Periodicity Detection in Lifelog Data with Missing and Irregularly Sampled Data

Feiyan Hu, Alan F. Smeaton, Eamonn Newman
Insight Centre for Data Analytics
Dublin City University
Glasnevin, Dublin 9
Email: feiyan.hu4@mail.dcu.ie, alan.smeaton@dcu.ie, eamonn.newman@insight-centre.org

***Abstract*—Lifelogging is the ambient, continuous digital recording of a person's everyday activities for a variety of possible applications. Much of the work to date in lifelogging has focused on developing sensors, capturing information, processing it into events and then supporting event-based access to the lifelog for applications like memory recall, behaviour analysis or similar. With the recent arrival of aggregating platforms such as Apple's HealthKit, Microsoft's HealthVault and Google's Fit, we are now able to collect and aggregate data from lifelog sensors, to centralize the management of data and in particular to search for and detect patterns of usage for individuals and across populations. In this paper, we present a framework that detects both low-level and high-level periodicity in lifelog data, detecting hidden patterns of which users would not otherwise be aware. We detect periodicities of time series using a combination of correlograms and periodograms, using various signal processing algorithms. Periodicity detection in lifelogs is particularly challenging because the lifelog data itself is not always continuous and can have gaps as users may use their lifelog devices intermittingly. To illustrate that periodicity can be detected from such data, we apply periodicity detection on three lifelog datasets with varying levels of completeness and accuracy.**

I. Introduction

Lifelogging is a phenomenon whereby people digitally record their own daily lives in varying amounts of detail, for a variety of purposes [1]. In a sense lifelogging represents creating a "black box" of an individual human's life activities and may offer the potential to mine or infer knowledge about how we live our lives. Lifelogging can capture data from either wearable sensors such as cameras, accelerometers, GPS or iBeacon locators, or sensors built into our environment such as energy usage meters, temperature sensors or passive infra-red sensors to detect the presence of other people.

Once sensor data has been captured it is typically uploaded to a cloud-based server where it can be analysed, stored, and visualised by the user who created the data and this is what constitutes the lifelog. Various applications can then use this data and at present, most of them are based around personal healthcare or wellness. There are several cheap products on the market which log caloric energy expenditure and the types of human physical activity being performed including the FitBit One, Lark, and the Nike Fuelband. With built-in accelerometers and gyroscopes and a fairly simple algorithm these can be used to count the number of steps the wearer takes in a day. Similar products can record the duration and quality of sleep.

More advanced lifelog devices such as the SenseCam or Autographer, can record visual lifelogs where images and/or video from a first person (wearer) viewpoint can capture more detail about the wearer's daily activities. Such visual lifelogs can support more advanced applications other than health monitoring including memory recall for people with early-stage dementia [2], diet monitoring, smoking cessation, or even some job-specific lifelogs for example from medical practitioners. All of these applications of lifelogging, and most others, are based around accessing the very recent past from the lifelog and accessing discrete units of that past, corresponding to events and as such, most lifelogs are structured in a manner that reflects the way the brain is believed to store memories, i.e. based around events. For example [3], [4], [5], all refer to Cohen and Conway's model of episodic memory [6] which suggests that our memories of specific events and experiences are autobiographical and personal, and can be used to recall dates, times, places, people, emotions and other contextual facts.

Recording lifelogs and analysing them into events so we can recall specific events is a first-order use of creating lifelogs, i.e. it represents the immediate benefit. But what happens when we build up lifelogs over extended periods of time, such as years? Companies are now offering long-term storage and aggregation of lifelog data. Apple recently announced their HealthKit framework for iOS 8 at WWDC 2014 while Google announced Google Fit at around the same time. Microsoft have their HealthVault which is a similar health data aggregation and management platform. With such longitudinal and multimodal data, we can not only detect events, but we can try to identify and mine lifestyle patterns by considering time correlation within one sensor stream, and/or cross-correlation across different sensor types. To address such a challenge, we consider both low-level data and high-level data. Low-level here means the raw data stream we capture using ambient or wearable sensors. After applying machine learning techniques, low-level data can be mapped onto a higher semantic level, so accelerometer values are turned into activity levels, or categorised into activities like sitting, walking, or jogging.

Researchers in lifelogging are just now starting to realise the potential that aggregated lifelogs which bring together data from multiple sensors, all for a single individual, can offer. We know that current research into lifelogging does not fully exploit time relationships when dealing with data [1]. In [7] time series analysis methods were used to study chronologicallypresented lifelogging images. The authors concluded that DFA (Detrended Fluctuation Analysis) shows lifelogging data is not a random walk but is closer to a time series with a cyclic fluctuation. The work presented in this paper builds upon this finding. Detecting patterns of periodicity would give huge insights and reveal aspects of a persons lifestyle. However, periodicity detection usually relies on data which is both complete and has no missing values, and is accurate with no probabilities associated with the data. With lifelogging, this isn't always the case as people can simply decide not to switch on their logging devices or there can be calibration errors with the lifelog sensors. In this paper we address how to detect repeating patterns of lifestyle from lifelogs when the underlying data has missing or incomplete data, or even data which is erroneous. Once such patterns and periodicities have been detected it is beyond the scope of this paper as to how to use them or present them back to users. To illustrate our work on detecting from such noisy data we work with real lifelog datasets[1] which have in-built gaps and noise. Our work demonstrates that even with very noisy data which is also far from being continuous, we can detect repeating patterns and periodicities.

