Original Paper

An Environmental Uncertainty Perception Framework for Misinformation Detection and Spread Prediction in the COVID-19 Pandemic: Artificial Intelligence Approach

Jiahui Lu[1,2], PhD; Huibin Zhang[2], BE; Yi Xiao[2], PhD; Yingyu Wang[2], BE

[1]State Key Laboratory of Communication Content Cognition, People's Daily Online, Beijing, China
[2]School of New Media and Communication, Tianjin University, Tianjin, China

Corresponding Author:
Huibin Zhang, BE
School of New Media and Communication
Tianjin University
Number 92, Weijin Road
Tianjin, 300072
China
Phone: 86 15135154977
Email: zhanghb@tju.edu.cn

Abstract

Background: Amidst the COVID-19 pandemic, misinformation on social media has posed significant threats to public health. Detecting and predicting the spread of misinformation are crucial for mitigating its adverse effects. However, prevailing frameworks for these tasks have predominantly focused on post-level signals of misinformation, neglecting features of the broader information environment where misinformation originates and proliferates.

Objective: This study aims to create a novel framework that integrates the uncertainty of the information environment into misinformation features, with the goal of enhancing the model's accuracy in tasks such as misinformation detection and predicting the scale of dissemination. The objective is to provide better support for online governance efforts during health crises.

Methods: In this study, we embraced uncertainty features within the information environment and introduced a novel Environmental Uncertainty Perception (EUP) framework for the detection of misinformation and the prediction of its spread on social media. The framework encompasses uncertainty at 4 scales of the information environment: physical environment, macro-media environment, micro-communicative environment, and message framing. We assessed the effectiveness of the EUP using real-world COVID-19 misinformation data sets.

Results: The experimental results demonstrated that the EUP alone achieved notably good performance, with detection accuracy at 0.753 and prediction accuracy at 0.71. These results were comparable to state-of-the-art baseline models such as bidirectional long short-term memory (BiLSTM; detection accuracy 0.733 and prediction accuracy 0.707) and bidirectional encoder representations from transformers (BERT; detection accuracy 0.755 and prediction accuracy 0.728). Additionally, when the baseline models collaborated with the EUP, they exhibited improved accuracy by an average of 1.98% for the misinformation detection and 2.4% for spread-prediction tasks. On unbalanced data sets, the EUP yielded relative improvements of 21.5% and 5.7% in macro-F1-score and area under the curve, respectively.

Conclusions: This study makes a significant contribution to the literature by recognizing uncertainty features within information environments as a crucial factor for improving misinformation detection and spread-prediction algorithms during the pandemic. The research elaborates on the complexities of uncertain information environments for misinformation across 4 distinct scales, including the physical environment, macro-media environment, micro-communicative environment, and message framing. The findings underscore the effectiveness of incorporating uncertainty into misinformation detection and spread prediction, providing an interdisciplinary and easily implementable framework for the field.

KEYWORDS

XSL•FO
RenderX

Introduction

Background

The World Health Organization and the United Nations have issued warnings about an "infodemic," highlighting the spread of misinformation alongside the COVID-19 pandemic on social media [1]. Misinformation is characterized as "factually incorrect information not backed up by evidence" [2]. This misleading information frequently encompasses harmful health advice, misinterpretations of government control measures and emerging sciences, and conspiracy theories [3]. This phenomenon has inflicted detrimental impacts on public health, carrying "severe consequences with regard to people's quality of life and even their risk of mortality" [4].

Automatic algorithms are increasingly recognized as valuable tools in mitigating the harm caused by misinformation. These techniques can rapidly identify misinformation, predict its spread, and have demonstrated commendable performance. The state-of-the-art detection techniques exhibit accuracy ranging from 65% to 90% [5,6], while spread-prediction techniques achieve performance levels between 62.5% and 77.21% [7,8]. The high accuracy of these techniques can be largely attributed to the incorporation of handcrafted or deep-learned linguistic and social features associated with misinformation [9-11]. Scholars have consistently invested efforts in integrating theoretically relevant features into algorithmic frameworks to enhance accuracy further.

Scholars have introduced diverse frameworks for misinformation detection and spread-prediction algorithms. Nevertheless, existing frameworks have predominantly concentrated on the intricate post-level signals of misinformation, emphasizing linguistic and social features (such as user relationships, replies, and knowledge sources) associated with misinformation. Notably, these frameworks have often overlooked the characteristics of the information environment in which misinformation originates and proliferates [12]. This neglect could potentially result in diminished performance for misinformation detectors when applied in various real-world misinformation contexts. This is due to the fact that different misinformation contexts possess unique characteristics within their information environment, influencing the types of misinformation that can emerge and thrive [13]. An indispensable characteristic of the information environment concerning misinformation is uncertainty. Uncertainty arises when the details of situations are ambiguous, complex, unpredictable, or probabilistic, and when information is either unavailable or inconsistent [14]. In uncertain situations, individuals tend to generate and disseminate misinformation as a means of resisting uncertainty and seeking understanding amid chaotic circumstances [15,16]. The COVID-19 pandemic serves as a notable example, marked by a lack of understanding of emerging science [17], uncertainties surrounding official guidelines and news reports [18], and unknown impacts on individuals and society [19]. Hence, in this study, we recognize uncertainty as the pivotal feature in the information environment of misinformation. Our objective is to formulate a novel framework for perceiving environmental uncertainty, specifically tailored for the detection and spread prediction of misinformation during the COVID-19 pandemic.

Our contributions can be outlined as follows. Theoretically, we provide a comprehensive exploration of uncertainty across 4 distinct scales of the information environment, namely, the physical environment, macro-media environment, micro-communicative environment, and message framing. These scales collectively contribute to the emergence and dissemination of misinformation. Furthermore, we hold the distinction of being the pioneers in integrating Environmental Uncertainty Perception (EUP) into the realms of misinformation detection and spread prediction. In terms of methodology, we introduce the EUP framework, designed to capture uncertainty signals from the information environment of a given post for both misinformation detection and spread prediction. Our experiments conducted on real-life data underscore the effectiveness of the EUP framework.

This paper unfolds as follows: In the "Related Work" section, we provide a concise review of the related work. The "Proposed Theoretical Framework" section elucidates uncertainty features within the information environment, which are pertinent to misinformation detection and spread prediction. Moving on to the "Research Objectives" section, we outline our study objectives. The "Methods" section details our methodology for testing the proposed framework. In the "Data Set and Experiment" section, we present our data set, experiments, and comprehensive analyses. The "Discussion" section delves into discussions on our findings, unraveling the theoretical and practical implications of our work. Finally, the "Conclusions" section concludes with a summary and outlines directions for future research.

Related Work

Detecting misinformation on social media represents a burgeoning research field that has garnered considerable academic attention. Multiple frameworks have been put forth for this task, primarily falling into 2 approaches: the post-only approach and the "zoom-in" approach [12]. In the former, frameworks focus on studying post features to differentiate misinformation from general information. Linguistic features, including novelty, complexity, emotions, and content topics, are frequently explored [6,11]. Additionally, researchers have delved into multimodal features, particularly those based on visuals [20,21]. Deep learning models in natural language processing have also proven beneficial for the misinformation detection task [5,22].

The "zoom-in" approach places emphasis on socio-contextual signals, centering on users' networking aspects (eg, user relationships, number of replies, number of created threads; [23,24]) and network characteristics (eg, degree centrality [25]). Another line of research underscores the significance of relevant knowledge sources, including fact-checking websites [26] and knowledge graphs [27], which can be used to validate specific claims of interest.

Recently, Sheng et al [12] introduced a "zoom-out" approach, concentrating on the information environments of misinformation that can offer signals for detection. In their

approach, they incorporated the news environment into fake news detection. Their hypothesis posited that fake news should not only be relevant but also novel and distinct from recent popular news, enabling them to capture audience attention and achieve widespread dissemination. Their findings revealed that signals of popularity and novelty can enhance the performance of state-of-the-art misinformation detectors.

In the realm of misinformation detection, misinformation spread prediction represents another challenging task, albeit one that has received limited attention. This task involves predicting whether a piece of misinformation is likely to be disseminated to a broader audience through actions such as likes, comments, and shares. Within this context, our specific focus is on predicting whether misinformation is likely to be retweeted. This can be viewed as a binary classification task, akin to misinformation detection. Frameworks for this task typically incorporate linguistic and social features, which may overlap with or differ from those used in misinformation detection. Linguistic features such as persuasive styles, emotional expressions, and message coherence prove valuable in predicting the spread of misinformation [28,29]. Additionally, social features, including user metadata (eg, number of friends, verification) and tweet metadata (eg, presence of images and URLs), are identified as relevant factors for predicting misinformation spread [25].

