Visualizing Personal Lifelog Data for Deeper Insights at the NTCIR-13 Lifelog-2 Task

Qianli Xu
Institute for Infocomm Research
qxu@i2r.a-star.edu.sg

Vigneshwaran Subbaraju
Singapore Bioimaging Consortium
vigneshwaran_subbaraju@sbic.a-star.edu.sg

Ana Garcia del Molino
Institute for Infocomm Research
Nanyang Technological University
stugdma@i2r.a-star.edu.sg

Jie Lin
Institute for Infocomm Research
lin-j@i2r.a-star.edu.sg

Fen Fang
Institute for Infocomm Research
fang_fen@i2r.a-star.edu.sg

Joo-Hwee Lim
Institute for Infocomm Research
joohwee@i2r.a-star.edu.sg

ABSTRACT
We present a method for finding insights from personal lifelogs. First, we create minute-wise annotation of the users' activities with respect to the given topics (e.g. socialize, exercise, etc.). This is achieved by performing image retrieval using deep learning, followed by the fusion of multi-modality sensory data. Second, we generate insights of users' activities that include facts of activity occurrence, temporal and spatial patterns, associations among multiple activities, etc. Finally, we build a prototype mobile app to visualize the insights according to selected themes. We discuss challenges in the process, and opportunities for research and applications in lifelog information access.

Keywords
Lifelogging; quantified-self; visualization; insight

Team Name
VCI2R

Subtask
LIT - Lifelog Insight (sub) Task

1. INTRODUCTION
This paper presents a method to generate and visualize lifelog insights under the NTCIR-13 Lifelog-2 Task, Lifelog Insight (sub) Task (LIT). The aim is to gain insights into the lifelogger's life, following the Quantified Self (QS) movement. QS focuses on the visualization of knowledge generated from self-tracking data to provide knowledge of the self in numbers [1]. Five topics are included, namely *diet*, *exercise*, *social*, *where*, and *compare* [2]. In this paper, we show the procedure to generate insights related to the topics, and a suite of tools to visualize the result. It is to our belief that presenting visual insights will help build stronger value profile of lifelogging for better user acceptance [3, 4].

2. METHOD
The procedure of insight visualization is shown in Figure 1. First, we create minute-wise annotation of the users' activities with respect to given topics (e.g. eating, exercising, etc.). This is achieved by performing image understanding using deep learning, followed by fusion of multi-modality sensory data using a retrieval model trained on ground truth data. The outcome is annotations that describe if a user is engaged in an activity at any given time. Second, we generate insights of users' activities that include facts of activity occurrence, temporal and spatial patterns, and associations among multiple activities. Finally, we visualize the result in a themed diary in the form of a prototype mobile app that allows users to browse and search information, as well as being advised on health lifestyle.

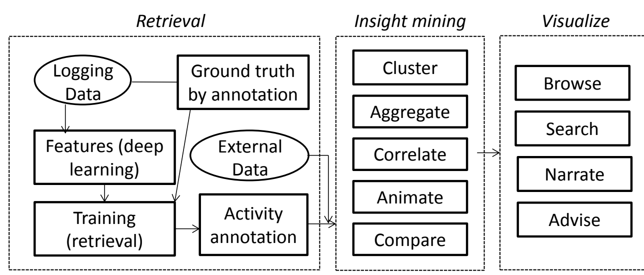


Figure 1. Framework of insight generation and visualization

2.1 Retrieval: annotating user activities
Activity is the basic unit for describing user behaviors. For any given topic, we define one or a few activities that can be inferred from the lifelog data, e.g. eating, driving a car, working on a computer. In this paper, we adopt a process in [5] to annotate user activities with respect to the given topics. Table 1 shows the interpretation of topics and the corresponding activities.

For topics not directly related to activities, one can use metadata (e.g. food log, health log, and GPS locations) and external resources (e.g. food nutrition database). In this study, dietary information is extracted directly from the textual food log of u1. It is observed that the food log is not strictly structured. Multiple food items may appear in a single log item (e.g. "*homemade bolognaise with wholewheat pasta and red wine cashew*"); and different names may refer to the same type of food ("*cerial*", "*all-bran cerial*", "*all-bran*"). To allow for automated analysis, each item in the food log is broken down into individual foods. For example, four individual foods are extracted from above, namely *homemade bolognaise*, *whole-wheat pasta*, *red wine*, *cashew*. Next, similar foods are clustered into groups, where each group refers to a unique food. From this, the statistics related to food consumption can be computed with respect to the unique foods, e.g. number of times the user ate cereal. Finally, by searching in an open dataset [6], the nutritional information (glycemic load - GL and glycemic index - GI) is computed as an aggregation of multiple food items consumed in a day.