In the next section we examine how lifelogs are usually analysed and structured into events and following that we present an overview of the mathematical tools we use to detect repeating patterns in lifelogs.

II. Detecting Events in Lifelogs

Most applications of lifelogs benefit from automatically structuring the lifelog into discrete events. The challenges of effective structuring, searching and browsing of a lifelog in order to locate important or significant information has been addressed as a media process which is based on 1) capture and upload of sensor data, images or video 2) post processing of uploaded data and 3) access to processed data. This has been described in detail in [8] which presents the lifelog as a repository from which information – events of importance – can be retrieved, and this has been the access paradigm for the lifelog.

In [9], a method that can automatically segment a collection of lifelog images captured from a wearable camera is described. The features used to compare the similarity between images were MPEG-7 descriptors namely colour layout, colour structure, scalable colour and edge histogram; similarity scores across adjacent images were calculated using those features. The authors used a technique called peak scoring to enlarge the dissimilarity and some automatic thresholding methods were applied to determine the boundaries between discrete events. In the final step of this process, event boundaries that are too close to each other are merged. Following this approach, other researchers would apply machine learning techniques such as support vector machines (SVM) to train a classifier which would be used to identify the boundaries between events in a sequence of lifelog images. External data from other sensor sources such as accelerometers, GPS co-ordinates or metadata, could also be used in the segmentation process.

[1]As scientists our philosophy is always to make our research data openly available to others in the interests of transparently and reproducibility but because this is *personal* data from a personal lifelog we cannot publish this easily.

Once images have been segmented into events, a single image is selected to represent the whole event in order to facilitate event queries from users. Several selection methods have been investigated including selecting the middle image, selecting the image that is most representative, and selecting the image that is most representative but also most different to other events. Image quality was also considered as an important criterion in selecting key frame images and different image quality measures have been evaluated.

When a lifelog is segmented into events for event-based access, by default we get date and time, and perhaps location, as keys by which we can access those events but we also need to analyse lifelog content because of the huge benefits that content-based access can bring. A standard approach to multimedia access is to build a set of classifiers for a set of pre-defined semantic concepts and to train each classifier so that it assigns images from the lifelog, a score as to the confidence of that semantic concept being present in the image. In [10], thresholds were applied to determine whether a lifelog image belongs to a concept or not. One of the most important statistics for concept detection is the author-calculated average number of concepts detected for each event and compared among users.

While indexing lifelog events by the presence or absence of a set of concepts is useful, [11] described a way that a user can retrieve events by using queries which have far more semantics and which can encapsulate different aspects of an information need, specifically the *when, where, who, what* aspects. This also allows for similar events to be retrieved by computing and ranking similarity between events. Other lifelogging research [12] has shown an interest in building an ontology of semantic concepts that occur in everyday activities and which can be detected in lifelogging image collections. Wang [13] used Markov chains to model the probability distribution of objects and of semantic concepts detected in lifelog image events.

Despite all the research carried out into applications of lifelogging and into post-processing of lifelog data, especially visual lifelogs consisting of images from wearable cameras, research concentrating on analysis of lifelogs which investigates longitudinal aspects and the causality and impact of patterns detected from longitudinal analysis on lifestyle, is not apparent. This is our particular interest and is what we focus on in this paper.

III. Background Methodology

Our aim is to detect and report longitudinal patterns in lifelogs which we can regard as a form of time series, and these patterns can be referred to as periodicities. Signal processing theory tells us that in order to detect low-level periodicities in any time-series, we calculate its power spectral density (PSD or power spectrum) [14]. The PSD essentially tells us how strong is the expected signal power at each possible frequency of the signal. Because frequency is the inverse of period, we wish to identify frequencies that carry most of the energy and then from that to detect the most dominant periods. Two estimators of the PSD could be used to detect and present periodicities; the *periodogram* and the *circular autocorrelation* or *full cross correlation*. The power spectral density can be computed using the DFT (Discrete Fourier Transform) or FFT

(Fast Fourier Transform). PSD is also called periodogram and we can detect and visualise periodicity using a periodogram. The periodogram was first proposed in 1898 (Schuster, A., ”On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena,” Terrestrial Magnetism, 3, 13-41, 1898.) and is visualised as a 2D plot with spectral frequencies on the x-axis and the strength of the pattern at each frequency measured on the y-axis.

In terms of lifelogging, the periodogram can be used to detect the natural cycles that occur in lifestyle, behaviour, and activities. Periodicity can be observed in many natural phenomena, such as circadian rhythms associated with our sleep, annual seasons and so on. Intuitively, we think of our routine daily lives as composed of various forms of recurring events with obvious periodicities around daily, weekly, monthly, seasonal and annual cycles. In any kind of spectral analysis of a lifelog we expect to see periodicity around these frequencies. However, without the help of lifelogging devices and the resulting lifelog, analyzing the periodicity of human life is not a practical proposition.

We now define the tools we use to detect periodicity in lifelogs.

