Exploring Cyber Physical Data Streams Using Radial Pixel Visualizations

M. Hao[1], M. Marwah[1], S. Mittelstadt[2], H. Janetzko[2], D. Keim[2],
U. Dayal[1], C. Bash[1], C. Felix[1], C. Patel[1], M. Hsu[1], Y. Chen[1], M. Hund[2]
[1]Hewlett-Packard Laboratories, Palo Alto, CA [2]University of Konstanz, Germany

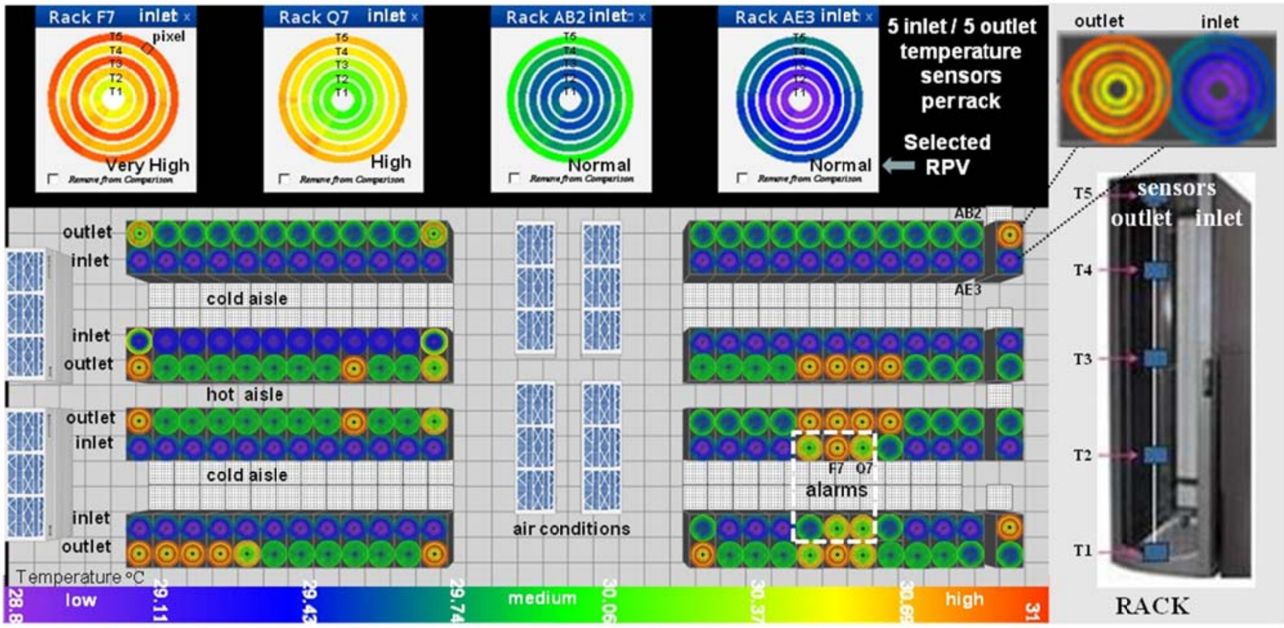


Figure 1: This figure shows Radial Pixel Visualization (RPV) of rows of racks and hot/cold aisles in a 500 kilowatt data center over the past 24 hours in real-time. Each rack has a pair of RPV's depicting data from 5 inlet/outlet temperature sensors which should always be below a threshold and usually also maintain an ascending sequence, that is, T1<T2<T3<T4<T5, where T1 is closest to the floor. Each measurement is represented by a pixel cell. Color depicts temperature from low (purple) to high (red). By correlating thermal alarms and their physical locations and by looking at temperature patterns in the recent past, administrators are able to quickly identify problems (e.g., Rack F7 has out of sequence sensor: T3>T4) and find the root causes of those alarms.

Abstract

Cyber physical systems (CPS), such as smart buildings and data centers, are richly instrumented systems composed of tightly coupled computational and physical elements that generate large amounts of data. To explore CPS data and obtain actionable insights, we construct a Radial Pixel Visualization (RPV) system, which uses multiple concentric rings to show the data in a compact circular layout of small polygons (pixel cells), each of which represents an individual data value. RPV provides an effective visual representation of locality and periodicity of the high volume, multivariate data streams, and seamlessly combines them with the results of an automated analysis. In the outermost ring the results of correlation analysis and peak point detection are highlighted. Our explorations demonstrate how RPV can help CPS administrators to identify periodic thermal hot spots and understand data center energy consumption.

