Ads-InSite: Location based Advertising Framework with Social Network Analyzer

A. A. G. A .K Perera¹, R. P. E. T Jayarathe², B.Y Thilanthe³, S. I. G Kalupahana⁴, P. S. Haddela⁵,
A. Kirupananda⁶, E. A. T. D. Edirisinghe⁷

Department of Information Technology, Faculty of Computing, Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
¹garushiperera@gmail.com, ²erandajayarathne@gmail.com, ³yasitha.bangamu Wage@gmail.com,
⁴i.shara26@live.com, ⁵prasesna@sliit.lk, ⁶abarnah.k@sliit.lk, ⁷mailthari@gmail.com

Abstract—The massive amount of valuable data circulated in social media is nowadays identified as a very effective source in marketing. The content of social media can be used to identify the business needs and preferences of people as the data is user related. This research is carried out with the aim of providing suitable advertisements for people based on their preferences by analysing their social media content. Users' demographic details are also considered in providing suitable advertisements of the shops that are most conveniently located for a particular user. The primary goal of this research is to build an advertisement framework that supports targeted advertising by analysing social media content. The information extracted by analysing the content of social networks is used to predict the advertisement categories that interest a particular user. The framework applies location based services to filter advertisements based on the location of the shop.

Keywords— Social Networks, Support Vector Machines, Target Advertising, Location based Services.

I. INTRODUCTION

Identifying the potential customers who are most likely to make a purchase is one of the major concerns in marketing a product or a service. This intention is not achieved most of the time as it is not easy to identify who are really interested in purchasing a particular product. Therefore, businesses spent a huge amount of money in vain for advertising their products and services to people who have no intention in buying them. On the other hand, people receive advertisements of products and services of which they have no interest in buying at all. With the rapid changes in the commercialized world, there are many places available for the customers to buy everything they need under one roof. But as such shopping malls consists of a huge amount of shops, customers often find it difficult to locate the shops that are really useful for them, so that they do not have to go from shop to shop in search of the product they want to buy. Another concern is the location where the product or service is available, as it is not convenient for customers to travel halfway around the country to buy a product. Nowadays location based advertising is used by many businesses, but most of them are limited to country filtering and even the rich location based systems work only when GPS (Global Positioning System) connection is available.

There are many concerns regarding the sources that can be used to identify the business needs of people more accurately. Social networks has gained a very powerful foothold on today’s society and the pool of data that is been shared through these networks contain valuable and accurate information about the users and most importantly about the necessities and preferences of people based on their opinions, experience and other demographic factors.

This research is been carried out with the goal of identifying the business value of such data and using them in a marketing aspect that will benefit both the supplier and the consumer. ‘Ads-InSite’ is been implemented with the objective of supplying a user with a well customized and filtered list of advertisements by analyzing the social network content of that particular user. The framework has been integrated with the API of Watapita Ads [1], which is a location-based advertising framework for mobile and web application developers in Sri Lanka. One of the APIs of the framework is been developed based on the weighing methods mentioned in Intellemo [2], [3]. The intended outcome is a rich advertising framework that supports both targeted and location based advertising. The framework is been applied in creating a mobile application that supports indoor navigation inside a shopping mall.

The organization of this paper is as follows; Section 2 explains the background and a review of related work. Section 3 describes the prototype system overview and the main modules of the framework. Section 4 contains the discussion. Section 5 concludes the research along with future work.

II. BACKGROUND AND RELATED WORK

A. Social Networks and Location based Advertising

As social media contains a vast amount of data about people, their opinions and preferences many researches have been conducted in identifying different aspects of the effective usage of social network data.

Some of these researches and findings are based on identifying the market value of data available in social networks. According to [4], many organizations and companies are looking for ways to use the content posted in social networks for business purposes. Main issue faced by such companies is identifying the relevancy of posted content [5]. It is important to get behavioral data on personal preferences and interests for more appropriate targeting of advertisements, as well as the use of demographic and historical customer data [5]. Reference [6] states that Facebook is currently working with a data mining company to track purchases of users. They have intentions to identify whether the users purchase a particular product or a service after searching for it online.
ForSight [7] is a tool that facilitates marketers and market researchers to listen to online conversations and understand the opinions of consumers, shoppers, and audiences expressed on social networks. ForSight uses an opinion analysis algorithm called BrightView [8].

In addition to these, there are many advertising services available and all of the major advertising services are considered in carrying out this research.

Target advertising is much popular among most of the advertising frameworks. Admob [9], Millennial Media [10] and iAds [11] find their targets by analyzing past buying patterns of the user. Facebook [12] find targets by analyzing the behavior of the user on Facebook profile.

