An Integrated Inferencing Framework for Context Sensing

Surapa Thiemjarus, Julien Pansiot, Douglas Mcllwraith, Benny Lo, and Guang-Zhong Yang

Abstract— **This paper presents the use of distributed inferencing with resource optimisation and Spatio-Temporal Self-Organising Map (STSOM) for effectively combining the wearable and ambient sensors. STSOM is an efficient local processing technique which is also suitable for enhancing the temporal behaviour of the distributed inferencing model. To reduce the complexity of the distributed model, a multi-objective Bayesian framework for feature selection has been proposed for model learning. The validation of the techniques has been conducted with activity recognition with both wearable and ambient sensors in a lab-based home monitoring setting.**

I. Introduction

BODY Sensor Networks (BSNs) [1] represent the latest evolution of diagnostic tools from the traditional episodic management to continuous monitoring of patients' physical and biochemical parameters under their natural physiological conditions. This allows the detection of transient but life threatening abnormalities and the early prediction of adverse events. BSNs can also be used for monitoring the health and general well-being of elderly individuals from their Activities of Daily Living (ADL) [2, 3]. In healthcare, BSNs are essential for resource allocation and management. In a mass casualty event, BSNs can be used to prioritise and locate those who will benefit the most from trauma care and rapid surgical intervention [4].

Reliable data fusion in BSNs is a challenging task that has drawn extensive research interests in recent years. Context-aware sensing is an important topic in that the environment under which the BSNs are deployed can change constantly. A context-aware system therefore needs to adapt its functionality according to the changing environment. Information about the context under which the biological and physical signals are collected is also important to the detection of clinically relevant episodes. This is because similar sensory signals can be interpreted differently depending on the context of the patient. One of the pre-requisites of context-aware sensing is reliable feature extraction and multi-sensor fusion.

While the human body is a self-contained system with a complicated internal environment, it can also respond to and interact with the external surroundings in a complex manner. It has been illustrated in previous studies [5, 6, 7] that the use of wearable sensors can provide an effective means of inferring a user's activity. In a home environment, however, there are cases when the use of wearable sensors alone will not be sufficient [8]. Poor recognition accuracy can be observed in certain classes due to the intrinsic ambiguities of the information acquired. To achieve truly pervasive health monitoring, additional information from ambient sensors can be used to enhance the recognition accuracy.

Previous work has shown that data transmission is the most power-hungry task in BSNs [9]. Continuous transmission of raw signals can lead to excessive power consumption, and therefore reduces the battery life. It can also impose a significant burden on other resources such as wireless bandwidth utilisation. Balanced local and distributed inferencing is essential for resolving bandwidth and power consumption issues in BSNs. To address this issue, we adopt a two-tier network architecture based on a hybrid network topology as shown in Fig. 1. At the bottom level, simple configurations of the local sensors (mostly in a star topology) allow low latency data aggregation, and thus are ideally suited for wearable on-body sensors. At the higher level, a mesh network is typically used to ensure scalability and a large spatial coverage. Under the BSN environment, this can be used to link ambient sensors in a home or hospital environment. At this level, the network can rely on existing heterogeneous network infrastructure such as ZigBee, WiFi or GPRS. The real-time requirement at this level is less critical and the data rate is relatively low as the abstraction of the raw data has already been achieved at the lower level.

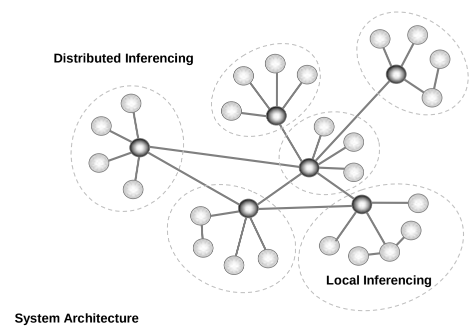

Fig. 1 A system architecture for data inferencing in BSNs.

One of the significant advantages of this two-tier network architecture is in the simplicity of managing mobile users. For patient monitoring, each subnet can be regarded as a collection of sensors worn on the body, where the central node can be regarded as the local processing unit, such as a

Manuscript received February 20, 2008. This work was supported in part by UK DTI SAPHE project, EU FP6 WASP project, and the Ananda Mahidol Foundation.

