Optimisation of targeted digital advertising in shopping complexes using trends identified from human activity heat maps

[1]Nikhil Mollay
[1]Undergraduate Student, Undergraduate Student
[1]School of Computer Science and Engineering, VIT University, Vellore, India

***Abstract*—Developing a system that aims to track human presence and movement through commercial shopping complexes, such as malls, to create heat maps of activity and use this data to optimise advertising. This process is conducted over a period to identify activity patterns with respect to multiple factors. Temporal factors like time of day, day of week or month, Cultural Factors like During festivals or other notable occasions etc. By analysing the variation of the heat map with respect to changes in these factors, patterns are identified by the system. By using the identified patterns targeted advertising using digital advertisement signs and billboards can be optimised.** *(Abstract)*

***Index Terms*—Targeted Digital Advertising, digital advertising, heatmaps, human activity frequency, temporal catalogue.** *(key words)*

I. Introduction (Heading 1)

The system will comprise of a hardware component and a software component. The hardware component will comprise of multiple specially designed sensor devices that will be installed in the targeted building. These devices will be able to sense human presence and movement in an area around them and in turn relay this data to the software backend. To deploy the system one must perform an exhaustive survey of the building to create a detailed digital map. Based on this map the dedicated hardware sensors are deployed across key locations throughout the building.

The hardware detectors operate autonomously and keep pushing the movement and tracking data to the software server component. The server tracks all the received requests, logs them and creates heatmaps after fixed time intervals which are then overlaid onto the digital map created during the survey phase.

The ultimate goal of the system is optimisation of Targeted Digital advertising. Targeted digital advertising is by far the most effective form of advertising. Most advertising done these days is online, and almost all of it targeted. Targeted advertising is a method by which advertisers are able to provide more specific and relevant advertisements to users based on data they have collected on the users from various sources- Services like Google Ad Words and Facebook advertising provide this type of user specific data based advertising options. The reason this type of advertising is extremely effective is because it utilises data that is most pertinent to the user- taker. For instance, the case of a user liking Facebook posts about one of their friends going on vacation, is effectively used by Facebook. It takes this data and based on this, informs this user advertisements about vacation packages. Unlike the user specific targeted advertising presented in the example, the system proposed by this paper is based on analysing the movement of large groups of people – catalogued based on temporal factors like time of day, day of week/ month.

Finally, the heatmaps are analysed, and the patterns are identified. This step will commence once a large dataset of heatmaps is gathered, so that there is a reliable basis for the identified patterns. Once the patterns are identified, the data is passed onto the digital advertising partners who then utilise this information to serve the targeted advertising. An example of this process is described as follows- If patterns identified from the data indicate that on Wednesdays from 12.30 pm to 3.30 there is a relatively large presence of people at the food court then the advertising during and leading up to that time period will focus on food and food related products.

II. Work Related Review

History of the Cluster Heat Map [1]:

In a data matrix, the row and column hierarchical cluster structures can simultaneously be displayed using cluster heat map. This is done through rectangular tiling, in which each tile is shaded on a colour scale to represent the value of the corresponding element of the data matrix. The rows (columns) of the tiling are arranged in such a way that similar rows (columns) are placed next to each other. At the vertical and horizontal margins of the tiling are placed the hierarchical cluster trees. The cluster heat map then becomes an amalgamation of several different graphic displays that have been developed by statisticians over more than a century. The earliest source of this display can be traced back to the late 19th century publications, and a subsequent thread can also be seen in the diverse 20th century statistical literature that provided a base for this very commonly used bioinformatics display.

[Permut Matrix] Arranging gene expression profiles, in the best possible manner, in a linear order, in a graphical environment. [2]:

PermutMatrix is a work space that has been designed to graphically explore gene expression data. It utilizes the graphical approach created by Eisen and further it also provides various methods for the optimal reorganization of rows and columns of

a numerical dataset. To give an illustration, several methods are proposed for optimally reorganizing the leaves of a hierarchical clustering tree, simultaneously allowing many serrations or one dimensional scaling method that does not require any initial hierarchical clustering. The program, that has been developed for MS Windows, using MS-Visual C++, has a clear and efficient graphical interface. Large datasets can be systematically and speedily analysed.

IBM Research Centre – Visualisation of data [3]:

One of the often discussed topics at the annual IEEE Computer Society Visualization conferences, has been the importance of visual representation. This idea was first published by Huff [1954] in his book How to Lie with Statistics. Both in the book and at the How to Lie with Visualization Sessions at those conferences, the major focus is on how the interpretation of data can be undermined by manipulating the data representation. In this article, the moot point discussed is, how can the interpretation of data be enhanced? To answer this question, what is considered is firstly the structure of the data, secondly, the visual dimensions used in visualization, and thirdly the task the analyst is trying to solve – the solution lies in using applications of heatmaps.