Proposed Theoretical Framework

Uncertainty as a Central Aspect in Misinformation

Our study builds upon Sheng et al's [12] "zoom-out" approach, adopting an interdisciplinary perspective that centers on the uncertainty within the information environment of misinformation. The realms of communication and psychology literature have conceptualized uncertainty as a fundamental aspect of misinformation. Uncertainty is said to prevail "when details of situations are ambiguous, complex, unpredictable, or probabilistic; uncertainty is also present when information is unavailable or inconsistent, and when individuals feel insecure about their own state of knowledge or the general state of knowledge" [14]. Confronted with uncertainty, individuals are driven to alleviate it by constructing their understanding of the situation [16]. This constructive process is known as sensemaking, which encompasses how individuals impart meaning to their surroundings and use it as a foundation for subsequent interpretation and action [30]. Sensemaking entails the utilization of information by individuals to fill gaps in their understanding [31]. Yet, the utilization of information in this manner does not always guarantee truth. In situations where information is slow to emerge, individuals are driven to comprehend uncertain situations by relying on their existing knowledge and heuristics for judgment. Unfortunately, this process often leads to the formation of false beliefs and misinformation [32]. Additionally, individuals may "turn to unofficial sources to satisfy their information needs," potentially exposing themselves to inaccurate information [33]. As suggested by Kim et al [34], exposure to misinformation has the potential to diminish feelings of uncertainty. Moreover, as individuals integrate more information into their comprehension of a situation, there is a tendency to seek plausibility, which

may lead to the generation and acceptance of misinformation [16,35].

The aforementioned tendencies are notably prominent in the context of the COVID-19 pandemic, as the pandemic represents a time of heightened uncertainty. The emergence of the pandemic was marked by a mysterious disease with previously unseen symptoms. Fundamental questions regarding the origins of the disease, measures for self-protection, and strategies for containing the outbreak were not immediately evident. As the pandemic progressed, uncertainty persisted regarding how and when the outbreak would be fully contained, as well as the long-term impact it would have on individuals and society. The uncertainty stemming from the pandemic, coupled with the surge of social media as a primary source of information, has facilitated the spread of misinformation [16].

Although many studies have identified "uncertainty" as a central aspect of misinformation, they have not thoroughly elucidated how uncertainty, as a crucial feature of the information environment, can aid in the detection of misinformation and the prediction of its spread. The literature frequently treats uncertainty as a static and holistic feature of a situation. However, the level of uncertainty within a situation can be dynamic, evolving as the situation progresses. For instance, uncertainties about the virus and the initial life changes induced by the COVID-19 pandemic would have been considerably higher at its onset than they are at present [36]. Moreover, uncertainty can manifest differently across various scales of the information environment. The information environment has become increasingly intricate with the proliferation of the internet and communication technologies. Individuals may be exposed to a substantial volume of information about trending topics through mainstream mass media (eg, newspapers, TV, social media trends) within a short time frame, constituting a macro-media environment. Simultaneously, they may selectively engage in detailed communications on a specific issue provided by self-media (eg, subscription accounts, self-broadcasting), shaping a micro-communicative environment. Uncertainty manifested in these 2 environments may independently or interactively influence people's sensemaking processes and, consequently, their outputs (eg, misinformation). Additionally, uncertainty can be inherent in the misinformation itself, providing cues for its detection and spread prediction. We will elaborate on the features of uncertainty in the information environment in the following section.

Uncertainty in the Information Environment

Uncertainty in the Physical Environment

Uncertainty prevails in the physical environment when unknown risks pose potential threats to our societal systems [15,16]. Scholars refer to such threats as "crises," which can encompass natural disasters, large-scale accidents, social security incidents, and public health emergencies such as the pandemic [37]. Crises are marked by the existence of uncertainty and the imperative for timely decision-making [38]. Therefore, a crucial process during crises is sensemaking. However, the efforts needed for sensemaking will vary as a crisis progresses through stages. The Crisis and Emergency Risk Communication Model delineates 5 common stages in the crisis life cycle, spanning

"from risk, to eruption, to clean-up and recovery, and on into evaluation [38]." The eruption of the crisis, also known as the breakout stage, occurs when a key event triggers the crisis [39]. This is the period when the public becomes initially aware of the crisis, characterized by mysteries and heightened motivation to make sense of it. Evidence indicates that the breakout stage of a crisis harbors the highest level of uncertainty and demands extensive sensemaking efforts (eg, government updates [40]; social media communication [41]), consequently leading to a higher incidence of misinformation [42]. This evidence implies that misinformation is more likely to surface and proliferate in tandem with uncertainty in the information environment during the breakout stage compared with other stages throughout a crisis. These insights offer valuable cues for the detection and prediction of misinformation during the COVID-19 pandemic.

Uncertainty in the Macro-Media Environment

The macro-media environment encompasses recent media opinions and public attention to trending topics [12]. Governments and mainstream media play a pivotal role in setting the agenda for public attention. During crises such as the COVID-19 pandemic, governments frequently make swift and crucial decisions to safeguard the public. However, these decisions are often made without sufficient transparency, leading to potential uncertainties surrounding their rationale [43]. Such decisions inevitably draw media and public attention, quickly becoming trending topics in mainstream media outlets [44,45]. Regrettably, these rapid decisions often leave audiences with a high level of uncertainty about the reasons behind and the processes involved in making these decisions, potentially paving the way for misinformation. Supporting this notion, Lu [3] identified a correlation between the swift decision to quarantine Wuhan city and the emergence of misinformation regarding government control measures during the early stages of the COVID-19 pandemic in China. The evidence presented indicates that when public attention is directed toward a trending topic that carries uncertainty, misinformation is likely to emerge and spread. In simpler terms, it can be anticipated that when a piece of information is associated with a trending topic characterized by high uncertainty (as opposed to low uncertainty), there is a higher probability that the information could be misinformation and disseminated.

Uncertainty in the Micro-Communicative Environment

Differing from the macro-media environment, which offers a macro perspective on what mass audiences have recently read and focused on, the micro-communicative environment provides a micro view of the communication surrounding a specific issue. Both media and individuals tend to communicate using frames or terms imbued with uncertainty when discussing matters that lack evidence or consensus, such as those stemming from emerging science during the COVID-19 pandemic [32,46]. As an illustration, in the initial phase of the pandemic, when Hong Kong officials reported the first instance of a dog testing "weakly positive" for COVID-19 infection, subsequent media reports highlighted that "Hong Kong scientists aren't sure [emphasis added] if the dog is actually infected or if it picked up the virus from a contaminated surface [47]." Experimental evidence has shown that such uncertainty frames about scientific matters can diminish people's trust in science [48]. Empirical evidence from real-life social media data further indicates that a communication style marked by ambiguity can potentially lead audiences to generate and disseminate misinformation [32]. This body of findings implies that if information is embedded in uncertain (as opposed to consensus) communication, it is more likely to be misinformation and disseminated.

Uncertainty in Message Framing

Uncertainty can also manifest within the message through its framing or word choice. Uncertainty frames are prevalent in misinformation [15,49]. Oh et al [15] illustrated that source ambiguity and content ambiguity are 2 significant features of misinformation. When individuals create a piece of misinformation that lacks evidence and credibility, they often use uncertain words to describe the unreliable source (eg, someone) or the potential rationale (eg, possible, likely) behind the statement. The incorporation of uncertain words can indeed facilitate the spread of misinformation [29,50]. The inclusion of uncertainty expressions in messages leads individuals to perceive the information as more relevant and suitable for themselves [51]. Consequently, if misinformation exhibits a higher level of uncertainty, it is more likely to be accepted and disseminated by the public.

Research Objectives

Our research objective is to explore whether uncertainty features within the information environment can enhance the effectiveness of misinformation detection and spread prediction. To achieve this, we introduce a novel EUP framework specifically designed for both tasks. We seek to assess the standalone effectiveness of the EUP and anticipate that it can augment the capabilities of existing state-of-the-art misinformation detectors and predictors. Therefore, we conducted experiments to answer the following research questions:

- *Research question 1*: Can EUP be effective in misinformation detection and spread prediction?
- *Research question 2*: Can EUP improve the performances of the state-of-the-art algorithms for misinformation detection and spread prediction?

Methods

Overview

Figure 1 offers an overview of the EUP pipeline. The model consists of 4 uncertainty extraction components. Upon receiving a post (denoted as p), the initial step involves constructing its macro-media environment and micro-communicative environment. This is accomplished by extracting recent news and social media data, respectively. Subsequently, we use a probabilistic model and a similarity calculation method to derive the uncertainty information for the 2 environments mentioned above, denoted as I_M and I_C. Likewise, we utilized the probabilistic model to capture the uncertainty of the post p itself, resulting in the representation of message framing denoted as I_F. Simultaneously, the operationalization of uncertainty in the physical environment entails using the number of COVID-19 cases and the volume of news as key indicators, denoted as I_P.

Lastly, the 4 vectors are integrated using a gate guided by the extracted post feature o (which may not necessarily equal p) from the misinformation detector, such as bidirectional encoder representations from transformers (BERT) [52]. The fused

vectors I and o are then input into the final classifier, typically a multilayer perceptron (MLP), to predict whether p is fake or real in task 1 and low or high in task 2.

