

## Research Article

# Multimedia Fusion for Public Security in Heterogeneous Sensor Networks

Jiangfan Feng,<sup>1,2</sup> Wenwen Zhou,<sup>1</sup> and Kaixin Sun<sup>3</sup>

<sup>1</sup> College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

<sup>2</sup> Key Laboratory of Instrument Science and Dynamic Test, North University of China, Taiyuan 030051, China

<sup>3</sup> Key Laboratory of Virtual Geographic Environment, Nanjing Normal University, Nanjing 210046, China

Correspondence should be addressed to Jiangfan Feng; [fengjf@cqupt.edu.cn](mailto:fengjf@cqupt.edu.cn)

Received 9 May 2014; Revised 7 September 2014; Accepted 21 September 2014; Published 27 October 2014

Academic Editor: Pietro Siciliano

Copyright © 2014 Jiangfan Feng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Public security is a widespread disastrous phenomenon that constitutes a grave threat. Although information fusion of video sensor networks for public security has been studied extensively, multimedia fusion in heterogeneous sensor networks or its application in public security remains a challenge and central goal in the field of information fusion. In this study, to realize the detection, monitoring, and intelligent alarm of such hazards, we develop a graph-based real-time schema for studying the dynamic structure of heterogeneous sensors for public security. In the proposed schema, data fusion algorithms based on data-driven aspects of fusion are explored to locate the optimal sensing ranges of sensor nodes in a network with heterogeneous targets. In addition, we propose a framework incorporating useful contextual and temporal cues for public security alarm, explore its conceptualizations, benefits, and challenges, and analyze the correlations of the target motion elements in the multimedia sensor stream. The experimental results show that the new method offers a better way of intelligent alarm that cannot be achieved by existing schemes.

## 1. Introduction

Public security is a thorny issue concealed almost everywhere and threatens our physical safety at any time. The threats are influenced by some unpredicted reasons, such as unruly crowd aroused by emergencies or natural hazards from nowhere. Unfortunately, the most tragic catastrophes are allegedly due to humans, and the entrenched invisible terrorists always find a way to break the social regulatory life. In such case, any leak may lead to various tangible and intangible losses. Accordingly, there is a critical need for offering early detection and alarm, and such guarantee may enable instant reaction to winning most time for evacuation and deployment and will also reduce the possibility of loss.

Information fusion, since born in 1970s as a military requirement, shows considerable potential, and later multi-sensor data fusion (MSDF) arises at the historic moment [1]. The past experience of MSDF shows that it contributes to integration and analysis of sensors data and to more accurate understanding of the situation and how to react to it [2, 3],

which makes the disaster prevention possible. Mostly, it was restricted to combine vast correlation sensor data in some particular detection, and some of the topics of interest are fire detection [4, 5], surveillance of transportation situation [6, 7], and novel for the specialty as data management [8, 9]. On the other side, approaches are devoted to elimination of redundancy in the sensor networks [10]. Considering the complicacy of public security, maximizing relevant multimedia sensors data as well as exclusion of redundancy should be involved in this case. That means strong spatial-temporal correlation multimedia sensor data are needed to effectively redeploy a multiview and multisense monitor system [11]. As the equivalent intention, a coordinated scheduling method [12] was presented following the divisible load theory (DLT) to minimize both finish time and consumption in sensors network. Also topology rebuilding algorithm (TRA) [13] is addressed as the solution to data dissemination, in which the workloads meet the specified requirements.

Vision sensors that afford successive video frames conceal some correlation semantic knowledge transferred by series

interactional target motion elements [14, 15]. Besides, the series multimedia sensors could be used to detect much associated information, such as the noise caused by the crowd, the smell which may be the clue of some toxic gas, and even some detected suspicious words captured by the sound sensor. As a result, the fusion of these levels of environmental data may feasibly be responsive to any situation that happens to the region without actually being here. Moreover, an alarm is made if any hidden danger is included. Hence, to pursue the goal, technologies are required to achieve immediate intelligent alarm, including the following: (1) capture multisense environmental data in one scene from different perspectives; (2) dig the spatial relations and interaction underlying the multimedia sensors; (3) achieve the decision of multimedia sensors data fusion.

## 2. Related Works

Recent information fusion researching shows the expressive activity and is highlighted. Most of them concentrate on making the utmost use of the comprehensive sensor resources, which are based on the reasonable disposal of multiple sensors information, and includes several distinctive uses of the correlation and independence, contextual information, and modality selection between different sensors [16, 17]. The methods are required for robust and parallel processing ability. Furthermore, complementary and redundant information from different sensors which vary in the modalities and spatiality asks for an optimizing description on consistency of aimed objects.

The fusion of various modalities is generally performed at the feature level (early fusion) and decision level (late fusion) [16]. The feature level fusion [18, 19] extracts the features from the original information referred by sensors and then analyzes and processes the extracted modality information which in our case may include visual feature (Color, Shape, or Texture), audio feature (frequencies or Texture), and motion features (motion trace or motion pattern). The feature level fusion achieves impressive compression for real-time processing and also provides considerable feature information for the decision analysis, but it may not perform well in the time synchronization between multimodal features. The decision level fusion [20, 21] employs numerous sensors to observe one object, and each sensor finishes its own local processing (e.g., pretreatment, feature extraction, and recognition), and then the local decisions are combined to a fused decision based on the correlation. This approach is advantageous in that it utilizes unique representations for the diverse modality features. Besides great scalability in fusion process rather than at the feature level, it costs extra time caused by obtaining the local decisions.

Based on diverse specific application background, there are many widely used fusion approaches involved in or concerned with relevant studies. In the information retrieval domain, data fusion is provided for the A rank/score function [22, 23]. In the application of current tendency, several approaches employ temperature, infrared, and smoke sensors for environmental data detection [24, 25] and video sensor network (VSN) is utilized by intelligent transportation

systems or surveillance systems [26, 27], while information fusion is widely used in the fields of robotics, image processing, or some interconnected fields [28, 29]. For instance, the environmental data (temperature and humidity) are adopted as in-field lower-level data, in which the change at each sensor node can be observed by a chosen threshold, and then the probability of fire is confirmed by using Dempster-Shafer evidential reasoning (D-S) to fuse VSN data [5]. However, the supply of the vision sensors is limited to corroborate the assumption proposed as the first step; much associated video semantic knowledge was barely effective in timely warning and thus was wasted. In addition, the combination of the Kalman filtering (KF) with D-S [30] conquers the insufficiency of both algorithms and results in a matching model. A strategy of camera topology estimation was proposed to prevent the probability of the overlap that occurs in the reality, and also the fusion of dynamic features of the targets between successive video frames was still excluded in the theory. Moreover, an information fusion model devoted to tracking and locating materials was adopted into automated identification and location estimation [31]. The author provided a multilevel data fusion model based on the modified JDL model which joined D-S theory and weighted mean. The hybrid fusion method succeeded in fusing data from disparate sensor nodes in a noisy and dynamic environment and was advanced in uncertainty and imprecision.

Recent information fusion research shows the expressive activity and is highlighted. The research concentrates on making the utmost use of comprehensive sensor resources, based on reasonable disposal of multisensor information, and also focuses on distinctive uses of the correlation, independence, contextual information, and modality selection between sensors [16, 17]. Methods are required to improve the robust and parallel processing ability. Furthermore, the complementary and redundant information from different sensors which vary in the modality and spatiality asks for an optimized description with consistency of aimed objects.

## 3. Multisensors Network and Directed Graph

As far as concerned, we will begin by looking at some minimal models of public security in crowds. Regardless of how data from different sensors are organized, the underlying fusion algorithms must ultimately fuse the input data. The aim of public security-oriented fusion applications is to deal with several data-related challenges. As a result, we will explore data fusion algorithms from the data-driven aspects and seek the optimal sensing ranges of sensor nodes in a network with some targets. It is assumed that a sensor is able to dynamically adjust its sensing data by the arbitrary amount.

