

Browsing Pattern Analysis: What user browsing Patterns Indicate

Sai Sagar Peri
SoCSE
University of Petroleum and
Energy Studies, Dehradun

Nikhil Srivastava
SoCSE
University of Petroleum and
Energy Studies, Dehradun

Nitin Arora
Assistant Professor(SS)
University of Petroleum and
Energy Studies, Dehradun

ABSTRACT

With the advent of the Internet of Things (IoT) the number of users and the number of websites in the World Wide Web increase exponentially by the day. Consequently, the users have issues searching required content in the plethora of available websites. Optimisation, speed, and convenience of browsing through data have never been more important. This is where the role of user preference mapping through the process of browsing pattern analysis comes in. The process of user browsing pattern analysis using the experimental setup of a self-constructed tool LAZY has been done to show a sample of browsing pattern analysis and the effects it might have on today's applications. Why is it needed? The tool analyses the browser data of the user and places the user into certain demographic which is greatly beneficial to the user who now receives suggestions for relevant content. It is also beneficial to the content developers who can now receive data about user preferences which could further boost the navigability of the website and thereby the profitability of the website.

Keywords

Internet of Things (IoT), optimisation, user preference mapping, browsing pattern analysis, demographic, relevant content, navigability, profitability, website.

1. INTRODUCTION

In this digital era, a website not only provides a cost-effective and timely platform for communication with stakeholders but also helps to shape and present its image on the internet [1]. This vital role played by the websites provides data of immense value and an opportunity of using this platform to shape an image of the global society.

Analysing user browsing patterns although a complex task yields beneficial results which could change the way many industries such as advertising, search engine optimisation, Organisational Security work. It provides critical input data which on processing can yield user demographics leading to efficient demographic based, user-targeted applications. [2][3] Analysing user browsing patterns is an emerging concept using which it is possible to categorize users into certain demographics. Through this process, delivering user-specific content can be optimised since the deliverables are for a demographic and not a single user.

Browsing pattern analysis can be helpful in the following ways:

- It can prove to be critical in the identification of preferred content of a part of the population which can be considered geographically. This can be helpful in customising the user deliverables according to the geographical preference of that area

which increases the probability of the user liking it considerably.

- It can find a major application in the field of Search Engine Optimisation. The results displayed by the search engine can be further filtered based on user preference and the regional preferences.

In this paper, user browser pattern analysis has been used to determine the preferences of the users in various geographical regions developing demographic data which maps preferences to a group of users. For this purpose, a tool LAZY has been developed which is a standalone, cross-platform tool. The feedback sent from this tool provides the browsing pattern analysis of the user.

The paper has been organized as the follows: the next section will highlight the research background and the user browser analysis process. Section 3 highlights the experimental setup. In Section 4 of the paper, the results are discussed and the formulated framework is displayed. In the end, paper is concluded.

2. RESEARCH BACKGROUND

Analysing user browsing patterns can be a complex process due to much ambiguity existing in the field. There are various methods through which it can be done such as collecting data on a single user basis, collecting data on a geographical basis and collecting data on a preference basis for a group of people. These methods can also be combined with each other. This field although having critical applications in many fields is not given enough credit. This concept finds its application in various fields such as Search Engine Optimisation, Security and most importantly the field of advertising.

Despite there being approaches trying to utilize this concept to boost the efficiency of web applications to their full capacity, the full potential of this concept has yet to be utilized. The traditional approaches perform the analysis on a single user basis making it more complex and time taking to find the results. The research done in this paper tries to show how the categorical differentiation of users on a geographical level prevails over the traditional single user approach.

For each search on the World Wide Web, there are many websites available which provide information related to the search. The problem arises when there is a necessity to choose a site from the many sites available. The closest answer to the problem may not be at the top of the results. Thus, a tool has been prepared which collects user categorization and can help display the most relevant content to the user, thereby increasing the much-deserved navigability of the selected site.

