

# Analysis and Cluster Based Modelling and Recognition of Context in a Mobile Environment

Agathe Battestini and John A. Flanagan

Nokia Research Center  
P.O. Box 407, FIN-00045, NOKIA GROUP, Finland  
agathe.battestini@nokia.com, adrian.flanagan@nokia.com

**Abstract.** <sup>1</sup>Our intuitive understanding of context has been formulated by Dey, as the information describing the situation of an entity. In the case of simple information sources, determining the relevant information describing the situation or context of an entity is straightforward. This is not the case when several information sources can be used, essentially to generate high level contexts. We set out to show that extracting the relevant context information from the combination of multiple information sources can be achieved by considering contexts as clusters in the data. We demonstrate this by analyzing a set of real measured data. A cluster structure in the data is visualized and it is shown how different user contexts are associated with different clusters. The cluster structure of context data can be modelled using a dynamic mixture model which gives insight into which properties a clustering algorithm should have in order to be used in a context recognition application.

## 1 Introduction

In order to understand context and develop applications that can be used for context recognition the first most important step is to define context. Dey's definition of context [1], where context is the information that can be used to describe the situation of an entity, is a formal statement of our intuitive understanding of context. However, in order to understand this definition of context and apply it in a practical situation we are obliged to define information. In the case where only a single information source is considered, for example location, then the information in the source is easily identified. On the other hand when given a set of information sources that could potentially be used to characterize a user's situation the problem remains of identifying which information from the different sources is relevant in describing the context. In other words, Dey defines context as information which for a simple information source is obvious and can be considered low level context but for several sources the definition of relevant information or high level context is not obvious. In what follows we investigate how high level contexts can be generated by combining or fusing the output of several different information sources and our proposition is to define contexts as

---

<sup>1</sup> Appeared in Proc. of MRC 2005, Edinburgh, Scotland, 2005, pp 85–96

clusters in the data. This leads to a very practical definition of context in the case where information is available from several different sources. Furthermore algorithms for clustering data are robust in the presence of noise and hence we can also include the case of imperfect information sources in our definition of context. To illustrate what this proposition means in practise, a set of recorded context sensor data is analyzed. First the direct correlation between the outputs of the information sources is calculated, a basic indicator of a cluster structure in the data, and shown to vary between different user contexts. Second, the data is clustered in an unsupervised manner using the Self-Organizing Map (SOM) [2] in such a way that the cluster structure can be visualized using the U-Matrix representation [3]. In this visual representation it is clear that data samples from the same user context are classified to the same cluster.

Data clustering [4] is a widely used method for feature extraction from noisy data and as such, based on our proposal, can be used for context recognition where the extracted features correspond to high level contexts. As pointed out in [5], [6] many current context aware applications make unrealistic assumptions about the quality of the available context information. Typically these approaches, often based on ontologies [7], [8], [9] make the inherent assumption that the lower-level context information, typically coming from sensors, is error-free. Hence, our clustering proposal based on data clustering algorithms which can deal with very noisy and corrupted information in a robust manner can be considered a novel approach to context recognition. Clustering of sensor data with application to context recognition has been described by [10]. In this case the clustering was mostly used as a preprocessing stage to generate cues that could be used in symbolic Artificial Intelligence (AI) techniques and not considered as a means of fusing information sources to generate higher level contexts.

In Sec. 2 we briefly describe the set of measured sensor signals used in our cluster analysis and the different user contexts in which they were recorded. In Sec. 3 we look at some of the individual sensor signals and their characteristics. Here we identify and describe some different user contexts recorded in the data. The correlation and cluster structure of the result of fusing the data from the different sources is analyzed in Sec. 4. In Sec. 5, based on the discussion in Sec. 4 a dynamic mixture model is suggested as an appropriate means of modelling context data. In a very brief manner it is pointed out that data which is modelled by a dynamic mixture model presents problems for classic clustering algorithms. Sec. 6 contains the conclusions.

