Fake News, Real Emotions: Emotion Analysis of COVID-19 Infodemic in Weibo

Mingyu Wan, Yin Zhong, Xuefeng Gao, Yat Mei Lee, and Chu-Ren Huang

Abstract—**The proliferation of COVID-19 fake news on social media poses a severe threat to the health information ecosystem. We show that affective computing can make significant contributions to combat this infodemic. Given that fake news is often presented with emotional appeals, we propose a new perspective on the role of emotion in the attitudes, perceptions, and behaviors of the dissemination of information. We study emotions in conjunction with fake news, and explore different aspects of their interaction. To process both emotion and 'falsehood' based on the same set of data, we auto-tag emotions on existing COVID-19 fake news datasets following an established emotion taxonomy. More specifically, based on the distribution of seven basic emotions (e.g., *Happiness, Like, Fear, Sadness, Surprise, Disgust, Anger*), we find across domains and styles that COVID-19 fake news is dominated by emotions of *Fear* (e.g., of coronavirus), and *Disgust* (e.g., of social conflicts). In addition, the framing of fake news in terms of gain-versus-loss reveals a close correlation between emotions, perceptions, and collective human reactions. Our analysis confirms the significant role of emotion *Fear* in the spreading of the fake news, especially when contextualized in the loss frame. Our study points to a future direction of incorporating emotion footprints in models of automatic fake news detection, and establishes an affective computing approach to information quality in general and fake news detection in particular.**

Index Terms—**Emotion, fake news, COVID-19, infodemic, weibo, gain-versus-loss framing.**

I. Introduction

The digitization of content and the accessibility of information over the web and on social media have changed how people receive information and perceive the world. The ease of posting content online that becomes almost instantaneously accessible to web users all over the world does not allow time for either fact-checking or damage control refutation. This unprecedented open accessibility and sharability of information in human history ironically heralded the proliferation of fake news, which is falsified or reconfigured information, containing unverified, false, and often malicious content that is usually created to mislead the readers [1], [2], [3].

Logically the definition of fake news is simply the contents that are either proven to be false or cannot be verified. Fake news, on the web and in human interactions, is far more complex and often described as "purposefully crafted, sensational, emotionally charged, misleading or fabricated information that mimics the form of mainstream news" in [3]. Fake news overlaps with other two types of information distortion, i.e., misinformation (false or misleading content that is unintentionally spread) and disinformation (false or inaccurate information that is intentionally spread) [4], [5]. From this perspective, a subtler but perhaps more crucial aspect of fake news involves the roles played by the content provider, and the targeted audience, especially in terms of deliberateness, emotion, purpose, and malleability. That is, fake news is intended to affect the readers and can have serious implications for society. Last but not least, the acceleration and large scope of the spread of fake news is also one of its signatures as web-based content. This is why the spread of information over the web is often modeled as an epidemic or memetic [6], thus the effect of the spread of fake news is sometimes called an infodemic [7][1] As emotion is basically how people affect each other, a feasible, and often observed, explanation of the 'easy' appeal of fake news is its emotional content.

Emotion, typically serving as either cause of or reaction to almost all human actions and decisions [8], has been shown to play a critical role in the outbreak and proliferation of the infodemic. Fake news is often packaged in simplistic and emotional formats that usually capture attention with eye-catching titles [8], [9]. For instance, the *fear* emotion happens in response to a physically or psychologically threatening object resulting in an avoidance motivation [10]. Such mechanisms encourage our acceptance of false information without much cognitive effort. In other words, emotions are the result of meaning-making that give rise to action tendencies, and each emotion presents a core relational theme that may guide our responses [11].

[1] [83] defines 'infodemic' as "an overabundance of information some accurate and some not that makes it hard for people to find trustworthy sources and reliable guidance when they need it".

Manuscript received 10 March 2022; revised 3 June 2023; accepted 29 June 2023. Date of publication 17 July 2023; date of current version 2 September 2024. This work was supported in part by Hong Kong RGC under Grants GRF #15611021 (Lee) and PDFS2122-5H01 (Huang) and in part by the Hong Kong Polytechnic University under Grants #YWA7 (Huang), and #ZZKE (Wan, Huang). Recommended for acceptance by C. Clavel. *(Corresponding author: Mingyu Wan.)*

Mingyu Wan is with the School of Continuing Education, Hong Kong Baptist University, Kowloon Tong, Hong Kong (e-mail: pku.clara@gmail.com).

Yin Zhong is with the Center for Language Education, The Hong Kong University of Science and Technology, New Territories, Hong Kong (e-mail: lcyinzhong@ust.hk).

Xuefeng Gao, deceased, was with the Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hung Hom, Hong Kong (e-mail: xuefeng.gao@polyu.edu.hk).

Yat Mei Lee is with the Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hung Hom, Hong Kong (e-mail: ym.lee@polyu.edu.hk).

Chu-Ren Huang is with the Department of Chinese and Bilingual Study, The Hong Kong Polytechnic University, Hung Hom, Hong Kong (e-mail: churen.huang@polyu.edu.hk).

This article has supplementary downloadable material available at https://doi.org/10.1109/TAFFC.2023.3295806, provided by the authors.

Digital Object Identifier 10.1109/TAFFC.2023.3295806

In reaction to the COVID-19 pandemic, people's quick responses to fake news also demonstrate a triggering effect by certain emotions, such as *fear* and *anxiety* [12]. Emotions, therefore, are likely to influence people's attitudes toward health information and affect our judgments subconsciously. For example, *disgust* elicited by a message about fecal microbiota transplants can increase people's risk perceptions and influence their attitudes toward policy and regulation [13]. *Fear* and *anger* toward videos from the Discovery Channel's Shark Week have also been found to drive shark conservation behaviors [14].

As such, analyzing emotion becomes a crucial issue for combating infodemic, another hot topic of affective computing [15], [16], [17]. The synergizing of these two issues, has the potential of providing important and practical insights for a wide range of affective computing practices in AI and Natural Language Processing (NLP) communities. Up to date, however these two issues are typically treated as two separate tasks, i.e., emotion analysis [18] and fake news detection [2], [5], [19], with a limited intertwining of the two as one integral system. This paper postulates the significance of the affective status of people in triggering the transmission of fake news and proposes to investigate fake news through the lens of emotions.

Given the characteristics of fake news as described above and given the typical context of impactful events for fake news, a key issue would be how affective content is leveraged for the expected outcomes, i.e., how did fake news change people's decisions and/or actions in the face of the threat of a pandemic. To account for this causal relation, we adopt one of the most influential theories of human action and decision-making: Prospect Theory [20]). The prospect theory shows that, with an identical set of facts, different decisions can be made depending on the Gain versus Loss framing of the issue and facts. For instance, a gambler or a stock investor is more likely to decide to 'raise' or 'buy' given a Gain Frame that focuses on maximizing the gain. On the other hand, a 'bail' or 'sell' decision will be reached with the same facts given a Loss frame that focuses on minimizing the loss. In accordance with the theory and the human behaviors, we hypothesize that fake news is introduced to effect certain behavior or behavioral changes with the facilitation of Gain versus Loss framing. Hot pandemic topics such as facemasks and other PPEs, or vaccinations are clear issues of pertaining collective human behaviors and hotly debated among people holding different perspectives. Recall that emotion is generally defined as the aroused state or reaction given actual or perceived external stimuli. As such, emotions must be involved in the collective human reactions and decisions in a severe global health crisis that is highlighted by threats and insecurity. Hence we further hypothesize that the framing interacts with the emotions aroused in the processing of decision making and action taking.

In this work, we refer to the literature on emotional framing effects [21], [22], to understand how different frames can elicit different emotional states by triggering different cognitive evaluations of events.

In particular, we assess how framing the pandemic in terms of gains (e.g., non-infection or better recovery) versus losses (e.g., infection or death) can affect the emotional states, and eventually action, of information receivers. To better illustrate the relations among the key concepts in this work, we use Fig. 1 to represent the core schema of this paper. It categorizes COVID-19 fake news into topics of gains and losses, which correspond to different stages of COVID-19 epidemiology in parallel. At the same time, we study the emotion distributions and correlations with the different frames.

Our paper mainly focuses on the emotion analysis for fake news, and deals with several issues regarding emotion taxonomy that are specific to gain-versus-loss framing, and fake news topic modeling. The major emotion information (type and valence) are labeled using lexicon-based resource and emotion categories following the Aristotelian paradigm but most specifically cnsenti [23].