In general, a varying circumstance will diversify the work condition of each sensor and may cause many sensors to directly respond to the change. This means, under the scope of the effected sensors, it captures considerable correlative information about the detected environment. Meanwhile, considering the spatial relations, as one sensor detects the environmental factors and some other factors vary in time, the corporate sensors also respond to the instant scene.

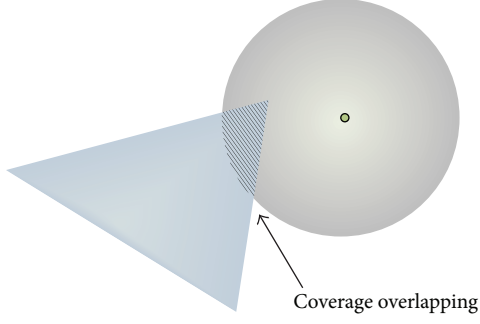


FIGURE 1: Overlapping relationship.

In this paper, we define it as the interaction between two sensors that interprets the probability when one sensor is triggered while the other responds to the same initiator. And a quantization method is used to dynamically define the interaction relationship, and directed graph is adopted to describe this related tendency, which leads to the maximum correlation information from diverse sensors, and eventually sensor data are adaptively dispatched for optimal fusion aimed at the dynamic environment.

**3.1. Multisensor Attribute Node.** Multisensor attribute mainly represents the perceived multimodal data, such as the image information captured by video sensors. The aims of uniform description of the interaction are to seek for the state analysis and quantization of multimodal sensor data. Besides the effects from temporal and spatial factors, we also define and illustrate the related notions of involved elements and thus write the problem below.

**Definition 1 (sensor state).** The state of sensor is the diverse modal sensor data, that multi-sensors perceive one scene and is to give one description of the condition from one sensing method at some point. In this paper, the multisensor state is defined as follows: state = {Decibel, Smokescope, Cohesion}, which separately indicates the results that the audio sensor, odor sensor, and video sensor quantify the basic state of the detection range. Specifically, Decibel and Smokescope quantify the real-time environment, while Cohesion measures the compactness index of motion elements in the video images.

**Definition 2 (spatial dependence).** Measurement of spatial relationship aims to measure the spatial scale between two entity elements. In this case, the spatial dependence is associated with the range distribution of the sensors' coverages, rather than the linear distance between them, and in physics, it may refer to the overlapping area of any two sensors (Figure 1). Apparently, the overlapping area directly reveals the dependence of one sensor on the other, as targets can be detected by both in the overlapping area. Thus, the proportion of coverage overlap accounts for the range of one sensor's coverage that indicates the probability that the other sensor can detect the target at the same time. Therefore, the dependence can be described by scaling the proportion of the overlapping area in one sensor's geometric coverage under

the unified geographic space. Hence, we assume any two sensors coverages as two nonempty sets of spatial elements, and based upon that, we utilize mathematical methods to describe the spatial dependence degree as follows.

The overlapping area of sensors A and B is  $D_{ab}$ , and the sensor coverages of A and B, respectively, are  $D_a$ ,  $D_b$ . Pursuant to the definition we make for spatial dependence (SD), we can measure the SD between A and B as value  $h$ :

$$h_{ab} = \frac{D_{ab}}{D_a} \quad (1)$$

$$h_{ba} = \frac{D_{ab}}{D_b}.$$

Among them  $h_{ab}$  denotes the spatial dependence sensor A to B; likewise  $h_{ba}$  is the spatial relationship in which B affects A. In this case, as the proportion of the two sensors' coverage overlapping area is considerable to the sensor coverage, which equals the SD that is high in value, it represents the strong spatial dependence between them.

Accordingly any multisensor attribute node can be defined as follows:  $S_i = (\text{Time}, \text{State}, \text{SD})$ , which indicates that at temporal space Time the sensor  $S_i$  is at State and the SD ( $h_{i,j}$ ) refers to the spatial relationship with another sensor  $S_j$ .

**3.2. Interaction of Multisensors.** According to the definition of sensor attribution, the interaction between two sensors is mainly affected by the State and SD. Therefore, the values  $P$  as the probability of sensor B relating to sensor A at Time can be computed on the basis of the real-time multisensor node attribute.

Considering the rate of State change and SD as a two-dimension random variable ( $S$ , SD), when the value of the quantized State approaches the infinity, the interaction probability  $P$  only limits to SD (3) and also is a linear relationship. Similarly, as the two spatial element sets entirely match each other as  $h$  approaches the infinity,  $P$  only depends on State (4). The expressions can be formulized as follows:

$$P_{SD}(h) = P\{SD \leq h\} = P\{SD \leq h, S < \infty\} = P(SD, \infty)$$

$$P_S(S) = P\{S \leq s\} = P\{S \leq s, SD < \infty\} = P(S, \infty). \quad (2)$$

Accordingly we can define the interaction probability  $P$  as a two-dimension distribution function  $P(S, SD)$ , and  $P(\infty, \infty) = 1$ , as random parameters  $S$  and SD are mutually independent, and thus the probability function  $P(s, h)$  is expressed as follows:

$$P(s, h) = p_S(s) p_{SD}(h), \quad (3)$$

where  $p_{SD}(h)$  and  $p_S(s)$ , respectively, denote the marginal function based on SD and  $S$ , respectively, and the relevant probability formulas are expressed as follows:

$$P_S(s) = 1 - \exp\left(-\frac{s}{a_1}\right), \quad a_1 > 0, s > 0 \quad (4)$$

$$P_{SD}(h) = h, \quad h > 0. \quad (5)$$

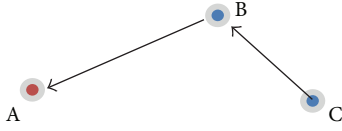


FIGURE 2: Graph evolution rule.

Thus, (3) equals

$$P(s, h) = h \left[ 1 - \exp\left(-\frac{s}{a_1}\right) \right]. \quad (6)$$

And  $s > 0$ ,  $h > 0$  and  $a_1 > 0$ . Hence it can be achieved that, at any time, if the SD between sensors A and B stays constant  $h_{BA}$  and the State of sensor B change is  $S_B$ , then the probability sensor B related to A is  $P_{BA} = P(S_B, h_{BA})$ .

### 3.3. Directed Graph

**3.3.1. Graph Evolution Rules.** Graph  $G$  is defined as  $G = (V, E, \Omega)$ , where  $V$  is a vertex set,  $E$  is an edge set, and  $\Omega$  is the probabilistic space. Any two sensors connect with probability  $P_{i,j}$  while sensor  $S_i$  relates to sensor  $S_j$  which means  $P(e_{ij} \in E) = P_{ij}$ ,  $P_{ij} \in (0, 1)$ . As  $P_{i,j}$  varies in time, at time  $T$  the graph evolution rule can be expressed as as Figure 2.

**3.3.2. Graph Syntax.** Like the graph evolution rule mentioned above, the multisensor network may be diagrammatized as the graph. In the graph, the sensor nodes act as the vertexes, and any two sensors construct an interaction relationship connecting with the independent directed edge, and the edges perform at probability  $p_{ij}$ . Since  $p_{ij}$  is affected by sensor state and also dynamically varies with time factor, the graph syntax based on multisensor network and the interaction among sensor nodes can be defined as five meshes:

$$G = (N_0, N_n, P, C_0, N_g), \quad (7)$$

where  $N_0$  is the initial vertex which is the node trigger of the event, and the initial one is exclusive; the vertex set  $N_n = \{N_1, N_2, N_3, \dots\}$  represents all the sensor nodes involved in the multisensors network;  $P = \{P_{00}, P_{01}, \dots, P_{ij}\}$  is the quantized interaction between two sensors ( $N_i, N_j$ ) and is also the probability that two nodes can construct incidence relation;  $C_0$  in the random graph indicates the clustering coefficient of the initial node  $N_0$  which means the amount of the sensor nodes that the initial node related to;  $N_g$  denotes the clustering vertex set starting from node  $N_0$ .