A. Autocorrelation

In statistics, correlation is basically measuring how similar two sequences are. This quantitative measurement of similarity of signal 1 and signal 2 can be defined as:

$$r_{12} = \frac{1}{N} \sum_{n=1}^{N-1} x_1[n] x_2[n]$$

Cross correlation between time shifted sequences, can be defined as:

$$r_{12}(k) = \frac{1}{N} \sum_{n=1}^{N-1} x_1[n] x_2[n+k]$$

All possible k-shifted time series could generate another sequence of numbers only changing with k, which is called full cross-correlation. The correlation between a signal and time shifted version of itself is called an auto-correlation. A lag operator is used to generate the time shifted signal and ‘0 lag’ equals to mean-square signal power. Auto-correlation can be defined as:

$$r_{11}(k) = \frac{1}{N} \sum_{n=1}^{N-1} x_1[n] x_1[n+k]$$

B. Periodogram

The normalized Discrete Fourier Transform (DFT) of a sequence $x(n)$, $n = 0, 1, \ldots, N-1$ is a sequence of complex numbers $X(f)$:

$$X(f_{k/N}) = \frac{1}{\sqrt{n}} \sum_{n=0}^{N-1} x(n) e^{-\frac{j2\pi kn}{N}}$$

where the subscript k/N denotes the frequency that each coefficient captures. Suppose that X is the DFT of a sequence $x(n)$. The periodogram P is provided by the squared length of each Fourier coefficient:

$$P(f_{k/N}) = \|X(f_{k/N})\|^2 \quad k = 0, 1, \ldots, \lceil \frac{N-1}{2} \rceil$$

Notice here that k ranges from 0 to $\frac{N-1}{2}$. In order to find the k dominant periods, we need to pick the k largest values of the periodogram. This works well for short to medium length periods but for long periods or low frequencies, performance is worse because each value in the periodogram indicates the power at frequency interval $[\frac{N}{k}, \frac{N}{k-1}]$ which is too wide to capture large periodicity. Thus the accuracy of periodicity detection at low frequency will be lower than at higher frequency. For lifelogging, this means there is difficulty in detecting patterns measured in years. Another difficulty when using periodograms is *spectrum leakage*, which causes frequencies that are not integer multiples of the DFT bin width to disperse over the entire spectrum which could result in false alarms being detected in the periodogram. However, the periodogram is still a good way to guarantee the accuracy of detected periods with short to medium frequency.

In the context of our work on periodicity detection from lifelogs, one of the challenges we are faced with is missing or erroneous data from the lifelog. For such a scenario, the Lomb-Scargle periodogram [15] can be used to detect periodicity in signals with missing, unevenly or unequally spaced data. This is defined formally as

$$P_X(\omega) = \frac{1}{2} \left\{ \frac{[\sum_{n=1}^{N} y(t_n) \cos(\omega(t_n - \tau))]^2}{\sum_{n=1}^{N} \cos^2(\omega(t_n - \tau))} + \frac{[\sum_{n=1}^{N} y(t_n) \sin(\omega(t_n - \tau))]^2}{\sum_{n=1}^{N} \sin^2(\omega(t_n - \tau))} \right\}$$

where τ is defined as:

$$\tan(2\omega\tau) = \frac{\sum_{n=1}^{N} \sin(2\omega t_n)}{\sum_{n=1}^{N} \cos(2\omega t_n)}$$

IV. Datasets

The purpose of this work is to determine how well periodicity can be detected in lifelog data, focussing specifically on how the tools perform in the scenario of missing data and gaps in the lifelog. In this section we describe the datasets which we have used.

A. Sleep Dataset

The first dataset represents 2.5 years of continuous nightly sleep monitoring for an individual with a +80% capture rate. Data was collected using the wrist-worn Lark sleep sensor[2] and contains the following information:

1) Time to sleep – represents the time taken between going to bed and falling asleep;
2) Time to rise – represents the time taken between waking and getting out of bed;

[2] http://www.lark.com

3) Time asleep – represents the duration of sleep;
4) Quality – a numeric indicator of sleep quality computed as a function of how well the night's sleep mapped to the circadian sleep (90-minute) rhythm and how many cycles of that rhythm were completed;
5) Times woken up – represents the number of instances of a wake-up during sleep, where "wake up" represents even a turning over in the bed;

The distribution of some of these parameters (3 and 4) is shown in Figure 1 and the frequency of data capture is shown in Figure 2 where a black line represents an instance of captured data. An obvious periodicity we would hope to detect is based on the weekly cycle where the subject tends to sleep longer at weekends than during workdays because he has a regular work schedule of Monday to Friday.

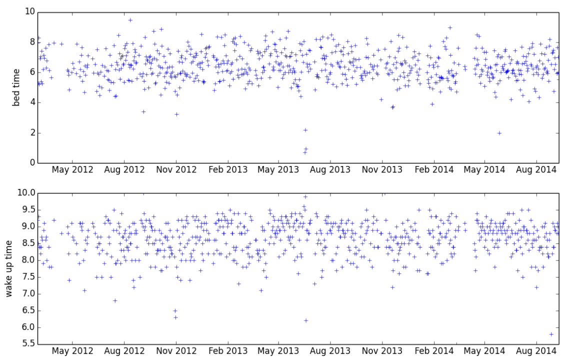


Fig. 1. Some of the raw sleep data

Fig. 2. Frequency of capture of sleep data

B. Sports Dataset

The second dataset represents a 10-year log of physical exercise activities including running, cycling and swimming, from an international tri-athlete (now retired from competition). The log contains a daily entry for distance covered for 1 or more of the sports as well as daily text comments which can indicate mood, training effort, relative performance, weather, etc. and these can be analysed for sentiment. sports dataset capture 100% of activity log in 10 years. Obvious periodicities to be detected from this data include seasons, performance at targeted sports events, perturbations caused by occasional injury and overall decline over the decade from ageing.