Keywords: Radial pixel visualization, cyber physical system, correlations, peaks, time-varying data.

Index Terms: CR Categories and Subject Descriptors: 1.3.3 [Computer Graphics]: Picture/Image Generation – Display Algorithms; H.5.0 [Information Systems]: Information Interfaces and Presentation – General.

1 Motivation

Cyber physical systems (CPS) [1] are systems which are characterized by a tight coupling between computational and physical infrastructures. Examples of such systems include smart buildings, data centers, smart electric grids, etc. As illustrated in Figure 1, data center administrators are interested exploiting this data in real-time to identify thermal problems, infer the root causes of anomalies, determine under- or over-utilized resources, etc. Specifically, relevant questions include: Are temperature alarms that originate from different servers related? Are the alarms from servers in close physical proximity due to the same underlying causes?

2 Our Approach and Contributions

To meet CPS challenges, we propose a radial pixel visualization which enables the visualization of large amounts of spatio-temporal sensor data. The basic idea is to visualize a data stream in multiple concentric rings to represent a single measurement in a time interval. The pixel cells (measurement values) are aligned by time across different rings which allow an easy detection of patterns and anomalies. They are also accessible to the user for query and drilldown. For example, in Figure 1, administrators are able to adjust the vent tile to fix the thermal problem, when they discover that multiple alarms are in fact co-located around rack F7. Furthermore, the results of correlation and peak detection are highlighted in the outermost rings to help users to quickly identify the important information in large data streams (see Figure 2). These unique features do not exist in today's radial visualizations [2] but are important for many CPS applications.

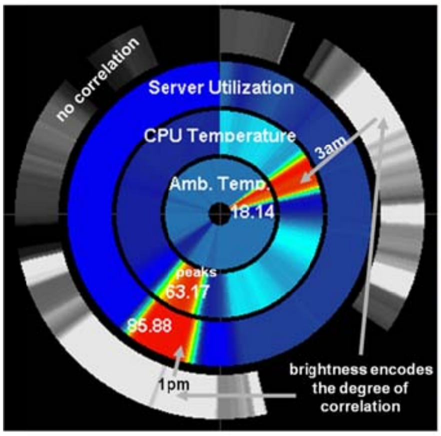


Figure 2: Daily Thermal State Analysis
(3 rings show: Server, CPU, Amb. Temp)

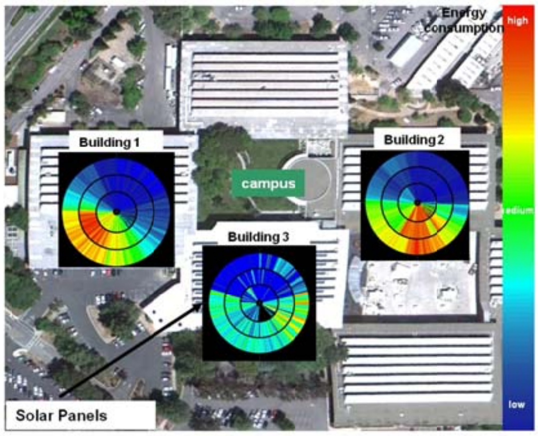
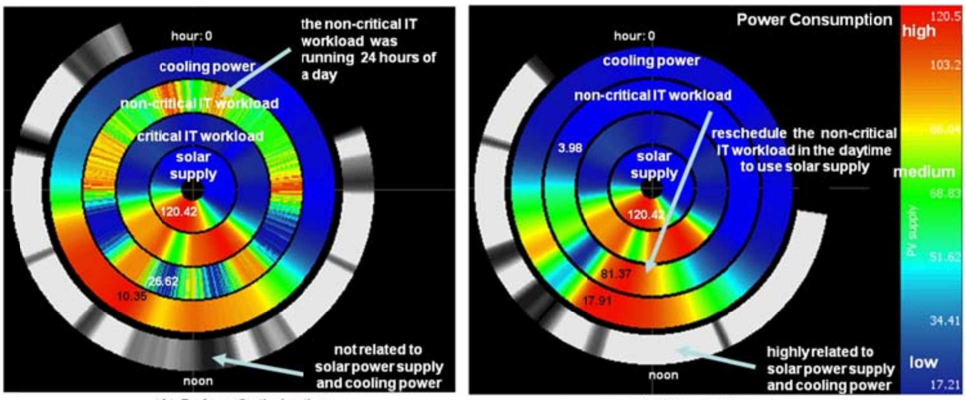


