

## Location based Advertising for Mass Marketing

Vignesh Kandasamy S<sup>1</sup>, Anjali Madhu<sup>1</sup>, Prasun Kumar Gupta<sup>1</sup>, Niveditha A<sup>1</sup>, Kusumbor Bordoloi<sup>1</sup>

<sup>1</sup> Geoinformatics Department, Indian Institute of Remote Sensing, Indian Space Research Organisation, Department of Space, Dehradun, Uttarakhand, India

{kandasamy, madhu}@student.utwente.nl, prasun@iirs.gov.in, {m.s.ajithkumar, bordoloi}@student.utwente.nl

### Commission V, WG V/8 - Promotion of Open Source on Geospatial Technology

**KEY WORDS:** Location based Advertising, Open Source GIS, Tesseract OCR

#### ABSTRACT:

GIS and machine learning (ML) are powerful ICT tools in retail industry which helps the sellers understand their markets. For the consumers, however, there always lies an ambiguity with respect to the quality and quantity of the product to be purchased, vis-à-vis the price paid for it. Most retail businesses today adopt "Discount Pricing Strategies" or "Offers" to make new customers and increase sales. Owing to several establishments selling the same product and offering a variety of offers, the process of identifying the shops where the consumer can get the best value for his money, requires a lot of manual effort. A prototype has been developed in this study to allow the consumers to locate such prospective shops based on advertisements in newspapers. This solution has a two-pronged approach. First, all the offers advertised in the newspaper are pre-processed and text extraction is performed using a ML algorithm named Tesseract OCR. Second the location of shops is collected and stored in a geodatabase. Finally, the advertisement is matched to the respective geo-located shop based on its name and location. Further based on the location of the consumer and his purchase choice, shops offering discounts are shown on a web based map. This prototype provides the consumer, a platform for geo-discovery of establishments of interest through the clutter of unrelated endorsements, by the use of Open Source GIS, Python programming and ML techniques.

### 1. INTRODUCTION

The integration of geographical information system (GIS) with machine learning (ML) has paved the way to explore many research opportunities. New avenues in computing such as ML and artificial intelligence are broadening the capabilities of GIS. GIS, combined with the power of global navigation satellite system (GNSS), can now be effectively used as a decision support system that can analyse and visualize geospatial data. The shift of geospatial applications from scientific communities to the business enterprises has become possible only due to availability of large consumer and product consumption datasets.

With the digitization of cable television in India and the entry of multi-national research and consultancy groups, large amounts of spatial data on consumer patterns such as product preference, consumption patterns with respect to gender, age and location is now readily available. Hence the industry is exploring spatial data to do targeted marketing and understand customer behaviour, for maximizing their profits ('How Geographic Information Science Can Help Businesses | USC', n.d.). Non-spatial databases of product sales are incapable of highlighting spatial patterns of potential customers. GIS can answer questions on where, why and how based on such location information (Km & Patil, 2011). The understanding of geo-demographic characteristic of the customers became a basic necessity for the competitive business environment. Mass marketing and direct marketing are the two kinds of approaches that industry follows to advertise their product and services. Television, newspaper and broadcasting agency falls under mass marketing category. This approach is very much

beneficial when the products have good demand in public (Ling & Li, 1998). On the other hand direct marketing deals with direct promotion with the customers. This is beneficial for retailers who are good at promoting their products. In order to expose the product to the people, mass marketing is the common strategy that is practiced widely.

Though the number of people buying the product after seeing the advertisement is proportionally low but the implementation of mass marketing through the channel of geo-marketing is overcoming those odds. Geo-marketing is preferred by the industry to successfully implement the marketing strategy and to find the profitable selling locations of the product, in order to get the maximum circulation of the product. Even the best advertisement strategy is found by analyzing the consumer behavior with the point of sale of the product and comparing it with the nearby advertisement published. Geo-marketing has been rightly termed as the origin of market intelligence. Yrigoyen defined geo marketing "as a set of techniques that analyses the economic and social realities of a geographical point, by ... space statistical tools" (Nunes, Santana, Bezerra, & Sobral, 2014). Like GIS, geo-marketing answer the questions such as who buys and where he buys from.

