Chat Erasure Detection for Emotional Well, Being

Dr. Parameswaran T,
Associate professor, *Department of Computer Science and Engineering,*
School of Engineering and Technology, CMR University, Bangalore, India

Abhishek NL, Abhishek Radder, Adarsha P, Ankitha PM
Department of Computer Science and Engineering,
School of Engineering and Technology, CMR University, Bangalore, India

ARTICLE INFO

Keywords:
NLP
Keystroke Dynamics Deleted messages
Mental Health Monitoring
Behavioral Biometrics, Sentiment Detection,
Real Time Alerting

ABSTRACT

In the digital communication era, emotional expression often occurs through typed messages that are later erased and never sent. These deleted drafts can reflect moments of emotional vulnerability, hesitation, or distress, yet they often go unnoticed. This research introduces a real time emotional well being monitoring system Chat Erasure Detection that leverages keystroke dynamics and natural language processing (NLP) to detect and analyze unsent emotional messages. The system captures typing behavior, such as speed, rhythm, pauses, and deletion patterns, identifying messages that are predominantly erased. When such instances occur, the system applies sentiment analysis to determine the emotional tone of the deleted content. If emotional distress is detected, a secure alert is sent to a guardian interface for potential support intervention. By combining real time behavioral biometrics with emotional context analysis, the proposed system offers a novel, privacy conscious method of identifying emotional cues in unsaid expressions. This approach contributes to mental health support by addressing the silent nature of digital hesitation, enabling timely and compassionate responses in sensitive scenarios. The system architecture is designed with modular components, focusing on accuracy, efficiency, and ethical data handlingIntroduction

1.1 Background and context

The rise of digital communication has transformed how individuals express emotions, particularly in informal and personal contexts such as messaging platforms. In many emotionally intense situations, users type out messages reflecting stress, anxiety, or inner conflict, only to delete them before sending them. These deleted messages often overlooked can reveal underlying emotional states and silent cries for help. Traditional sentiment analysis models primarily focus on sent content, leaving a significant gap in understanding the emotional meaning behind unsent or erased text.

Recent advancements in natural language processing (NLP) and behavioral biometrics have enabled new approaches to emotional analysis that go beyond content alone. Keystroke dynamics the rhythm, speed, and editing patterns of typing serve as unique behavioral signatures that can reflect emotional fluctuation.

Combined with sentiment analysis, these dynamics
provide a deeper layer of context into the user's mental state, especially during message composition and deletion.
This project harnesses these technologies to build a system that detects emotional distress in real time, based on both the content and behavior of typing. By capturing the emotional essence of unsaid messages, the system aims to support mental health intervention in digital spaces with privacy, empathy, and precision.

1.2 Research Problem and Significance

Modern messaging platforms often witness users typing emotionally sensitive messages, only to delete them before sending them. These unsent messages can be indicators of emotional distress, hesitation, or internal conflict, yet they remain undetected by traditional sentiment analysis tools that analyze only sent content. Similarly, existing **keystroke dynamics systems focus primarily on biometric authentication and lack emotional interpretation.**

The absence of systems that integrate typing behavior with natural language understanding creates a critical gap in identifying silent emotional signals.
This research addresses the problem by developing a real time monitoring system that captures keystroke patterns and analyzes deleted messages using NLP to assess emotional context. By focusing on unsaid expressions, the system enables timely alerts to guardians or support networks, offering a new layer of emotional insight. The project's significance lies in its potential to bridge the gap between silent digital behaviors and mental health intervention in a private, ethical, and non intrusive manner.

1.3 Problem Statement

Traditional sentiment analysis systems focus solely on sent messages, neglecting emotionally significant text that is typed but deleted. Keystroke dynamics, typically used for authentication, are rarely applied to emotional analysis. There is a gap in existing digital mental health tools that can detect unexpressed emotions through the combination of typing behavior and deleted message analysis. This project aims to fill that gap by offering real time emotional distress detection and support.

