AMPLIO VQA implements several similarity measures for the automated pattern search. In addition, the degree of similarity between the target and the found patterns are displayed to help users gain better understanding of the result. The tool also implements a number of normalization methods for data preprocessing.

Figure 2: *The AMPLIO VQA tool combines line charts and recursive patterns visualizations to explore the data (A). The line chart shows the typical battery storage cycle over two days (normalized with Z-score). In the recursive patterns the same time frame is marked with blue rectangles and black rectangles show the 4 most similar time frames. Searching for similar time frames is controlled with (B) and the most similar time frames can be analyzed in detail (C).*

4. Evaluation

We analyze a simulation result for an energy provider of a major German city. The supplier wants to find out whether it is sensible to install batteries for energy storage in the future. The assumption is that batteries can serve as an energy buffer to reduce production costs. For example, we can charge batteries when large amounts of energy are generated by the local system or the import price is low, and discharge batteries for local supply when the import price is high or export energy when the export price is high. However, the assumption will not be valid unless the cost of CO_2 emission is also taken into consideration. The current cost of €7 to €13 per tCO_2 emission is expected to increase in the future.

To find out whether it is a good strategy to install batteries, the supplier simulated the system capabilities in the year 2015 and set the CO_2 emission costs to €30 per tCO_2, which is the highest value in historical data. Figure 2 shows the storage level of the battery over the year in the simulation data. In the recursive patterns we can see that the battery is typically charged in the morning and then discharged during midday. During winter it is common that a second charging and discharging phase happens in the afternoon (1). Interestingly, this second phase is not present in the summer term except for Fridays (2). Using a different normalization, we found out that European Energy Exchange (EEX) price rises in the morning and drops from noon, which may explain the pattern. However on Friday afternoon the EEX price usually increases slightly. The battery exploits this pattern by adding another charge/discharge cycle.

The analysis also reveals that batteries are usually charged when the imported energy amount is high and the EEX price is low, and discharged when the export amount and the EEX price are high. To find out whether the EEX price has a real influence on this pattern, we use the search function to look for similar patterns in the data. Figure 3 shows a part of the search result. The black rectangles in the recursive pattern visualization highlight the 20 most similar time frames. The highlighted regions show a common pattern that correlates with the low energy demand during night and high demand in the morning: during night the fossil power stations are often shut down, and imported energy as well as renewable energy becomes the main energy supply. The fossil power stations are usually powered up in the morning and often generate more energy than locally demanded (also for exporting). Finally, we came to the conclusion that in general the provider imports energy during night to charge batteries when the EEX price is low and sells stored energy in the morning when the energy price increases.

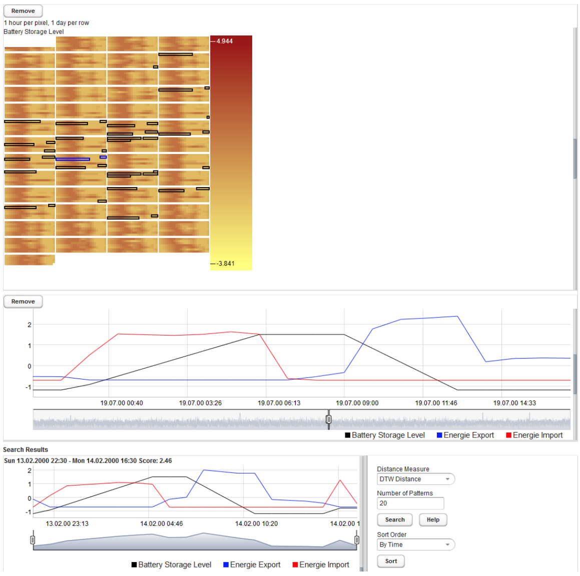

Figure 3: *Energy is imported and stored in the batteries over night and exported during the morning.*