We adopt the expanded 7 emotion types that have been widely followed in recent literature on Chinese emotion analysis. They are: 好 'Like,' 乐 'Happiness,' 哀 'Sadness,' 怒 'Anger,' 惧 'Fear,' 惧 'Disgust,' and 惊 'Surprise'. Emotion Valence represents the intensity of emotion in terms of sentiment polarity: neg(ative), pos(itive), which is based on the sentiment lexicon in HowNet [24]. There are two strong motivations for adopting this seven-category system. First, the seven emotions (i.e., 七情, seven-emotion) is a tenet of Chinese culture that can be traced back as early as the Book of Rites (ca. 2nd century BC). There are several variations adopted in the history, but all versions include 爱ai 'love/like', and 恶wu 'dislike/disgust'. Second, although we are not aware of a modern taxonomy that contains the above pair, 'Love/like' is one of the 6 main classes of emotions according to Shaver et al. [25]. Plutchik [26], on the other hand, has 8 categories that include 'disgust' and 'anticipation,' 'trust', but not 'like/love'. Noting that Since these two can be viewed as polarity opposites, not unlike Happiness/Sadness, inclusion both in essence do not differ significantly from the Shaver or Plutchik systems. Note also the use of 好hao 'like' instead of 爱ai 'love/like' in the recent studies can be viewed is simply an issue of strength of the emotion, similar to love/like in English and adorer/aimer in French.

The two main research questions we will address are:

- What is the distribution of the basic emotions in COVID-19 fake news? Which emotions are most closely associated with fake news?
- Is fake news typically put in a Gain or a Loss frame, or both? Do the selection of the Gain/Loss and the expression of a dominant emotion correlate?

By addressing these two research questions we aim to establish or update the following hypothesis that would lay a foundation for future affecting computing approaches to the process of fake news.

- Hypothesis One: Fake news typically relies on the expression of one or two specific emotions to mislead. (And, if multiple emotions are involved, each emotion would be associated with either a Gain Frame or a Loss Frame.)
- Hypothesis Two: Whether fake news is framed as a Gain or a Loss Frame is according to the type of behavior of the readers that the creators want to influence. Each frame will strongly correlate with one (or two) emotions as crucial framing devices.

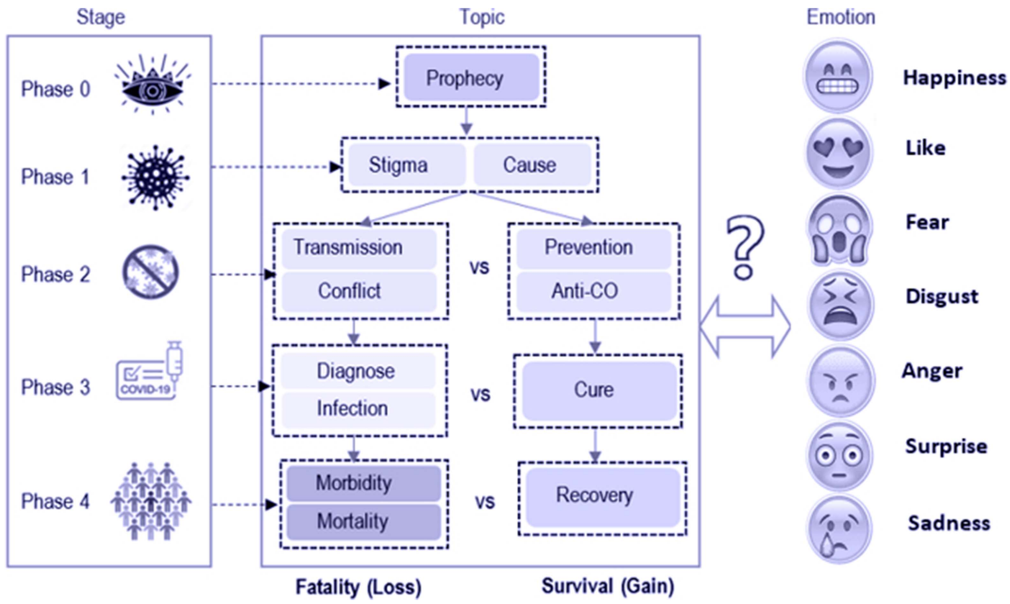

Fig. 1. Gain-versus-loss framed fake news topics intertwined with emotions.

II. Related Work

Our research questions and hypotheses focus on the correlations among COVID-19 framing, emotion, and the changes in collective social behaviors. We will focus on the literature on fake news studies, emotion analysis, and emotion for fake news detection, in this section.

A. Fake News Studies

High-stake fake news may have severe consequences for society and its people. For instance, the health fake news that ammonium chloride is effective for COVID-19 treatment led to a significant number of people ingesting ammonium chloride, an industrial material that can potentially harm human bodies [2]. Identifying fake news as well as its telltale features are, therefore, important NLP tasks with social implications. The recent surge of NLP research on fake news covers a wide range of topics, such as fake news detection [27], fact-checking [28], the spread of health-related fake news during the COVID-19 pandemic [29], factors that influence the dissemination of fake news [30], so on and so forth. [31] consolidated recent fake news research articles and suggested an intersection research line by integrating computational linguistics and networks.

Some state-of-the-art studies that are relevant to the current research include [32]'s study of the COVID-19 infodemic by fake news detection. The main data are tweets and the methodology involves classifier vote ensembles formed by base classifiers SMO, Voted Perceptron, Liblinear, Reptree, and Decision Stump. A set of 81,456 bag-of-words were input to the model for prediction, which encapsulated 2964 COVID-19 tweet instances and 3,169 extracted numeric vector attributes. Alternatively, [33] introduced a Markov-inspired computational method to identify the "fake news" trends in Twitter accounts. Clusters of words were employed to identify the topics that emerged in the investigated period. Lastly, [34] proposed a transformer-based language model [35] for fake news detection using RoBERTa [36] and domain-specific model CT-BERT [37], which were fused by one multiple-layer perception to integrate fine-grained and high-level specific representations[2]

Fake news detection remains a challenging task due to its complex nature [38]. For instance, neither human intelligence nor fact-checking programs can possess or have access to all relevant facts needed for checking. In addition, different perspectives may lead to different evaluations of facts [39], [40]. As such, current fact-checking technologies could not identify more subtle fact news that cannot be detected by truth-value alone, such as those driven by emotions, distrust, cognitive biases, racism, and xenophobia. These factors make individuals more vulnerable to certain types of fake news and also may lead to failure in future correction attempts among these affected receivers. In this study, we explore and investigate additional measures beyond fact-checking that may help to mitigate the effects of fake news in the current infodemic.

B. Emotion Analysis

Emotion analysis is sometimes considered analogous to applications of opinion mining and sentiment analysis from a psycho-linguistics perspective that evaluates the relationship between psychological processes and linguistic behaviors. In contrast to opinion mining and sentiment analysis where polarity is the focus, emotion analysis associates text with a taxonomy of emotions, which are treated as a set of psychological models as determined by multiple dimensions, e.g., valence, dominance, activation, control, etc.

[2]These representations are denoting more subcategories of information which are known as containing more in-depth information and are more informative and fine-grained.

Classical NLP studies of emotion analyses adopted automatic classification techniques, mostly based on lexicons and machine learning algorithms, to extract the psychological associations between words and emotions [41]. This line of methodology, also known as emotion classification, predicts emotion categories of emotive texts automatically. To facilitate the emotion classification task, different emotion lexica were developed. For example, EmoLex [42] contains terms that can be used to annotate documents with the emotions expressed such as Anger and Sadness. [43] constructed normative data of 372 Chinese emotion words based on the subjective ratings from native speakers of Chinese. Moreover, [44] built a Chinese valence-arousal resource, named "Chinese EmoBank," to support dimensional emotion classification. Other researchers [45] focused on emotional reactions (*Love, Joy, Surprise, Sadness, Anger*) triggered by news posts and comments on social media.

Current psychological theories have gone beyond the classical discreet categorical model of human emotions. Instead, emotions are modeled with a set of features that map emotion in a two or three-dimensional space. For instance, Plutchik's 3-dimensional "wheel of emotions," [26] contains primary emotions,[3] intensity dimensions,[4] and compound emotions. Among them, compound emotions are analyzed as being the composition of two primary emotions. All non-primary emotions are analyzed as a combination of the primary emotion. Turner [46] extended the model with a more detailed account of the compound emotions. He defined "secondary dyad" (mixed emotions that are one step apart on the wheel) and "tertiary dyad" (emotions that are two steps apart on the wheel) emotions. In the Chinese emotion literature, a commonly adopted emotion taxonomy "Turner-Plutchik" includes *Happiness*, *Sadness*, *Fear*, *Anger*, and *Surprise* [47]. This refined emotion taxonomy, incorporated with emotion words in Chinese [48], was later utilized to detect emotion causes [49], explore the relationship between emotion-causing event types and factivity [50], and apply it to the study of metaphors and emotion causes [51].

What we can learn from modern cognitive models of emotion is that a complex emotion can contain two or more basic emotions, and in some cases, a complex emotion can combine with a basic emotion to form another complex emotion. Hence emotion categories, with the possible exception of the five basic emotions, are not discreet categories [52]. For English, the compositional nature of complex emotions is vividly visualized based on an aggregation of the emotion-annotated corpus by PyPlutchik [53]. A multi-feature model fits well with the vector space representation and machine learning approaches in recent advances in the fields of machine learning and NLP. They improve bag-of-words techniques by supervised machine learning algorithms over a complex set of features, such as n-grams, word embeddings, and affect lexicons [15]. Machine learning techniques are then able to categorize and predict the appropriate emotion category for text. Recently, many state-of-the-art methods have utilized pre-trained word embeddings to extract features using unsupervised machine learning [54]. Through these embeddings, words can be projected into a space such that they are represented as a function of their context words.