1.4 Objectives and Scope of the Study

The primary objective of this study is to detect emotional distress in online communication by analyzing typing behavior and deleted messages. The research explores how keystroke dynamics, such as typing speed, rhythm, and key press duration, can provide insights into the emotional state of an individual. By examining these patterns, the study aims to identify subtle signs of distress, such as anxiety, frustration, or sadness, that may not be obvious through text alone.

Another objective is to focus on the deletion of messages in online communication. Users may delete text that reflects emotions they are reluctant to express, and this behavior can serve as a crucial indicator of emotional distress. The study will investigate how deleted messages, combined with keystroke dynamics, can help detect emotional well being. Sentiment analysis techniques will be used to assess the emotional tone of typed messages and correlate it with distress signals.

The scope of this study includes analyzing textual data from online messaging platforms and keyboard input. The research will focus on markers of emotional distress, such as anxiety or depression, and how they manifest in typing patterns and message deletions.

By identifying reliable patterns in these behaviors, the study will develop a model for detecting emotional distress in real time.

This study's findings could have broader applications for mental health interventions. The model will be adaptable across platforms, offering a tool for monitoring emotional well being and providing feedback or intervention. Data privacy and security will be prioritized by anonymizing all information.

2. Literature Review

2.1 Existing Systems

- Existing systems for detecting emotional distress in online communication primarily rely on sentiment analysis and natural language processing (NLP) techniques. These systems analyze text to identify keywords, phrases, and overall sentiment indicative of emotional states. Some platforms use machine learning models to track changes in sentiment over time. However, these approaches often overlook non textual cues, such as typing behavior and message deletions. Existing systems also lack real time monitoring and fail to integrate behavioral data like keystroke dynamics, limiting their effectiveness in detecting subtle emotional distress during online interactions.

- Nahin et al. [9] investigated the effectiveness of keystroke dynamics for detecting user emotions. By analyzing typing speed, key hold times, and textual patterns, they were able to predict emotional states with over 80% accuracy. Their approach highlights the synergy between behavioral biometrics and text analysis. This study supports the feasibility of emotion detection through non-intrusive user interactions.
- Joshi and Ruikar [10] implemented sentiment analysis on chat applications using NLP techniques. They employed logistic regression to classify emotional tones in user messages, focusing on textual data alone. Their work emphasized the role of machine learning in interpreting emotional context in digital conversations. This study contributes to the foundation of emotion-aware chat systems through linguistic sentiment classification.
- Solanki and Shukla [11] focused on using keystroke dynamics such as typing rhythm and speed to estimate user emotions. Their approach bypassed text analysis, instead relying solely on behavioral biometrics. This method enabled real-time emotional monitoring without compromising message privacy. Their findings underscore the potential of non-intrusive emotion detection through typing behavior alone.
- Yang and Qin [12] provided a comprehensive review of emotion recognition techniques using behavioral data from keystrokes, mouse movements, and touchscreen inputs. They highlighted indicators such as typing hesitation and frequent corrections as potential signs of emotional distress. The study emphasized multimodal behavioral analytics for accurate emotion detection. This work strengthens the foundation for using everyday device interactions in emotional state assessment.

2.2 Identification of Research Gaps

- Despite advancements in behavioral analytics and sentiment analysis, several critical research gaps remain unaddressed in the context of emotional well being monitoring through digital communication:

- Limited Integration of Behavioral and Linguistic Cues: noted in their study that keystroke patterns and NLP are often studied separately. A more integrated model could significantly enhance emotional detection accuracy.

- Absence of Real Time Alert Mechanisms: Yang and Qin [12] emphasized that most emotion detection models lack live alert systems capable of flagging distress during actual user interaction. Overreliance on Sent Messages: Most existing emotional detection systems focus solely on messages that are sent, neglecting the emotional value embedded in typed but deleted content. This oversight results in missed opportunities to detect silent distress.