## 2 Sensor Signal Data Set

One of the contributing reasons to the increasing interest in context awareness is the availability of small, cheap, low power sensors that can be distributed in an environment or on the user. These sensors can be used to sense different environmental characteristics such as temperature, acceleration, humidity etc. Location

in terms of GSM Cell IDs and location area codes<sup>2</sup> can also be considered as a signal, in essence, sensed by a mobile phone. However, despite the fact that such sensor signals are considered as the basis of context awareness, surprisingly little real measured data exists. The Nokia Context Dataset is a set of measured context data and is available at [11]. In what follows the raw signals of this data set, publicly, available at this location and described in [12] is analyzed. However, the public version of this data is only available in a quantized form. A complete description of the data is available in [11], but in brief, the data set consists of a set of feature files for 43 different recording sessions numbered from 0–42. In each recording session the same user carried a mobile phone, sensor box and laptop PC, going from home to the workplace or vice-versa. During the journey the user walks, takes a bus and Metro and sometimes uses a car. On occasion slightly different or very different routes/modes of transport are taken. During the session, sensors recorded 3-axis acceleration, atmospheric pressure, temperature, humidity etc. The ambient audio was recorded on the laptop using a microphone and sound card. On the mobile phone, the user’s location was recorded as Cell ID and Location Area Code (LAC) as defined by the GSM network. After each recording session the sensor signals were low pass filtered where appropriate and sampled once every second. The audio signal energy was furthermore averaged over a 10 second interval and the average sampled every second. Similarly the accelerometer signal magnitude was averaged and sampled every second to give the average user activity. In what follows different signals and sets of signals from these recordings are analyzed.

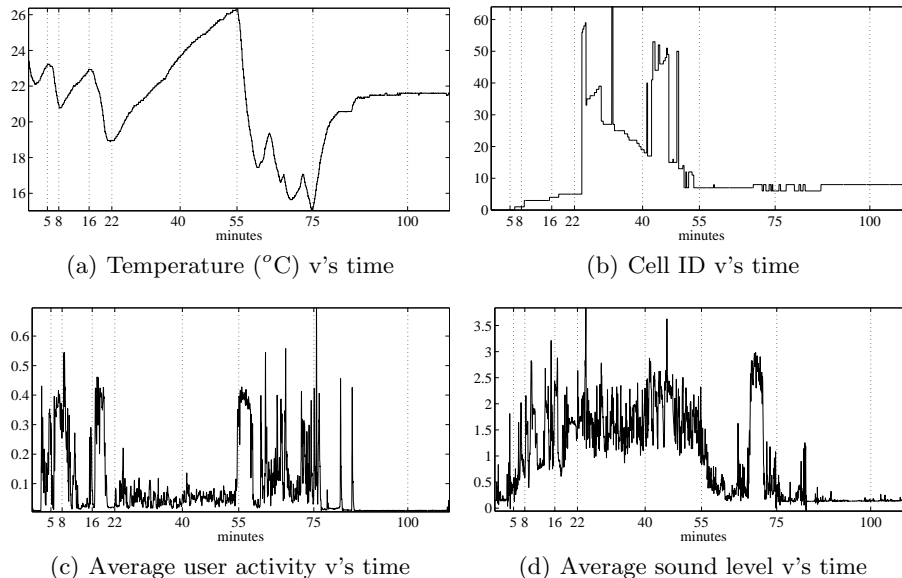
### 3 Single Sensor Signals

In order to appreciate the diverse nature and characteristics of the individual sensor signals it is useful to look at an example subset of sensors in a single recording session. Figure 1 shows the temperature, Cell ID, average user activity and average background sound level during recording session 35 in the Nokia Context Dataset. Referring to Fig. 1 (a) it is possible to identify different 7 distinct user contexts which are summarized in Table 1. The user is inside from 0–5 minutes and then briefly walks outside to the metro station between 5–8 minutes. From 8–16 minutes the user is inside the station and takes the Metro. From 16–22 minutes the user walks to the bus and waits at the bus stop. Between 22–55 minutes the user is in a crowded bus. From the bus the user walks outside again, briefly walking into to a shop and then continuing home arriving there at approximately 75 minutes. Some time after 75 minutes the sensor box is removed from the pocket and placed on a table. The recording took place during the winter with a substantial difference between the inside and outside temperature. It is clear there is a certain time lag between a change in real temperature (i.e. between inside and outside) and the time it takes the thermometer to reach this real temperature. The reasons for this is due to the location of the thermometer

