Optimal time window for temporal segmentation of sensor streams in multi-activity recognition

Jonathan Liono, A. K. Qin, and Flora D. Salim

School of Science

Department of Computer Science and Information Technology

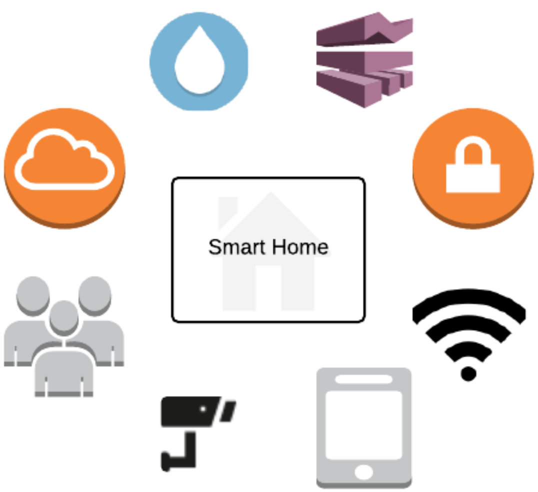
RMIT University

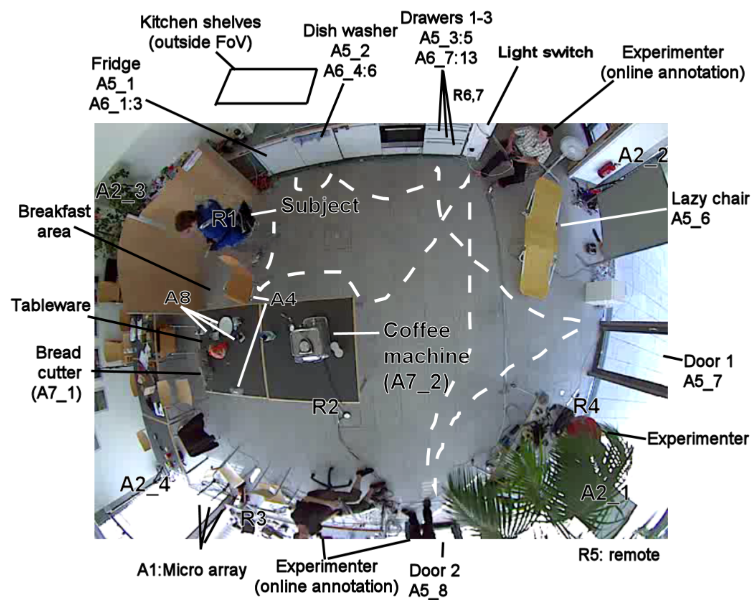
Outline

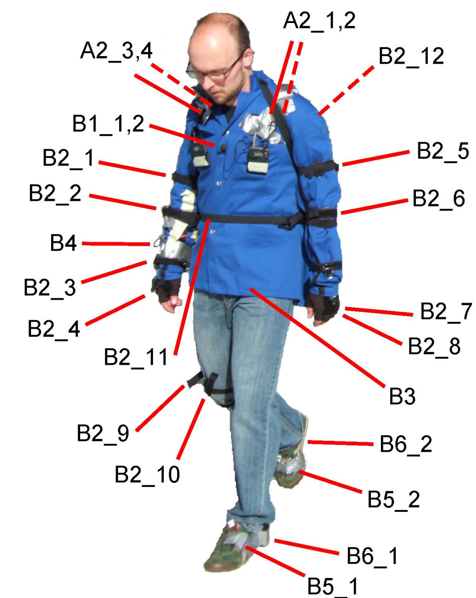
- Introduction
- Background
- Problem Definition
- OPTWIN algorithm
- Experiment and Evaluation
- Conclusion
- Future Works
- Q&A

Introduction

- Ubiquitous computing offers limitless applications to support everyday life.
- These applications are enabled through the availability of mobile sensors that can be used to recognise and track activities.
- Emerging challenges are derived from continuous streaming of data from heterogeneous sensors.



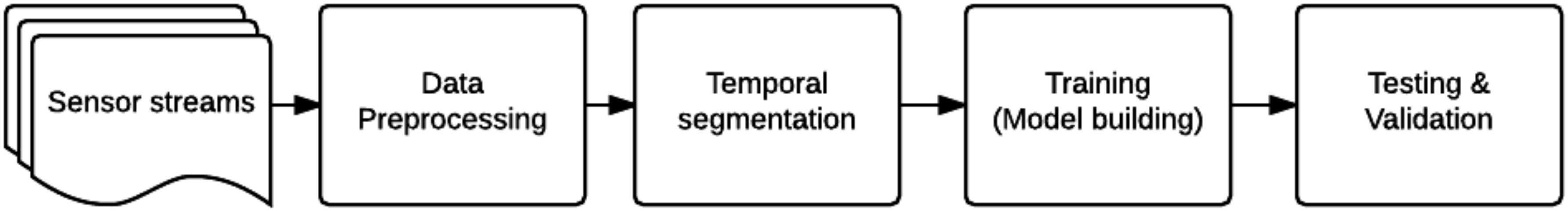




Smart home scenario from OPPORTUNITY dataset [1]

