Temperature Control Framework Using Wireless Sensor Networks and Geostatistical Analysis for Total Spatial Awareness

Charles Castello and Jeffrey Fan  
Department of ECE  
Florida International University  
Miami, FL, USA

Asad Davari  
Department of ECE  
WVU Institute of Technology  
Montgomery, WV, USA

Ruei-Xi Chen  
Dept. of Comp. Sci. and Info. Eng.  
St. John’s University  
Taipei, Taiwan

Abstract — This paper presents a novel framework for intelligent temperature control in smart homes using Wireless Sensor Networks (WSN) and geostatistical analysis for total spatial awareness. Sampled temperature readings from sensor nodes have the ability to inform the system on temperatures at specific locations. However, these locations where readings are taken do not completely describe the area under consideration. To solve this issue, geostatistical techniques are utilized, which include variography and kriging to predict temperature where measurements are not available. This added information would allow the system to control cooling and heating mechanisms in buildings at every location for improved user comfort.

Keywords – Smart Home; Intelligent Temperature Control; Wireless Sensor Network; WSN; Context Awareness; Geostatistical Analysis; Classical Variography; Ordinary Point Kriging

I. INTRODUCTION

Research involving smart homes has flourished in recent years due to advancements in computational resources including processing power, system memory, and transmission bandwidth. These improvements allow for state-of-the-art electronic devices to be seamlessly integrated into homes for unparalleled comfort, control, and energy efficiency. In order to see these benefits come to fruition, there are three elements that must be considered: (1) internal positioning, (2) intelligent control, and (3) home automation [1]. Home automation is the use of products to link services and systems outside the home. These product types include environmental, security, home entertainment, domestic appliances, information and communication, and health [1]. This paper will focus on environmental services, mainly developing a novel framework for intelligent temperature control in smart homes using geostatistical analysis. This work will add to knowledge in context awareness modeling for smart home services.

There has been many quality research work [2-4] done on smart homes, but improvements are needed, particularly in context awareness modeling for smart home services. This paper seeks to accomplish this by developing a novel framework for intelligent temperature control using geostatistical analysis for total spatial awareness. Described in Section II of this paper is previous research in temperature control, where decision-making is based primarily on temperature data from a handful of sensors, not taking into account locations lacking temperature readings. This gives an inaccurate depiction of the area being monitored, thereby controlling temperature of the home in an inefficient manner. In order to improve monitoring techniques, geostatistical analysis is used, particularly variography to determine the spatial dependency of referenced observations and kriging, to interpolate across space for predicted values between observed locations. This information is used in the proposed framework to aid the temperature controller in making more informed decisions for user comfort.

The remainder of this research paper is organized as follows: Section II reviews previous research pertaining to smart homes, temperature control, and context awareness techniques. Section III reviews geostatistical analysis methodologies being used in the proposed framework, Section IV describes the proposed framework and analytical algorithm, Section V discusses the experimental dataset, Section VI examines generated results, and Section VII concludes with final thoughts and future work.

II. PREVIOUS RESEARCH

Past research in smart homes covers a wide range of topics including control systems, platform and architecture, middleware, security, and context awareness. Research in [5] used Laboratory Virtual Instrument Engineering Workbench (LabVIEWTM) and Personal Digital Assistants (PDA) to control and monitor smart home services including elevator positioning, power, gas, light, water temperatures, and ambient temperatures. A threshold is set by the user for ambient temperature using a PDA where temperature is monitored with a single sensor. Work accomplished in [6] developed a cost-effective home automation system having the capability to adjust security, cooling, heating, and lighting services for added comfort, energy conservation, and user convenience. The SmartHome unit developed was based on the Rabbit 3000 microprocessor, particularly, the RMC3700 RabbitCore designed for Internet applications. The temperature control of this unit utilized sensors with user-defined programs. Work in [7] described a lightweight, user-controlled smart home system geared toward the elderly living for longer periods of time at home. This was done by concentrating on three key topics: privacy, control of personal space, and enjoyment within the home. The developed system was composed primarily of a multimodal robot, for use as a personal assistant, a mobile assistant (PDA...
or a smart phone), and an interactive TV, which were both utilized for control and monitoring. The multimodal robot has the ability to travel throughout the home to fulfill various tasks, such as measuring temperature and checking whether lights are on or off. Control of the multimodal robot is either through voice commands, the mobile assistant or TV. Cooling and heating temperatures can also be monitored or changed through the mobile assistant or TV. A recurring characteristic of research just mentioned is the lack of intelligent temperature control, where the majority base decisions primarily on a single data sample, giving an inaccurate depiction of the environment in question. To improve system awareness, geostatistical methods are used in the proposed temperature control framework described in Section IV. To the best knowledge of the authors, there are no other temperature control schemes in smart home design using geostatistical analysis methods for improved spatial awareness.

