Proceedings of the 23rd IEEE International Conference on Advanced Information Networking and Applications (AINA 2009), May 26-29, 2009, Bradford, UK.

Street CORNERS: Real-time Contextual Representation of Sensor Network Data for Environmental Trend Identification

Patricia Morreale, Feng Qi±, Paul Croft±, George Chang, Justin Czarnik, Nick Doell, Jhon Espin, and Ryan Suleski
Department of Computer Science, ±Department of Geology and Meteorology
Kean University
1000 Morris Avenue
Union, NJ 07083
{pmorreal, pcroft, fqi, gchang, jczarnik, ndoell, jespin, rsuleski}@kean.edu

Abstract— **Sensor networks have been deployed for a range of rural and environmental applications. Well-regarded for the volume and range of data which can be obtained, wireless sensor network applications capable of using the data gathered have not been fully realized, particularly in urban settings. Street CORNERS is a wireless sensor network application which supports the contextual presentation of data gathered from an urban setting. The Street CORNERS application offers real-time data display, and provides support for predictive algorithms suitable for anticipating, detecting, and defending urban communities, among others, from environmental threats such as declining air quality and urban flash floods. Street CORNERS is presented in two parts. The network design and deployment is outlined, followed by a discussion of the design of the network application, which is involved in data pre-processing and the contextual presentation of the data gathered for trend identification.**

Keywords-wireless sensor networks; application design

I. INTRODUCTION

Urban ecosystems, such as those defined by street corners and other public spaces in our cities and towns, are interconnected with both well-being and our urban lifestyles. Vehicle emissions and air borne contaminates can impact air quality [1] for the pedestrians while roadways and underpasses can become flooded in the event of a sudden rainstorm [2]. Such events may have unexpected and rapid onset, or be a chronic and recurring problem. In all circumstances, the events affect the environmental sustainability of the urban ecosystem, as a range of impacts and interactions from these events are possible. The *Street CORNERS* project, where *CORNERS* stands for "*Cor*relation of *N*etworked *E*nvi*r*onmental *S*ensors" is an application which presents real-time sensor information in an appropriate geographical context for the identification of environmental trends in support of environmental sustainability.

II. URBAN ENVIRONMENTAL SUSTAINABILITY

Environmental sustainability encompasses several stages. Initially, environmental sustainability referred to development that minimized environmental impact. However, in established areas, such as urban communities, environmental sustainability includes *detection* of potential problems, *monitoring* the impact of potential or actual problems, and *working to reduce* adverse impact of identified threats to environmental sustainability.

The repetitive nature of threats to environmental sustainability in urban environments, such as the underpass that consistently floods during heavy rains, or the air quality that predictably degrades over the course of a workday, is the stuff of urban legend. Neighborhood residents and regular visitors to the area may generally know the hazards of a particular urban spot, but sharing the knowledge of destructive or hazardous patterns with organizations which might be able to remediate or prevent such regular environmental degradation is not easily done. Rather, at each instance of an environmental threat, the flooded underpass or poor air is addressed as a public safety crisis and personnel and resources are deployed in an emergency manner to provide appropriate traffic re-routing or medical attention.

In addition to critical events which threaten urban environmental conditions, more insidious, slowly evolving circumstances which may result in future urban crises are not monitored. For example, traffic volume on highly used intersections or bridges is not monitored for increasing noise or volume, which might result in a corresponding increase in hazardous emissions or stress fractures. When urban threats are identified, the response process can be aggravated as traffic comes to a standstill or is rerouted, hindering or delaying emergency response personnel.

Both immediate and slower threats to urban environmental sustainability are dealt with on an 'as occurring' basis, with no anticipation or preventive action taken to avert or decrease the impact of the urban threat. The result, over the past years, has been an increase in urban crisis management, rather than an increased understanding of how our urban environments could be better managed for best use of all our resources – including resources for public safety and environmental sustainability.

The increasing age of urban infrastructures, and the further awareness of environmental hazards in our midst highlights that the management of environmental threats on an 'as needed' basis is no longer feasible, particularly as the cost of managing an environmental crisis can exceed the cost of preventing an environmental crisis. With the potential to gather data in a real time manner from urban sites, the opportunity for anticipatory preparation and preventive action prior to urban environmental events has become possible. Specifically, street level mapping, using real-time information, is now possible, with the integration of new tools and technology, such as geographic information systems and sensors.

