

City University of Hong Kong  
香港城市大學

Diffusing Event Monitoring in  
Sensor Networks  
基于传感器网络的  
扩散型事件监测技术研究

Submitted to  
Department of Computer Science  
電腦科學系  
In Partial Fulfillment of the Requirements  
For the Degree of Doctor of Philosophy  
哲學博士學位

by

Cui Xiao Ning  
崔筱宁

July 2010  
二零一零年七月

## Abstract

The research is motivated with the background of an increasingly pervasive sensing world today. Being the key role in the applications of pervasive computing, sensor networks have made ubiquitous sensing, ad-hoc networking, automatic computing and strategic communication possible and effective. The term, *diffusing event*, summarizes the common characteristic of a series of environmental problems such as fire, typhoon, flood and fume diffusion. Being a flexible reactive network, the sensor network serves as the best choice of monitoring diffusing events and sketching the diffusion trend in the real time. Therefore, the diffusing-event monitoring problem is researched based on sensor networks. The major contributions are as follows:

Firstly, the data processing architecture is designed and a data discrimination framework is presented. Specifically, in the node scope, the data processing architecture is designed to support dynamic correlation discovery and adaptive data sampling. The experiments on real-world datasets have proved the energy consistency and data accuracy of the architecture. In the network scope, in order to reduce the influence of sensor data uncertainty, a systematic discrimination framework is presented to partition the raw data set into event, error and ordinary subsets through node-level temporal processing, neighbor-level spatial processing, cluster-level ranking and network-level decision fusion. The experiment has increased the distinction ratio to as high as 97% with the error occurrence rate up to 50% in the network. Compared with traditional event/anomaly detection, the proposed framework can considerably reduce false-alarm rate and keep miss-hit rate at an acceptable low level.

Secondly, diffusing events are modeled. A *TSEC* conceptual model is presented to completely express the temporal, spatial and event-related correlations. Based on *TSEC* model and the domain knowledge of diffusing events, association modeling is presented, including the establishment of homogeneous/heterogeneous data correlation models, the event-abstraction model and their associations. This work is the first to integrate the temporal, spatial and event relations into the same conceptual model, and the first to apply such a conceptual model into the diffusing-event monitoring in sensor networks.

Thirdly, the general process of diffusing-event monitoring is analyzed, the cor-

responding algorithms are presented for efficient monitoring, and both theoretical analysis and case studies are proposed. Based on the *TSEC* model, the general monitoring process is analyzed and the available solutions are proposed for the key points in the general process. Following the general process, a series of window-based in-network cooperation algorithms are presented for diffusing-event monitoring, involving the linear regression on sensor nodes, the intra-/inter-cluster information exchanging for event-boundary detection and the judgment of diffusion trend. The experiments on real-world datasets have demonstrated the energy efficiency, report reliability and scalability of the proposed algorithms. The theoretical analyses on the comparison of single-source vs. multi-source diffusion and pure diffusion vs. mixed diffusion have indicated that the basic diffusing-event monitoring strategies can be easily extended to monitoring the two types of complex diffusing events. In addition, the scenario-based case studies have been conducted to investigate the wind effect and geographical influence on diffusing-event monitoring. For the wind-constrained diffusing-event monitoring, a heterogeneous method is proposed to cooperate wind nodes (for diffusion-trend determination) and concentration nodes (for event-boundary detection). The experimental results have shown the event coverage rate at 80%~95% with the heterogeneous method, while 60%~25% without such a method. For the geography-constrained diffusing-event monitoring, GA-deployment is presented to enhance the event detection probability. A Voronoi diagram is generated from the layout of geographical obstacles and the network area are allocated with different node densities according to the edge set of the Voronoi diagram. The typical network tessellations have been extended based on the idea of GA-deployment, and such extension has considerably saved energy, with low miss-hit rate, high fault tolerance and strong report reliability.

In summary, the major contributions of this work involve the design of data processing architecture for dynamic correlation discovery, the data discrimination framework to improve data accuracy, the modeling for diffusion-event monitoring, the presentation of a series of diffusing-event monitoring algorithms with theoretical analysis and case studies on typical applications. The ideas, models and approaches presented in this thesis can also be extended and applied to the related domains.