---

<sup>2</sup> In an urban area a cell may have a radius of several 100 meters with a LAC consisting of several cells.

in the user's outside pocket which could be considered semi-insulated. This is compared to the temperature variation after 75 minutes where it would seem that when the sensor box is left on the table and in contact with free air then the temperature stabilizes much quicker.



**Fig. 1.** Variation of temperature, Cell ID, average user activity and average sound level during the single recording session no. 35 from the Nokia Context Dataset at [11].(See Table 1).

In the measurement of location, as determined by the Cell ID of the GSM network, Fig. 1 (b) shows a plot of the changes in the Cell ID for the same recording session. After 75 minutes when the physical location of the device is static it is clear however the Cell ID changes in a somewhat random manner between Cell IDs 7,9 from 75 to approximately 85 minutes and then remains stable. This of course is related to the the conditions of the GSM network, GSM network load, and radio transmission environment. Figure 1 (c) shows the average user activity, as determined from the accelerometer, during the recording. The activity feature seems quite noisy and even during the bus trip when the user is stationary there is an elevated activity level, obviously related to the movement and vibration of the bus. However, even when walking the activity level changes quite significantly. Figure 1 (d) shows the average sound level during the recording. The randomness of this signal is not unlike that of the average user activity.

Table 2 shows the mean and variance of the average user activity and sound levels in the different user contexts. It is clear that changes in the characteristics

**Table 1.** Different user contexts and the associated time intervals for session 35 of the Nokia Context Dataset at [11]. See Fig. 1.

Context No.	1	2	3	4	5	6	7
minutes	0–5	5–8	8–16	16–22	22–55	55–75	>75
Detail	Inside office	Walk outside to Metro	Inside Metro station, travelling on Metro	Walk to and wait at bus stop	Sit in a crowded bus	Walk outside, briefly inside	Arrive home

of a signal are correlated with changes in the user’s context. For example the average user activity in contexts 1, 2, 3, 4, 6 when the user is moving is greater than when the user is in the bus in context 5, which is in turn greater than when the user is at home. Similar difference between the average and variance of the sound level in different user contexts is also noticeable. For example there is a factor of 10 difference in the mean and variance between contexts 5, 7, the bus and home context.

**Table 2.** Mean and variance of the average user activity and average sound levels for different user contexts from session 35. See Fig. 1.

Context No.		1	2	3	4	5	6	7
Activity level	mean	0.1017	0.2747	0.1250	0.1826	0.0541	0.1731	0.0318
	var	0.0133	0.0157	0.0173	0.0285	0.0014	0.0161	0.0056
Sound level	mean	0.0003	0.0007	0.0014	0.0015	0.0017	0.0009	0.0002
	var ( $\times 10^{-6}$ )	0.0552	0.1559	0.3430	0.2373	0.1915	0.6905	0.0146

## 4 The Fusion of Multiple Sensor Signals

It is straightforward to detect the variation in the characteristics of a single sensor signal during changes of a user context. However it is interesting to analyze the correlation between the changes in the characteristics of several sensor signals and changes in user contexts. In fact the combination of a group of sensors or sources can be defined as a higher level context. In order to obtain a large data sample and an even distribution of different types of recording sessions the data from 12 sessions numbered 27, 28, 30, 31, 34, . . . , 39, 41, 42 in the Nokia context dataset are used. These sessions are pairs of ”going to” and ”coming from work” recording sessions on the same day. For example session 40 is a coming from work recording without an associated going to work recording session on the same day and so is not included. During the following analysis the values of the Cell ID, LAC, temperature, relative humidity, dew point, average user activity, average