Advertising is normally viewed as an art (Adi Ignatius, 2013), but with advent of digital marketing, business intelligence and space based inputs, advertising has now become a scientific exercise. The location based advertisement is a kind of direct marketing which uses the mobile technology and location based services to approach the specific target audience. When a mobile user searches for a shop near to him, by factorizing his needs or previous purchases, current offers, special events etc., advertisements are now-a-days tailor made for the consumers.

Though the technology is making the advertiser to approach the consumer directly, the final decision to choose the product lies with the customer. To make them to choose the product lies in the basics of marketing principles "Discount pricing" or "Offers". Human psychology, which gets attracted to offers, holds a key factor to circulation of the products and to the growth of the business.

In absence of location linked distribution of shops or establishments proposing such discounts or offers, the process for finding the best value for the money requires lot of manual effort. It generally leads to the consumer paying a higher price for the product. In this study, the solution to this problem is given by developing a prototype as a map of discounts/offers. The map displays the shop and its ongoing discount pricing or offers. Therefore the objective of the study is to provide a platform for the consumers, who can search for the shops which have offers and choose them based on their preference.

## 2. METHODS

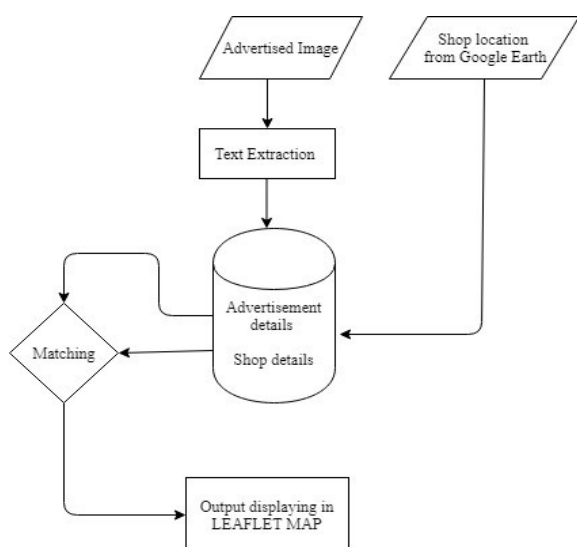


Figure 1. Methodology for Visualizing the offers advertisements in the map

Figure 1 shows the proposed methodology for this study. It has 3 main segments (1) data collection, (2) text extraction and (3) adding offers to its location and visualizing it in a map.

### 2.1 Data Collection:

The data collections are done by two methods; first the collection for advertisement is done based on the newspaper. The latest Offers from the newspaper are taken. Second the locations for the advertised shops are taken from Google Earth. The advertisement for current date is collected. First these advertisements are categorized depending on the product type such as jewellery, foot wear, cosmetics etc. Second, the shops location for the respective advertisement are taken from Google Earth as KML. These KML files are converted to the shape file and ingested in a PostgreSQL / PostGIS spatial table named as Shop\_Details.

### 2.2 Text Extraction

The information such as name and address of the shop needs to be extracted from the advertisement, so that it can be added with its location information. The machine learning package, named Tesseract OCR which is sponsored by Google (Smith, 2007), is used in extracting the text from the image. The advertised image which contains offers or Discount pricing may have multiple colours and different stylized fonts. In order to get higher accuracy in text extraction, the advertisement images are enhancement using image enhancement techniques (such as RGB to grayscale conversion, positive to negative image conversion, image sharpening etc.) The enhanced image is fed in to the ML algorithm and the text is retrieved. These text are stored with the respective image in PostgreSQL / PostGIS spatial table named as Advertisement\_Details.

### 2.3 Adding offer to its location and visualizing it in a map

The shop name from the Advertisement\_Details table is searched in the extracted text. If the shop name is matched, then the same is used to match with its location in the Shop\_Details table. If both the conditions are satisfied, the image is linked with the shop's location. When the consumer searches for the offers the location of the shop with the advertised image is displayed using Leaflet.

## 3. RESULTS

The methodology which is described in the section 2 is used to accomplish the desire result. The result is explained in three subsections 3.1 Text extraction, 3.2 adding the image with its location and 3.3 Leaflet.

