



Identifying the Activities Supported by Locations with Community-Authored Content

-Written By-

David Dearman and Khai N. Truong

-Presented By-

Scott Mitchell

CS Department





Problem Domain

- **Determine types of activities which are possible at a given location**
 - **The set of activities is dynamic**



“Traditional” Context Aware

- **Low cost, integrated into environment**
 - **RFID, infra-red, accelerometer**
- **Designed to correlate specific sequence of actions to a specific event**
 - **Scalability**
 - **Recognition of dynamic nature tasks**



Alternative Context Aware

- **Traditional methods do not apply well when activities are “intertwined”**
- **Location activities can not be determined a priori**
- **Use content provided by the community**
 - **Scalability**
 - **Dynamic in nature**
 - **Determine potential user activities**



Natural Language Processing

- **From: Yelp**
 - popular community driven location review site
- **How: Verb-Noun Pairs**
 - Check zoo
 - Play chess

Architecture

- **Harvest**

- Name, URL, latitude, longitude, number of reviews

- **Parse**

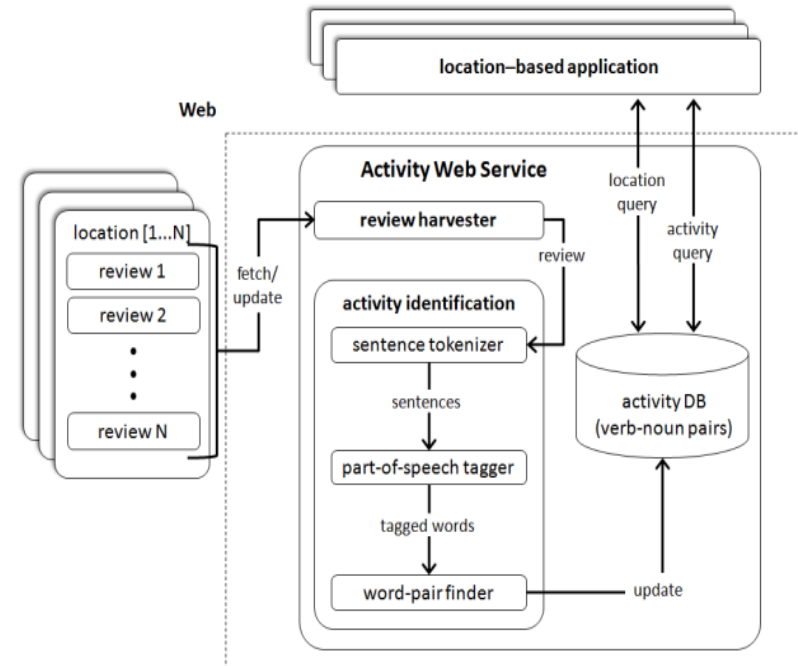
- Stanford Part-Of-Speech Tagger (English maximum entropy sentence tokenizer)

- **Tag and Extract**

- Activity finder pairs verbs with nouns if < 5 words away
- Perspective (1st I, we, 2nd you, 3rd he, she)
- Original and base words retained

- **Populate and Update**

- Quick access of word-pairs





Experimental Approach

- **14 diverse locations**
- **Participants**
 - provide activities performed/experienced at locations
 - validate 40 most common verb-noun pairs
 - True Positive – participant validated
 - False Positive – participant rejected
 - False Negative – not in most common

Questions / Comments

- **More details coming up...wake up**



Measurement Tools

- **Precision = False Positive / True Positive**
- **Recall = True Positive / False Negative**
- **Filter applied to noun-verb pairs to reduce number of false positives**
 - None, 1st Person, Frequency > 1
- **Known activity to identified verb-noun pairs**
 - Exact Terms
 - Similar Terms – statistically similar permutations of base words
 - Synonyms

Results

- Precision – Averaged across 14 locations
- Average Precision – Considers ranked order of noun-verb relevance
- 57 average known activities per location (participant provided + participant validated)
 - Limits recall to a max of 70.2%.
 - Observed 55.5% recall rate.

