Spatial Interpolation for a Wireless Sensor Network of Chemical Point Detectors: Preliminary Results

Robert J. Schafer

robert.schafer1@us.army.mil

SAIC (US Army Edgewood Chem/Bio Center Operations)
PO Box 68, Gunpowder Branch, Aberdeen Proving Ground, MD 21010, USA

Robert J. Hammell II

rhammell@towson.edu

Department of Computer & Information Sciences, Towson University
Towson, MD 21252, USA

Abstract

The US Army Edgewood Chemical/Biological Center has developed the Spectrometric Point Ionizing Detector ARray, a networked array of small, rugged ion mobility spectrometers and meteorological sensors. The Spectrometric Point Ionizing Detector ARray was developed to provide real-time ground truth measurements of chemical vapor cloud releases during acceptance testing of chemical analysis devices. A point sampling array of chemical detectors is an essential tool for chemical vapor sensing in an outdoor environment, as it provides higher sensitivity and selectivity than stand-off detectors, and has the ability to sample much closer to the ground than stand-off detectors. However, the data acquired from a wireless network of point sensors like the Spectrometric Point Ionizing Detector ARray is spatially limited to the specific locations of the sensors. Spatial interpolation of point sampling wireless sensor networks is a common problem that has seen recent activity in the literature. However, most interpolation techniques do not consider the effect of underlying environmental phenomena, such as micrometeorological conditions, upon their results. It has been demonstrated that such phenomena can have a significant effect on the behavior of a vapor cloud; therefore, it is desirable to have an interpolation model that can account for local meteorological conditions at each chemical point detector.

Keywords: spatial interpolation, wireless sensor networks, chemical vapor point detection, point sampling sensor networks, chemical detector array

1. BACKGROUND

Acceptance testing of chemical analysis systems frequently includes large scale outdoor field testing to verify that a system behaves properly in a setting outside of the laboratory. These outdoor field tests may challenge the system's ability to respond to a chemical release qualitatively and/or quantitatively. Ground truth instrumentation is required in such cases to provide a referee

against which the system in question is tested. The US Army Edgewood Chemical/Biological Center (ECBC) has developed the Spectrometric Point Ionizing Detector ARray (SPIDAR) as part of the Range Test Validation System (RTVS). SPIDAR is a wireless sensor network (WSN) of chemical point detectors and meteorological sensors connected via TCP/IP to a command and control software system called SPIDARweb. SPIDAR was developed to provide real-time ground truth measurements of chemical vapor cloud releases during acceptance testing of chemical analysis devices (Schafer, 2008).

SPIDARWeb, the software portion of SPIDAR, is an object oriented, event driven system developed in C# for Microsoft Windows XP. SPIDARWeb provides an integrated platform for monitoring all SPIDAR Stations in real time, as well as archiving all data from the SPIDAR array. It also provides real-time semi-quantitative concentration analysis for compounds which exist in its library. Concentration reports are generated based on extensive laboratory regression analysis, as detailed in Harden et al. (2008). SPIDARweb also provides a server module. Clients to this server can be simple data viewers to allow multiple views of an experiment in progress, or they can provide the SPIDAR data to external systems such as data fusion systems for combining point detector data with stand-off detector data (Schafer, 2008).

An example screenshot of SPIDARweb playback is presented in figure 1. Several features of the interface are visible in the figure, including playback controls, playback time display, the comments box, the SPIDAR station list, the map display, a typical SPIDAR station display window, and the color scale. The comment box shows any timestamped comments recorded by an operator during data recording. The color scale on the left side of the interface gives the user the chemical concentration range in milligrams per cubic meter (mg/m^3) for each icon color. The SPIDAR station list, just to the right of the color scale, shows all of the stations, with icon colors corresponding to the color bar. Each SPIDAR station item in the list can be expanded to show details for that station, including chemical detector

