**Monitoring user behavior through Session Web Mining**

**(Standalone application)**

**Group Members :**

**Roll No. Name**

**17402 Swarangi Aher**

**17420 Misba Sayyad**

**17422 Aafiya Sheikh**

**Project**

**Proposal**

**Rationale**

The Project **Monitoring user behavior through session web mining (standalone)** will be designed to track websites visited on a computer and maintain a record of the web session. The project will help the office, institute, department, class in monitoring web usages of students. This aims to solve the problem of accessing games, social media and pornography at work. The system has the solution to the problem as it is a technology so there will be no need of manual monitoring. The whole monthly report generation process will be managed by the standalone Windows application. So, by which there is no need of the authorities inducing the work of finding out what each student, employee is surfing on the internet.

The manual work of checking the browser history and CCTV cameras will be reduced and the records will be generated by the application itself which can later be accessed by the admin with ease.

The report generation like most visited website, web history and sessions along with their duration and timing will be generated on a monthly basis, etc.

The Project will help us to brighten our programming, technical skills like development, database management, data analysis also some non-technical skills like communication, team work, time management etc.

The system will boost the standard of the school, office on promoting high quality education, productivity through the development of the system.

**Introduction**

This software tracks the web sessions and websites visited and clicks a picture of the user accessing a website he/she isn’t supposed to be accessing in the respective school or office. The proposed software also generates a monthly report which included the most visited website and the duration spent on each website. A reason for developing this software is to generate the report automatically at the end of the session or in the between of the session. The application is compatible with Windows OS.

**Systems which are currently available:**

* Websites tracking system on server.
* Android Based website tracking system.
* Applications to block certain websites on a server.

**Purpose**

This software will be used in college, school, office to monitor the web activity. It record session & generate a monthly report session report based on the web session. In school or at work place users often brows website which they are not suppose to be accessing during work or school hours using. This software we can track, monitor web activity and click a picture of user accessing sites like social media, online games, porn sites etc. Using this software we can overcome this problem.

**Scope**

This software will be used for monitoring user activities on the web. It can be used in office, schools to increase productivity. In future, this software can be made more efficient and improved in ways to monitor and track user behavior better way.

**Literature survey**

This software for Monitoring user behavior through session web mining (standalone) is a software solution for monitoring the web activity. The system maintains a daily record and a monthly report. The proposed system database is an contains details regarding web sites visited and their sessions. The implementation plan will be amended and updated as new actions are introduced and as developments are implemented to the system. The system tracks the websites visited on a Computer along with their durations. Background of the Study The proposed system intended for tracking web activity to increase productivity at work and schools.

**Abstract**

The Project **Monitoring user behavior through session web mining (standalone)** will be designed to track websites visited on a computer and maintain a record of the web session. The project will help the office, institute, department, class in monitoring web usages of students. This aims to solve the problem of accessing games, social media and pornography at work. The system has the solution to the problem as it is a technology so there will be no need of manual monitoring. The whole monthly report generation process will be managed by the standalone Windows application. So, by which there is no need of the authorities inducing the work of finding out what each student, employee is surfing on the internet.

The manual work of checking the browser history and CCTV cameras will be reduced and the records will be generated by the application itself which can later be accessed by the admin with ease.