sound level are used to form a 7 dimensional data vector, the numerical values denoted by  $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5, x_6, x_7) \in \mathbb{R}^7$ , at every second of the recording. The vector  $\mathbf{x}$  represents the fusion of the information sources. Over the 12 recording sessions this corresponds to a total of over 59,000 data samples corresponding to over 16 hours of recorded context data. Note, that strictly speaking, representing Cell ID and LAC by real numbers as part of a vector is dubious in general. However, in the case of these recordings it turns out for the most part, rather by accident than design, the consecutive Cell ID and LAC numbers do actually indicate the relative physical proximity of the Cells and LACs to each other as can be understood from Fig. 1 (b) and for this reason only, they are treated as real numbers. Each of the variables  $x_i, i = 1, \dots, 7$  is normalized before processing by removing the average and dividing by the variance. Table 3 shows the cross correlation matrix between the different variables over all data samples and hence user contexts. Some general observations include an expected general high correlation of 0.7 between  $x_3, x_5$  the temperature and dew point and 0.87 between  $x_4, x_5$  the humidity and dew point. The average sound level is generally significantly correlated with all variables except the average user activity. On the other hand the average user activity is small but significantly negatively correlated with the relative humidity and dew point. These correlations are over the whole data set and do not reflect the correlations that may exist between variables in different user contexts. In Table 4 the correlation coefficients between the  $x_i$ 's in context 2 (8–25 minutes) of session 35 are illustrated, corresponding to the user walking to the Metro station. Clearly now there is a change in the correlation coefficients with respect to the general case and it seems that more of the variables are strongly positively correlated. The correlation is strong between the Cell ID and LAC as the distance covered is quite small which also means these two variables are strongly correlated with all other variables. Apart from the Cell ID and LAC compared to the general case of Table 3 there is now a strong negative correlation of  $-0.53$  between the temperature and user activity. This most likely reflects the case of the user walking outside and being stationary when inside. The strong positive correlation between average sound level and user activity is also noticeable reflecting also that the user is walking outside on a busy street. Table 5 shows the correlation coefficients between the  $x_i$ 's in context 5 (22–55 minutes) of session 35, corresponding to the user in a moving bus. Clearly now there is a change in the correlation coefficients with respect to both the general case and the case of Table 4. As an example, there is a decreased correlation between the Cell ID and LAC compared to the previous 2 cases, most likely due to the fact that context 5 is when the user is travelling in a bus. There is a very large change in the correlation between temperature and average sound level at  $-0.52$  in context 3 to  $0.17$  in context 5. It is clear from further comparison of the correlations that they do change, sometimes quite significantly between different user contexts.

These cross correlations between the  $x_i$ 's in different contexts are now interpreted as the existence of clusters in the data [4]. One well known approach to visualization of a cluster structure in high dimensional data is based on the Self-

**Table 3.** Cross correlation coefficient matrix for the sensor data from the 12 recording sessions with  $(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$  representing (Cell ID, LAC, Temperature, Relative Humidity, Dew Point, average User Activity, average Sound Level) respectively.

	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$
$x_1$	1.00	0.53	0.11	0.35	0.32	-0.10	0.43
$x_2$	0.53	1.00	0.25	0.48	0.50	-0.06	0.35
$x_3$	0.11	0.25	1.00	0.27	0.70	0.00	0.20
$x_4$	0.35	0.48	0.27	1.00	0.87	-0.15	0.42
$x_5$	0.33	0.50	0.70	0.87	1.00	-0.12	0.41
$x_6$	-0.10	-0.06	0.00	-0.15	-0.12	1.00	0.02
$x_7$	0.43	0.35	0.20	0.42	0.41	0.02	1.00

**Table 4.** Cross correlation coefficient matrix for the sensor data recording session 35 for the time interval 5 – 8 minutes, with  $(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$  representing (Cell ID, LAC, Temperature, Relative Humidity, Dew Point, average User Activity, average Sound Level) respectively.