status, global positioning system (GPS) coordinates, and meteorological data. On the right side of the interface in figure 1 is the map display. This is a geographical representation of the SPIDAR stations. The icon colors correspond to the color scale. The green and black arrows point into the wind, as measured by the meteorological sensor at the station. Near the center of the interface is an example station display, in this case for station NE1. This window gives details about the status of the station, including meteorological conditions, temperature and pressure readings internal to the chemical detector, GPS coordinates, the time of the last reported chemical detector data, any chemicals currently detected, and the raw ion mobility spectrum. The spectrum displayed here is typical of methyl salicylate. In the upper right hand of the station window is a rocker switch that toggles the spectrum display between positive and negative mode. Certain classes of chemicals form positive ions while others form negative ions, and this switch allows the user to view either.
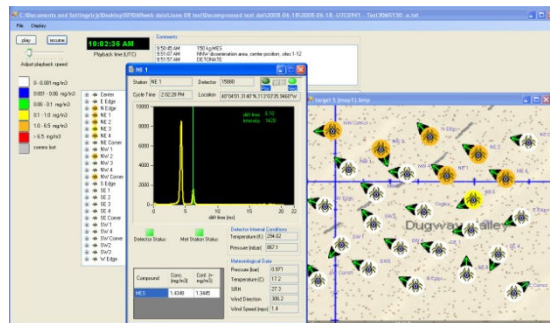

Figure 1 - SPIDARWeb playback screen-shot displaying ion mobility spectrum and geographic map display.

One scenario for conducting acceptance testing in an outdoor setting is to define a "truth box" across which the chemical vapor concentration of the analyte in question is known to reasonable degree of confidence at any point. Typically, between 12 and 25 chemical point detectors are deployed across the truth box. Because the truth box covers a relatively large amount of space, for instance 1 km^2, there will be wide areas of space between the detector stations.

A   B

Figure 2 - SPIDAR Station hardware. A. The LCD-3 chemical agent detector, Smiths Detection. B. The Airmar Weatherstation PB100/LB100.

The chemical detection technology employed by SPIDAR is Ion Mobility Spectrometry (IMS). An ion mobility spectrometer is an analytical device that identifies chemical vapors by ionizing the compound and then separating the product ions based on their size to charge ratio. Raw IMS data is in the form of a pair of ion mobility spectra, one showing negatively charged ion signatures (negative mode), and the other showing positively charged ion signatures (positive mode). Each spectrum is a plot of drift time in milliseconds versus amplitude of the detected signal in arbitrary units, where drift time is a measure of the amount of time required for a signal producing ion to pass through the IMS instrument's drift tube. The Lightweight Chemical Detector LCD-3 (specifically version LCD-3.2e), manufactured by Smiths Detection, was selected as the chemical detector for SPIDAR (see Figure 2A). IMS devices have a long history of use for chemical detection, and the LCD-3 has been selected as the Joint Chemical Agent Detector (JCAD) for the US military. It is a small device (less than 1 pound) and provides data output every 5 seconds. The operating temperature range is -30°C to +40°C, which is important considering that outdoor field tests frequently take place in harsh environments. Details relating to the calibration of these detectors for use in SPIDAR are related by Harden et al. (2008).

Each SPIDAR field station also includes a meteorological sensor, The Airmar WeatherStation LB100/PB100, manufactured by Airmar Technologies Corporation (see Figure 2B). The WeatherStation is an ultrasonic solid state device that provides measurements for several meteorological parameters, including air temperature, wind speed, wind direction, barometric pressure, and relative humidity. In addition, it has an integrated Wide Area Augmentation System (WAAS) compliant Global Positioning System (GPS) receiver. Selected meteorological specifications are as follows: wind speed range: 0.6 – 114 mph ± 0.1 mph; wind direction resolution: 1°; air temperature range: -30°C to +50°C ± 1.5°C; relative humidity: 10 – 95% ± 5%; barometric pressure: 850 – 1150 millibar ± 1.5% (WeatherStation Owner's Guide, 2006).

In order to connect each SPIDAR Station's serial devices to the TCP/IP network, a serial port server device is required. The port server provides TCP/IP network access to remote serial devices. Each SPIDAR Station is equipped with a PortServer TS-2-H-MEI port server, from Digi International, Inc. The TS-2-H-MEI provides two RS-232/422/485 serial ports, and is hardened for operation in harsh environments. The chemical point detector and the meteorological sensor at each station are connected to the TCP/IP network through the TS-2-H-MEI. Virtual serial port software running on the base station host allows the SPIDARweb software to connect to the remote serial devices by way of the port servers as if they were physical serial ports of the host.