Environmental sustainability in an urban environment is challenging. While numerous measures of environmental sustainability, including air quality, rainfall, and temperature, are possible, the group assessment of these measured parameters is not as easily done. Air quality alone is composed of a variety of measurements, such as airborne particulate matter (PM10), nitrogen dioxide (NO_2), Ozone (O_3), carbon monoxide (CO) and carbon dioxide (CO_2). Road traffic is the main cause of NO_2 and CO. While simple environmental solutions such as timing traffic lights are identified as saving billions in fuel consumption and reducing air-pollution (i.e. improving air quality) by as much as 20%, the technological underpinnings to accomplish this, outlined in *Street CORNERS*, have not been developed and deployed on an appropriate scale for urban data gathering and correlation. Furthermore, the use of predictive models and tools, such as data mining, to identify patterns in support or opposed to environmental sustainability is not commonly done in an urban setting. While hurricane, earthquake, and other extreme weather events occur and the aftermath is dramatically presented, more mundane but not less impacting events such as urban flash flooding, chemical spills on city roads, and other environmental events are not anticipated or measured while occurring or developing, with intent to reduce in scope and damage. By gathering data locally, for assessment and prediction, the *Street CORNERS* team is able to identify areas and events that might harm environmental sustainability, and use the knowledge gathered to avoid or disable what might have previously an urban environmental disaster.

III. DATA MINING FOR TREND IDENTIFICATION

Data mining [3] can be referred to as 'knowledge discovery in databases', and is a key element of the *Street CORNERS* project. Of the four core data mining tasks:

- Cluster analysis
- Predictive modeling
- Anomaly detection
- Association analysis

both predictive modeling and anomaly detection are used in the *Street CORNERS* project.

"Predictive modeling" can be further defined by two types of tasks: classification, used for discrete target variables, and regression, which is used for continuous target variables.

Forecasting the future value of a variable, such as would be done in a model of an urban ecosystem, is a *regression task* of predictive modeling, as the values being measured and forecast are continuous-valued attributes. In both tasks of predictive modeling, the goal is to develop a model which minimizes the error between predicted and true values of the target variables. By doing so, the objective is to identify crucial thresholds that can be monitored and assessed in real-time so that any action or alert may be automatic and high responsive.

"Anomaly detection" is also crucial to the success of the *Street CORNERS* model. Formally stated *anomaly detection* is the task of identifying events or measured characteristics which are different from the rest of the data or the expected measurement. These anomalies are often the source of the understanding of rare or infrequent events. However, not all anomalies are critical events, meriting escalation and further investigation. A good anomaly detection mechanism must be able to detect non-normal events or measurements, and then validate such events as being outside of expectations – a high detection rate and low false alarm rate is desired, as these define the critical success rate of the application.

IV. SENSOR NETWORKS AND VISUALIZATION

Sensor networks have become part of our everyday lives and attract wide interest from industry due to their potential diversity of applications, with a strong expectation that outdoor and environmental uses will dominate the application space [4]. However, actual deployment experience is limited and application development has been further restricted [5]. Previous work in sensors for structure monitoring [6], urban flash flood awareness [7], and mobile emissions monitoring [8] has been initiated, but not to the extent and geographical contextual presentation illustrated by *Street CORNERS*,

The overall objective of the *Street CORNERS* network implementation is the gathering of environmental information in real-time and storing the data in a database so that the data can be visually presented in a geographic context for maximum understanding. Ideally, *Street CORNERS* is an implementation of a wireless environmental sensing network for urban ecosystem monitoring and environmental sustainability. By measuring environmental factors and storing the data for comparison with future data gathered, the changes in data measured over time can be assessed. Furthermore, if a change in one measured variable is detected, examination of another measured variable may be needed, to correlate the information and determine if the measured conditions are declining or advancing over time. Known as ‘exception mining’, this assessment can also be visually presented in a geographical context, for appropriate understanding and preventive or divertive action.

V. NETWORK DESIGN AND DEPLOYMENT

Kean University is situated outside of New York City, near main thoroughfares, and environmental monitoring was an ideal application for initial implementation. The urban environmental measurements obtained, part of the Street CORNERS network architecture [9], provide a good exercise in wireless sensor network design and deployment. In addition the urban setting, which includes traffic fumes and extreme weather events such as urban flash flooding, the Kean University campus has calmer locations, including a stream which cross-cuts campus and significant green space. This intersection of a wide variety of environmental systems provides a perfect incubator for *Street CORNERS*.