Background

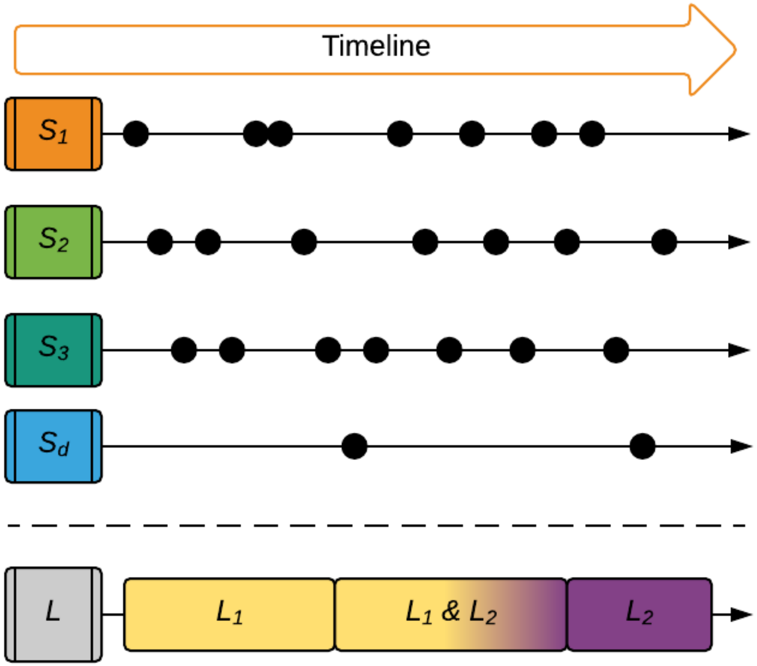
- Human annotation data may not be aligned consistently in terms of boundary between activities.
- General process in activity recognition:



- Real-time activity recognition requires feature extractions from temporal segmentation.
- Since sensor data is represented in time series, time interval based temporal segmentation is typically performed on the sensor streams.
- Therefore, impurity during temporal segmentation process can inflict a significant impact on classifier's results.