[3]Other literature on emotion taxonomy: [55] classified human emotions into eight primary emotions: *Surprise*, *Interest*, *Joy*, *Rage*, *Fear*, *Disgust*, *Shame*, and *Anguish*. [56] put forward another primary emotion model with six emotions, namely, *Anger*, *Disgust*, *Fear*, *Happiness*, *Sadness*, and *Surprise*.

[4]The intensity dimension is the degree of a certain emotion. It is one of the three dimensions of emotion in [26].

C. Emotions for Fake News

Fake news and emotions are closely intertwined. Successful fake news is reported to be able to trigger intense emotions in an attempt to be disseminated in social networks [45]; emotional appeal thus is considered one crucial factor distinguishing fake and real news [9], [57]. For instance, fake news titles were found to contain much higher and more negative emotional content than those in real news; content in the fake news also showed more negative emotion types such as *Anger* and *Disgust* than *Joy* [57]. In a similar vein, fake news inspired a high probability and intensity of more negative emotions in their responses (e.g., *Fear*, *Disgust*, and *Surprise*), while true news elicited more positive emotions in the replies [58]. In addition, [9] found falsehood more novel than truth and diffuses significantly farther, faster, deeper, and more broadly in all categories of information due to emotional appeals. In particular, false rumors trigger *Fear*, *Disgust* and *Surprise* in their replies, whereas true rumors trigger *Joy*, *Sadness*, *Trust* and *Anticipation*. It is discussed that the degree of novelty and the emotional reactions of recipients may be responsible for such differences. The most recent comprehensive study of fake news spreaders found that Fear and Sadness are the most commonly attested emotions [59]. Note that although versions of the Aristotelian basic emotions are adopted almost in consensus, there is still no consensus on the correlation between emotion and fake news.

Most current fake news detection studies assume that the emotion and fake news correlation is either a downstream task or a problem of emotion and sentiment analysis, showing limited awareness of the coordination of the two lines of research. Only a handful of pioneer studies attempted to utilize emotion analysis and the relevant features to improve existing machine learning architectures in detecting fake news. Some recent works [60], [61], [62], [63] incorporated emotions and other affective information to address the detection of fake news, conspiracy theories and disinformation respectively. In general, these studies added affective information as features to a current machine learning algorithm.

In other words, they typically involved the creation of emotion vectors to augment in-system lexical features and machine learning. Their common objective is to build systems that are capable of understanding the patterns of deceiving information flow to inform and educate the user [64]. For instance, [45] proposed the EmoCred approach, an LSTM-based model that combines information from the claims' text with emotional signals (emotion lexicon and intensity) for credibility assessment.

In sum, while there is consensus on the effectiveness of leveraging emotion and affective detection systems for studying fake news, the role of emotional appeals and responses to the COVID-19 infodemic has not yet been fully explored, especially

TABLE I
INFORMATION ABOUT THE FOUR SUB-DATASETS IN THIS STUDY

Sub-corpus	Domain	Style	Source Data	Size
CovClaim	COVID-19	Headline	CrossFake, Infodemic2019, CHECKED	2,429
CovBlog	COVID-19	Post	CHECKED	2,094
GenClaim	General	Headline	WeFEND-AAAI20	20,727
GenBlog	General	Post	EANN-KDD18	8,973

in Chinese. In addition, the identification of smaller text claims, such as social media posts, have not received the same amount of coverage as other forms of fake news (i.e., propaganda, falsified news articles, etc.), particularly regarding COVID-19.

III. Materials and Methods

A. Data

Given that to the best of our knowledge, there is no available dataset that accommodates both emotion tags and fake news (such as COVID-19 news that fails to fact-check), we choose to add emotion annotation to a dataset of fake news. This approach is chosen because the alternative approach of marking fake news from an emotion-annotated corpus would not generate enough reliable contexts of emotion versus fake news interaction. This is because existing emotion datasets are generally not designed to include or highlight fake news [43], [65], [66], [67]

The data used for this study was curated from a wide collection of fake news datasets extracted from Weibo [68], [69], [70], [71], [73]. We group these datasets into four categories according to whether it is related to COVID-19 or not, and whether the news is short headlines or microblog posts. That is, CovClaim contains short headlines regarding COVID-19; CovBlog contains microblog posts regarding COVID-19; GenClaim contains short headlines of general texts; GenBlog contains microblog posts of general texts (listed as 'Sub-corpus' in Table I). The four groups of data can be used to support the comparative studies attesting to the domain-style variances of emotions in fake news. Details about the data in terms of domain, style, source data, and size (by items) are provided in Table I.

Note that all the target datasets are from Weibo for data homogeneity and to avoid compounding conditions such as genres and domains. To normalize the data, we randomly extracted 2,000 (1,000 FALSE and TRUE each) samples from the four groups of datasets to conduct data annotation and comparative analysis.

B. Emotion Annotation

The annotation of emotion in the fake news datasets is conducted automatically and supported by existing NLP tools and emotion lexicons. The quality of the first round of annotation is then enhanced by rule-based validations and automatic weightings on emotion vectors. Algorithm 1 outlines the major procedure of emotion annotations.

We adopt six tag system for each emotion event, expanding and elaborating the framework first proposed by [74]. The six tags, i.e., $\vec{ty}$(pe), $\vec{v}$(alence), $\vec{s}$(ignals), $\vec{c}$(ause), $\vec{o}$(wner), $\vec{ta}$(rget), are each assigned values that are translated into vectors. This way, each emotion event is tagged and represented by a vector space with six vectors, which in term contains all the essential

Algorithm 1: Emotion Annotation.

Input: dataset = $[m_0,\dots, m_n]$
Output: emotion_vectors = $[m_0 : v_0,\dots, m_n : v_n]$,
where $v_i = [\vec{ty}$(pe), $\vec{v}$(alence), $\vec{s}$(ignals), $\vec{c}$(ause), $\vec{o}$(wner), $\vec{ta}$(rget)], and for $0 \leq$ k ≤ 5 type (v_i) is float vector
d1 = de_duplicate(dataset);
out = { }
dst = { }
I. Data Pre-processing
for m_i in d1 **do**
if len$(m_i) > 5$ and last_char(m) != '?':
dst[m_i]['sent'] = sent_seg(m_i)
dst[m_i]['sent']['tok'] = { }
for tok in tokenize(dst[m_i]['sent']):
dst[m_i]['sent']['tok'].append(tok)
II. Emotion Type Tagging
for tok in dst[m_i]['sent'] **do**
if tok in emotion dictionary:
signal $\leftarrow$ token.mapped
$\vec{s} \leftarrow$ {emotion_vectors[m_i]['sent'][signal]}
if tok in valence dictionary:
valence $\leftarrow$ token.mapped
$\vec{v} \leftarrow$ {emotion_vectors[m_i]['sent'][valence]}
end for
$\vec{ty}$ = auto_weight($\vec{s}, \vec{v}$) for tok in emotion_vectors[m_i]['sent'] according to TF-IDF of signal and valence in $\vec{s}$ and $\vec{v}$
III. Emotion Participant Tagging
P=ner(m_i), ner: name entity recognition
for x in P **do**
if x in **emo_exp**:
$\vec{o} \leftarrow$ {emotion_vectors[P]}
else $\vec{o} \leftarrow$ {emotion_vectors['public']}
if x in **emo_pro**
$\vec{ta} \leftarrow$ {emotion_vectors[P]}
$\vec{ta} \leftarrow$ {emotion_vectors['null']}
end for
IV. Emotion Cause Tagging
C=BIO(m_i) pre-trained in RECCON [76]
for chunk in m_i **do**
if chunk in C:
$\vec{c} \leftarrow$ {emotion_vectors[chunk]}
end for
end for
out = array(emotion_vectors) for each $m_i : v_i$
refine out with auto-filling missing values
return out

information we need to study the correlation between emotion and fake news framing. More details about these tags are elaborated in Sections III-B1 and III-B2.

Data pre-processing includes filtering and segmentation. Data filtering excludes claims that are too short (less than five words), or cannot be converted to a true-false assertion statement, such as questions. These are filtered for the obvious reason of not being able to be fact-checked. Data segmentation takes a sentence and segments it into tokens, using NLTK (the nltk.tokenize package) [75].

The first step of data processing is a word-level annotation of emotion vector space and valence for each emotion-signaling word based on the augmented emotion lexicon and the polarity lexicon. Each m_i is then assigned a dominant emotion type based on the auto-weighting of the combined emotion vector spaces of all words. The auto-weighting is incorporated mainly to pinpoint the key signal word and provides a higher weight for that signal word in the process of determining the emotion type of the entire sentence. The value assigned by the automatic weighting is calculated based on a pre-labelled set of homogeneous data by a standard TF-IDF (Term Frequency-Inverse Document Frequency) algorithm, The TF-IDF based automatic weighting gives higher weight to important words, defined as those with higher frequency (TF), and appear in fewer text (IDF, i.e., more topical). Our preliminary experiment shows that automatic weighting improved performance.

