engineering proceedings

MDPI

Proceeding Paper

Classification Crisis Communication: Semiotic Approach with Latent Semantic Analysis †

Richard G. Mayopu , Long-Sheng Chen and Venkateswarlu Nalluri *

Department of Information Management, Chaoyang University of Technology, Taichung 413310, Taiwan; s11014905@gm.cyut.edu.tw (R.G.M.); lschen@cyut.edu.tw (L.-S.C.)
* Correspondence: nallurivenkey7@gmail.com

† Presented at the 3rd IEEE International Conference on Electronic Communications, Internet of Things and Big Data Conference 2023, Taichung, Taiwan, 14–16 April 2023.

Abstract: Previous crisis communication research has been based on qualitative methods such as interviews or questionnaires, which require considerable manpower, material resources, and time to focus on specific topics. The current situation needs to be reflected timelier. With the rise of social communities, community users' comments have gradually become an important reference for other community members. Twitter is one of the most popular social media in the world. During the COVID-19 pandemic, people were restricted by rules and government policies, such as wearing masks, maintaining social distancing, and avoiding crowding. This led people to spend time on devices. By using devices, most people are involved in social media activities. This study aims to discover the awareness Indonesians display in the text they upload to Twitter. Using the Twitter crawling technique, we collected data. We also analyzed the text with text mining techniques and latent semantic analysis (LSA) with semiotic methods. The crisis communication was classified, and the definition of crisis terminology was improved in social media.

Keywords: text mining; latent semantic analysis; crisis communication; social media; semiotic

Citation: Mayopu, R.G.; Chen, L.-S.; Nalluri, V. Classification Crisis Communication: Semiotic Approach with Latent Semantic Analysis. *Eng. Proc.* **2023**, *38*, 9. https://doi.org/10.3390/engproc2023038009

Academic Editors: Teen-Hang Meen, Hsin-Hung Lin and Cheng-Fu Yang

Published: 19 June 2023

1. Introduction

Crisis communication has become very important in the last three years, during which the world struggled with the COVID-19 pandemic since the first time this virus appeared and affected society, the economy, culture, politics, and technology. During the pandemic, people often used technology to communicate with families and friends and work from home.

Social media is one of the most used technologies in daily life in modern society. Workers use it for their jobs, students for studying, and parents for parenting. Technologies support many activities and change people's behavior. At the same time, people offer advice to the developer to improve the technology. Social media has rapidly increased the interaction of humans and the awareness of understanding how to use technology wisely. People can use social media for e-commerce to promote their products and share content. Social media helps a community build close relationships. The interaction among the members of the community is an interesting phenomenon to observe. This interaction occasionally becomes a new approach or a new habit in society, which is named the "new normal".

The new normal is the concept provided by scientists to describe social activity during the pandemic. People had to do social distancing and wear masks often. In this situation, billions of people went online where they could interact without facing their friends or family directly, to feel safer and healthier. The new normal has been created to describe the world on the "crisis road". Crisis during the pandemic became a global issue. This is the challenge in the digital and technology era.

A crisis must be communicated to the public with a strategy. This can be conducted by using technology as the channel. In conventional communication theory, the channel can be the newspaper, television, or radio. However, in the modern era, the channel can be the new social media such as Twitter, Facebook, YouTube, and many others. Twitter is one of the most popular social media that was used during the pandemic. Crisis communication is the term for messages delivered from the communicator to the communicant (recipient), considering crisis requirements such as fast response communication, clear communication, and informative fact. The crisis messages contain meaning based on the understanding of the text and context. In this view, the semiotic approach is the best approach to finding the meaning behind the text related to context. Previous research in crisis communication has always used qualitative approaches [1–3] with semiotics. Thus, we aim to provide a combination of semiotics and latent semantic analysis as a new approach in this field of research.

Our research question is "how to classify crisis communication using a semiotic approach with semantic analysis". This research was focused on crisis communication on a large scale not limited to corporations, institutions, or organizations. We attempted to find out the awareness that Indonesians display in the text they upload to Twitter. Using the Twitter crawling technique, we collected data and analyzed the text with latent semantic analysis (LSA) combined with semiotic methods. The results provide classified crisis communication to redefine crisis terminology in social media.

2. Literature Review

2.1. Crisis Communication

Scholars have defined crisis communication for decades and discussed it through an interdisciplinary approach. Several have described crisis communications from a management perspective, while others have explained it through a cultural approach. Crisis communication is also dealt with in communication science. Based on this background, we reviewed crisis communication from the communication and technology perspective. This approach is expected to enrich the understanding of crisis communication for the development of knowledge. We classified crises into small and large scopes depending on the crisis. Most crises occur in the institution or corporation [4] and have local impacts. Nevertheless, a crisis with a serious impact, such as the global crisis of COVID-19, is a large-scope crisis, and the strategy to deal with it is different.