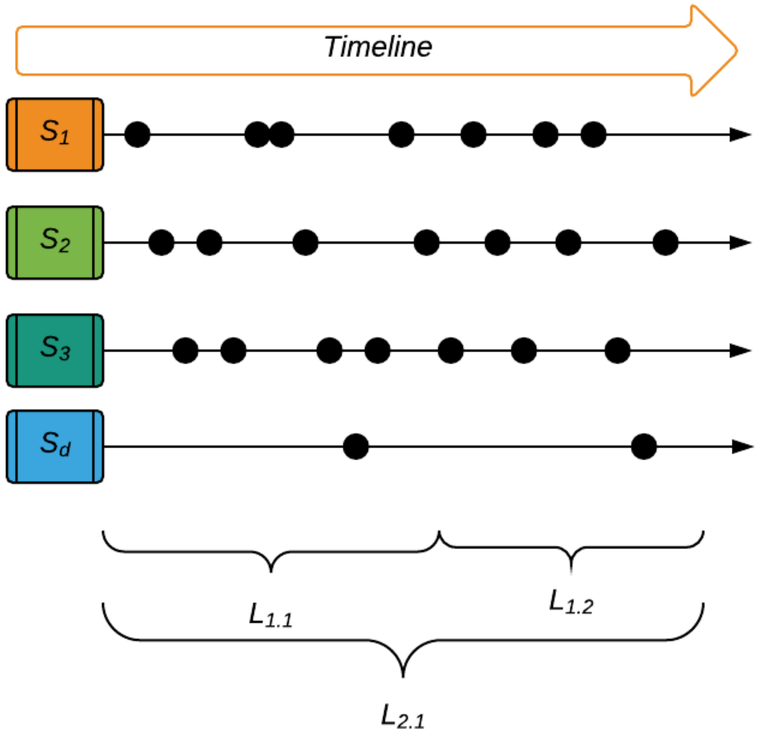
Problem illustration

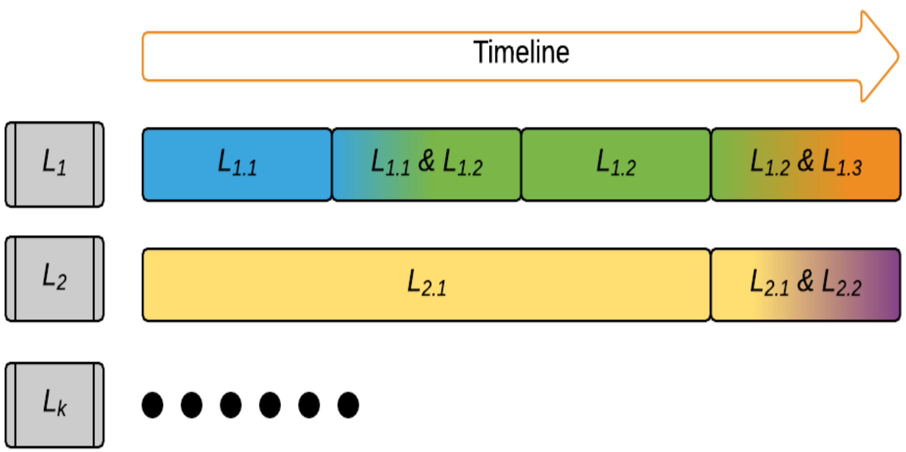
– Impurity in temporal segmentation
  An impure segment is composed of more than one label.



Problem illustration

- Multi-label problem in temporal segments





Problem definition

- Typical process in data processing involves temporal segmentation.
- Common approach in finding a suitable time window is to perform sensitivity analysis.

- Alternatively, this can be achieved through the assumption of 1 second window size as the heuristic [2,3,4,5].
- In addition, many of current studies are restricted for single activity recognition scenarios (i.e. single-label problem) [6,7,8,9,10].

- In order to find optimal time window size for temporal segmentation, multi-objective approach is proposed by gaining the balance between:
  - Minimising impurity of segments in data streams
  - Maximising factor for class separability based on k number of label sets.

- To achieve the objective, two metrics are used:
  - Impurity proportion of segments
  - Class separability score

Metric 1: impurity proportion

- An impure segment contains a mixture of multiple annotations (which can indicate important information of transition between activities).
- The impurity proportion can be computed via:

$$pimpure = \frac{m_{impure}}{m}$$

where m_{impure} is the count of impure segments over total m segments in data streams.

Metric 2: class separability – KL divergence

- Class separability refers to distinct separation between features with respect to each class label.
- The objective of computing class separability score in data streams:
  - Identifying which features are dominant towards a class label.
- Kullback-Leibler divergence [11] is used to calculate class separability score. It is a measure of difference between two probability distributions *P* and *Q*.
- Assumption: streaming sensor data is constrained by numeric features.
- For each feature f, we calculate maximum score of the class separability (i.e. maximum divergence of f) as:

$$mdivergence_f = \max_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} KL_f(i, j),$$

where n is number of class labels in a label set L_k and $KL_f(i, j)$ is the KL distance between two distributions.

- The KL distance between two distributions (histograms) class labels i and j corresponds to the following:

$$KL_f(i, j) = \sum_{m=1}^{b} Pr_f(m \mid i) \, log\left(\frac{Pr_f(m \mid i)}{Pr_f(m \mid j)}\right)$$

where b is the number of bins in the histograms.

OPTWIN: Finding optimal time window size

- The OPTWIN algorithm (refers to our paper) requires three main input parameters: ws_{start}, ws_{end}, and ws_{step}.
- In order to find the optimal time window size for each label set, the previously computed metrics are used: impurity proportion and normalised squared deviation of divergence score.
- As both values are normalised between 0 and 1, the first intersection point at a time window size is derived as the optimal one.

- Normalised squared deviation of divergence score ndd_i can be computed as the following:

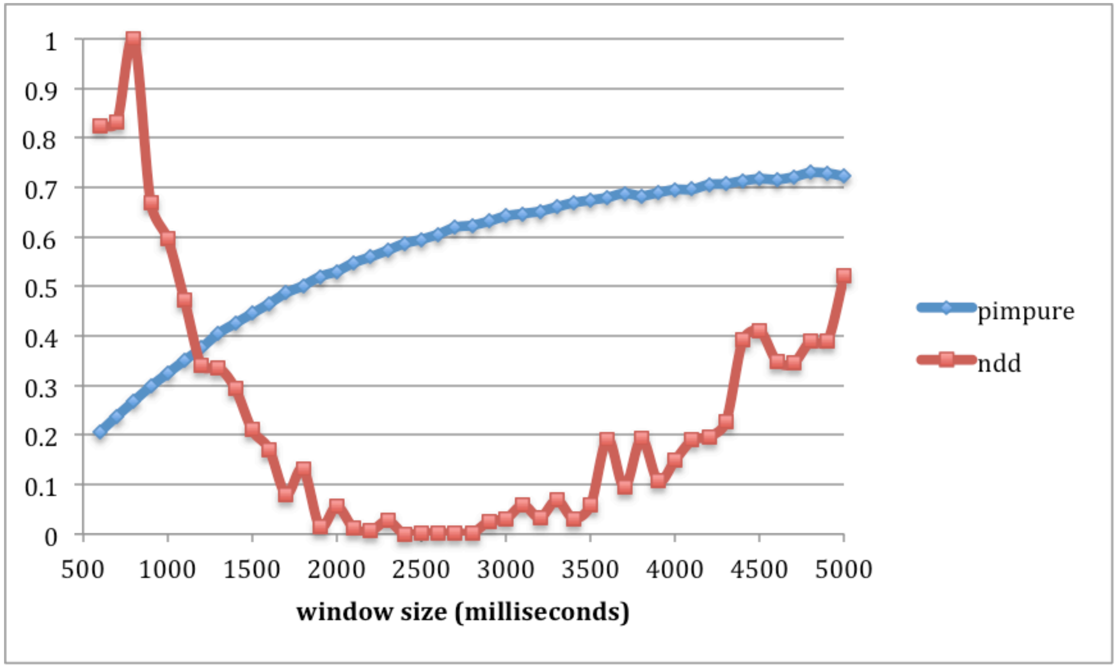
$$ndd_i = \frac{dd_i - \min(dd)}{\max(dd) - \min(dd)}$$