## 4. Multisensor Data Processing Framework

The framework incorporates useful contextual and temporal cues for public security alarm, which constitute an integrated platform of associated sensors and compute infrastructure with capability of delivering valuable real-time information regarding the natural hazards to public security. In this case, each involved sensor acts as an independent local data processing unit, and each output of them represents a set

of individual decisions and is reported to the fusion center, which contributes to self-interaction or cooperates with others during the fusion process. There are two parts during this process: the first is local fusion center, which is in charge of fusing the single modality data so as to judge the surrounding change in identity-sensor. The output can utilize the other local fusion center's capabilities of confirming the result or cooperating with others to achieve the final judgments in the results in the next part, the fusion center. This part is capable of global estimation based on multimodalities data, which finally give a comprehensive review of the environment.

Before the data fusion process, the information of all the sensors requires a unified expression, which means the decision from each other can be obtained between the homogeneous or heterogeneous sensors. In this case, all the sensor data are first quantized and then transferred to the information for decision-making that acquires the semantics from various media sources. With the public monitor system, the quantization of correlation Cohesion between motion elements can express the semantic knowledge underlying the video stream from each video sensor node. In our case, the data of acoustic sensor and odor sensor are fused to detect the same scene from different aspects. Figure 3 depicts the fusion process and the multisensor data flow, and each modality node owns an exclusive judgment standard.

**4.1. Audio Sensor.** The sensor senses the voice occurring in the detection region and mainly aims to capture the strange noise or irregular sound change and capture the optimal data to deploy these sensors in a decentralized pattern. Meanwhile, feature redundancy for robustness and simultaneous elimination of data redundancy are required during the voice capturing process, in case of data blocking or energy limitation [32].

The audio information performs as temporal variation analog signal, because the diversity of the sounds is to change with the circumstances. For instance, at the same spot, the voice at 10:00 a.m. differs from that at 10:00 p.m., since the first one in the rush hour may contain the noise of busy crowd, while the other is usually a peaceful night and any isolated voice can be easily caught by sensors. Thus, measuring the voice in one region is designed to collect the mean per hour  $T_N$  ( $N = 1, 2, \dots, 24$ ) and then compare it with the real-time value  $T_0$ , and the unusual case can be excluded if the difference does not exceed a reasonable range  $D$ . which denotes the mean of the ambient voice 12 per hour. Therefore, the disparity of the two values landing in the range indicates a normal situation, which is formularized as  $|T_0 - T_N| \in D$ , and otherwise the node would send a warning.

Additionally, the voice also transmits fertile underlying semantics. Thus, recognizing the hidden semantics in the ambient sounds can realize the acquaintance of the scenarios. With this purpose, the processing includes two steps at each sensor node: eliminate the background noise and proceed to speech recognition. In particular, speech denoising as a mature technology has considerable academic results [33, 34] and also is used in diverse realms. In comparison, speech recognition as a raising one is proposed to analyze associated semantics. The speech consists of brain-made phoneme

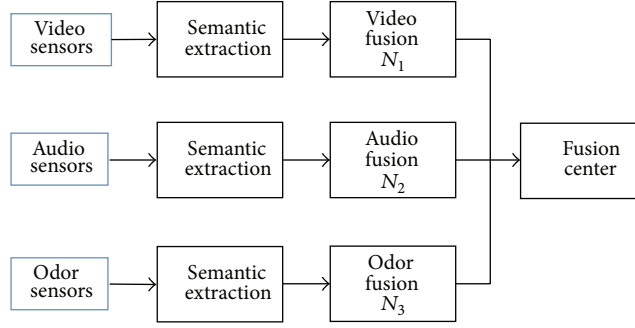


FIGURE 3: The fusion process.

flow parameters and is actually a double stochastic process. Meanwhile, the voice signal is an observable time-varying sequence. The statistical model HMM is founded by the structure of time series that well simulates this process. It is a suitable model for speech recognition by vector quantization (VQ) speech signal and abundantly training the model and ultimately achieves the process of speech recognition.

As we intend to capture those particular words that may cause unwanted results, the process of speed recognition mainly contributes to their extraction from the background sounds. Considering the indeterminacy of the daily language, somehow, the speech semantics analysis may not be a decisive attribute of the event detected but plays a role of coefficient attribute with other features to intensify the judgment.

**4.2. Odor Sensor.** The odor sensors as environment data monitoring devices are deployed as well as the audio sensors to work as the smoke scope detection and observe the real-time ambient data. Since the odor in specific confines of an environment mostly stays at one level, from the abundant statistic data and research on the unknown environmental change, it can approach exception events information by analyzing normal distribution of sensed periodic output, which assumes that it obeys normal distribution. The probability density function is described as follows:

$$f(x_i; \mu_0, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - \mu_0)^2}{2\sigma^2}}, \quad (8)$$

where  $\mu_0$  is the average of the given odor value and  $\sigma$  is standard deviation indicated as  $N(\mu_0, \sigma^2)$ , and then the test statistics shows the law of probability distribution of the odor state. The absence of emergencies occupies the most probability, whereas the rare events are the matters that should be paid attention to. If the sensor captures odor sample  $X_1$ , the value drifts from the normal event region in the probability distribution, leading to the judgment whether  $X_i$  triggers the target event.

**4.3. Video Sensor.** Video sensors set at different angles for surveillance of one observation region can transfer successive video frames affording video stream data and analyze associated semantic knowledge. Such information instantly supplies the potential for detecting the unsafe issue and with

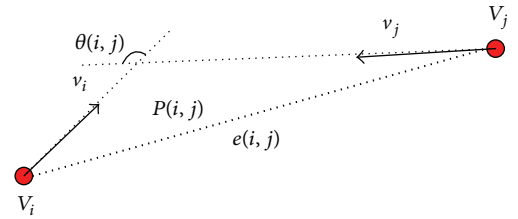


FIGURE 4: Dynamic spatial topology relationship.

substantive stochastic factors, the whole motion elements can be ascribed to an analyzable model. In this case, Cohesion is adopted to measure the compactness of motion elements [35], as the intense activity directs higher Cohesion that shows the active movement of the crowd.

Motorial entity objects shown in the video monitor system are described as video motion element (VME), such as pedestrians appearing in the observation regions. They usually consist of relevant states (*Appear*, *Disappear*, *Stop*, and *Move*), spatial relationships (*Measure*, *Direction*, and *Topology*), visual features (*Color*, *Texture*, *Shape*, and *Size*), and behavior attributes (*Location* and *Velocity Vector*) in a series. In particular, *Appear*, *Disappear*, *Stop*, and *Move* represent the four basic states of a motion element. We observe the relation that correlative effects on the interactive objects acquire this interaction as the clue to express the intrinsically link among these elements and finally estimate the possibility of changing the motion states.

