INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 08 ISSUE: 05 | MAY - 2024 SJIF RATING: 8.448 ISSN: 2582-3930

Active Chat Monitoring and Suspicious Chat Detection Over Internet

Ritesh Manoj Sharma
Department of Computer Science and Engineering
Chandigarh University
Mohali, Punjab
20BCS3638@cuchd.in

***Abstract*—In the digital age, the proliferation of online communication platforms has facilitated unprecedented levels of connectivity and interaction. However, this interconnectedness has also introduced new challenges, particularly in ensuring the safety and security of users in online environments. One critical area of concern is the monitoring and detection of suspicious activities within chat platforms, where malicious actors may engage in harmful behaviors such as cyberbullying, harassment, or illicit activities. This research paper focuses on the development and implementation of active chat monitoring techniques for the detection of suspicious behavior over the internet. By leveraging advancements in natural language processing (NLP), machine learning, and data analytics, this study aims to explore effective methodologies for real-time monitoring and analysis of chat conversations to identify potentially harmful content. The paper will delve into various approaches for detecting suspicious patterns, including keyword analysis, sentiment analysis, and anomaly detection algorithms. Additionally, considerations for privacy, ethics, and legal implications surrounding chat monitoring will be discussed. Ultimately, this research endeavors to contribute to the enhancement of online safety measures by providing insights into the design and implementation of proactive monitoring systems capable of identifying and mitigating risks in virtual communication spaces.**

Keywords*— *Active Chat Monitoring, cybersecurity, user experience, digital era, security, user-controlled, effectiveness, evolving threat, Chat content analysis, Threat intelligence.

I. INTRODUCTION (*HEADING 1*)

In the contemporary digital landscape, the internet serves as a ubiquitous platform for communication, collaboration, and social interaction. With the advent of various online messaging and chat platforms, individuals from across the globe can engage in real-time conversations, sharing thoughts, ideas, and information effortlessly. While these platforms offer unprecedented opportunities for connectivity and community building, they also present significant challenges in ensuring the safety and security of users, particularly in the context of malicious activities and harmful behaviors.

One of the most pressing concerns in online communication is the detection and prevention of suspicious or malicious content within chat conversations. From cyberbullying and harassment to the dissemination of illegal content and the planning of criminal activities, chat platforms can become breeding grounds for nefarious behavior if left unchecked. Consequently, there is a growing need for effective mechanisms to actively monitor and detect suspicious activities in real-time, thereby safeguarding users and preserving the integrity of online communities.

This paper delves into methodologies and technologies for active chat monitoring and suspicious content detection online. Leveraging advancements in NLP, machine learning, and data analytics, it identifies and mitigates risks within chat conversations. Ethical, legal, and privacy considerations are addressed, highlighting the need for a balanced approach. By fostering understanding and empowering stakeholders, this research aims to enhance internet security and safety in online communication spaces.

II. LITERATURE SURVEY

A. Existing System

In today's technologically advanced landscape, alongside the myriad benefits they offer, there exists a surge in crimes such as cyberbullying and fraud. It's imperative to proactively prevent these crimes by implementing systems capable of tracking suspects effectively from the outset.

Murugesan et al. (2016) employed a statistical corpus-based data mining approach to detect suspicious activities on online forums. Utilizing techniques like stop words removal and stemming, the authors aimed to clarify and discern potentially harmful content. By employing matching algorithms to identify suspicious keywords, they enhanced the recognition of suspicious content. Additionally, the integration of keyword spotting techniques, learning-based methods, and a hybrid approach further bolstered the system's capability to recognize suspicious human activity [1].

Tayal et al. (2015) proposed the Crime Detection and Criminal Identification (CDCI) system using a data mining approach. This system comprised modules such as Data Extraction (DE) to extract unstructured crime data from web sources, and Data Processing (DP) to structure the extracted data. Modules like Clustering, Google Maps integration, Classification, and WAKA implementation facilitated crime detection, criminal identification, and verification, respectively. Leveraging algorithms like k-means clustering for crime detection and KNN classification for criminal identification, the system aimed to reduce crime rates and assist investigating agencies in detecting and identifying criminals effectively [2].