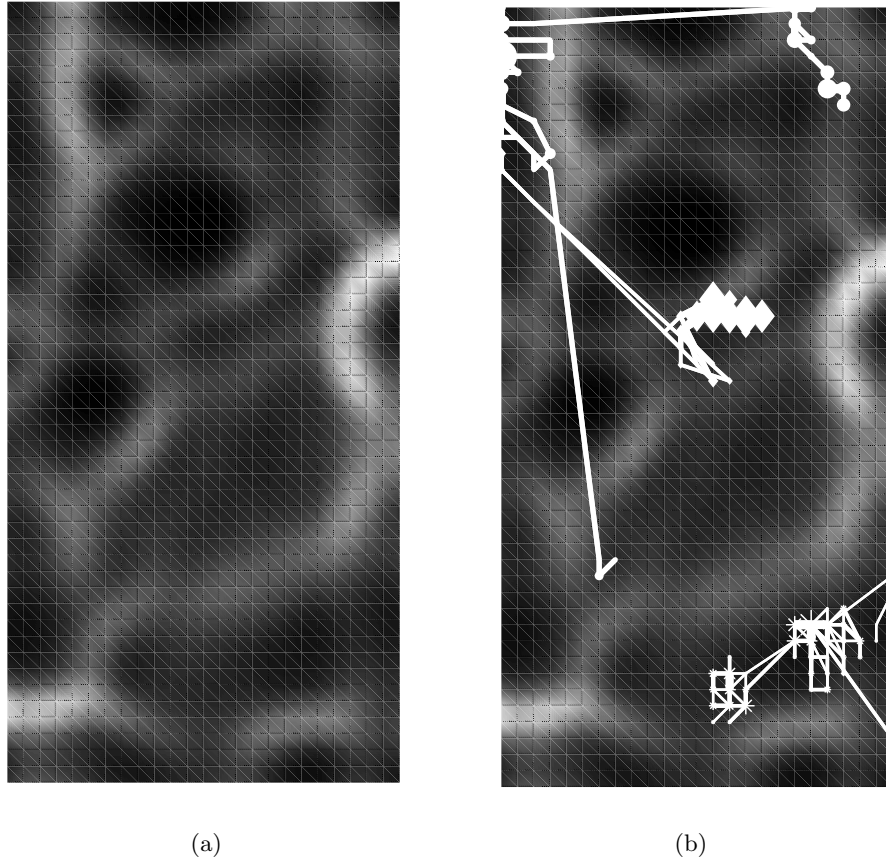
	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$
$x_1$	1.00	1.00	-0.69	0.16	-0.31	0.75	0.60
$x_2$	1.00	1.00	-0.69	0.16	-0.31	0.75	0.60
$x_3$	-0.69	-0.69	1.00	-0.71	-0.22	-0.53	-0.52
$x_4$	0.16	0.16	-0.71	1.00	0.84	0.21	0.43
$x_5$	-0.31	-0.31	-0.22	0.84	1.00	-0.14	0.17
$x_6$	0.75	0.75	-0.53	0.21	-0.14	1.00	0.59
$x_7$	0.60	0.60	-0.52	0.43	0.17	0.59	1.00

**Table 5.** Cross correlation coefficient matrix for the sensor data recording session 35 for the time interval 22 – 55 minutes, with  $(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$  representing (Cell ID, LAC, Temperature, Relative Humidity, Dew Point, average User Activity, average Sound Level) respectively.

	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$
$x_1$	1.00	-0.02	-0.27	-0.01	-0.21	0.17	0.21
$x_2$	-0.02	1.00	-0.07	0.38	0.07	-0.20	-0.11
$x_3$	-0.27	-0.07	1.00	0.72	0.98	-0.11	0.17
$x_4$	-0.01	0.38	0.72	1.00	0.85	-0.20	0.18
$x_5$	-0.21	0.07	0.98	0.85	1.00	-0.14	0.19
$x_6$	0.17	-0.20	-0.11	-0.19	-0.14	1.00	0.47
$x_7$	0.21	-0.11	0.17	0.18	0.19	0.47	1.00

