Recognizing Activities of Daily Living from Patterns and Extraction of Web Knowledge

Isibor Kennedy Ihianle
University of East London,
London, United Kingdom
u1051232@uel.ac.uk

Muhammad Awais Azam
University of Engineering and
Technology, Taxila, Pakistan
awais.azam@uettaxila.edu.pk

Usman Naeem
University of East London,
London, United Kingdom
u.naeem@uel.ac.uk

Abdel-Rahman Tawil
University of East London,
London, United Kingdom
a.r.tawil@uel.ac.uk

Ubicomp/ISWC'16 Adjunct, September 12-16, 2016, Heidelberg, Germany
ACM978-1-4503-4462-3/16/09.
http://dx.doi.org/10.1145/2968219.2968440

ABSTRACT
The ability to infer and anticipate the activities of elderly individuals with cognitive impairment has made it possible to provide timely assistance and support, which in turn allows them to lead an independent life. Traditional non-intrusive activity recognition approaches are dependent on the use of various machine learning techniques to infer activities given the collected object usage data. Current activity recognition approaches are also based on knowledge driven techniques that require extensive modelling of the activities that needs to be inferred. These models can be seen as too restrictive, prescriptive and static as they are based on a finite set of activities. In this paper, we propose a novel "top down" approach to recognising activities based on object usage data, which detects patterns associated with the activity-object relationship and utilizes web knowledge in order to build dynamic activity models based on the objects used to perform the activity. Experimental results using the Kasteren dataset shows it is comparable to existing approaches.

Author Keywords
Activity Recognition; Pattern Analysis; Topic Model; Web Extraction; Ontology Model.

ACM Classification Keywords
I.5.2 Pattern Recognition: Design Methodology: Pattern Analysis; I.2.4 Knowledge Representation Formalisms and Methods: Representations (Procedural and Rule-Based)

INTRODUCTION
With the rising cost of providing assistance to the elderly and the cognitively impaired, it has become imperative to consider technology driven solutions which can help provide activity anticipatory solutions to independent and autonomous living. This area of research has attracted enormous attention and has seen efforts in the use of video [1], wearable sensors [2, 3] and wireless sensor networks [4, 5]. The use of wireless sensor networks has proven to be promising due to their low cost, ease of installation and most importantly being non-intrusive [4, 5]. Systems built to recognize, track and anticipate Activities of Daily Living (ADL) of the cognitively impaired can be of significant importance in the monitoring of their wellbeing and provision of assistive interventions. These approaches require object usage data to be classified using various machine learning techniques or modelling activities for recognition in knowledge driven approaches. The eventual activities recognized are dependent on labelling the acquired data to a finite set of regularly conducted activities or modelling the activities in the knowledge driven models based on generic "*know hows*". The classification process to recognizing and anticipating activities usually involves learning and inferring, which is dependent on the prior knowledge of activity patterns. But, human activities have been known to be diverse and complex. They can be performed in different ways. So, the interactions of the various objects in a home setting can result to a number of different activities which may not belong to the set of regularly conducted activities. As such, modelling activities solely reliant on a finite sets of activities can be seen to be restrictive, prescriptive or even static. This poses a big challenge and hence a gap in the recognition of the activities considering the boundless number of activities that could result from the interactions of objects within a home environment. In this paper, we propose an ADL recognition framework from activity-object patterns and web knowledge of object usage. We believe that the combination of the activity-object patterns from a topic model process can be complemented by the web knowledge enriched activity models to recognize infinite range of activities that could evolve from the object usage interactions. This paper describes how this proposed approach is able to recognize an infinite set of activities. The remainder of the paper is organized as follows. The related work section provides an overview of previous work related to the proposed approach. This is then followed by a description of the proposed recognition approach. The paper also provides a set of experimental results, which show that this approach is robust and comparable with existing work. This is then followed by discussion and conclusion.

RELATED WORK
The proposed framework builds on previous work done in the area of ADL recognition. This section reviews these efforts focusing on mainly activity patterns discovery, web knowledge extraction and ontology activity recognition models.