Location based advertising has been achieved by Admob [9], iAds [11] and Millennial Media [13] by using GPS. But if the GPS connection is not available, these systems does not able to find the user’s current locations. Facebook locates users by the available location details of the user’s profile [12]. So that it is not perfect on street level location filtering.

This research is carried out as an extension to Watapita Ads [1] which is a location based advertising framework aimed at developers of Sri Lanka. The framework provides an API for developers to build their web and mobile applications on top of it, which enables them to provide advertisements for users who are using the developer’s application. The category of the advertisements received through a particular application is chosen by its developer. The application will prompt users with advertisements of the developer preferred category by tracking down their current location.

This research is carried out with an aim to build an advertising system combining both location based advertising and targeted advertising together, with extending the researches done in Watapita Ads [1]. And authors of this research expect that, this research will exceed the limitations of the similar systems that are currently available.

B. Indoor navigation

There are many indoor navigation technologies [14] that are used by the available shopping mall navigation applications.

WiseBy [15], Mally, GeniusMatcher and WiFiSLAM [16] use technologies such as indoor GPS, computer vision, 3D technology and Wi-Fi signals to navigate. Any of these applications do not integrate with social network accounts of users and provide a unique experience to the user or in other words targeted advertising. Decade old technologies such as Bluetooth Low Energy are also used for indoor navigation using mobiles [17]. Some of these applications do provide a social networking experience within the mall [18]. However none of these available applications has the ability to track shopping patterns through the user’s social network and provide advertisements that are relevant to the user.

Targeted advertising is not available within indoor location based mobile applications. The prototype mobile application will contain features where the user can navigate within the shopping mall while searching for advertisements. The results received will be filtered to retrieve the closest shops to the user and they will match the user’s shopping perspective and taste will be sent.

The authors of this research have applied target advertising based on social network content analysing to provide relevant advertisements to users when they are navigating inside a shopping mall as well as travelling through the country. Therefore, the user will get advertisements related to a category preferred by user. The advertisements are filtered based on location specific data. The possible current location of the user can be obtained through predicting, even though GPS is unavailable.

III. Ads-IN SITE : PROTOTYPE SYSTEM OVERVIEW

The research is focused at developing an ‘Advertising Framework’ that will consist of four vital components, namely, ‘Social Network Connector API’ that will provide the complete functionality required to access the major social network accounts of a particular user and retrieve information, ‘Social Network Content Analyser’ to analyse the data retrieved from one’s social network account and provide a list of advertisement categories which interests the him the most, ‘Advertising Web Portal’ that facilitates anyone to search and retrieve advertisements of products that will hold his interest, ‘Shopping Mall Mobile Application’ that assists the user in finding shops of interest inside a shopping mall, along with the advertisements and the path directions to the exact location.

Fig. 1 shows the high level architecture diagram of the Advertising Framework.

A. Social Network Connector API

One of the major outcomes of this research is an API with the combination of all major social network APIs to access any social network account that belongs to a user. Social Network Connector API gathers all the information that is required by the Content Analyser API to do the extraction. The API connects with the users’ profiles online or offline and gets information to the system. To access user information almost all the major social networks requires creating a public application to interact with social network users. These public applications are created inside each social network and also combined with the implemented API. Developers simply have to create their own applications in the relevant social network and get credentials and configure developed API. The implemented API has used all social network software development kits (sdk) [19] to communicate with public application. The API has to interact with the developer’s application and register users to developer specified social network. The implemented services of the API can be accessed by calling a restful link.
and it will return a JSON result. The API can have a smooth interaction with mobile and web applications due to the techniques used in development. When a user gives his permission, the API gets all latest user details of his social network accounts such as the latest user profile status, home page details, likes, share details, and these will be used for information extraction.

B. Social Network Content Analyser API

The Social Network Content Analyser API is responsible for extracting relevant information provided by Social Network Connector API and analysing the extracted information. The main objective is to provide the user with a customised list of advertisements based on their preferences, which is known as targeted advertising or personalized advertising. The API consists of two components, namely the Tokenizer module and the Text analysis and Prediction module.

1) Tokenizer Module:

Information extraction techniques are used to extract information from user posted status in social networks. For each status of the user, the API splits the status into very simple tokens such as numbers, punctuation and words of different types and these tokens are used for identifying patterns and making predictions about the advertisements that might interest the user. This module is responsible for providing a vector of extracted information to the Text Analysis and Prediction module. All user-related social network data is obtained based on the date and the following steps are used for creating the vector.

Step 1 is identifying the most likely location of the user [20] by applying Stanford Name Entity Recognizer (NER) on the collected information.

Step 2 is creating the Social Feed Term Map and the following list explains the procedural steps.