For the version of SPIDAR described in this paper, an Engenius EOC-3220+ wireless client bridge with integrated 16 dBi patch antenna was deployed at each SPIDAR Station. This is a high powered (up to 24 dBm at 11 Mbps) 802.11 b/g device. As with all 802.11 b/g radios, it operates in the 2.4 GHz range. SPIDAR has been tested and operated using this wireless bridge, as well as similar (but lower powered) devices from other manufacturers. At the base station, an Engenius EOC-3220-EXT bridge configured to access point mode was deployed. The EOC-3220-EXT is similar to the EoC-3220+, except that it does not contain an integrated antenna. An external 20 dBi 120 degree

sector panel antenna was used with the access point.

Since IMS is a point sampling technology, the array provides data only for the points in space at which a detector is located. There is no information available to a user about the chemical vapor concentration profile for the majority of the test area. However, it is important to have a sense of the concentration profile across the entire test grid, in order to provide full ground truth for acceptance testing. Although data from stand-off detectors, such as infrared spectrometers, can be used to supplement point detection data, these stand-off devices typically do not provide information about the vapor cloud's concentration profile close to ground level. Infrared spectrometers are essentially cameras that detect heat signatures, and rely on temperature differentials to provide information about vapors that they are sensing from some distance. These devices have difficulty detecting vapors close to the ground due to the heat of the ground itself. Also, since a stand-off detector looks through space toward the horizon, it cannot provide a chemical concentration value for a single point in space, but rather provides an integrated chemical density over distance. Therefore, in order to make the truth box more useful at ground level and at specific points, it is advantageous to have an algorithm or algorithms that may be leveraged to provide an estimation of the chemical profile at points in the truth box between point detector stations. This provides an operator with more information about the vapor cloud's profile, and allows the truth box to more successfully provide ground truth data for acceptance testing.

2. DATA INTERPOLATION

Data interpolation methods date back to at least Hipparchus of Rhodes (190-120 BC), who used linear interpolation to construct tables of the chord function for computing the positions of celestial bodies (Meijering, 2002). Interpolation has been applied, in recent years, to wireless sensor networks (WSN), usually with the goal of increasing network efficiency and power savings. These techniques can also be used to improve sensitivity or coverage of the WSN.

There are two broad and distinct classes of applications for sensor networks: event detection networks and continuous spatiotemporal sampling networks. The goal of an event detection network is to notify when a particular event takes place. Each node has a sensing range, which determines the distance over which an event can be detected. Reliable network coverage is obtained by ensuring that each point in space falls into the sensing range of a minimum number of sensors. In contrast, a spatiotemporal sampling network consists of sensors each reading a single sample in time and space. The goal of this network is to create a continuous map of the physical phenomenon by interpolation the values between sensor readings. There is no concept in this case of sensing range. For this reason, techniques that leverage sensing redundancy are not applicable to spatiotemporal sampling networks (Liaskovits, 2007). It is clear that a sensor network of chemical point detection devices falls into the category of spatiotemporal sampling networks.

Liaskovits and Schurgers (2007) describe using a finite-dimensional Hilbert space for managing sensing topology. The Hilbert space allows both the sensor network topology and the underlying physical process to be modeled in the same framework. Sensor locations map onto vectors in the Hilbert space, and the inner products between vectors are defined by the correlation structure of the sensed physical process. This representation allows the physical phenomenon to be reconstructed to a set accuracy without a priori knowledge of the physical process. The key aspect of Liaskovits and Schurgers' work is that the assumption can be made that point reconstruction is linear between the measured values; this allows the spatial map of the phenomenon to be reconstructed from the sensor data. This assumption is widely accepted (Vuran, 2006) (Perillo, 2004) (Guestrin, 2004) (Krause, 2006) (Tynan, 2005).

Spatiotemporal interpolation can considered as two separate but related concepts: spatial interpolation and temporal interpolation. Spatial interpolation provides an estimate of values between two sensor nodes in space. Temporal interpolation provides an estimate of the value at a particular node at an in-

stant in time, even if the node was not sensing at that time. This requires that the node must have had at least one previous and one subsequent measurement. Interpolation in a wireless sensor network can be conducted using a linear or cubic weighted average algorithm. Alternatively, other functions such as B-spline, cubic spline, or kriging functions, could be used in place of the weighted average (Tynan, 2005). In order to calculate the interpolated error for each sensor, the actual observed sensor value is compared to the value that would be interpolated at that point if the sensor is ignored. The average of these errors gives the average error for the function. This is known as the leave-one-out validation technique (Tynan, 2005).