III. GEOSTATISTICAL ANALYSIS

There have been many studies in environmental sciences that use geostatistical analysis to help understand a substance or a parameter’s concentrations on a spatial plane. This type of analysis could provide valuable information for improved scientific understanding, risk assessment of contaminants or pollutants, and decision support. Examples of such work include using geostatistics to assess mercury in soils around a coal-fired power plant in Baoji, China [8], analyzing spatial correlation of nitrogen dioxide (NO2) concentrations in Milan, Italy [9], studying spatial distribution of soil lead in the mining site of Silvermines, Ireland [10], and development of a spectrophotometer in generating detailed soil maps for rice paddy fields using geostatistical analysis [11].

The geostatistical methods used in this paper will now be described which include variography and kriging, briefly summarized from [12]. The theoretical basis of both techniques can be found in greater detail in [13-15].

A. Classical Variogram

In geostatistical analysis, the variogram is used to describe the spatial dependency between referenced observations within the analyzed plane where the true variogram is generally unknown. Therefore, an estimation of the variogram can be calculated through known observations. This is accomplished by first determining the experimental variogram, given by the semivariance:

\[ \gamma(h) = 0.5 \cdot (z_x - z_{x+h})^2 \]  

where \( z_x \) is the observed value at point \( x \) and \( z_{x+h} \) is an observed value at another point within distance \( h \). \( \gamma(h) \) is also known as a semivariogram or variogram and the distance between observations is known as lag distance.

The next step is to summarize the experimental variogram with a variogram estimator, which determines the central tendency and is similar to descriptive statistics derived from univariate observations [12]. The variogram estimator is calculated by Equation (2) where \( N(h) \) represents the amount of pairs within lag interval \( h \).

\[ \gamma_e(h) = \frac{1}{2 \cdot N(h)} \sum_{i=1}^{N(h)} (z_{x_i} - z_{x+h})^2 \]  

Lastly, a variogram model or parametric curve is fitted to the variogram estimator. This is similar to frequency distribution fitting where the frequency distribution is modeled by a distribution type and its parameters [12]. The three commonly used models in variography are: spherical, exponential, and linear.

B. Ordinary Point Kriging

In order to interpolate observations on a regular grid, ordinary point kriging is used, which is considered the most commonly utilized method. Ordinary point kriging uses weighted averages and neighboring observations to predict unobserved points:

\[ \hat{z}_m = \sum_{i=1}^{N} \lambda_i \cdot z_i \]  

where \( \lambda_i \) are estimated weight values. In order to guarantee that estimates are unbiased, the sum should equal one.

\[ \sum_{i=1}^{N} \lambda_i = 1 \]  

Equation (5) shows the average estimation error must equal zero where \( z_{i0} \) is the unknown value.

\[ E(\hat{z}_{i0} - z_{i0}) = 0 \]  

Calculating the mean-square error using Equations (3) through (5) in terms of the variogram and using a Lagrange multiplier \( \nu \) for optimization yields a linear kriging system of \( N+1 \) equations and \( N+1 \) unknowns which is calculated by:

\[ \sum_{i=1}^{N} \lambda_i \cdot \gamma(x_i, x_j) + \nu = \gamma(x_i, x_j) \]  

where \( \lambda_i \) is the weight for the \( i^{th} \) data point, \( \gamma(x_i, x_j) \) is the variogram between data points \( x_i \) and \( x_j \), and \( \gamma(x_i, x_j) \) is the variogram between the data point and unobserved point. The kriging variance is calculated by:

\[ \sigma^2(x_0) = \sum_{i=1}^{N} \lambda_i \cdot \gamma(x_i, x_0) + \nu(x_0) \]  

IV. TEMPERATURE CONTROL FRAMEWORK

The temperature control framework will now be described, which is shown in Figure 1. In Step 1, data is sampled throughout the WSN nodes at a user-defined rate.
This data must first undergo an initial training sequence in Step 2 where the appropriate distribution model is determined for the variogram model. It is assumed the selected distribution model along with sill and range values are acceptable for subsequent sets of data samples being analyzed. Once the variogram model has been determined, kriging analysis is executed in Step 3 to interpolate observations for the monitored area. Results from kriging are then divided into user-defined sub-regions in Step 4 where each sub-region must possess a heating or cooling element with the ability to be controlled separately from other elements in the system. The mean is then calculated for each sub-region. The calculated mean for each is then compared with user-defined thresholds. If the calculated mean is less than the lower threshold, $T_{\text{low}}$, or greater than the upper threshold, $T_{\text{high}}$, a signal is sent to the heating or cooling system respectively to engage the sub-region in question. This process minus the initial training step repeats at a user-defined basis (e.g., 10 minutes).