In Figure 3, the raw distances for running, cycling, swimming and for aggregated activity effort is shown. The latter of these plots accounts for days where the athlete would exercise or compete in more than one discipline and aggregated activity is computed according to the metabolic equivalent (MET) where the unit of MET is 1 kcal/kg*h. To calculate this the average speed for each of the three sports activities of the athlete is used. In [16], the MET for each sport activity at the average speeds indicated by the athlete are shown in Table I.

TABLE I. MET TABLE

Activity	Speed (kph)	MET
Running	13	12.9
Cycling	25	8.4
Swimming	3	8.9

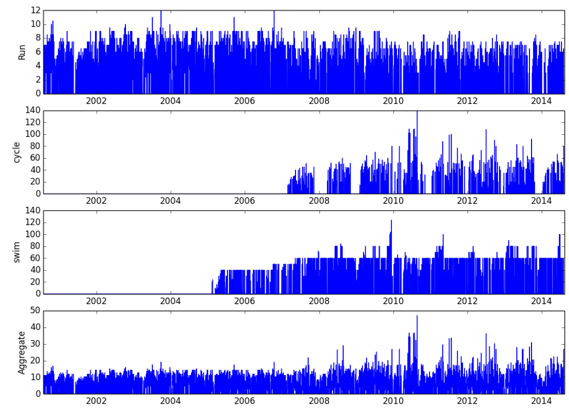
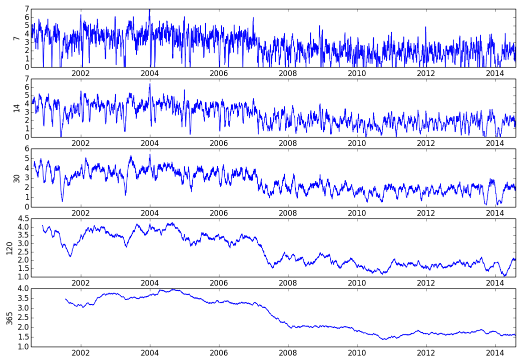


Fig. 3. Visualization of raw data in the sports activity dataset

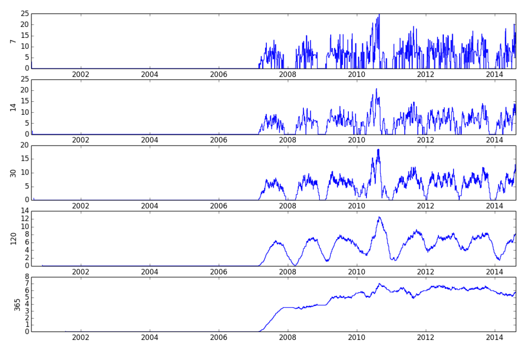
In the running, cycling, swimming and aggregated data visualized in Figure 3 the X-axis represents time, while Y-axis is the distance for the corresponding activity. From the visualization, no obvious periodicity can be observed in running, swimming or aggregated data but there seems to be an annual periodicity in the cycling data.

For each sporting activity and for the aggregated data, we applied window sizes of 7, 14, 30, 120, 365 days to calculate the moving averages. Figure 4 shows the results of this. Running, cycling and swimming start from 2000, 2007 and 2005 respectively. Moving average calculates the mean value of a fixed size window and then moves the window one day forward to get the new value. Moving average works like a low-pass filter; the bigger the window size, the lower the frequency can pass. Because of this, it is easier to find long-term trends using a larger window size because short terms shocks in the data (competitions, vacation, short-term injuries) will be smoothed. From the moving average results, we can see that running distance decreased over time, while the cycling and swimming distances increased. The total amount of energy expenditure according to MET fluctuates and no obvious trends can be seen in the aggregated data. We can infer from this data that after the athlete started to train for swimming in 2005 and for cycling in 2007, he adapted himself to this by reducing the amount of training for running.

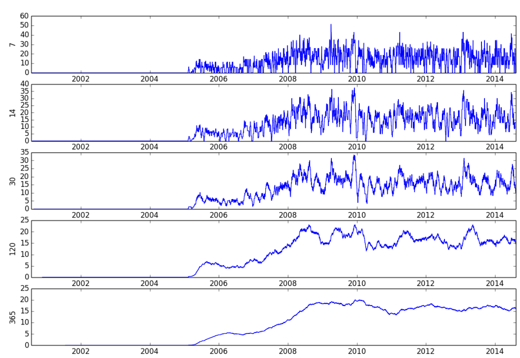
One major difference between the sleep and sports datasets is that the sports dataset has 100% capture rate of activity over 10 years, while the sleep dataset captures just over 80% of the nights in a 2.5 year period. The raw figures on sporting activities are augmented by the athlete annotating most days with text comments which summarise the day and occasionally report on performance or mood. These reports are infrequent (25–30%), and so provide sparse data which we can also examine for periodic patterns.