**Key Words:** diffusing-event monitoring, sensor network, major diffusion trend.

## Table of Contents

Chapter 1	Introduction .....	1
1.1	Background .....	1
1.2	Problem Statement .....	2
1.2.1	Problem Scope .....	2
1.2.2	Problem Definition .....	4
1.3	Major Contributions .....	6
1.4	Thesis Organization .....	10
Chapter 2	Literature Review .....	13
2.1	Overview of Sensor Networks .....	13
2.1.1	Sensor Nodes .....	13
2.1.2	Sensor Networks .....	15
2.2	Sensor Data Mining and Processing .....	19
2.2.1	Sensor Data Mining .....	19
2.2.2	Sensor Data Pre-Processing .....	20
2.2.3	Sensor Data Modeling and Processing .....	21
2.3	Event Monitoring Strategies .....	22
2.4	Domain Knowledge of Diffusing Events .....	24
2.4.1	Domain Knowledge in Fluid Mechanics .....	24
2.4.2	Domain Knowledge in Geometry .....	25
2.4.3	Domain Knowledge in Graph Theory .....	27
Chapter 3	Data Processing for Efficient Diffusing-Event Monitoring .....	29
3.1	The Architecture Design of the Sensor Node for Dynamic Correlation Discovery .....	29
3.1.1	Data Correlation Problem .....	29
3.1.2	Correlation Exploiting Architecture .....	31
3.1.3	Core Solution Module .....	34
3.1.4	Performance Evaluation .....	37
3.1.5	Summary .....	42
3.2	Systematic Discrimination of Sensor Data for Accurate Extraction of Event Data Samples .....	43

3.2.1	Problem Analysis .....	43
3.2.2	Discrimination Framework .....	45
3.2.3	Performance Evaluation .....	47
3.2.4	Summary .....	57
Chapter 4	Diffusing-Event Modeling .....	59
4.1	Network Assumptions .....	60
4.2	TSEC Modeling .....	60
4.2.1	Attribute Classification .....	61
4.2.2	Conceptual Correlation Models .....	61
4.3	Knowledge-Based Association Modeling .....	64
4.3.1	Domain Knowledge of Fume Diffusion .....	64
4.3.2	Homogeneous and Heterogeneous Correlation Models .....	64
4.3.3	Event-Abstraction Model .....	66
4.3.4	Association Modelling .....	68
4.4	Summary .....	71
Chapter 5	Diffusing-Event Monitoring Strategies .....	73
5.1	General Monitoring Process .....	73
5.1.1	Selection of Representative Sensor Nodes .....	74
5.1.2	Generation of <i>ag-samples</i> .....	76
5.1.3	Establishment of Event-Related Links .....	76
5.1.4	Event Confirmation .....	77
5.2	Specific Monitoring Strategies for Diffusing Events .....	77
5.2.1	Monitoring Strategies .....	78
5.2.2	Performance Evaluation .....	80
5.3	Summary .....	83
Chapter 6	Theoretical Analysis on Complex Diffusing-Event Monitoring Strategies	85
6.1	Single-Source Diffusion vs. Multi-Source Diffusion .....	85
6.2	Pure Diffusion vs. Mixed Diffusion .....	87
6.3	Summary .....	89
Chapter 7	Case Studies on Constraint-Based Scenarios .....	90
7.1	Scenario Description .....	90
7.2	Wind-Constrained Diffusing-Event Monitoring .....	91
7.2.1	Network Configuration .....	91