Surapa Thiemjarus, Julien Pansiot, Douglas Mcllwraith, Benny Lo, and Guang-Zhong Yang are with the Department of Computing, Imperial College London. (corresponding author to provide phone: 0044-20-75948337; fax: 0044-20-75818024; e-mail: st01@doc.ic.ac.uk).

PDA or mobile phone, for aggregating the sensor information and performing local data inferencing. As the patient moves around, this subnet can leave and re-join different parts of the top-tier network, thus permitting seamless patient monitoring anywhere, everywhere. Fig. 2 illustrates how a BSN moves around and re-connects to different parts of the ambient network.

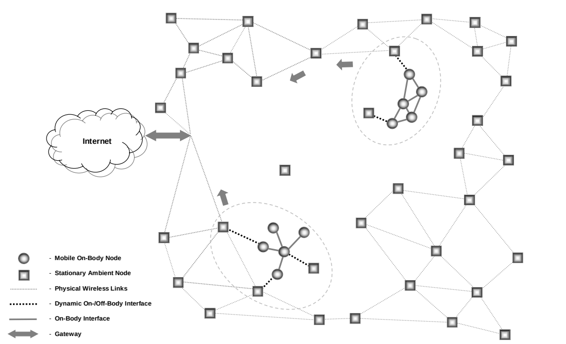

Fig. 2 Inferencing with wearable and ambient sensors when the subject moves around in an ambient sensing environment.

In our previous studies, we discussed the technical challenges of context aware sensing and proposed a number of machine learning techniques for context recognition for BSNs [10]. The techniques include the use of feature selection for optimal sensor placement [11], an autonomic sensing framework for distributed inferencing [12], and a hierarchical Self-Organising Map (SOM)-based model for incorporating temporal information [1, 13]. Each technique has been individually validated by experiments on activity recognition with wearable sensors In this paper, we will present the full integration of the previously proposed techniques and demonstrate how to use the framework for effectively combining the wearable and ambient sensors.

II. An Integrated Inferencing Framework

The design of the integrated inferencing framework can be summarised as follows. In the first layer, the raw signals are locally pre-processed and features extraction is performed with appropriate feature detectors. In the second layer, the local inferencing models generate quantised outputs from the continuous input features. In the third layer, the distributed model aggregates outputs of the local inferencing model to infer the final context decision. This section provides brief descriptions of the local and distributed inferencing models used in the framework.

A. A Model for Local Inferencing

Since the processing power of each sensor node is limited, only simple algorithms are suitable for on-board implementation. To incorporate the use of temporal information into the integrated inferencing framework, we propose the use of Spatio-Temporal Self-Organising Maps (STSOMs) [1, 13] for local inferencing.

STSOM extends the traditional SOM with a temporal layer and an adaptive mechanism for class separation and node expansion. Based on the signal characteristics of the input data and the sequence of node activations produced by the bottom layer (static map) of the STSOM, the corresponding temporal behaviour can be extracted and fed into the dynamic layer (dynamic map) of the STSOM.

The introduction of a temporal layer simplifies the recognition of cyclic patterns and enhances the discrimination power of the network while in the same time significantly reducing the number of neurons involved. The simplicity of the processing involved means the technique is ideally suited for local processing where on-board computational resource is limited. STSOM can therefore be used in the pre-processing step for extracting high level features and enhancing the temporal behaviour of the distributed inferencing model.

B. A Model for Distributed Inferencing

In [12], we have presented a framework for developing a distributed inferencing model for context recognition with BSNs. A novel approach to model learning and inferencing based on feature analysis was proposed. Central to the proposed concept of distributed inferencing is the use of probabilistic graphical models for data abstraction and inferencing with belief propagation. The key elements of the proposed framework include dependency graph construction by using feature selection, causal direction assignment based on dependency analysis, computation of model parameters, creation of the factor graph representation, and model inference by the use of the sum-product algorithm [14]. The model structure is formulated by learning the dependencies and causalities among the user context and observed variables. The separation in the representation of the actual measurements and the functional storage in a factor graph facilitates the mapping of a logical structure onto a physical network, and thus greatly enhances the practical value of the proposed technique. A parallel message passing scheme has been proposed for the FG inference engine. A summary of steps involved in model construction is illustrated in Fig. 3.