2.2. Social Media

Social media is a digital platform to enable users to communicate or share content such as texts, photos, videos, and files [5]. Earlier, social media was applied in web-based technology; however, nowadays, most users use social media on mobile apps [6] such as smartphones or tablets.

2.3. Text Mining

In the past few decades, text mining has become popular in computer science as a part of deep learning focusing on how to detect or identify text. Text mining is used to find the meaning of the text by extracting meaningful information from it. According to experts, text mining is used to find new information about human character [7]. Moreover, text mining provides flexible analysis depending on the research objects. In this research, we investigated crisis communication.

2.4. Latent Semantic Analysis

Many published works utilized LSA to identify relevant semantic spaces. For example, Das and Sultana [8] performed a semantic analysis of the Bengali language to propose a predictive model [8]. Hsiao and Hsiao [9] employed LSA for online hotel reviews to understand how satisfied or dissatisfied guests felt. We employed LSA to extract important concepts from a huge number of documents.

3. Methodology

We used a semiotic approach combined with latent semantic analysis (LSA). Previous research usually used qualitative research to find models or strategies for crisis communication. We mixed semiotics in social research with latent semantic analysis for information technology research.

3.1. Pre-Processing

We pre-processed [10] collected data using Python for tokenization and lemmatization and obtained the TDM using MATLAB.

3.2. Natural Language Processing

Natural language processing (NLP) usually considers the interesting applications of text mining and is able to synthesize information from text. Messages on Twitter have an intention for objective communication. However, if a complicated sentence appears, information can be extracted through NLP [11]. The implementation procedure is described as follows.

3.2.1. Step 1: Tokenization

Textual tokenization uses the Natural Language Toolkit (NLTK) of the Python language.

3.2.2. Step 2. Data Cleaning

We deleted stop words such as "the", "and", and other low-importance words in this step.

3.2.3. Step 3. Lemmatization

This step is to reduce complex forms of a single word to its most basic form, e.g., "ate" to "eat".

3.2.4. Step 4. Counting Word Frequency

We counted word frequency and deleted words with a frequency of less than 5.

3.2.5. Step 5. Constructing a Term–Document Matrix (TDM)

In this step, we used TF–IDF (term frequency–inverse document frequency) in Equation (1) to create a term–document matrix (TDM) for further analysis.

$$\text{TF–IDF} = TF(t_i\ , d_i) \times \log\left(\frac{N}{N(t_i)}\right) \tag{1}$$

3.3. Latent Semantic Analysis

We performed singular value decomposition (SVD) on TDM to investigate the relationship between words and their concepts.

SVD

The process of SVD is shown in Figure 1. The SVD function results in three matrices "terms matrix $U_{t\times r}$", "orthogonal matrix $S_{r\times r}$", and "document matrix $V^T_{r\times n}$", where t refers to the number of words, n is the document term, and r is the number of concepts in the semantic space.

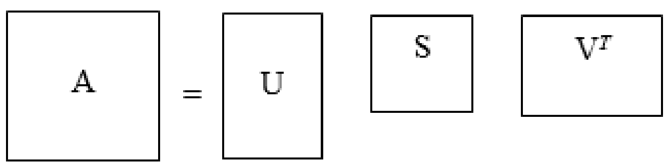

Figure 1. Singular Value Decomposition.

Even after running SVD, it still contains a lot of unimportant information, so it is necessary to reduce the dimensional space. To affect the original characteristics, the feature value k must be chosen. In this study, the scree test was used to determine the k value. In Figure 2, we selected k concepts, so we took the S_k matrix to reduce the dimensionality.

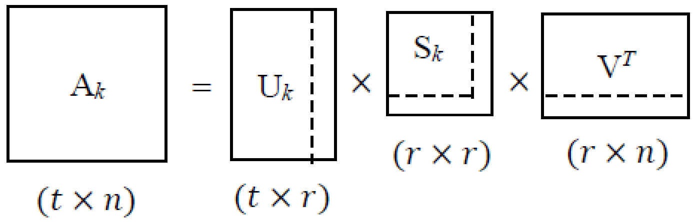

Figure 2. SVD after dimension reduction.

