International Journal of Smart Home
Vol.7, No.6 (2013), pp.351-358
http://dx.doi.org/10.14257/ijsh.2013.7.6.34

Discriminating Event Information for Calculation of Basic Probability Assignment

Donghyok Suh and Juhye Yook[1]

Department of Multimedia Communication, Far East University,
[1]Department of Rehabilitation Technology, Korea Nazarene University
{hanhwaco@naver.com},[1]{jhyook@kornu.ac.kr}

Abstract

This study proposed a novel way to calculate Basic Probability Assignment(BPA), which is crucial in the data fusion. The study mapped sensing signal with a linear discriminant function analysis and assessed the sizes with time. This was beneficial to get a clue for context inference with using the Dempster-Shafer Theory and to determine BPA based on the size changing of mapping data in time intervals. The study provides with the way of context inference for fast detecting a local event that affects the whole area.

Keywords: *Context inference, Basic Probability assignment, Data fusion*

1. Introduction

Data fusion process with multi-sensors has caught increasingly more attention nowadays. Fuzzy Theory (FT) and Dempster-Shafer Theory (DST) have been used much as a method to process heterogeneous signals and data fusion although these theories were designed originally to express the ambiguity and uncertainty of events occurring in the real world. Multi-sensors data fusion process is applied to acquire information with high quality by using some kinds of sensors in biometrics, geographic information, network security, robot control, etc. Multi-sensors data fusion is also important to acquire context information with high quality in the Ubiquitous Sensor Network(USN) data process.

Multi-sensors data fusion for acquiring context information in the sensor network has become big data process in the data streaming conditions. The studies in the data streaming conditions first consider searching, classifying, and clustering the characteristics of continuous data that are not saved in a server. However, it is necessary to infer the changing context in real time through the data patterns acquired from the real time data. For example, we need to infer situations urgently and efficiently without holding to analyze when a blaze in a large area occurs, objects move around the road, or some risk factors threat the security in our residential environment. To do so, it is necessary to infer the context based on the rapid analysis of data patterns by time intervals in the data stream. We could obtain a clue of the simple and early context inference from the changing patterns with time of the sensing data detected and reported by sensors.

The study presents a way to infer context information by discriminating event information reported to a host through sink nodes in the data streaming environment. It is also suggested that the results from discriminating event information are applied to BPA determination with using DST. It is reasonable to determine BPA based on the detection of patterns appeared in

ISSN: 1975-4094 IJSH

the acquired data because the BPA determination for context awareness in the data fusion is based on the various characteristics of data.[1]

Another consideration for context awareness in the data fusion is the changes with time in the process. The data fusion for context inference should consider changes with time because the target situation to recognize by sensor network changes with time. The changes with time in the data fusion of other fields have not been taken into consideration much because they deal with static information such as geographic information and biometrics. However, the changes with time should be applied to controlling robots' condition with acquiring variable information, detecting trespass in the network security, as well as recognizing context in sensor network. The context inference based on discriminating event information in the study should adopt the time flow because the results by discriminating event information with time could be compared. The study comes to context inference by seeking the way to determine BPA based on quantifying the characteristics.

The study includes reviewing basic theories and related research in Chapter 2, explaining the core of the theory from the study in Chapter 3, conducting an experiment with the method and showing the results in Chapter 4, and describing the conclusion in Chapter 5.

2. Related Research

Various studies on BPA have been conducted. W. Jiang *et al.*, [1] proposed a new method to obtain BPA that based on the distance measures between the sample data under test and the model of attributes of species. Zhou *et al.*, [2] proposed transforming BPA to probability function. Their transformation method can acquire better balance between easy decision-making and high risk. J. Cao *et al.*, [3] presented an improved method that redefined the distribution of conflicts. It determined the allocation actors of conflicts by calculating the distance vector of two evidences. Then, it used the improved method to fusion video multiple features in speaker tracking system and compared the results with the standard DST. Their proposed method is suitable for the speaker tracking system in a complex environment. X. Deng *et al.*, [4] showed a novel method for the fusion of multi-sensors information. Within their method, the sensor report has been represented by using DST. Then, an evidence-driven method is proposed to obtain the relative credibility of each sensor based on the power average operator. The weighted balance evidence theory is employed to combine the sensor reports. The proposed method is efficient for the representation of uncertain information and the fusion of conflicting sensor reports. R. Feng *et al.*, [5] proposed a trust management scheme based on revised DST. A trust propagation mechanism including conditional trust transitivity and dynamic recommendation aggregation is developed for obtaining the recommended trust values from third part nodes. They adopted a flexible synthesis method that uses recommended trust only when no direct trust existed to keep a good trust-energy consumption balance. A.M. Lonea *et al.*, [6] found a quantitative solution for analyzing alerts generated by the IDSs, using the DST operations in 3-valued logic and the fault-tree analysis for the previously mentioned flooding attacks. At the last step, his solution uses the Dempster's combination rule to fuse evidence from multiple independent sources.