Evaluation of Artery Visualizations for Heart Disease Diagnosis [4]:

The paper uses a task taxonomy for hemodynamics that is based on a formative user study involving domain experts. On the basis of the results achieved after the study, the authors developed HemoVis, an interactive visualization application for heart disease diagnosis that uses a novel 2D tree diagram representation of coronary artery trees. The paper gives the results of a quantitative user study with domain experts that compares the effect of 2D versus 3D artery representations and of colour maps on identifying regions of low ESS. The majority of the results showed that the 2D visualizations are more accurate and efficient than 3D representations and it also showed that a perceptually appropriate colour map leads to fewer diagnostic mistakes than a rainbow colour map.

Large Interactive Visualization of Density Functions on Big Data Infrastructure [5]:

In lots of visualization techniques, point set visualization is required. Scatter plots and geographic heat-maps are clear examples. Data analysts are increasingly using such visualization techniques. The need to work with larger and larger volumes of datasets requires to make these techniques work as fast as the data grows. The Big Data Infrastructure offers the possibility to scale horizontally. A critical requirement is to be able to design point set visualization methods that can handle the new changes. In this paper, what is presented is a complete architecture which fully fits into the Big Data paradigm and therefore allows interactive visualization of heatmaps at ultra-scale.

III. WORKING

The system is composed of two components – the hardware detectors and the software backend. The hardware components have been prototyped using an Arduino development board with PIR sensors. These are fixed at certain key positions which are mapped and isolated during the planning and surveying phase.

The planning and surveying phase of the deployment is extremely important. Since this determines the scale of the project deployment and the number of hardware detectors required. The survey has 2 essential responsibilities – firstly isolating the key locations of interest that would be important for the analysis and identification of meaningful patterns. Secondly the mapping and scaling of 2D map onto which the heat map will be overlaid. Flow of the system is described on the following page.

Hardware Details

The hardware detector comprises of an Arduino development board, an ESP Wi-Fi module and a PIR sensor. The code on the Arduino is responsible for using the PIR to detect human presence and log it. Once the log is large enough it is packed into a data packet – multiple different formats may be used to do this however this project achieves this simply passing the information in the request URL- that is then transferred to the server over the network using the TCP/IP protocol.

Software Details

This program is written in python, as it provides an easy way to implement the requirements of the project i.e. multithreading, image processing, stable networking and graphics capabilities. The software leverages the multithreading library of python to enable the server application to be more streamlined and efficient; this also ensures that none of the data packets being sent to the server are dropped.

IV. ALGORITHM

Figure 1 shows the algorithm for the working model related to this paper.

