Human Routine Change Detection using Bayesian Modelling

Yangdi Xu
VI-Lab, University of Bristol, Bristol, UK
Email: Yangdi.Xu@bristol.ac.uk

Dima Damen
VI-Lab, University of Bristol, Bristol, UK
Email: Dima.Damen@bristol.ac.uk

***Abstract*—Automatic discovery of changes in a human's routine is one of the requirements for the future of smart home living, and its contribution to the E-health of the community. In this paper, a Bayesian modelling approach is used which models routine change discovery as a pairwise model selection problem. The method is evaluated on a collected office kitchen dataset that captures snapshots of the routine of the same person over multiple years (2014-2017). The results show that our method is able to detect not only the presence of routine changes, but also which activity patterns have been changed, fully automatically, and in a fully unsupervised manner. Moreover, changes within the same activity pattern can be discovered. Interestingly, discovered changes demonstrate subtle variations that are missed by the visual inspection of a human observer.**

I. Introduction

Human routines contain useful information for understanding people's behaviour patterns. **Routine**, as a term, is used by the general public to refer to *the common, regular or standard course of activity patterns*. It is tightly linked to temporal regularity. For example, routine *morning coffee* and *evening coffee* both contain the same act of making coffee. However, since they regularly happen at a different times of day, they make two different routines.

Generally speaking, most people tend to follow their routines unless they are influenced by other factors. It is the human nature for a person to stay in his/her comfort zone. Once a person has stepped out of the comfort zone, it indicates that something may have changed in his/her life. Primarily, if we are able to discover the routine changes between routine models of the same person, captured at independent durations (e.g. this year vs next year), we can detect health-related changes of prime importance to long-term monitoring. For example, spending lengthened periods in front of the TV and less on the dining table could be a sign of depression [1].

Routine changes can be **obvious** or **subtle**, based on their potential for visual discovery. An obvious change is when the change is visible and possible to identify by observations, and the point at which the change takes place is rather explicit. This type of routine change can either be caused by the change of external environments or the person himself is actively willing to change. On the other hand, a subtle change is significantly harder to visually identify and in most cases, people do not recognise when the change begins. Such type of change occurs as natural behaviour shifts and usually requires significantly longer-term recordings to identify. In this paper, we proposed a method to discover the obvious routine changes regardless of whether it is actively changed by will or passive change by external factors. The proposed method is built on top of our previously proposed unsupervised routine modelling approach using Dynamic Bayesian Network(DBN) [2].

II. Related Work

Pattern modelling allows us to understand the laws of the world surrounding is. Discovering changes in learnt patterns always remains an interesting research topic because we are naturally attracted by difference, and change indicates extra information that has not been captured before. The methods used to discover change in visual patterns vary significantly based on the different tasks people want to perform, e.g. environmental changes in remotely-sensed data [3], image change detection algorithms [4]. In these works, the aim is to separate 'unimportant' changes from those that require detection, primarily using some threshold over changes. More advance methods include probabilistic mixture models [5], minimum description length [6], predictive models (spatial and temporal) [7], [8], shading models [9] as well as background modelling [10]. All the change detection methods above focus on detecting pixel-level changes, but incapable of detecting changes in complex activities and routine patterns.

In action and activity recognition problems, change detection is usually an extension of the proposed modelling method. In supervised learning, the training dataset represents the 'normal' or standard action/activity we are interested in. The features are extracted for the best description of the activity properties. Therefore, if part of the feature is abnormal compared to the rest, the corresponding activities represented by these features should be considered as changed or abnormal [11]. [12] proposes a method to detect abnormal events in group activities. An energy potential is calculated for detecting abnormal events based on the group's position and velocity relative to its neighbours. As people move closer, the energy increases slowly whereas the magnitude and direction of the velocity remain unchanged. A bag of words representation and SVM are used to detect abnormal activities. Similarly, [13] introduces a Social Force model to detect abnormal behaviours in crowd videos. A grid of particles is placed on top of the image with the space-time average of optical flow. If the particles are treated as individuals, the interaction force is estimated and mapped onto image plane to calculate force flow per pixel per frame. The abnormalities are localised in the frame by locating the regions of high force flow.