Table 1. Decoding topics into activities

Topics	Activities
T1: Diet and blood sugar level	Eating: user is eating food
T2: Exercise and physical activity	Walk, Run, Hiking*, Gym/Yoga.
T3: Social	User is facing one or more persons in a conversation
T4: Transportation	Driving a car or taking a taxi, taking a bus, taking a train, taking a plane

* For u1, hiking refers to moments when he is standing or walking in the mountain area.

2.2 Theme-finding: templates for insight visualization

A few common strategies are applied for finding insights, which is further visualized in a few templates.

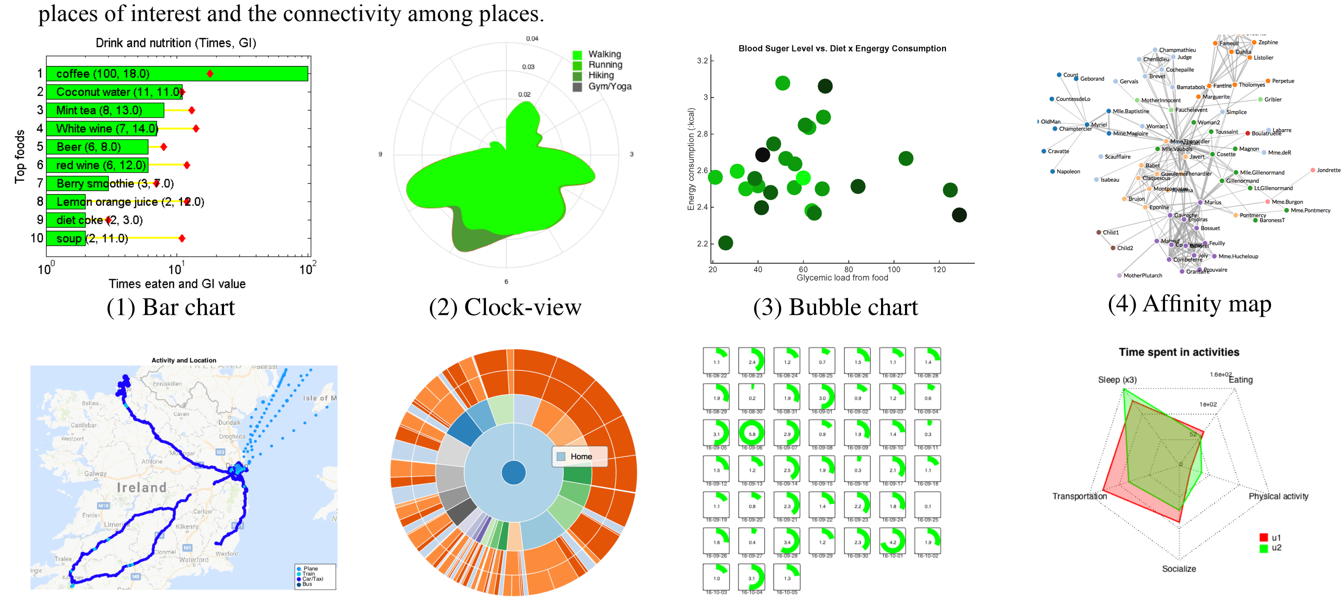
2.2.1 *Theme-finding*

(1) Aggregate: This is a process whereby one selects data points according to certain criteria and derives the statistics (e.g. mean, variance, frequency, etc.), e.g. the number of times that a user socializes, the total time spent in cycling.
(2) Cluster: Clustering can be used to group data points into categories based on specific distance measures. It reduces the dimensionality of the data, allowing users to review and identify the associations among data. In the LIT task, we apply clustering to the location information, so as to find places of interest and the connectivity among places.
(3) Associate: This is intended to discover correlation and/or causal relationships between different measurements. It answers important questions such as how diet affects blood sugar level, what factors influence a user's mood. We use scatter plots and bubble charts to demonstrate the correlations.
(4) Animate: Animation is good for demonstrating the dynamics of data. That is, how things change over time. It can be combined with other information such as location and multimedia to provide richer viewing experience.
(5) Compare: By juxtaposing the outcome of the above process, one can examine how people differ along the specific dimensions.