In this case, we use the vertex set to describe motion entities and define the independent edge as the relation of any two vertexes connected by probability  $P$ . Hence, the interactive spatial relationship in the motor process that dynamically changes with time can be approached in a united representation model in the particular region. The vertex set  $V_N = \{V_1, V_2, V_3, \dots, V_n\}$  denotes all the objects in the region;  $V_i$  and  $V_j$  connect with each other via  $P(V_i, V_j) = P(t)$ . To bring in these correlative coefficients, relevant information of these motion elements in the observation region can be extracted from the video stream and then specified as some particular description. The two objects (shown as a vertex)  $V_i$  and  $V_j$  indicate their dynamic topology relationship as in Figure 4, and  $[P(i, j) = P(e_{ij} \in E)] \in [0, 1]$  represents the interaction probability, and the value relates

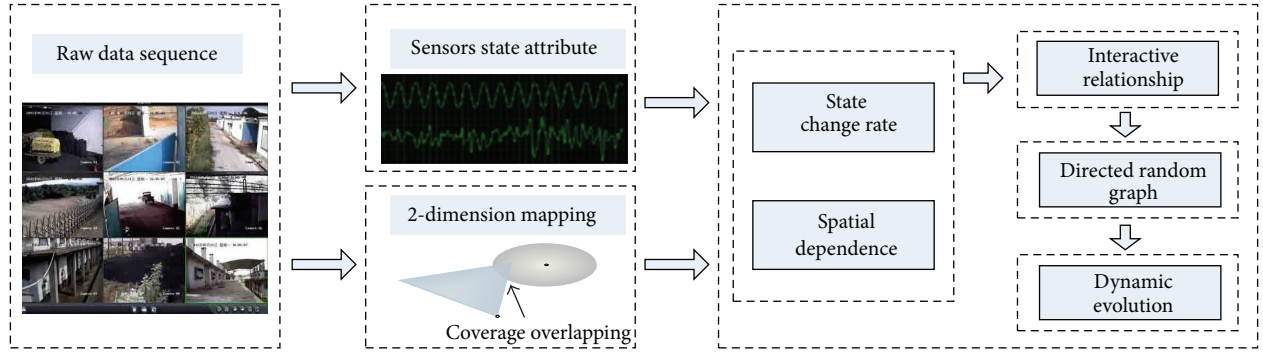


FIGURE 5: Frame diagram of implementation progress.

to the compactness between the two objects controlled by the distance( $i, j$ ) and other behavior attributes. We define the initial value as  $P(0) = \rho_1 \times \text{distance}(i, j) + \rho_2 \times \theta(i, j)$  and  $\sum P(i, j)$  in the improbability space  $\Omega$ . The  $P(i, j)$  contained by  $V_i$  and  $V_j$  varies with Time.

The intensification of linking between motion elements shows the activity and disorder degree of the elements in the frames and is indicated by Cohesion ( $C_0$ ). Let  $N(i) = \sum_{i=1}^n P(V_i, V_j)$  and  $\xi(i) = N(i) / \sum_{j=1}^n N(j)$ , and then the Cohesion of  $G$  can be expressed as

$$\text{Cohesion}(G) = - \sum_{i=1}^n \frac{N(i)}{n} \times \frac{\xi(i) \ln \xi(i)}{\ln n}. \quad (9)$$

The value of Cohesion ranges from 0 to 1.

## 5. Simulations

In this section, we will propose a concrete implementation scheme based on the conceptual model and analyze the simulation results from the hypothetical scenarios. We assume that all targets within the node sensing ranges are sampled at an equal rate, independent of their distance to sensor nodes. The progress frame diagram is indicated in Figure 5.

The scenarios are based on the assumption that there is a known scene deployed with ordered multisensors  $N_i$  including audio, odor, and video sensors, and among them,

the sensors detect target events constituting a set  $N_j$ ,  $N_j \subseteq N_i$ . In scenarios, we map the range of the sensor coverage onto a two-dimensional coordinate plane, and thus the coverage of sensors can be regarded as the closed bounded points set. Meanwhile, we consider the sampling period of each sensor as  $T$ , and consequently  $S$  can be expressed as the average rate of State change in period time  $nT$  ( $n > 2$ ), which is calculated as

$$S = \left( \sum_{n=1}^{n-1} \frac{\Delta \text{State}_n}{n} \right) = (|\text{State}_{nT} - \text{State}_{(n-1)T}| + |\text{State}_{(n-1)T} - \text{State}_{(n-2)T}| + \dots + |\text{State}_{2T} - \text{State}_T|) \times (nT)^{-1}. \quad (10)$$

**5.1. Simulation.** In this simulation, we assume that the effect coverages of the audio, odor, and video sensors can be diagramed as circles with radius of 4, 3, and 5, respectively, and deploy the parameters of  $N_0$  in the coordinate plane listed in Table 1.

Known as Definition 2, we can approach the spatial dependence between any two sensors from their pixel proportion and overlapping proportion based on (1); accordingly, we can obtain the spatial dependence measuring value  $h_{i,j}$  of entire  $N_j$  listed as the matrix  $m$ :

	$N_1$	$N_2$	$N_3$	$N_4$	$N_5$	$N_6$	$N_7$	$N_8$	$N_9$	$N_{10}$	$N_{11}$	$N_{12}$	$N_{13}$
$N_1$	1.0000	0.0259	0.0160	0.0994	0.2556	0.0568	0	0	0.1293	0.2226	0	0.1293	0
$N_2$	0.0259	1.0000	0	0	0	0	0.2052	0.0829	0.1961	0	0.1569	0	0
$N_3$	0.0160	0	1.0000	0	0	0.2294	0	0.1490	0	0	0	0	0.3600
$N_4$	0.0994	0	0	1.0000	0.2294	0	0	0	0	0.0991	0	0	0
$N_5$	0.3993	0	0	0.3584	1.0000	0	0	0	0	0.5625	0	0.1321	0
$N_6$	0.0888	0	0.3584	0	0	1.0000	0	0	0	0	0	0.0969	0.0268
$N_7$	0	0.3206	0	0	0	0	1.0000	0	0.0479	0	0	0	0
$N_8$	0	0.1295	0.2327	0	0	0	0	1.0000	0	0	0	0	0.0422
$N_9$	0.3593	0.5448	0	0	0	0	0.0852	0	1.0000	0	0	0	0
$N_{10}$	0.6185	0	0	0.2753	1.0000	0	0	0	0	1.0000	0	0.1482	0
$N_{11}$	0	0.4357	0	0	0	0	0	0	0	0	1.0000	0	0
$N_{12}$	0.3593	0	0	0	0.2349	0.1722	0	0	0	0.1482	0	1.0000	0
$N_{13}$	0	0	1.0000	0	0	0.0477	0	0.0750	0	0	0	0	1.0000

(11)

TABLE 1: The coordinate of sensors.

$N_i$	$N_1$	$N_2$	$N_3$	$N_4$	$N_5$	$N_6$	$N_7$	$N_8$	$N_9$	$N_{10}$	$N_{11}$	$N_{12}$	$N_{13}$
$X$	16	9	21	12	17	23	5	15	11	17	6	21	21
$Y$	15	9	7	22	20	12	13	5	13	19	5	17	6
$r$	5	5	5	5	4	4	4	4	3	3	3	3	3

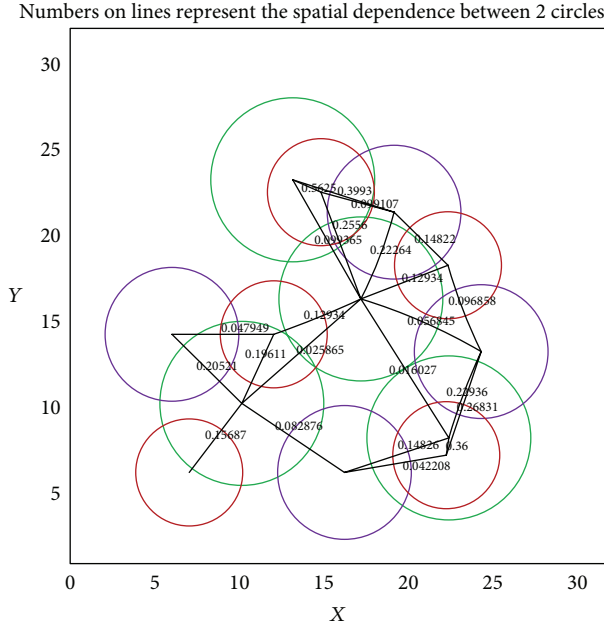


FIGURE 6: Spatial dependence.

TABLE 2: The index of sensors.