The abnormal events and change detection methods described above cannot be applied directly to discovering change in routine patterns because the definition of change in routine and the change in activities/events are fundamentally different. In routine patterns, a change would not be considered unless the change itself has become a new routine pattern. In other words, rather than looking for abnormalities within a single routine model, routine change discovery should be achieved by identifying the difference between two or more routine models. In the next section, we propose a method based on Bayesian statistics to perform pairwise model comparisons.

III. Method

Before we propose a method for routine change discovery, we must understand why and how routine changes. As the person's routine itself is a long-term activity pattern, the changes in routine can take relatively long time to build up. For example, if *morning coffee* is a routine of a person for months, not drinking coffee one morning does not indicate the routine has changed. That is simply a one-off out-of-routine activity. It is considered as a routine change only if he continues to refrain from drinking coffee in the morning for a period of time, making 'not drinking coffee' a routine in itself, or replacing it with a different beverage. Thus a **routine change is a replacement of one activity pattern in a routine by another, which becomes in its own right a routine activity pattern**. Replacement here should be understood in its general term as the activity might cease to exist, or might be newly introduced in a routine. The precondition of routine change detection is that the detected change is an eligible routine pattern. In other words, it has to be part of the routine itself.

We build on our previous work of unsupervised routine modelling [2], where a Dynamic Bayesian Network (DBN) is able to model daily human routines using spatial, pose and time-of-day information as sources of input. Our previous work assessed variations of the independence assumptions within the DBN model. Compared to other approaches where a DBN is trained in an unsupervised way, we automatically select the number of hidden states for fully unsupervised discovery of a single person's indoor routine. We emphasize unsupervised learning as it is practically unrealistic to obtain ground-truth labels for long term behaviours. Thus in the next section, we assume that a Bayesian routine model can be built using the method in [2] and focus on detecting changes in human routine. Our routine detection method would be applicable to other types of Bayesian routine models.

A. Bayesian Model Selection

In Bayesian model selection, an efficient approach is to compute the posterior over the models.

$$p(\mathcal{M}|D) = \frac{Pr(D|\mathcal{M})Pr(\mathcal{M})}{\sum\limits_{j=1..K} Pr(\mathcal{M}_j, D)} \tag{1}$$

where $\mathcal{M}$ is an abstract expression of the model, D is the set of observations from the data, and K is the number of models. From this, we can compute the MAP model $\hat{\mathcal{M}} = argmax_{\mathcal{M}} p(\mathcal{M}|D)$. If we use uniform prior over models where $p(\mathcal{M}) \propto 1$. Thus we just need to maximise $p(D|\mathcal{M}) = \int p(D|\theta)p(\theta|\mathcal{M})d\theta$ which is called the marginal likelihood or the evidence for model $\mathcal{M}$. One might think that using $p(D|\mathcal{M})$ to select models will always favour the one with more parameters, eventually over-fitting. However, it is not the case if we use marginal likelihood. This is called the Bayesian Occam's razor effect [14] which states that one should pick the simplest model that adequately explains the data. Another way to explain the Occam's razor effect is that the probabilities must sum to 1. Although complex models are able to fit the data better, they need to spread their probability mass thinly, hence will not obtain as large probability as a simpler model, based on the conservation of probability mass principle.

In the case of routine model detection, assume two pre-trained models $\mathcal{M}_1, \mathcal{M}_2$, and two observed data sets or recordings D_1, D_2. If both sets of recordings have been used in training $\mathcal{M}_1$, it is expected that the likelihood of the data is higher when tested using that model, namely $Pr(D_1|\mathcal{M}_1) \geq Pr(D_1|\mathcal{M}_2)$ and $Pr(D_2|\mathcal{M}_1) \geq Pr(D_2|\mathcal{M}_2)$. While this is generally true for any case, we now consider the case where D_1 represents an activity pattern that persists throughout both recordings while D_2 represents an activity pattern that is only frequent in the first recording. One would expect that the drop in likelihood will be noticeably less than that for the activity pattern represented by the data D_2.

$$\frac{Pr(D_1|\mathcal{M}_2)}{Pr(D_1|\mathcal{M}_1)} \gg \frac{Pr(D_2|\mathcal{M}_2)}{Pr(D_2|\mathcal{M}_1)} \tag{2}$$