Several advantages of the wireless sensor network design include the reporting of GPS parameters by the sensors, providing positional coordinates. The ease of relocation of a wireless sensor and the adjustment of the network to the new reporting location of the relocated sensor is very significant. Previous generation network applications expected conventional, stationary network reports. However, this particular implementation and application, known as variational data assimilation, widely used in weather modeling, supports and encourages sensor mobility, for the greatest value in information gathered.

Standard wireless sensor network (WSN) implementations consist of a group of dispersed sensors or motes that have the expectation of covering a geographical area, referred to as the sensor field, in terms of some measured parameters. Sensor nodes are expected to have wireless communication capabilities and some associated logic. The type of sensor used here, a WSN with physical sensing of parameters such as temperature, light, or seismic events combined with computational and networking capabilities, is widely used.

For implementation purposes, the wireless sensor network kit from Crossbow Technology was selected, along with several support components. The objective of the deployment aspect of the project was to develop a wireless sensor network that would permit measurement of a variety of different environmental factors, such as temperature, light, and humidity. The Crossbow sensors operate on an xMesh network design, such that the resulting network is self-forming and self-healing. This was a significant advantage, as the deployed network has the ability to route traffic around unresponsive nodes without external intervention.

Currently, seven nodes are deployed, which provide an adaquate testbed. Future plans are to scale the network up. The sensors report new readings every 3 minutes at a data rate of approximately 2

Kbps. Each sampling provides updated voltage, humidity, pressure, and light readings.

VI. NETWORK APPLICATION DESIGN

The network design used was that of a system of distributed sensors, reporting to a base station, which was then connected to a server between the network and the application. The focus was on total application design, from data storage in the relational database, to final interpretation and presentation in the geographical context.

A. Visual Presentation

Application development included collected data, which was archived into a database. This was accomplished by using SQL, a relational database system. To clearly understand and visualize the importance, functionality and advantages provided by the wireless sensor network, the data must be clearly represented. In order to do so, a programming language or framework is needed that provides the ability to quickly gather data and accurately represent each of our sensors.

After consideration of availability, scalability, and recognition, as well as understanding the well-defined API and number of tutorials available, the Google Maps framework was selected. The Google Maps API allows the user to use Google Maps on individual websites, with JavaScript. In additional, a number of different utilities are available.

Google Maps provides an additional advantage, as the nodes represented on the map and the data contained in each one of the nodes can be set up using XML and additional nodes or markers can be added with ease. Once the decision to use Google Maps was made, the software development effort shifted from data representation efforts to working on accessing the actual, real time data from the wireless sensor network server. For Google Maps to process this new data and properly represent it, an XML file is generated. Figure 1 shows the exchange between the web server and the wireless sensor network server.

As data is sent in from the nodes, it is passed through the base station to the Perl XML Parser, which parses the incoming data and filters out unwanted packets. The result remaining is the desired data packet set.

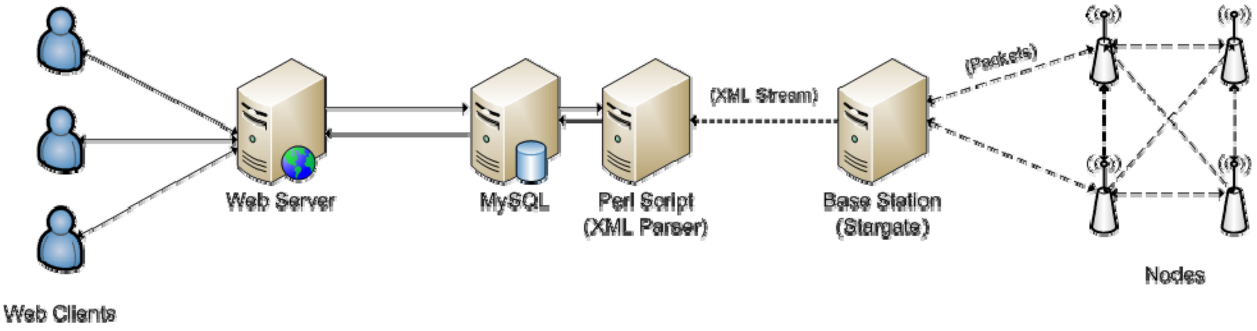

Figure 1. Request-response data exchange between web server and wireless sensor network server.