According to Umer et al. (2008), coverage strategies for wireless sensor networks that rely solely on the deployment of additional nodes are frequently unrealistic, particularly in hostile environments or when costs are prohibitive. An accurate representation of the physical phenomenon which the network is sensing requires augmentation of the reported data through interpolation. Typical interpolation techniques are not readily available to wireless sensor networks due to the reliance of these techniques on global knowledge of the network, which is prohibitively expensive to obtain due to the large scale of such networks. However, in the case of the SPIDAR system, global spatial knowledge will be available because the interpolation will occur at the base station, which does have full knowledge of the network and access to all data.

3. APPROACH

Two interpolation algorithms were examined during the preliminary experimentation phase: a linear method and a cubic method. These methods are options for the 'griddata' function found in Matlab R2009a. Both methods use the Quickhull algorithm to perform Delaunay triangulation of the data (Matlab 7 Function Reference, 2009). The Delaunay triangulation of a set of points is defined as a set of triangles drawn connecting the points to their neighbors such that the circle circumscribing the vertices of any one triangle does not enclose the vertices of any other triangle (Daintith, 2004). The results

from these early experiments will be considered the baseline for continuing research on alternative interpolation algorithms. In addition, the baseline determined here will be used to develop interpolation techniques which consider the underlying micrometeorological conditions at each detector site as an additional input.

The LCD-3 detector at each SPIDAR site generates ion mobility data approximately once every five seconds. The SPIDARweb software system receives this data as it is issued. SPIDARweb normalizes the spectra and calculates the reduced mobility (K_0) of each peak. Reduced mobility is derived from the drift time of the peak, and factors out such environmental variables as temperature, pressure, and detector voltage. Reduced mobility values of the detected peaks are compared to a library of known peak signatures, and the normalized peak intensities are used to calculate the chemical concentration based on regression analysis performed in the laboratory (Harden, 2008). Since no synchronization between SPIDAR sites is performed, concentration reports for each SPIDAR site were quantized to 5 second intervals prior to spatial interpolation.

Data recorded at Dugway Proving Ground during the Range Test Validation System (RTVS) event in June 2008 was used to evaluate interpolation methods. This data set includes a series of outdoor, open air releases of chemical warfare agent simulants, specifically triethyl phosphate (TEPO) and methyl salicylate (MES). For each test event, a known quantity of one of the simulants (or in some cases, both simulants together) was released either explosively or through a vapor stack. The SPIDAR WSN employed during these tests was comprised of 25 sensor sites arranged across a 1 km^2 "truth box", as shown in figure 3. Truth box coordinates for each site were obtained by converting GPS coordinates reported by each site's meteorological sensor into Universal Transverse Mercator (UTM) coordinates, and then normalizing to an arbitrary 1 km scale, with grid center defined as (0,0).

For each 5 second timeframe, a table was created which contained each of the 25 sensor site's coordinates and the reported con-

centration of chemical simulant in mg/m^3. Spatial interpolation of concentration across the truth box was performed to a resolution of 10 meters for each timeframe.

Interpolation results were compared using the leave-one-out validation technique. For each data set, interpolation is performed using all but one data point. The error is calculated at the missing point by comparing the interpolated value to the true value. This is repeated for each data point in the data set. The interpolation error is then defined as the average of the errors for the leave-one-out interpolations.
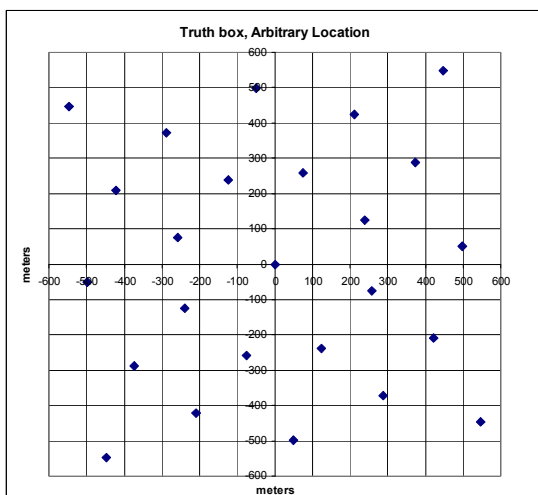

Figure 3 - Truth box, 1 km^2. Note that the sides of the truth box are not parallel with the axes. This is because the truth box was aligned with the prevailing winds during field tests.