**Review on Literature**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr. No** | **Title** | **Author Name** | **Year** | **Published At** | **Findings** | **GAP** | **Future Direction** |
| 1 | Micro-policies for Web Session Security | Stefano Calzavara,[Riccardo Focardi](https://ieeexplore.ieee.org/author/37324204400),[Niklas Grimm](https://ieeexplore.ieee.org/author/37085660085),[Matteo Maffei](https://ieeexplore.ieee.org/author/37424515600). | 2016 | IEEE | In this paper, they presented the first application of micro-policies to web security, by studying its effectiveness at securing web sessions. | In this paper, they presented the first application of micro-policies to web security, by studying its effectiveness at securing web sessions. | It will help researchers to make more secure web sessions. |
| 2 | Automate session setup based on machine learning | [Jagannath Putrevu](https://ieeexplore.ieee.org/author/37847921400),  [XiaotaoWu](https://ieeexplore.ieee.org/author/37293513600),  [Venkatesh](https://ieeexplore.ieee.org/author/37297108100)  [Krishnaswamy](https://ieeexplore.ieee.org/author/37297108100) | 2010 | IEEE | This paper proposes a system of using machine learning algorithms to extract communication session information, such as conference bridge number and participant code, from users' emails or appointments. | The implementation is complex.  The authors were so correct that this research doesn’t have any limitations. | It will help the researchers choose the appropriate algorithm to do the specific task. |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr. No** | **Title** | **Author Name** | **Year** | **Published At** | **Findings** | **GAP** | **Future Direction** |
| 3 | SHPF: Enhancing HTTP(S) Session Security with Browser Fingerprinting | [Thomas Unger](https://ieeexplore.ieee.org/author/37086741709),  [Martin Mulazzani](https://ieeexplore.ieee.org/author/37400147900),  [Dominik Frühwirt](https://ieeexplore.ieee.org/author/37086738045),  [Markus Huber](https://ieeexplore.ieee.org/author/37392400500),  [Sebastian Schrittwieser](https://ieeexplore.ieee.org/author/37889064900),  [Edgar Weippl](https://ieeexplore.ieee.org/author/37400148600) | 2013 | IEEE | In this paper, they presented their framework SHPF which is able to raise the bar for session hijacking. | The implementation is complex.  The authors were so correct that this research  doesn’t have any limitations. | Web security can be increased |
| 4 | Adaptive Rule Loading and Session Control for Securing Web-Delivered Services | [Yu Zhang](https://ieeexplore.ieee.org/author/38008219700)  [Vugranam Sreedhar](https://ieeexplore.ieee.org/author/37338224600)  [Lin Luo](https://ieeexplore.ieee.org/author/38106023100)  [Shun Xiang Yan](https://ieeexplore.ieee.org/author/38106762100)g | 2009 | [IEEE](https://ieeexplore.ieee.org/) | In this paper, they present Arctic, an adaptive reinforcement learning control technique for Web intrusion check. | The implementation is complex.  The authors were so correct that this research doesn’t have any limitations. | It will be helpful in making web more secure. |
| 5 | Modelling User Behaviour for Web Recommendation Using LDA Model | Guandong Xu, Yanchun Zhang, and Xun Yi | 2008 | [IEEE](https://ieeexplore.ieee.org/) | In this paper, we propose a collaborative Web recommendation scheme based on Latent Dirichlet Allocation (LDA) model | The authors were so correct that this research doesn’t have any limitations. | It helps improving web recommendation using user behavior |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr. No** | **Title** | **Author Name** | **Year** | **Published At** | **Findings** | **GAP** | **Future Direction** |
| 6 | CARENA: a tool to capture and replay Web navigation sessions | [I.J. Nino](https://ieeexplore.ieee.org/author/37562946600)  [B. de la Ossa](https://ieeexplore.ieee.org/author/37562946400)  [J.A. Gil](https://ieeexplore.ieee.org/author/37275210200)  [J. Sahuquillo](https://ieeexplore.ieee.org/author/37274197100)  [A. Pont](https://ieeexplore.ieee.org/author/37267144400) | 2005 | [IEEE](https://ieeexplore.ieee.org/) | In this paper, they have proposed a tool to help web performance evaluation studies | .The implementation is not given | It will help improve web performance |
| 7 | System for Web information monitoring | [Wieslaw Lubaszewski](https://ieeexplore.ieee.org/author/38667816800)  [Krzysztof Dorosz](https://ieeexplore.ieee.org/author/38667938800)  [Michal Korzycki](https://ieeexplore.ieee.org/author/38667815500) | 2013 | [IEEE](https://ieeexplore.ieee.org/) | This paper describes the functionality and the architecture of the MPI system, which was developed for Web information monitoring | The implementation is not given | I |
| 8 | A Web Service Monitoring Indicator and Model System and Performance | Yinsheng  Zhang  Xiaodong  Qiao  Feng Han  Jitian Wang  Jian Liang  Peng Li | 2009 | [IEEE](https://ieeexplore.ieee.org/) | This paper analyses the technical limitations in the monitoring web services | Implementation is not given | It will be used to give better web services. |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr. No** | **Title** | **Author Name** | **Year** | **Published At** | **Findings** | **GAP** | **Future Direction** |
| 9 | Monitoring QoS parameters of composed web services | [J. Karthikeyan](https://ieeexplore.ieee.org/author/37085461725)  M. Suresh Kumar | 2014 | [IEEE](https://ieeexplore.ieee.org/) | In this paper, Web service is a software system designed to support interoperable machine-to-machine interaction over a network. | Equations given in this paper are difficult to understand. | It will be used for machine to machine interaction. |