3.4. Ortoghonal Rotation of Axes

The concept load L_T was then calculated by multiplying the dimensionally delimited concept U_k with the concept S_k (Equation (2)) to obtain the concept load L_T. Each feature word was ranked according to the load, and the word concept was named.

$$TL_T = U_k \times S_k \tag{2}$$

4. Results and Discussion

4.1. Data Processing

Crisis communication has increased for decades, especially in the last three years during the COVID-19 pandemic, going beyond conventional crisis communication. Traditional crisis communication usually uses conventional media such as television, radio, newspaper, and word of mouth. Nowadays, crisis communication uses the internet. There are many platforms to deliver and receive messages that contain crisis communication. Moreover, it can be an interaction medium between the communicator and the communicant.

The data has been investigated with SVD and charted to show the information. We selected the k value from the top five (k = 5) as follows. It is possible to ignore the variation gap after the decision point since it decreases after the plot slope, as shown in Figure 3.

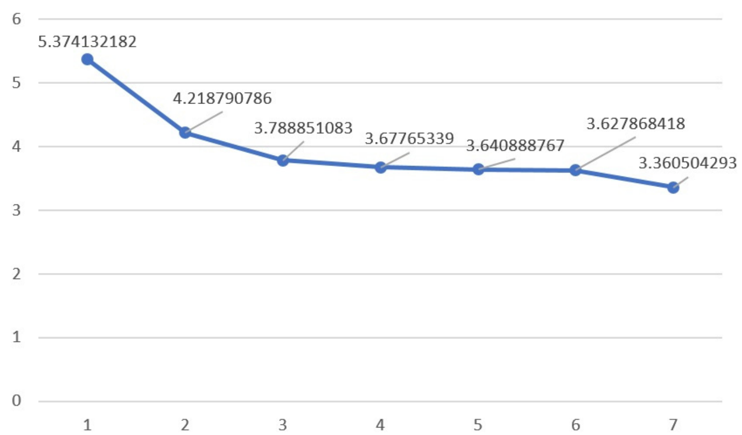

Figure 3. Scree plot.

4.2. LSA and Concepts

In the LSA process, we obtained the five concepts using the semiotic approach. The data are shown in Table 1.

Table 1. Keywords and Loadings.

1st Concept		2nd Concept		3rd Concept	
Keyword	**Loading**	**Keyword**	**Loading**	**Keyword**	**Loading**
Abramovic	1.7554	Crisis	0.2504	Injury	0.0863
Army	0.3468	Destroyed	0.1601	Damage	0.0824
Battle	0.2258	Pandemic	0.0773	Fighting	0.0750
Lockdown	0.0645	Death	0.0637	Military	0.0346
4th Concept		**5th Concept**			
Keyword	**Loading**	**Keyword**	**Loading**		
Civilian	0.0264	Rescuer	0.0223		
Killing	0.0248	Volunteer	0.0212		
Suffered	0.0242	Community	0.0199		
Violence	0.0222	Protester	0.0023		

Based on Table 1, we formulated the concepts of the LSA experiments with five concepts. The top 50 words from each loading were chosen to name the concept after ranking the defining words based on their loadings. Furthermore, we found the fixed concepts to be analyzed further. Table 2 shows the concepts for crisis communication.

Table 2. Concepts of crisis communication using latent semantic analysis.

No.	Concept	Representative Terms
1st	The Antagonist	Abramovic, army, battle, lockdown
2nd	The Phenomenon	Crisis, destroyed, pandemic, death
3rd	The Destructions	Injury, damage, fighting, military
4th	The Victims	Civilian, killing, suffered, violence
5th	The Protagonists	Rescuer, volunteer, community, protester

4.3. Semiotics in Crisis Communication

Roland Barthes' model is used in this research to find the signifier in crisis communication. Figure 4 shows the model to explain the concept.

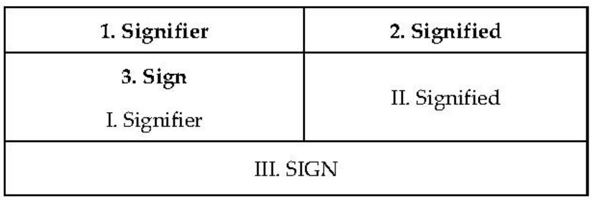

Figure 4. Barthes' model of semiotics.

Figure 4 shows the base model of the semiotics. We put the concept of LSA into the diagram for analysis. The dataset is about the crisis caused by the invasion of Russia into Ukraine. Therefore, the context of the data is about invasion. The context must be clear to analyze the text, and we found several concepts in this research.

