A DISTRIBUTED BAYESIAN FRAMEWORK FOR BODY SENSOR NETWORKS

Surapa Thiemjarus, Benny P. L. Lo and Guang-Zhong Yang
Department of Computing, Imperial College London, UK

ABSTRACT – Due to the dynamic nature of Body Sensor Networks (BSNs), both sensor nodes and communication links are prone to noise interference and failure. This makes integration of multi-sensory data for BSNs a significant technical challenge. The purpose of this paper is to describe a noise resilient distributed Bayesian framework for robust distributed sensing. The method relies on the introduction of hidden nodes to incorporate intrinsic redundancies of the sensor network and provides a way of node failure detection. The strength of the framework is demonstrated by experiments on context aware activity recognition.

Key Words – Bayesian networks, hidden node, noise resilience, distributed processing, context awareness

INTRODUCTION

Continuous monitoring of patients under their normal physiological condition is an important area of research for Body Sensor Networks (BSNs). It provides a unique mechanism for context aware sensing with which transient abnormalities and their associated environment factors can be reliably captured. BSN is set to transform the traditional *episodic* way of patient management and allows the prediction of the onset of the adverse events. This is particularly important for the effective management of chronic diseases. For example, many cardiac diseases are associated with episodic rather than continuous abnormalities such as transient surges in blood pressure, paroxysmal arrhythmias, or induced or spontaneous episodes of myocardial ischaemia. These abnormalities are important but their timing cannot be predicted and much time and effort is wasted in trying to capture an "episode" within a controlled environment. Important and even life threatening disorders can go undetected because they occur only infrequently and may never be recorded objectively.

Reliable activity recognition for BSN is an active research topic and a number of classification frameworks have been proposed. They include naïve Bayesian Network (BN) (1), SOM (1), Hidden Markov Models (HMMs) (2), Neural Network (NN) (3) and decision tree (4). For context aware sensing with these approaches, the classification process is typically made by a centralised processing unit based on the information received from the context sensors around the body. Reliable data fusion, in this case, is a challenging task as BSNs are dynamic in nature. Sensor noise, node failure, and motion artefact can introduce significant errors to data inferencing results. Furthermore, continuous sensing may involve a large number of sensors, and directly sending all sensory data to the centralised unit can impose significant burden on resource and bandwidth utilisation. To overcome this problem, it is necessary to introduce a distributed framework both for sensor data fusion and noise resilience. This will permit processing tasks being effectively delegated across the network. By introducing local redundancies, the overall performance of the network in terms of noise resilience and accuracy can be significantly improved. The purpose of this paper is to introduce a distributed Bayesian framework for BSN and we demonstrate how to increase the robustness of the network by the use of hidden nodes.

A DISTRIBUTED BAYESIAN FRAMEWORK

BNs represent a graphical probabilistic model that consists of both qualitative and quantitative components. Qualitatively, their structures are directed acyclic graphs in which directed arcs signify causal dependency among nodes that represent random variables. Quantitatively, they provide a compact representation of joint probability distributions. By following the assumption of conditional independence, the joint probability distribution can be decomposed into a product of distributions involving a smaller subset of variables based on the chain rule.

For calculating the posterior probabilities for BN, the variables are assumed to be conditionally independent given their parent(s). In practical applications, however, this assumption may not always hold when there is high correlation between child nodes. This is particularly true when local redundancy (typically heterogeneous) is introduced to avoid motion artefact and sensor failure. In this case, hidden nodes need to be introduced to neutralise the over-weighted contribution by the dependent child nodes (5). Unlike ordinary BN where the conditional probabilities in the link matrices are obtained from the data distribution, the close-form representation for deriving link matrices does not exist for a BN with hidden nodes. The link matrices associated with the unobservable node have to be learned from the data through an iterative error minimising process. When all observable child nodes can be instantiated, the backward propagation method as proposed by Kwoh *et al* (6) can be used for parameters learning. In this case, the squared error cost function computed at the root node is defined as:

$$\xi = \sum_{i=1}^{|A|} \left(d(a_i) - P'(a_i) \right)^2$$

By following this framework, training the new link matrices when new sensors are introduced will involve only data from the root and those associated with the subnet. A small probabilistic weight is typically added to each element of the link matrices to ensure convergence, avoid local minima and cater for unseen cases.