Activity Pattern Discovery
Probabilistic topic models inspired by the text and natural language processing community have been applied to discover and recognise human activity

Activity	Instances
Sleeping	33.42%
Toileting	0.65%
Go Out	49.6%
Showering	0.70%
Breakfast	0.23%
Dinner	1.00%
Drink	0.10%
Idle	14.12%

Table 1: Kasteren House A Dataset.

routines [2, 7]. The work we propose in this paper, extends our initial work Ihianle et al [17] by the inclusion of web knowledge extraction and ontology activity model. Our activity-object pattern discovery process is similar to Huynh et al [2] and Katayoun and Gatica-Perez [7]. Huynh et al [2] applied the "bag of words" model of the Latent Dirichlet Allocation (LDA) to discover activities like dinner, commuting, office work etc. The process involved activity discovery of partitioned sensor segments of each time window. Also an LDA topic model approach was applied by Katayoun and Gatica-Perez [7] to discover routines from mobile phone data. While Huynh et al [2] used wearable sensors attached to the body parts of the user, Katayoun and Gatica-Perez [7] captured their data from a single mobile phone by the user. While it is not feasible to use only a single mobile phone or phones as in Katayoun and Gatica-Perez [7] to capture low level every day ADL, our work uses multiple state-sensor tagged to every day home objects to capture object use and user activities in the home setting. Our work also significantly differs from Katayoun and Gatica-Perez [7] and Huynh et al [2] with the web knowledge of object usage and the ontology activity model.

Web Knowledge Extraction

We apply a web knowledge extraction inspired by the Natural Language Processing (NLP) community in Named Entity Recognition NER [8] and Relation Extraction [9]. This process was also applied by Perkowitz et al [10] to mine entities for their proposed model from a single web page. Palmes et al [11] mined the web to extract the most relevant objects according to their normalized usage frequency. They also approached activity modeling process by relying on the relevance weights of objects as the basis of activity discrimination rather than sequence information. Wyatt et al [6] extracted from the web a set of objects used to perform named activities. They show in their work that object-usage does not necessarily rely on a prescriptive set of activities following a "bottom-up" process of genre classification of activities, which they use to build their model. Our proposed framework also significantly differs by following a "top down" approach and an ontology activity model.

Ontology Activity Recognition Models

Ontology models follow Description Logic for the specification of conceptual structures and their relationships [13]. The authors of [14] and [15] followed generic activity knowledge to develop an ontology model for the smart home users. Whilst these approaches are commendable, they do not use evidenced patterns of object usage and activity evolution but rely on generic "*know hows*" and "*hows to*" to build ontology models. The ontology activity model which forms a major part of this proposed framework relies on activity topics, activity-object patterns initiated by the object use interactions and retrieved results of object use instances from the web.

ACTIVITY RECOGNITION MODEL

Activities are carried out by the interactions of objects within various locations in a home environment. Specific objects tend to be used in specific places for routine activities following patterns which are common to the user. In most cases, certain objects in specific locations have been known to be linked to particular activities. For the approach we propose through this framework, we use the Kasteren dataset [5]. This dataset was generated using a set of 14 state-change sensors. Table 1 show the annotated ADL including "idle" which corresponded to times when no significant activity took place. The non-intrusive nature of binary sensors suits the privacy and acceptability of the home occupants whilst object interactions are ongoing for activities. Given this dataset, we aim to recognize possible activities from object usage progressively along their timelines. Towards this, we form a "bag of sensor observations" from a continuous 60 seconds' partitions of the sensor data. We use the topic model Probabilistic Latent Semantic Analysis (pLSA) to discover the underlying activity topics and the activity-object patterns. We then perform web knowledge extraction of the activities associated to the different object usage. Further, we determine activities intersection using the discovered patterns on the extracted activities which we then assemble on an ontology activity model for recognition. An overview of the framework is as illustrated in Fig 1.

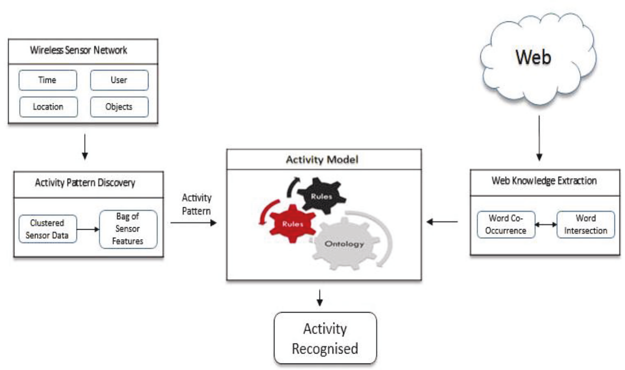

Figure 1. Conceptual Overview of the proposed Activity Recognition Framework.