2.2.2 *Templates for insight visualization*

A suite of tools are developed to implement the aforementioned theme-finding strategies, and importantly to visualize the outcome for sense making. Some of the visual illustrations are shown in Figure 2. The actual usage of these templates for insight finding is presented in Section 3.

(1) Bar chart shows rank-ordered frequency/occurrence.
(2) Clock-view of activity intensity shows the temporal information of activities.
(3) Bubble chart illustrates association/correlation. If there are two variables, the x-axis is the explanatory variable and y-axis is the response variable. If there are three variables, both x- and y- axes capture the explanatory variables, the response variable is shown as the size and/or color of the bubbles.
(4) Affinity map illustrates the clustering and relationship of items according to certain distance metrics.
(5) Map is user to show geographical information. It can be combined with time and multimedia for a more comprehensive view.
(6) Sunburst view of integrated activity profile uses a series of concentric rings to illustrate activities according to time and location information. It also allows a user to view the lifelog information at varying levels of granularity.
(7) Calendar view shows activity intensity over time.
(8) Radar chart compares two users along selected dimensions.



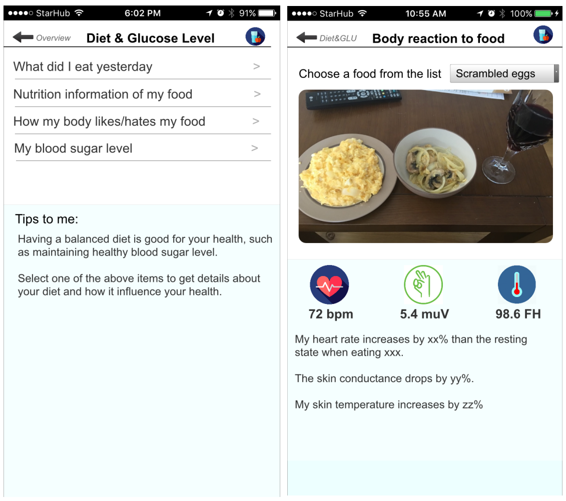
(1) Bar chart (2) Clock-view (3) Bubble chart (4) Affinity map

(5) Activity on geographical map (6) Sunburst chart (7) Calendar view (8) Radar chart

Figure 2. Templates of insight visualization

2.3 Mobile UI for lifelogging visualization

The output of theme finding and visualizations are integrated in a mobile app according to the selected topics. The purpose is to give consistent user experience in exploring the visual insights. Under each theme, there is a list of questions that is considered to be relevant to a user. For example, four items are included in the "Diet & Glucose Level" page (Figure 3a). One can select an item to get details, e.g. "how my body like/hats my food" shows the physiological value when that food was being consumed.



(a) Theme questions (b) Physiological responses

Figure 3. Sample user interface of insight visualization

3. RESULT

Insights are presented following the themes of the LIT topics. Under each LIT topic, multiple visualization tools are used. The comparison between two users (T5) is included under each of the other LIT topics where applicable.

3.1 Diet and blood sugar level

This topic is only reported for u1, whose dietary log and blood sugar level are available.

A total of 111 unique foods and 18 drinks are extracted from the log. The top foods and all drinks are shown as the bar chart in Figure 4. The GL value of each food/drink is shown as a red diamond marker. From the chart, one can observe that u1 had a balanced diet, with a variety of food and varying levels of GL. His favorite foods included cereal, vegetables and boiled eggs and cake. He drank coffee frequently (two coffees a day), and occasionally drank coconut water, mint tea, and white wine.

A bubble chart is generated to illustrate the possible factors that contribute to blood sugar level (Figure 5). The glucose level is color-coded as circles: lighter green means lower glucose level and darker green means higher level. Both x-axis (GL from food) and y-axis (energy consumption) are factors that may modulate glucose level. It is observed that the circles become darker as x increase, whereas no such trend is observed along the y-axis (energy consumption). Through correlation analysis, it is found that food intake contributes to GL with a correlation coefficient of 0.41 (p=0.03). Meanwhile, energy consumption and quality of sleep is only weakly correlated with glucose level, and it is not statistically significant.

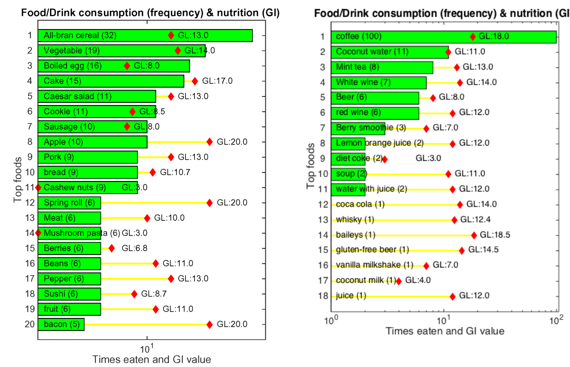


Figure 4. Nutritional information (glycemic load) of frequent food (u1).