The notion of comparing ratios of likelihoods is referred to as the Bayes factor (BF). If we assume we have uniform priors on all the models, then the model selection is equivalent to selecting the model with the highest marginal likelihood. Let's suppose we are considering only two models the **null hypothesis** ($\mathcal{M}_1$), and the **alternative hypothesis** ($\mathcal{M}_2$). The Bayes factor is defined as the ratio of marginal likelihoods,

$$BF_{1,0} \equiv \frac{p(D|\mathcal{M}_2)}{p(D|\mathcal{M}_1)} \frac{p(\mathcal{M}_1)}{p(\mathcal{M}_2)} \tag{3}$$

With the uniform prior, $\frac{p(\mathcal{M}_2)}{p(\mathcal{M}_1)} = 1$. If $BF_{1,0} > 1$, then $\mathcal{M}_2$ is favoured, otherwise $\mathcal{M}_1$ is preferred. However, when BF is only slightly greater than 1, we are not confident to judge that $\mathcal{M}_1$ better explains the data. [15] proposed a scale of evidence indicating which model should be favoured based on the BF value.

B. Proposed Method and Implementation

We propose to discover routine changes based on the theory of Bayesian model selection. First, we would like to find out whether we could detect any routine change between the two time periods. Second, we would like to identify the particular activity (or activities) that has undergone change.

Let us assume there are two routine models $\mathcal{M}_1$ and $\mathcal{M}_2$ which are built using different data captured at different durations (weeks/months), namely D_1 and D_2. In general, we assume that D_1 should fit $\mathcal{M}_1$ the best, as the model is trained using this data. To determine exact routine patterns that have changed, we examine the BF value (Equation 3) on a per frame basis instead of the overall trend. The log-likelihood value is on a point estimate basis so that we can understand how well each data point fits the model. In theory, if there is a drop in log-likelihood on certain frames in $\mathcal{M}_2$, it may indicate that the pattern in those frames has changed compared to $\mathcal{M}_1$. We use this point estimate to discover which activity patterns within $\mathcal{M}_1$ have changed comparing to $\mathcal{M}_2$.

If two Bayesian models $\mathcal{M}_1, \mathcal{M}_2$ share the same independence assumptions, we additionally assume that they share the same number of parameters for $\mathcal{M}_1$ and $\mathcal{M}_2$ to avoid any marginalisation or normalisation concerns. Thus, the models are built using exactly the same complexity but different parameter values resulting from different training data.

We use the Bayes factor (BF) value on the point estimates for model selection. For frames that produce significantly low likelihoods when tested on the alternative model, we believe that this is comparable to 'decisive' evidence mentioned in [15], and we consider that as evidence of routine change in those frames. However, these extremely low likelihood values, in fact, result in a low likelihood of all subsequent frames in the sequence due to the Markovian assumption.

$$Pr(D_1|\mathcal{M}_2) = Pr(d_1|\mathcal{M}_2)\prod_{t=2}^{T} Pr(d_t|d_{t-1},\mathcal{M}_2) \tag{4}$$

The first order Markov assumption states that the estimation of the immediate past captures everything we need to know about the entire history, expressed in Equation 4. This means that when a single frame's likelihood, in a sequence, approaches zero, the following data points will also report an accumulated MLE that approaches zero. This applies to frames that could, independently, result in high MLE when tested solely. Figure 1 (left) shows an example of the situation we described above. Once a single frame reports a low likelihood, below the fixed constant, the output continues to report failure cases for the remainder of the sequence.

In order to circumvent the problem, and evaluate the likelihood of fitness for longer sequences, instead of inputting the entire sequence at once, we use a sliding window approach. We choose a window length W with a jump interval L. By doing this, we limit the failure cases within the window as a new start point is introduced each time a new window is taken as input. Among the overlapping windows, we select the maximum likelihood value as the final result of that particular frame. By selecting the maximum, we are allowing each frame to fit the new model the best, using past or future frames as part of the sliding window. It is worth mentioning that there is a trade-off in using the sliding window approach. Each time a

