INTELLIGENT WEB HISTORY MANAGEMENT

Musthaq Mohamed Aadhil

A dissertation submitted in partial fulfilment of the requirement for
Bachelor of Engineering (Honours) degree in Software Engineering

Department of Computing

Informatics Institute of Technology, Sri Lanka
in collaboration with
University of Westminster, UK

2021

Abstract

In the 21st century, the web and the internet are almost part of the human lifestyle. According to the Internet Usage Stats 2019, On average, people are spending 6 hours and 42 mins on the web each day. Although this web browsing was tracked through several web industries, the manageable aspect of web browsing history is very poor. Mostly those browser trackings were used for advertisement purposes and other business purposes that are less valuable to the users.

The main issue with current web browser history recording is, it is being just a simple log of web URLs. If a user browses around 20 pages per day, they will possibly lose the track of their history record by a day or 2. Because it will be very harder to identify what were the key websites, they visited recently. Maybe a user might have spent few hours on a blog that might be important to their recent work. What if that user forgets to bookmark that web page and wanted to revisit for some reason? Then it will be a challenging task for the user to go through their browsing history to find the correct link.

Also, the modern bookmarking system not much efficient as people will be finding a bunch of web pages that they might think it's important to them. And their bookmark tabs will be stacking up daily and ultimately making it harder to find what's more important for their recent work. It because people usually will not remove any unused bookmarks or too lazy to organize their bookmarks efficiently.

So in this research paper, we are trying to find a way to classify and organize users browsing history in a way that users can quickly identify the important web pages they have visited in a machine learning way. The challenging part of this research will be identifying the browsing behavior of the user when they found an interesting web page.

Keywords: Web History, History Management, Internet, Browser Extension, Rankings