Twitter User Sentiments Analysis: Health System Cyberattacks Case Study

Muhammad Abusaqer
Department of Math and Computer Science
Minot State University

M. Benaoumeur Senouci
The Faculty of Engineering, SDU Mechatronics (CIM)
University of Southern Denmark

Kenneth Magel
Department of Computer Science
North Dakota State University

Abstract

Social media, notably Twitter, plays a crucial role in how the public interacts with and responds to significant events, including emergent situations where it's often used by officials for public communication. Despite sentiment analysis being extensively explored across various domains, the specific area of public sentiment during cybersecurity breaches is not well-documented. Addressing this research gap, this study delves into the public's sentiment in reaction to the cyberattacks on Universal Health Services (UHS) in September 2020. Over a span of five weeks, a robust dataset of tweets was meticulously compiled using targeted keywords and hashtags relevant to the cyberattacks. The research team applied advanced deep learning techniques to classify sentiments, achieving noteworthy accuracies. An exploratory data analysis was undertaken to trace the trajectory of sentiment over time, providing a nuanced understanding of the public's evolving reactions.

Introduction & Research Questions

- The COVID-19 pandemic amplified vulnerabilities in healthcare cybersecurity, leading to a 273% increase in cyberattacks in 2020.
- Many cyberattacks remain unreported, hinting at an even more extensive threat landscape.
- Existing research on sentiment analysis via social media is extensive, yet the niche of cybersecurity breaches in healthcare settings, particularly during pandemics, is less understood.
- This study bridges this gap by analyzing the sentiment of tweets related to the cyberattacks9 on Universal Health Services (UHS).
- **Key research questions guiding this analysis:**
  i. How does public sentiment evolve following a cyberattack on a healthcare provider?
  ii. Are public reactions to the cyberattacks generally positive or negative over time?
  iii. What is the duration of public engagement with the cyberattack incident?
  iv. Which sentiments are more frequently expressed by the public in the wake of cyberattacks: positive or negative?

UHS Ransom Cyberattacks

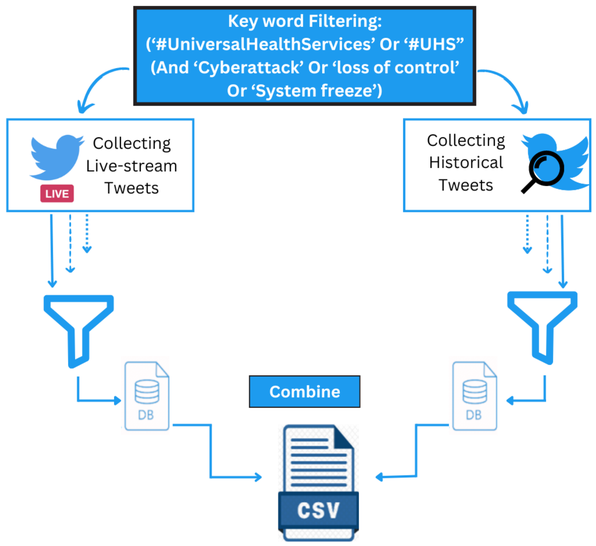
- UHS, a major healthcare provider in the US and UK, suffered a cyberattack in September 2020.
- Operational disruptions in over 400 facilities; critical processes reverted to pen and paper.
- The incident resulted in a substantial financial loss of $67 million pre-tax.
- These events stress the significance of public sentiment in shaping patient trust and organizational reputation.

References:

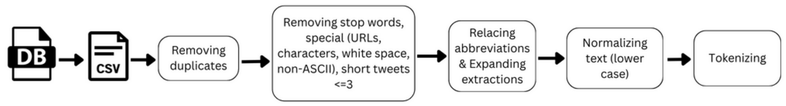
- Healthcare IT News: Universal Health Services faces $67 million loss after cyberattack
- UHS Official Statement: Statement from Universal Health Services

Experiment: Data collection

- In this study, researchers collected tweets about the UHS ransom cyberattack using keywords and hashtags over a period of five weeks.
- To collect the tweets, the researchers used two techniques:
  i. The Twitter live stream API.
  ii. Data scraping.
- In total, the researchers were able to collect approximately 1 million tweets related to the UHS ransom cyberattack.
- The use of these techniques allowed:
  i. Collection of a large amount of data
  ii. Enabling the analysis of the sentiment of a diverse range of Twitter users.