		Precision				Average Precision	
		no filter		1 st person		no filter	1 st person
		n	%	n	%	%	%
Validated		444	79.3	438	78.2	88.3	88.9
Provided	exact	32	5.7	29	5.2	23.2	24.7
	similar	66	11.8	62	11.1	26.7	26.7
	synonym	73	13.0	73	13.0	34.6	31.6

Results Continued

- **Participant verb-noun pair recognition relatively low**
 - 16.4% using synonymous terms
 - 83.6% false negatives
- **Number of reviews considered influences recognition**

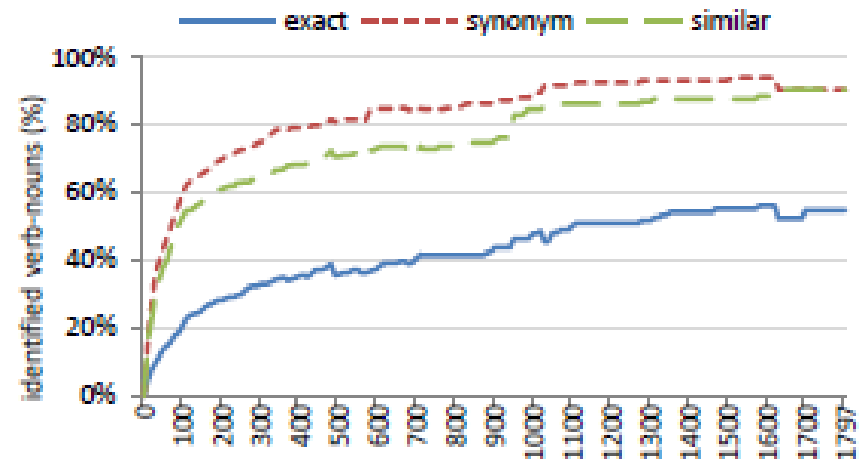


Figure 5. The percentage of participant-provided activities that are identified in the set of verb-noun pairs when varying the number of reviews processed, averaged across the 14 locations. X-axis is the number of review processed. The values are for no filter.



Clustering

- **Grounded Theory Affinity Clustering**
 - **Abstract activities into very high level**
 - **Physical (buy a book)**
 - **Cognative (enjoy art...)**
 - **Perceptual (watch people...)**

Real Life Applications



Figure 6. Activity Compass – a mobile application that characterizes the activities available in the user's vicinity.

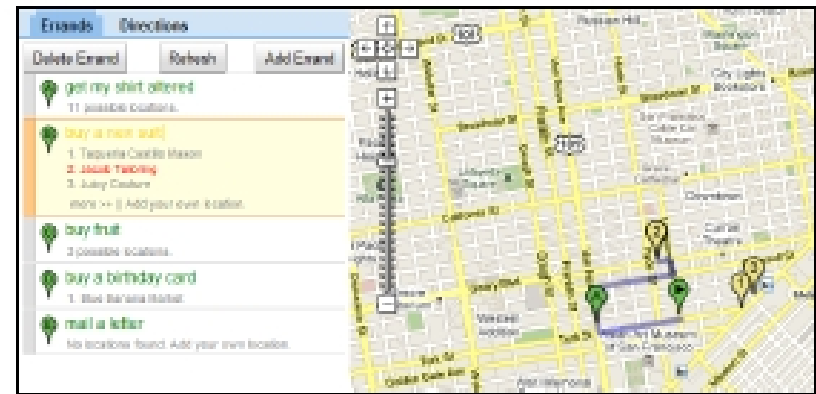


Figure 7. Better Errands – a web application that helps users discover nearby locations to perform their tasks and errands.



Questions / Comments

- **Natural Language Limitations?**
 - **Single sentence analysis**
- **Simplistic Frequency Analysis?**
 - **40 most common verb-noun pairs**