A prototype application has been developed. The wireless sensor network data is gathered in a database, which is then presented in a visual context. Google maps is used to present the information. By clicking on the sensor node (marked by a red 'pin head' on the Google map), real-time information is be presented to the viewer, in a geographical context. The raw data, collected from the sensor, has been appropriately converted to standard units for display and understanding.

In addition to real-time data presentation in a geographical context, a temporal presentation, using time and date information, was also developed. An illustration of this can be seen in Figure 2. A query by sensor presents the specific details of that sensor at one point in time, as well as providing a comparison of the sensor's status for prior dates and times. More than one sensor measurement can be overlaid on the chart, which permits correlation of events and times precisely with sensors.

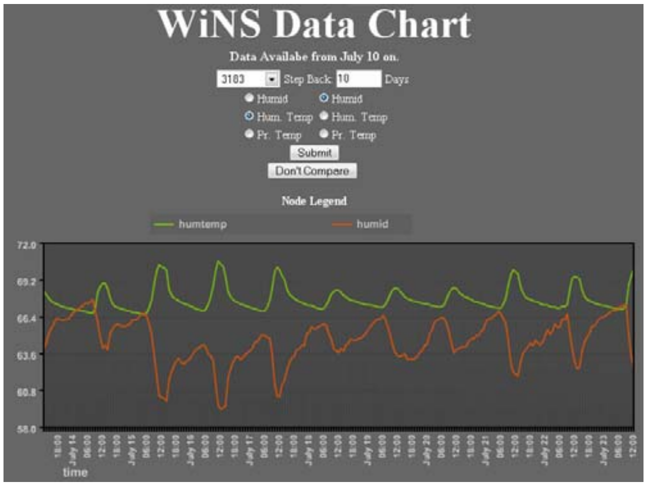

Figure 2. Visual presentation of sensor data with date.

While data has been gathered by sensors before, the correlation and presentation of this environmental real-time information in a geographical context, in additional to providing support for historical and temporal comparison, is innovative.

B. Integration of Geovisualization and Data Mining

The presentation of data in real-time for contextual understanding is only one aspect of the *Street CORNERS* application. The archived information in the database permits before and after animations to be developed, using time gradients to show how the measured variables have changed in the preceding time, or are expected to change in the future, based on predictive algorithms, taking into account the reported variables, such as wind speed, in the case of a chemical spill and the potential migration over an urban community, for example.

Once the data of the multiple network nodes has been collected for a long period of time, it is desirable to reveal patterns, if any, in the data. For example, with regard to spatial-temporal variations of air quality along the streets or at intersections, patterns regarding daily and seasonal cycles of air quality, change, and decay over distances from the intersection can be directly visualized through an interactive and animated visualization environment. Furthermore, the visualization tool can be used together with numerical data mining algorithms to model quantitative relationships between air quality and other factors such as weather and traffic conditions. Numerical data mining algorithms such as supervised learning can be easily integrated here. The findings can then be applied to simulate the spatial-temporal air quality variations given arbitrary weather and traffic conditions for the purpose of predictive modeling.

Human visual perception offers a broadband channel for information flow and excellent pattern recognition capabilities that facilitate knowledge discover and the detection of spatial-temporal relations [10]. An effective tool to explore geographic data and communicate geographic information to private or public audiences, geo-visualization has long been used for data exploration and pattern recognition. The approach presented and discussed here integrates geo-visualization with data mining to reveal spatial-temporal patterns embedded in the data collected by the sensor network over time. While prior work [11] has approached such an idea, the work presented here is the first to visually present the data correlations.

VII. CONCLUSIONS AND FUTURE RESEARCH

Wireless sensor network applications are an emerging area of technology which will benefit organizations and governments with valuable real-time data. In order to properly use such data, a strong, dynamic and user-friendly interface is needed which allows individuals to clearly see how measured conditions, such as environmental circumstances, are changing over time. The visual depiction of urban environmental events, for example, will permit anticipatory or preventive actions to be taken in advance of adverse human and ecological impact. The use of data mining techniques on the data gathered by the wireless sensor network permits the identification of past patterns and developing trends in air quality or urban flooding, for example. The network and interface illustrated here accomplishes the goals of real-time information gathering and display for environmental sustainability and further work is underway to improve and refine the solution presented here.