Activity Pattern Discovery

The aim of the activity and pattern discovery process is to determine the activity-object patterns and the activity topics. This process takes advantage of the pLSA topic model assumption that there are hidden themes or latent topics which have associations with the words contained in a corpus of documents. It involves the use of "bag of words" in the corpus of documents which are generatively classified to latent themes or topics [16]. We conversely apply this assumption to the activity-object discovery context that latent activity topics would have associations with the features of sensor data or objects in the partitioned sequences of the "bag of sensor observations". For the "bag of sensor observations", we partition the Kasteren dataset of sensor observations into sequences using a sliding window of 60 seconds intervals to form a sensor-segment matrix. We also set the observed sensors to be represented as aliases such as Bedroom (B), Microwave (M), Groceries Cupboard (G), Fridge (F), Cup Cupboard (C), Toilet (T), Shower (S) etc. to be encoded onto the partitioned sensor sequences. The documents are presented in the form of sensor sequences $d_1 \ldots d_D$ composed of co-occurring sensor data observations along their timelines. If D is composed of sensor sequences $d_1 \ldots d_D$, d_i would be made of sensors represented as $x_{i1} \ldots x_{in}$ from X sensors of $x_1 \ldots x_n$. The pLSA assumes that a latent activity topic $\boldsymbol{z}$ from topics $z_1 \ldots z_k$ can be classified from $d_1 \ldots d_D$ as contained in D i.e for a sensor x_i contained $x_{i1} \ldots x_{in}$. there is a marked probabilistic relationship with the activity topic $\boldsymbol{z}$. In principle, there is a joint probability over $\boldsymbol{D}$ x $\boldsymbol{X}$ which corresponds to the sensor-segment matrix mentioned earlier such that a conditional independence assumption that d and x are independently conditioned on the state of the associated activity latent topic. The proposed framework is trained to infer activity topic probabilities by the iterative Expectation Maximisation (EM) algorithm. The first step, Expectation (E step) computes the posterior probabilities of the latent variables P(z_i|d_i) from the activity topic probabilities P(z_i) and the conditional probabilities of the sensors given the activity topics P(x_i|z_i). The second step, Maximisation (M step) updates the parameters from the E step by computing the new values for P(z_i) and P(x_i|z_i). The posterior inference of the EM iterative process can be computed from P(z_i|d_i) for each d_i. This computes the activity topic from the given sensor sequence. P(x|z) computes the probability of the sensors given activity topics. In the context of pattern discovery, the sensors or objects linked to an activity topic are computed from P(x|z) and this defines the activity-object pattern. Modelling an activity for recognition would rely on P(x|z) to define the compositional object usage that are linked to specific activity topic and P(z|d) defines the activity topics therein. The learning process of the proposed framework involves dividing the dataset into training subset (90%) and test subset (10%). The EM posterior inference is used on the training subset and then progressively on the test subset. Modelling an activity for recognition would rely on P(x|z) to define the compositional object usage that are linked to specific activity topic and P(z|d) which defines the activity topics therein.

Extraction of Web Knowledge

Given a set of objects within the home environment, the proposed approach retrieves web instances of activities associated with a specific object usage. The web knowledge extraction process we propose follows the W3C Resource Description Framework (RDF) assertions of subject (*s*), predicate or property (*p*) and object (*o*) [12]. We rely on the background web knowledge to extract statements having predicate knowledge of interest linked to a referenced subject or object. Using Google as the preferred search engine, instances of activities were retrieved from the first 100 pages of search. The predicate or property used for this process is limited to "*used for*" and object names as query threads. From the search results, all irrelevant word tokens were removed leaving only noun words (activities) associated to the objects and the "*used for*" property. Further, we formed a term-document matrix to determine the word co-occurrence and the vocabulary overlap in the constituent document. From initial investigation, we observed objects given their usage could have use for different tasks and activities as with `Cup` in `Making Tea`, `Making Coffee` and `Making Orange Juice`.

Computing Activity from Object Use Pattern.