Organizing (SOM)[2], an unsupervised Artificial Neural Network (ANN). The SOM consists of a lattice of nodes, in this case a two dimensional lattice, and associated with each node  $m$  is a weight vector  $w_m \in \mathbb{R}^n$  with  $n$  the dimension of the input  $\mathbf{x}$ . After training with all the data samples  $\mathbf{x}$  the  $w_m$  form a topological mapping of the input data which is a vector quantization of the support  $S$  of the input data  $\mathbf{x}$ . For any given data sample  $\mathbf{x}$  there is a node  $v$  of the SOM for which  $\|\mathbf{x} - w_v\| < \|\mathbf{x} - w_j\|, \forall j \neq v$  hence  $v$  is the best matching or winning node for  $\mathbf{x}$ . The distance between the SOM weights,  $\|w_i - w_j\|$ , of neighboring lattice nodes  $i, j$  is inversely proportional to the probability distribution  $p_{\mathbf{x}}$  of  $\mathbf{x}$  in the region of  $S$  adjacent to  $w_i, w_j$ . This distance, and hence the representation of  $p_{\mathbf{x}}$ , is formalized for all nodes in the lattice in terms of the so called U-Matrix (Unified distance matrix) [3]. Each entry  $k$  of the U-Matrix is the average of the distance  $\|w_k - w_j\|, \forall j \in N_k$ , with  $N_k$  the indices of the immediate neighboring nodes of  $k$  on the lattice. For the purposes of visualization the magnitude of each component of the U-Matrix is coded on the SOM lattice by a color which for our purposes is a gray-scale. Nodes with lower valued U-Matrix components (i.e. higher values of  $p_{\mathbf{x}}$ ) have a darker color relative to the nodes with higher valued (i.e. lower values of  $p_{\mathbf{x}}$ ) U-Matrix components. Darker colored regions represent peaks of  $p_{\mathbf{x}}$  and hence cluster centers while lighter colored regions represent valleys of  $p_{\mathbf{x}}$  and hence boundaries between clusters. A Matlab toolbox for implementing the SOM and different U-Matrix visualization functions can be found at <http://www.cis.hut.fi/projects/somtoolbox/>.

Figure 2 (a) shows the U-Matrix of an SOM trained with data from all the 12 sessions. Despite the limitations of the gray-scale coloring it is possible to visually distinguish a certain cluster structure. Figure 2 (b) shows the same U-Matrix as Fig. 2 (a) but with highlighted nodes connected by trajectories, indicating winning nodes and frequencies, for data samples taken from session 35. In the first case consecutive data samples  $\mathbf{x}$  from context 4, in the time interval 16–22 minutes (i.e. walking outside context) were each classified in the SOM to determine the winning node for each  $\mathbf{x}$ . The winning nodes are denoted by  $\bullet$  and the size of the  $\bullet$  indicates the number of samples classified to that node. The trajectories connect the winning nodes of consecutive data samples, the width of a trajectory line indicating how many times the winner node changed between the node at each end of the line. In context 4 the winning nodes are confined to clusters in the top right and top left of the lattice with one trajectory going to the lower left. In the second case consecutive samples from context 5 in the time interval 30–40 minutes (i.e. travelling in the bus context) are classified and the winning nodes denoted by  $\star$ . In this case the winning nodes are confined to the bottom right hand corner of the lattice. Finally for consecutive samples in context 7 in the time interval 80–90 minutes (i.e. at home context) the winning nodes are denoted by  $\diamond$  and they are mainly confined to the center of the lattice. Hence the data samples representing each context are classified to different nodes in the SOM and to different clusters. It is clear also that different clusters are very much confined to defining data samples from a single context with no samples from context 5 classified to the same cluster as samples from contexts 4 or 7.



**Fig. 2.** U-Matrix for a SOM trained with data from 12 sessions in the Nokia context dataset indicating the cluster structure of the data. (a) shows the U-Matrix. (b) shows the same U-Matrix along with the trajectories of the winner nodes for consecutive data samples from 3 different time segments of session 35 (i.e. ● = 16–22 mins, ★ = 30–40 mins, ◇ = 80–90 mins), illustrated in Fig. 1.