- where squared deviation dd_i of divergence score ds_i is defined as:

$$dd_i = (ds_i - \overline{ds})^2$$

OPTWIN: Finding optimal time window size

- As previously mentioned, the optimal time window size is found based on the intersection of impurity proportion and normalised squared deviation of divergence score.
- For example, 1.2 seconds is derived as optimal window size for multi-activity label set:



Data and Features

OPPORTUNITY
Activity and Context Recognition with Opportunistic Sensor Configurations

- Dataset: OPPORTUNITY [1] from UCI repository (4 human subjects)
- 101 features (from on-body sensors) are used out of total 242 features
- Each instance is associated with labels from the following:
  - High Level Activity (HLA) denoted as L_1
  - Locomotion Activity (LA) denoted as L_2
  - Gesture Activity (GA) denoted as L_3

$$L_1 = \left\{ \begin{array}{l} \text{Relaxing,} \\ \text{Coffee time,} \\ \text{Early morning,} \\ \text{Cleanup,} \\ \text{Sandwich time,} \\ \text{None} \end{array} \right\}$$

$$L_2 = \{\text{Stand, Walk, Sit, Lie, None}\}$$

$$L_3 = \left\{ \begin{array}{l} \text{Open Door 1, Close Door 1,} \\ \text{Open Door 2, Close Door 2,} \\ \text{Open Fridge, Close Fridge,} \\ \text{Open Dishwasher, Close Dishwasher,} \\ \text{Open Drawer 1, Close Drawer 1,} \\ \text{Open Drawer 2, Close Drawer 2,} \\ \text{Open Drawer 3, Close Drawer 3,} \\ \text{Clean Table, Drink from Cup, Toggle Switch,} \\ \text{None} \end{array} \right\}$$

