Effective Prediction of Kid's Behaviour Based on Internet Use

Neetu Anand

Dept. of Computer Sc., Maharaja Surajmal Institute, Janakpuri, Delhi.

Abstract

Internet use has become essentially universal among adolescents. Young people are exposed to an increasingly open and collaborative online social culture, whichallows them to access information and maintain friendships and relationships. Young people are, however, at a dynamic stage of development in which risk-taking behaviours and immature decision making capacities can lead to negative outcomes. Parents' involvement in safe use of technology starts from a child's first use, and they are a critical part of ensuring their teenage children's responsible and safe use of online services. Monitoring Internet use and communicating with teenage children are important strategies. The aim of the paper is to help parents for monitoring their child computer and Internet usage details and provide them with the tools to understand the online behaviour of their teenage children.

Keywords: Web mining; monitoring; classification; pre-processing

1. Introduction

Nowadays Internet and computer usage grow dynamically among youngsters. Monitoring the Internet usage of their teenage children is a strategy that is often suggested for parents, and the literature indicates that the majority of parents do engage in monitoring behaviours at least some of the time. Monitoring behaviours include checking that sites are appropriate for their child's use, and keeping an eye on the screen, with checks more likely to occur at younger ages [8].

Certain factors are seen as making it difficult for parents to monitor and manage children's Internet use including:

a) Not being able to keep an eye on the screen or what the child is doing all of the time. This is especially applicable in the age of wireless connections and

Internet-enabled mobile phones, and access that is occurring within school time [9];

b) The amount of time children and teenager spend on Internet-related activities;
c) Children's resistance to time limits;
d) Difficulty of preventing exposure to inappropriate content;
e) Children's own control of details (e.g., through use of passwords, phone locks and private browsing;
f) Difficultyparents have in keeping up with the pace of change on the Internet, particularly social networkingand virtual reality sites.

By considering all of these factors the objective of this paper is to record all the usage data of the child and store it for the parents for the future use and apply the different mining techniques to check the web behaviour of the child. The rest of the paper is structured as follows. Section 2 reviews the related work on behavioural analysis .In Section 3 we shortly explain the basics of web mining; Section 4 presents the way to collect data. In Section 5 weexplain the workings of our approach and finallydraw the conclusions and future directions in Section 6.

2. Related Work

Dembczyński et al. [1] have created dataset consist of 4882 different users and 20 different Web page categories. They have explored the week and 24-hour periodicity in users' behaviour, among others. All of their algorithms scale well with large data and have linear time complexity, which is the smallest possible complexity for such problems.Kellar[2] propose two further studies: an exploration of monitoring behaviour on the Web through semi-structured interviews and the development and validation of recommendations for the design of future Web browser tools. Morrison et al. [3] present three taxonomic classification schemes based on Web users' responses to what Web activities significantly impacted their decisions and actions. The taxonomic classifications focus on three variables: the Purpose of people's search on the Web, the method people use to find information, and the Content of the information for which they are searching. Sellen et al. [4] studied the Web activities of 24 knowledge workers over two days .Participants were interviewed in front of theirweb history at the end of the second day and described the different activities in which they engaged. Activities wereclassified into six main categories: finding, information gathering, browsing, transacting,communicating,and housekeeping. Then compared the duration of all. Shahabi [5] have introduced a remote agent(javascript and applet) thatacquires the user interactions from the browser at the client side. Vespan [6] describe the process of discovering internet user behaviourmodels by analysing proxy server raw data and they emphasize theimportance of such models for the e-learning environment. They have usedURL as source of input to analyse Web user behaviour. They have identify six different behavioural models namely study and research,daily life, shopping and products/services information,on-line communication,job, and entertainment. Krzysztof [7] describes a method for classifying users' interests based on an analysis of the site-keyword graph. The method

is for extracting sub-graphs representing users' main interests from a site-keyword graph which is generated from augmented Web audience measurement data (Web log data).Their solution has low time and space complexity; scales well with large datasets and, at the same time, produce high-quality results.

3. Mining the Web

Web mining is a very broad research area emerging to solve the issues that arise due to the WWW phenomenon. The Web mining research is a converging research area from several research communities, such as Databases, IR and AI.[10] Web mining can be defined roughly as data mining using data generated by the web and includes the following sub areas: web content mining, web usage mining, and web structure mining. Web usage mining is an important and fast developing area of web mining where a lot of research has been done already. Web usage data includes web server access logs, proxy server logs, cookies, user profiles, registration data, user sessions, user queries, bookmark data, mouse clicks and scrolls and any other data as the results of interactions.

The analysis of the user's click data is known as Web usage mining.Web usage mining literature focuses on techniques that could predict user behaviour patterns. Web usage mining studies could be classifiedinto two categories: those based on the server side and those based on the client side. Studiesbased on the client side retrieve data from the user using cookies or other methods.The other application area of web usage mining includes effective web site management, creating adaptive sites, business and support services, personalization, and network traffic flow analysis and so on. Web usage mining consists of three main phases, namely Data Pre-processing, Pattern Discovering and Pattern Analysis.

4. Creating Usage Dataset

Behavioural profiling involves collecting data (recording, storing and tracking) and searching it for identifying patterns (with the help of data mining algorithms).
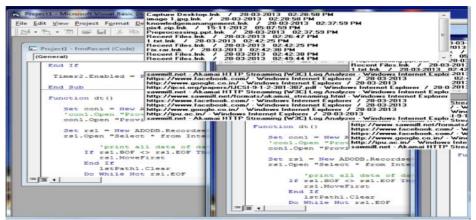

Fig. 1: Software for collecting usage data.