Z. Zuo *et al.*, [7] presented a method of rough set theory based on random set and BP neural network to obtain the basic probability assignment. Their study makes use of the ability of rough set attribute reduction to reduce the neural network input dimension.

[1] Corresponding Author

Thus, their method is applied to bearing fault diagnosis of wind turbine, and better results are to be achieved.

However, a study on BPA determination considering changes with time for context reference is rarely conducted. Therefore, studies to determine BPA considering time flow as well as to discriminate the streaming from big data streaming acquired by many sensors and reflecting the result on the BPA determination are necessary.

3. Event Information Discrimination and Context Inference

The study presents a new way to determine BPA, which is fundamental in the data fusion for context inference. The presented method to determine BPA is reflecting time flow and based on discriminating event information continuously acquiring in the form of streaming in big data streaming conditions.

3.1. Linear Discriminant Analysis (LDA)

LDA is the way to reduce a dimension of feature vector of the data by maximizing the ratios of between-class scatter and within-class scatter.

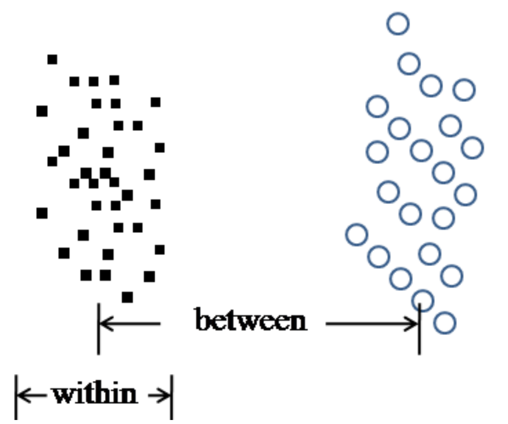

Figure 1. Good Class Separation; Data Distribution that is Easy to Discriminate

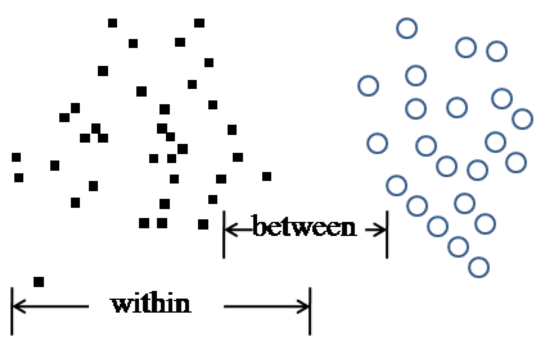

Figure 2. Bad Class Separation; Data Distribution that is Difficult to Discriminate

The Figure 1 is easier to distinguish than the Figure 2 because the data in the Figure 1 are clustered closely together to each central point, and the two central points are separated far away. LDA is a way to reduce a dimension by canceling based on the criteria of a pivot for maximal gab in the feature space so that discriminant information between the classes can be upheld to the maximum.