Figure 1. An environmental uncertainty perception (EUP) framework for misinformation detection and spread prediction in the COVID-19 pandemic.
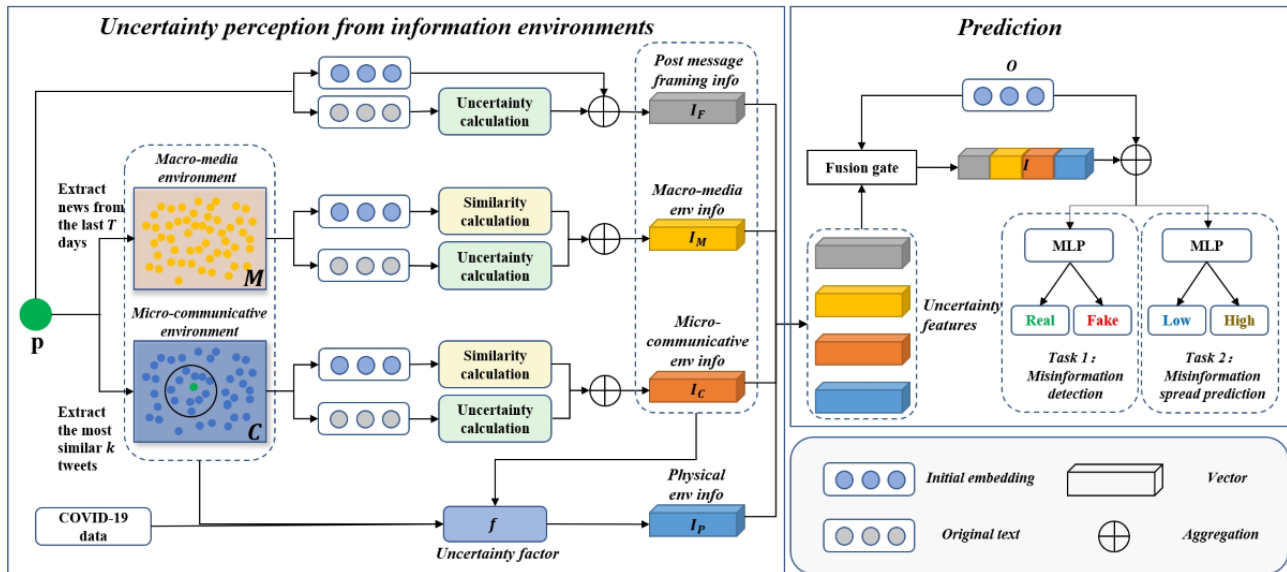
Uncertainty Detection Model

For detecting uncertainty in natural language [53], we used a probabilistic model that considers the local n-gram features of sentences. Each n-gram is assigned a weight that reflects its tendency to convey uncertainty. The definition of each feature involves a quadruplet (type, size, context, and aggregation). "Type" signifies the type of n-gram considered, such as lemma or morphosyntactic pattern. "Size" indicates the size of the n-gram. "Context" serves as an indicator, specifying whether the weight is based on the occurrence frequency of the n-gram in an uncertain sentence or on the occurrence frequency of the n-gram as an uncertainty marker. "Aggregation" refers to the method used to consolidate different scores of the n-grams within a sentence. Multimedia Appendix 1 [49,54-57] furnishes a summary of the diverse features, denoted as F_i, that are scrutinized in the uncertainty detection model.

Next, we exemplify the calculation of uncertainty using 1 of these features, F_1, as an illustration. F_1 is defined by the quadruplet (*Lemma, 1, uncertainty marker,* and *sum*). For each lemma w, we can compute the number of occurrences in the corpus, the number of occurrences in uncertain sentences, and the number of occurrences as an uncertainty marker, denoted as F_s, F_u, and F_m, respectively. The conditional probability of a lemma w becoming an uncertainty marker is calculated using the following equation:

$$p(c|w) = F_m / F_s \quad (1)$$

where c represents the class of context uncertainty under analysis, specifically whether it pertains to being an uncertainty marker. Additionally, we introduce a confidence score linked to the probability of mitigating the impact of instances where certain lemmas occur infrequently in the corpus yet yield a high probability:

$$conf(w) = 1 - (1 - F_s) \quad (2)$$

F_1 takes into account both the conditional probability of each lemma w and the corresponding confidence score in the sentence s, and the formula is calculated as follows:

$$F_1(s) = \sum_{k=1}^{|s|} p(c|w_k) \times conf(w_k) \quad (3)$$

Similarly, other features F_i can be derived using the above method. We generated the uncertainty of the whole sentence by mean pooling to represent the average uncertainty signals of F_i:

$$F^{A,\text{Mean}}(s) = \text{Mean}(\text{Norm}(\{F_i(s)\}^{|F|}_{i=1})) \quad (4)$$

where $\text{Norm}(\cdot)$ denotes the normalization.

Representation of the Macro-Media Environment

We collect news reports from mainstream media outlets released within T days before the post p is published to construct a *macro-media environment* according to the following definition:

$$M = \{e: e \in E, 0 \le t_p - t_e \le T\} \quad (5)$$

where E denotes the set of all collected news items, M denotes the set of news items in the macro-media environment of the post p, and t_p and t_e represent the release time of post p and news e, respectively. For post p or each news item e, the initial representations are the output of a pretrained language model (eg, BERT [52]), denoted as p and e, respectively.

The macro-media environment is expected to reflect the impact of a trending topic with high uncertainty on the veracity of a post. That is, if a post is related to a trending topic with (vs without) high uncertainty, it is then expected to be more likely misinformation and disseminated. To this end, the representation of the macro-media environment should consider both the

correlation between the post and the environment and the uncertainty of the environment. We first calculate cosine similarity between $\boldsymbol{p}$ and each news item $\boldsymbol{e}$ in E:

$$S(p,e) = (p \cdot e)/(|p| \cdot |e|) \quad (6)$$

We combine the similarity and environment representations to represent the similarity representation of a post $\boldsymbol{p}$ to the environment:

$$S_M = \sum_{i=1}^{|M|} \left(S(\boldsymbol{p}, \boldsymbol{e}_i^M) \odot \boldsymbol{e}_i^M \right) \quad (7)$$

where $\boldsymbol{e}_i^M$ represents each news item in M and $\odot$ is the Hadamard product operator.

We then measure the uncertainty of the macro-media environment using the model described in the "Uncertainty Detection Model" section. The uncertainty representation of the macro-media environment, denoted as U_M, can be expressed by the following equation:

$$U_M = \sum_{i=1}^{|M|} \left(F^{A,Mean}(e_i^M) \odot \boldsymbol{e}_i^M \right) \quad (8)$$

Finally, the macro-media environment of a post p is represented as an aggregation of the similarity representation of p to the environment (S_M) and the uncertainty representation of the environment (U_M) using an MLP, denoted as I_M:

$$I_M = \text{MLP}(S_M \oplus U_M) \quad (9)$$

where $\oplus$ is the concatenation operator. The integration of an MLP is instrumental in the dual objective of retaining crucial information while concurrently achieving data dimensionality reduction. All MLPs are individually parameterized. We omit their index numbers in the above equations for brevity.

Representation of the Micro-Communicative Environment

We collected tweets from Twitter (X; X Corp.) published within T days before the post p was published to construct the micro-communicative environment. We calculated the similarity of all tweets to the post p and selected the top k of them, using them as a *micro-communicative environment* (C), which is defined as follows:

$$C' = \{v : v \in V, 0 \le t_p - t_v \le T\} \quad (10)$$

where V denotes the set of all collected tweet items and t_v represents the release time of the tweet v.

$$C = \{v : v \in \text{Top}k(p, C')\} \quad (11)$$

where $\text{Top}k(\cdot)$ represents the operation of selecting the k tweets that have the highest similarity to p, $k = r \cdot |C'|$, and $r \in (0,1)$ represents the percentage of extraction.

Using the same approach as in the previous 2 sections, we derive the similarity representation of the post p to the micro-communicative environment and the uncertainty representation of the environment:

$$S_C = \sum_{i=1}^{|C|} \left(S(\boldsymbol{p}, \boldsymbol{v}_i^C) \odot \boldsymbol{v}_i^C \right) \quad (12)$$

$$U_C = \sum_{i=1}^{|M|} \left(F^A(\boldsymbol{v}_i^C) \odot \boldsymbol{e}_i^C \right) \quad (13)$$

Finally, the micro-communicative environment of a post p is represented as an aggregation of the similarity representation of a post p to the environment (S_C) and the uncertainty representation of the environment (U_C) using an MLP, denoted as I_C:

$$I_C = \text{MLP}(S_C U_C) \quad (14)$$

Message Framing

To perceive the uncertainty in the message framing of post p, we used the same approach as described in the "Uncertainty Detection Model" section to construct the uncertainty representation of the post p:

$$I_F = \text{MLP}[\boldsymbol{F}(p) \oplus \boldsymbol{p}] \quad (15)$$

Physical Environment

To measure uncertainty in the physical environment, we collected the daily number of new cases from the start of the COVID-19 outbreak and counted the number of daily news items related to the outbreak, denoted as N^{Cases} and N^{News}, respectively. Intuitively, the higher the number of new cases and news items for a day, the more sensitive the public is to the social environment and the more uncertain the environment is on that day. Thus, the uncertainty factor in the physical environment is defined as follows:

$$f^{ph}_i = \text{Norm}(\log(1+\text{abs}(N_i^{\text{Cases}} - N_{i-1}^{\text{Cases}})) \times \log(1+\text{abs}(N_i^{\text{News}} - N_{i-1}^{\text{News}}))) \quad (16)$$

where f^{ph}_i denotes the uncertainty factor at day i and *abs* is the absolute value operation. For each post, we can obtain the uncertainty factor for its corresponding date $f^{ph}(p)$.