## 5 Mixture Models for Modelling Context Data

Having analyzed some of the properties of the signals from individual sensors as well as their cross correlation in different user contexts, it is now interesting to consider how it is possible to model these signals. The ability to model a signal or a set of correlated signals provides insight to the nature and characteristics of the signals and it can also prove useful for providing simulated input for testing context recognition systems in the case where real measured data is difficult to access. Even more important is to gain insight into the characteristics a clustering algorithm applied to the context recognition algorithm should have in order to be useful.

In the previous sections where we looked at individual sensors and sets of sensors, we mainly looked at the statistical properties of the signals the mean, variance, cross correlations and cluster structure as determined from the vector quantization abilities of the SOM. Our approach to modelling the signals is then mostly based on modelling the probability distribution  $p_{\mathbf{x}}$  of the signals. We use a very basic and general probability model which is based on mixtures of probability distributions, referred to as mixture modelling [13]. A mixture model is of the form,

$$p_{\mathbf{x}}(\mathbf{x}) = \sum_{j=1}^K p(\mathbf{x}|\theta_j)\pi_j \quad (1)$$

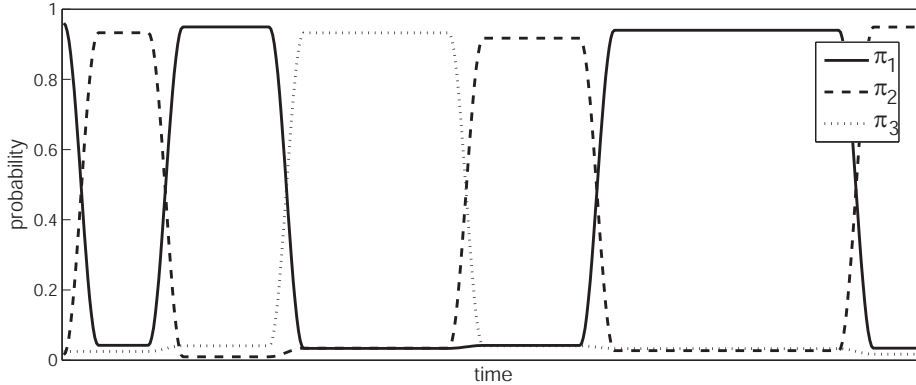
Where  $\pi_j$  are the mixture probabilities and the  $p(\mathbf{x}|\theta_j)$  are the uni-modal component distributions with parameters  $\theta_j$ . The distribution  $p_{\mathbf{x}}(\mathbf{x})$  is assumed to be a K-modal probability distribution and describes a cluster structure with  $K$  clusters and  $\pi_j$  are such that  $\sum_{j=1}^K \pi_j = 1$ .

The assumption we use is that each cluster corresponds to some user context. When working with context data it is assumed that at any given time a user is in a context/cluster and remains there for a finite period of time. However the formulation of the mixture model in (1) does not reflect this assumption, as the value of each  $\pi_j$  is the probability that at a given time a data sample  $\mathbf{x}$  is sampled from cluster/context  $j$ . In order to adapt (1), which we call a static model, to our understanding of context we allow the  $\pi_j$  to vary with time. Figure 3 illustrates how this variation might occur for  $K = 3$ , where for example when the user is in context 1 then  $\pi_1 \approx 1$  and  $\pi_2, \pi_3 \approx 0$ . This situation lasts for a finite time period and then the values of the  $\pi_j$  change again. The static mixture model of (1) becomes dynamic and expressed as,

$$p_{\mathbf{x}}(\mathbf{x}) = \sum_{j=1}^K p(\mathbf{x}|\theta_j)\pi_j(t), \quad (2)$$

where now the  $\pi_j$  vary over time and we interpret the variations in the  $\pi_j$  as variations in context.

The dynamic nature of the mixture model has an important implication when applying classic clustering algorithms to the context recognition problem. For



**Fig. 3.** Variation of mixture probabilities  $\pi_i$  over time for  $K = 3$ .

example if the SOM algorithm is applied to clustering context data in a serial on-line manner the choice of learning parameters becomes very difficult when the probability distribution of the input data is described by the model in (2), which in effect means that data samples are not independent and identically distributed. This is related to the fact that the robust convergence and self-organizing properties of the SOM, in serial mode operation, are dependent on the independent and identically distributed nature of the input data samples.