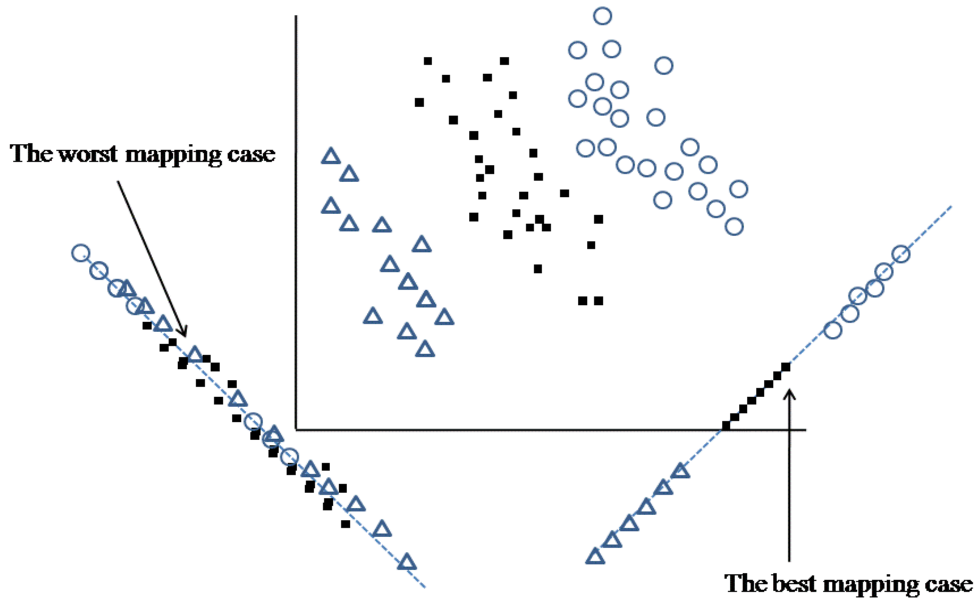

Figure 3. Rearrangement by Pivot Transformation after LDA

It is possible to reduce a dimension by canceling into 1 dimensional subspace through various transformation matrices with 3 classes of 2 dimensional data as the figure 3. For example, it is difficult to discriminate between classes because red, blue, and green data are all mixed in 1 dimension for the worst 1D subspace reduction. However, we could see that it is easy to distinguish each class even with the data in 1 dimension for the best 1D subspace reduction. Therefore, the LDA is the way to reduce dimensions based on maximizing the class separation.

3.2. Event Occurrence and Sensing Data in the Sensor Network

Many sensor motes play a role of terminal and relay nodes and construct a network. Ideal sensor network is spread in a large area. The sensing data reported in real time in the large area are obtained with eliminating overlaps by data aggregation policy. It is necessary to cluster many sensing data come from the wide area by a criterion of set times. Then, the discrimination of sensing data clustered by the set times is conducted. The study has the background with getting an optimum discriminant function.

The discrimination by a discriminant function does not recognize noticeable context when a particular event does not occur in the large area. However, sensors began to send event information from the center of the area occurring events when particular events occur. The data distribution becomes denser when a discriminant function is applied to the sensing data with event information. The sensing data with event information are gathering towards a particular value when noticeable context is developing.

The distance between the centers of distribution or between the average values could be great when the sensor network detects multi-contexts. The study infers a context by using the changing patterns of event data distribution related to these situations in time intervals as a solution clue. The figure 4 below shows the distribution changes of event information data when events occur in the sensor network. The figure 4 displays the data distribution in the environment in which a particular context does not occur. The data distribution changes when a particular event occurs.