Third, the data is processed with name entity recognition and then fed into a pre-trained emotion experiencer model (emo_exp) and an emotion projectee model (emo_pro) based on the keyframe arguments of the Emotion Class in Mandarin VerbNet.[5] Besides, covert name entities were labeled 'public' for the $\vec{o}$(wner) vector, and 'null' for the $\vec{ta}$(rget) vector. Covert name entities are implicit name entities that are especially prevalent in Chinese when a pronoun is used instead of a previously mentioned person's name, or when a paraphrase is used to refer to a particular individual or place. The semantic roles are still effective even when name entities are covert, so marking them is also crucial for identifying the emotion participate roles.

Lastly, the emotion cause for each emotion is tagged using the BIO (Beginning-Inside-Outside) sequence labeling pre-trained in RECCON [76]. Each emotion cause for each the claim is also automatically labeled with a vector space (based on averaged word representation). We use an averaged vector space to represent the emotion cause. A similar cause of certain emotion will show a closer distance in the vector space. Finally, all the vectors represent the concatenated emotion embedding for each sentence, which constitutes an emotion array for the whole data.

1) Emotion Types, Valence and Signals: The vector space information derived from annotation and utilized in this study includes several important components that have been discussed extensively in cognitive psychology and linguistic theories of emotion. In addition to the commonly adopted concept of emotion type, there are also emotion valence [77], and emotion signal. They are explicated below in addition to showing the methods used for automatic emotion annotation.

- *Emotion Type*: Two emotion types are added to the basic emotions that are a common modern interpretation of the Aristotelian emotions. That is, (to) like and surprise are added to anger, disgust, fear, happiness, and sadness. The addition of surprise is well-attested, given that it is often referred to as the sixth emotion in the literature. The addition of 'to like' does not have a clear cognitive justification but is justifiable in affective computing as it shows the effect of positive emotion. That is, instead of a pure emotion state, there is also mental activity involved. Similarly, it is important to note that 恶'Disgust' is a verb in Chinese, and is more appropriately translated as '(to) dislike.' That is, it is the effect of a negative emotion as a mental activity. Having like/dislike labels in affective computing is crucial as they explicitly mark the relation between the emotion owner and a target. Note that these 7 types of emotion taxonomy have also been widely adopted in Chinese emotion processing shared tasks, such as NLPCC2014 [78] and cnsenti [23]). These seven emotion-type labels: i.e., 好 'Like,' 乐'Happiness,' 哀'Sadness,' 怒'Anger,' 惧'Fear,' 恶'Disgust,' and 惊'Surprise', are annotated at sentence-level based on the auto-weighting of the emotion space of signal words.
- *Emotion Valence*: The general goodness/badness dichotomy of emotions can be traced back to Aristotle and has been codified in most modern theories of emotion, e.g., Plutchik, Frijda. It is not only crucial in terms of emotion theory but also instantiated in affective computing as sentiment analysis. It is reasonable to assume that such polarity may play a role in the presentation of fake news. We annotate emotion valence with the polarity of neg(ative), and pos(itive). Following cnsenti's method, the valences of these Chinese emotion words are extracted from the sentiment lexicon in HowNet.
- *Emotion Signal*: Emotion signals are generally defined as the physical signals indicating emotion states given by the body, including verbal and non-verbal ones [79]. In text-based emotion processing, we label lexical emotion expressions as emotion signals. The annotation is based on emotion lexicons. We take the emotion lexicon from the cnsenti package[6] that is developed by [3] as the core and augment the lexicon with emotion words from [78] and [43].[7] The size and coverage of these lexicons are provided in Table II.

TABLE II
EXPANDED EMOTION LEXICON SIZE

	Like	Disgust	Sadness	Happiness	Fear	Anger	Surprise
cnsenti [23]	11081	10262	2312	1967	1176	389	227
NLPCC2014 [78]	446	201	219	224	25	116	66
Lin and Yao [43]	32	35	80	18	89	37	86
all (deduplicate)	11469	10443	2544	2202	1236	525	297

Note that by the above design, the emotion type and emotion valence are sentence-level features, and the emotion signal is a lexical-level feature. To ensure that topic/domain-specific emotion signals are well covered, we extracted the high-frequency emotion signals occurring exclusively in the CovClaim, and CovBlog sub-corpora and manually checked them to be well justified. Examples of the COVID-19-specific emotion signals are provided in Appendix B, available online.

[5] http://verbnet.lt.cityu.edu.hk/##/frame/EMOTION

[6] https://github.com/thunderhit/cnsenti

[7] In Lin and Yao [43], there is no *Like* emotion. We divide the *Happiness* category into *Like* and *Happiness* according to the current study's definition.

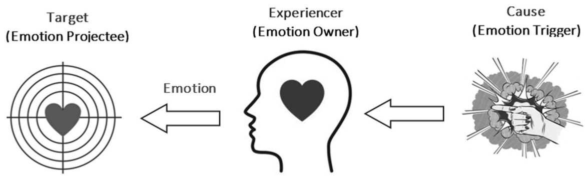

Fig. 2. The three key social elements for emotions.

2) Emotion Triple Elements: In addition to the above tags, we annotate three key elements for emotions to capture the cause-response-appeal chain essential for understanding emotions comprehensively. Note that emotions could be varied according to different participants or from different perspectives. For instance, *Happiness* emotion for Donald Trump could be *Disgust* emotion for Nancy Pelosi. Therefore, the triple-element tags can be very crucial for modeling emotion triggers and appeals. We use Fig. 2 to represent the concept of emotion triple elements in this paper. The methods of tagging such elements are briefly introduced in Algorithm 1.

We use the Example 1 for concept illustration.

E1:	震惊哈佛大学教授：新型冠状病毒诞生于人为基因改造 (Professor in Harvard astonished: Coronavirus was caused by genetic modification.
Chunk:	震惊/哈佛大学教授/ (Professor in Harvard astonished: /新型冠状病毒/诞生/于/人为基因改造 Coronavirus was caused by genetic modification.
Tag:	震惊_emo(surprise)/[哈佛大学教授]_owner/: /[(新型冠状病毒)_target/诞生/于/人为基因改造]_cause

Example 1 shows that the primary emotion in this statement is *Surprise* as signaled by the word ‘震惊’ (astonished); the emotion owner (Experiencer) is explicitly expressed as ‘哈佛大学教授’ (Professor in Harvard); the emotion is projected to the target ‘新型冠状病毒’ (COVID-19) due to ‘人为基因改造’ (Genetic Modification). With the three elements of emotions tagged, we can look into the social networks of emotions distributed in both sensational and recipient aspects, and at the same time can investigate the triggering and appeal effects.

C. Fake News Framing

We frame the fake news in the datasets with topics of gains and losses regarding COVID-19 based on the Prospect Theory [20]. Frames are generally labeled as either ’gain’ or ’loss’ (or ’shared’ if unidentifiable) based on the judgment of our model of how likely an expression is denoting a beneficial meaning to people, similar to a binary classification task based on keyword expressions. Such frames are automatically identified by modeling COVID-19 topics following the method in [80]. Minor manual validation is conducted when necessary to ensure the annotation quality. To quantify the impact of human validation, we adopt inter-rater reliability analysis between two raters, which measures the degree of agreement between human validators, using a statistical measure Cohen’s kappa coefficient (k). It takes into account the possibility that the raters may agree by chance, and provides a measure of how much agreement beyond chance exists between the two raters. Out supplementary test on 20 samples[8] shows that the agreement between the two annotators is good as kappa is 0.67.

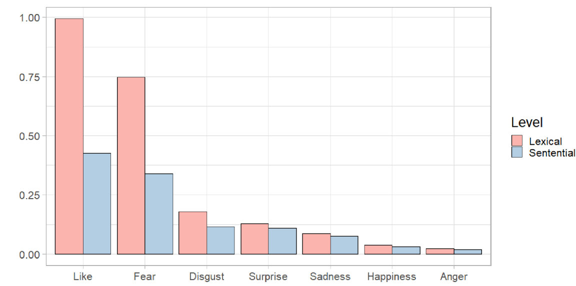

Fig. 3. Emotion distribution of lexical and sentential levels in COVID-19 fake news.

An interesting pattern is observed indicating a salient divergence regarding COVID-19 fake news, where information about survival-versus-fatality could be inferred. Such topics framed at the two ends are exemplified in Table 3, (available online) for concept illustration. The various topics also correspond to four commonly investigated stages of the COVID-19 life cycle. The entire set of concepts and the correlations in framing the COVID-19 fake news are demonstrated in Fig. 1.

IV. RESULTS

This section presents the results of the analysis addressing the issues raised in the Introduction.