	Video	Audio	Odor
$\mu$	$\mu_v = 0.5$	$\mu_a = 45$	$\mu_o = 83$
$\sigma$	$\sigma_v = 1.5$	$\sigma_a = 1$	$\sigma_o = 1$

The  $m_{ij}$  in the matrix equals value of  $h_{ij}$ , also the entire spatial relationship among sensors can be diagramed in general as in Figure 6.

In this situation, we assume the data of the sensors are subject to normal distribution and the simulation can correspond to the status of the sensors change, which we define as the expectation and variance of each kind of sensors based on actual circumstances as in Table 2.

And particularly as the example shown in Figure 7, on the top of the fig pictures the simulation results which denote the Decibel value detected by audio sensors in a certain period time (equal to 200 ms, and in which  $T = 2$  ms,  $n = 10$ ) besides the outgo discrete points on the bottom half show the average rate values of its State change in every sampling period.

Based on the measurable parameters  $S_i$  and  $h_{ij}$ , the established incidence relation between two sensors ( $N_i, N_j$ ) probability  $P_{ij} = \{p_{00}, p_{01}, \dots, p_{ij}\}$  at time  $nT$  can be approached by (6), and Figure 8 shows this kind of incidence relation at a certain time point.

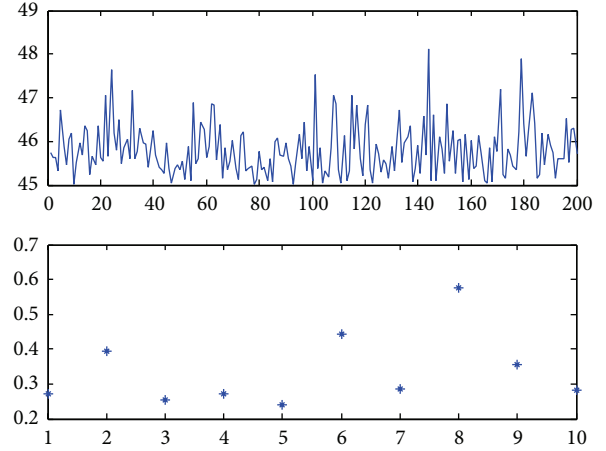


FIGURE 7: Sample index.

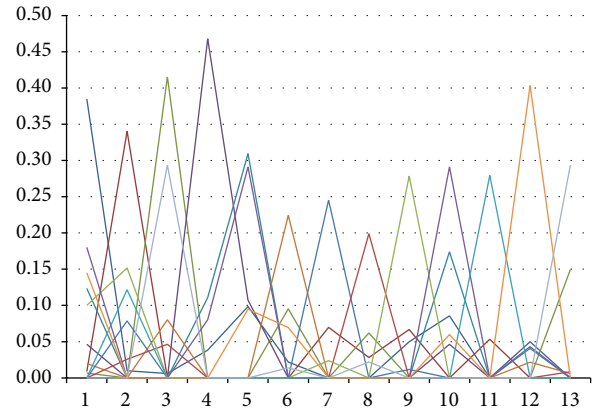


FIGURE 8: Incidence relational time point  $t$ .

Each continuous line in Figure 8 denotes the connection probability of one sensor relating to the others (also including itself). According to the graph evolution rule defined before, this mutual incidence relation can be expressed as graph  $G$  (Figure 9(a)). On the other side, we can easily estimate the average cluster coefficient (ACC) of a network node set, and at time  $t$ , it is equal to 0.1625, and the clustering coefficient of each sensor is demonstrated in Figure 9(b).

### 5.2. Scenario

5.2.1. Scenario 1. In this scenario we create a target event  $E$  that occurs in the spheres of the sensors, whose coordinate locates at (7, 17) in the mapping two-dimension coordinate.

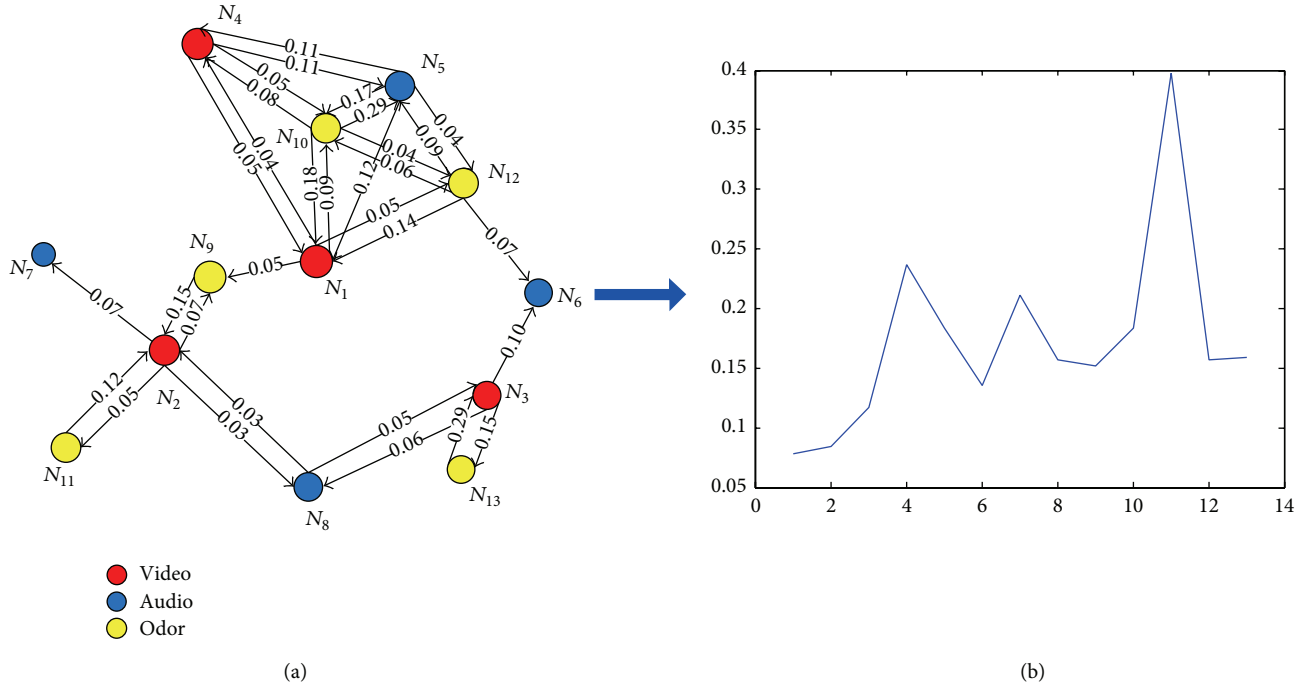


FIGURE 9: (a) Generated directed graph at time  $t$ . (b) Clustering coefficient at time  $t$ .

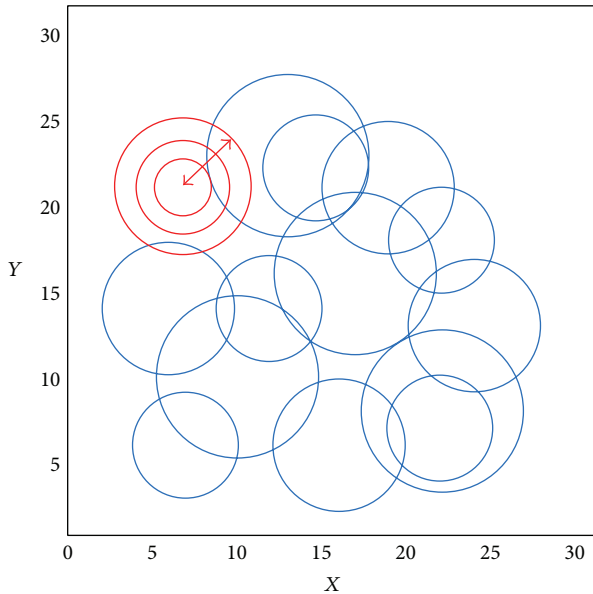


FIGURE 10: The process event E works.