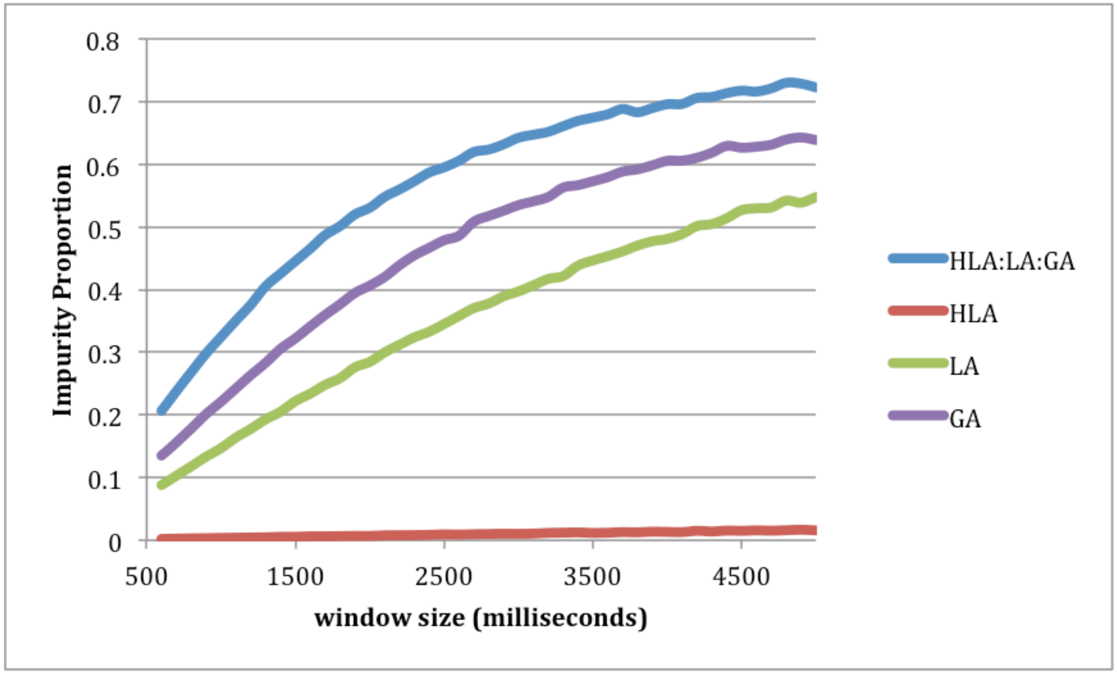
Feature Extraction

- For each window (temporal segment), the following statistical features are extracted:
  1. Mean
  2. Minimum
  3. Maximum
  4. Standard deviation
  5. IQR (interquartile range)
  6. Median
  7. RMS (root-mean-square)

- In total, 707 features are used in model building to recognise activity sets.

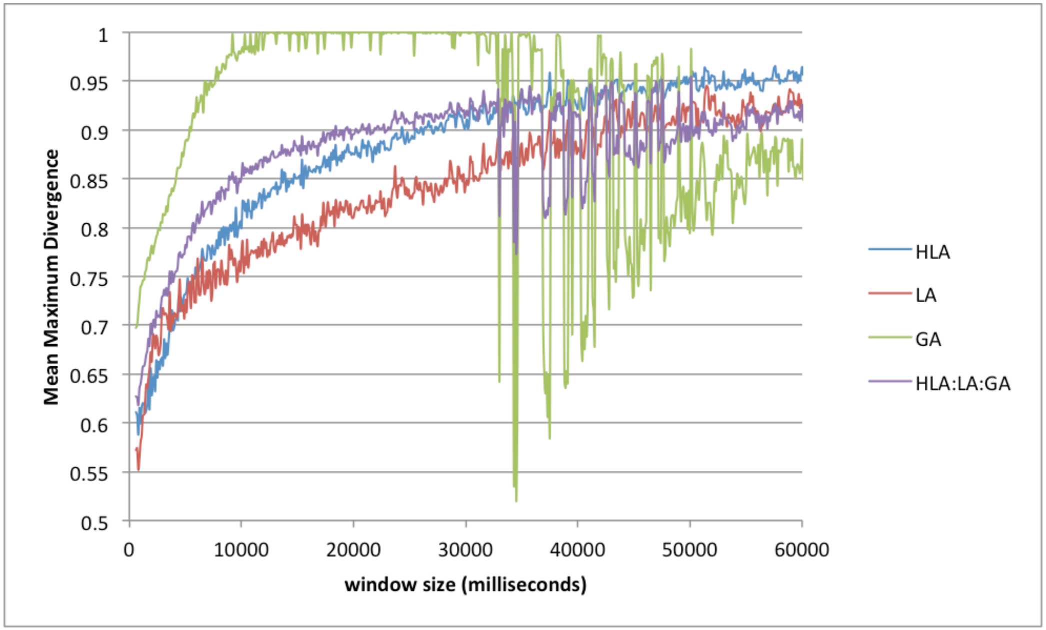
Data characteristics: impurity proportion

- Impurity proportion for multi-activity labels is generally higher than single activity labels.



Data characteristics: mean maximum divergence

- Mean maximum divergence increases as time window is expanded.
- It is expected to degrade after reaching the peak at a certain time window size.
- However, it should be noted that there is a potential degradation of label set dimension as the window size grows.



Experiments: classifiers

- To recognise the activities, the following classifiers are leveraged:
  - Naïve Bayes (**NB**) classifier
  - Decision Tree (**DT**) classifier
  - Random Forests (**RF**) classifier

- Experiments are performed using WEKA 3.8, with default parameters for each classifier.

- For multi-activity recognition, a new label set (denoted as **HLA:LA:GA**) is generated from composition of high level activity, locomotion activity, and gesture activity.