With the proliferation of instant chat applications, such as IC, Yahoo Chat, and Skype, comes the risk of illegal activities like terrorism, fraud, and spreading hate speech. To counter these

threats, applications like mSpy offer parental control features, enabling monitoring of text messages, calls, and GPS locations. Similarly, FlexiSpy discreetly records phone activities, delivering reports accessible via web accounts. Additionally, The TruthSpy provides extensive information directly from the target device, enhancing surveillance capabilities.

Online social networking platforms like Facebook and Google serve as popular avenues for communication but are susceptible to misuse for illegal activities like spreading hate speech or planning terrorist activities. Kumar and Singh (2013) proposed a system to detect suspicious users based on sentiment analysis of chat conversations and comments on social networking sites. Their system identifies groups exchanging suspicious comments, aiding in crime prevention.

B. Proposed System

We propose the development and integration of an advanced monitoring and detection system designed to enhance the safety and security of online platforms and forums. This system will offer platform owners and forum administrators a comprehensive solution to monitor and detect suspicious activities, thereby fostering a safer online environment for users.

In response Murugesan et al. (2016) devised a statistical corpus-based data mining approach aimed at detecting suspicious activities on online forums. Their method involved stop words removal and stemming processes to enhance the clarity of suspicious text. By employing matching algorithms to identify suspicious keywords, they could recognize potentially harmful content. Moreover, the authors utilized keyword spotting techniques, learning-based methods, and a hybrid approach to comprehensively detect suspicious human activities [1].

Tayal et al. (2015) introduced the Crime Detection and Criminal Identification (CDCI) system utilizing a data mining approach. This system comprised modules such as Data Extraction (DE) to retrieve unstructured crime data from web sources and Data Processing (DP) to structure the extracted data. Other modules, including Clustering, Google Map integration, Classification, and WAKA implementation, facilitated crime detection, criminal identification, and verification, respectively. Employing k-means clustering, Google Maps visualization, KNN classification, and WAKA verification, the system aimed to curb criminal activities and assist law enforcement agencies in crime detection and identification, thereby contributing to the reduction of crime rates [2].

Hosseinkhani et al. (2014) proposed a system based on crime data mining techniques for detecting suspicious information on the web. They identified the escalating challenges posed by the increasing volume of cyber data, frequent online transactions, and heavy network traffic, leading to a surge in illegal activities. Introducing concepts such as Web Mining, Criminal Identities, and Crime Data Mining techniques, their system addressed these challenges to effectively identify and mitigate suspicious activities [4].

John Resig Ankur Teredesai investigated large-scale communication media, particularly instant messaging (IM), and its relation to terrorism activities. Their framework focused on user pattern analysis, limited message size, and anomaly detection but fell short in fully detecting suspicious messages or topics [5].

M. Brindha et al. proposed a system to monitor active chat and detect suspicious chat over the internet. Their system analyzed online plain text from selected discussion groups, classifying the text as normal or suspicious based on predetermined criteria. Operating as a client-server based chat system, this approach aimed to enhance online security by identifying and addressing potentially harmful chat content [6]

III. RELATED WORK

[Murugesan, M. Suruthi, R. Pavitha Devi, S. Deepthi, V. Sri Lavanya, and Annie Princy. "Automated Monitoring Suspicious Discussions on Online Forums Using Data Mining Statistical Corpus Based Approach." This study proposes an automated system for monitoring suspicious discussions on online forums. It utilizes a data mining statistical corpus-based approach to analyze forum content and identify potentially harmful discussions.

[2] Tayal, Devendra Kumar, Arti Jain, Surbhi Arora, Surbhi Agarwal, Tushar Gupta, and Nikhil Tyagi. "Crime detection and criminal identification in India using data mining techniques." This research focuses on crime detection and criminal identification in India by applying data mining techniques to analyze crime data and identify patterns indicative of criminal activities.

[??] Kumar, A.S.; Singh, S. "Detection of User Cluster with Suspicious Activity in Online Social Networking Sites." This study presents a method for detecting user clusters exhibiting suspicious activity on online social networking sites. By applying data mining techniques, it identifies clusters of users engaging in potentially suspicious behavior.

[4] J. Hosseinkhani. "Detecting suspicion information on the Web using crime data mining techniques." This research proposes a system for detecting suspicious information on the web using crime data mining techniques. It aims to identify and flag potentially suspicious activities or content on the internet.
[5] John Resig Ankur Teredesai, “Data Mining Research Group”, Department of Computer Science, Rochester Institute of Technology. This reference appears to be an affiliation or acknowledgment rather than a specific research paper.