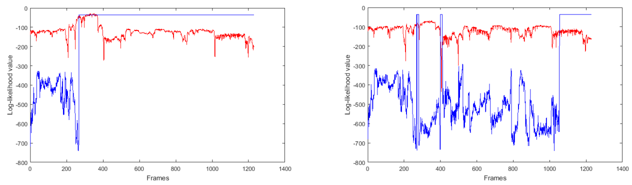

Figure 1: Left: An example of sequence from office kitchen1 directly fitting into routine model $\mathcal{M}_1$(red) and $\mathcal{M}_2$(blue). Right: An example of likelihood value of office kitchen1 fitting routine model $\mathcal{M}_1$(red) and $\mathcal{M}_2$(blue) using sliding window approach.

sequence of data is taken as input, the Bayesian model treats the first data point as the start of the sequence, and has a separate initial probability associated with it. Therefore, in theory, the likelihood estimate will be different for the first data point of each sliding window because they are not considered as an initial point when the entire sequence is taken as input directly. However, we compared estimations of the start point of sliding window using both initial and transition probabilities for $W = 300$, and found them to be only marginally different. However, it can only be true for large enough W, as for shorter sequences the initial probability will have a significant effect on the estimation results.

Using the sliding window approach, we calculate the MLE, Pr_s, of each frame:

$$Pr_s(d_t|D,\mathcal{M}_2) = \max_{i=1:\frac{W}{L}} \{Pr(d_{t-iL}|\mathcal{M}_2) \prod_{j=d_t-iL+1}^{d_t} Pr(d_j|d_{j-1},\mathcal{M}_2)\} \tag{5}$$

We can then assess the BF for each frame, and report on abnormalities resulting in significantly large BF values, namely:

$$BF_s(d_t|D,\mathcal{M}_1,\mathcal{M}_2) = \frac{Pr_s(d_t|D,\mathcal{M}_2)}{Pr_s(d_t|D,\mathcal{M}_1)} \tag{6}$$

$$C(d_t,\mathcal{M}_1,\mathcal{M}_2) = \begin{cases} 1 & BF_s(d_t|D,\mathcal{M}_1,\mathcal{M}_2) \ll 0.01 \\ 0 & otherwise \end{cases} \tag{7}$$

Thus a frame will only be considered as abnormal $C(d_t|\mathcal{M}_1,\mathcal{M}_2) = 1$ when its estimation fails for every sliding window. All other cases are regarded as the effect of the Markovian assumption. Figure 1 (right) shows an example of a log-likelihood plot for discovering routine changes using the sliding window approach. Compared to figure 1 (left), most of the abnormal estimations caused by Markov assumption are successfully removed. Red represent D_1 fitting $\mathcal{M}_1$, blue represent the same data fitting $\mathcal{M}_2$. The remaining glitches imply that, as noted in Equation 7, the data fits $\mathcal{M}_1$ but does not fit $\mathcal{M}_2$.

While this method enables us to find frames in the various sequences that represent routine change, these do not directly correspond to routine activity patterns. Thus, we need to associate theses frames to the activity pattern. In this way, we can find out which pattern has changed. To evaluate the change in routine patterns, we calculate the percentage of frames of each activity pattern with $BF_s << 0.01$, as follow:

$$R_c(A^k, \mathcal{M}_1, \mathcal{M}_2) = \frac{1}{\sum_t \mathcal{M}_1(A_t) = A^k} \sum_{t:\mathcal{M}_1(A_t)=A^k} C(d_t, \mathcal{M}_1, \mathcal{M}_2) \tag{8}$$

where A is the symbol of a routine pattern; A^k is the k^{th} routine pattern in $\mathcal{M}_i$; $\mathcal{M}_i(A_t)$ is the prediction of the activity at t using DBN model $\mathcal{M}_i$, and $=$ is the Boolean equality operator. Note that the method uses the *null hypothesis* model $\mathcal{M}_1$ for discovered patterns, as it is the model on which the data was trained. A threshold could be introduced to judge whether the discovered activity pattern has changed. A routine pattern is believe to have changed if $R_c(A^k, \mathcal{M}_1, \mathcal{M}_2) > \alpha_{R_c}$.