A. Emotion Distributions in Fake News

We first describe the overall distribution of emotions in fake news with comparisons across datasets and then conduct the intersection study between fake news framing and emotions. For the overall distribution of the seven types of emotion in COVID-19 fake news, focusing on the CovClaim and CovBlog sub-datasets, please see Fig. 3. The Figure shows that the distribution of emotions at lexical (emotion signal) and sentential (emotion type) levels are rough in parallel. Two emotions are most often used, the emotion (to) Like emotion is the most frequent, followed by the emotion Fear. Note that the *Like* and *Fear* emotions are the primary two emotions for both word and sentence levels, which show similar distributions. That Fear associated with fake news is not only attested by a recent analysis of COVID-19 fake news [59], [72], but also by well-respected studies on rumor and disinformation [81], [82]. The inclusion of *Like* is somewhat surprising. The frequent uses may arise from either its representation of hope and/or its prominence as a mental state verb ‘to like’ in social media. This is subject to further analysis.

[8]Samples accessible via https://github.com/ClaraWan629/emotion-in-covid-19-fake-news/blob/main/supplementary_kappa.xlsx.

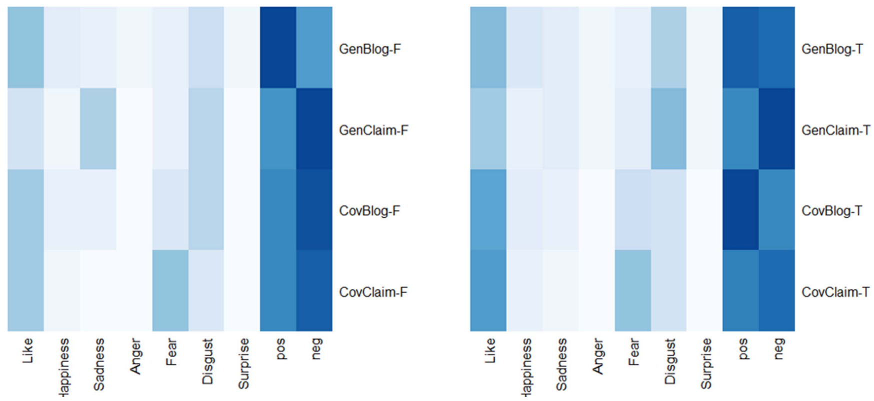

Fig. 4. Emotion Distribution (with Valence) across Datasets. Value ranges in [0, 1], with higher degrees of correlation, reprented by darker shades of blue.

B. Cross-Dataset Comparisons

In addition to the overall distribution of emotions above, we use heatmaps (correlation matrices) for cross-dataset comparison. The correlation matrix we display is a visual representation of the correlation coefficients between different emotions and data in a dataset. Correlation coefficients (with a range -1 to 1) are statistical measures that indicate the strength and direction of the linear relationship between two variables. In a correlation plot, the variables are represented by both rows and columns, and the cells of the plot contain the correlation coefficients between each pair of variables (the darker cell, the stronger correlation). A correlation plot is a useful tool in exploratory data analysis and can help identify relationships between variables that may not be immediately apparent from looking at the raw data. Two sister heatmaps are plotted based on the distribution of both emotion types and emotion valences, one for true news and one for fake news, to highlight distinctive features of the distribution of emotions in fake news. Overall, Fig. 4 visualizes the cross-dataset comparisons.

It is shown in Fig. 4 that there is a general within-dataset similarity, while there are clear variations among the datasets for either fake news or true news. Hence the most meaningful comparison should be between true and fake news of the same dataset. Overall, the generalization is that different fake news contents do not share a consistent set of emotion distribution characteristics in contrast to true news. Instead, there is a high degree of consistency in the distribution of emotion within the same data set. There are, however, a few direct contrasts that can be observed within the same datasets. This is consistent with the nature of fake news and previous studies. That is, to be effective, fake news needs to appear as similar to trustworthy news as possible. And that previous studies based on different datasets did report inconsistent, though somewhat similar results.

Based on observations of both our data and emotion distribution contrast from previous studies on fake news, we may also hypothesize that there could be a small set of subtly distinctive emotion distribution features of fake news that could be utilized to identify it and that such distributional features can be attested in some, but not all, datasets. In other words, there are a few distinctive emotion distribution characteristics that different datasets may optionally exhibit, but not all of them. For instance, Sadness is overused in fake news in the GenClaim dataset, but not in others. Negative valence is overused and positive valence is underused in CovBlog, but not in other datasets. Overall, '(to) Like' tends to appear more frequently in true news compared to fake news. This resolves the puzzle of the high percentage of 'to Like' emotion overall mentioned above, and also intuitively as 'to Like' typically marks positive situations.

If we look further into the results of the four sub-datasets, there are obvious distribution discrepancies (color gradation changes) for different text domains and styles. Therefore, we suppose that the variation of text domain and style may influence the emotion distributions. To find out to what extent these discrepancies are, we plot the normalized emotion ratios among the seven types of emotions for each dataset using the following radar plots in Figs. 5 and 6.

The figures show clearly that two emotions are manipulated in fake news. In terms of topical domains, Fear is profiled and overused (compared to trustworthy news) for false COVID-19 news, while Sadness is profiled and overused for false general news, while Disgust (to dislike) is also slightly overused for false general news. In terms of genre, both Fear and Sadness are profiled and overused in false titles/claims, and Disgust (to dislike) is also slightly overused for false bog news.

In sum, the radar plots confirmed our early hypothesis that false news would profile and manipulate certain emotions, but the emotion choices may vary in different types of texts. It is shown that Fear and Sadness are the two emotions most often profiled in fake news and that Disgust (to dislike) may also be used. It should not be surprising to find that the blogs have the most similar footprints for 'true' and false news, as it is well known that blogs are typically not fact-checked before being published.

From Fig. 5 we can see an obvious mismatch of the *Fear* and *Sadness* emotions in the false group in terms of the two domains (COVID-19 versus General). In contrast, the true group shows an obvious mismatch between the *Fear* and *Disgust* emotions. In other words, *Fear, and Sadness* are most vulnerable to topical change in fake news, while *Fear, and Disgust* are most vulnerable to topical change in true news. It also demonstrates that positive emotions (such as *Like* and *Happiness*) are consistent across the two domains. In addition, the statistics seem to suggest that

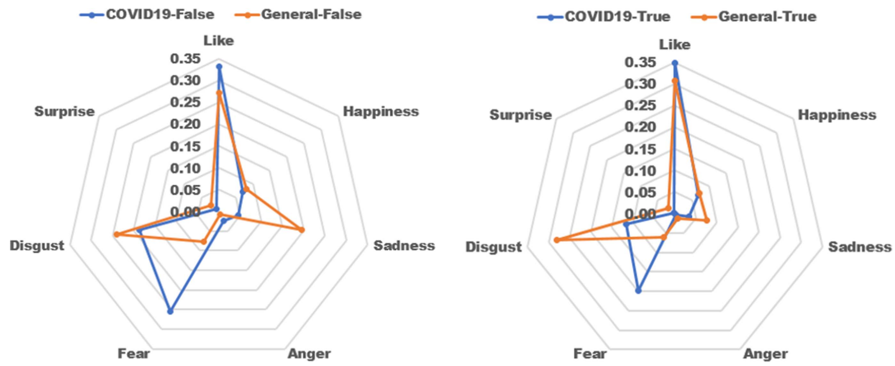

Fig. 5. Domain Comparison. Value ranges in [0, 0.5], larger the higher distribution.

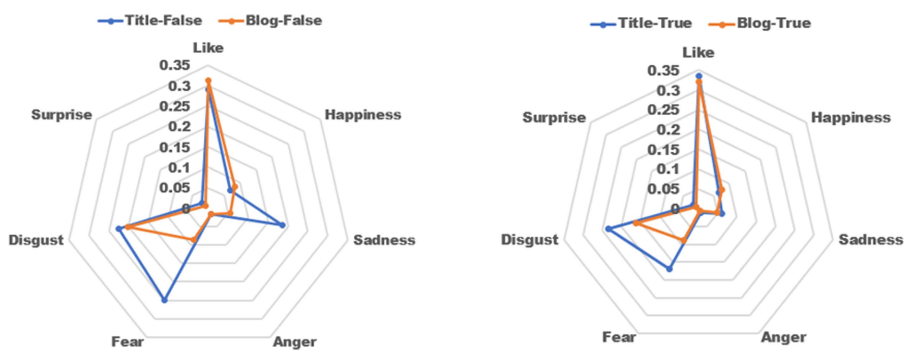

Fig. 6. Style Comparison. Value ranges in [0, 0.5], larger the higher distribution.