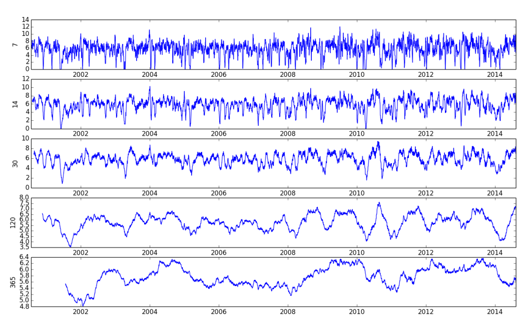
(a) Run



(b) Cycle



(c) Swim



(d) Aggregated

Fig. 4. Moving average values for sports dataset

We annotated the reports for mood and for performance to create a third dataset. Four annotators were asked to annotate the text for mood by following the following strategy: if a comment provides an indication of mood (*"feeling great"* or *"not well"*, *"ok"*), give a rating between 1 and 5, where 1 indicates the worst feelings and 5 indicates the best feelings. If there is no indication of mood in the text, give a rating of 0. For annotation of performance the four annotators were given the following instruction: when a comment provides an indication of performance (*"personal best"*, *"strong finish"*, *"stopped early"*), give a rating between 1 and 5 where 1 indicates poorest performance and 5 indicates best performance. If there is no indication of performance in the text, give a rating of 0.

Comments made by the athlete during the year 2007 were randomised and presented to 4 annotators. Because the marks for mood and performance given by annotators are highly subjective and have biases, inter-annotator agreement namely Cohen's Kappa co-efficient [17] was calculated across the annotators and is presented in Table II and III.

TABLE II. INTER ANNOTATION AGREEMENT FOR MOOD

Annotator	A	B	C	D
A	1.00	0.47	0.60	0.47
B	0.47	1.00	0.41	0.48
C	0.60	0.41	1.00	0.36
D	0.47	0.48	0.36	1.00

TABLE III. INTER ANNOTATION AGREEMENT FOR PERFORMANCE

Annotator	A	B	C	D
A	1.00	0.15	0.12	0.17
B	0.15	1.00	0.35	0.37
C	0.12	0.35	1.00	0.36
D	0.17	0.37	0.36	1.00

Cohen's Kappa coefficient ranges from 0 to 1, where a value of 1 indicates complete agreement between a pair of annotators, and 0 denotes complete disagreement. For mood, we can see that annotator A highly agrees with annotator C, while C and D are agree least with each other, though all values are greater than 0.3. For annotation of performance, it is obvious that annotator A has low agreement with all three other 3 annotators. Based on this assessment of inter-annotator agreement, we apply the following fusion strategy:

- For *Mood*, for each annotated comment discount the value which is the greatest outlier and average the remainder;
- For *Performance*, discount annotator A completely and then for the other (B,C,D) annotations on each comment, discount the one who is the greatest outlier, then average the remainder.

The fused mood and performance data from the 4 annotators are sparse and have large amounts of missing data and gaps as shown in Figure 5 where a black line represents a mood or performance value while whitespace indicates there is either no comments made by the subject for that day's activity or the mood and/or performance indicators are absent. The gap sizes for fused mood and performance vary between 1 and 19 days, while mood has a mean gap of size 4.15 days compared to 3.15 for performance. This unevenly sampled data makes it a real challenge to detect periodicity from this data and an ideal target for the Lomb-Scargle periodogram.

(a) Fused mood

(b) Fused performance

Fig. 5. Distribution of fused data from annotators

V. Experimental Results

We applied periodograms and correlations to both datasets to see if periodicities were apparent even with missing data and irregular sampling. The periodogram reveals the energy carried by each frequency across a range and is plotted as a graph where the x-axis is frequency and the y-axis is energy. If there is statistically significant energy carried by one frequency or different frequencies, this will be revealed graphically.

A. Results on Sleep Dataset

Each of the parameters from sleep logging (duration, quality, number of wakes, time in bed, etc.) has been analysed for periodicity but rather than present all of them, we limit ourselves to just two. For time asleep, a weekly periodicity is clearly detected as can be seen in Figure 6. This can be explained by the weekday/weekend cycle which is the basis for the subject's lifestyle of working during weekdays and having to get up early to commute to work and then leisure activities with later rising at the weekend. There is also a periodicity at around the 120 day frequency, about every 4 months but without going back to the subject to investigate, this remains unexplained for the moment.

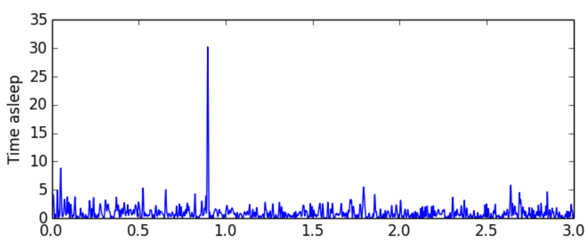


Fig. 6. Time asleep periodogram

For sleep quality as shown in Figure 7 there is no weekly periodicity which tells us that even though the subject sleeps more at weekends, he doesn't actually sleep with better quality. We also observe a periodicity around 128 days (ca. 4 months) for sleep quality but at the time of writing, without conferring with the subject, this is something we cannot explain.