4. PRELIMINARY RESULTS

Approximately one minute of data was extracted from a SPIDAR data file. This file was recorded during an explosive dissemination of 150 kg of methyl salicylate (oil of wintergreen, or MES). This one minute of data was quantized to 5 seconds intervals or frames, and spatial interpolation of MES concentration data was performed for each of these time frames with both the linear method and the cubic method. Figure 4 presents the mean interpolation error for these interpolations versus time, as calculated using the leave-one-out cross valida-

tion technique. From these results, it is shown the cubic interpolation method provides better interpolation (the cubic mean errors are closer to zero) than the linear method in approximately 60 per cent of the time frames examined.
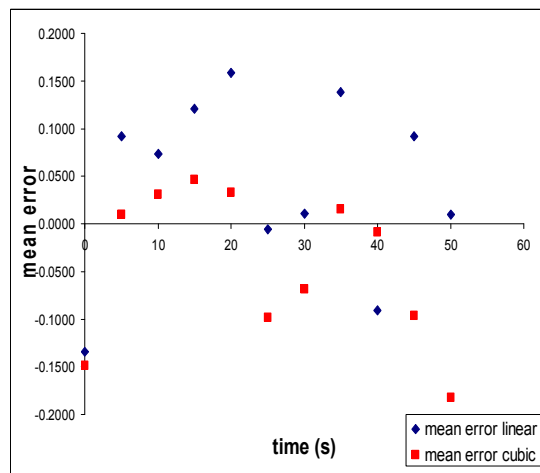

Figure 4 - Mean Interpolation Error for a series of interpolations, calculated by the leave-one-out cross validation technique. Linear and cubic interpolation errors are displayed.

Figure 5 shows the linearly interpolated concentration profile of the time frame at 20 seconds as a color map, with the positions of the chemical detectors overlaid as red dots. The color bar on the right side of each of these figures is the legend to the color map, with red colors corresponding to high concentrations of MES, and blue colors corresponding to low concentrations of MES. Note that empirically observed concentration values are only available at the detector sites (the red dots). All area between red dots is interpolated. Figure 6 is a similar plot for the cubic interpolation of the same time frame. By comparing these representative plots, it is clear that the cubic method provides a smoother interpolated surface than the linear method. This is evidenced in the northern parts of the test grid, where the linear interpolation produces steep gradients between the high and low detected levels, while the cubic method produces a more gradual transition, with a more Gaussian shape.
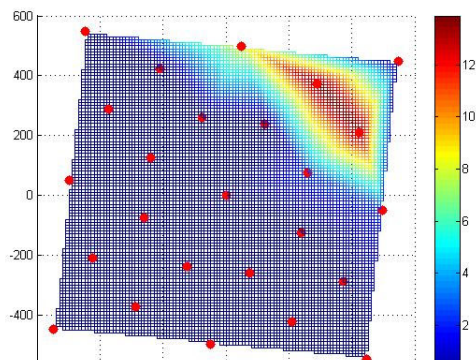
Figure 5 - Linear interpolation of time-frame at 20 seconds, overlayed with the positions of the chemical detectors (red circles). The x and y axes are distance in kilometers, and the color scale represents the concentration of MES in mg/m^3.
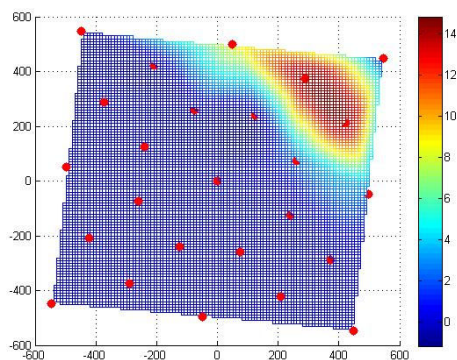

Figure 6 - Cubic interpolation of time-frame at 20 seconds, overlayed with the positions of the chemical detectors (red circles). The x and y axes are distance in kilometers, and the color scale represents the concentration of MES in mg/m^3.

5. CONCLUSIONS AND FUTURE RESEARCH

Based on the preliminary results shown here, it appears that the cubic method of interpolation is more accurate (lower error) than the linear method. In addition, it is

reasonable to assume that the more Gaussian shape produced by cubic interpolation better models the real world chemical cloud distribution than the linear interpolation. However, it must be stressed that these conclusions are based upon a limited amount of data from a single chemical release. Additional experiments will be required. In addition to the linear and cubic interpolation methods presented here, a kriging algorithm will be examined. Interpolation errors for the three methods will be compared for a variety of test cases, including other triethyl phosphate (TEPO) releases and mixed MES and TEPO releases. Test cases with fewer point detectors available will also be examined.