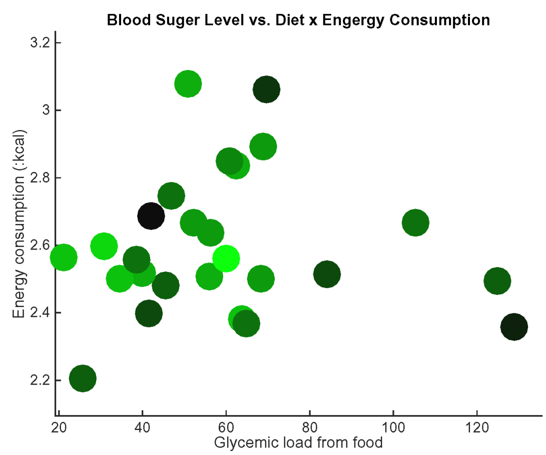


Figure 5. Factors that may affect blood sugar level.

3.2 Exercise and physical activity

This paper shows four modes of exercises, namely walking, running, hiking, gym/yoga. None of the users engaged in regular exercises other than normal walking. On average, u1 walked 19.2 minutes per day, and u2 walked 47.4 minutes per day. Considering the low frequency of the other activities (running, hiking, gym/yoga), they are summarized as the number of occurrence and average duration of an occurrence, as shown in Figure 6. User 1 ran twice during the recording period (59 days) and he only spent about 1.3 minutes in the run. User 2 ran 3 times in 31 days, each run last about 4 minutes. Both users had one hiking experience, where u1 spent 22.5 minutes and u2 spent 79 minutes. U1 practiced yoga once for about 2.6 minutes, and u2 went to the gym 4 times, he spent an average of 110 minutes for a gymnastic exercise.

Using the clock-view of activates, one can study the temporal pattern of users' exercises. In Figure 7, each clock shows activity patterns (color-coded according activity type) over 12 hours. One can easily tell when a user is likely to engage in an activity. For example, the hiking activity of u1 is around 6-7:30pm; u2 is most likely to go to the gym 14:00-16:00 or 19:30-21:00. Comparing top-row (u1) and bottom-row (u2), u2 spends more time exercising.

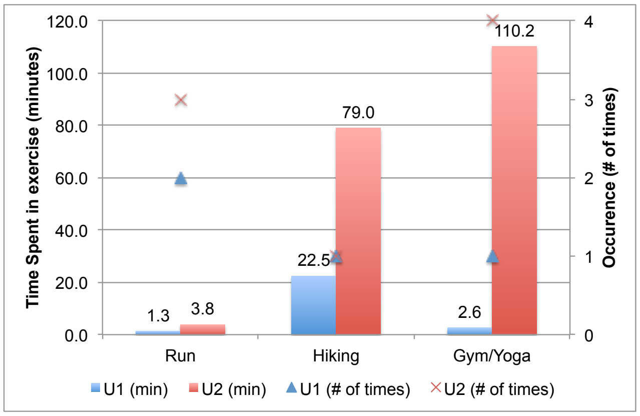


Figure 6. Exercise of two users.

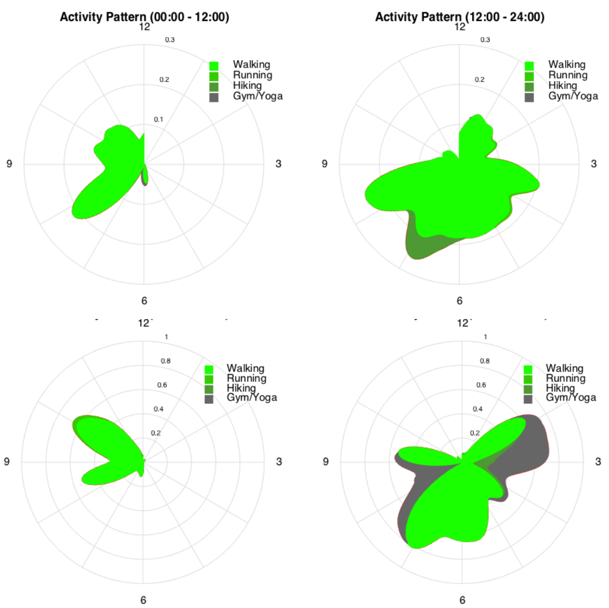


Figure 7. A clock-view of exercises/physical activities (top-u1, bottom- u2). The radius of the vector shows the likelihood of the user in an activity (calculated as the number of times it happens over a week). The left column shows the AM hours, and the right column shows PM hours.