We added the uncertainty factor of the physical environment to the representations of *macro-media environment* (I_M), *micro-communicative environment* (I_C), and *post message framing* (I_F) to get the representation of the physical environment, denoted as I_P:

$$I_P = (f^{ph} \times I_M) \oplus (f^{ph} \times I_C) \oplus (f^{ph} \times I_F) \quad (17)$$

Prediction

Prediction With EUP Alone Without Baseline Models

We concatenate the above 4 environment uncertainty features and feed the result into an MLP layer and a softmax layer for the final prediction:

$$I_{\text{EUP}} = I_M \oplus I_C \oplus I_F \oplus I_P \quad (18)$$

$$\hat{y} = \text{softmax}(\text{MLP}[I_{\text{EUP}}]) \quad (19)$$

XSL•FO

RenderX

Prediction With Baseline Models

We expect that our EUP is compatible with and can empower various misinformation detection and prediction algorithms. Therefore, we used an adaptive feature selection approach based on a gate mechanism to accommodate different misinformation detectors:

$$I = g_M \odot I_M \quad + \quad g_C \odot I_C \quad + \quad g_F \odot I_F \quad +$$
$$g_P{}^{\hat{y} \,=\, \text{softmax}(\text{MLP}[I_{\text{EUP}}])} I_P \; (20)$$

where o denotes the last-layer feature from the misinformation baseline algorithm. The gating vector g_M=sigmoid(Linear($o \oplus I_M$)) and g_C, g_F, and g_P are obtained in the same way. Then, we concatenated o and I, and fed the result into an MLP layer and a softmax layer for the final prediction:

$$\hat{y} = softmax(MLP[o \oplus I]) \; (21)$$

During training, we minimize the cross-entropy loss.

Ethical Considerations

The study is exempt from ethical review for human subject research for the following reasons. First, the study uses data from 2 publicly available Twitter data sets collected through the official application programming interface (API) of the Twitter platform for gathering tweets. The news data set was obtained from the official websites of news media. Second, the data used in this study are anonymized and do not contain any personally identifiable information. It is also impossible to reidentify individuals from the data set. The data set is stored on a dedicated secure data server, and the analysis is conducted on the platform's designated site. This process is undertaken for research purposes and adheres to Chinese data privacy laws and regulations. Third, this study does not involve any experimental manipulation of human individuals or other ethical concerns. For instance, it does not include data on children under 18 years of age, which require legally mandated parental or guardian supervision. It also does not encompass sensitive aspects of participants' behavior or pose any physical, psychological, or economic harm or risk to the research participants.

Data Set and Experiment

Data Set

The statistics and description of our experimental data set are shown in Tables 1 and 2, respectively.

Table 1. Statistics of the data set.[a,b]

Data set	Misinformation detection, n		Spread prediction, n		Total, n
	Real	Fake	Low	High	
Train	901	1324	1054	1171	2225
Value	312	430	360	382	742
Test	310	432	358	384	742

[a]News items in M=58,095. The corresponding mean and range are 988 and 10-2511, respectively.

[b]Tweet items in C=321,656. The corresponding mean and range are 793, 138-1214, respectively.

Table 2. Descriptions of the data set.

Data	Features	Size, n
Post	Content, created time, retweet count, veracity label, retweeted label	3709
News	Content, created time	58,095
Tweets	Content, created time	321,656

Post

We processed and integrated 2 existing COVID-19 data sets, FibVID [58] and CMU_MisCov19 [59], for our experiments. Both data sets have been labeled for veracity by experts, providing ground-truth labels for our experimental evaluations. For FibVID, we extracted data related to COVID-19, assigning veracity tags as 0 (COVID true) or 1 (COVID fake). We relabeled CMU_MisCov19, classifying *calling out or correction, true public health response,* and *true prevention* as *real* tags, and *conspiracy, fake cure, sarcasm or satire, false fact or prevention, fake treatment,* and *false public health response* as *fake* tags. Furthermore, we used the Twitter API to retrieve the number of retweets for all tweets in both data sets. Subsequently, we categorized the retweet labels as low (when the retweet count is 0) and high (when the retweet count is >0) following an analysis of the distribution of retweet numbers. The data revealed that misinformation was predominantly observed from January to July 2020, coinciding with the period of heightened uncertainty during the pandemic outbreak. Consequently, our focus was directed solely to this specific period, resulting in the extraction of 3709 posts from January to July of 2020.

Macro-Media Environment

We gathered all the news headlines and brief descriptions from the Huffington Post, NPR, and Daily Mail from January to July 2020, as per the methodology outlined previously [12]. Notably, these 3 outlets represent the left-, center-, and right-wing perspectives, contributing to the diversity of news items for our analysis. We then used the keywords "covid," "coronavirus,"

"pneumonia," "pandemic," "epidemic," "infection," "prevalence," and "symptom" to filter these data to ensure that the collected data were relevant to COVID-19. We ended up with 58,095 news items from January to July 2020.

Micro-Communicative Environment

We obtained the tweet IDs associated with COVID-19 from an ongoing project [60]. Given the substantial volume, we randomly sampled 1% of these IDs (amounting to approximately 205,581,778 records). Subsequently, using the Twitter API, we retrieved the content associated with these IDs, resulting in a data set comprising 321,656 tweets spanning from January to July 2020.

Physical Environment

We compiled the daily count of new worldwide COVID-19 cases starting from January 2020, utilizing the Our World in Data database. Additionally, the daily volume of news data corresponds to the information we gathered during the same period.

Experimental Setup

Tasks

We used the proposed model for 2 tasks:

Task 1. Misinformation Detection

The objective was to analyze the text content of a tweet and ascertain whether it contained misinformation.

Task 2: Spread Prediction

The objective was to evaluate the text content of a tweet to determine whether it is likely to be retweeted.

Uncertainty Features

Following Jean et al [53], we used WikiWeasel [61], a comprehensive corpus consisting of paragraphs extracted from Wikipedia, to compute the frequency of each lemma. The uncertainty score for each sentence is determined using mean pooling $F^{A,\mathrm{Mean}}$. We leverage [62] to acquire sentence representations, relying on pretrained BERT models [52] and subsequent posttraining on news items. In the macro-media environment and the micro-communicative environment, we set $T=3$, $r=0.1$, $|C|_{\min}=10$.

Baseline Models

The baseline models considered are listed in Textbox 1.

Textbox 1. Baseline models.

1. **Bidirectional long short-term memory**

   Bidirectional long short-term memory (BiLSTM) [63] is a type of recurrent neural network architecture designed for sequence modeling tasks, particularly in natural language processing. It processes input sequences in both forward and backward directions simultaneously, allowing the model to capture information from both past and future contexts.

2. **Event adversarial neural networks**

   Event adversarial neural networks (EANN$_T$) [64] is a model using adversarial training to eliminate event-specific features derived from a convolutional neural network for text (ie, TextCNN).

3. **BERT**

   Bidirectional encoder representations from transformers (BERT) [52] is a pretrained language model based on deep bidirectional transformers.

4. **BERT-Emo**

   BERT-Emo [65] is a fake news detection model that integrates multiple sentiment features into BERT.

Evaluation Metrics

For both tasks, we used accuracy and macro-F_1-score as evaluation metrics. Additionally, in task 1, we used F_1-scores for fake ($F_{1\mathrm{fake}}$) and real ($F_{1\mathrm{real}}$), while in task 2, we considered F_1-scores for low ($F_{1\mathrm{low}}$) and high ($F_{1\mathrm{high}}$). Further implementation details can be found in Multimedia Appendix 1.

Results

Overview

Tables 3 and 4 showcase the performances of the EUP without baseline models and those of various baseline models, with and without EUP, for the misinformation detection and spread prediction tasks, respectively. The results indicate that the performances of EUP are comparable to those of state-of-the-art baseline models in both tasks. Moreover, it is noteworthy that all baseline models exhibit performance improvements when incorporating EUP for both tasks. These observations suggest the effectiveness of our proposed EUP.

Table 3. Model performance comparison on the misinformation detection task without the baseline algorithm or without the EUP[a] module.[b]

Model	Accuracy	Macro-F_1-score	$F_{1\ fake}$	$F_{1\ real}$
EUP	*0.753*	*0.739*	*0.800*	*0.677*
BiLSTM[c]	0.733	0.729	0.783	0.683
BiLSTM + EUP	*0.755*	*0.743*	*0.798*	*0.688*
EANN$_T$[d]	0.745	0.730	0.795	0.664
EANN$_T$ + EUP	*0.767*	*0.765*	*0.806*	*0.708*
BERT[e]	0.755	0.743	*0.797*	0.689
BERT + EUP	*0.771*	*0.767*	0.796	*0.738*
BERT-Emo	0.749	0.740	0.789	0.691
BERT-Emo + EUP	*0.768*	*0.763*	*0.799*	*0.726*

[a]EUP: Environmental Uncertainty Perception.

[b]The best result in each group is in italics.

[c]BiLSTM: bidirectional long short-term memory.

[d]EANN$_T$: event adversarial neural networks.