7.2.2	Monitoring Strategies with Wind Effect .....	92
7.2.3	Comparative Performance Evaluation .....	93
7.3	Geographical Constrained Diffusing-Event Monitoring.....	96
7.3.1	The Geographical Relations and Influence on <i>MDT</i> .....	96
7.3.2	Geography-Aware Network Deployment.....	99
7.3.3	Experimental Study .....	103
7.4	Summary .....	110
Chapter 8	Conclusions and Future Work .....	112
8.1	Conclusions .....	112
8.2	Future Work.....	114
	List of Papers Published during PhD Study .....	116
	Bibliography.....	117
	Appendix 1: Dataset of Center Mattress Fire.....	127
	Key points .....	127
	Performance of the linear-regression of the dataset.....	128
	Appendix 2: Dataset of Intel Lab Data .....	132
	Node deployment .....	132
	Data format .....	132
	Sample dataset.....	132

## List of Tables

Table 1.1 The classification of events .....	3
Table 1.2 Geometrical and physical meaning of $f$ .....	5
Table 1.3 Differences between traditional event monitoring and diffusing-event monitoring.....	7
Table 2.1 Energy Comparison among communication, storage, and processing.	17
Table 3.1 Gas concentrations and event states .....	30
Table 3.2 Homogeneous-data processing in Case 3.1 .....	39
Table 3.3 Alert messages in Case 3.2 .....	42
Table 3.4 Event and error types.....	44
Table 3.5 Parameter configuration of Scenario 3.1 .....	48
Table 3.6 Parameter configuration of Scenario 3.2.....	52
Table 3.7 Parameter configuration of the experiment.....	54
Table 4.1 Classification of sensor attributes .....	61
Table 5.1 Basic format of an event report from a sensor node .....	78
Table 5.2 The experiment configuration on event states.....	81

## List of Figures

Figure 1.1 Thesis organization .....	11
Figure 2.1 The logical structure of a sensor node .....	14
Figure 2.2 Sensor board [29].....	14
Figure 2.3 Processor board [29].....	15
Figure 2.4 2-D topological relations [88].....	26
Figure 2.5 Directional relations [88].....	27
Figure 2.6 Directional relations for extended spatial objects of arbitrary shape [88].....	27
Figure 3.1 Homogeneous-data processing framework and data flow.....	32
Figure 3.2 Heterogeneous-data processing framework and data flow.....	33
Figure 3.3 Core solution module.....	34
Figure 3.4 Comparison of raw data and processed results.....	39
Figure 3.5 Comparison of original sensed and output samples for the precipitation dataset of Aug. 1 <sup>st</sup> , 1998 in China, with initial correlation interval 10 and maximum sampling rate four times of the minimum sampling rate .....	40
Figure 3.6 Homogeneous-data processing with/without correlation adjustment.	41
Figure 3.7 The energy cost with different initial correlation intervals and sampling rates .....	41
Figure 3.8 The workflow of the discrimination framework.....	46
Figure 3.9 Distinction ratio in each step of Case 3.1.1 .....	49
Figure 3.10 Comparison of distinction ratio of error ( $r_7/(r_1+r_2+r_7)$ ), event ( $r_8/(r_5+r_6+r_8)$ ) and ordinary samples ( $r_9/(r_3+r_4+r_9)$ ) of Case 3.1.1 .....	49
Figure 3.11 False-alarm and miss-hit rate of Case 3.1.1 .....	49
Figure 3.12 Distinction ratio in each step of Case 3.1.2 .....	50
Figure 3.13 The distinction ratios of error, event and ordinary samples of Case 3.1.2.....	51
Figure 3.14 False-alarm and miss-hit rate of Case 3.1.2.....	51
Figure 3.15 The change pattern of the acceleration in a car accident .....	52