V. EXPERIMENTAL DATASET

The experimental dataset that is used to test the proposed temperature control framework for smart homes is from Intel Berkeley Research Lab [16]. Collected data consists of temperature, humidity, light, and voltage for 54 sensor nodes in a WSN. Samples were collected every 31 seconds between February 28th and April 5th, 2004, totaling 2.3 million readings. Locations are also given for each node relative to the upper right corner of the lab, which is shown in Figure 3. Temperature data will be focused upon in this research paper.

VI. RESULTS

The first dataset used to analyze the proposed framework consists of a data sample from each of the 54 sensors minus samples from nodes 5 and 20 due to irregular readings. Samples were taken approximately at 01:10:00am on February 28th, 2004. Due to the use of linear kriging techniques, temperature samples should be Gaussian in nature, which is confirmed in the histogram shown in Figure 4.

The experimental variogram and lag values are then calculated for the initial dataset and plotted in Figure 5. Analysis of the variogram cloud conveys the dispersion of variogram values at different lags. A variogram model is fitted to the variogram estimator in Figure 6 where linear, exponential, and spherical models are fitted. By visual inspection, the spherical model was chosen for kriging.
analysis where the sill is set to 0.73 and range to 13. Kriging results for the initial dataset from Intel Berkeley Research Lab are shown in Figure 7 where predictions are calculated for locations without sensor nodes. Figure 8 displays the sensor node locations along with variance values for the entire grid. It is shown that variance tends to be lower at or near sensor node locations rather than predicted locations. This demonstrates variance being a measure of information density [17]. It should be noted the nugget effect was not taken into consideration during this study. Lastly, temperature trends are plotted in Figure 9 for data from Intel Berkeley Research Lab every 10 minutes from 01:10:00am to 04:20:00 a.m. on February 4th, 2004 minus samples from nodes 5 and 20. The mean and standard deviation values were taken from the initial dataset analysis where \( \mu - \sigma = 17.5749 \) is defined as \( T_{low} \), and \( \mu + \sigma = 19.2633 \) is defined as \( T_{high} \). \( \mu \) and \( \sigma \) values were used to define \( T_{low} \) and \( T_{high} \) values in this study, although other user-defined amounts are acceptable.

Temperature trends for each sub-region, defined in Figure 7, are shown in the graph. Following the temperature control framework of Figure 1, the binary control variable \( e_{4h} \) would change from a ‘0’ to ‘1’ value to produce heat for sub-region 4 at 01:20:00am. The same would occur for sub-region 1 at 02:40:00am, sub-regions 2 and 8 at 02:50:00am, sub-regions 6 and 5 at 03:00:00am, and sub-regions 3 and 7 at 03:50:00am. Heating would continue for each sub-region until the averaged temperature reaches \( \mu = 18.4191 \).

VII. CONCLUSION

A novel temperature framework for smart homes has been designed and developed using geostatistical analysis tools, particularly variography and kriging. This allows the monitored area to be described in greater detail for improved temperature control, thereby improving user comfort. This
work adds to knowledge in context awareness modelling for smart home services. To the best knowledge of the authors, there are no other temperature control schemes in smart home design using geostatistical analysis methods for improved spatial awareness.

An important observation in the discussed temperature framework that should be mentioned involves the user-defined time for sampling. This is an extremely important parameter because if set too high, temperature can travel from below $T_{low}$ to above $T_{high}$ values or vice versa in less than a system cycle. If the sampling time is set too low, unnecessary analysis and system resources are wasted. Although this issue is beyond the scope of this paper, further analysis would be beneficial in determining the most optimal values for sampling times.

Future work involving the discussed temperature control framework includes further research and analysis of user-defined sampling rates. Other research efforts can be invested in the possible affect on power consumption and accuracy analysis of kriging methods. Log-normal distributions using log-normal kriging can also be studied, which is not uncommon in spatial analysis. Lastly, further research is possible in improving context awareness for various aspects in smart homes using spatial analysis besides temperature control.

REFERENCES