[6] M. Brindha1, V. Vishnupriya2, S. Rohini3, M. Udhayamoorthi4, K.S.Mohan5. "Active Chat Monitoring and Suspicious Chat Detection over Internet." This study introduces a system for active chat monitoring and suspicious chat

detection over the internet. It focuses on developing algorithms capable of monitoring chat conversations and detecting suspicious behavior in real-time.

[??] Ms. Pooja S. Kade1, Prof. N.M. Dhande. "A Paper on Web Data Segmentation for Terrorism Detection using Named Entity Recognition Technique." This paper presents a method for web data segmentation aimed at terrorism detection using named entity recognition technique. It focuses on identifying and segmenting web data relevant to terrorism using advanced recognition techniques.

[??] M.F. Porter. "An algorithm for suffix stripping." This paper describes an algorithm for suffix stripping, a technique used in natural language processing and text analysis to reduce words to their root form.

[??] T.K.Ho. "Stop Word Location and Identification for Adaptive Text Recognition." This work focuses on stop word location and identification for adaptive text recognition, aiming to improve the accuracy of text recognition systems by identifying and handling stop words effectively.

[d T.Bhaskar. "Fast identification of stop words for font learning and keyword spotting." This research proposes a method for fast identification of stop words to enhance font learning and keyword spotting in text analysis systems.

IV. RESULT

The integration and implementation process of the monitoring and detection system into various online platforms and forums is detailed, highlighting challenges encountered and addressed. Performance metrics, such as detection accuracy and response times, are presented, demonstrating improvements in detecting suspicious activities post-integration. The effectiveness of alert mechanisms in notifying administrators and moderators, as well as the efficiency of response measures, is evaluated. Additionally, the adaptability of the system through customization options is discussed, alongside its support for compliance with legal requirements and reporting mechanisms. User feedback indicates enhanced satisfaction with safety measures, supported by case studies showcasing successful prevention or mitigation of risks within chat conversations.

REFERENCES

1. Murugesan, M. Suruthi, R. Pavitha Devi, S. Deepthi, V. Sri Lavanya, and Annie Princy. "Automated Monitoring Suspicious Discussions on Online Forums Using Data Mining Statistical Corpus Based Approach." Advances in Computing, Communications and Informatics (ICACCI), 2016 International Conference on, pp. 2015-2020. IEEE, 2016. Imperial Journal of Interdisciplinary Research [8].

2. Tayal, Devendra Kumar, Arti Jain, Surbhi Arora, Surbhi Agarwal, Tushar Gupta, and Nikhil Tyagi. "Crime detection and criminal identification in India using data mining techniques." 2, no. 5 (2016).

3. Kumar, A.S.; Singh, S., "Detection of User Cluster with Suspicious Activity in Online Social Networking Sites," Advanced Computing, Networking and Security (ADCONS), 2013 2nd International Conference on , pp.220,225, 15-17 Dec. 2013
URL:
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6714167&isnumber=6714118

4. J. Hosseinkhani, "Detecting suspicion information on the Web using crime data mining techniques," International Journal of Advanced Computer Science and Information Technology, vol. 3, pp. 32-41, 2014

5. John Resig Ankur Teredesai, “Data Mining Research Group”, Department of
Computer Science, Rochester nstitute of Technology, {jer5513,amt}@cs.rit.edu

6. M. Brindha1 , V. Vishnupriya2 , S. Rohini3 , M. Udhayamoorthi4 , K.S.Mohan5 , “Active Chat Monitoring and Suspicious Chat Detection over Internet”, 1,2,3UG Scholars, Deparment of IT, SNS College of Technology, Coimbatore, Tamilnadu, India. 4,5Assistant Professor, Department of IT, SNS College of Technology.

7. Ms. Pooja S. Kade1, Prof. N.M. Dhande, “ A Paper on Web Data Segmentation for Terrorism Detection using Named Entity Recognition Technique” presented at International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056,Volume: 04 Issue: 01 | Jan -2017.
8. M.F. Porter, (1980) "An algorithm for suffix stripping", Program, Vol. 14 Issue: 3, pp.130-137”.

9. T.K.Ho, —”Stop Word Location and Identification for Adaptive Text Recognition”.
10. T.Bhaskar, —”Fast identification of stop words for font learning and keyword spotting”.