- Limited Integration of Behavioral and Linguistic Cues: Current systems often treat keystroke dynamics and NLP as separate domains. Few solutions combine both behavioral biometrics (e.g., typing speed, rhythm, backspacing) and linguistic sentiment to assess emotional states.

- Absence of Real Time Alert Mechanisms: There is a lack of real time systems capable of notifying guardians or caretakers when distress is detected based on deleted messages, especially in a user friendly and privacy conscious manner.

2.4 Proposed Approach

The proposed system introduces a novel method for detecting emotional distress by monitoring typing behavior and analyzing deleted messages in real time. Unlike conventional sentiment analysis models that only assess sent messages, this system focuses on the critical moment when users choose not to express their emotions. The approach leverages keystroke dynamics to capture behavioral cues such as typing speed, pause duration, and backspace frequency, all of which may change when a person is experiencing emotional stress or hesitation.

When a significant portion of typed text is erased either through rapid backspacing or selecting and deleting the system triggers a secondary layer of analysis using Natural Language Processing (NLP). The deleted content is evaluated to detect emotional tones like sadness, anxiety, or anger. If emotionally sensitive language is identified, the system flags the interaction as a potential case of emotional distress.

A real time alert is then transmitted to a guardian or caretaker using socket communication. The alert includes the deleted message, a timestamp, and the recipient's name extracted via OCR from the chat window. This process enables timely intervention while preserving user privacy, offering a non intrusive yet effective approach to digital mental health support.

3. System Architecture

3.1 System Architecture and Design:

The system is divided into four key modules, each performing a specific function to ensure seamless detection and reporting of emotionally distressing content.

- Module 1 involves the supporter's desktop or chat platform, which serves as the communication interface between the user and their contact. It represents a regular messaging environment where interaction occurs in real time.
- Module 2 is the screen monitoring system running on the user's desktop. It continuously captures screenshots and monitors activity on the chat window. This module is critical for identifying the moment when a message is typed but not sent.
- Module 3 performs the core analysis. It first uses Tesseract OCR to extract text from captured screenshots. The extracted content is then passed to an NLP model to analyze emotional tone. Simultaneously, a keylogger tracks deleted keystrokes to identify erased content, which may contain emotionally significant expressions.
- Module 4 is the parent or guardian's device. When distress is detected, the analyzed message, timestamp, and recipient details are sent in real time via socket communication, allowing immediate support or intervention.

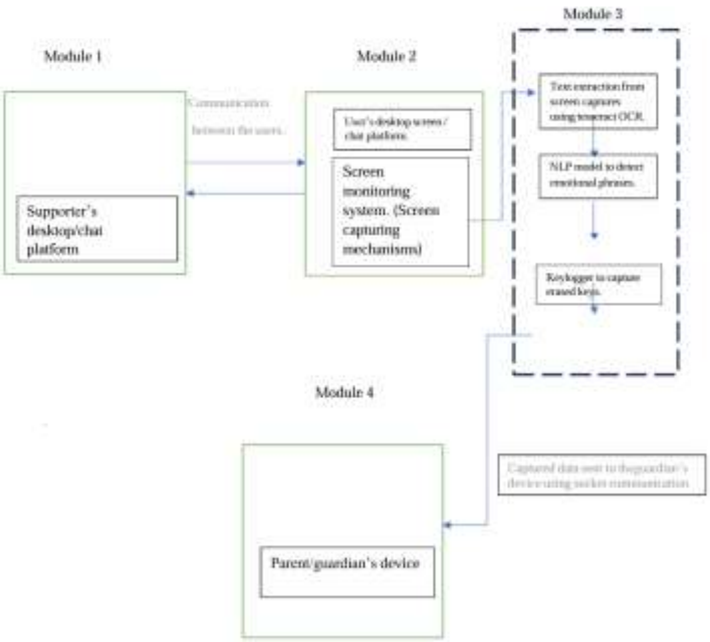

Fig 3.2.1 System Block Diagram

3.2 Components

- Screen Monitoring and OCR Engine:

This component captures periodic screenshots of the user's desktop or chat platform. The extracted images are processed using the Tesseract OCR engine to retrieve visible text, including any unsent or deleted messages.