EXPERIMENTS ON ACTIVITY RECOGNITION

Model Training

To evaluate the performance of the proposed method, a reference dataset was constructed based on the ETH dataset (7) after temporal feature extraction. The ETH dataset contains acceleration sensor readings of major joints of the human body during eight different activities. We used six representative and two highly correlated features in this experiment. To ensure model generalisation, only 10% of the dataset was used for training and the rest was used for model accuracy evaluation. Since the range of the acquired signal for each sensor could be contrastingly different, the overall standard deviation was used as the quantisation scale and each feature was quantised into eight discrete states. Through causality and local dependency analysis, the parameters and structure of the BN can be constructed automatically as shown in Fig. 1. Hidden nodes were inserted when local dependency was detected and the parameters of the link matrices were derived through backward propagation.

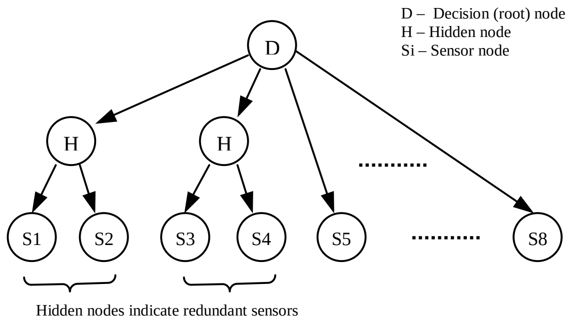

Fig. 1. The structure of the BN with hidden nodes learned from the training dataset.

To assess the numerical stability of the optimisation process, two initialisation schemes were investigated in this study, which include random initialisation and naïve BN initialisation. With naive BN initialisation, the link matrices connecting the hidden nodes and their respective child nodes were initialised with the link matrices of the corresponding naïve BN, whereas the link matrices connecting the hidden nodes and the root node were initialised to an identity matrix. Fig. 2 shows a comparison of the accumulative error over different iterations by using the above and random schemes. It is evident that with the proposed initialisation method, the model accuracy was equivalent to a naïve BN prior to parameters learning. Furthermore, compared to the random initialisation scheme, the training process with naïve initialisation scheme converged significantly faster. This is beneficial in a dynamic environment when new sensors are introduced and the topology of the network is altered.

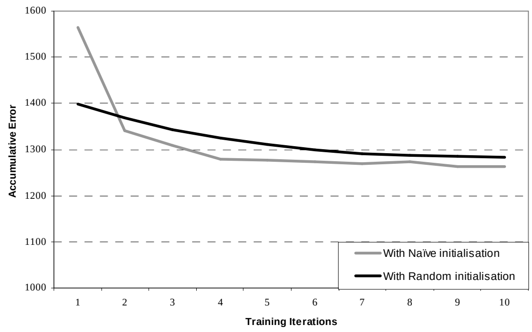

Fig. 2. Comparison of accumulative error over iterations of hidden nodes training with different initialisation schemes

Noise Resilience

To assess the noise behaviour of the model, different levels of Gaussian and white noise were introduced to the above datasets. The Gaussian noise was generated from a normal distribution with zero mean and standard deviation (SD) being different multiples of the feature SD of the original dataset, *i.e.*, ¼, ½, and 1, respectively. Figs. 3 and 4 illustrate the performance comparison of BNs with and without hidden nodes in the presence of single and dual channel noise. It is evident that with single channel noise, the average accuracy of naïve BNs dropped from 74% to 64%, whereas the introduction of hidden nodes ensured the overall accuracy being maintained. When dual channel noise was introduced, the average accuracy of naïve BNs deteriorated further to 53%, whereas the average accuracy of BNs with hidden nodes was maintained at 70%.

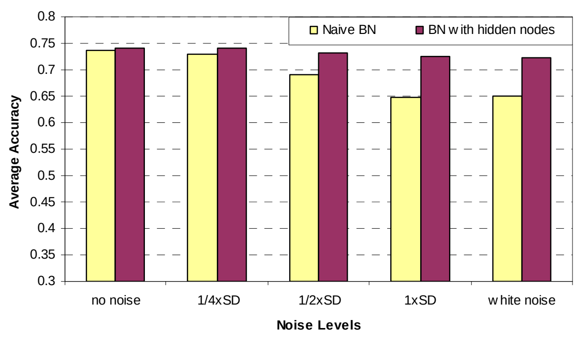

Fig. 3. Performance of naïve BNs and BNs with hidden nodes when different noise levels are introduced in a single redundant channel.

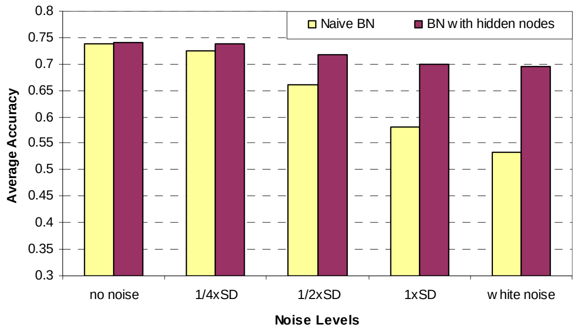

Fig. 4. Performance of naïve BNs and BNs with hidden nodes when different noise levels are introduced in dual redundant channels.