IV. Routine Change Dataset

To our knowledge, no datasets are publicly available for long-term routine modelling using visual input, let alone routine change discovery. In most published datasets, instructions are given to the participants so that they perform activities in some exact given orders such as [16] and [17]. These scripted activities cannot exhibit realistic patterns of a person's routine or behaviour. When people are told to do something, their mind will be more focused on following the exact order each time. In this case, the dataset will be far from natural. A couple of datasets do attempt non-scripted or long-term recording of multiple activities are publicly available [18], [19]. The TUM Kitchen Dataset [18] uses a multiple cameras system in a simulated kitchen environment. Each recording consists of a single person performing a single activity - 'prepare the dining table', multiple times but recorded over a single day. It is thus not usable for routine modelling. [19] uses a human morning routine dataset [20] for activity analysis. However, the work focuses on using motion capture data, and the visual information using RGBD cameras is not publicly available. The recently released EPIC-KITCHENS dataset [21], captured using wearable cameras records non-scripted kitchen activities for three consecutive days, however it is not long-term and thus cannot be used for routine change discovery.

The self-recorded *Bristol Routine Change Dataset - BRCD*[1] (Fig 2), captured in office kitchen, is viable for assessing changes in routine. It contains three different recordings for the same individual over several years, specifically years 2014, 2016 and 2017. The video sequences, recorded using a single RGB-D sensor (PrimeSense) captured a single individual in an office kitchen and each annual recording lasted for 6 consecutive working days. The camera is set at the entrance of the kitchen, and the recording is manually started each time the person enters the kitchen and is stopped when the subject leaves. We have no control over when the subject enters the kitchen as well as what the subject does. All activities are non-scripted in order to obtain natural behaviour patterns. Over the three years, we do not script any changes of pattern or ask the subjects to perform anything differently. In addition, We do

[1]Bristol Routine Change Dataset, publicly available from: http://people.cs.bris.ac.uk/∼damen/Routine/

Figure 2: Bristol Routine Change Dataset (BRCD)

Table I: Semantic activities for routine change dataset

Oct 2014	**Jan 2016**	**May 2017**
Prepare tea	Get hot water	Use microwave
Wash	Wash	Wash
Get hot water	Put cup	Get hot water
Get cold water	Prepare tea	Get cold water
Use fridge	Use fridge	Prepare tea
Put cup	Throw tea bag	Use fridge
Make porridge	Make coffee	
	Use microwave	

not remind the subject of her previous routines over the years, and simply capture her normal office kitchen activities over time. As the time gap between each recording is large, we believe that interesting routine changes would be observed, as the change in routine in real life is expected.

Table I shows the different ground-truth activity labels for office kitchen 2014, 2016 and 2017. The labels are listed in a descending order in terms of appearance frequency, in other words, how many times an activity took place.

V. Evaluations and Results

In this section, we demonstrate our result of routine change discovery. In terms of understanding what actually happens when a routine change occurs, semantic ground-truth is required, for which we report $R_c(A^k_{gt}, \mathcal{M}_1, \mathcal{M}_2)$. We show the results of routine change discovery among the routine models for 2014, 2016 and 2017, namely $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_3$. Based on the concept of marginal likelihood, all three models are trained using the same parameters settings.

Figures 3 plots the log-likelihood value for the first recording of the dataset. It shows that recording fits $\mathcal{M}_1$ the best as it has the highest overall likelihood value despite the glitches in $\mathcal{M}_2$ and $\mathcal{M}_3$. We can conclude that in terms of routine similarity, both the $\mathcal{M}_2$ and $\mathcal{M}_3$ show significant differences than $\mathcal{M}_1$ with lower log-likelihood estimates. However, $\mathcal{M}_3$ may be more similar to $\mathcal{M}_1$ as its MLE value is generally higher than $\mathcal{M}_2$ for the majority of the time.