Meanwhile the effect coverage of the event varies over time which expands then gradually narrows down, apparently as it invades the spheres of sensors that will change the state parameters of these sensors, and the process works like what Figure 10 shows.

Hence we define the event E radius of effect coverage

$$R = \begin{cases} rT, & T < t \\ r(2t - T), & T > t, \end{cases} \quad (12)$$

$(0 < t < T);$

$r$  denotes the initial effect radius and  $T$  means the effecting time as event E works. Accordingly the expectation of sensors  $\mu$  varies with  $R$  and can be expressed as

$$\mu = \begin{cases} \mu_0 T, & T < t \\ \mu_0 (2t - T), & T \geq t, \end{cases} \quad (13)$$

$(0 < t < T),$

where  $\mu_0$  is the initial expectation and the event E will affect the sensors  $N_4$ ,  $N_7$ , and  $N_9$  during its affecting time, and Figure 11 shows the incidence relation among 13 sensors at the sampling time point of this period. Also Figure 12 refers to the clustering coefficient of each sensor at the sampling time  $T$ , and the corresponding ACC of the sensor node set ( $Nn$ ) and the clustering affected nodes set ( $N_g$ ) are listed in Table 3.

**5.2.2. Scenario 2.** This scenario simulates the target event  $E'$  as an effective point moves across the spheres of sensors with a constant velocity; the movement shows as in Figure 13.

Equally, as we defined the event  $E'$ , it works on the expectation of the affected sensor  $\mu$ . The effective point will move across the sensors  $N_1$ ,  $N_9$  and  $N_3$ ,  $N_{13}$  successively, and the working status of these correlative sensors will change correspondingly as the event works. Thus, the value of  $\mu$  floats during the period, which can be expressed as function (14), and the period from  $T_1$  to  $T_2$  is when the event works on the sensor. Consider

$$\mu = \begin{cases} \mu_0, & t \in \text{other} \\ \mu_1, & T_1 < t < T_2, \end{cases} \quad (14)$$

$0 < t < T.$



TABLE 3: The ACC of the sensor nodes set and affected nodes set.

ACC	$T_0$	$T_1$	$T_2$	$T_3$	$T_4$	$T_5$	$T_6$	$T_7$	$T_8$	$T_9$
$N_n$	0.1768	0.1853	0.1800	0.1874	0.1907	0.1848	0.1944	0.1885	0.1558	0.1794
$N_g$	0.2096	0.2062	0.2421	0.2209	0.2507	0.2318	0.2413	0.2049	0.1762	0.1982

TABLE 4: The ACC of the sensor nodes set and affected nodes set.

ACC	$T_0$	$T_1$	$T_2$	$T_3$	$T_4$	$T_5$	$T_6$	$T_7$	$T_8$	$T_9$
$N_n$	0.1640	0.1761	0.1680	0.1845	0.1725	0.1688	0.1912	0.1560	0.1763	0.1696
$N_g$	0.2228	0.2262	0.2274	0.2407	0.2223	0.2126	0.2697	0.2130	0.2557	0.2166

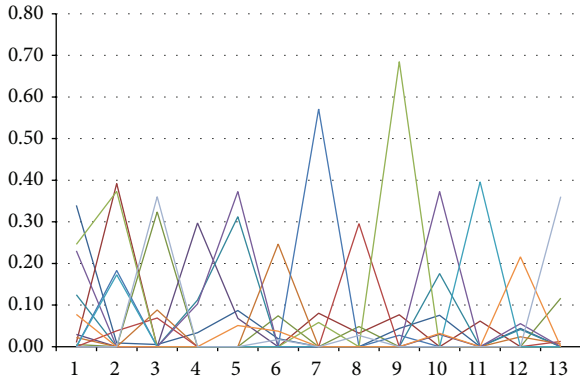


FIGURE 11: Incidence relation at affecting time point.

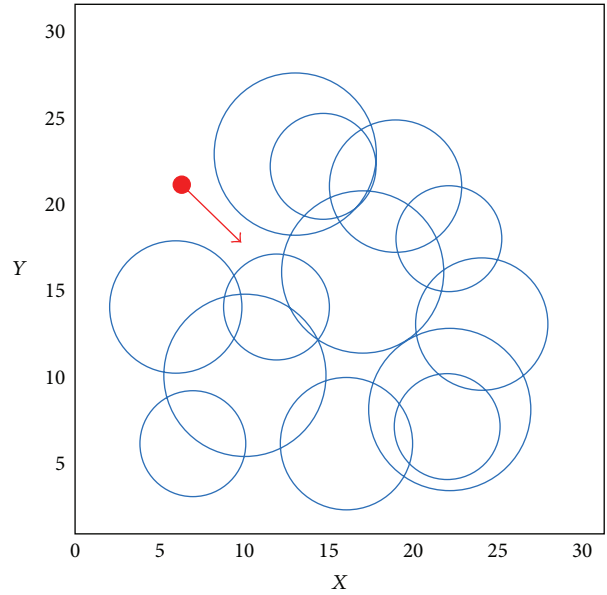


FIGURE 13: The process event  $E'$  works.

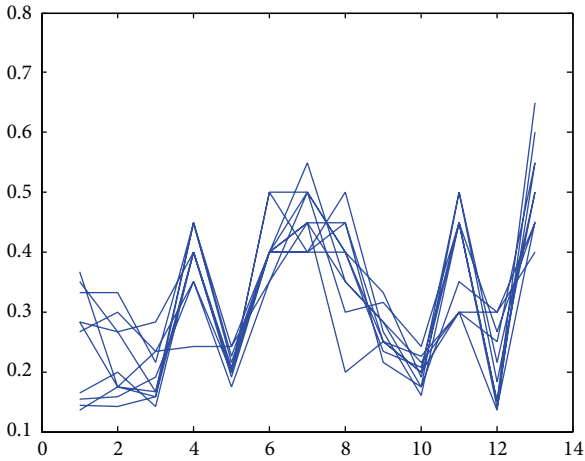


FIGURE 12: Clustering coefficient during affecting time.

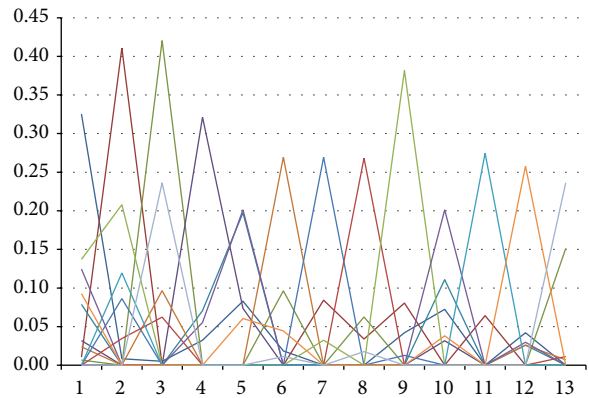


FIGURE 14: Incidence relation at affecting time point.

During this process, the incidence relation of the sensor network can be demonstrated as in Figure 14 at the sampling time  $T$ , and meanwhile at each time point in this period, we can approach the clustering coefficient of each sensor and the results are showed in Figure 15.

With the method we proposed, the acquired clustering vertex set ( $N_g$ ) and the whole sensor network set ( $N_n$ ) can be computed with the corresponding ACC. Table 4 demonstrates the results that the ACC index has a particular enhancement compared with the original complete graph  $N_n$ .