### 3.1 Text extraction

The advertisement which is collected is enhanced using image enhancement. The initial images Figure 2 (a & b) shows the original input image, which is converted into grayscale and stretched. In order to get good accuracy of the converted image is again converted as negative as sharpened. Figure 2 (c & d) shows these enhanced images. The enhanced image is fed into tesseract OCR and the extracted text shown in Table 1 is stored in the database.





Fig 3 (a) Visualizing the jewellery shop which have offers in New Delhi, (b) visualizing the sports shop which containing offers.

Figure 2. (a) Original Image; (b) After applying Grayscale to original image; (c) After applying invert to grayscale image (d) After applying sharpness to negative image

Figure 2 shows the steps of image enhancement techniques which is used to improve the accuracy of the text extraction.

... KALYAN ... GOLD JEWELLERY ... DIAMOND & POLKI JEWELLERY ... PRECIOUS JEWELLERY ... 5% FLAT OFF ... PITAMPURA - Netaji Subash Place. / DEFENCE COLONY - Block B. No. 42 / KAROL BAGH - Near Metro Station, Pusa Road GURGAON - Empire Square, JMD Mall, MG Road. / NOIDA - Near Metro Station, Sector 18. ...

Table 1: Text extracted using Tesseract OCR from Figure 2 (d)

### 3.2 Adding the image with its location

The program searches for shop and its location from the extracted text which need to be matched with the shop details using SQL queries. Example:

<<<SQL query >>>

```
SELECT ST_AsGeoJSON(loc_point), sd1.shop_name FROM
shop_details sd1,offer_details od WHERE od.offer_text \
like ('%||sd1.shop_name||%') AND od.offer_text like
('%||sd1.location||%') AND sd1.shop_category = '{0}'"
```

### 3.3 Leaflet

The most popular Open Source JavaScript library to make interactive maps is Leaflet. Leaflet have all the necessary features that a developer need and its weight about 33 KB of JavaScript (Wang, Pi, Zhou, & Zhou, 2015). The offer locations were downloaded from PostgreSQL / PostGIS spatial table as geoJSON objects and visualized in leaflet with open street map as base layer. The corresponding advertisement images were added as pop up in the map so that user could view the advertised image when they hover over the Offer location. Figure 3 visualizes the shops in New Delhi which have Offers.

## 4. CONCLUSION

The prototype described in this paper is an attempt to visualize the discount pricing or Offer in a geo-demographic way. This minimizes the efforts of people who are searching for the Offers and makes a meaningful value for their money. The combination of ML with GIS laid the foundation for the accomplishment of this task. In this paper only pictorial advertisements in newspapers were taken into consideration, but various other ways of publishing the advertisement can be utilized for further expanding this work. With digital advertisement and online targeting of ads, the task of gathering advertisements also becomes a challenging area.

## ACKNOWLEDGEMENTS

Thanks for the Open Source community for developing wonderful tools and we sincerely thanking our professors for giving this opportunity.

## REFERENCES

- Adi Ignatius. (2013, March). Advertising Is an Art—and a Science. Retrieved 25 June 2018, from <https://hbr.org/2013/03/advertising-is-an-art-and-a-science>
- How Geographic Information Science Can Help Businesses | USC. (n.d.). Retrieved 19 June 2018, from <https://gis.usc.edu/blog/how-geographic-information-science-can-help-business/>
- Km, P. K., & Patil, P. G. (2011). Geographic Information System (GIS) – for Business Analytics. *International Journal of Scientific & Engineering Research*, 2(11), 1–6.
- Ling, C., & Li, C. (1998). Data Mining for Direct Marketing: Problems and Solutions. *Proceedings of KDD 98*, 98, 73–79. <https://doi.org/10.1.1.332.1803>
- Nunes, A., Santana, C., Bezerra, F., & Sobral, N. (2014). Knowledge Acquisition Based on Geomarketing Information for Decision Making: A Case Study on a Food Company. *International Journal of Innovation, Management and Technology*, 5(6), 422–427. <https://doi.org/10.7763/IJIMT.2014.V5.552>
- Smith, R. (2007). An overview of the tesseract OCR engine. *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, 2, 629–

633. <https://doi.org/10.1109/ICDAR.2007.4376991>

Wang, L., Pi, R., Zhou, X., & Zhou, H. (2015). The Construction of Off-line Map Based on OpenStreetMap and Leaflet, (Iccmcee), 1478–1482.