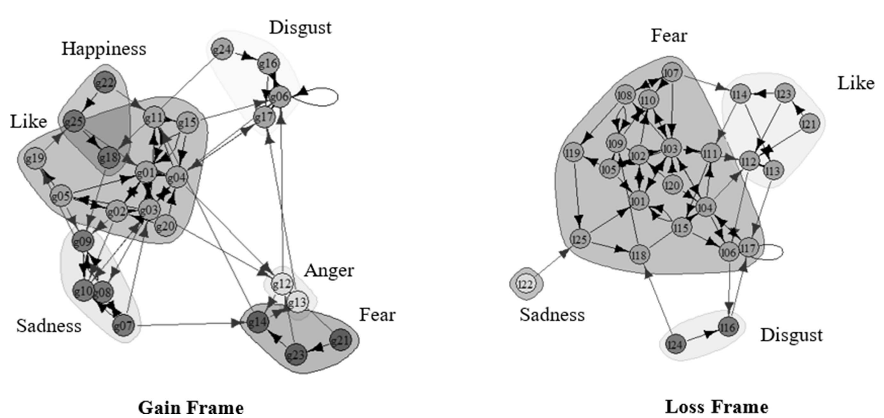

Fig. 7. Emotion triple-element clustering of gain and loss communities.

the *Fear* emotion is inherent to COVID-19 regardless of the information's credibility.

From Fig. 6 we can see that the *Fear* and *Sadness* emotions also vary the most in the false group in terms of the two styles (Title versus Blog). Similarly, the true group also shows an obvious mismatch between the *Fear* and *Disgust* emotions. It tends to show that the *Fear*, and *Sadness* emotions are vulnerable to Text-span change in fake news, while the *Fear*, and *Sadness* emotions are most vulnerable to Text-span change in true news. Again, the positive emotions are consistent across Text-span. Besides, the titles show condensed sentiment (stronger emotions) compared to entire blogs. In general, both Figs. 5 and 6 suggest the dominance of the *Fear* emotion for COVID-19 fake news regardless of the style differences.

C. Fake News Framing and its Relation to Emotions

To understand the role of gain and loss framing in fake news, we further examine fake news in terms of gain/loss topics and the associated emotion attested in the text. We plot the relations of the major emotion elements according to their gain/loss frame topics, with the nodes clustered according to their major emotion types. The result, shown in Fig. 7, demonstrates the emotion communities distribution for the gain and loss frames

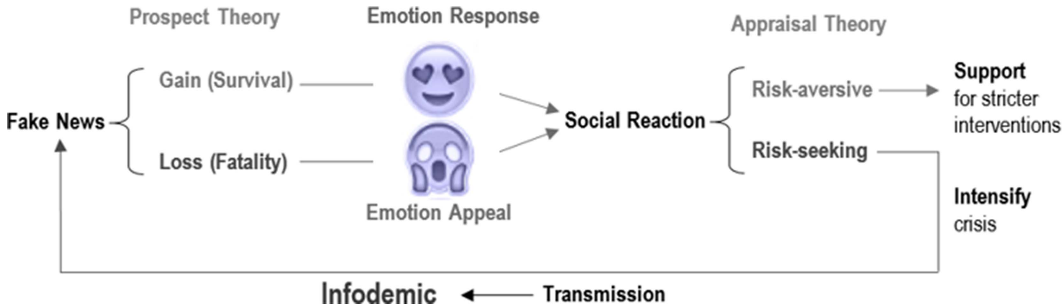

Fig. 8. A causal-action chain graph representing the major findings of this work.

respectively. In particular, the emotion experiencers and targets are taken as the nodes [marked by l(oss) or g(ain) in the initial letter], and the emotion projecting directions between the two emotion elements in the data are taken as the edges, e.g., person a is angry at event b. We can see that some nodes are shared by multiple emotions, as the same entity can receive and/or experience both positive and negative emotions. The cause element is not included because of data sparseness.

Fig. 7 shows two distinct patterns of clustering of emotion communities for the gain and loss frames. The loss frame is dominated by Fear, supplemented by Disgust and Sadness. Other than (to) Like, there is no other emotion with positive emotional valence. Note that the interpretation of the Like emotion words in the Chinese lexicon as the mental verb ‘to like,’ then basically the loss frame is driven by Fear, with other emotions feeding into it. On the other hand, the gain frame seems to be more heterogeneous in terms of emotion. That is, the gain (something ‘to like’ or happiness) is framed in contrast to emotions to avoid, such as Disgust, Sadness, and Anger. In combination with the distribution of fake versus false news, the loss frame is coherent and powerful for fake news, while the gain frame is not as effective or efficient.

V. Discussion

Results from different studies reported earlier will be synthesized with discussions in this section. The results of the distribution of emotions, the heatmaps, and the radar plots of emotions have one common feature. That is, the overall patterns of news from the same topical domain and/or genre remain stable for both trustworthy and fake news. Two possible contributing factors can be derived from common assumptions: that the same type of events shares a pattern of the footprint of emotions, and that the fake news mimics the event types of trustworthy news as much as possible to be convincing. Based on these two assumptions, texts expressing the same type of events, regardless of true or fake, would share similar patterns of emotion footprints.

In addition, given similar footprints, we also showed that fake news can often be differentiated by parochial variation(s) when compared with the common pattern of the text from the same topic and/or genre. Such variations are parochial in the sense that 1) it does not override the overall shape of the footprint of the emotions, and 2) the particular emotion involved in variations may differ according to the topic and/or genre.

Given these results, we predict that 1) types of events can be classified and/or identified by their emotion footprints and that 2) the emotion footprints of fake news can be used to differentiate it from other news from the same genre by specific emotions. That is, given distinctive spikes of specific emotions over norms in the same text types, fake news can be identified by its increased footprint of that specific emotion. For instance, COVID-19 fake news often shows an increase in the footprint of Fear and/or Sadness. Our study also shows that there may also be a secondary, and moderate footprint increase of another emotion, such as Disgust/Dislike. This explanation accounts for the similar and overlapping but inconsistent results from previous studies, as the particular emotion of increased footprint in any particular fake news varies according to text types and cannot be predicted in n certain. This unpredictability of fake news could be one of the design features of fake news, as predictable fake news would be easily exposed and would not be effective.

In addition, in Fig. 8, we use a causal-action chain graph to illustrate the differences in emotion mapping according to gain/loss framing. It shows that even though fake news can be broadly framed as gains or losses based on the Prospect Theory, the loss frame driven and dominated by Fear is the most common, and assumed to be most effective. Besides, the distinct frames interact with different emotional responses of the emotion owners, due to the intensified crisis perceived by information consumers.

Compared to existing research on using emotional signals for fake news assessment, our work may provide some unique insights into studying the interactions between emotions and COVID-19 fake news, as highlighted in Table 4, available online. In addition to showing the general distributions of emotions in COVID-19 fake news in China social media, we are one of the first studies looking in-depth into the correlation between emotion footprints and gain-versus-loss framing of COVID- 19 fake news. We observed that the COVID-19 infodemic is driven by *Fear*, which in turn triggers risk-seeking interventions in social reactions, such as social conflicts and the spreading of fake news.

VI. Conclusion

This paper reports a first attempt to capture the role of emotion in COVID-19 fake news with theory-grounded empirical evidence. It aims to establish affective computing approach to information quality in general and fake news detection in particular.

Grounded in the cognitive theory of emotions and the Prospect Theory, this study probes the interactions between emotions and COVID-19 fake news. It is shown that Fake news does not have categorical patterns of emotion, instead a specific emotion is profiled and reinforced in the context of the norm of specific event types. We found that *Fear* is the most often profiled, followed by *Sadness*. Yet, other emotions may also arise if the situation calls for it. The result of the study of emotion distributional patterns and framing both suggest that fake news detection cannot be treated as a single homogeneous target category but must be dealt with in contrast with the default baseline of similar events.

More specifically, the two research questions are addressed positively. We established that 1) Fake news as a type of texts does not have a typical signature emotion footprint vis-a'-vis non-fake-news in general. Instead, fake news typically shows a spike in negative emotions when compared with other news reporting the same event type. The most commonly involved emotions are Fear, and/or Sadness. This is consistent with the popular labeling of fake new spreaders are 'fear-mongers'. 2) Fake news can be expressed with either a Gain or Loss frame. Loss framing is shown to be strongly correlated with Fear, hence plays a critical role in fake news given the signature spike of Fear just observed. Gain-framing, on the other hand, does not seem to strongly correlate with a specific motion.

Based on these results, we confirm and refine the originally two hypotheses as the following two conclusions:

- First: Fake news typically relies on the expression of one or two specific emotions, especially Fear, to mislead. It is important to note that such emotion footprints of fake news should be establish when compared with news of the same type of even, instead of news in general.
- Second: Fake news can be presented in either a Gain or a Loss Frame. The Loss frame in particular is associated with Fear and figures prominently in Fake news. Other patterns of associations between emotions and Gain/Loss framing are likely to be emotion and event type specific.

Our findings suggest that fake news should be dealt with differently according to event types described in the text. Given a specific event type described, emotion footprints, especially the spike of a specific emotion over the norm, would be a reliable cue for fake new detection. For instance, even though blogs form a single genre, they describe a wide range of types of event. Hence it is not surprising that emotion footprints of true and fake news in blogs do not differ significantly. This result simply underlines the need to detect emotion footprints vis-à-vis targeted event types.

In sum, our study points to the rich potential of affective computing in predicting how human behaviors are effected.

Data Availability Statement

The current data is publicly available at https://github.com/ClaraWan629/emotion-in-covid-19-fake-news, and the major codes of the Python project are also provided.