5.3. Analysis. The results shown in the figures and the recorded data demonstrate that the incidence relation index and directed graph parameters significantly change during the event acting period. The interactive relationship among the sensors reveals diversities that compared to the general case depicted by Figure 8, especially in the affected sensor

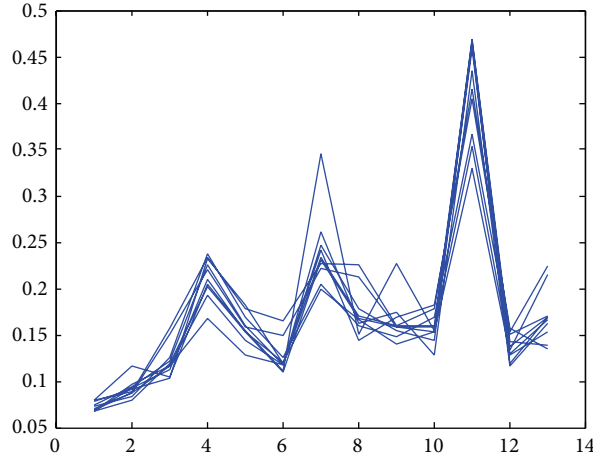


FIGURE 15: Clustering coefficient during affecting time.

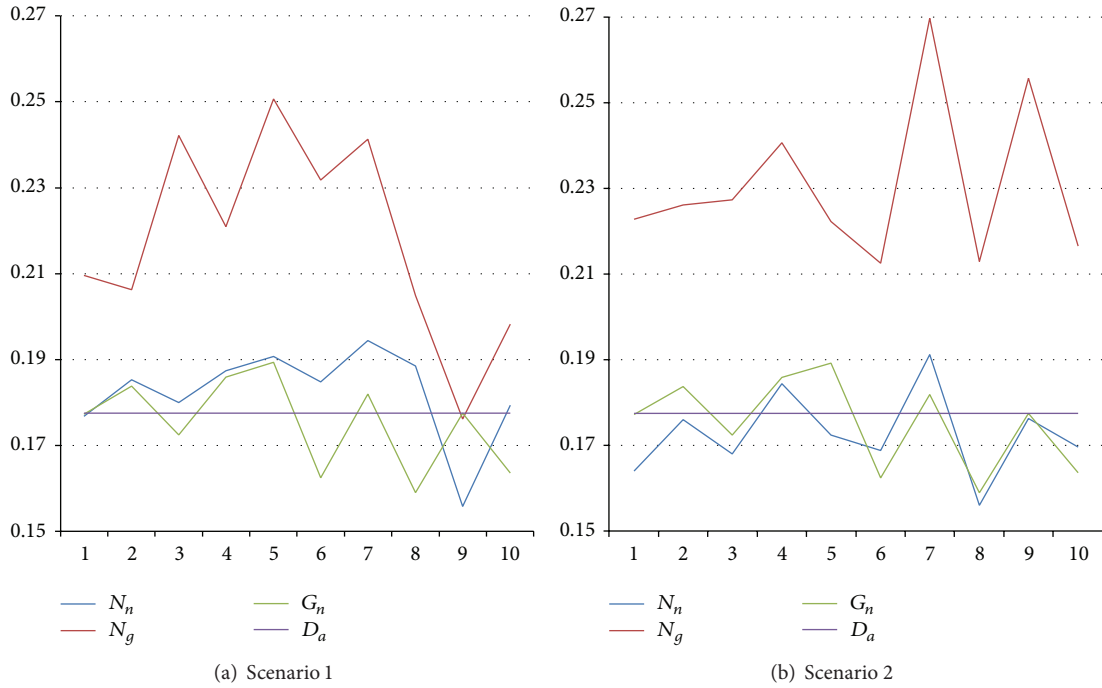


FIGURE 16: The comparison indexes in scenarios.

node set. ACC generally increases from 0.1625 to about 0.18, which points to the intense interaction influence among sensors. However, in the clustering affected node set  $N_g$ , the ACC (around 0.22) shows the more intense interaction compared to the whole sensor network node set  $N_n$ , which also means the method excluded the excess redundancy nodes and obtained maximum correlation information.

The authors in [36] highlight the spatial correlation of sensor nodes and measure the correlation between a sensor and its neighboring sensor nodes to approach the data aggregation. The usage of topology is to diagram the distribution of sensors, and thus in this graph, we can acquire

the correlative graph parameters recorded as  $D_a$  and compare with the indexes acquired by our method from scenarios 1 and 2 (shown in Figure 16). Due to the exclusion of real-time sensor state, the spatial correlation remains unchanged; we can deduce that the involved sensor state would develop the affirmative function to better analyze the interaction of multiple sensors.

In Figure 16, the  $G_n$  and  $N_n$  lines denote the ACC index of the whole sensor network in the general condition and in the event acting condition, respectively. The figure indicates that the indexes are resemblant except in the effective process of the event; the  $N_g$  line clearly shows that the index of the

clustering affected nodes increases compared to the index of the whole network, which means the nodes we approach are in a close correlation. Besides, all the indexes indicate the real-time change of the interactive relation among sensors except  $D_a$  and also show certain growth compared with this constant.

Thus, based on the incidence relation, it makes it feasible to approach the optimal associated sensors which are derived by the events. Moreover, the spatial range of sensors enhances the dependence on sensed data, which raises the feasibility value of the final fusion process.

## 6. Conclusions and Future Work

We introduced a new sensor data fusion method of a heterogeneous multimedia sensor network, explored the associated frameworks and algorithms for incorporating useful contextual and temporal cues for public security alarm, and thereby analyzed the correlations among the target motion elements in the multimedia sensor stream. We investigate a multimedia intelligent processing method based on spatial relations of heterogeneous sensors, which is able to achieve scalable recognition. We also develop a dynamic structure of multimedia sensor representation and thereby realize the fast generation of multimedia stream, highly heterogeneous networks, and complicated alarm needs in public safety.

We further propose establishing a general framework for incorporating useful contextual and temporal cues for public security alarm and thereby analyze the correlations of the target motion elements in the multimedia sensor stream. The graph-based framework will serve as a useful tool for designing multimedia in-network processing schemes in MSNs. Simulation results verify the analysis of the proposed techniques.

This new multisensor data fusion scheme will inspire a number of interesting topics in this field for future research. For instance, fuzzy set theory is widely recognized as a critical issue for public security with multisensor fusion, but in the present paper, we have a light touch on this. In the future, we aim to advance implemented fusion algorithms by integrating alternative combination rules. We also attempt to deal with public events of alarm that are conducted to be optimized to multimedia sensed data and set appropriate fusion parameters as alarm thresholds, fusion weights, and others. Also, further validation through real data is of extreme importance and will also be conducted in future work validation.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

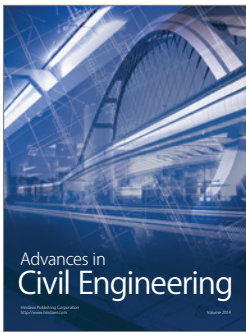
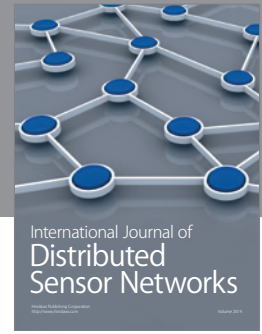
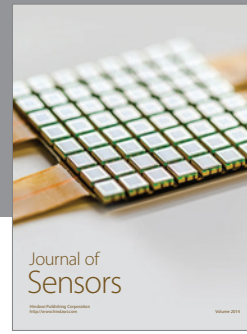
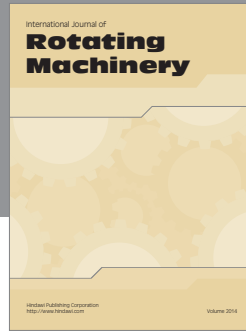
The work is supported by the National Nature Science Foundation of China (nos. 41201378 and 41101432), the Science Foundation of China for Postdoctors (no. 2014M561212), and

the Natural Science Foundation Project of Chongqing CSTC (no. 2011jjA30014).