The percentage of change per pattern R_c (Equation 8) is shown in table II. We set two heuristic thresholds to represent the different levels of pattern change: $\alpha_{R_1} = 25\%$ represents a moderate level of pattern change (in bold), $\alpha_{R_2} = 50\%$ represents a major pattern change (in red). In

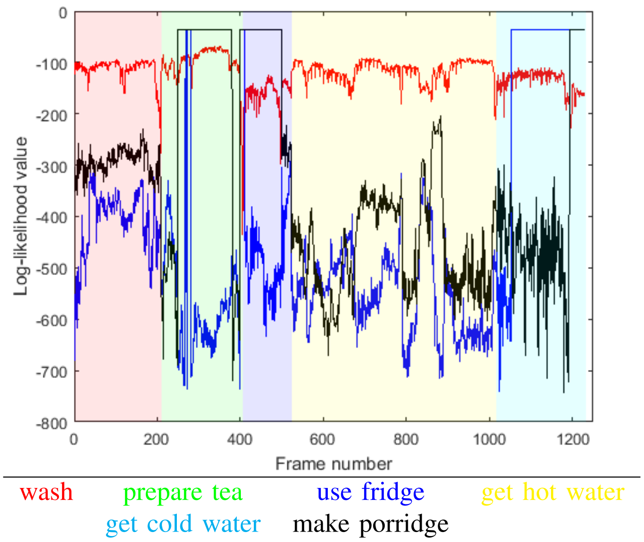

Figure 3: Log-likelihood result of sequence 1 in the 2014 recording fitting routine model $\mathcal{M}_1$(red), $\mathcal{M}_2$(blue) and $\mathcal{M}_3$(black).

Table II: The Evaluation of routine change discovery among routine model $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_3$ with semantic ground-truth.

	$\mathcal{M}_1 \rightarrow$		$\mathcal{M}_2 \rightarrow$		$\mathcal{M}_3 \rightarrow$	
	$\mathcal{M}_2$	$\mathcal{M}_3$	$\mathcal{M}_1$	$\mathcal{M}_3$	$\mathcal{M}_1$	$\mathcal{M}_2$
Prepare tea	12.0%	13.1%	**53.7%**	5.8%	**59.5%**	8.2%
Wash	7.7%	1.0%	**36.4%**	0.1%	**35.4%**	17.1%
Get hot water	10.3%	6.8%	**42.1%**	18.9%	**94.8%**	**21.5%**
Get cold water	**48.6%**	**32.0%**			**82.4%**	11.3%
Use fridge	**38.3%**	**80.5%**	**77.7%**	**47.6%**	**81.1%**	15.0%
Put cup	0%	0%	**58.6%**	**33.4%**		
Make porridge	**28.3%**	**28.3%**				
Throw tea bag			**57.8%**	**54.9%**		
Make coffee			**81.6%**	6.4%		
Use Microwave			**100%**	**100%**	**53.6%**	8.8%

$\mathcal{M}_1 \rightarrow \mathcal{M}_2, \mathcal{M}_3$, we see little pattern change for three of the most frequent patterns: 'prepare tea', 'wash' and 'get hot water'. The percentage of change is not significant as they are all below the threshold α_{R_1}. A close to major change has been identified for the activity pattern 'get cold water' in $\mathcal{M}_2$. Theoretically speaking, the change rate should be 100% because there is no such label in office kitchen 2016. The result shows that the routine model $\mathcal{M}_2$ is able to explain nearly half of the frames labelled as 'get cold water' from office kitchen 2014. By inspecting all recordings in office kitchen 2016, we conclude that some of the recordings in office kitchen 2016 may be similar to 'get cold water' due to the similarity in spatial and pose features. In $\mathcal{M}_3$ the change rate for 'get cold water' is 32.0%. More importantly, we detect one major change (in red) for the routine pattern 'use fridge' in $\mathcal{M}_1$ when tested using $\mathcal{M}_3$. When inspecting Table I, we notice that this pattern moves into the least frequent pattern in office kitchen 2017. The subject indeed stops adding milk into her tea in 2017, despite using it frequently in 2014 and 2016.

The table also shows pattern changes evaluation of office kitchen 2016($\mathcal{M}_2$) in $\mathcal{M}_1, \mathcal{M}_3$ as well as office kitchen