Additional research underway includes a case study where a number of exploratory spatial data analysis (ESDA) techniques will be tested to facilitate the visual detection of spatial-temporal patterns of air quality in relation to weather conditions. ESDA techniques being tested in the case study include temporal brushing [12] and temporal focusing [13], temporal re-expression through multi-scale data aggregation [14] and static visual bench marking [15]. Animation of the temporal data will enable common users to visualize the change of air quality over space and time. With temporal brushing and focusing, the use is not only a passive viewer of the information, but can interact with the animation and learn actively. Temporal re-expression through multi-scale data aggregation provides an opportunity to directly visualize the daily and seasonal cycles of air quality change. Finally, using static visual benchmarking, the air quality level at any recorded time spot can be compared visually with health standards and give the viewer a direct alert on how high air quality is affecting human health. Efforts continue to integrate the geo-visualization environment with a knowledge discovery procedure for data mining.

REFERENCES

[1] Aldrin, M. and Hogaek-Haf, I., “Generalized additive modeling of air pollution, traffic, volume, and meteorology”, *Atmospheric Environment*, Vol. 39, No. 11, April 2005, pp. 2145-2155.

[2] Ashley, R.M., Balmforth, D.J., Saul, A.J., Blanksby, “Flooding in the future – predicting climate change, risks, and responses in urban areas”, *Water Science and Technology*, Vol. 52, No, 5, 2005, pp. 265-274.

[3] Thuraisingham, B., Khan, L., Clifton, C., Maurer, J. and Ceruti, M., “Dependable Real-time Data Mining”, *Proceedings of the 8th IEEE International Symposium on Object-Oriented Real-Time Distributed Computing (ISORC’05)*, pp. 158-165.

[4] Martinez, K., Hart, J.K., and Ong, R. “Environmental Sensor Networks”, *IEEE Computer*, August 2004, pp. 50-56.

[5] Lewis, F.L., “Wireless Sensor Networks”, *Smart Environments: Technologies, Protocols, and Applications*, D. J. Cooke and S.K. Das, editors, John Wiley, New York, 2004.

[6] Zimmerman, A.T., Lynch, J.P., “Data Driven Model Updating using Wireless Sensor Networks”, *Proceedings of the 3rd Annual ANCRiSST Workshop*, May 2006.

[7] Chang, N., and Guo, D., “Urban Flash Flood Monitoring, Mapping, and Forecasting via a Tailored Sensor Network System, Proceedings of the 2006 IEEE International Conference on Networking, Sensing and Control, April 2006, pp. 757-761.

[8] Cordova-Lopez, L.E., Mason, A., Cullen, J.D., Shaw, A., and Al-Shamma’a, A.I., “Online vehicle and atmospheric pollution monitoring using GIA and wireless sensor networks”, *Journal of Physics: Conference Series*, Volume 76, Issue 1, 2007.

[9] Morreale, P.A. *“Street Corners:* An Architecture for Correlation of Real-time Networked Environmental Sensors”, *Proceedings of the 2nd IEEE Systems Conference*, Montreal, Quebec, Canada, April 2008, pp. 387-392.

[10] Gahegan, M., Wachowicz, M., Harrower, M., and Rhyne, T-M., “The Integration of geographic visualization with knowledge discovery in databases and geocomputation”, *Cartography and Geographic Information Science*, 28(1): 29-44, 2001.

[11] Arici, T., Akgu, T., and Altunbasak, Y., “A Prediction Error-Based Hypothesis Testing Method for Sensor Data Acquisition”, *ACM Transactions on Sensor Networks*, Vol. 2, No. 4, November 2006, pages 529-556.

[12] Monmonier, M., “Geographic brushing: Enhancing exploratory analysis of the scatter plot matrix”, *Geographical Analysis,* 21(1): 81-84, 1989.

[13] MacEachren, A. M., Polsky, C., Haug, D., Brown, D., Boscoe, F., Beedasy, J., Pickle, L., and Marrara, M., “Visualizing spatial relationships among health, environmental, and demographic statistics: interface design issues”, *18th International Cartographic Conference* Stockholm, June 23-27, 1997, pp. 880-887.

[14] Monmonier, M, “Strategies for the visualization of geographic time-series data”, *Cartographica* 27(1): 30-45, 1990.

[15] Harrower, M., *Visual Benchmarks: Representing Geographic Change with Map Animation.* Ph.D. dissertation, Pennsylvania State University, 2002.