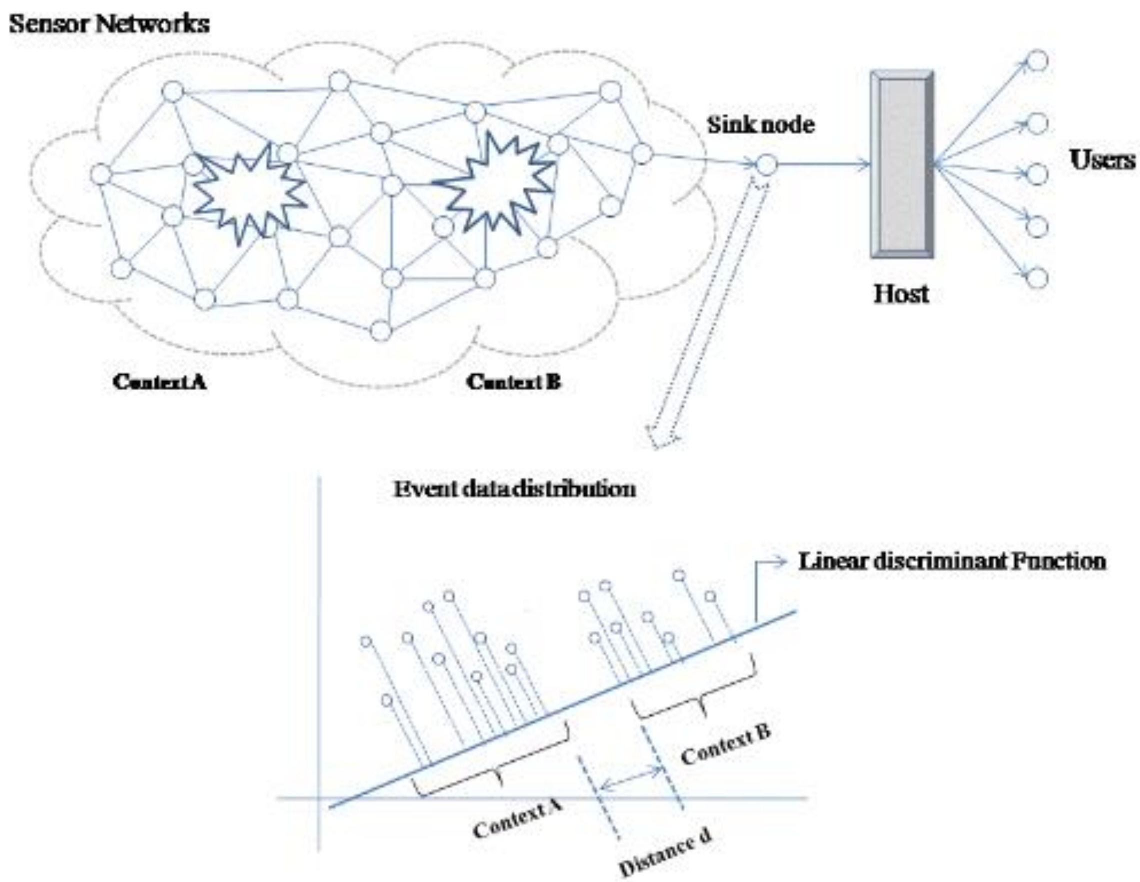

Figure 4. Internal Variance and External Variance of Mapping Data

Groups are formed with the measured values related to their contexts, and the distance exists between the groups as a result that event information data gathered into a sink node through eliminating overlaps are mapping with a discriminant function when the Context A and the Context B are different events. The study suggests a way to infer a context based on the changes of event data distribution corresponding to such event occurrence.

The data in the Context A group are called as A, the data in the Context B group are called as B, and the time is named as t_1, t_2.... t_n when a discriminant function is applied to the event data. The optimal transformation matrix **W** is as follows when **W** is a transformation matrix to minimize data distribution within groups and to maximize the distribution between groups. $\mathbf{S_W}$ is a distribution matrix within classes, $\mathbf{S_B}$ is a distribution matrix between the classes, $\mathbf{u_1}$ is an average vector of group A, and $\mathbf{u_2}$is an average vector of group B.

$$\mathbf{w}^* = \arg\max_{\mathbf{w}} \left\{ \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_w \mathbf{w}} \right\} = \mathbf{S}_w^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$$

The projective $[y_1,y_2,...,y_{c-1}]$ with the number of (C-1) by the projective vector $\mathbf{w_i}$ with the number of (C-1) instead of a single projective vector are calculated. These projective vectors are arrayed with the projective matrices of $W=[w_1|w_2|...|w_{c-1}]$. The focus of the study is the size of the data distribution for Group A with mapping (projecting) a linear discriminant function. The formula for calculating with the linear discriminant function is as follows.

$$\mathbf{y} = \mathbf{W}^{\mathbf{T}}\mathbf{x}$$

x is data matrices, and y is projective vectors. The absolute values of these are shown below.

$$|\,\mathbf{y}\,| = |\,\mathbf{W}^{\mathbf{T}}\mathbf{x}\,|$$

A context can be inferred based on the changes of the absolute values in each time interval. That is, the context becomes clearer as the variation values of the projective data are getting smaller. The context inference can be used for a particular event that occurs locally in a wide area. The procedure so far is organized as the following.

Step 1: Time intervals for data evaluation are set.
Step 2: The averages and variations for sensing data including event information in each time interval are calculated.
Step 3: Optimal transformation matrices are obtained.
Step 4: The sizes of projective data groups are computed with projective vectors of data groups.
Step 5: The sizes of the projective data groups are compared by the times; t_1~t_2, t_2~t_3, t_3~t_4… t_{n-1}~t_n.
Step 6: Contexts are inferred based on the size changes of the projective data groups.

4. Experiment and Evaluation

The experiment is conducted with the method shown in Chapter 3. Thirty sensor motes, a temperature sensor, and a sound sensor were used. The experiment environment is described next.

Experiment procedure

- Wireless sensor network is constructed with the temperature sensor and the sound sensor.
- Sensing data are obtained in every 10 seconds.
- Averages and variations are calculated with the obtained data, and the sizes of projective data sets with a linear discriminant function are computed.
- Then, the differences are looked with repeating the processes in every 10 seconds.
- The BPA with the differences and the *belief* and the *uncertainty* with the DST are produced.
- The contexts are inferred by comparing the *belief* and the *uncertainty* of focal elements.

The Table 1 shows the difference values of the sizes from projective data groups in time intervals.

Table 1. Reported Values

Time(s)	Innervariance	Outervariance
t_1	4.50	2.30
t_2	3.00	2.43
t_3	2.79	2.69
t_4	2.45	3.00
t_5	2.36	3.09

The Table 2 displays the BPA for the DST based on those difference values.

