Crowdsourcing Based on GPS

Ketan Vazirabadkar

Department of Computer Science

Pune Institute of Computer Technology

Savitribai Phule Pune University, Pune, India

ketan.vazir@gmail.com

Siddhant Gadre

Department of Computer Science

Pune Institute of Computer Technology

Savitribai Phule Pune University, Pune, India

sidgadre@gmail.com

Rajeev Sebastian

Department of Computer Science

Pune Institute of Computer Technology

Savitribai Phule Pune University, Pune, India

Rajeev.hothead@gmail.com

Nikhil Dwivedi

Department of Computer Science

Pune Institute of Computer Technology

Savitribai Phule Pune University, Pune, India

dwivedinikhil12@gmail.com

Abstract

Machines cannot answer all the queries. There are some queries which need human input for providing information that is missing from the database, to perform computationally difficult procedures, and for cross checking, ordering, or aggregating results based on fuzzy criteria. Crowd sourcing can be used to solve such queries. Smartphones can unfold the full potential of crowdsourcing, allowing users to transparently contribute to complex and novel problem solving. We present the intrinsic characteristics of smartphones, a taxonomy that gives classification of the emerging field of mobile crowdsourcing and application that optimize location-based queries and similarity services over data generated by a crowd. This paper will focus on our application which uses Internet as a source of getting information which is so relevant to us that it is about what we want to do right where we are right now in the best possible way. We have implemented two ideas over the course of the project – first, a Location-Based crowd sourced forum and secondly, a platform where one can get location based relevant information. The forum will enable a user to ask relevant and location based question which will be answered by the people who are familiar with that area and the platform will monitor user's location at regular intervals and pop up relevant information as soon as s/he enters a particular radius of an area of his/her preferred choices.

Introduction

This paper is focused on a project for creating a forum for answering questions and then using this data to gather information for the GPS-based task assistant. Most current applications on to-do lists were based on time rather than location. Generally, the users set their tasks depending up-on time. Our application will provide a feature for users to set their to-do list depending upon their GPS location. If user has set a task that he has to do in a particular geographical area, whenever the user is in the proximity of the area, our application will provide notification that will remind the user of his/her task.

Through crowdsourcing we are able to generate information which is relevant to a particular user rather than depend on the vast but often incorrect information available on the internet. For this, we have used KEA(Key-word Extraction Algorithm). Users will be able to search questions/information category-wise as well as location-wise. Store-owners can provide offers to customers present in the vicinity within a radius. On the other hand, Users can gain information about the offers provided by various stores, particular stores in their vicinity.

Program Structure:

There are two stages in KEA's extraction algorithm:

1. Training: create a model for identifying key-phrases, using training documents where the author's key- phrases are known.

2. Extraction: Using the above model, choose key-phrases from a new document.

The KEA makes use of the following steps in order to remove the word from the documents:

Input cleaning: ASCII input files are filtered to regularize the text and determine initial phrase boundaries. The input stream is split into tokens (sequences of digits, letters and internal periods), and then several modifications are made:

• Phrase boundaries re-place punctuation marks, brackets, and numbers.

• Apostrophes are removed.

• Delete the remaining non-token characters, also delete any tokens that do not contain letters.

Key-phrase Identification:

KEA then considers all the subsequences in each line and determines which of these are suitable candidate phrases.

1. Candidate phrases are limited to a certain maximum length (usually three words).

2. Candidate phrases cannot be proper names (i.e. single words that only ever appear with an initial capital).

3. Candidate phrases cannot begin or end with a stop-word.

Feature Calculation

Two features are calculated for each candidate phrase and used in training and extraction. They are: TF × IDF, a measure of a phrase's frequency in a document compared to its rarity in general use; and the distance into the document of the phrase's first appearance, which is the first occurrence.

Android Application

The android application we created consists of a forum for question and answers. We were able to extract keywords from questions to categorize the questions and from answers to extract the places (proper nouns). We used Google-Maps to verify the existence of the places and to get their co-ordinates. Whenever there is similar question posted, our system will generate an answer (Suggested Places) for that using previous data without anyone answering the question. We also provided navigation feature to provide notifications whenever the user is in the vicinity of his suggested places.

Challenges and Solution

1. How to retain users?

Many encouragement and retention (E&R) schemes already exist. We have provided ways like up-votes and down-votes to establish measure, and show trust/fame/reputation. We can provide ownership situations, where a user may feel he or she "owns" a part of the system, and thus is compelled to "cultivate" that part.

2. What contributions can users make?

In many CS systems, the kinds of contributions users can make are limited. For example, to evaluate, users' upvote, review, rate, or tags. The challenge is to define a range of possible contributions (and design the system such that it can gather a critical crowd of such contributions).

3. How to combine User Contribution?

Many CS systems do not combine contributions.

If they do, it is in a loose fashion. For example, the present evaluation systems do not combine reviews and ratings, they combine numeric ratings using relatively simple formulas. Sometimes human users as well as machines make contributions. Combination of such contributions is difficult to process.

Conclusion and Future Work

In this paper, we have presented a GPS based task assistant. The application has been deployed on Android OS. We have developed a one of its kind mobile application which aims at providing users only the most relevant information. By using the concept of crowdsourcing, we have been able to generate information about a location based on user's preferences. We will be working on providing a hierarchical categorized search and implementing more efficient keyword search algorithm on a distributed environment. We intend to build similar apps for different platforms like Windows OS and iOS. Also, our future work will include the exciting world of analytics to discover more about the ever changing trends of user behavior.

Acknowledgement

References

Chatzimilioudis G.; Konstantinidis A; Laoudias C.; Zeinalipour-Yazti D., 2012 – "Crowdsourcing with Smartphones", Internet Computing, IEEE (Volume:16 , Issue: 5).

Doan A; Ramakrishnan R; Halevy Y., 2011 "Crowdsourcing

Systems on the World-Wide Web", Communications of the ACM.

Witten I.H; Paynter G.W; Frank E, Gutwin C; Neville Manning C.G, 1999 - "KEA: Practical Automatic Keyphrase Extraction", Proc. Australasian Computer Science Conference, Auckland, New Zealand.

Ma L.; Gu L.; Wang J, 2014 "Research and Development of Mobile Application for Android Platform", International Journal of Multimedia and Ubiquitous Engineering.

Turney, P, 1999 "Learning to extract keyphrases from text." Submitted to J Information Retrieval.