Noise Detection

With the proposed framework, it is also possible to detect node failure. To this end, a measure of node decoupling within each subnet with redundant features is introduced. Provided that there is enough discriminatory information within the subnet to produce a reliable prediction of the hidden states, the noisy node can be detected from the asynchronising child-parent dependency. In the case where the subnet is dominated by noise, noise interference can still be indicated by a continuously high difference in child-parent dependency between two nodes in the subnet. Since the SD of a static signal is usually zero, any dependency measure that relies on SD is difficult to be computed. To avoid this problem, the L1 dependency measure is used for this paper, *i.e.*,

$$\mathrm{Dep}(A,B) = \sum_{AxB} \left| P(a_i \,\&\, b_j) - P(a_i)P(b_j) \right|$$

where the joint probabilities $P(a_i\&b_j)$ and the prior probabilities $P(a_i)$ and $P(b_j)$ are calculated from the data within a shifted window for each node pair within the subnet. By continuous probability updating, online noise detection can be achieved. Fig. 5 demonstrates the difference between child-parent dependency by using a shifted window size of 30 before and after the introduction of 1xSD Gaussian noise into the subnet.

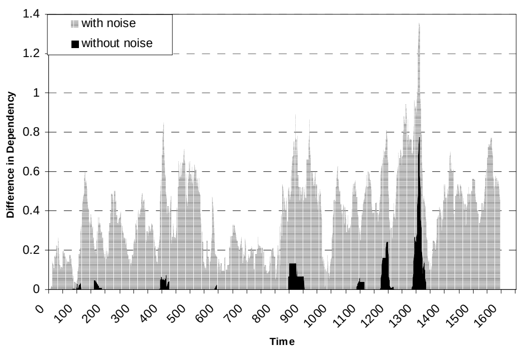

Fig. 5. Comparison of difference in dependency measure before and after 1xSD Gaussian noise is introduced.

The Introduction of Temporal Constraints

Human activity involves body movements that are continuous in nature. By enforcing the smoothness constraint, the model accuracy can be significantly improved. In this paper, this is achieved by averaging the instantaneous model beliefs over a fixed size temporal window. By taken activity transitions into consideration, the test datasets were created such that it simulated nines repetitions of a sequence of eight activities. Recognition was performed with the same model as used in the previous experiment. The average posterior probabilities of the root node over a window of 50 samples was calculated, and with single noise channel the overall accuracy improved from 64-74% to 80-85% for the naïve BNs. For BNs with hidden nodes, the corresponding accuracy improved from 72-74% to ~87%, and there was no decrease in recognition accuracy despite a higher level of noise was introduced.

DISCUSSIONS AND CONCLUSIONS

This paper demonstrates that distributed processing and inferencing with BNs is effective for BSNs. The structure of the network with hidden nodes enables the computations to be distributed and is ideal for utilising the resource in a multi-hop network structure. The experimental results have shown that hidden node can effectively make use of intrinsic redundancy within the subnet to filter out unreliable information to maintain accuracy. In previous work, Paskin and Guestrain (8) illustrated the use of message passing in distributed sensing systems by proposing a robust message passing algorithm which can be used for inferencing in a junction tree representation of a multiply connected BN. The hidden node insertion adopted by this paper is another way of transforming a multiply connected network into a singly connected network, allowing simple Pearl's message propagation algorithm to be used for model inferencing (9).

REFERENCES

1. Van-Laerhoven, K., Kern, N., Gellersen, H.W. and Schiele, B., 2003, IEE Eurowearable
2. Lukowicz, P., Ward, J., Junker, H., Stager, M., Troster, G., Atrash, A. and Starner, T., 2004, PERVASIVE, LNCS 3001, 18-32
3. Randell, C. and Muller, H., 2000, ISWC
4. Bao, L. and Intille, S.S., 2004, PERVASIVE, LNCS 3001, 1-17
5. Sucar, L. E., Gillies, D. F., 1994, IVC, 12, 42-60
6. Kwoh, C. K. and Gillies, D. F., AI, 1-38
7. Kern, N., Schiele, B. and Schmidt, A., 2003, EUSAI
8. Paskin, M.A. and Guestrain, C.E., 2004, UAI
9. Pearl, J., 1998, "Probabilistic Reasoning in Intelligent Systems, Networks of Plausible Inference", Morgan KaufMann Publishers, Inc.