| 10 | A Survey on Web Tracking: Mechanisms, Implications, and Defenses | Tomasz Bujlow   Valentín Carela-Español  Josep Solé-Pareta  [Pere Barlet-Ros](https://ieeexplore.ieee.org/author/38270117900) | *2017* | [IEEE](https://ieeexplore.ieee.org/) | In this survey, they review the existing literature on the methods used by web services to track the users online as well as their purposes, implications, and possible user's defenses. | The authors were so correct that this research doesn’t have any limitations. | It will be used to protect privacy of users |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr. No** | **Title** | **Author Name** | **Year** | **Published At** | **Findings** | **GAP** | **Future Direction** |
| 11 | Linux Network Security Access & Monitoring Service Tools | Mahesh Kumar | 2018 | International Journal of Engineering Research and Technology (IJERT) | In this topic security in Linux Environment implementation and research of enhancing network security is done. | Implementation is not given. | Network security can be inhanced in Linux |
| 12 | An Effect of Cyber-Attacks on Session Management | K Bhupathi Reddy,  P Ashok Kumar,  Dr G. V. Ramesh Babu | 2018 | International Journal of Engineering Research and Technology (IJERT) | This describes Web application design flaws that could be exploited for session management attacks and discusses these flaws' current prevalence. | The authors were so correct that this research doesn’t have any limitations. | It will be helpful is reducing session management attacks. |
| 13 | Mobile Tracking System using Web Application and Android Apps | Mia Md. Karimul Hoq,  Mohammad Jahangir Alam,  Md. Nurul Mustafa | 2017 | International Journal of Engineering Research and Technology (IJERT) | In this paper*,* This Mobile Tracking System has been designed | The authors were so correct that this research doesn’t have any limitations. | It will help in tracking web applications of android. |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr. No** | **Title** | **Author Name** | **Year** | **Published At** | **Findings** | **GAP** | **Future Direction** |
| 14 | Predicting user behavior through Sessions using the Web log mining for an E commerce application | Sushmeendra N Rao,  Rakesh B,  Pallavi N Hegde,  Anusha R Kotur | 2019 | International Journal of Engineering Research and Technology (IJERT) | It is the method to extract the user sessions from the given log files. | The authors were so correct that this research doesn’t have any limitations. | It will help to understand Web log mining ,User identification , Session identification. |
| 15 | Usage Analysis of Web Access Behavior | Ankita Gaur | *2018* | International Journal of Engineering Research and Technology (IJERT) | In this paper, the various concepts of data mining are used and also some traditional data mining concepts for exploring the web usage | *The implementation is complex as it deals with the user web behavior.* | The work presented here belongs in the research area of data mining as applied to data on the web. |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr. No** | **Title** | **Author Name** | **Year** | **Published At** | **Findings** | **GAP** | **Future Direction** |
| 16 | A New Multilevel Hierarchial Architecture Approach for Secure User Authentication and Session Management | Akshatha B M, Harshini B H, Kavya S B, Kavya M S, Sanjay Kumar N V | 2019 | International Journal of Engineering Research and Technology (IJERT) | This paper explores promising alternatives offered by applying new approach to user session management. | The implementation is complex and difficult to understand. | It helps to provide better session management and secure user approach |
| 17 | Curlcrawler Optimization: A Framework for Crawling the Web with URL Tracking and Canonicalization | Ashok Kumar  Dr. Saurabh Mukherjee  Manisha Garhwal | 2013 | International Journal of Engineering Research and Technology (IJERT) | The aim of this paper is to raffle a framework, which will elevate search engine’s dexterity. | It is very complex and hard to understand. | It will be used to improve search engines. |
| 18 | Precognition of Users Web Browsing Behaviour | M. Trupthi, Dr. Suresh Pabboju | 2014 | International Journal of Engineering Research and Technology (IJERT) | In this paper, they proposed and accelerator to let a website know when a user will visit it. | The implementation is complex as it deals with tracking user web behavior. | It will be used for user behavior prediction. |
| **Sr. No** | **Title** | **Author Name** | **Year** | **Published At** | **Findings** | **GAP** | **Future Direction** |
| 19 | Analyzing the user Action for Web Content Augmentation | Dheeraj Kumar N,  Joyce Vinnarasi A ,  Sheethal Chinnappa C ,  Tejaswini S,  Dr. Mahesh T R | 2018 | International Journal of Engineering Research and Technology (IJERT) | In this paper, we have studied about exploring the users action for web content augmentation | The implementation is complex as it deals with tracking user web behavior. | It will be used for user behavior prediction. |
| 20 | A Survey on Incorporating User Behaviour to Enhance the Web Search | S. Phadke  C. Banchhor | 2013 | International Journal of Engineering Research and Technology (IJERT) | This paper statesmany techniques by which we can improve the searching process | The implementation is complex as it deals with tracking user web behavior. | It will be used for user behavior prediction. |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr. No** | **Title** | **Author Name** | **Year** | **Published At** | **Findings** | **GAP** | **Future Direction** |