In this research, the signifier was the people who provided the text to the mass media, the victims, the soldiers, and the politicians regarding the invasion. The signified is the text that the signifier has produced. Based on the text we found, the sign was used for analyzing the text. In the model, denotation and connotation were defined. Denotation is the system of level signification in Barthes' semiology at the first level, while connotation is at the second level. It means that once readers see the sign (text or image), they consider using denotation to analyze the sign. Then, they find the denotation from the sign.

In this research, denotation has a more closed meaning. With the theory of signs, we understand that literal meaning is natural. According to this theory, the meaning of signs is expanded with meaning that takes place in two stages as shown in Figure 5.

<table>
<tr><td>1. Signifier</td><td>2. Signified</td><td rowspan="2">Denotation: Invasion, civil war, economic restriction</td></tr>
<tr><td>3. Sign
I. Signifier: Mass media, Politicians, etc.</td><td>II. Signified: text</td></tr>
<tr><td colspan="2">III. SIGN</td><td>Connotation: Conspiration, Intelligent movement, etc.</td></tr>
</table>

Figure 5. Implementation of Barthes' models of semiotics.

5. Conclusions and Future Directions

We found a classification of crisis communication in a large scope. Previous research found a small scope of crisis communication, while a large scope was found on Twitter. We found several concepts from LSA that divided texts into small and large scopes. Semiotic analysis was conducted to provide a new approach as a research method. Barthes' model was implemented to analyze the concepts, and semiotics was used in this research. For future research, researchers can consider the semiotic approach in text mining to develop the utilization of semiotic models and provide new models of the denotation and connotation concepts.

Author Contributions: Conceptualization, R.G.M. and V.N.; methodology, R.G.M. and V.N.; software, R.G.M. and V.N.; validation, R.G.M.; formal analysis, V.N. and L.-S.C.; writing—original draft preparation, R.G.M.; writing—review and editing, V.N. and L.-S.C.; visualization, R.G.M.; supervision, L.-S.C.; project administration, L.-S.C.; funding acquisition, L.-S.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Science and Technology Council, Taiwan (Grant No. MOST 111-2410-H-324-006).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The secondary data that were used for analysis in this study are available from the corresponding author upon request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Arendt, C.; LaFleche, M.; Limperopulos, M.A. A qualitative meta-analysis of apologia, image repair, and crisis communication: Implications for theory and practice. *Public Relat. Rev.* **2017**, *43*, 517–526. [CrossRef]
2. Coady, A.; Lainchbury, K.; Godard, R.; Holtzman, S. What twitter can tell us about user experiences of crisis text lines: A qualitative study. *Internet Interv.* **2022**, *28*, 100526. [CrossRef] [PubMed]
3. Lu, X.; Jin, Y. Information vetting as a key component in social-mediated crisis communication: An exploratory study to examine the initial conceptualization. *Public Relat. Rev.* **2020**, *46*, 101891. [CrossRef]
4. Huynh-Cam, T.T.; Nalluri, V.; Chen, L.S.; Yang, Y.Y. IS-DT: A New Feature Selection Method for Determining the Important Features in Programmatic Buying. *Big Data Cogn. Comput.* **2022**, *6*, 118. [CrossRef]
5. Wang, H.; Xiong, L. Understanding the antecedents of mobile social media acceptance during the public health crisis maintenance stage. *Int. J. Disaster Risk Reduct.* **2022**, *74*, 102925. [CrossRef]
6. Morgan, A.; Wilk, V. Social media users' crisis response: A lexical exploration of social media content in an international sport crisis. *Public Relat. Rev.* **2021**, *47*, 102057. [CrossRef]

7. Jung, H.; Lee, B.G. Research trends in text mining: Semantic network and main path analysis of selected journals. *Expert Syst. Appl.* **2020**, *162*, 113851. [CrossRef]
8. Das, P.; Sultana, N. Sentiment Analysis on Comments in Bengali Language Using Text Mining & Machine Learning Approach. In Proceedings of the IEEE 7th International conference for Convergence in Technology, Mumbai, India, 7–9 April 2022; pp. 1–4.
9. Hsiao, Y.-H.; Hsiao, Y.-T. Online review analytics for hotel quality at macro and micro levels. *Ind. Manag. Data Syst.* **2020**, *121*, 268–289. [CrossRef]
10. Öztürk, N.; Ayvaz, S. Sentiment analysis on Twitter: A text mining approach to the Syrian refugee crisis. *Telemat. Inform.* **2018**, *35*, 136–147. [CrossRef]
11. Martinelli, D.D. Evolution of Alzheimer's disease research from a health-tech perspective: Insights from text mining. *Int. J. Inf. Manag. Data Insights* **2022**, *2*, 100089. [CrossRef]