## 6 Conclusion

In order to understand context and develop applications that can be used for context recognition the first most important step is to define context. Dey has formulated an intuitive understanding of context in terms of information. This intuitive definition does not indicate how to determine which information is relevant or how to extract his information in the case of multiple sources of information. We make the proposal that the high level context information can be extracted from multiple information sources by clustering data from the information sources. The clusters in the data are contexts. First by analyzing a set of measured data, different correlations between the information sources are found in different contexts. These correlations indicate the existence of clusters in the data which are visualized using the SOM and a U-Matrix representation. Furthermore it is shown that sequences of data sampled from the same user context are classified to the same cluster, with sequences of data sampled from another user context classified to a different cluster. This analysis supports our proposal to define high level contexts as clusters in data generated from the fusion of multiple information sources.

Based on this analysis the obvious conclusion is that clustering algorithms could be used as context recognition algorithms. The probability distribution of a set of data with a cluster structure can be modelled using a mixture of

probability distributions. In the case of context data this model needs to be dynamic in nature. Classic clustering algorithms such as the SOM, used in serial mode, are limited in the case of data which is modelled by a dynamic probability mixture model. The implication is that context recognition based on clustering requires some adaptation of classic clustering algorithms.

## Acknowledgements

This work has been performed in the framework of the IST project IST-2004-511607 MobiLife, which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues, although the views expressed are those of the authors and do not necessarily represent the project.

## References

1. Dey, A.: Understanding and using context. *Personal and Ubiquitous Computing* **5** (2001) 4–7
2. Kohonen, T.: *Self-Organizing Maps*. 3rd edn. Springer, Berlin (2001)
3. Ultsch, A., Siemon, H.: Kohonen's self organizing feature maps for exploratory data analysis. In: *Int. Neural Network Conf.(INNC'90)*, Dordrecht, Netherlands, Kluwer (1990) 305–308
4. Jain, A., Dubes, R.: *Algorithms for Clustering Data*. Prentice Hall, Englewood Cliffs, NJ (1988)
5. Henriksen, K., J.Indulska: Modelling and using imperfect context information. In: *Pervasive Computing and Communications Workshops(PERCOMW)*. (2004) 33–37
6. Dey, A., Mankoff, J., Abowd, G., Carter, S.: Distributed mediation of ambiguous context in aware environments. In: *UIST '02: Proceedings of the 15th annual ACM symposium on User interface software and technology*, ACM Press (2002) 121–130
7. Staab, S.: *Handbook on Ontologies*. Springer (2004)
8. Korpipää, P., Mäntyjärvi, J.: An ontology for a mobile device sensor-based context awareness. In: *Proc. Context03*. LNAI no. 2680, Springer-Verlag (2003) 451–459
9. Chen, H., Finin, T., Joshi, A.: An ontology for context-aware pervasive computing environments. *Knowledge Engineering Review* **18** (2004) 197–207 Special Issue on Ontologies for Distributed Systems.
10. Schmidt, A., Aido, K.A., Takaluoma, A., Tuomela, U., van Laerhoven, K., van de Velde, W.: Advanced interaction in context. In: *Proc. Intl Symp. Handheld and Ubiquitous Computing*, LNCS 1707, Springer-Verlag (1999) 89–101
11. Mayrhofer, R.: Context database. [http://www.soft.uni-linz.ac.at/Research/-Context\\_Database/](http://www.soft.uni-linz.ac.at/Research/-Context_Database/) (2004)
12. Flanagan, J., Murphy, D., Kaasinen, J.: A Nokia context recording database with synchronized user interaction. In: *Benchmarks and a Database for Context Recognition: Workshop Proceedings Pervasive 2004*, ETHZ, Switzerland (2004)
13. Hand, D., Mannila, H., Smyth, P.: *Principles of Data Mining*. MIT Press, Cambridge, US (2001)