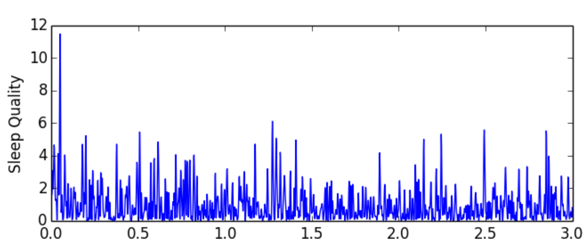


Fig. 7. Sleep quality periodogram

While the other sleep parameters such as *time spent in bed*, *time going to bed* have also yielded interesting results, the point we wanted to make is already made, namely that we can detect credible periodicities from lifelogs even though there is missing data and irregular sampling.

B. Results on Sports Dataset

Since the sampling rate of our sports activity dataset is 1 day, the minimum periodic pattern of this dataset we can detect is 2 days.In Figure 8, periodograms for the sports dataset which does not have missing data and is consistently and regularly sampled for the three sport activities and for the aggregated data MET levels shows interesting results. We can observe three significant energy levels carried by three different frequencies consistently across all 4 subplots. These three frequencies are around 0.14, 0.28, 0.43, which corresponding to periods of 7 days, 3.5 days and 2.3 days. Moreover, if we look at the plots more thoroughly, there exists a frequency at circa 0.0027 located the near the left end of the cycling and aggregated data subplots. This frequency corresponds to the annual period (ca. 365 days) that we observed in the visualization of the cycling data.

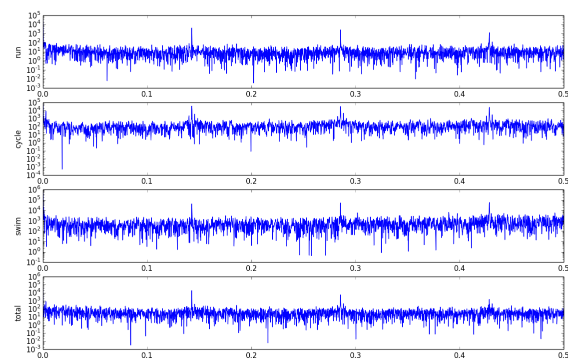


Fig. 8. Sports dataset periodograms

In order to investigate periodicity in irregularly sampled data, a second tool we use is autocorrelation. Autocorrelation of 10 years data is plotted in Figure 9.

Autocorrelation computes the correlation between the signal and a time-shifted version of the same signal. The x-axis of the autocorrelation plot is time lag and the y-axis is a measure of the correlation of the original signal and lagged signal. If the original signal is periodic then the autocorrelation of the signal should also be periodic and the periods will be located at the peaks the autocorrelation plot. From Figure 9, there are no periodicities observed in the running, swimming or MET score aggregated data, but an annual periodicity can be found in the autocorrelation of the cycling data. Curious as to where the periodicities over 7, 3.5 and 2.3 days which were found in periodograms from running, swimming and the aggregated data, we took one year of data from 2007 to see if we could detect periodicity in periodograms for just that year. An autocorrelation plot for data from the year 2007 is shown in Figure 10.

The autocorrelation plot of sports data from 2007 shows that there is a very regular weekly periodicity in running, cycling