Figure 3: Building 3 has flat energy consumption (Solar)
(Each ring shows one day consumption, 9/1-9/3)



(A) Before Optimization
The distribution of the non-critical workload
does not use the solar power supply

(B) After Optimization
Reschedule the non-critical workload in the daytime
to use the solar power supply

Figure 4: Use RPV to visually validate the benefits of rescheduling the non-critical IT workload

3 Visual Exploration

3.1 Thermal State Analysis

In this application scenario, we compute the average of all pair-wise correlation-results shown by brightness in the outermost ring. As illustrated in Figure 2, CPU temperatures are high in two different time intervals (around 3am and 1pm). When observing only the temperature ring, the root cause of the high temperature would remain unclear. By plotting the ambient temperature and server utilization in Figure 2, it becomes evident that the first increase in CPU temperatures at 3am is related to a period of high ambient temperature while the second increase at 1pm is related to high server utilization. With this knowledge, the data center administrators are able to manage their resource consumption more effectively.

For thermal state analysis, peak points also have a high level of significance. Peaks in temperatures could reduce hardware reliability. We can use the outermost ring to show the significant peaks in the corresponding time intervals instead of showing correlations.

3.2 Building Energy Consumption Characterization

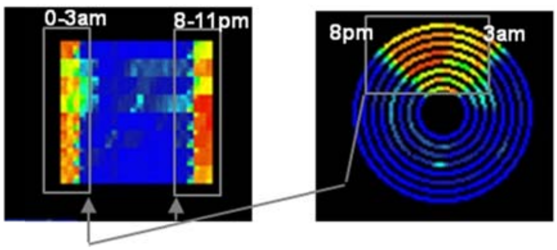
Figure 3 shows the power consumption of three buildings at an urban campus over three days. The daily usage patterns are easy to see for Buildings 1 and 2. As expected, consumption is high during the working hours (9 am to 6 pm). Building 3's consumption is flat, due to the presence of solar panels that offset part of the demand during the day. The visualizations allow administrators to compare usage patterns between buildings, validate the impact of solar panels under different weather conditions (e.g. sunny, cloudy), etc.

3.3 IT Workload Optimization

Figure 4 shows attributes pertaining to IT workload and cooling power in a data center over a 24-hour period. Administrators are able to use RPV to visualize the power consumption differences and reschedule the non-critical IT workload in the daytime to use the solar supply. While before optimization, the non-critical workload was spread throughout the day in Figure 4(A), after optimization, it is concentrated during the period of solar panel power generation (supply) as shown in Figure 4(B). The visualization confirms the benefits of rescheduling the workload and reducing the overall daily cooling power consumption.

4 Evaluation and Conclusion

In this poster, we presented RPV, a new approach for visualizing large amounts of multi-attribute sensor data. Our approach combines a pixel cell based radial visualization with efficient and effective knowledge discovery techniques (e.g., correlation analysis and peak point detection). RPV maps each data point to one pixel cell in the ring and displays their correlations between multiple rings and their peak points. We have applied the RPV idea to real data sets from data centers and smart buildings. The resulting Redial Pixel Visualizations provide significantly more information than radial visualizations without using pixel cells. Our explorations also indicate that RPV is more intuitive to visualize a continuous periodic daily pattern using radial coordinates than using Cartesian coordinates as shown in Figure 5.



(A) Shows a split period pattern (B) Shows an entire period pattern

Figure 5: Cartesian coordinates versus RPV in a data center (4/1-4/8)

[1] R. H. Campbell, et al: Cyber-Physical Systems: position paper: CPS Environments. University of Illinois, Urbana-Champaign. 2012.

[2] S. Diehl, et al.: Uncovering Strength and Weaknesses of Radial Visualizations – an Empirical Approach. IEEE Transactions on Visualization and Computer Graphs. November/December 2010.