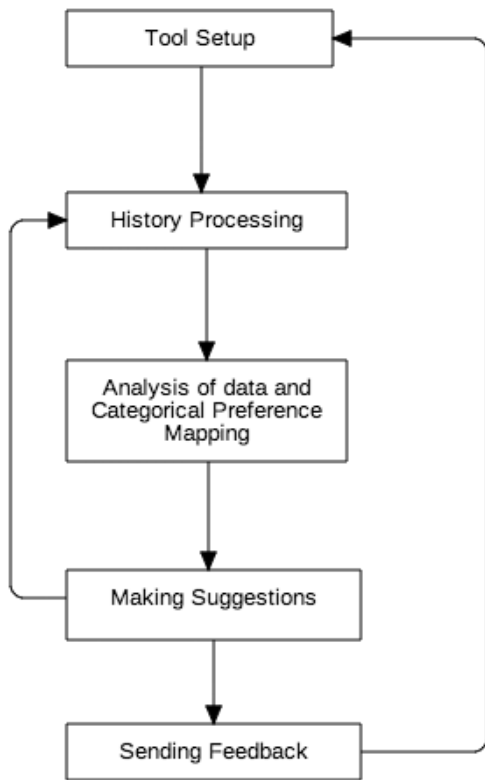


Fig. 1: Browsing Pattern Analysis Process

Browsing pattern analysis follows the following iterative process: [4][5]

2.1 Tool Setup

A pattern analysing tool is set up and installed on the system.

2.2 History Processing

The user history is extracted from the browser and is converted to a usable state.

2.3 Analysis of data and Categorical Preference mapping.

The processed history is then provided to the tool as an input. The tool processes the input and maps the user to the category most likely preferred by the user.

2.4 Suggestions

The user is suggested content based on the mapped category which makes the user browsing experience much better.

2.5 Sending Feedback

The mapped category of each user is sent anonymously (user identification is not disclosed).

This helps understand the preferences of a demographic.

3. EXPERIMENTAL SETUP AND PRACTICAL IMPLEMENTATION

To perform this study, a sophisticated tool has been developed which can process the user history and analyse the search patterns of the user and suggest sites to the user based on preferences. This is done through the implementation of four modules respectively:

- Preparation of generic categories
- URL Filtration
- String Comparison and website validation
- Selecting the preferred category and making suggestions
- Sending Feedback

Preparation of generic categories

In this module, the list of most preferred websites worldwide taken from Network Mapper which is a security tool which produces this list periodically is sorted into categories based on their genre. These lists are then stored into files for further processing.

URL Filtration

The URL filtration is done to the user history, converting it into a usable, sophisticated list of websites comprising of the user browser history. The raw input file is obtained from the user browser via a JSON extension, Export History to JSON.

String Comparison and website validation

The output of the URL Filtration phase serves as an input to the String Comparison and website validation phase of the project. The browser history file is compared to the category files created in module 1 (Preparation of generic categories) using the Rabin Karp String Comparison algorithm [6]. Through this, the number of hits (the number of matches in the category and history files) received by each category are calculated.

Selecting the preferred category and making suggestions

The number of hits corresponding to each category are sorted using Quick Sort [7] the results of which provide the preferred category of the user. Then suggestions and user likable predictions can be made. The suggestions are made based on the most visited websites and the predicted websites which might be preferred by the user based on the category of preference.

Sending Feedback

Feedback is sent to the developer which can be processed in order to prepare a global preference of all the users which can have various applications and be very helpful in many fields which require a general outlook of their user base.

Practical Implementation

This experimental setup has been tested on the university class level and the results were found to be promising. The setup consisted of 45 individual users who used the tool. The tool was installed on each of their systems and their browser history was extracted and refined. The tool processed lakhs of characters received as input and mapped each user to one category based on a maximum number of hits to that category. The categories considered for this setup were:

- Educational
- Entertainment
- Social Media
- News
- Sports

Out of the 45 users 10 preferred Social Media, 6 preferred sports, 3 preferred news, 6 preferred educational while a whopping 20 preferred entertainment. This showed out of the 45 tested the most preferred category was entertainment which accounted for about 50%. This category consisted of sites such as YouTube, Hotstar, Netflix etc.