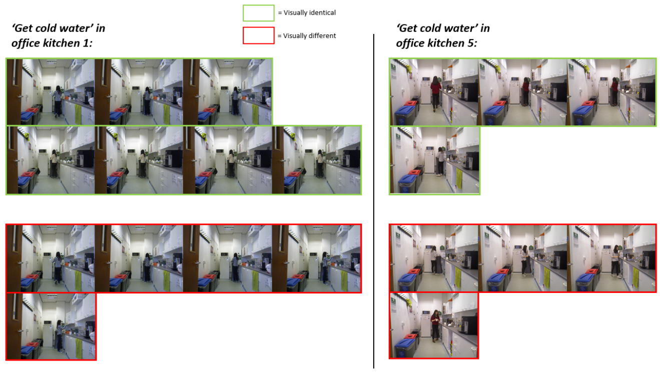

Figure 4: Visual comparison of the changes in activity pattern 'get cold water'. Left: Example recordings from office kitchen 2014 that is identified as pattern change when testing with $\mathcal{M}_3$. Right: Example recordings of pattern 'get cold water' in office kitchen 2017. Green box: Visually identical by human inspection. Red box: Visually different by human inspection.

2017($\mathcal{M}_3$) in $\mathcal{M}_1, \mathcal{M}_2$. The pattern 'throw tea bag' shows change in both models. All patterns in $\mathcal{M}_3$ are explained by $\mathcal{M}_2$ showing no pattern changes in that model. However, almost all patterns in both $\mathcal{M}_2$ and $\mathcal{M}_3$ exhibit major changes when tested using $\mathcal{M}_1$. We can conclude that $\mathcal{M}_1$ is a more specific model compared to later pattern models.

Qualitatively, we inspect the pattern 'get cold water' in $\mathcal{M}_1 \rightarrow \mathcal{M}_3$ by examining the frames of the routine change discovery (Fig 4). The example recordings from office kitchen 2014 that are identified as pattern changes are on the left. On the right are example recordings of 'get cold water' from office kitchen 2017. All images in the green box show they are visually identical, images in red are visually different. Thus, for all the recordings that are identified as a change in office kitchen 2014, only part of them can be visually identified as different. Note that we are comparing individual frame similarities. Differences could also arise from the Markovian assumption of ordering activities. This is harder to inspect.

Interestingly, the method is able to discover some minor changes in routine patterns that may go unnoticed by a human observer. A typical example we present here is related to pattern 'prepare tea', which exists in all three routine models across the years, is a frequent activity and has thus been part of the three routine models. However, a minor change takes place between recordings in the kitchen layout. In office kitchen 2014, there is a small compost bin on the worktop. The subject discards the used tea bag into the bin during her 'prepare tea' process. In the recordings of both office kitchen 2016 and 2017, the bin has been removed. The frames when the tea bag is being thrown is picked up by the method as a change in routine. Although the overall difference in prepare tea, when compared to $\mathcal{M}_2$ and $\mathcal{M}_3$ is only 12% and 13.1% respectively, such action difference within activities can still be meaningful and worth discovering. Figure 5 top show frames identified as change by our algorithm when comparing to $\mathcal{M}_2$ and $\mathcal{M}_3$. These show the person using the compost bin within

Figure 5: Top: Example frames of activity 'prepare tea' that is recognised as change when compare to $\mathcal{M}_2$ and $\mathcal{M}_3$. Bottom: Example frame of activity 'prepare tea' in office kitchen 2016 and 2017.

the activity pattern 'prepare tea'. However, as the kitchen layout changes in the latter years, the compost bin does not exist anymore during the recording of office kitchen 2016 and 2017 (area circled in red - bottom).

One of the important characteristics of our dataset is unscripted behaviours, which leads to performing the same activity in different ways. The Bayesian statistical method is able to find out the change of actions within the same activity. It is proven to be useful in discovering routine changes and somewhat powerful as it can discover the detail which a human observer may ignore.

VI. Conclusion

In this paper, we aim to discover the long-term routine changes in an unsupervised manner. Unlike changes in actions and activities, routine change is a replacement of one activity pattern in a routine by another, which becomes in its own right a routine activity pattern. We propose a routine change detection method using Bayesian statistics. We use the concept of the Bayesian factor to convert the routine change discovery problem into a model selection problem. Furthermore, we can discover which activity pattern has been changed by looking into the data fitness on a per frame basis. The method is tested on a newly introduced dataset of routine for one individual over multiple years, that exhibits natural changes in a person's routine. The results are evaluated using semantic ground-truth and qualitatively. The evaluation shows that routine changes have been successfully discovered at activity pattern level. The minor changes with the same activity pattern can also be discovered.