Figure 3.16 Distinction ratio in each step of Scenario 3.2.....	53
Figure 3.17 The comparison of distinction ratios .....	53
Figure 3.18 The comparison of different distinction metrics.....	54
Figure 3.19 The comparison of different steps in processing .....	55
Figure 3.20 The comparison of different error occurrence frequency .....	56
Figure 3.21 Statistical result of false alarm and miss hit rates.....	56
Figure 3.22 Comparison of false-alarm rate .....	57
Figure 3.23 Comparison of miss-hit rate .....	57
Figure 4.1 The relation of event and network .....	59
Figure 4.2 TSC-plane .....	63
Figure 4.3 TSEC-cube.....	63
Figure 4.4 The correlation model for a sensor node with homogeneous neighbor nodes .....	65
Figure 4.5 The correlation model for a cluster with heterogeneous sensing functions.....	66
Figure 4.6 An example of the event-abstraction model .....	68
Figure 4.7 An association model for diffusing-event monitoring.....	70
Figure 5.1 2-partitioned FDS on square tessellation, where the nodes intersecting the vertices of the dotted grids constitute a dominating set. ....	75
Figure 5.2 2-partitioned FDS on triangular tessellation, where the nodes at the center of the dotted hexagons constitute a dominating set.....	75
Figure 5.3 2-partitioned FDS on hexagonal tessellation, where the nodes intersecting the vertices of the dotted triangles constitute a dominating set. ....	75
Figure 5.4 An example of fixed transmission-direction for internal boundary determination .....	80
Figure 5.5 Adaptive linear regression on each node .....	81
Figure 5.6 <i>HAL</i> comparison of two cooperation methods; Methods 1 and 2 represent Methods 5.1 and 5.2 for short.....	82
Figure 5.7 Scalability evaluation of Methods 5.1 and 5.2 .....	83
Figure 7.1 Diffusion modes.....	90
Figure 7.2 In-network behaviors of sensor nodes .....	92
Figure 7.3 Energy comparison of heterogeneous and homogeneous methods ....	94
Figure 7.4 Comparison of event coverage rate .....	95

Figure 7.5 Energy comparison of $12 \times 12$ and $25 \times 25$ networks with fume source at (3,7) and major diffusion direction in the south-east .....	95
Figure 7.6 Energy comparison of $25 \times 25$ and $81 \times 81$ networks with fume source at (3,7) and major diffusion direction in the south-east .....	96
Figure 7.7 Comparison of event coverage rate of three kinds of network scales	96
Figure 7.8 1-D geographical style: (a) point style; (b) line style .....	97
Figure 7.9 2-D geographical styles: (a) flat style; (b) convex style; (c) concave style .....	97
Figure 7.10 The disjoint relations for different 2-dimensional styles of geographical obstacles: (a) flat-flat; (b) flat-concave/convex; (c) convex-convex; (d) concave-concave; (e) convex-concave. ....	98
Figure 7.11 Using boundary pair to determine <i>local MDT</i> : (a) parallel pair; (b) intersecting pair .....	99
Figure 7.12 Allocating node density in 1-dimensional space $R$ .....	100
Figure 7.13 1-dimensional GA deployment .....	102
Figure 7.14 2-dimensional GA deployment: (a) the Voronoi partition of four obstacles $A$ , $B$ , $C$ and $D$ (represented by light-gray area); (b) an example of GA deployment .....	102
Figure 7.15 Layout examples with two deployment methods: (a) random deployment; (b)(c)(d) GA deployment with grade=2, 3, 4 .....	104
Figure 7.16 The comparison of the amount of witness nodes with (a) event range=100, (b) event range=200, and (c) event range=300 .....	105
Figure 7.17 <i>ECR</i> comparison of the two deployment methods with (a) event range=100, (b) event range=200, and (c) event range=300 .....	105
Figure 7.18 The view of a diffusing event in different tessellations (a) square tessellation; (b) triangular tessellation; (c) hexagonal tessellation .....	106
Figure 7.19 Comparison of DS nodes in square tessellation .....	107
Figure 7.20 Comparison of MF nodes in square tessellation .....	107
Figure 7.21 Comparison of reporting nodes in square tessellation .....	107
Figure 7.22 Comparison of reporting nodes in triangular tessellation .....	108
Figure 7.23 Comparison of reporting nodes in hexagonal tessellation .....	108
Figure 7.24 Comparison of false-alarm errors based on <i>OTDS</i> in square tessellation .....	108
Figure 7.25 Comparison of Miss-hit errors based on <i>OTDS</i> in square tessellation	

.....109

Figure 7.26 Comparison of false-alarm errors based on *OTDS* in uniform deployment.....109

Figure 7.27 Comparison of miss-hit errors based on *OTDS* in uniform deployment.....109