The second phase of the research will leverage meteorological data recorded concurrently with chemical vapor concentration data in order to improve the results of interpolation. Algorithms will be explored and developed to more closely consider the physical phenomena underlying the WSN. In the third phase, meteorological inputs will be considered fuzzy terms, and fuzzy methods will be examined to further improve interpolation results by providing a mechanism for handling the innate ambiguity of micrometeorological conditions. Additionally, temporal interpolation may be explored.

Acceptance testing of chemical detection systems is the primary application for which this interpolation system is being developed. The integration of meteorological considerations into the interpolation algorithms is what will make this system unique, and will enhance the point detector array's ability to evaluate chemical detection systems. With the inclusion of the meteorological effects, this system could be applicable to a variety of real world scenarios. Potential applications could be discovered in the military force protection or homeland security domains, including monitoring military installations and high risk terrorism targets like sporting events or transportation hubs. Additional uses may include hazard evaluation of industrial chemical accidents or tracking and predicting dispersion of smoke from natural disasters. It may also be useful for the continuous background monitoring of emissions from chemical plants for early warning of leaks or accidental releases.

6. ACKNOWLEDGEMENTS

7. REFERENCES

"griddata," in Matlab 7 Function Reference Volume 2: F-O. Natick, MA, 2009, pp 2-1585 – 2-1588.

"WeatherStation Owner's Guide," Airmar Technology Corporation, 2006.

Daintith, John, "Delaunay triangulation," A Dictionary of Computing. 2004. Encyclopedia.com. 13 Aug 2009.

Guestrin, C., P. Bodik, R. Thibaux, M. Paskin, and S. Madden, "Distributed regression: an efficient framework for modeling sensor network data," 3rd International Symposium on Information Processing in Sensor Networks, Berkeley, CA, 2004, pp. 1-10.

Harden, C. S., V. M. McHugh, B. S. Ince, R. J. Schafer, G. E. Blethen, and E. L. Maloney, "Debunking the myth: "Ion mobility spectrometers are unsuitable for quantitative analyses of chemical plumes"," Chemical Biological Defense Physical Science and Technology Conference, New Orleans, LA, 2008.

Krause, A., C. Guestrin, A. Gupta, and J. M. Kleinberg, "Near optimal sensor placements: maximizing information while minimizing communication cost," Information Processing in Sensor Networks, Nashville, TN, 2006, pp. 2-10.

Liaskovits, P. and C. Schurgers, "Levaring redundancy in sampling - interpolation applications for sensor networks," Distributed Computing in Sensor Systems. vol. 4549 Berlin / Heidelberg: Springer, 2007, pp. 324-337.

Meijering, E., "A Chronology of Interpolation: From Ancient Astronomy to Modern Signal and Image Processing," Proceedings of the IEEE, vol. 90, pp. 319-342, March 2002.

Perillo, M., Z. Ignjatovic, and W. Heinelman, "An energy conservation method for wireless sensor networks employing a blue noise spatial sampling technique," 3rd International Symposium on Information Processing in Sensor Networks, Berkeley, CA, 2004, pp. 116-123.

Schafer, R.J., V.M. McHugh, C.S Harden, G.E. Blethen, E.L. Maloney, and B.S. Ince, "A Detector Aray for Real-time Ground Truth Measurement and Reporting of Chemical Concentration and Micrometeorological Conditions – part I: Development," Chemical Biological Defense Physical Science and Technology Conference, New Orleans, LA, 2008.

Tynan, R., G. M. P. O'Hare, D. Marsh, and D. O'Kane, "Interpolation for Wireless Sensor Network Coverage," 2nd IEEE Workshop on Embedded Networked Sensors, Sydney, Australia, 2005, pp. 123-131.

Umer, M., L. Kulik, and E. Tanin, "Kriging for localized spatial interpolation in sensor networks," 20th International Conference on Scientific and Statistical Database Management, Hong Kong, China, 2008, pp. 525-532.

M. C. Vuran and O. B. Akan, "Spatio-temporal characteristics of point and field sources in wireless sensor networks," IEEE International Conference on Communications, 2006, pp. 234-239.