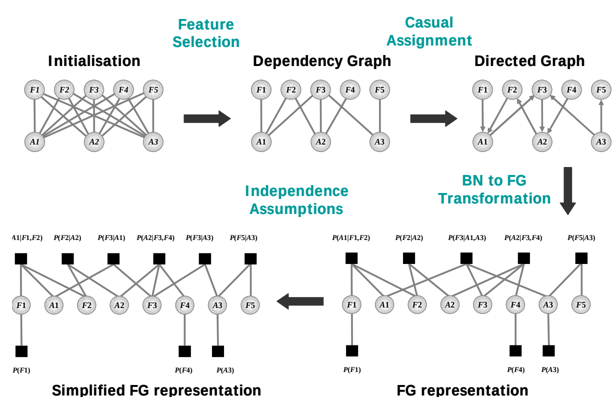

Fig. 3 An illustration of distributed model construction

In BSN applications, a good feature depends on a number of factors including the quality and availability of the sensors, as well as the communication costs between them. A multi-objective feature selection framework has been developed. The framework extends the original Bayesian Framework for Feature Selection (BFFS) algorithm [15] by differentiating redundant and irrelevant features through a numerically efficient optimisation strategy. Besides feature selection, a successful use of multi-objective BFFS for defining model structure has also been demonstrated in [12].

III. Experiments and Results

A. Data Acquisition and Feature Extraction

To assess the performance of the proposed integrated framework, a similar experiment setting as presented in [8] was used for collecting data from more subjects. Both wearable and ambient sensors were used. The datasets were collected from 5 subjects performing 9 different activities in a lab-based home environment. The activities include 1) walking, 2) standing, 3) standing (head tilted), 4) sitting, 5) reading, 6) eating, 7) sitting (sofa), 8) lounging, and 9) lying down. Each activity lasted for approximately one minute. The sampling rate used in this experiment was 25 Hz and each dataset was divided into 80% for training and 20% for evaluation.

For wearable sensing, an ear-worn Activity Recognition (e-AR) sensor is used [16]. The sensor consists of a 3-axis accelerometer, allowing angular acceleration, as well as linear acceleration and gravity to be detected. The output of the accelerometer is a combined measure of all the above static and dynamic components. The orientation of the sensor can be detected when the sensor is static. This allows the estimation of body posture and activity changes. In addition to the acceleration values, we also extract the tilt of the sensor and a moving window Fast Fourier Transform (FFT) [17] features from the wearable sensor.

For ambient sensing, blob-based vision sensors were used. With this technique, the captured video is processed on-node in real-time so as to remove any appearance information of the subject being monitored. This follows the concept of '*from blobs to personal metrics to behaviour profiling*' [18, 19] such that sufficient information about the user's activity is captured without exposing his/her appearance details. The silhouette of the subject, also known as a *blob*, are first extracted from the video signal using an Gaussian Mixture Model-based statistical background segmentation scheme [20]. The binary blob provides the information about the subject overall shape and articulation of the body. This information alone, however, may not provide enough information for the detection of detailed activities. Therefore, local motion of the subject is estimated using optical flows (apparent motion of the visual features within the images).

Due to the high dimensionality of the images and varying spatial relationship between the moving subject and the sensor under perspective projection, a direct use of the information from the binary blob and optical flow is not appropriate. Principal Component Analysis (PCA) [21] is applied to the region of interest to derive the Oriented Bounding Box (OBB), from which blob-based features can be calculated.

A summary of the features extracted from the raw feature set is provided in Table 1. It can be seen that in the second input layer, there exist 101 features in total. Features 1-95 are extracted from the wearable sensor and the rest are from the blob sensor.