| 19 | Analyzing the user Action for Web Content Augmentation | Dheeraj Kumar N,  Joyce Vinnarasi A ,  Sheethal Chinnappa C ,  Tejaswini S,  Dr. Mahesh T R | 2018 | International Journal of Engineering Research and Technology (IJERT) | In this paper, we have studied about exploring the users action for web content augmentation | The implementation is complex as it deals with tracking user web behavior. | It will be used for user behavior prediction. |
| 20 | A Survey on Incorporating User Behaviour to Enhance the Web Search | S. Phadke  C. Banchhor | 2013 | International Journal of Engineering Research and Technology (IJERT) | This paper statesmany techniques by which we can improve the searching process | The implementation is complex as it deals with tracking user web behavior. | It will be used for user behavior predictio. |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr. No** | **Title** | **Author Name** | **Year** | **Published At** | **Findings** | **GAP** | **Future Direction** |
| 21 | Client-Side Monitoring for Web Mining | Kurt Fenstermacher and Mark Ginsburg | *2003* | Citeseerx  Normal paper | This paper statessurvey server-side analysis, describing the collection of the data and its shortcomings for inferring Web user behavior | The authors were so correct that this research doesn’t have any limitations. | It will be used for client side monitoring |
| 22 | On Mobile User Behaviour Patterns | Milan Vojnovic | *2008* | Citeseerx  Normal paper | This paper states empirical analysis of human searches for information from mobile devices, focusing on temporal dynamics, semantics, and topics of queries. | The authors were so correct that this research doesn’t have any limitations | It will be hepful in the study of mobile pattern behavior. |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr. No** | **Title** | **Author Name** | **Year** | **Published At** | **Findings** | **GAP** | **Future Direction** |
| 23 | Modeling User Behavior by Planning | Alfredo Milani, and Silvia Suriani | 2008 | Citeseerx  Normal paper | In this paper, A planning approach to user behaviour recognition has been introduced. | The equations are hard to understand. | It will be helpful in studying user behavior on web. |
| 24 | Learning rules from user behavior | Domenico Corapi , Oliver Ray , Alessandra M. Russo1 , Arosha K. Bandara , and Emil C. Lupu | 2008 | Citeseerx  Normal paper | *This paper states that* We have shown that by extending a non-monotonic ILP learning system it is possible to learn incrementally and revise rules describing user behaviour. | The implementation is not given | It can be used to study and analyze user behavior |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr. No** | **Title** | **Author Name** | **Year** | **Published At** | **Findings** | **GAP** | **Future Direction** |
| 25 | Data Preparation for Mining World Wide Web Browsing Patterns | Robert Cooley, Bamshad Mobashr and Jaideep Srivastava | *1999* | Citeseerx  Normal paper | This paper has presented the details of preprocessing tasks that are necessary for performing Web Usage Mining, the application of data mining and knowledge discovery techniques to WWW server access logs. | Session identification is quite complex | It can help in preprocessing tasks to WWW server access logs. |
| 26 | Web Mining Research: A Survey | Raymond Kosala,  Hendrik Blockeel | *2000* | Citeseerx  Normal paper | In this paper, they survey the research in the area of web mining. We point some confusions regarding the usage of Web mining | The authors were so correct that this research doesn’t have any limitations | It helps in web mining and information extraction*.* |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr. No** | **Title** | **Author Name** | **Year** | **Published At** | **Findings** | **GAP** | **Future Direction** |
| 27 | Modeling and Predicting the Task-by-Task Behavior of Search Engine Users | Claudio Lucchese,  Salvatore Orlando, Raffaele Perego, Fabrizio Silvestri, Gabriele Tolomei | *2013* | Citeseerx  Normal paper | In this paper, they presented the first proposal for a novel task recommender system, which generates task recommendations on the basis of the task-by-task behavior of several web search engine users. | The authors were so correct that this research doesn’t have any limitations | It helps in enhancing the user web search engine experience |
| 28 | Enhancing Web Search User Experience: from Document Retrieval to Task Recommendation | Gabriele Tolomei,  Tutore del Dottorando,  Direttore della Scuola | *2011* | Citeseerx  Normal paper | In this dissertation, they have presented two research challenges that they claim next generation Web search engines should deal with in order to enhance the overall user search experience. | The calculation and analysis is compex*.* | It helps in improving user web search experience |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr. No** | **Title** | **Author Name** | **Year** | **Published At** | **Findings** | **GAP** | **Future Direction** |
| 29 | Surviving the Web A Journey into Web Session Security | Stefano Calzavar, Riccardo Focardi, Marco Squarcin,  Mauro Tempesta | 2017 | Citeseerx  Normal paper | This paper takesa retrospective look at different attacks against web sessions and we surveyed the most popular solutions against them. | The authors were so correct that this research doesn’t have any limitations | It helping in finding solutions against web attacks. |
| 30 | TaintDroid: An Information-Flow Tracking System for Realtime Privacy Monitoring on Smartphones | William Enck,  Peter Glibert,  Landon P.rox,  Jaeyeon Jung,  Patrick McDanie,  Anmol N.Shenth | 2010 | Citeseerx  Normal paper | This paperpresents TaintDroid, an efficient, system-wide information flow tracking tool that can simultaneously track multiple sources of sensitive data. | The authors were so correct that this research doesn’t have any limitations | It helps in enabling protection of sensitive data. |