We relied on the activity-object pattern discussed above and the web term-document of subject *(s),* predicate/property *(p)* and object *(o)* to prune and converge the retrieved results to an activity intersection or overlap. That is to say, if we have `Object_1`, `Object_2`, `Object_3` and `Object_4` in a pattern resulting to `Activity Topic 1`, the web retrieved result (activities) which is common or intersects all the objects in this pattern becomes the activity label for `Activity Topic 1`. This is also as illustrated in Fig 2 for the activity-object pattern having "*Microwave*", "*Fridge*", "*Freezer*", "*Groceries Cupboard*", "*Cups Cupboard*" and "*Plates Cupboard*" as objects would have "*Food*" as the activity overlap and the activity label for the activity topic. To determine activity overlap with regards to activity-object pattern and web retrieved activities (subjects), we computed the activities common to a set of objects otherwise known as activity overlap A. Recalling from

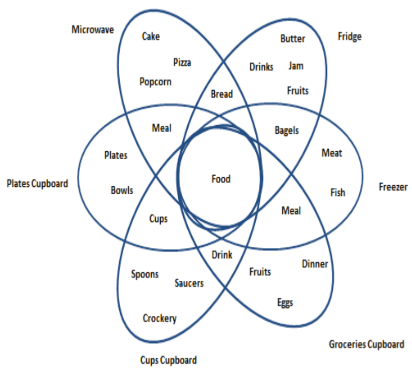

Figure 2. "Food" as the activity intersecting all the web retrieved results of object usage for "Microwave", "Fridge", "Freezer", "Groceries Cupboard", "Cups Cupboard", "Plates Cupboard".

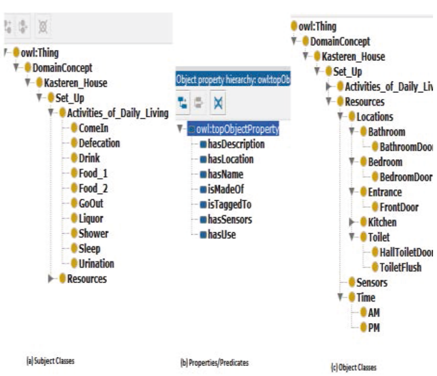

Figure 3. Context overview of Subject and Object Classes and the Properties in the Activity Ontology Model.

above that document d_i has sensor sequence $x_{i1} \dots x_{in}$. If each of these sensors in d_i have activities $A_1 \dots A_n$ they are used for with regards to the retrieved results from the web, the activity overlap A is then computed from $A_1 \cap \dots A_n$ for all the sensors $x_{i1} \dots x_{in}$. The activity overlap A would be performed by the use of all these objects corresponding to sensors $x_{i1} \dots x_{in}$ and this defines the activity topic for this sensor sequence. We use these to build the activity model for recognition.

Ontology Activity Model for Recognition

For the activity recognition process, we modelled the computed activity overlap as corresponding activities following the activity-object patterns into an ontology activity model. We followed an activity hierarchy formalism enabled by the ontology editor Protégé [18] to create subject and object class nodes. For each of the pattern, the subject and object are modelled as class and individual entities. Object and data properties which represent the predicates are modelled for each of the subject and object classes. A subclass in this structure is seen to have all properties of a super class. The sensor classes modelled in this process are made to abide by the object/data property domain and observing the range restrictions of the associated subjects and objects as the case may be. This process then allows for a sensor and object based query linked to subjects or objects which retrieves the most similar activity label. Activities are described through class equivalence axiom which links them to object usage. A Description Logic DL reasoner (e.g. Fact ++, Pellet) uses these modelled instances relative to object usage to classify ongoing activity. The specification of an activity in this process is built on the theories of description logic DL and reasoning which supports consistencies, subsumption, satisfiability, equivalence and disjointness [14]. Theoretically, if a subject is an instance of an activity to be recognised from the observation of an object *(o)* with its relationship specified as *(p)*. The reasoner checks for the equivalency and the subsumption of *(o)* in all *(s)* in the model with the specification of *(p)* to determine *(s)* as the recognised activity. The activity recognition process is enhanced by assembling activity-object patterns and retrieved web results in an ontology activity model as illustrated in Fig 3.

EXPERIMENTS

The experimental process followed the steps outlined in sections above on the Kasteren dataset: Construction of the "bag of sensor observation", activity-object patterns discovery using the pLSA, web knowledge extraction, computing activity intersection and ontology activity model. To test the learning process, the dataset was divided into training and test subsets. Furthermore, the performance based on accuracy and precision were determined using the true positives TP, true negatives TN, false positive FP and false negatives FN.