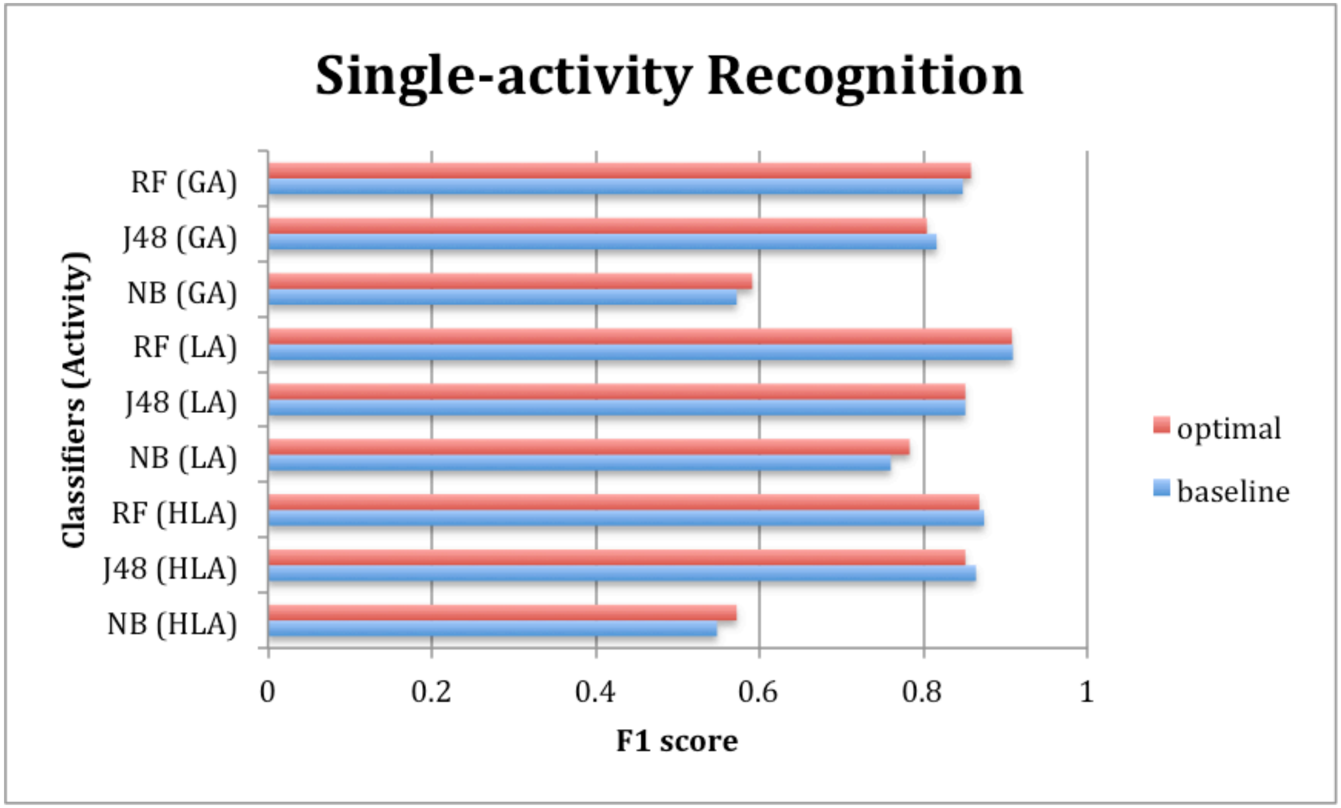
Training and Testing

- Baseline: 1 second heuristic.
- Performance of classifier is evaluated based on F1 score (harmonic mean between precision and recall).
- Validation process: 10-folds CV (Cross Validation).

- The following time window sizes are produced from OPTWIN:
  - **HLA**: 2.4 seconds - single-activity recognition
  - **LA**: 1.5 seconds - single-activity recognition
  - **GA**: 1.4 seconds - single-activity recognition
  - **HLA:LA:GA**: 1.2 seconds - multi-activity recognition