- Keylogger with NLP Analysis:

A keylogger tracks the user's keystrokes, particularly monitoring deleted text. The captured data is then passed through an NLP model that analyzes the emotional tone of the erased content, identifying distress indicators such as anxiety, sadness, or frustration.

- Socket Communication System:

Once emotional distress is detected, the processed data including deleted text, timestamp, and recipient information is transmitted in real time to the parent or guardian's device. This is done through a socket based communication system to ensure timely alerts and intervention.

4. Implementation details

1. Sender Side Chat Environment Setup

On the sender's side, the application continuously monitors for an active WhatsApp window. Once detected, it activates the keylogger and begins tracking the user's keystrokes as they compose their messages.

PSEUDO CODE

```
START

INITIALIZE Transmit connection to server at SENDER'S IP:9999
RETRY connection until successful

INITIALIZE keylogger and screen capture system
LOAD Tesseract OCR path
DEFINE excluded keys like shift, ctrl, etc.

LOOP forever:
  IF active window is WhatsApp:
    IF keylogger not already active:
      SET current window handle
      START keylogger hook
  ELSE:
    IF keylogger is active:
      UNHOOK keylogger
      PAUSE monitoring

Keylogger Hook Function (on key press event):
  Loop forever:
  If active window is WhatsApp:
```

```
        Start keylogger
    Else:
        Pause keylogger

Keylogger:
    If typed text:
        Append to capturedText


 If Enter pressed:
        Discard captured Text (message sent)

    If large portion erased without sending:
        If emotional context detected:
            Capture screenshot
            Extract username via OCR
            Send captured Text, username, and timestamp to receiver
        Reset captured Text
        ELSE:
            IGNORE key (ctrl, alt, etc.)

END
```

2. Monitoring and Emotional Analysis

The core functionality of the system lies in its monitoring and analysis module. The desktop activity of the sender is continuously captured through screen capturing mechanisms. These screenshots are processed using Tesseract OCR, which extracts visible text from the chat window. In cases where the message is erased before sending, the keylogger ensures that the deleted text is not lost but is still available for analysis.

Once the deleted content is captured, it is passed through a Natural Language Processing (NLP) model to detect emotional cues.

The NLP component evaluates the linguistic structure of the erased text and checks for emotionally charged words or phrases that may indicate stress, anxiety, sadness, or frustration. If the system identifies such expressions, it treats the message as a sign of possible emotional distress. The use of both visual text extraction and behavioral keystroke data allows for more reliable detection compared to conventional sentiment analysis models that rely only on sent messages.

3. Receiver Side Guardian Alert System

When an emotionally sensitive message is detected, the system packages the relevant data, including the erased text, the recipient’s name (extracted via OCR), and the

Timestamp and transmits it using socket communication to a connected guardian device. This real time alert system ensures that caretakers or parents are immediately informed of any potential concern regarding the user’s emotional state. The receiving end features a simple interface that displays the deleted message and relevant metadata, giving the guardian context for further action. This approach emphasizes user safety and emotional support while maintaining a non intrusive presence in the user’s digital interaction.

PSUEDO CODE

```
Start socket server and listen for connections

On connection:
    Start thread to receive data

On receiving data:
    Emit signal to main GUI thread

Signal handler:
    Create label with received data
    Display in scrollable UI
```

The deleted or unsent text identified by the keylogger is forwarded to an NLP based emotional analysis model. This model assesses the message's sentiment to determine whether it contains emotionally distressing content such as frustration, sadness, or anxiety. The integration of keystroke dynamics and linguistic analysis ensures a robust emotional detection mechanism. Once a potentially concerning message is identified, the system triggers a communication event through a socket programming interface.