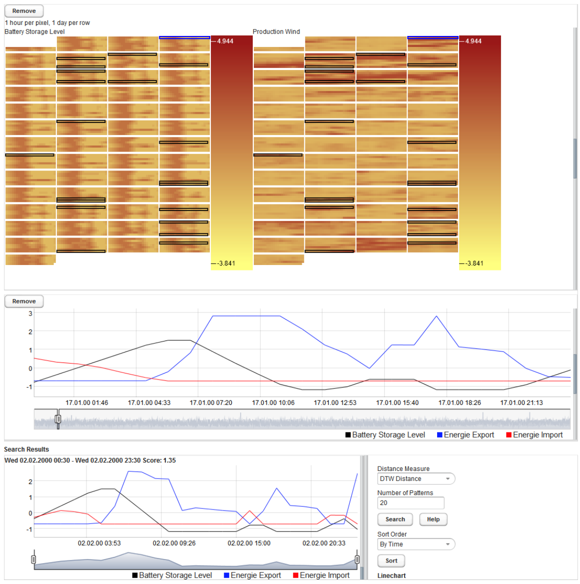

Figure 4: *When large amount of wind energy is generated the battery is used to shift wind energy.*

We observe that the battery power level changes when much wind energy is generated, especially during winter. For instance, in Figure 4 a large amount of wind energy was produced on the 17th of January. On that day the battery was charged and import was reduced. When the battery was full, the wind energy was exported. The energy saved by the battery was later sold during high EEX price phases. So we made the assumption that the battery helps to shift energy supply, but the planning is often based on production costs and sales price rather than local demand. To test our hypothesis, we searched for similar time frames in the data to find out whether this is a singular artifact or a common pattern. The similar patterns in the query result shown in Figure 4 indicate that it is indeed a common pattern.

For further evaluation we analyze the same simulation result with two other existing systems, Tableau and TimeSearcher. The former is one of the most popular commercial data analysis and visualization tool, and the latter is one of the most well-known open source tool for query based time series data analysis. We found that with the aggregation functionality of Tableau, it is easy to find out the global dependencies between the different variables. However, it is difficult to identify rare (small) patterns in Tableau, for instance, finding time periods when much wind energy is generated and additional energy is imported. Such small effects often get lost due to aggregation. The user has to drill down in the original data to find them. Verifying whether this effect is a common pattern is also hard, because Tableau does not support similarity based pattern search.

TimeSearcher supports visual query analysis. Thus it is easy to define a target pattern and search for similar patterns. But unlike Tableau the system does not implement an aggregation functionality; hence it is hard to analyze the query result. For example, verifying the hypothesis about high wind energy generation leading to high energy export is hardly manageable with TimeSearcher. Another drawback is that the search function only supports searching over all dimensions, which in this particular case leads to many false positive search results.

5. Conclusions

Designing an energy mix plan for a new micro grid system is complex. Often the designer has to simulate different scenarios based on a large number of parameters. Here we presented a web application that integrates with the existing energy planning framework to help designers to analyze the simulation result. The system combines complimentary visualization techniques with advanced pattern search methods such that users can mine interesting patterns from the result and understand the interrelationship between parameters. Design choices were carefully made to guarantee the responsiveness of the system while linking and brushing different views. The system was evaluated by testing with real data and comparing with existing tools.

In the future we will support interactive parameterization such that hypothesis testing becomes more straightforward. We also plan to design and integrate novel and effective visualization techniques to enable examination of the data and simulation results from different perspectives. More advanced data analysis methods, for example, automatic anomaly detection algorithms will be integrated to support deeper level analysis.

References

[AKK95] ANKERST M., KEIM D. A., KRIEGEL H.-P.: Recursive Pattern: A Technique for Visualizing Very Large Amounts of Data. In *Proc. Visualization '95, Atlanta, GA* (1995), pp. 279–286. 2

[CDA] http://www2-pcmdi.llnl.gov/cdat. 2

[DLRmC09] DELISLE P., LUEHE J., ROTH M., MAN CHUNG K.: JavaServer Pages specification. `http://www.jcp.org/en/jsr/detail?id=245`, 2009. 2

[HS04] HOCHHEISER H., SHNEIDERMAN B.: Dynamic query tools for time series data sets, timebox widgets for interactive exploration. *Information Visualization 3*, 1 (2004), 1–18. 1, 2

[Mat] http://www.mathworks.de/products/matlab/. 2

[Mic] http://www.shimz.co.jp/english/theme/sit/technology_03.html. 1

[Qli] http://www.qlikview.com/. 2

[R] http://www.r-project.org/index.html. 2

[Reu] http://thomsonreuters.com/. 2

[Spo] http://spotfire.tibco.com/. 1, 2

[Tab] http://www.tableausoftware.com/. 1, 2