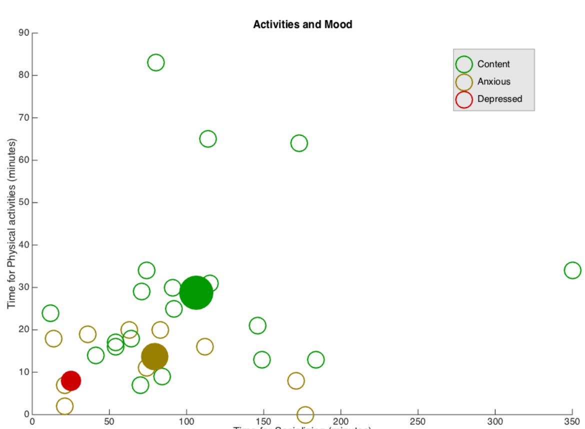


Figure 8. Mood vs. socializing and physical activity (u1)

3.3 Social

In activity annotation, socializing is retrieved as moments when a user was interacting with one or a few persons who appeared in front of the user at a social distance. Such distance is estimated based on the size (in pixels) of detected persons in an image. User 1 spent 94 minutes per day in socializing, and u2 spent about 54 minutes. For deeper insights, it is investigated if socializing activities contributed to the user's emotional wellbeing. To do so, a bubble chart is drawn where the x-axis is the time spent in socializing in a day, y-axis is the time spent in exercising. Moods are drawn as empty circles in different colors (Figure 8). Solid circles show the centroid of the variables, e.g. the solid green circle shows the average time spent in socialization and average time in physical exercise when the mood is "Content". There is an obvious trend that both socializing and physical activities lead to better mood. It should be noted that the sample size is pretty small, only 31 days has mood log, and only one "Depressed" mood is reported. Therefore, the conclusion may not be strong.

3.4 Where: location and movement

3.4.1 Modes of commuting

A summary of the modes of commuting is show in Figure 9. Car driving is the main commuting mode of u1, where he spent around 108 minutes in the car everyday. User 2 adopted diverse modes of commuting, spending about 15 minutes per day in bus, car and train. User 1 took planes more often, flying three times in the period totaling to 17 hours' in the cabin. User 2 flied only once, which was a 6-hour trip from Athens to Dublin.

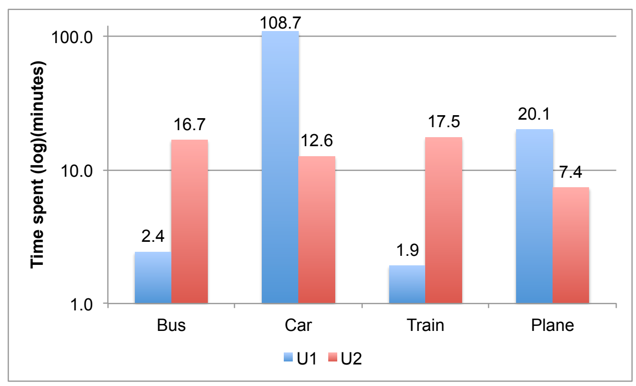
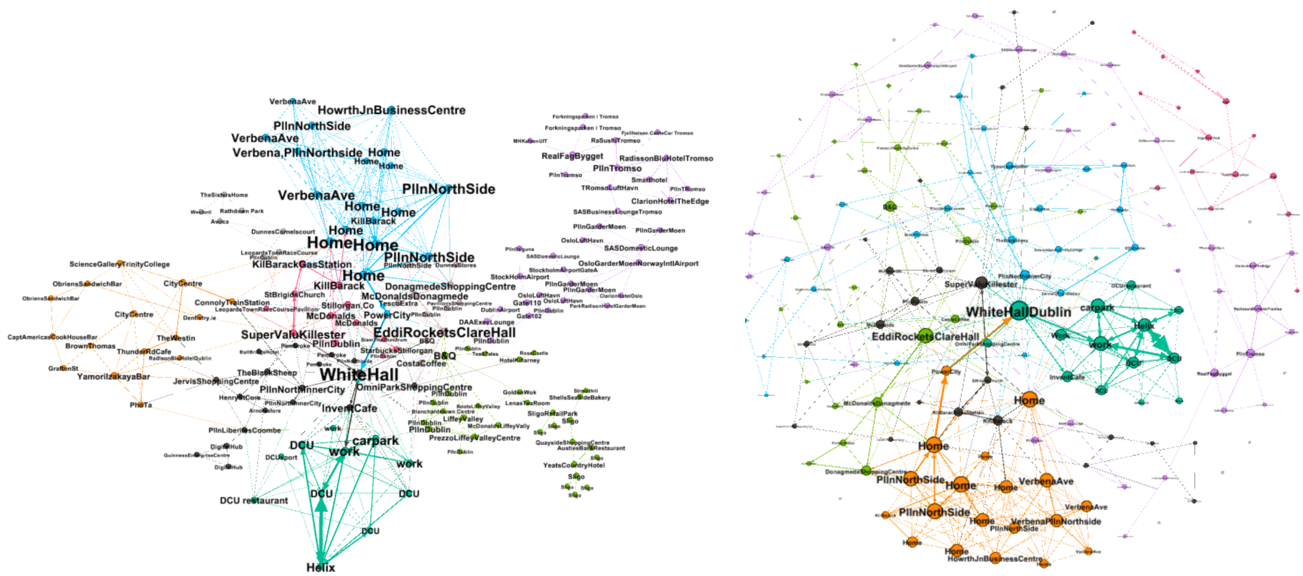