**Problem Definition**

This software will be used in college, school, office to monitor the web activity. It record session & generate a monthly report session report based on the web session. In school or at work place users often browse website which they are not supposed to be accessing during work or school hours using. This software we can track, monitor web activity and click a picture of user accessing sites like social media, online games, porn sites etc. without having to manually check the browser history or continuously keep a watch on the user.

**Disadvantages of existing system**

1. It is a standalone application.

2. It can only be installed on Windows OS.

**Proposed methodology**

**Aim:**

To track the user web sessions and maintain a record of it and also click a picture using the web cam of the user trying to access any site they are not supposed to such as gaming sites, pornography sites, etc. The aim is to eliminate the manual work in checking the web history which can easily be deleted.

**Objective:**

To track the web sessions along with their time stamps and durations and maintain a record of it, to click picture of user trying to use any site a student or employee is not supposed to access at workplace or school.

**Resources required**

Software Requirement

* Front end: Java
* Back end: MS Access
* Platform: Windows 10
* IDE: Netbeans

Hardware Requirement

* Processor: Intel i5 core
* Hard disk: 80 GB
* RAM: 4 GB
* Other hardware: Webcam

**Action plan:**

1. **Designing of the app:**

* The design of **Monitoring user behavior through session web mining (standalone)** is having two categories of plan in it. The first plan is based on technical aspects, medium of communication of Web Session and Database. 1. Design of the software is based on Technological Plan. Computer (PC/Laptop), Database, Web sessions involved, details on websites accessed, and maintenance details of the entire system.

The second plan is rest with implementing the principles of management information system to make the entire system effective by which the management users can access the system according to their need and produce sufficient soft reports for analysis and take decision. The decision-making effectiveness using the software is proposed in the second category of plan.

1. **Technological Design Plan**

The requirement analysis is the main step in any design process. The software architecture-based base design is the first step in technological design. The web sessions on Computers based on Windows platform like laptops and PCs are tracked. An MS Access database stores the data / information, collected Computers (front end part) and fed into databases (back end part).

**MIS based Design Plan**

The proposed software is mainly designed, developed and implemented in any institution, for needful effective decision making by the management users. As the parts of management function, the existing manual mode of planning, controlling and decision-making activities are taken into account and focused on computer-based manipulation during design. Then the management specific information system (IS) and its database schema is developed. The users of the system are defined, and their role on the system access and the restrictions on the user accesses are defined.

* 2. **Development and Implementation of the Software for Monitoring user behavior through session web mining (Standalone).**

This new automation system is installed into a Windows application. In proposed system, the websites visited and web sessions are tracked on their respective Windows OS based device. The reports is made available in the database server along with the duration of the web sessions.

* **Characteristics of the Software for Monitoring user behavior through session web mining (Standalone).**

It always runs in the background monitoring and tracking web activity.

1. The data collected can be viewed only by the admin, other cannot access it.
2. The software cannot be uninstalled without the admin login
3. The data collected can be viewed only by the admin , other cannot access it.
4. The software cannot be uninstalled without the admin login.
5. The database can be accessed offline.