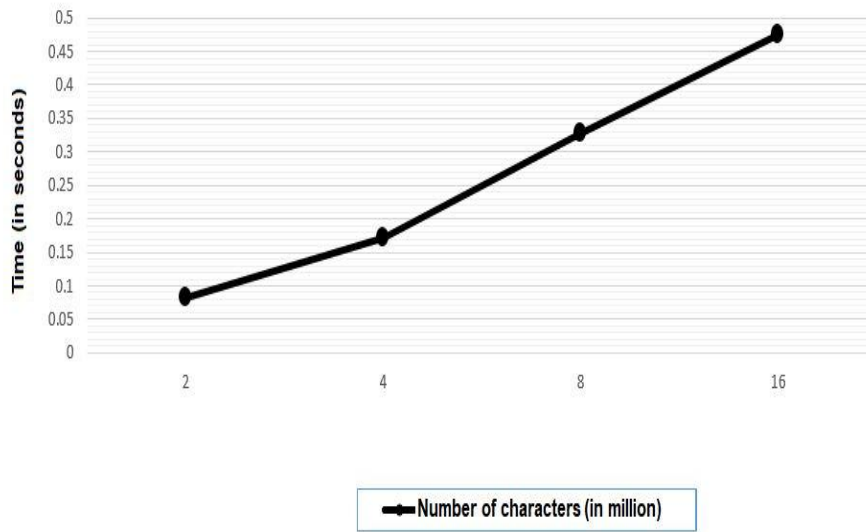


Fig. 2: Input Processing Efficiency

The current batch under experimentation preferred the entertainment and on being enquired answered that they would like to receive relevant content.

4. RESULTS

This experimental setup goes on to show that user browsing patterns can be predicted and through appropriate measures and utilization of this data the browsing experience of any user can be improved by several times. The experiment of the user at hand yields data which states that the user prefers the entertainment category over others and so the suggestions made are based on this result.

From this setup, it is evident that the user browsing patterns can be analysed and predicted by using user browser data and the list of the most popular websites (from the generic user point of view). The benefit and strength of user browsing pattern analysing lies in its many real-world applications. These patterns obtained from many users can be processed and can help make generic assumptions for a large group of people which can have various applications. Some such applications include:

Advertising

The field of advertising can be very ambiguous when it comes to user preferences. Through the analysis of user browsing patterns and preferred categories this can be improved. The advertiser can be given information of the user's preference so advertisements can be given according to user preferred category. For example, if a user prefers education as a category then advertisements related to books, educational institutions, coaching centres can be given to that user.

Security

The power of the browsing pattern analysis is that it can filter out users who prefer malicious content and help in securing an organization, or any company in general. Through the creation of new categories, the tool adapts to the particular user and reads the user patterns and maps preferences. This can be vital as any user with a preference to malicious categories can be flagged and dealt with accordingly.

Search Engine Optimisation

The output dataset of this process can be used in search engine optimisation to order the search results. This is done by collecting user preferences regionally and ordering the search results according to the preferences generated.