Figure 9. Commuting statistics of two users

3.4.2 Affinity map of locations

An affinity map is built based on the connectivity of locations. Two locations are connected when they are in sequence and co-located within reasonable distance. Based on the connectivity, clusters are generated where locations that are co-located in close geographical distance and temporal distance are clustered together. The result is visualized in an affinity map.

The affinity map summarizing the transit profile of u1 is in Figure 10 (a). It is a directed graph, where each node corresponds to a unique GPS coordinate. Note that the location tag is extracted from Google Map instead of the original metadata, which are not accurate. An edge corresponds to a transition from one tagged location to tagged location in sequence. In Figure 10 (a), the thickness of the edge denotes the relative number of times a transition was seen. A directed graph provides a summary of the important travel patterns of u1 in a single snapshot and also adds necessary rigor for a formal analysis of the patterns. For example, the *modularity* of the graph can be computed to give a clustering



(a) Master view of the affinity map

(b) Affinity map showing walking as the commuting mode

Figure 10. Affinity map showing relationship among places

structure between the nodes. The color of the nodes denotes the 'modularity' of the graph and nodes related to 'Work', 'Home', 'Shopping', 'Cafés/Bars', 'Overseas travel', 'Vacation trails' etc. are clustered automatically. This graph can answer some questions at the macro-level such as, *What are the most frequent transitions made by the user?* and some micro-level questions such as, *Where does u1 shop before leaving overseas?*. In addition, the graph reveals some patterns that may be relevant to the user's habits. For example, the graph suggests that u1 normally went from Home to Work directly (passing through 'Power City' and 'WhiteHall' on the way) but the same path was not always followed when returning home from work. The graph shows that u1 often took a path via 'Super value Killester' and 'Kill Barack Gas station') on the way back home.

Affinity Maps for Specific Mode of Transit: The graph above raises interesting questions such as, *'Does the user usually drive between nodes A and B? What are the main locations that the user covers by walking?'* etc. To gain further insights into the travel patterns, we extracted graphs which had the same nodes as above, but the directed edges were formed only if the user adopted a particular transport mode when leaving a particular source node and used the same mode when entering the subsequent target node. As an example, when we consider "Walking" as the transport mode, the corresponding affinity maps is shown in Figure 10 (b).