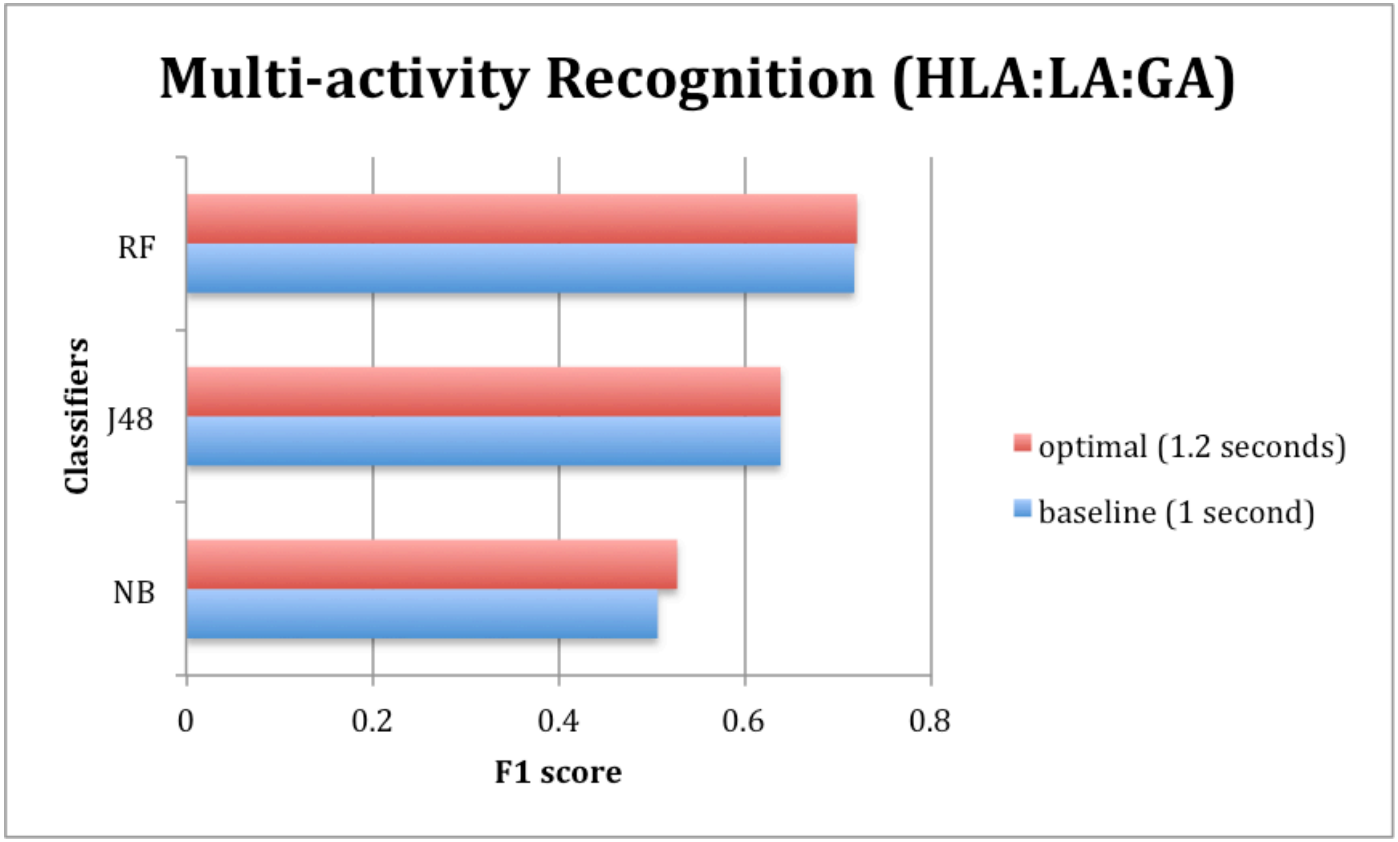
Evaluation

- Single Activity Recognition



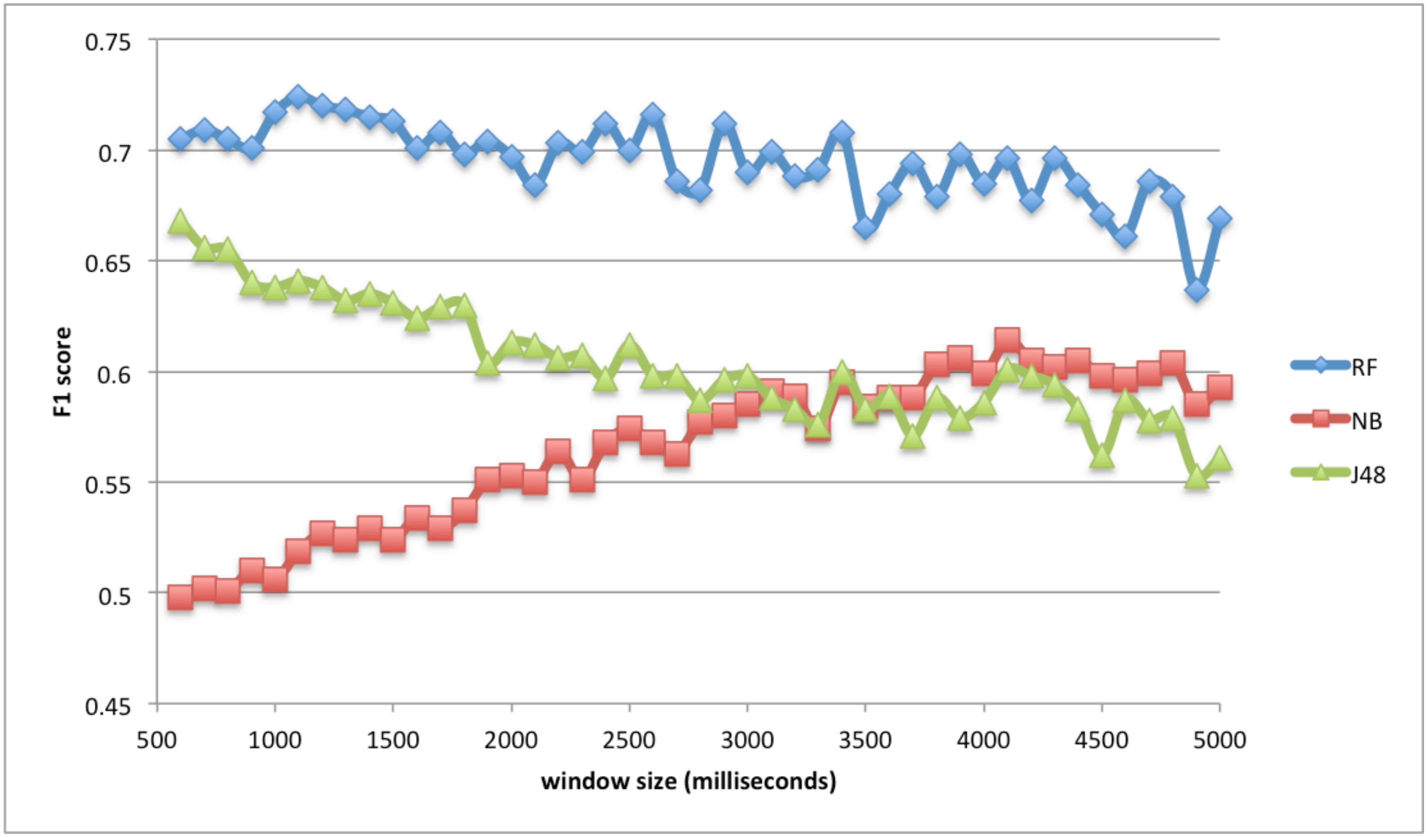
Evaluation (Cont'd)