Table 1
Features Extracted From The Wearable And Blob Sensors

ID	Feature
1	Acceleration on the Z axis
2	Acceleration on the X axis
3	Acceleration on the Y axis
4	Head tilt on Z axis
5	Head tilt on X axis
6-35	FFT coefficients on Z axis
36-65	FFT coefficients on X axis
66-95	FFT coefficients on Y axis
96	Speed estimation
97	Aspect ratio
98	Subject height estimation
99	Subject optical flow intensity
100	Subject optical flow correlation
101	Subject optical flow aspect ratio

B. Results

The integrated framework was then applied to the training dataset of each subject to derive subject-specific models. For efficiency, each set of FFT coefficients was clustered with a single STSOM, whereas other features were used to train each STSOM individually. As a result, a total of 14 STSOMs were created. In this experiment, a static map with 64 neurons and a dynamic map with 25 neurons were used. Since each STSOM contains a static map and a dynamic map, the full dimensionality of the feature vector in the third input layer is 28 (one quantised feature channel for each map). Node labels were assigned by using the frequency of the class activation. In this way, each feature channel in the third input layer can have a maximum number of states corresponding to the class number. With this method of label assignment, each STSOM can be seen as an on-node inference engine that provides the classification results based on the available input locally. For simplicity, all STSOMs in this experiment were created by one pass.

Since the datasets consisted of 9 activities, 9 binary decision nodes were used to represent the presence or absence of each activity. The most informative features for the classification of each activity were derived using the multi-objective BFFS algorithm with the following evaluation function:

$$\mathrm{D}_r\left(f_i\right) = -\left(1-\omega_1\right)\times E_{AUC}\left(\mathcal{G}^{(k)} - \left\{f_i\right\}\right) + \omega_1 \times E_{AUC}\left(f_i\right) \quad (1)$$

where ω_1 is the weighting factor from 0 to 1, $E_{AUC}(\;)$ is the function which returns the expected AUC given by its parameters, and $\mathcal{G}^{(k)} = \{ f_i,\ 1 \le i \le k \}$ denotes the feature set at the beginning of the iteration. With this bi-objective function, multi-objective BFFS takes into account both the discriminatory power and the amount of redundant information in the feature set for the selection process. To reduce the search time, variables ks and ω_1 were fixed to 2 and 0.1, respectively. The selection of feature set associated with each classification was made based on the criteria that an extra feature is added to the feature set with AUC over 0.99.

Consequently, the dependency graph can be directly derived. The results show that, compared to the initial fully connected model, the number of dependency links in the models reduced from 252 links to only 23, 29, 26, 23 and 26 for subjects 1 to 5, respectively. As the dependencies between feature and decision nodes are discarded, smaller numbers of conditional probability distributions are required, thus greatly simplifying the potential computation and inter-node communication involved.

Fig. 4 summarises the static and dynamic maps required for the distributed models for each subject. In addition to a significant reduction in size and model parameters in the distributed models, only half of the entire maps are used in the distributed models and not all the continuous features need to be calculated. Out of the 14 groups of continuous features, only 8 to 10 channels are required. It can be seen that for the derived inference model, the optical flow correlation feature is in fact not required at all. This indicates a significant amount of saving in computational resources compared to a fully connected model.

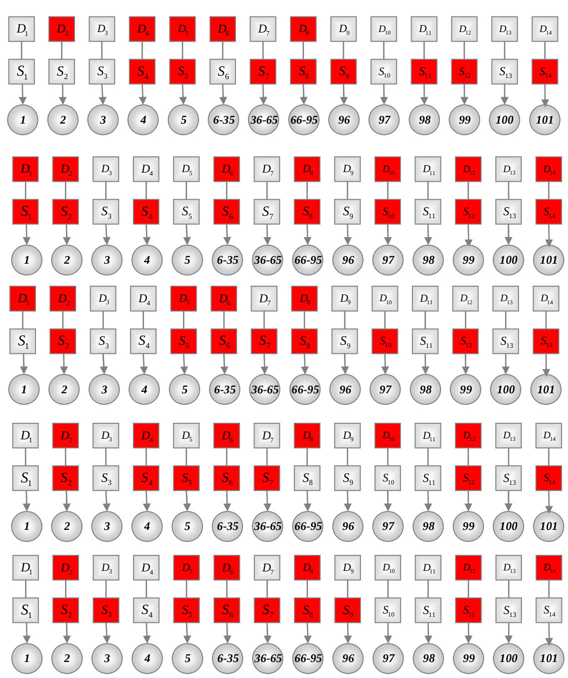

Fig. 4 A summary of the static and dynamic maps used in the distributed model (the red colour indicates map selection) for the five subjects studied in this chapter.