3.4.3 Sunburst chart

While the affinity map gives spatial ordering of the transitions, it misses some important temporal aspects and also some finer details such as activities performed in those locations. To obtain this, we use a sunburst representation (Figure 11) to summarize the distinctive activities performed in a spatio-temporal unit (e.g. *what are the activities performed at home in the early morning*?). To obtain this, we fist considered the different location spaces generated by the modularity values (color codes in Fig 10). The following were the chief location classes (Home, Work, Café/restaurant, Shop, Other Town, Overseas, Other-known, Unknown). Each color in the first concentric ring in the sunburst chart represents one of these locations. The second ring represents time in 1-hour intervals (e.g. 0-1AM, 1-2AM etc.). The third layer represents activities (Eat, Sleep, Travel, Exercise, Social). The final ring represents sub-activities (Travel: car, bus, train, plane, walk; Exercise: walk, ran, gym, hike, etc.). The width of each slice in the final ring represents the count of a particular tuple of [location, time, activity, sub-activity]. This representation gives us the spatio-temporal distribution of the activities performed by the user. Clicking on one of the slices in a ring would open up a sunburst chart specific to that slice. For example if the user is only interested at the distribution of activities at home, a new sunburst chart showing only three rings (time, activity and sub activity) would appear to enable the user have a closer look of the temporal distribution of activities at Home. This view provides an interactive way of slicing and dicing the data.

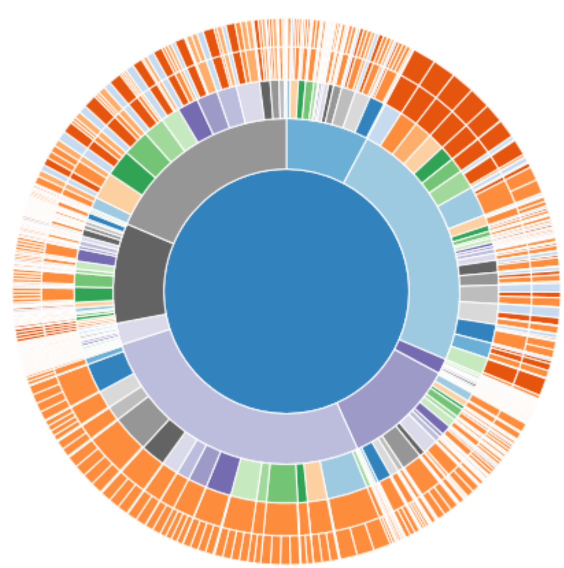


Figure 11. Sunburst view of spatio-temporal distribution of activities (u1)

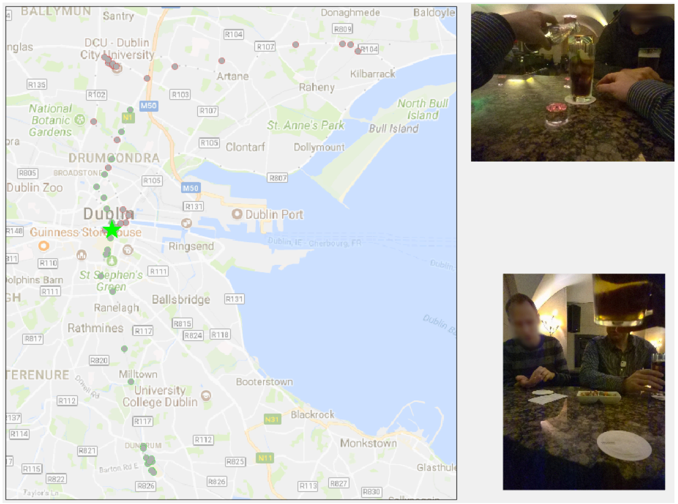


Figure 12. When-where-what: combining temporal, spatial, and multimedia information (u1&u2)

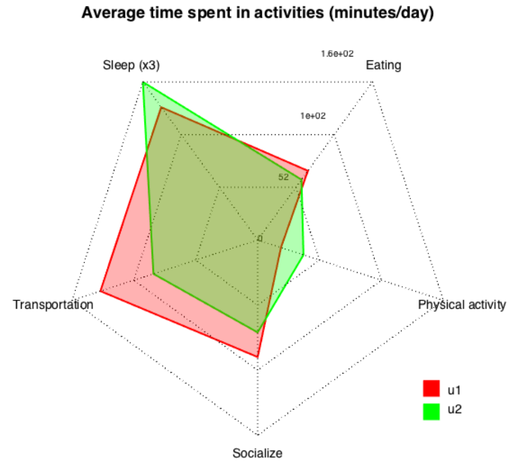


Figure 13. Comparing u1 and u2 on time spent in activities.

3.4.4 Co-location of two persons

Relying on the GPS information and merging the temporal information and multimedia into a unified view, the co-location of two persons can be visualized in an animation. By setting a threshold distance of 30 meters, it is found that two users meet each other at dinnertime in downtown Dublin on October 4th, 2016. Figure 12 shows a snapshot of the UI that marks the location of two users on that day, together with a slide show of lifelog photos of both users.