- Multi-activity recognition



Sensitivity Analysis

- The sensitivity analysis (for multi-activity recognition) is performed within the same boundary defined for OPTWIN parameters.



Conclusion

- OPTWIN technique is proposed to find optimal window size for time-interval based temporal segmentation.
- Hence, the contribution of the technique is achieved through:
  - Minimising label impurity while maximising class separability for time-interval based temporal segmentation.
  - Adaptation for multi-activity recognition scenario (multi-label problem)
- The optimal time windows are found to improve the quality of classifier performances from simple heuristic (such as 1 second window size).
- It is found to be simple and effective. As a result, it is less time consuming compared to typical sensitivity analysis tasks in finding reasonable time window size.

Future Works

- Adaptation of window size to expand and shrink in data stream according to activity context.
- More robust and accurate classification solution for inherent multi-label problem in multi-activity recognition, where the study of dependency between activity label sets should be considered.

References

1. Sagha, H., Digumarti, S.T., Millán, J.D.R., Chavarriaga, R., Calatroni, A., Roggen, D. and Tröster, G., 2011, October. Benchmarking classification techniques using the Opportunity human activity dataset. In *IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2011* (pp. 36-40). IEEE.
2. Pirttikangas, S., Fujinami, K. and Nakajima, T., 2006, October. Feature selection and activity recognition from wearable sensors. In *International Symposium on Ubiquitious Computing Systems* (pp. 516-527). Springer Berlin Heidelberg.
3. Zhu, C. and Sheng, W., 2009, May. Human daily activity recognition in robot-assisted living using multi-sensor fusion. In *IEEE International Conference on Robotics and Automation, 2009. ICRA'09.* (pp. 2154-2159). IEEE.
4. Wang, L., Gu, T., Tao, X. and Lu, J., 2009, November. Sensor-based human activity recognition in a multi-user scenario. In *European Conference on Ambient Intelligence* (pp. 78-87). Springer Berlin Heidelberg.
5. Gordon, D., Czerny, J. and Beigl, M., 2014. Activity recognition for creatures of habit. *Personal and ubiquitous computing*, 18(1), pp.205-221.
6. Chen, L., Hoey, J., Nugent, C.D., Cook, D.J. and Yu, Z., 2012. Sensor-based activity recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C* (Applications and Reviews), 42(6), pp.790-808.
7. Riboni, D. and Bettini, C., 2011. COSAR: hybrid reasoning for context-aware activity recognition. *Personal and Ubiquitous Computing*, 15(3), pp.271-289.
8. Hemminki, S., Nurmi, P. and Tarkoma, S., 2013, November. Accelerometer-based transportation mode detection on smartphones. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems* (p. 13). ACM.
9. Feng, T. and Timmermans, H.J., 2013. Transportation mode recognition using GPS and accelerometer data. *Transportation Research Part C: Emerging Technologies*, 37, pp.118-130.
10. Kim, Youngsung, et al. "Activity recognition for a smartphone based travel survey based on cross-user history data." *22nd International Conference on Pattern Recognition (ICPR), 2014*. IEEE, 2014.
11. Kullback, S. and Leibler, R.A., 1951. On information and sufficiency. *The annals of mathematical statistics*, 22(1), pp.79-86.

Thank you: Q&A

- Email: jonathan.liono@rmit.edu.au