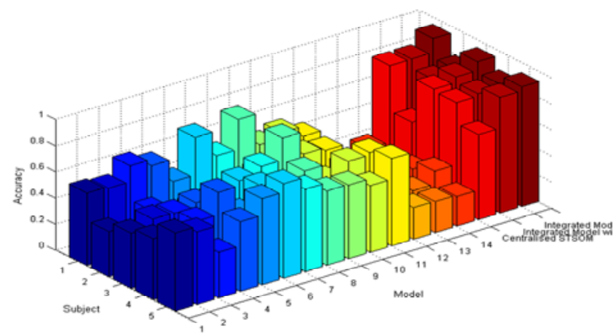

Fig. 5 A comparison of the classification accuracy of the 14 local STSOMs, the centralised STSOM, and the integrated models with and without feature selection for the five subjects studied.

To evaluate the classification performance, the integrated models are compared to those with fully connected inference model based on using all the low level features. By using feature selection for model simplification, the average accuracy only decreased modestly from 88.61% to 83.43% (~5%), whereas the amount of computational savings achieved is about 90%.

The proposed model can be viewed as a combining classifier which uses a higher probabilistic model for integrating the output of the local experts (STSOMs) to achieve a better accuracy. Fig. 5 shows a comparison of the classification results for different models of the test datasets. It can be observed that the integrated models (with and without feature selection) yield a better accuracy than individual STSOMs. It outperforms the centralised STSOM of the same dimension for all subjects. Similar to the problem found in other neural network models, with random weight initialisation, the training of an STSOM can sometimes be trapped in local minima. Model selection is normally required to solve this problem. Since the training of STSOMs in this experiment is achieved in a single pass, this could be the reason for the low classification accuracy of the centralised STSOM for Subjects 2 and 5. Due to the high dimensionality of the input vectors to the second layer and the aforementioned local minima problem, STSOM only yields a classification accuracy of 71.95% on average. This, however, does not seem to affect the accuracy of the combined model, thus illustrating the robustness of our proposed technique.

IV. Conclusion

In this paper, we have proposed a framework for integrating on-node processing with distributed inferencing for combining both wearable and ambient sensors in a laboratory based home activity monitoring experiment. In this framework, we have used STSOM for local inferencing and the sum-product algorithm in a probabilistic graphical model for distributed inferencing.

Different layers of the data processing pipeline have been described. These include the extraction of indices from the 3-axis accelerometer and image features from the video sequence. STSOM is then applied to each subset to extract both the static and dynamic contents. The experimental results have shown that STSOM can provide an effective abstraction of the high dimensional data, thus resulting in significant savings in data transmission. The local classification results of the selected STSOMs can then be combined by using the distributed inferencing model for final classification.

It has been shown that the multi-objective BFFS algorithm is effective in learning the structure of the distributed model and for achieving overall resource optimisation. Significant savings in terms of the number of features, the number of STSOMs, and the number of node connections to be used can be achieved only with minor deterioration in the classification accuracy.

The experimental results of the integrated model show that the distributed model provides an effective means of combining individual STSOMs, which is reflected from the consistently high classification accuracy achieved.

REFERENCES

[1] G.-Z. Yang, *Body Sensor Networks*. London: Springer-Verlag, 2006.

[2] M. P. Lawton, "Environment and Other Determinants of Well-Being in Older People," *The Gerontologist*, vol. 23, pp. 349-357, 1983.

[3] M. P. Lawton, "Aging and Performance of Home Tasks," *Human Factors*, vol. 32, pp. 527-536, 1990.

[4] D. Malan, T. Fulford-Jones, M. Welsh, and S. Moulton, "CodeBlue: An *ad hoc* Sensor Network Infrastructure for Emergency," *Proceedings of the First International Workshop on Wearable and Implantable Body Sensor Networks*, Imperial college London, United Kingdom, 2004.