3.5 Comparison

Beyond the above illustrations, a radar chart is used to compare multiple dimensions of user activity. In Figure 13, time spent in five activities is presented where two users' data are overlaid and color-coded. Apparently, u1 spent more time in commuting, eating and socializing, whereas u2 has more physical activity and enjoyed more sleep.

4. DISCUSSION

4.1 Data recording and processing

This study makes use of data recorded by active lifeloggers, with rich content and well-structured data format. Despite the good effort by the organizers and lifeloggers, it is inevitable to include missing and noisy data, and data contaminated by inconsistencies. This could have been caused by the fact that the data were recorded using multiple devices, which have adopted different data format and had varying levels of accuracy. For example, the GPS data may not by captured consistently especially when users went indoors; the location information may have varying level of granularity – from very specific (e.g. "*The Helix*" and "*Yamamori Izakaya Bar*") to very general (e.g., "*Place in Dublin*"); the physiological data may be contaminated when users had excessive body movement. In addition, the time stamp may not be perfectly synchronized in multiple devices. Moreover, when manual input was involved, e.g. writing the food log or health log, the input may not be complete and is subject to logger's preferences and bias. Meanwhile the person who does the analysis may not be aware of the loggers' behaviors, thus failing to interpret the data correctly. It is desirable to establish a common platform where different devices follow a common protocol, to minimize misinterpretation of the logging data.

4.2 Customization and personalization

Insights are highly individualized. It is probably less interesting when a user receives similar sets of insights. It is desirable to allow users to customize the templates to fit his/her personal interest.

4.3 Insight interpretation

The visualization results, while useful for a quick understanding of the data, is still subject to interpretation. Depending on the experience and technical proficiency of a user, it is not always straightforward to make sense from the visuals. It is useful to design templates with adequate instructions. Alternatively, some intelligent tools can be provided to interpret the visualization of the data.

4.4 Scientific rigor vs. user experience

From the perspective of creating engaging user experience, it is useful to organize insights as a set of themes as presented in Section 2.3. On the other hand, such insights may not strictly follow scientific principles, e.g. correlation between variables may be unwittingly understood as causal relationship, and a layman user may overgeneralize the result. Therefore, designers must be cautious not to mislead users into wrong conclusions and actions.

5. CONCLUSION

In this paper, we present a method for generating insight from personal lifeloggs. We first annotate the data automatically according to a set of activities relevant to the selected themes (i.e. topics). Next, we design a set of insight visualization templates to show insights with respect to various topics and themes. Through examples, we show how the templates are used to visualize and understand insights generated. We also present a prototype mobile app to organize the insights into themed diary, which provides a holistic view of the lifelog information and facilitate consistent user experience.

6. ACKNOWLEDGMENTS

The work is funded by the Singapore A*STAR JCO REVIVE Project (1335h0009).

7. ADDITIONAL AUTHORS

Liyuan Li (Institute for Infocomm Research, A*STAR, Singapore. email: lyli@i2r.a-star.edu.sg)

Vijay Chandrasekhar (Institute for Infocomm Research, A*STAR, Singapore. Nanyang Technological University, Singapore. email: vijay@i2r.a-star.edu.sg).

8. REFERENCES

[1] E. K. Choe, B. Lee, and M. C. Schraefel, "Characterizing visualization insights from quantified selfers' personal data presentations," *IEEE Computer Graphic and Applications,* vol. 35, pp. 28-37, 2015.

[2] C. Gurrin, H. Joho, F. Hopfgartner, L. Zhou, D.-T. Dang-Nguyen, R. Gupta., R. Albatal, "Overview of NTCIR-13 Lifelog-2 Task," in *NTCIR-13*, Tokyo, Japan, 2017.

[3] J. Meyer and S. Boll, "Digital Health Devices for Everyone!," *Pervasive computing,* vol. 13, pp. 10-13, 2014.

[4] A. Sellen and S. Whittaker, "Beyond total capture: A constructive critique of lifelogging," *Communications of the ACM,* vol. 53, pp. 70-77, 2010.

[5] J. Lin, A. G. del Molino, Q. Xu, F. Fang, V. Subbaraju, J. H. Lim, L. Li, V. Chandrasekhar, "VCI2R at the NTCIR-13 Lifelog-2 Lifelog Semantic Access Task," in *NTCIR-13*, Tokyo, Japan, 2017.

[6] F. S. Atkinson, K. Foster-Powell, and J. C. Brand-Miller, "International Tables of Glycemic Index and Glycemic Load Values: 2008," *Diab Care,* vol. 31, 2008.