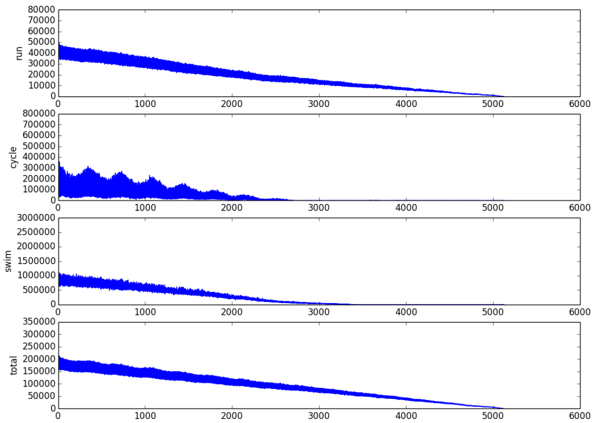


Fig. 9. Sports dataset autocorrelations

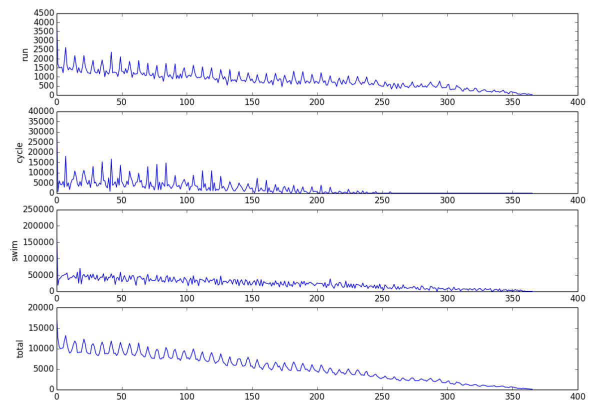
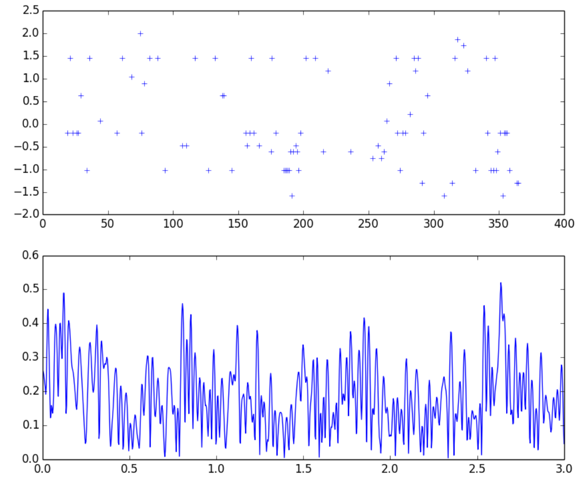


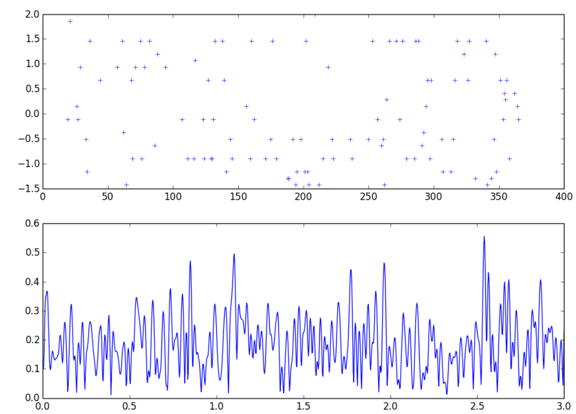
Fig. 10. Autocorrelation plot of sports data from year 2007

and in the total energy expenditure of activities, but a less regular weekly periodicity for swimming. We can also find smaller peaks between the two obviously large peaks from running and cycling data, which may correspond to the 3.5- and 2.3-day periodicities also detected in the periodogram. However there are no obvious smaller peaks found in the autocorrelation of aggregated data. A possible explanation may be that these detected periodicities indicate the lifestyle of the subject such as regular scheduled training sessions for running, cycling and swimming. Another explanation might be that there exists an inherent timetable that the subject follows in order to balance participation in the three different activities. For instance the timetable could be every 2 or 3 days run, cycle or swim once. Determining this requires going back to the subject to confirm this though this falls into the category of exploiting rather than determining periodicities which as mentioned earlier, is beyond the scope of this paper.

In order to detect periodicity in mood and performance, which are unevenly sampled and have gaps in the data, the Lomb-Scargle periodogram is applied to the mood and to the performance data. For the Lomb-Scargle periodogram, the period is $T = \frac{2\pi}{f}$. In Figure 11, we can see that there are no statistically significant energy levels carried by any of the frequencies in the LS periodogram for mood or for performance. In other words, no periodicity is detected in either mood or performance data which has been fused from the annotators. Trying to rationalise this by going back to the subject might reveal that his training schedule is oriented to have peak performance during the months of competition, typically the Summer months, so there could be an annual cycle for performance which could be tied to a mood performance cycle. There may also be peaks in performance, and in mood, around regular seasonal targets such as Winter, Spring, Summer and Autumn events. The fact that such periodicities did not appear does not mean that they do not exist, it just means that they were not detected, most probably because of the sparsity of our mood and performance data with large gaps and irregular sampling. Not even the Lomb-Scargle periodogram was able to overcome this disadvantage.



(a) Mood



(b) Performance

Fig. 11. Mood and performance data

VI. Conclusions

In the work presented in this paper, we applied periodicity detection on two longitudinal datasets, which include distances for athletic training and competition for an international triathlete, over a 10 year period, and sleep quality, duration and timing data from a subject over a 2.5 year period. The first dataset was augmented with a pool-based annotation of the triathlete's daily text commentary on his training and performance, from which we were able to get annotations for mood, and for performance. This gave us a collection of datasets which are rich in the variability of their regularity of logging, from consistent and regular daily entries to much more sporadic data with missing data and irregular sampling.

Applying moving average, we discovered that after starting cycling and swimming at a point several years ago, the subject decreased the amount of running while the distances for swimming and cycling kept increasing. The use of periodograms revealed that there are rhythms of repeating patterns at 7, 3.5 and 2.3 days for the running, cycling and swimming data, as well as for when the individual activity data is aggregated based on MET scores. An annual periodicity was also detected in the cycling data. Using an autocorrelation plot for data from year 2007, an obvious weekly periodicity was detected in running, cycling and aggregated MET data but the weekly pattern for swimming is weak suggesting less rigour and regularity associated with training in that sport. An autocorrelation plot of running and cycling shows an unexpected periodicity at a cycle of less than a week (2 or 3 days). This infra-week periodicity may be caused by training schedules for different sports in order to achieve a balanced exercise portfolio. There are no significant periodicities detected in the Lomb-Scargle periodogram for mood or for performance when fused from the annotations of a set of four annotators.