[5] J. A. Ward, P. Lukowicz, G. Troester, and T. Starner, "Activity Recognition of Assembly Tasks Using Body-Worn Microphones and Accelerometers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, pp. 1553-1567, 2006.

[6] K. Van Laerhoven and H.-W. Gellersen, "Spine versus Porcupine: A Study in Distributed Wearable Activity Recognition," *Proceedings of the Eighth International Symposium on Wearable Computers*, pp. 142-149, 2004.

[7] L. Bao and S. S. Intille, "Activity Recognition from User-Annotated Acceleration Data," *Proceedings of the Second International Conference on Pervasive Computing*, Vienna, Austria, pp. 1-17, 2004.

[8] J. Pansiot, D. Stoyanov, D. Mcllwraith, B. P. L. Lo, and G.-Z. Yang, "Ambient and Wearable Sensor Fusion for Activity Recognition in Healthcare Monitoring Systems," *the Fourth International Workshop on Wearable and Implantable Body Sensor Network*. RWTH Aachen University, Germany, vol. 13, pp. 208-212, 2007

[9] V. Shnayder, M. Hempstead, B.-R. Chen, G. W. Allen, and M. Welsh, "PowerTOSSIM: Efficient Power Simulation for TinyOS Applications," *Proceedings of The ACM Conference on Embedded Networked Sensor Systems*, Baltimore, Maryland, 2004.

[10] S. Thiemjarus, B. P. L. Lo, and G.-Z. Yang, "Context Aware Sensing - What's the Significance?," *Proceedings of the Perspective in Pervasive Computing*, pp. 163-170, 2005.

[11] S. Thiemjarus, B. P. L. Lo, and G.-Z. Yang, "Feature Selection for Wireless Sensor Networks," *Proceedings of the First International Workshop on Wearable and Implantable Body Sensor Networks*, Imperial College, London, 2004.

[12] S. Thiemjarus and G.-Z. Yang, "An Autonomic Sensing Framework for Body Sensor Networks," *Proceedings of the Second International Conference on Body Area Networks*, Florence, Italy, 2007.

[13] S. Thiemjarus, B. P. L. Lo, and G.-Z. Yang, "A Spatio-Temporal Architecture for Context-Aware Sensing," *Proceedings of the Third International Workshop on Wearable and Implantable Body Sensor Networks*, MIT Media Lab, pp. 191-194, 2006.

[14] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor Graphs and the Sum-Product Algorithm," *IEEE Transactions on Information Theory, Special Issue on Codes on Graphs and Iterative Algorithms*, vol. 47, pp. 498-519, 2001.

[15] G.-Z. Yang and X. P. Hu, "Multi-Sensor Fusion," *Body Sensor Networks*, G.-Z. Yang, Ed. London: Springer-Verlag, pp. 239-286, 2006.

[16] O. Aziz, B. P. L. Lo, R. King, A. Darzi, and G.-Z. Yang, "Pervasive Body Sensor Network: An Approach to Monitoring the Post-Operative Surgical Patient," *Proceedings of the Fourth International Workshop on Wearable and Implantable Body Sensor Networks*, MIT Media Lab, pp. 13-18, 2006.

[17] J. W. Cooley and J. W. Tukey, "An Algorithm for the Machine Computation of the Complex Fourier Series," *Mathematics of Computation*, vol. 19, pp. 297-301, 1965.

[18] J. Wang, B. P. L. Lo, and G.-Z. Yang, "Ubiquitous Sensing for Posture/Behavior Analysis," *Proceedings of the Second International Workshop on Body Sensor Networks*, Imperial College, London, pp. 112-115, 2005.

[19] G.-Z. Yang, B. P. L. Lo, J. Wang, M. Rans, S. Thiemjarus, J. W. P. Ng, P. Garner, S. Brown, B. Majeed, and I. Neid, "From Sensor Networks to Behaviour Profiling: A Homecare Perspective of Intelligent Building," *Proceedings of the IEE Seminar for Intelligent Buildings*, 2004.

[20] C. Stauffer and W. E. Grimson, "Learning Patterns of Activity using Real-Time Tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 747-757, 2000.

[21] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. Berlin: Springer-Verlag, 2002.