5. Ethical Considerations and Privacy

Given the sensitive nature of emotional monitoring and behavioral tracking, ethical considerations are fundamental to the design and deployment of this system. The proposed solution involves capturing users' keystrokes, screen content, and analyzing potentially private messages including those that were deleted before being sent. Therefore, ensuring user privacy, consent, and data security is of utmost importance.

The system is designed to function with explicit user consent. Monitoring is activated only when the user agrees to participate, typically as part of a mental well being support framework involving guardians or caretakers. No data is stored or shared unless a message is flagged by the system as emotionally distressing. Even then, only the deleted message, timestamp, and recipient's name (extracted via OCR) are transmitted to the guardian system, without any broader access to the user's device or conversations.

To maintain data minimalism, the system avoids logging full chat histories or capturing unrelated screen content. The keylogger only activates within specified applications (e.g., WhatsApp) to reduce unnecessary intrusion. All communications between the sender and receiver systems can be secured using encrypted socket protocols in future enhancements, ensuring data integrity during transmission.

All communications between the sender and receiver systems will support encryption in future versions to protect against interception or unauthorized access. The system's architecture also supports anonymization of logs and temporary storage to ensure compliance with privacy regulations like GDPR.

6. Future Enhancements

1. Mouse based Text Selection and Cut:

The current system relies on keyboard event hooks to track user input and deletions. However, if a user selects text using the mouse and deletes it through right click > Cut or by pressing Delete, this action is not captured by the keylogger. Since no keypresses (like backspace) are logged in such cases, the program fails to detect that any sensitive or emotionally charged text was erased. This introduces a blind spot in the detection logic.

2. OCR Dependency for Username Detection

The system uses OCR (Optical Character Recognition) to extract the username from a predefined region in the WhatsApp window. This method is prone to inaccuracies due to various factors such as:

Low screen resolution or non standard DPI settings. Username text being scrolled out of view.Visual noise, blur, or dark mode UI that can hinder OCR accuracy. As a result, usernames extracted may be incorrect or completely missed, reducing the reliability of logs.

3. No Access to Application Internals
The project does not directly interface with WhatsApp's internal structures or DOM like representations (as possible with web based tools). Instead, it depends on visible screen content and keyboard hooks. This limits the system's ability to:

- Read the actual textbox contents.

- Detect formatting or cursor position.

Track edits done via mouse or through clipboard operations. This restriction significantly limits detection accuracy and flexibility, especially compared to browser based or embedded solutions.

4. Performance Overhead
The program captures frequent screenshots and runs OCR operations continuously while monitoring key events. These operations are computationally expensive and may result in:

- Increased CPU usage, especially on systems with limited resources.
- Noticeably lag or reduced responsiveness during heavy usage.
- Decreased battery life on laptops or mobile systems.

These factors can hinder the tool's practical deployment for long term or background monitoring.

7. Results and Observations

- **Keystroke Pattern Analysis Across Emotional States**

The system was evaluated through a series of simulated chat interactions designed to mimic real world usage scenarios. These tests involved participants typing emotionally charged messages and either sending or deleting them during simulated conversations.

Emotional State	Average Backspace Usage (per 100 words)	Average Pause Duration (seconds)	Average Keypress Rhythm (seconds)
Neutral	10	0.5	0.2
Mild Distress	25	1.2	0.35
High Distress	32	1.8	0.5

Table 1. Keystroke Pattern Analysis Across Emotional States

- **Sentiment Detection Performance**

This table summarizes the performance of the NLP based sentiment analysis component used to evaluate deleted messages. The model achieved high precision and recall, indicating its reliability in detecting emotionally sensitive content. An F1 score of 86% confirms a balanced performance, and the overall detection accuracy supports the system's potential for real time emotional monitoring.