Results and Performance

The activity-object patterns computed from $P(x|z)$ specifies that Activity Topic 1 is recognized from the interaction of the objects (Microwave, Fridge, Freezer, Plates Cupboard, Pans Cupboard and Grocery Cupboard) and Activity Topic 3 is recognized from the interactions of Hall Toilet Door and Toilet Flush etc. as illustrated in Figure 4. Figure 5 illustrates the accuracy and precision performances of the proposed framework. The result suggests very good performance for "Defecation or Urination", "Go Out or Come In" "Sleep", and "Bath or Shower" all corresponding to Activity Topics 3, 4, 5 and 7 respectively due to no evidence of activity confusion and semantic ambiguity. Of particular interest are

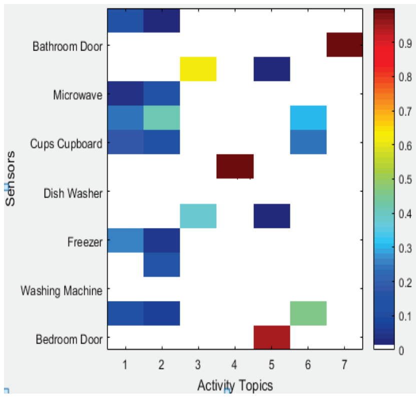

Figure 4. Activity-Object pattern and the Activity Topics

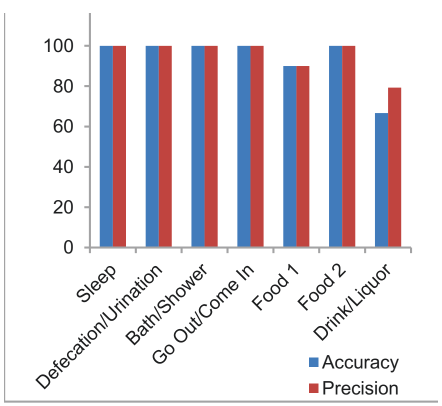

Figure 5. Accuracy and Precision performance of the framework.

Activity Topics 1, 2 and 6 corresponding to "Food 1", "Food 2" and "Drink or Liquor". These activities share same and similar objects interactions. While "Food 1" and "Food 2" are very similar, they are only different with "Pans Cupboard" used in all the instances of "Food 2". The use of "Pans Cupboard" is also evident as an object linked to "Food 2" as illustrated in Figure 8. "Food 1" and "Food 2" also have different temporal patterns. All instances of "Food 2" were recognized, but result suggests reduced accuracy of 89.7% for "Food 1" owing to two instances of confused with "Food 2" where "Pans Cupboard" was used. Accuracy and precision for "Drink and Liquor" were 66.7% and 79.3% respectively due confused recognition with "Food 1" and "Food 2". "Drink and Liquor" have activity-object pattern of interactions from "Fridge", "Cups Cupboard" and "Grocery Cupboard" which are also in the same activity-object patterns for "Food 1" and "Food 2". The overall accuracy and precision achieved was 93.8% and 95.6% which is significantly encouraging and comparable to results achieved using the same dataset.

Discussion

Activities recognized using our proposed framework includes "Defecation or Urination", "Go Out or Come In" "Sleep", and "Bath or Shower", "Drink or Liquor", "Food 1" and "Food 2" against "Toileting", "Go Out", "Sleep", "Shower", "Drink", "Breakfast" and "Dinner" specified in ground truth. With Thesaurus [19], the activities we recognized are synonymous with the activity labels of the ground truth. Because the activities recognized were more than specified in the ground truth, we are not able to carry out direct comparisons with the methods used on this dataset. With regards to the class of activities recognized "Food 1" and "Food 2" corresponds to "Breakfast" and "Dinner" respectively. "Food 2" was recognized in all the instances for which it occurred. "Food 1" was confused with "Food 2" in two instances resulting to a recognition accuracy of 90% and reduced precision of recognition. In these instances, "Pans Cupboard" was activated in addition to the objects in its pattern of occurrence. To distinguish the activities further, we have used their temporal attributes which also constitute a pattern of their occurrence. "Food 1" and "Food 2" are activities involving meal preparation at different times of the day. The recognition of "Drinks and Liquor" were confused with "Food 1" and "Food 2"in some instances because "Fridge", "Cups Cupboard" and "Grocery Cupboard" forms objects used for all three activities. Despite this, accuracy for "Drinks and Liquor" was 66.7%.