- Perform preprocessing on the social feed of the user and identify important text for vector on the social feeds gathered by the Social Network Connector API.
- Tokenize and remove stop words [21] by using a Java-based tokenizer and Part-Of-Speech Tagger (POS Tagger). Each social feed is separated into individual text using spaces or any other special character. Unimportant words that have no use in creating the vector are identified and removed.
- Instantiate the POS Tagger using Viterbi decoder and by using parts of speech tags [22] such as adjectives, adverbs, nouns, hash tags and proper nouns.
- Perform stemming for social feeds [23], the process of identifying words that have same stem but some variance in order to reduce the size of the vector and improve the performance.

At the end of each pre-processing, one vector related to user and particular date will be provided as the output of the module.

Step 3 is creating the Term-Weight Map Weighing Method, to determine the importance and relevance of each item in the vector. Each entry in the map contains terms related to a single social feed along with the weighing of that term. The weighing is performed by Term Frequency – Inverse Document Frequency [24] (TF-IDF).

II) Text Analysis and Prediction Module:

The main goal of this module is to predict a list of advertisement categories based on the information extracted through the Tokenizer module. This module contains the implementation of a classification model, and it is achieved via Support Vector machines (SVM) [25]. An algorithm called SMO (Sequential Minimal Optimization) [26] by J. C Platt is used to implement the SVM and to handle classification [27] of advertisement categories. Weka framework [28], which is a collection of machine learning algorithms designed for data mining tasks by the University of Waikato is been used for the implementation of the classification model for predicting the advertisement categories that the user will find interesting.

The prediction model is built with the aim of accurately classifying the advertisement categories out of the data extracted through social networks by modifying one of the algorithms provided in the Weka jar for training the Support Vector Machines. SVMs find an optimal solution out of all the possible solutions to identify the hyper plane that separates the different advertisement category classes available in the data and maximize the margin around the separating hyper plane [29]. The problem of finding the optimal hyper plane requires the solution of a very large quadratic programming (QP) optimization problem, which consumes a huge amount of time and resources. SMO breaks this large QP problem into a series of smallest possible QP problems to solve them analytically [26]. Weka provides the implementation of SMO algorithm in the jar they provide and it is been used to achieve the purpose of this research. Fig. 2 describes the composition of Social Network Connector API and Social Network Analyzer API.

Developers can use this Social Network Content Analyzer API to get user related advertisement category list as a JSON result (When the developer app integrated with other APIs). Both APIs are restful web services so any developer who likes to get a list user related advertisement categories for their application, can use the implemented APIs for both mobile and web applications. Furthermore, these APIs will be more useful to advertisement search queries as well.
C. Shopping Mall Mobile Application (SMAPP)

When a customer enters a shopping mall, the advertisements provided through the data extraction methodologies will be filtered by the shops registered within the shopping mall and provided to the customer. The technology of Wi-Fi has been used for the indoor location based purposes [30], [31]. For the prediction of the distance between the user and a router, the signal strength loss and frequency of the Wi-Fi signal is used. The signal received may be reduced by many environmental effects and the walls of the shopping mall mainly.

The target was to run an algorithm to add any reductions of ‘decibels’ using the idea that shopping malls have standard shop sizes. Therefore the signal strength will be equal to the value that is encountered in a free space. For the calculation of the distance of the user from each router the ‘Free Space Path Loss’ equation is used [32].

\[
FSPL (dB) = 20 \log_{10}(d) + 20 \log_{10}(f) - 27.55
\]

d - Distance in meters.

f - Frequency of signal in megahertz.

The frequency and path loss of each router can be used to determine the distance from the user to each router. When determining the location of the user within the shopping mall a minimum of 3 routers is required. And for determining the user position Wi-Fi triangulation algorithm is used. Initially we draw a triangle with the 3 routers. When the user position is connected to a router another triangle can be created as shown in Fig. 3. We create all as a right angled triangles. The next triangle is created as a connection between Router 2 and Router 3. It is also right angled.

\[
a_1 = \Theta \text{ is the angle between distance from router to user position and distance between 2 routers}
\]

![Fig. 3 Router Diagram](image)

This uses the Pythagorean Theorem. Another equation is that the distance between 2 routers (full distance -fd) get split between the 2 right angled triangles in Fig. 3.