The data collection phase is often referred to as *Behavioural Tracking*.The Web user behaviour has been contributed to a variety of factors from timeto time.Secondly, collecting rich and detailed user data is another challenge poised toour research. There are no direct means to collect both the computer usage information as well as Web browser interactions.The software for collecting and storing usage data(computer and internet both) has been created and the screenshot for the same is shown below.

The objectives of this work are to monitoring, observing and understanding how the computer and web is used and utilized by the children and teenager and then find their behaviour.

There are various advantages of finding user behaviour

1. Information so gathered could be used to offer feedback to parents of users.
2. To apply the mining techniques to the recorded dataset so that parents can put an eye on the usage record of their child.
3. To put restriction on the child computer and internet usage.

5. Process of Finding User Behaviour

Usage history refers to data recorded into a log file as a user (child) enters any internet address on a web browser's using URL bar. Next step is Pre-processing of web data to make it suitable for mining .It is identified as one of the key issues for Web mining. The collected internet log file includes path and browser,access date and access time.Out of path and browser the key items are get filtered in pre-processing step. To derive the behaviour roles of children, their activity patterns need to be compared against the behavioural characteristics of each role type. Afterwards log datacan be classified by user behaviour. We classify child behaviour on web browser as acquiring knowledge, download, buying, work, relationship and Enjoyment etc. Web user behaviour is also influenced by externalfactors. For example student is preparing for his/her exam, so the nature of websurfing is related to study and education based surfing. The table showing classification of child behaviour depending on the usage pattern is as shown below:

Table 1: Classifying child behaviour for Internet usage data.

BEHAVIOUR	SPECIFIC CLASSIFICATION	KEYWORDS
Acquiring Knowledge	Explore, Viewing the Document, Learning, Blog, Map, News sites Weather sites	Explore *.doc,* .ppt,* .xls,* .pdf, *.txtetc. Lecture, study, learn Blog Map News Weather ,rain ,sunny, temp
Buying	Shopping, Reservation	Shop, store, stock, collection Reserve, ticket, booking

Work	Mail,	Mail, email, e-mail
	Calendar,	Cal
	Banking	Bank
	Registration	Register
Relationship	Community,	forum, cafe, society, group, community, social sites
	Chat	address
		chat
Download/copy	ftp:// ,download	ftp:// ,download
Enjoyment	Game,	Game,
	Movie,	Movie,
	Music,	Music,
	Cartoon,	Cartoon,
	Sports,	Sports,
	Videos	Videos

The child behaviour is categorized by using the above stated rules.**Monitoring what the teen is searching and where they're going online will help the parents to**block content being sent from the computer, set limits on the time that children can go online, filter inappropriate content and check child**online behavior offline.**

6. Conclusion and Future Work

In this paper, we have proposed a new method for monitoring child's behavior during Internet use that can help the parents to lower the harmful or dangerous situations .Parents are need to be more aware of the times when young people are on-line, in contact with others and explain them which sites are inappropriate and off-limits, amount of time and times of day computer is used, and which information cannot be shared online. Further in future we will apply the mining techniques on the same data for finding the behavioral pattern,that can help the parents to easily guide their children to safe and rewarding online experiences.

7. Acknowledgement

References

[1] Dembczyński K., Kotłowski.W, Sydow.M (2008), Effective Prediction of Web UserBehaviour with User-Level Models , Fundamental Informaticae, IOS Press , 0169-2968 (Print) 1875-8681 (Online), Volume 89, Number 2-3 / 2008, page 189.

[2]   M. Kellar(2006),'An examination of user behaviour during Web information tasks',Proceeding CHI EA '06 CHI '06 extended abstracts on Human factors in computingsystems ACM New York, NY, USA ©2006.

[3]   Morrison, J. B., Pirolli, P. and Card, S. K. (2001).,'A Taxonomic Analysis of WhatWorld Wide Web Activities Significantly Impact People's Decisions and Actions.', InProceedings of the CHI '01 Extended Abstracts on Human Factors in ComputingSystems, Seattle, WA, 163-164.

[4]   Sellen, A. J., Murphy, R. and Shaw, K. L. (2002).,'How Knowledge Workers Use theWeb.', In Proceedings of the SIGCHI Conference on Human Factors in ComputingSystems (CHI 2002), Minneapolis, MN, 227-234.

[5]   Cyrus S., Farnoush B.K. (2001),'A Framework for Efficient and Anonymous WebUsage Mining Based on Client-Side Tracking' , Lecture Notes In Computer Science;Vol. 2356 archive Revised Papers from the Third International Workshop on MiningWeb Log Data Across All Customers Touch Points table of contents Pages: 113 − 144Year of Publication: 2001 ISBN:3-540-43969-2.

[6]   Dragoş M. Vespan,(2007),'Internet User Behaviour Model Discovery Process',Revista Informatica Economică, nr. 3 (43)/2007.

[7]   Krzysztof D., Wojciech K., Marcin S. (2008), Effective Prediction of Web UserBehaviour with User-Level Models , Fundamenta Informaticae , IOS Press , 0169-2968 (Print) 1875-8681 (Online), Volume 89, Number 2-3 / 2008 , page 189.

[8]   www.mass.gov/ago/...the...internet.../monitoring-internet-use.html

[9]   www.stopmedicineabuse.org/blog/.../monitoring-your-teens-online-behavior

[10] www.ieee.org.ar/downloads/Srivastava-tut-paper.pdf