[e]BERT: bidirectional encoder representations from transformers.

Table 4. Model performance comparison on the spread prediction task without the baseline algorithm or without the EUP[a] module.[b]

Model	Accuracy	Macro-F_1-score	$F_{1\ low}$	$F_{1\ high}$
EUP	*0.710*	*0.710*	*0.719*	*0.701*
BiLSTM[c]	0.707	0.705	0.684	0.726
BiLSTM + EUP	*0.734*	*0.733*	*0.738*	*0.729*
EANN$_T$[d]	0.717	0.716	0.734	0.698
EANN$_T$ + EUP	*0.726*	*0.726*	*0.736*	*0.716*
BERT[e]	0.728	0.728	0.728	0.728
BERT + EUP	*0.743*	*0.743*	*0.752*	*0.734*
BERT-Emo	0.733	0.733	0.730	0.737
BERT-Emo + EUP	*0.741*	*0.741*	*0.733*	*0.749*

[a]EUP: Environmental Uncertainty Perception.

[b]The best result in each group is in italics.

[c]BiLSTM: bidirectional long short-term memory.

[d]EANN$_T$: event adversarial neural networks.

[e]BERT: bidirectional encoder representations from transformers.

Ablation Study

We systematically eliminated individual components, namely, macro-media environment, micro-communicative environment, message framing, and physical environment, and assessed the modeling performances on the data set. Tables 5 and 6 illustrate that, under all experimental conditions, performance degrades when any of these components are removed. These results underscore the effectiveness of all 4 uncertainty features of the information environment for both misinformation detection and spread prediction.

XSL•FO

RenderX

Table 5. Ablation study on the misinformation detection task.[a]

Model	Accuracy	Macro-F_1-score	$F_{1\ \text{fake}}$	$F_{1\ \text{real}}$
EUP[b]	*0.753*	*0.739*	0.800	0.677
Without I_M	0.748	0.738	0.790	*0.687*
Without I_C	0.745	0.720	*0.803*	0.637
Without I_F	0.739	0.734	0.778	0.673
Without I_P	0.747	0.730	0.797	0.663
BiLSTM[c] + EUP	*0.755*	*0.743*	*0.798*	*0.688*
Without I_M	0.745	0.741	0.793	0.669
Without I_C	0.741	0.728	0.788	0.668
Without I_F	0.747	0.735	0.791	0.678
Without I_P	0.746	0.742	0.796	0.665
BERT[d] + EUP	*0.771*	*0.767*	*0.796*	*0.738*
Without I_M	0.762	0.754	0.801	0.707
Without I_C	0.764	0.761	*0.807*	0.696
Without I_F	0.761	0.752	0.800	0.705
Without I_P	0.758	0.751	0.795	0.707

[a]The best result in each group is in italics.

[b]EUP: Environmental Uncertainty Perception.

[c]BiLSTM: bidirectional long short-term memory.

[d]BERT: bidirectional encoder representations from transformers.

Table 6. Ablation study on the spread prediction task.[a]

Model	Accuracy	Macro-F_1-score	$F_{1\ low}$	$F_{1\ high}$
EUP[b]	*0.710*	*0.710*	0.719	*0.701*
Without I_M	0.697	0.696	0.715	0.676
Without I_C	0.695	0.694	0.712	0.677
Without I_F	0.702	0.702	0.714	0.689
Without I_P	0.708	0.707	*0.721*	0.692
BiLSTM[c] + EUP	*0.734*	*0.733*	0.738	*0.729*
Without I_M	0.724	0.723	0.735	0.711
Without I_C	0.721	0.721	0.716	0.726
Without I_F	0.717	0.716	0.731	0.702
Without I_P	0.726	0.723	*0.753*	0.693
BERT[d] + EUP	*0.743*	*0.743*	0.752	*0.734*
Without I_M	0.741	0.739	0.764	0.713
Without I_C	0.741	0.738	*0.766*	0.711
Without I_F	0.736	0.735	0.753	0.716
Without I_P	0.740	0.738	0.759	0.717

[a]The best result in each group is in italics.

[b]EUP: Environmental Uncertainty Perception.

[c]BiLSTM: bidirectional long short-term memory.

[d]BERT: bidirectional encoder representations from transformers.

The Effect of the Day Parameter T

To explore the impact of the day parameter (T) on the results during the construction of the macro-media environment and the micro-communicative environment, we experimented with different values of T. Specifically, we sequentially set T=1, 3, 5, 7, and 9 for the BERT + EUP model, and the experimental results are depicted in Figure 2. Despite the fact that increasing T results in larger macro-media and micro-communicative environments, the optimal performance was achieved when T=1.

Figure 2. The effect of the day parameter T. Lines show the accuracies of both tasks and bars show the average number of news and tweet items in the environments.
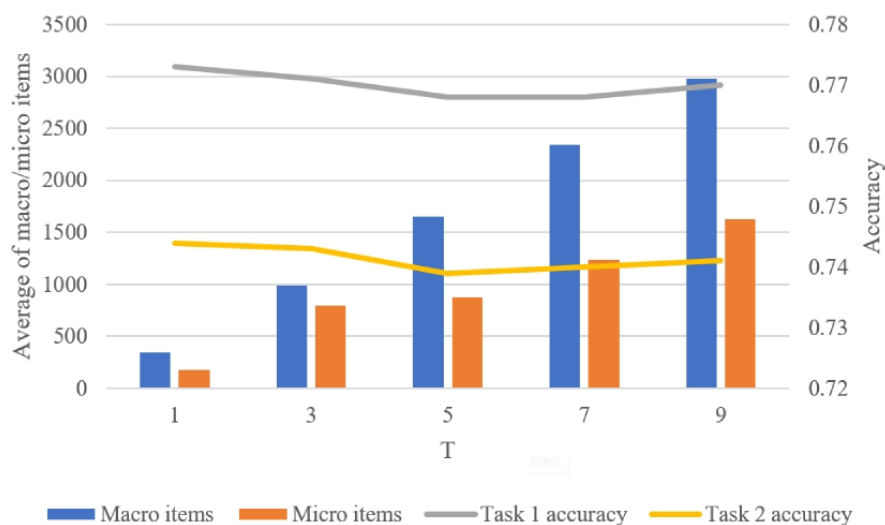
The Effect of the Rate Parameter r

We maintained the setting $T=3$ and systematically varied r, using values of 0.05, 0.1, 0.15, 0.2, 0.25, and 0.3 on the BERT + EUP model to examine the impact of r on the experimental results, as illustrated in Figure 3. The accuracy performance exhibited fluctuations with varying values of r. Notably, the highest accuracy for both tasks was observed when $r=0.1$.

Figure 3. The effect of the rate parameter r. Lines show the accuracies of both tasks and bars show the average number of tweet items in the environment.
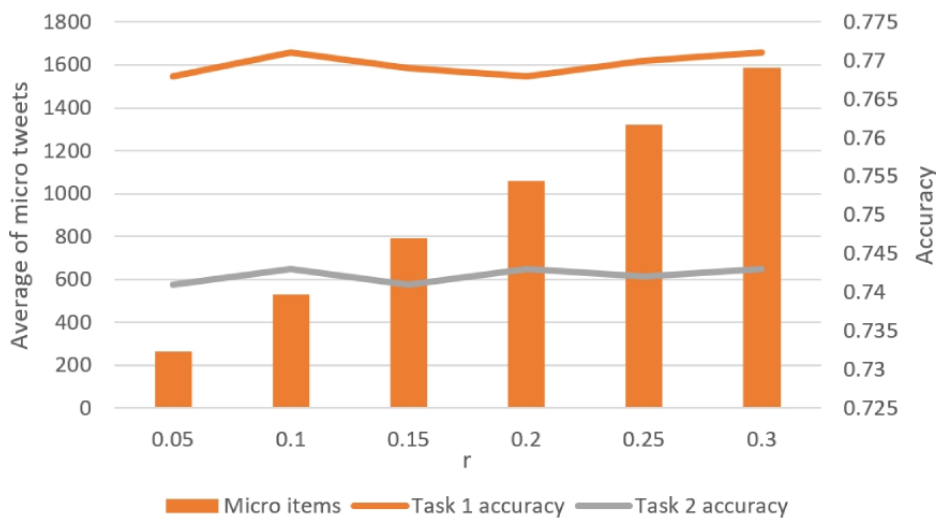
Evaluation on Imbalanced Data

In real-world scenarios, the distribution of real and fake information often exhibits significant imbalance. To evaluate the efficacy of our proposed EUP framework on unbalanced data sets, we conducted tests on data sets with varying ratios of real to fake data, ranging from 10:1 to 100:1. We measured and reported macro-F_1-scores and standardized partial area under the curve (AUC) with a false-positive rate of at most 0.1 (ie, spAUCFPR≤0.1 [66]) to assess the effectiveness of our EUP framework in handling nonbalanced data sets. As depicted in Figure 4, EUP yields relative improvements of 21.5% and 5.7% in macro-F_1-score and spAUCFPR≤0.1, demonstrating its effectiveness on unbalanced data sets.