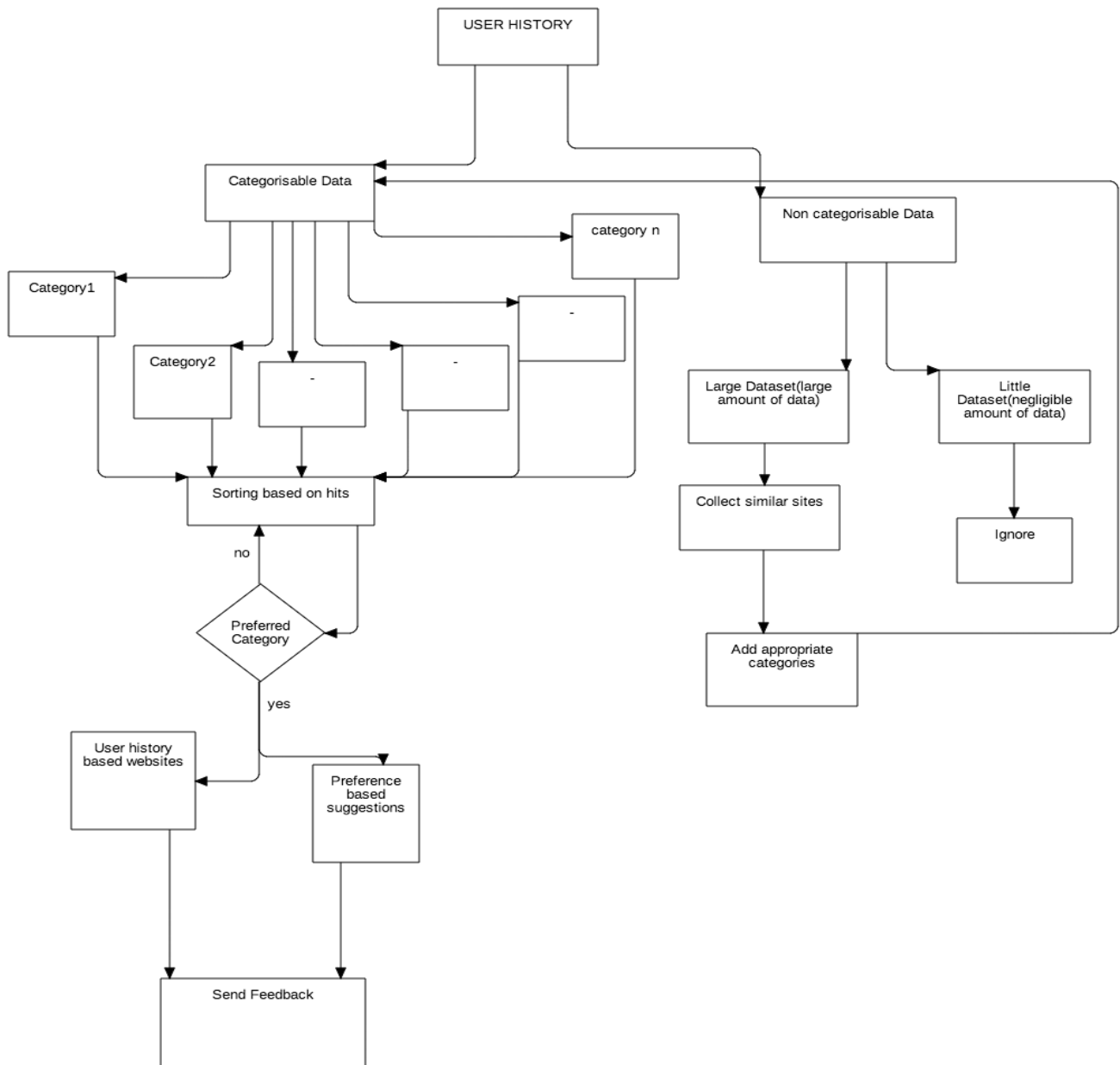


Fig. 3: Pattern Analysis Framework

5. CONCLUSION AND FUTURE SCOPE

In this paper, it has been stated how the developed tool works by analysing user browsing patterns and places users into certain demographics which is very beneficial for users as well as web developers and administrators. By the results obtained a framework has been composed which can be used for the analysis of user browsing patterns. This has been tested and proved to be useful. While the traditional models of browsing pattern analysis focus on the individual user this framework excels as it considers a group of users according to their preferences while processing data. An advantage of browsing pattern analysis is that it makes accessibility of content in the World Wide Web much easier and helps find related content faster. It also helps by providing crucial user based information which can be used to create interactive applications.

This work can be used to on a larger scale such as an entire nation in order to deliver likable content to the people and suggesting content which has a high probability of likability.

6. REFERENCES

- [1] Menten S. Ahmet, Aykut H. Turan, "Assessing the usability of university websites: an empirical study on Namik Kemal University", TOJET: The Turkish Online Journal of Educational Technology 11.3, 2012.
- [2] S.Roy, P. Pattnaik, R. Mall, "A quantitative approach to evaluate usability of academic websites based on human perception", Egyptian Informatics Journal, vol. 15, no. 3, pp. 159-167, 2014.
- [3] T. Stewart, "Websites- Quality and Usability", Behaviour & Information Technology, vol. 31, no. 7, pp. 645-646, 2012

- [4] D.Waisberg,A.Kaushik,"Web Analytics2.0: Empowering Customer Centricity" in The Original Search Engine Marketing Journal 2, vol. 2, no. 1, pp. 5-11, 2009.
- A. Dingli, S. Cassar, "An Intelligent Framework for Website Usability", *Advances in Human-Computer Interaction*, vol. 2014, pp. 1-13, 2014.
- [5] Er. Mohammad Shabaz, Er. Neha Kumari, "Advance Rabin Karp Algorithm for string matching" *International Journal of Current Research*, Vol 9, Issue 09, pp.57572-57574, September 2017
- [6] Abdulrahman Hamed Almutairi, Abdulrahman Helal Alruwaili, "Improving of QuickSort Algorithm Performance by Sequential Thread or Parallel Algorithms." *Journal of Computer Science and Technology Hardware & Computation*, Volume 12, Issue 10-Version-1.0,2012