CONCLUSION

In this paper, we presented activity recognition of sensor data using the "*bag of words*" topic model, web knowledge extraction assembled on an ontology model. We used the topic model to discover activity-object patterns and activity topics. We retrieved from the web, activities of object usage and determine activities intersecting the pattern of these object usage. Activities like "Defecation or Urination", "Go Out or Come In", "Sleep", and "Bath or Shower", "Drinks or Liquor", "Food 1" and "Food 2" were recognised. Given the performance of proposed framework, we think it can be exploited to improve recognition performances in other areas and further recognise abnormal activities.

REFERENCES

1. A. Umut, T. Pavan and C. Rama. 2008. Ontology based approach for activity recognition from video. In *Proceedings of 16th ACM international conference on Multimedia* (MM '08).

2. D. T. G. Huynh. 2008. *Human Activity Recognition with Wearable Sensors*. Ph.D Thesis. Technische Universität Darmstadt.

3. M. A. Labrador and O. D. L. Yejas. 2013. Human Activity Recognition: Using Wearable Sensors and Smartphones, *Computer and Information Science Series*.

4. J. Ordonez, P. d. Toledo and A. Sanchis. 2013. Activity Recognition Using Hybrid Generative/Discriminative Models on Home Environments Using Binary Sensors, *MDPI Sensor*, *Open Access*.

5. T. V. Kasteren, G. Englebienne and B. J. A. Krose. 2010 Human Activity Recognition from Wireless Sensor Network Data: Benchmark and Software. *Activity Recognition in Pervasive Intelligent Environments*, *Atlantis Press Book*.

6. D. Wyatt, M. Philipose and T. Choudhury. 2005. Unsupervised activity recognition using automatically mined common sense. In *Proceedings of the 20th national conference on Artificial intelligence* (AAAI'05).

7. F. Katayoun and D. Gatica-Perez. 2011. Discovering Routines from Large-Scale Human Locations using Probabilistic Topic Models. *ACM Transactions on Intelligent Systems and Technology* (TIST).

8. L. Ratinov and D. Roth. 2009. Design Challenges and Misconceptions in Named Entity Recognition. In *Proceedings of the Thirteenth Conference on Computational Natural Language Learning* (CONLL09).

9. Y. Yan, N. M. Y. Okazaki, Z. Yang and M. Ishizuka. 2009. Unsupervised Relation Extraction by Mining Wikipedia Text Using Information from the Web. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP* (ACL 09).

10. M. Perkowitz, M. Philipose, K. P. Fishkin and D. J. Patterson. 2004 Mining Models of Human Activities from the Web. In *Proceedings of the 13th international conference on World Wide Web* (www-04).

11. P. Palmes, H.-K. Pung, T. Gu, W. Xue and S. Chen. 2010. Object relevance weight pattern mining for activity recognition and segmentation *Pervasive and Mobile Computing* vol. 6, pp. 43-57.

12. RDF. Retrieved December 1, 2015 from http://www.w3.org/wiki/RDF

13. N. Yamada, K. Sakamoto, G. Kunito, Y. Isoda, K. Yamazaki and S. Tanaka. 2007. Applying Ontology and Probabilistic Model to Human Activity Recognition from Surrounding Things. *IPSJ Digital Courier*, vol. 3, pp. 506-517.

14. L. Chen, C. D. Nugent and H. Wang. 2012. A Knowledge-Driven Approach to Activity Recognition in Smart Homes. *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 6.

15. G. O. Okeyo, L. Chen, H. Wang and R. Sterritt. 2012. Dynamic Sensor Data Segmentation for Real-time Activity Recognition. *Pervasive and Mobile Computing*.

16. T. Hoffman. 1999. Probabilistic Latent Semantic Indexing. In *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval* (SIGIR'99)

17. I. K. Ihianle, U. Naeem and A.-R. Tawil. 2015. A Dynamic Segmentation Based Activity Discovery through Topic Modelling. In *Proceedings of the IET International Conference on Technologies for Active and Assisted Living* (TechAAL 2015), London.

18. Protégé. Retrieved December 1, 2015 from http://protege.stanford.edu/

19. Thesaurus. Retrieved April 4, 2016 from http://www.thesaurus.com