## References

- [1] F. Castanedo, "A review of data fusion techniques," *The Scientific World Journal*, vol. 2013, Article ID 704504, 19 pages, 2013.
- [2] O. H. Bray, "Data fusion for adaptive control in manufacturing: impact on engineering information models," *Engineering with Computers*, vol. 11, no. 3, pp. 145–156, 1995.
- [3] G. Navarro-Arribas and V. Torra, "Information fusion in data privacy: a survey," *Information Fusion*, vol. 13, no. 4, pp. 235–244, 2012.
- [4] A. Díaz-Ramírez, L. A. Tafoya, J. A. Atempa, and P. Mejía-Alvarez, "Wireless sensor networks and fusion information methods for forest fire detection," *Procedia Technology*, vol. 3, pp. 69–79, 2012.
- [5] E. Zervas, A. Mpimpoudis, S. Hadjiefthymiades, O. Sekkas, and S. Hadjiefthymiades, "Multisensor data fusion for fire detection," *Information Fusion*, vol. 12, no. 3, pp. 150–159, 2011.
- [6] Y. Ren, D. Peng, J. Wu, and Y. Zhou, "The research and application of multi-resource heterogeneous data fusion on Dynamic Traffic Routing System," *Advances in Intelligent Systems and Computing*, vol. 215, pp. 375–387, 2014.
- [7] B. Schick and S. Schmidt, "Evaluation of video-based driver assistance systems with sensor data fusion by using virtual test driving," in *Proceedings of the FISITA 2012 World Automotive Congress*, Lecture Notes in Electrical Engineering, pp. 1363–1375, Springer, 2013.
- [8] Z. Zhu and S. Donia, "Spatial and visual data fusion for capturing, retrieval, and modeling of as-built building geometry and features," *Visualization in Engineering*, vol. 1, p. 10, 2013.
- [9] R. Nuray and F. Can, "Automatic ranking of information retrieval systems using data fusion," *Information Processing and Management*, vol. 42, no. 3, pp. 595–614, 2006.
- [10] D. Gavalas, A. Mpitziopoulos, G. Pantziou, and C. Konstantopoulos, "An approach for near-optimal distributed data fusion in wireless sensor networks," *Wireless Networks*, vol. 16, no. 5, pp. 1407–1425, 2010.
- [11] D. P. Mandic, S. L. Goh, and K. Aihara, "Sequential data fusion via vector spaces: fusion of heterogeneous data in the complex domain," *Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology*, vol. 48, no. 1–2, pp. 99–108, 2006.
- [12] X.-L. Li and J.-N. Cao, "Coordinated workload scheduling in hierarchical sensor networks for data fusion applications," *Journal of Computer Science and Technology*, vol. 23, no. 3, pp. 355–364, 2008.
- [13] E. F. Nakamura, F. G. Nakamura, C. M. S. Figueiredo, and A. A. F. Loureiro, "Using information fusion to assist data dissemination in wireless sensor networks," *Telecommunication Systems*, vol. 30, no. 1–3, pp. 237–254, 2005.
- [14] J. Feng and H. Song, "Intelligent analysis for georeferenced video using context-based random graphs," *International Journal of Distributed Sensor Networks*, vol. 2013, Article ID 158569, 14 pages, 2013.
- [15] P. Perez, J. Vermaak, and A. Blake, "Data fusion for visual tracking with particles," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 1–18, 2004.
- [16] P. K. Atrey, M. A. Hossain, A. El Saddik, and M. S. Kankanhalli, "Multimodal fusion for multimedia analysis: a survey," *Multimedia Systems*, vol. 16, no. 6, pp. 345–379, 2010.

- [17] F. Caron, E. Duflos, D. Pomorski, and P. Vanheeghe, "GPS/IMU data fusion using multisensor Kalman filtering: introduction of contextual aspects," *Information Fusion*, vol. 7, no. 2, pp. 221–230, 2006.
- [18] Z. Wu, L. Cai, and H. Meng, "Multi-level fusion of audio and visual features for speaker identification," in *Proceedings of the International Conference on Advances in Biometrics (ICB '06)*, pp. 493–499, Hong Kong, January 2006.
- [19] W. Chen, P. Meer, B. Georgescu, W. He, L. A. Goodell, and D. J. Foran, "Image mining for investigative pathology using optimized feature extraction and data fusion," *Computer Methods and Programs in Biomedicine*, vol. 79, no. 1, pp. 59–72, 2005.
- [20] F. Castanedo, J. García, M. A. Patricio, and J. M. Molina, "Data fusion to improve trajectory tracking in a cooperative surveillance multi-agent architecture," *Information Fusion*, vol. 11, no. 3, pp. 243–255, 2010.
- [21] A. M. Aziz, "A new multiple decisions fusion rule for targets detection in multiple sensors distributed detection systems with data fusion," *Information Fusion*, vol. 18, no. 1, pp. 175–186, 2014.
- [22] D. F. Hsu and I. Taksa, "Comparing rank and score combination methods for data fusion in information retrieval," *Information Retrieval*, vol. 8, no. 3, pp. 449–480, 2005.
- [23] S. L. Wu, "Applying statistical principles to data fusion in information retrieval," *Expert Systems with Applications*, vol. 36, no. 2, pp. 2997–3006, 2009.
- [24] W. J. Sacks, D. S. Schimel, and R. K. Monson, "Coupling between carbon cycling and climate in a high-elevation, subalpine forest: a model-data fusion analysis," *Oecologia*, vol. 151, no. 1, pp. 54–68, 2007.
- [25] S. C. Popescu and R. H. Wynne, "Seeing the trees in the forest: Using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height," *Photogrammetric Engineering and Remote Sensing*, vol. 70, no. 5, pp. 589–604, 2004.
- [26] J. D. de Jesus, J. J. Villacorta Calvo, and A. I. Fuente, "Surveillance system based on data fusion from image and acoustic array sensors," *IEEE Aerospace and Electronic Systems Magazine*, vol. 15, no. 2, pp. 9–16, 2000.
- [27] F. Castaldo and F. A. N. Palmieri, "Data fusion using a factor graph for ship tracking in harbour scenarios," *Smart Innovation, Systems and Technologies*, vol. 26, pp. 189–196, 2014.
- [28] X. Zhao, Q. Luo, and B. Han, "Survey on robot multi-sensor information fusion technology," in *Proceedings of the 7th World Congress on Intelligent Control and Automation (WCICA '08)*, pp. 5014–5018, June 2008.
- [29] J. Xu, J. Li, and S. Xu, "Data fusion for target tracking in wireless sensor networks using quantized innovations and Kalman filtering," *Science China. Information Sciences*, vol. 55, no. 3, pp. 530–544, 2012.
- [30] *Information fusion and its application in transportation surveillance [Dissertation thesis]*, Shanghai Jiao Tong University, 2010.
- [31] S. N. Razavi and C. T. Haas, "Multisensor data fusion for on-site materials tracking in construction," *Automation in Construction*, vol. 19, no. 8, pp. 1037–1046, 2010.
- [32] L. E. Palafox and J. A. García-Macías, "Deploying a voice capture sensor network system for a secure ubiquitous home environment," *International Journal of Communication Systems*, pp. 1–14, 2009.
- [33] H. Attias, L. Deng, A. Acero, and J. C. Platt, "A new method for speech denoising and robust speech recognition using probabilistic models for clean speech and for noise," in *Proceedings of the European Conference on Speech Communication*, pp. 1903–1906, 2001.
- [34] K. W. Wilson, B. Raj, P. Smaragdis, and A. Divakaran, "Speech denoising using nonnegative matrix factorization with priors," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '08)*, pp. 4029–4032, April 2008.
- [35] J. Feng and H. Song, "Dynamic topology of motion element based on sparse random graph grammar," *Journal of Information and Computational Science*, vol. 10, no. 2, pp. 501–508, 2013.
- [36] F. Yuan, Y. Zhan, and Y. Wang, "Data density correlation degree clustering method for data aggregation in WSN," *IEEE Sensors Journal*, vol. 14, no. 4, pp. 1089–1098, 2014.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