Figure 4. Performance of macroF1 and spAUC values across datasets with varying ratios.
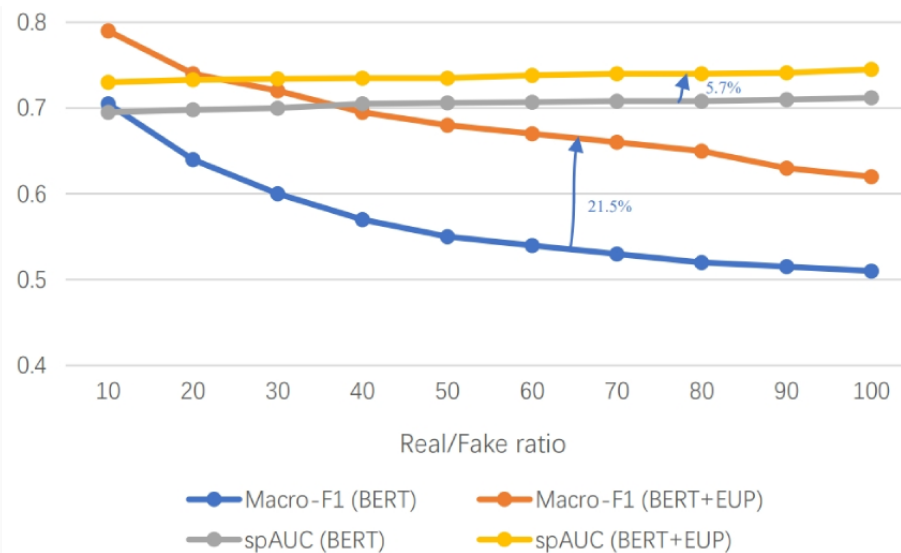
Discussion

Principal Findings

First, this study enhances scholars' comprehension of the misinformation detection and spread prediction problem by highlighting the significance of uncertainty in information environments. Notably, this research contributes to the literature by recognizing uncertainty features in the information environments of misinformation as a pivotal factor for improving detection and prediction algorithms during a pandemic. Our findings underscore that the EUP alone is sufficient for both tasks and has the potential to enhance the capabilities of state-of-the-art algorithms. In contrast to prior misinformation research that primarily concentrates on post content (such as post theme, sentiments, and linguistic characteristics, as seen in [6,11,29]) and network connections (eg, number of followers [25]) on social media, this study advances scholars' understanding of the misinformation problem by emphasizing the importance of uncertainty in information

environments. Recognizing and incorporating uncertainty as a fundamental concept in misinformation detection and spread prediction during crises hold theoretical significance. This is particularly relevant as a crisis is characterized by its unpredictable, unexpected, and nonroutine nature, inherently giving rise to uncertainty [38,67]. This uncertainty has been theorized to compel individuals to seek information as a coping mechanism for dealing with the anxiety and pressure generated by uncertainty. This process allows people to diminish uncertainty, restore a sense of normalcy, and alleviate anxiety [14,68]. Regrettably, this coping mechanism can inadvertently fuel the proliferation and dissemination of misinformation, particularly when there is a lack of timely and accurate information, contributing to the concurrent occurrence of an infodemic [6,11,50]. The current research seeks to advance the literature by establishing the legitimacy of uncertainty in the information environments of misinformation as a central indicator for the detection and prediction of misinformation during public health crises.

Second, this study delves into the intricacies of uncertain information environments for misinformation across 4 distinct scales, namely, the physical environment, macro-media environment, micro-communicative environment, and message framing. Our findings demonstrate the effectiveness of all 4 uncertainty features in misinformation detection and spread prediction. In contrast to prior misinformation literature during the COVID-19 pandemic, which often overlooked the role of the information environment in increasing the likelihood of misinformation dissemination, our research emphasizes the importance of considering uncertainty beyond the content of misinformation itself, such as ambiguous wording [29,50]. Our study broadens the concept of linguistic uncertainty in misinformation message framing to encompass a more comprehensive uncertainty across various information environments. We define uncertainty in information environments using a multiscale approach that highlights the significance of the interaction between the physical environment and macro-/micro-media environments. This approach diverges from focusing on a single dimension, such as ambiguities about official guidelines and news reports [18], or the misinformation framing strategy on social media [29].

Third, our findings indicate that uncertainties in information environments play a crucial role as motivators for the emergence and spread of misinformation. While previous studies have provided preliminary evidence suggesting that uncertainty stemming from government policies and news media could coincide with the occurrence of related misinformation during the COVID-19 pandemic, often relying on descriptive big data analyses [3,32], our study contributes stronger empirical evidence. We leverage machine learning techniques to demonstrate that uncertainty arising from the crisis and crisis communication through media can indeed incentivize individuals to generate and disseminate misinformation. Significantly, our findings revealed that the algorithm achieved its best performance for both detection and spread prediction tasks when incorporating items from the information environments published 1 day before the post (T=1). This discovery emphasizes the acute impact of uncertainty in the

information environment on the emergence and spread of misinformation, underscoring the importance of timely uncertainty reduction in crisis communication. Furthermore, the algorithm attained the highest accuracies when it included items highly relevant to the post but with an appropriate size (r=0.1). This rationale is reasonable, as a too-small r may fail to encompass enough misinformation-related items, while a larger r might include a significant amount of irrelevant information. The evidence theoretically establishes a connection between crisis communication research and misinformation research, reinforcing the notion that crisis communication and misinformation containment are 2 intertwined aspects of crisis management [3].

This study offers significant practical implications for misinformation detection and spread prediction. First, unlike previous studies that separately investigated computational frameworks for these tasks [24,29], this study introduces a unified uncertainty–based framework capable of addressing both tasks simultaneously. Second, our framework operates instantaneously, as it only requires easily accessible data such as posts, mainstream news, and relevant social media discussions published a few days prior. Moreover, the uncertainty detection algorithm has been trained using external data, rendering our algorithm easy to implement and capable of providing timely detection and prediction for streaming textual data. Third, this study affirms the effectiveness of uncertainty in various information environments for detecting and predicting misinformation on social media. Hence, the 4 proposed uncertainty components in information environments could be leveraged by social media platforms to improve the accuracy of misinformation detection and spread prediction, thereby safeguarding individuals from harm caused by infodemic. The benefits offered by our algorithm may serve as an impetus for integrating uncertainty components into practical systems.

Limitations and Future Work

This study is the first to incorporate the uncertainty present in the information environment of a post for both misinformation detection and spread prediction. However, it has some limitations. First, our framework concentrated solely on text-only detection and prediction. Future work should extend the framework to incorporate multimodal and social graph–based detection. Second, we used an uncertainty detection algorithm developed from a generic corpus sourced from Wikipedia. Nevertheless, past research has indicated that expressions of uncertainty may vary slightly across domains [53]. In other words, uncertainty expressions in the context of the COVID-19 pandemic may differ from those in general situations. Therefore, future work should aim to enhance our uncertainty measure by utilizing a corpus specifically designed for uncertainty detection in the discourse related to COVID-19.

Conclusions

We introduced an EUP framework for both misinformation detection and spread prediction. Our framework delves into uncertainty within information environments across 4 scales: the physical environment, macro-media environment, micro-communicative environment, and message framing. The experiments demonstrated the effectiveness of our proposed

uncertainty components in enhancing the performance of existing models. There are several directions for further investigation and extension of this work. First, we can explore the impact of different news and social media environments (eg, biased vs neutral; left wing vs right wing) on the emergence and spread of misinformation. Second, extending our algorithms to include multimodal misinformation detection could be beneficial, as misinformation increasingly incorporates images and videos. Third, investigating the interaction between misinformation detection and spread prediction using a multitask, transfer-learning model is a promising avenue, given the shared uncertainty framework identified in this study for both tasks.

Acknowledgments

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Uncertainty features.
[DOCX File , 18 KB-Multimedia Appendix 1]