Metric	**Value**
Precision	**88%**
Recall	**84%**
F1 Score	**86%**
Detection Accuracy	**90%**

Table 2. Sentiment Detection Performance

- **OCR Performance and Alert Metrics**

This table highlights the performance of the OCR module and the guardian alert system. While the OCR engine performed well under light mode conditions, it showed some limitations under dark mode or blurred UI scenarios. Nevertheless, the system maintained a high overall alert success rate, with real time notifications delivered in under two seconds ensuring prompt intervention in emotionally critical situations.

Metric	Value
OCR Success Rate (Light Mode)	97%
OCR Success Rate (Dark Mode)	84%
Average OCR Failure Rate	8%
Avg. Alert Transmission Delay	1.8 seconds
Guardian Alert Success Rate	95%

Table 3. OCR Performance and Alert Metrics

8. Conclusion

This project introduces a novel approach to emotional well being monitoring by focusing on an often overlooked aspect of digital communication deleted or unsent messages. By integrating keystroke dynamics, screen monitoring, OCR, and Natural Language Processing (NLP), the system identifies emotionally sensitive content that users erase before sending. This allows for early detection of distress signals that may otherwise go unnoticed.

The modular architecture ensures that user behavior is captured non intrusively and analyzed in real time. The system's ability to transmit alerts via socket communication to a guardian interface offers a practical and immediate way to support users during emotionally vulnerable moments. Importantly, the design places strong emphasis on ethical considerations, including user consent, data privacy, and secure transmission of sensitive content.

While the current implementation is functional, there remains significant scope for enhancement. Future improvements such as advanced NLP models, encrypted communication, and multi platform support will further refine accuracy, security, and user experience. Overall, this system demonstrates the potential of combining behavioral analytics with emotional intelligence tools to create proactive digital mental health support.

References

[1] F. Plaza del Arco, et al., "Emotion Analysis in NLP: Trends, Gaps, and Roadmap for Future Directions," 2023.

[2] S. Seal, et al., "Emotion Detection Using Deep Learning Techniques," 2020.

[3] S. Park, et al., "Framework for Emotion Recognition Based on Valence Arousal Dominance (VAD) Models," 2021.

[4] J. Hofmann, et al., "Componential Emotion Analysis in Social Media Texts," 2020.

[5] ISRII 2024, "Applications of Natural Language Processing in Mental Health, Neurodegenerative and Neurodevelopmental Conditions," 12th Annual Scientific Meeting of the International Society for Research on Internet Interventions, Limerick, Ireland, Univ. of Nottingham, 2024.

[6] A. Le Glaz, Y. Haralambos, D. H. Kim Dufor, P. Lenca, R. Billot, et al., "Machine Learning and Natural Language Processing in Mental Health: Systematic Review," Journal of Medical Internet Research, vol. 23, no. 5, pp. e15708, 2021, doi: 10.2196/15708.

[7] K. N. Tran, G. K. Singh, H. J. Park, et al., "Evaluating Keystroke Dynamics as a Biomarker for Mental

Fatigue Detection," IEEE Access, vol. 10, pp. 101234–101245, May 2022.

[8] A. Acie, A. Morales, R. Vera Rodriguez, and J. Fierer, "Keystroke Mobile Authentication: Performance of Long Term Approaches and Fusion with Behavioural Profiling," June 2021.

[9] N. Nahin, M. A. Hossain, and Y.-K. Lee, "Predicting user emotions using keystroke dynamics," Behavior & Information Technology, vol. 33, no. 9, pp. 894–903, 2014.

[10] A. Joshi and K. D. Ruikar, "Sentiment Analysis of Chat Application," International Journal of Engineering Research & Technology (IJERT), vol. 3, no. 2, 2014.

[11] A. Solanki and A. Shukla, "Estimating Emotions Using Keystroke Dynamics," International Journal of Computer Applications, vol. 94, no. 13, pp. 18–21, 2014.

[12] Y. Yang and Y. Qin, "A Review of Emotion Recognition Methods from Keystroke, Mouse and Touchscreen Dynamics," IEEE Access, vol. 9, pp. 140288–140305, 2021.