Experiment: Dataset cleaning



Experiment: Measuring Sentiment

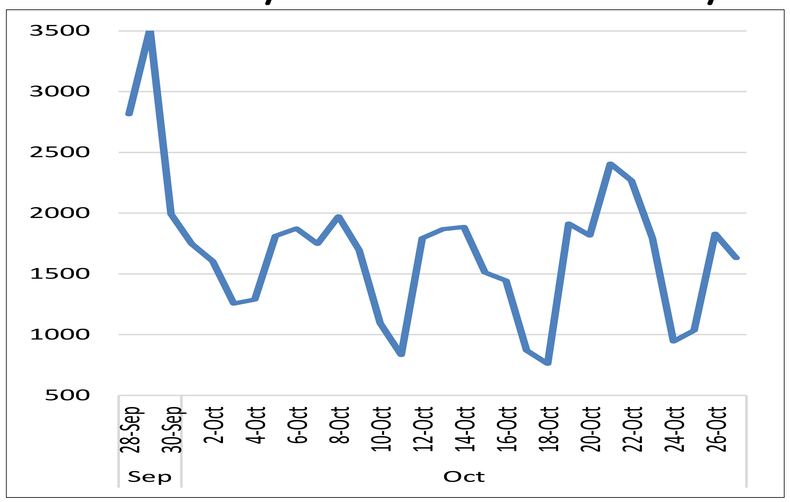
Sentiment Measurement Approaches

Started with TextBlob for preliminary sentiment analysis.
Employed LSTM (Long Short-Term Memory), a type of RNN, for modeling sequences and context in text data.
Utilized RoBERTa, a transformer model, for its advanced pretraining approach and effectiveness in NLP tasks.

Key Insights

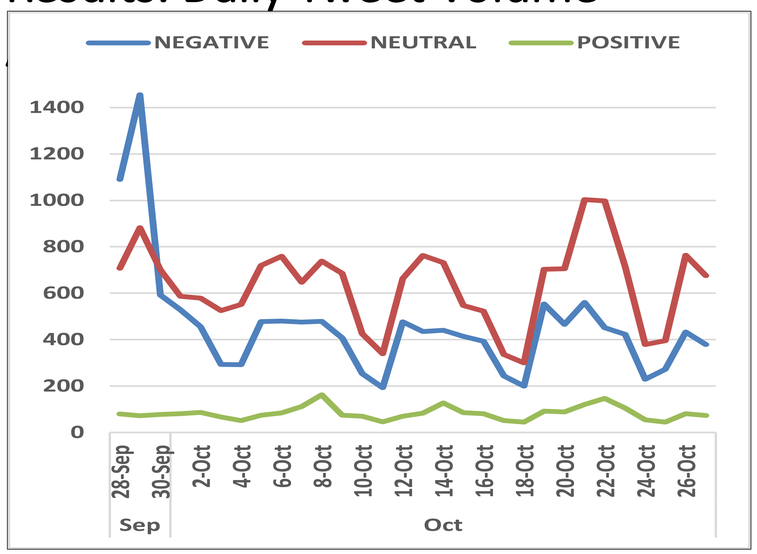
RoBERTa demonstrated superior performance in our tests, making it the preferred model for in-depth sentiment analysis.

Results: Daily Tweet Volume Analysis



This line graph depicts the daily volume of tweets related to the UHS cyberattack over a specific timeframe. Noticeable fluctuations correspond to key events during the incident, with peaks indicating heightened public response. The data illustrates how public engagement varies in reaction to developments in the cyberattack narrative.

Results: Daily Tweet Volume



This line graph delineates the trends in public sentiment on Twitter regarding the UHS cyberattack, categorized as negative, neutral, and positive over the observed period. The graph illustrates the dynamic nature of public sentiment, with the blue line indicating the frequency of negative sentiment, the red line showing neutral sentiment, and the green line for positive sentiment. Notable is the initial spike in negative sentiment, which gradually decreases, whereas neutral and positive sentiments exhibit variability, possibly reflecting the evolving public perception as the event unfolded.