Acknowledgements This work was partially funded by EPSRC-IRC SPHERE, Grant EP/K031910/1. The data collected as part of the study is publicly available from the authors' webpages, or the project website:
http://people.cs.bris.ac.uk/∼damen/Routine/

References

[1] I. Pantic, A. Damjanovic, J. Todorovic, D. Topalovic, D. Bojovic-Jovic, S. Ristic, and S. Pantic, "Association between online social networking and depression in high school students: behavioral physiology viewpoint," *Psychiatria Danubina*, vol. 24, no. 1., pp. 90–93, 2012.

[2] Y. Xu, D. Bull, and D. Damen, "Unsupervised long-term routine modelling using dynamic bayesian networks," *The International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, 2017.

[3] A. Singh, "Review article digital change detection techniques using remotely-sensed data," *International journal of remote sensing*, vol. 10, no. 6, pp. 989–1003, 1989.

[4] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: a systematic survey," *IEEE transactions on image processing*, vol. 14, no. 3, pp. 294–307, 2005.

[5] M. J. Black, D. J. Fleet, and Y. Yacoob, "Robustly estimating changes in image appearance," *Computer Vision and Image Understanding*, vol. 78, no. 1, pp. 8–31, 2000.

[6] Y. G. Leclerc, Q.-T. Luong, and P. Fua, "Self-consistency and MDL: A paradigm for evaluating point-correspondence algorithms, and its application to detecting changes in surface elevation," *International Journal of Computer Vision*, vol. 51, no. 1, pp. 63–83, 2003.

[7] K. Skifstad and R. Jain, "Illumination independent change detection for real world image sequences," *Computer vision, graphics, and image processing*, vol. 46, no. 3, pp. 387–399, 1989.

[8] Z.-S. Jain and Y. A. Chau, "Optimum multisensor data fusion for image change detection," *IEEE transactions on systems, man, and cybernetics*, vol. 25, no. 9, pp. 1340–1347, 1995.

[9] E. Durucan and T. Ebrahimi, "Change detection and background extraction by linear algebra," *Proceedings of the IEEE*, vol. 89, no. 10, pp. 1368–1381, 2001.

[10] M. Piccardi, "Background subtraction techniques: a review," in *Systems, man and cybernetics, 2004 IEEE international conference on*, vol. 4. IEEE, 2004, pp. 3099–3104.

[11] O. Popoola and K. Wang, "Video-based abnormal human behavior recognitiona review," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 42, no. 6, pp. 865–878, 2012.

[12] X. Cui, Q. Liu, M. Gao, and D. Metaxas, "Abnormal detection using interaction energy potentials," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 3161–3167.

[13] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009, pp. 935–942.

[14] D. J. MacKay, "Probable networks and plausible predictionsa review of practical bayesian methods for supervised neural networks," *Network: Computation in Neural Systems*, vol. 6, no. 3, pp. 469–505, 1995.

[15] H. Jeffreys, *The theory of probability*. OUP Oxford, 1998.

[16] O. Kliper-Gross, T. Hassner, and L. Wolf, "The action similarity labeling challenge," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 3, pp. 615–621, 2012.

[17] L. Gorelick, M. Blank, E. Shechtman, M. Irani, and R. Basri, "Actions as space-time shapes," *Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 12, pp. 2247–2253, December 2007.

[18] M. Tenorth, J. Bandouch, and M. Beetz, "The TUM kitchen data set of everyday manipulation activities for motion tracking and action recognition," *International Workshop on Tracking Humans for the Evaluation of their Motion in Image Sequences (THEMIS), in conjunction with ICCV*, 2009.

[19] M. Karg and A. Kirsch, "Low cost activity recognition using depth cameras and context dependent spatial regions," in *International conference on Autonomous agents and multi-agent systems*, 2014, pp. 1359–1360.

[20] ——, "A human morning routine dataset," in *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2014, pp. 1351–1352.

[21] D. Damen, H. Doughty, G. M. Farinella, S. Fidler, A. Furnari, E. Kazakos, D. Moltisanti, J. Munro, T. Perrett, W. Price, and M. Wray, "Scaling Egocentric Vision: The EPIC-KITCHENS Dataset," *arXiv preprint arXiv:1804.02748*, 2018.