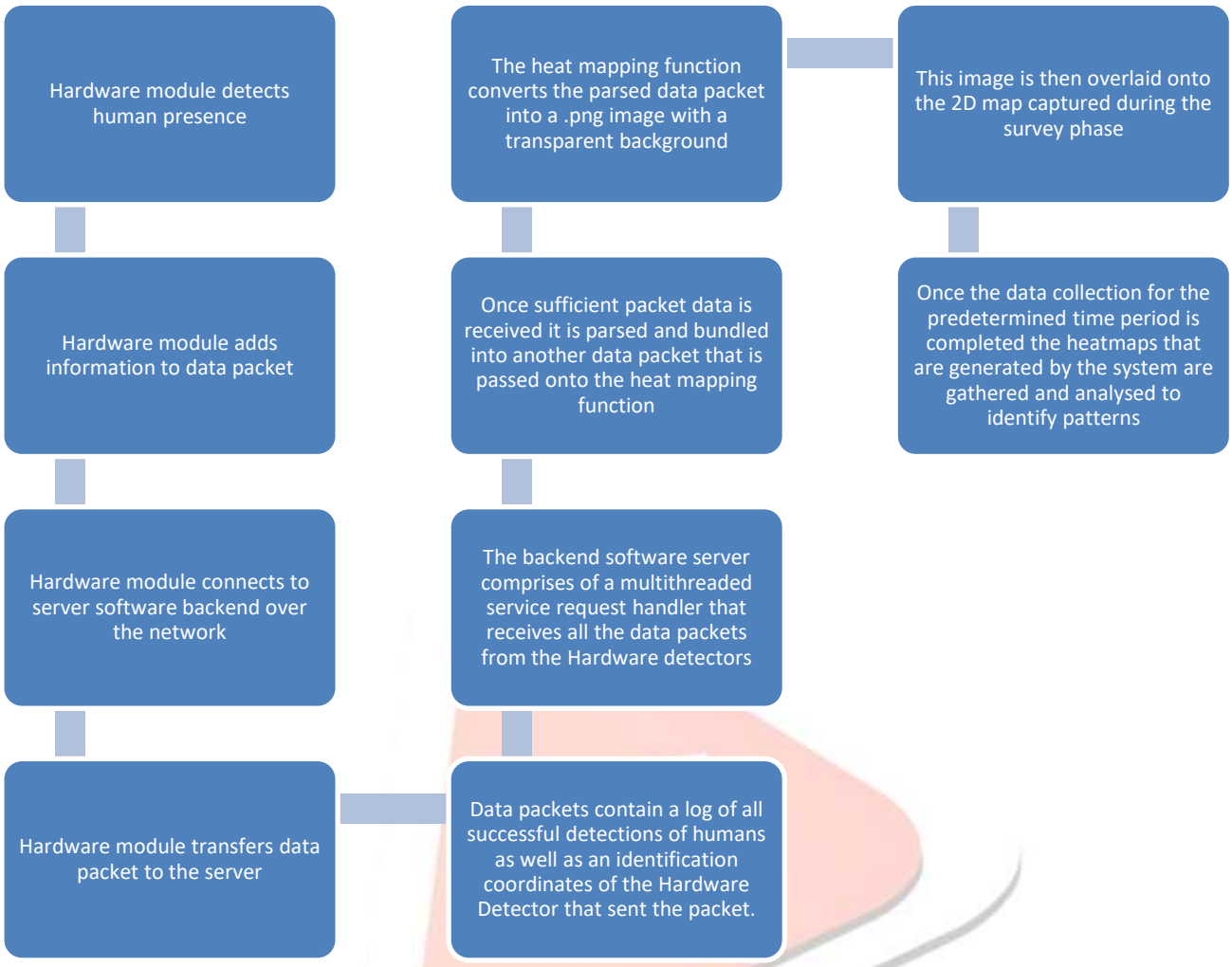

Fig. 1

V. SCREENSHOTS

The underlying figures, **Fig. 2, Fig. 3** and **Fig. 4** show the obtained results of the heatmaps produced for varying data.

Fig. 2

Fig. 3

Fig. 4

VI. RESULT AND AVENUES TO EXPAND RESEARCH

The system was successfully prototyped and tested with the results presented above; the server program was fed with coordinate and node data which it then successfully transformed into heat maps and then overlaid it onto a sample layout image. However, the coordinate positions do not correspond to certain areas of interest as that would only be required in a real world implementation of the system.

Due to the modular design of the system its individual software and hardware components can be easily upgraded to perform more complex and advanced tasks. Furthermore this system can be easily integrated with or developed into a security system for the location it is deployed at – due to its PIR sensor it can function as a motion detector at night when the system is offline and protect against intrusion.

By modifying the hardware module further with the addition of more sensors and servos to allow for movement of these sensors , we can upgrade the detection capabilities of the hardware detectors to create an ad hoc radar type system. Such a system would be able to measure the flow of people more accurately.

The analysis step in the system can be enhanced by using unique and efficient algorithms to identify more interesting patterns.

VII. ACKNOWLEDGMENT

Firstly, we would like to express our indebtness and appreciation to our guide, Prof. Rajkumar R, Assistant Professor, SCOPE, VIT University, Vellore. His constant guidance and advice played a vital role in making of this project and paper. We are indebted to our Head of Department, Dean (SCOPE) and VIT University to allow in their esteemed organization, the access to the required data and tools. We are grateful for their guidance, encouragement and suggestions which provided us with necessary insights to complete our research and paper successfully.

REFERENCES

[1] History of the Cluster Heat Map -http://www.datavis.ca/papers/HeatmapHistory-tas.2009.pdf
[2] [PermutMatrix]Graphical environment to arrange gene expression profiles in optimal linear order - http://bioinformatics.oxfordjournals.org/content/21/7/1280
[3] IBM research centre – Visualisation of data-http://www.cs.ubc.ca/~tmm/courses/533/readings/pravda/truevis.htm
[4] Evaluation of Artery Visualizations for Heart Disease Diagnosis http://ieeexplore.ieee.org/document/6065015/
[5] Large Interactive Visualization of Density Functions on Big Data Infrastructure http://ieeexplore.ieee.org/document/7348077/?reload=true