Conflicts of Interest

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

[1] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD Explorations Newslett.*, vol. 19, no. 1, pp. 22–36, 2017.

[2] Q. Su, M. Wan, X. Liu, and C. R. Huang, "Motivations, methods and metrics of misinformation detection: An NLP perspective," *Natural Lang. Process. Res.*, vol. 1, pp. 1–13, 2020.

[3] M. Zimdars and K. McLeod, *Fake News: Understanding Media and Misinformation in the Digital Age*, vol. 35. Cambridge, MA, USA: MIT Press, 2020.

[4] C. Leeder, "How college students evaluate and share "fake news" stories," *Library Inf. Sci. Res.*, vol. 41, no. 3, 2019, Art. no. 100967.

[5] K. Shu, S. Wang, D. Lee, and H. Liu, Eds., *Disinformation, Misinformation, and Fake News in Social Media*, Cham, Switzerland: Springer International Publishing, 2020.

[6] M. Jiang, X. Y. Shen, K. Ahrens, and C. R. Huang, "Neologisms are epidemic: Modeling the life cycle of neologisms in China 2008–2016," *PLoS One*, vol. 16, no. 2, 2021, Art. no. e0245984.

[7] M. Wan, Q. Su, R. Xiang, and C. R. Huang, "Data-driven analytics of COVID-19 'infodemic'," *Int. J. Data Sci. Analytics*, vol. 15, pp. 313–327, 2023.

[8] S. K. Yeo and M. McKasy, "Emotion and humor as misinformation antidotes," in *Proc. Nat. Acad. Sci. USA*, vol. 118, no. 15, 2021, Art. no. e2002484118.

[9] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," *Science*, vol. 359, no. 6380, pp. 1146–1151, 2018.

[10] C. A. Smith and R. S. Lazarus, "Emotion and adaptation," *Handbook of Personality: Theory and Research*, vol. 21, 1990, pp. 609–637.

[11] R. S. Lazarus, "Progress on a cognitive-motivational-relational theory of emotion," *Amer. Psychol.*, vol. 46, no. 8, pp. 819–834, 1991.

[12] J. J. Van Bavel et al., "Using social and behavioural science to support COVID-19 pandemic response," *Nature Hum. Behav.*, vol. 4, no. 5, pp. 460–471, 2020.

[13] Y. Sun, S. K. Yeo, M. McKasy, and E. Shugart, "Disgust, need for affect, and responses to microbiome research," *Mass Commun. Soc.*, vol. 22, no. 4, pp. 508–534, 2019.

[14] J. G. Myrick and S. D. Evans, "Do PSAs take a bite out of shark week? The effects of juxtaposing environmental messages with violent images of shark attacks," *Sci. Commun.*, vol. 36, no. 5, pp. 544–569, 2014.

[15] R. Xiang et al., "Affective awareness in neural sentiment analysis," *Knowl.-Based Syst.*, vol. 226, 2021, Art. no. 107137.

[16] S. Koelstra et al., "DEAP: A database for emotion analysis, using physiological signals," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 18–31, First Quarter 2012.

[17] H. Huang et al., "EEG-based brain computer interface for emotion recognition and its application in patients with disorder of consciousness," *IEEE Trans. Affect. Comput.*, vol. 12, no. 4, pp. 832–842, Fourth Quarter 2021.

[18] K. Sailunaz and R. Alhajj, "Emotion and sentiment analysis from twitter text," *J. Comput. Sci.*, vol. 36, 2019, Art. no. 101003.

[19] X. Zhang and A. A. Ghorbani, "An overview of online fake news: Characterization, detection, and discussion," *Inf. Process. Manage.*, vol. 57, no. 2, 2020, Art. no. 102025.

[20] A. Tversky and D. Kahneman, "The framing of decisions and the psychology of choice," *Science*, vol. 211, no. 4481, pp. 453–458, 1981.

[21] J. N. Druckman and R. McDermott, "Emotion and the framing of risky choice," *Political Behav.*, vol. 30, no. 3, pp. 297–321, 2008.

[22] K. Gross and L. D'Ambrosio, "Framing emotional response," *Political Psychol.*, vol. 25, no. 1, pp. 1–29, 2008.

[23] H. Xu, H. Lin, and Y. Pan, "Construction of sentiment lexicon in ontology," *Chin. J. Inf.*, vol. 27, no. 2, pp. 180–185, 2008.

[24] Z. Dong and Q. Dong, "HowNet–A hybrid language and knowledge resource," in *Proc. Int. Conf. Natural Lang. Process. Knowl. Eng.*, 2003. pp. 820–824.

[25] P. Shaver, J. Schwartz, D. Kirson, and C. O'connor, "Emotion knowledge: Further exploration of a prototype approach," *J. Pers. Social Psychol.*, vol. 52, no. 6, pp. 1061–1086, 1987.

[26] R. Plutchik, "The nature of emotions: Human emotions have deep evolutionary roots, A fact that may explain their complexity and provide tools for clinical practice," *Amer. Scientist*, vol. 89, no. 4, pp. 344–350, 2001.

[27] Y. Long, Q. Lu, R. Xiang, M. Li, and C. R. Huang, "Fake news detection through multi-perspective speaker profiles," in *Proc. 8th Int. Joint Conf. Natural Lang. Process.*, 2017, pp. 252–256.

[28] R. Kumari, N. Ashok, T. Ghosal, and A. Ekbal, "What the fake? Probing misinformation detection standing on the shoulder of novelty and emotion," *Inf. Process. Manage.*, vol. 59, no. 1, 2022, Art. no. 102740.

[29] M. Cinelli et al., "The COVID-19 social media infodemic," *Sci. Rep.*, vol. 10, no. 1, 2020, Art. no. 16598.

[30] T. Buchanan, "Why do people spread false information online? The effects of message and viewer characteristics on self-reported likelihood of sharing social media disinformation," *PLoS One*, vol. 15, no. 10, 2020, Art. no. e0239666.

[31] G. Ruffo, A. Semeraro, A. Giachanou, and P. Rosso, "Studying fake news spreading, polarisation dynamics, and manipulation by bots: A tale of networks and language," *Comput. Sci. Rev.*, vol. 47, 2023, Art. no. 100531.

[32] T. O. Olaleye, O. T. Arogundade, A. Abayomi-Alli, and A. K. Adesemowo, "An ensemble predictive analytics of COVID-19 infodemic tweets using bag of words," in *Data Science for COVID-19*, Cambridge, MA, USA: Academic Press, 2021, pp. 365–380.

[33] W. Ceron, M. F. de-Lima-Santos, and M. G. Quiles, "Fake news agenda in the era of COVID-19: Identifying trends through fact-checking content," *Online Social Netw. Media*, vol. 21, 2020, Art. no. 100116.

[34] B. Chen et al., "Transformer-based language model fine-tuning methods for COVID-19 fake news detection," in *Proc. Int. Workshop Combating Online Hostile Posts Regional Lang. During Emerg. Situation*, Cham, Springer, 2021, pp. 83–92.

[35] K. Han, A. Xiao, E. Wu, J. Guo, C. Xu, and Y. Wang, "Transformer in transformer," in *Proc. Adv. Neural Inf. Process. Syst.*, 2021, pp. 15908–15919.

[36] Y. Liu et al., "RoBERTa: A robustly optimized BERT pretraining approach," 2019. [Online]. Available: https://arxiv.org/abs/1907.11692

[37] M. Muller, M. Salathe, and P. E. Kummervold, "COVID-Twitter-BERT: A natural language processing model to analyse COVID-19 content on Twitter," 2020. [Online]. Available: https://arxiv.org/abs/2005.07503

[38] X. Zhou, R. Zafarani, K. Shu, and H. Liu, "Fake news: Fundamental theories, detection strategies and challenges," in *Proc. 12th ACM Int. Conf. Web Search Data Mining*, 2019, pp. 836–837.

[39] W. Y. S. Chou, A. Gaysynsky, and R. C. Vanderpool, "The COVID-19 misinfodemic: Moving beyond fact-checking," *Health Educ. Behav.*, vol. 48, no. 1, pp. 9–13, 2021.

[40] P. Vossen and A. Fokkens Eds., *Creating a More Transparent Internet: The Perspective Web (Studies in Natural Language Processing)*. Cambridge, U.K.: Cambridge Univ. Press, 2022, doi: 10.1017/9781108641104.

[41] M. Giatsoglou, M. G. Vozalis, K. Diamantaras, A. Vakali, G. Sarigiannidis, and K. C. Chatzisavvas, "Sentiment analysis leveraging emotions and word embeddings," *Expert Syst. Appl.*, vol. 69, pp. 214–224, 2017.

[42] S. M. Mohammad and P. D. Turney, "Crowdsourcing a word–emotion association lexicon," *Comput. Intell.*, vol. 29, no. 3, pp. 436–465, 2013.

[43] J. Lin and Y. Yao, "Encoding emotion in chinese: A database of chinese emotion words with information of emotion type, intensity, and valence," *Lingua Sinica*, vol. 2, no. 1, 2016, Art. no. 6.