References

1. Thomas Z. WHO says fake coronavirus claims causing "infodemic". BBC. 2020. URL: https://www.bbc.com/news/technology-51497800 [accessed 2022-09-08]
2. Bode L, Vraga EK. See something, say something: correction of global health misinformation on social media. Health Commun. Sep 16, 2018;33(9):1131-1140. [doi: 10.1080/10410236.2017.1331312] [Medline: 28622038]
3. Lu J. Themes and evolution of misinformation during the early phases of the COVID-19 outbreak in China—an application of the crisis and emergency risk communication model. Front Commun. Aug 14, 2020;5:57. [doi: 10.3389/fcomm.2020.00057]
4. Swire-Thompson B, Lazer D. Public health and online misinformation: challenges and recommendations. Annu Rev Public Health. Apr 02, 2020;41(1):433-451. [FREE Full text] [doi: 10.1146/annurev-publhealth-040119-094127] [Medline: 31874069]
5. Jiang G, Liu S, Zhao Y, Sun Y, Zhang M. Fake news detection via knowledgeable prompt learning. Information Processing & Management. Sep 2022;59(5):103029. [doi: 10.1016/j.ipm.2022.103029]
6. Kumari R, Ashok N, Ghosal T, Ekbal A. Misinformation detection using multitask learning with mutual learning for novelty detection and emotion recognition. Information Processing & Management. Sep 2021;58(5):102631. [doi: 10.1016/j.ipm.2021.102631]
7. Babic K. Prediction of COVID-19 Related Information Spreading on Twitter. New York, NY. IEEE; Presented at: 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO); May 24-28, 2021, 2021;395-399; Opatija, Croatia. [doi: 10.23919/MIPRO52101.2021.9596693]
8. Ghina Khoerunnisa; Jondri; Widi Astuti. Prediction of retweets based on user, content, and time features using EUSBoost. J RESTI (Rekayasa Sist Teknol Inf). Jun 30, 2022;6(3):442-447. [doi: 10.29207/resti.v6i3.4125]
9. Islam MR, Liu S, Wang X, Xu G. Deep learning for misinformation detection on online social networks: a survey and new perspectives. Soc Netw Anal Min. 2020;10(1):82. [FREE Full text] [doi: 10.1007/s13278-020-00696-x] [Medline: 33014173]
10. Shu K, Sliva A, Wang S, Tang J, Liu H. Fake news detection on social media. SIGKDD Explor Newsl. Sep 2017;19(1):22-36. [doi: 10.1145/3137597.3137600]
11. Su Q, Wan M, Liu X, Huang C. Motivations, methods and metrics of misinformation detection: an NLP perspective. NLPRE. 2020;1(1-2):1. [doi: 10.2991/nlpr.d.200522.001]
12. Sheng Q, Cao J, Zhang X, Li R, Wang D, Zhu Y. Zoom out and observe: news environment perception for fake news detection. New York, NY. Association for Computational Linguistics; Presented at: Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers); May 22-27, 2022, 2022;4543-4556; Dublin, Ireland. [doi: 10.18653/v1/2022.acl-long.311]
13. Rosnow R. Rumor as communication: a contextualist approach. Journal of Communication. 1988;38(1):12-28. [doi: 10.1111/j.1460-2466.1988.tb02033.x]
14. Bradac JJ. Theory comparison: uncertainty reduction, problematic integration, uncertainty management, and other curious constructs. Journal of Communication. 2001;51(3):456-476. [doi: 10.1111/j.1460-2466.2001.tb02891.x]

15. Oh O, Agrawal M, Rao HR. Community intelligence and social media services: a rumor theoretic analysis of tweets during social crises. MISQ. Feb 2, 2013;37(2):407-426. [doi: 10.25300/misq/2013/37.2.05]

16. Tandoc EC, Lee JCB. When viruses and misinformation spread: how young Singaporeans navigated uncertainty in the early stages of the COVID-19 outbreak. New Media & Society. Oct 25, 2020;24(3):778-796. [doi: 10.1177/1461444820968212]

17. Capurro G, Jardine CG, Tustin J, Driedger M. Communicating scientific uncertainty in a rapidly evolving situation: a framing analysis of Canadian coverage in early days of COVID-19. BMC Public Health. Nov 29, 2021;21(1):2181-2114. [FREE Full text] [doi: 10.1186/s12889-021-12246-x] [Medline: 34844582]

18. Zhang YSD, Young Leslie H, Sharafaddin-Zadeh Y, Noels K, Lou NM. Public health messages about face masks early in the COVID-19 pandemic: perceptions of and impacts on Canadians. J Community Health. Oct 20, 2021;46(5):903-912. [FREE Full text] [doi: 10.1007/s10900-021-00971-8] [Medline: 33611755]

19. Dietrich AM, Kuester K, Müller GJ, Schoenle R. News and uncertainty about COVID-19: survey evidence and short-run economic impact. J Monet Econ. Jul 2022;129:S35-S51. [FREE Full text] [doi: 10.1016/j.jmoneco.2022.02.004] [Medline: 35165494]

20. Cao J, Qi P, Sheng Q, Yang T, Guo J, Li J. Exploring the role of visual content in fake news detection. In: Shu K, Wang S, Lee D, Liu H, editors. Disinformation, Misinformation, and Fake News in Social Media. Cham, Switzerland. Springer; Jun 18, 2020;141-161.

21. Qi P, Cao J, Li X, Liu H, Sheng Q, Mi X, et al. Improving fake news detection by using an entity-enhanced framework to fuse diverse multimodal clues. In: MM '21: Proceedings of the 29th ACM International Conference on Multimedia. New York, NY. Association for Computing Machinery; Presented at: The 29th ACM International Conference on Multimedia (MM '21); October 17, 2021, 2021;1212-1220; Chengdu, China. [doi: 10.1145/3474085.3481548]

22. Liu C, Wu X, Yu M, Li G, Jiang J, Huang W, et al. A two-stage model based on BERT for short fake news detection. Cham, Switzerland. Springer; Presented at: International Conference on Knowledge Science, Engineering and Management (KSEM 2019); August 28-30, 2019, 2019;172-183; Athens, Greece. [doi: 10.1007/978-3-030-29563-9_17]

23. Vo N, Lee K. Hierarchical multi-head attentive network for evidence-aware fake news detection. In: Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics. New York, NY. Association for Computational Linguistics; Presented at: The 16th Conference of the European Chapter of the Association for Computational Linguistics; April 1, 2021, 2021;965-975; Online. [doi: 10.18653/v1/2021.eacl-main.83]

24. Silva A, Han Y, Luo L, Karunasekera S, Leckie C. Propagation2Vec: embedding partial propagation networks for explainable fake news early detection. Information Processing & Management. Sep 2021;58(5):102618. [doi: 10.1016/j.ipm.2021.102618]

25. Zhao Y, Da J, Yan J. Detecting health misinformation in online health communities: incorporating behavioral features into machine learning based approaches. Information Processing & Management. Jan 2021;58(1):102390. [doi: 10.1016/j.ipm.2020.102390]

26. Shaden S, Nikolay B, Giovanni DSM, Preslav N. That is a known lie: detecting previously fact-checked claims. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. New York, NY. Association for Computational Linguistics; Presented at: The 58th Annual Meeting of the Association for Computational Linguistics; July 5-10, 2020, 2020;3607-3618; Online. URL: https://aclanthology.org/2020.acl-main.332.pdf [doi: 10.18653/v1/2020.acl-main.332]

27. Cui L, Seo H, Tabar M, Ma F, Wang S, Lee D. DETERRENT: knowledge guided graph attention network for detecting healthcare misinformation. In: KDD '20: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York, NY. Association for Computing Machinery; Presented at: KDD '20: The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining; July 6-10, 2020, 2020; Virtual Event. [doi: 10.1145/3394486.3403092]

28. Kumar KPK, Geethakumari G. Detecting misinformation in online social networks using cognitive psychology. Hum Cent Comput Inf Sci. Sep 24, 2014;4(1):1-22. [doi: 10.1186/s13673-014-0014-x]

29. Zhou C, Li K, Lu Y. Linguistic characteristics and the dissemination of misinformation in social media: the moderating effect of information richness. Information Processing & Management. Nov 2021;58(6):102679. [doi: 10.1016/j.ipm.2021.102679]

30. Keller AC, Ansell CK, Reingold AL, Bourrier M, Hunter MD, Burrowes S, et al. Improving pandemic response: a sensemaking perspective on the spring 2009 H1N1 pandemic. Risk Hazard & Crisis Pub Pol. Aug 10, 2012;3(2):1-37. [doi: 10.1515/1944-4079.1101]

31. Genuis SK. Constructing "sense" from evolving health information: a qualitative investigation of information seeking and sense making across sources. J Am Soc Inf Sci Tec. Jun 29, 2012;63(8):1553-1566. [doi: 10.1002/asi.22691]

32. Lu J, Zhang M, Zheng Y, Li Q. Communication of uncertainty about preliminary evidence and the spread of its inferred misinformation during the COVID-19 pandemic—a Weibo case study. Int J Environ Res Public Health. Nov 13, 2021;18(22):11933. [FREE Full text] [doi: 10.3390/ijerph182211933] [Medline: 34831688]

33. Heverin T, Zach L. Use of microblogging for collective sense‐making during violent crises: a study of three campus shootings. J Am Soc Inf Sci. Oct 24, 2011;63(1):34-47. [doi: 10.1002/asi.21685]

34. Kim HK, Ahn J, Atkinson L, Kahlor LA. Effects of COVID-19 misinformation on information seeking, avoidance, and processing: a multicountry comparative study. Science Communication. Sep 13, 2020;42(5):586-615. [doi: 10.1177/1075547020959670]

35. Vos SC, Buckner MM. Social media messages in an emerging health crisis: tweeting bird flu. J Health Commun. Dec 31, 2016;21(3):301-308. [doi: 10.1080/10810730.2015.1064495] [Medline: 26192209]

36. Wood S, Michaelides G, Daniels K, Niven K. Uncertainty and well-being amongst homeworkers in the COVID-19 pandemic: a longitudinal study of university staff. Int J Environ Res Public Health. Aug 22, 2022;19(16):10435. [FREE Full text] [doi: 10.3390/ijerph191610435] [Medline: 36012069]

37. Longstaff PH, Yang S. Communication management and trust: their role in building resilience to "surprises" such as natural disasters, pandemic flu, and terrorism. E&S. 2008;13(1):3-3. [doi: 10.5751/es-02232-130103]

38. Reynolds B, W Seeger M. Crisis and emergency risk communication as an integrative model. J Health Commun. Feb 23, 2005;10(1):43-55. [doi: 10.1080/10810730590904571] [Medline: 15764443]