Table 2. Determined BPA

Time(s)	InnerBPA	OuterBPA
t_1	0.05	0.02
t_2	0.11	0.10
t_3	0.18	0.25
t_4	0.31	0.45
t_5	0.35	0.18

The Table 3 provides the *belief* and the *uncertainty* of focal elements based on the BPA.

Table 3. All Focal Element's *Belief* and *Uncertainty*

	t_1~t_2		t_2~t_3		t_3~t_4		t_4~t_5	
	bel	*uncer*	*bel*	*uncer*	*bel*	*uncer*	*bel*	*uncer*
I+O	1.00	0.00	1.00	0.00	1.00	1.00	1.00	0.00
I	0.48	0.28	0.34	0.29	0.29	0.28	0.38	0.28
O	0.24	0.28	0.37	0.29	0.43	0.28	0.33	0.28

The variances of projective values from data group A are decreasing from t_2 as seen in the Table 1. This indicates a particular context occurrence. To infer a cause for a particular context, the Table 3 is used. The context occurred in t_1~t_2 is by focal elements I.

5. Conclusion

The variance values of event information data detected and reported by sensors in sensor network environment distributed and constructed in a large area are changing with time. The study mapped sensing data with a linear discriminant function and checked the sizes with time. This was to get a clue for context inference with using the DST and to calculate the

BPA based on the size changing of mapping data in time intervals. The study provides with the way of context inference for early detecting a local event that affects the entire area.

References

[1] W. Jiang, Y. Deng and J. Y. Peng, "A New Method to Determine BPA in Evidence Theory", Journal of Computers, vol. 6, no. 6, **(2011)**, pp. 1162-1167.

[2] Z. Yang Zhou and Chunjie, "Belief function to probability: A tradeoff between easy decision-making and high risk", Control and Automation (ICCA), 2013 10th IEEE International Conference, **(2013)** June, pp. 233-237.

[3] J. Cao, X. Meng, J. Wang and N. Zhu, "Information Fusion Based on the Improved D-S Theory in the Speaker Tracking System", Journal of Computational Information Systems, vol. 9, no. 9, **(2013)**, pp. 3571-3578.

[4] X. Deng and Y. Deng, "Multisensor Information Fusion Based on Dempster-shafer Theory and Power Average Operator", Journal of Computational Information Systems, vol. 9, no. 16, **(2013)**, pp. 6417-6424.

[5] R. Feng, S. Che, X. Wang and N. Yu, "Trust Management Scheme Based on D-S Evidence Theory for Wireless Sensor Networks", International Journal of Distributed Sensor Networks, **(2013)**, pp. 9.

[6] A. M. Lonea, D. E. Popescu and H. Tianfield, "Detecting Distributed Denial of Service (DDoS) Attacks in Cloud Computing Environment", International Journal of Computers, Communications& Control, vol. 8, no. 1, **(2013)**, pp. 56-64.

[7] Z. Zuo, Y. Xu and G. Chen, "A New Method of Obtaining BPA and Application to the Bearing Fault Diagnosis of Wind Turbine", International Symposium on Information Processing (ISIP'09), **(2009)** August, pp. 368-371.

[8] B.-Gon Kim, K.-W. Kim and Y.-K. Kim, "A Dynamic Queue Management for Network Coding in Mobile Ad-hoc Network", International Journal of Multimedia and Ubiquitous Engineering, vol. 8, no. 4, **(2013)**, pp. 421-430.

[9] S.-J. Park, T.-K. Yang and Y.-H. Seo, "Development of an Electronic Ordering and Payment System with Embedded Devices", International Journal of Smart Home, vol. 7, no. 4, **(2013)**, pp. 393-402.

[10] J.-Y. Hong, S.-H. Kim and J.-H. Kim, "Cyberspace and Intercultural Strategy", International Journal of Smart Home, vol. 7, no. 3, **(2013)**, pp. 111-120.

Authors

Donghyok Suh, he received the M.S. degrees in computer engineering from Hoseo University in 2005 and the Ph.D. in computer science from Chungbuk National University in 2012. His research interests included in stream data processing and data fusion in wireless sensor network.

Juhye Yook, she received the M.Ed. in 1995 and the Ph.D. in 2000 in Special Education from the University of South Carolina in the U.S. She worked for the Korea Employment Promotion Agency for the Disabled (2000~2004) before working at the Korea Nazarene University up until now. Her research interests are computer/information access and assistive technology.