Distance between Router 2 and middle line = z * fd

Distance between Router 3 and middle line = (1 – z) * fd

Table 1

<table>
<thead>
<tr>
<th>Triangle</th>
<th>Equation Generated from Fig 3 Triangles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Router 2 triangle</td>
<td>[ c^2 = a^2 + b^2 ]</td>
</tr>
<tr>
<td></td>
<td>[ a^2 = c^2 - (z \times fd)^2 ]</td>
</tr>
<tr>
<td>Router 3 triangle</td>
<td>[ c^3 = a^3 + b^3 ]</td>
</tr>
<tr>
<td></td>
<td>[ a^3 = c^3 - ((1 - z) \times fd)^3 ]</td>
</tr>
</tbody>
</table>

This forms one Equation which can be used to find the value of “z”

\[
c^2 - (z \times fd)^2 = c^3 - ((1 - z) \times fd)^3
\]

Now the angle in router 2 triangle in Fig. 3 can be calculated as;

\[
\cos a_1 = \frac{b}{c^2}
\]

\[
a_1 = \cos^{-1} \left( \frac{z \times fd}{c^2} \right) \text{(Replace } b = z \times fd)\]

From the triangles shown below, Triangle 1 in Fig. 4 can be used to find the height and width of the triangle.

\[
a_2 = \Theta \text{ is the angle between distance from Router 2 to Router 3 and the horizontal line in Triangle 2 in Fig. 4}
\]

This can be found by using Triangle 2 in Fig. 4, as it creates a right angle there also. As we know x and y points of router 2 and router 3 we can find the horizontal distance and vertical distance of Triangle 2 in Fig. 4 and find the angle a2.

![Fig. 4 Position of User](image)

For this process to take place the users have to be within range of the shopping mall routers and also all routers have to be up and running. For navigation within the shopping mall the currently used concept is to allow the mall owner to save the path of each floor. SMAPP will then access the path to find shortest path to the shop he requires using Dijkstra’s algorithm. Initially the shopping mall owner will be given an image format that the image should look like when uploading at registration. Later the floor map can be edited to add the navigation path as shown below as the rectangle around the 4 slots and the pointers where shops within the shopping mall can register.

D. Advertising Portal and GPS Tracker (Location Tracker)

Out of the various location tracking systems, GPS is the most successful and accurate location tracking system when considering outdoor navigation [33], [34]. Therefore in this research, GPS is used as the technique to track the locations of the user. Fig. 5 contains an overview of the Location Tracker component.

The developed solution is a mobile application which runs on a smart phone, so that it can be easily used to identify the current location of the user. Using a simple algorithm built by Haversine formula [35], closely located shops can be identified when the location of the user is identified.

Usage of smart devices has risen dramatically during the past few years, so it is possible to implement the above algorithm, but still, there is no guarantee that the user’s device will be able to keep the contact with the system continuously. Therefore, a solution is required to keep the accuracy of the system at the expected level. This research proposes a semi-offline location prediction method to address this problem.

This location prediction method will run in three steps:

- Collect and store location data when the mobile application is connected to the system.
- Find the predicted locations using collected data and store data.
- If the device is offline, use predicted location details in selecting related advertisements.
At step one, the system will receive location details from user’s mobile device whenever the user is online. These details are stored in the system so that it can be accessed at step two.

At step two, predicted locations will be identified using the collected details at the past. Implementation of this module is done based on decision tree [36], a well-known algorithm used in data mining. Weka framework [28] is used to implement the decision tree algorithm, which is a Java based framework which consists of many data mining algorithms. Predicting will differ from person to person, therefore many facts should be considered while implementing the module. Details such as user’s occupation, monthly income, country, gender and day of the week are also used to maximize the accuracy of the module.

In step three, the system will try to get the location from user’s mobile device, if the device is not connected with the system, predicted location details will be used in finding the location related advertisements for the user.

By using this method, the system will be able to predict user’s locations more accurately. Therefore the system can identify the location of the user, when GPS connection is not available and the raised research gap will be filled with the usage of this method.

IV. DISCUSSION

The developed Advertising Framework is currently in the stage of evaluation. The framework as a whole is aimed at providing targeted and location based advertisements with the analysis of social network content. The next phase of this research is to evaluate the accuracy of the framework. Once evaluated, the authors of this paper are expecting to pursue ways to improve the accuracy of the framework.

V. CONCLUSION AND FUTURE WORK

This research is carried out with the aim of building a system that can reach the consumers with suitable advertisements, on their preferences based on social media contents. Since the mobile and the web application usage are rapidly been increasing these days, these technologies are used to fulfill the gap between the advertisers and consumers. In order to maximize the efficient usage of the system, new algorithms are implemented to reach consumers such as social network analyzer and the daily movement identifier.

The implementation and the comparison are performed and presented where the system accuracy and the user satisfaction is not tested. These characteristics will be tested and analyzed and the research will be continued to increase the efficient usage of the system.

The SMAPP module is to be developed further with the aim that shopping mall navigation can be automatically predicted when the customer is on the move using image edge segmentation. Currently there are packages such as OpenCV that provide such functions inbuilt. The aim is to scan the image and identify the edges in the image. Next from the user’s current location predict the path such that no edges are intersected and it becomes the shortest path.

REFERENCES