39. Fink S. Crisis Management: Planning for the Inevitable. New York, NY. AMACOM; 1986.

40. Lwin M, Lu J, Sheldenkar A, Schulz P. Strategic uses of Facebook in zika outbreak communication: implications for the crisis and emergency risk communication model. Int J Environ Res Public Health. Sep 10, 2018;15(9):1974. [FREE Full text] [doi: 10.3390/ijerph15091974] [Medline: 30201929]

41. Lwin MO, Lu J, Sheldenkar A, Cayabyab YM, Yee AZH, Smith HE. Temporal and textual analysis of social media on collective discourses during the Zika virus pandemic. BMC Public Health. May 29, 2020;20(1):804-809. [FREE Full text] [doi: 10.1186/s12889-020-08923-y] [Medline: 32471495]

42. Al-Zaman MS. Prevalence and source analysis of COVID-19 misinformation in 138 countries. IFLA Journal. Aug 27, 2021;48(1):189-204. [doi: 10.1177/03400352211041135]

43. Rajan D, Koch K, Rohrer K, Bajnoczki C, Socha A, Voss M, et al. Governance of the Covid-19 response: a call for more inclusive and transparent decision-making. BMJ Glob Health. May 05, 2020;5(5):e002655. [FREE Full text] [doi: 10.1136/bmjgh-2020-002655] [Medline: 32371570]

44. Ahmed MS, Aurpa TT, Anwar MM. Detecting sentiment dynamics and clusters of Twitter users for trending topics in COVID-19 pandemic. PLoS One. Aug 9, 2021;16(8):e0253300. [FREE Full text] [doi: 10.1371/journal.pone.0253300] [Medline: 34370730]

45. Zhao Y, Cheng S, Yu X, Xu H. Chinese public's attention to the COVID-19 epidemic on social media: observational descriptive study. J Med Internet Res. May 04, 2020;22(5):e18825. [FREE Full text] [doi: 10.2196/18825] [Medline: 32314976]

46. Featherstone JD, Zhang J. Feeling angry: the effects of vaccine misinformation and refutational messages on negative emotions and vaccination attitude. J Health Commun. Sep 01, 2020;25(9):692-702. [doi: 10.1080/10810730.2020.1838671] [Medline: 33103600]

47. Higgins-Dunn N. A dog in Hong Kong tests positive for the coronavirus, WHO officials confirm. CNBC. 2022. URL: https://www.cnbc.com/2020/02/28/a-dog-in-hong-kong-tests-positive-for-the-coronavirus-who-confirms.html [accessed 2020-02-28]

48. van der Bles AM, van der Linden S, Freeman ALJ, Spiegelhalter DJ. The effects of communicating uncertainty on public trust in facts and numbers. Proc Natl Acad Sci U S A. Apr 07, 2020;117(14):7672-7683. [FREE Full text] [doi: 10.1073/pnas.1913678117] [Medline: 32205438]

49. Zhang X, Ghorbani AA. An overview of online fake news: characterization, detection, and discussion. Information Processing & Management. Mar 2020;57(2):102025. [doi: 10.1016/j.ipm.2019.03.004]

50. Zhou C, Xiu H, Wang Y, Yu X. Characterizing the dissemination of misinformation on social media in health emergencies: an empirical study based on COVID-19. Inf Process Manag. Jul 2021;58(4):102554. [FREE Full text] [doi: 10.1016/j.ipm.2021.102554] [Medline: 36570740]

51. Liu Y, Ren C, Shi D, Li K, Zhang X. Evaluating the social value of online health information for third-party patients: is uncertainty always bad? Information Processing & Management. Sep 2020;57(5):102259. [doi: 10.1016/j.ipm.2020.102259]

52. Devlin J, Chang M, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of NAACL-HLT 2019. New York, NY. Association for Computational Linguistics; Presented at: 2019 Conference of the North American Chapter of the Association for Computational Linguistics; June 2-7, 2019, 2019;4171-4186; Minneapolis, MN. URL: https://aclanthology.org/N19-1423.pdf [doi: 10.18653/v1/n18-2]

53. Jean PA, Harispe S, Ranwez S, Bellot P, Montmain J. Uncertainty detection in natural language: a probabilistic model. In: WIMS '16: Proceedings of the 6th International Conference on Web Intelligence, Mining and Semantics. New York, NY. Association for Computing Machinery; Presented at: WIMS '16: International Conference on Web Intelligence, Mining and Semantics; June 13-15, 2016, 2016;1-10; Nîmes, France. [doi: 10.1145/2912845.2912873]

54. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. Advances in Neural Information Processing Systems. 2019:8037.

55. Loshchilov I. Decoupled Weight Decay Regularization. Presented at: International Conference on Learning Representations; 2018, 2017; online.

XSL•FO
RenderX

56.   Pennington J, Socher R, Manning C. GloVe: Global Vectors for Word Representation. Presented at: Proceedings of the
       Conference on Empirical Methods in Natural Language Processing (EMNLP); 2014, 2014; Doha, Qatar. [doi:
       10.3115/v1/d14-1162]

57.   Hartigan JA, Wong MA. Algorithm AS 136: a K-means clustering algorithm. Applied Statistics. 1979;28(1):100. [doi:
       10.2307/2346830]

58.   Kim J, Aum J, Lee S, Jang Y, Park E, Choi D. FibVID: comprehensive fake news diffusion dataset during the COVID-19
       period. Telemat Inform. Nov 2021;64:101688. [FREE Full text] [doi: 10.1016/j.tele.2021.101688] [Medline: 36567815]

59.   Memon S, Carley K. Characterizing COVID-19 misinformation communities using a novel twitter dataset. In: Proceedings
       of the CIKM 2020 Workshops. Presented at: CIKM 2020 Workshops; October 19-20, 2020, 2020;1-9; Galway, Ireland.
       URL: https://ceur-ws.org/Vol-2699/paper40.pdf

60.   Chen E, Lerman K, Ferrara E. Tracking social media discourse about the COVID-19 pandemic: development of a public
       coronavirus twitter data set. JMIR Public Health Surveill. May 29, 2020;6(2):e19273. [FREE Full text] [doi: 10.2196/19273]
       [Medline: 32427106]

61.   Farkas R, Vincze V, Szarvas G, Móra G, Csirik J. Learning to detect hedges and their scope in natural language text. In:
       CoNLL '10: Shared Task: Proceedings of the Fourteenth Conference on Computational Natural Language Learnin. New
       York, NY. Association for Computational Linguistics; Presented at: CoNLL '10: Shared Task: The Fourteenth Conference
       on Computational Natural Language Learnin; July 15-16, 2010, 2010;1-12; Uppsala, Sweden. [doi: 10.3115/1596409]

62.   Gao T, Yao X, Chen D. SimCSE: simple contrastive learning of sentence embeddings. In: Proceedings of the 2021 Conference
       on Empirical Methods in Natural Language Processing. Presented at: The 2021 Conference on Empirical Methods in Natural
       Language Processing; November 7-11, 2021, 2021;6894-6910; Online and Punta Cana, Dominican Republic. [doi:
       10.18653/v1/2021.emnlp-main.552]

63.   Graves A, Schmidhuber J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures.
       Neural Netw. Jul 2005;18(5-6):602-610. [doi: 10.1016/j.neunet.2005.06.042] [Medline: 16112549]

64.   Wang Y, Jin Z, Yuan Y, Xun G, Jha K, Su L, et al. EANN: event adversarial neural networks for multi-modal fake news
       detection. In: KDD '18: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data
       Mining. New York, NY. Association for Computing Machinery; Presented at: KDD '18: The 24th ACM SIGKDD
       International Conference on Knowledge Discovery and Data Mining; August 19-23, 2018, 2018;849-857; London, UK.
       [doi: 10.1145/3219819.3219903]

65.   Zhang X, Cao L, Li X, Sheng Q, Zhong L, Shu K. Mining Dual Emotion for Fake News Detection. Presented at: WWW
       '21: Proceedings of the Web Conference 2021; 2021, 2021;3465-3476; New York, NY, United States. [doi:
       10.1145/3442381.3450004]

66.   McClish DK. Analyzing a portion of the ROC curve. Med Decis Making. 1989;9(3):190-195. [doi:
       10.1177/0272989X8900900307] [Medline: 2668680]

67.   Xiao Y, Cauberghe V, Hudders L. Moving forward: the effectiveness of online apologies framed with hope on negative
       behavioural intentions in crises. Journal of Business Research. Mar 2020;109:621-636. [doi: 10.1016/j.jbusres.2019.06.034]

68.   Brashers D. Communication and uncertainty management. Journal of Communication. 2001;51(3):477-497. [doi:
       10.1111/j.1460-2466.2001.tb02892.x]

Abbreviations

API: application programming interface
AUC: area under the curve
BERT: bidirectional encoder representations from transformers
BiLSTM: bidirectional long short-term memory
EANNT: event adversarial neural networks
EUP: Environmental Uncertainty Perception
MLP: multilayer perceptron
spAUCFPR: standardized partial area under the curve with a false-positive rate
TextCNN: convolutional neural network for text