Our future work will concentrate on evaluating the detected periodicities using some form of qualitative evaluation. Evaluation is always a challenging part of this research. The relevance of any detected periodicities is quite subjective since every individual has his/her own understanding of their own periodicity. Both qualitative and quantitative evaluations will be used. In qualitative analysis for case studies, researchers usually study case-by-case independently and then draw findings separately. An interview is a common way to collect sufficient data to for qualitative analysis. After case-by-case study, a cross case analysis could be conducted to discover common phenomena. Future work will also focus on developing algorithms to increase the accuracy of periodicity detection, i.e., to more precisely compute the energy spectrum and locate periodicity. We also intend to closely investigate computational biology algorithms such as time series motif detection.

We have demonstrated in this paper that automatic detection of periodicites from lifelog data can be achieved, even when there is substantial missing data. We have shown that methods based on periodograms and autocorrelation can be used to detect periodicity on complete datasets, while Lomb-Scargle periodograms can be used to detect periodicity on datasets with missing data. Experiment conducted on three datasets with different level of sparsity shows that we are able to detect periodicity in these datasets.

Acknowledgements

The authors would like to thank Science Foundation Ireland for funding under grant number SFI/12/RC/2289, and the European Community 7th Framework Programme (FP7/2007-2013) for funding under grant agreement 288199 Dem@Care.

References

[1] C. Gurrin, A. F. Smeaton, and A. R. Doherty, "Lifelogging: Personal big data," *Foundations and Trends in Information Retrieval*, vol. 8, no. 1, pp. 1–125, 2014. [Online]. Available: http://dx.doi.org/10.1561/1500000033

[2] E. Berry, A. Hampshire, J. Rowe, S. Hodges, N. Kapur, P. Watson, G. B. G. Smyth, K. Wood, and A. M. Owen, "The neural basis of effective memory therapy in a patient with limbic encephalitis," *Neurology, Neurosurgery, and Psychiatry with Practical Neurology*, vol. 80, no. 3, pp. 582–601, 2009.

[3] A. R. Doherty, K. Pauly-Takacs, N. Caprani, C. Gurrin, C. J. A. Moulin, N. E. O'Connor, and A. F. Smeaton, "Experiences of Aiding Autobiographical Memory Using the SenseCam," *Human-Computer Interaction*, vol. 27, no. 1-2, pp. 151–174, 2012.

[4] K. Pauly-Takacs, C. J. Moulin, and E. J. Estlin, "SenseCam as a rehabilitation tool in a child with anterograde amnesia," *Memory*, vol. 19, no. 7, pp. 705–712, Oct 2011.

[5] A. R. Silva, S. Pinho, L. Macedo, and C. J. Moulin, "Does SenseCam improve general cognitive performance?" *Am J Prev Med*, vol. 44, no. 3, pp. 302–307, Mar 2013.

[6] G. Cohen and M. A. Conway, *Memory in the real world*. Routledge, 2008.

[7] N. Li, M. Crane, and H. J. Ruskin, "Automatically detecting "significant events" on sensecam," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 11, no. 06, 2013.

[8] H. Lee, A. F. Smeaton, N. E. O'Connor, G. Jones, M. Blighe, D. Byrne, A. Doherty, and C. Gurrin, "Constructing a sensecam visual diary as a media process," *Multimedia Systems*, vol. 14, no. 6, pp. 341–349, 2008.

[9] A. R. Doherty and A. F. Smeaton, "Automatically segmenting lifelog data into events," in *Image Analysis for Multimedia Interactive Services, 2008. WIAMIS'08. Ninth International Workshop on*. IEEE, 2008, pp. 20–23.

[10] D. Byrne, A. R. Doherty, C. G. Snoek, G. G. Jones, and A. F. Smeaton, "Validating the detection of everyday concepts in visual lifelogs," in *Semantic Multimedia*. Springer, 2008, pp. 15–30.

[11] A. R. Doherty, C. Ó Conaire, M. Blighe, A. F. Smeaton, and N. E. O'Connor, "Combining image descriptors to effectively retrieve events from visual lifelogs," in *Proceedings of the 1st ACM international conference on Multimedia information retrieval*. ACM, 2008, pp. 10–17.

[12] P. Wang and A. F. Smeaton, "Semantics-based selection of everyday concepts in visual lifelogging," *International Journal of Multimedia Information Retrieval*, vol. 1, no. 2, pp. 87–101, 2012.

[13] ——, "Using visual lifelogs to automatically characterize everyday activities," *Information Sciences*, vol. 230, pp. 147–161, 2013.

[14] M. Vlachos, S. Y. Philip, and V. Castelli, "On periodicity detection and structural periodic similarity." in *2005 SIAM International Conference on Data Mining*, vol. 5. SIAM, 2005, pp. 449–460.

[15] J. D. Scargle, "Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data," *Astrophysical Journal*, vol. 263, pp. 835–853, Dec. 1982.

[16] M. Jette, K. Sidney, and G. Blümchen, "Metabolic equivalents (mets) in exercise testing, exercise prescription, and evaluation of functional capacity," *Clinical cardiology*, vol. 13, no. 8, pp. 555–565, 1990.

[17] J. Cohen, "Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit." *Psychological bulletin*, vol. 70, no. 4, p. 213, 1968.