[44] L. H. Lee, J. H. Li, and L. C. Yu, "Chinese EmoBank: Building valence-arousal resources for dimensional sentiment analysis," *Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 21, no. 4, pp. 1–18, 2022.

[45] A. Giachanou, P. Rosso, and F. Crestani, "Leveraging emotional signals for credibility detection," in *Proc. 42nd Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2019, pp. 877–880.

[46] J. Turner, *On the Origins of Human Emotions: A Sociological Inquiry Into the Evolution of Human Affect*. Redwood City, CA, USA: Stanford Univ. Press, 2000.

[47] S. Y. M. Lee, *Emotion and Cause: Linguistic Theory and Computational Implementation*. Berlin, Germany: Springer, 2019.

[48] X. Xu and J. Tao, "On classification of emotion lexicons in Chinese [Hanyu qinggan xitong zhong qinggan huafen de yanjiu]," in *Proc. 1st Chin. Conf. Affect. Comput. Intell. Interact.*, 2003, pp. 199–205.

[49] S. Y. M. Lee, Y. Chen, C. R. Huang, and S. S. Li, "Detecting emotion causes with a linguistic rule-based approach," *Comput. Intell.*, vol. 29, no. 3, pp. 390–416, 2013.

[50] X. F. Gao, C. R. Huang, and S. Y. M. Lee, "Epistemic marker, event type and factivity in emotion expressions," in *Proc. 33rd Pacific Asia Conf. Lang. Inf. Comput.*, Future University Hakodate, Japan, 2019, pp. 29–36.

[51] X. F. Gao, C. R. Huang, and S. Y. M. Lee, "Conceptual metaphor in emotion expressions in Mandarin Chinese," in *From Minimal Contrast to Meaning Construct: Corpus-Based, Near Synonym Driven Approaches to Chinese Lexical Semantics*, Q. Su and W. Zhan, Eds., Berlin, Germany: Springer, 2020, pp. 211–222.

[52] Y. Chen, S. Y. Lee, and C. R. Huang, "Are emotions enumerable or decomposable? And its implications for emotion processing," in *Proc. 23rd Pacific Asia Conf. Lang. Inf. Comput.*, 2009, pp. 92–100.

[53] A. Semeraro, S. Vilella, and G. Ruffo, "PyPlutchik: Visualising and comparing emotion-annotated corpora," *PLoS One*, vol. 16, no. 9, 2021, Art. no. e0256503.

[54] A. Chiorrini, C. Diamantini, A. Mircoli, and D. Potena, "Emotion and sentiment analysis of tweets using BERT," in *Proc. Int. Conf. Extending Database Technol. Workshops*, 2021. [Online]. Available: https://ceur-ws.org/Vol-2841/

[55] S. Tomkins, *Affect Imagery Consciousness: Volume I: The Positive Affects*. Berlin, Germany: Springer Publishing Company, 1962.

[56] P. Ekman, "An argument for basic emotions," *Cogn. Emotion*, vol. 6, no. 3/4, pp. 169–200, 1992.

[57] J. Paschen, "Investigating the emotional appeal of fake news using artificial intelligence and human contributions," *J. Product Brand Manage.*, vol. 29, no. 2, pp. 223–233, 2016.

[58] B. Ghanem, P. Rosso, and F. Rangel, "An emotional analysis of false information in social media and news articles," *ACM Trans. Internet Technol.*, vol. 20, pp. 1–18, 2020.

[59] I. Russo, "Sadness and fear: Classification of fake news spreaders content on Twitter," in *Proc. Work. Notes CLEF Conf. Labs Eval. Forum*, Thessaloniki, Greece, 2020. [Online]. Available: http://ceur-ws.org/Vol-2696/

[60] B. Ghanem, S. P. Ponzetto, P. Rosso, and F. Rangel, "FakeFlow: Fake news detection by modeling the flow of affective information," in *Proc. 16th Conf. Eur. Chapter Assoc. Comput. Linguistics*, 2021, pp. 679–689. [Online]. Available: https://aclanthology.org/2021.eacl-main.56/

[61] A. Giachanou, E. Ríssola, B. Ghanem, F. Crestani, and P. Rosso, "The role of personality and linguistic patterns in discriminating between fake news spreaders and fact checkers," in *Proc. 25th Int. Conf. Appl. Natural Lang. Inf. Syst.*, Springer-Verlag, 2021, pp. 181–192.

[62] A. Giachanou, P. Rosso, and F. Crestani, "The impact of emotional signals on credibility assessment," *J. Assoc. Inf. Sci. Technol.*, vol. 72, pp. 1117–1132, 2021. [Online]. Available: https://doi.org/10.1002/asi.24480

[63] A. Giachanou, B. Ghanem, and P. Rosso, "Detection of conspiracy propagators using psycho-linguistic characteristics," *J. Inf. Sci.*, vol. 49, pp. 3–17, 2021. [Online]. Available: https://doi.org/10.1177/0165551520985486

[64] A. L. Mackey, S. Gauch, and K. Labille, "Detecting fake news through emotion analysis," in *Proc. 13th Int. Conf. Inf. Process Knowl. Manage.*, Nice, France, 2021, pp. 65–71.

[65] X. Cheng, Y. Chen, B. Cheng, S. Li, and G. Zhou, "An emotion cause corpus for chinese microblogs with multiple-user structures," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 17, no. 1, pp. 1–19, 2017.

[66] J. Li and F. Ren, "Creating a chinese emotion lexicon based on corpus Ren-CECps," in *Proc. IEEE Int. Conf. Cloud Comput. Intell. Syst.*, 2011, pp. 80–84.

[67] Q. Li, "Clickbait and emotional language in fake news," 2019. [Online]. Available: https://www.ischool.utexas.edu/ml/papers/li2019-thesis.pdf

[68] C. Yang, X. Zhou, and R. Zafarani, "CHECKED: Chinese COVID-19 fake news dataset," *Social Netw. Anal. Mining*, vol. 11, no. 1, pp. 1–8, 2021.

[69] J. Du, Y. Dou, C. Xia, L. Cui, J. Ma, and P. S. Yu, "Cross-lingual COVID-19 fake news detection," 2021. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9679918

[70] Y. Wang et al., "Weak supervision for fake news detection via reinforcement learning," in *Proc. AAAI Conf. Artif. Intell.*, 2020, pp. 516–523.

[71] Y. Wang et al., "EANN: Event adversarial neural networks for multi-modal fake news detection," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2018, pp. 849–857.

[72] P. Luo, C. Wang, F. Guo, and L. Luo, "Factors affecting individual online rumor sharing behavior in the COVID-19 pandemic," *Comput. Hum. Behav.*, vol. 125, 2021, Art. no. 106968.

[73] J. Luo, R. Xue, J. Hu, and D. El Baz, "Combating the infodemic: A chinese infodemic dataset for misinformation identification," *Healthcare*, vol. 9, no. 9, 2021, pp. 1094. [Online]. Available: https://doi.org/10.3390/healthcare9091094

[74] J. Li, Y. Xu, H. Xiong, and Y. Wang, "Chinese text emotion classification based on emotion dictionary," in *Proc. IEEE 2nd Symp. Web Soc.*, 2010 pp. 170–174.

[75] S. Bird, E. Klein, and E. Loper, *Natural Language Processing With Python: Analyzing Text With the Natural Language Toolkit*. Sebastopol, CA, USA: O'Reilly Media, Inc., 2009.
[76] S. Poria et al., "Recognizing emotion cause in conversations," *Cogn. Comput.*, vol. 13, no. 5, pp. 1317–1332, 2021.
[77] N. H. Frijda, *The Emotions*. Cambridge, U.K.: Cambridge Univ. Press, 1986.
[78] M. Wang, M. Liu, S. Feng, D. Wang, and Y. Zhang, "A novel calibrated label ranking based method for multiple emotions detection in chinese microblogs," in *Proc. CCF Int. Conf. Natural Lang. Process. Chin. Comput.*, Berlin, Germany, Springer, 2014, pp. 238–250.
[79] J. J. Campos, D. Mumme, R. Kermoian, and R. G. Campos, "A functionalist perspective on the nature of emotion," *Japanese J. Res. Emotions*, vol. 2, no. 1, pp. 1–20, 1994.
[80] A. Abd-Alrazaq, D. Alhuwail, M. Househ, M. Hamdi, and Z. Shah, "Top concerns of tweeters during the COVID-19 pandemic: Infoveillance study," *J. Med. Internet Res.*, vol. 22, no. 4, 2020, Art. no. e19016.
[81] R. H. Knapp, "A psychology of rumor," *Public Opin. Quart.*, vol. 8, no. 1, pp. 22–37, 1944.
[82] R. L. Rosnow, "Rumor as communication: A contextualist approach," *J. Commun.*, vol. 38, no. 1, pp. 12–28, 1988.
[83] J. Donovan, "Here's how social media can combat the coronavirus 'infodemic'," *MIT Technol. Rev.*, vol. 17, 2020. [Online]. Available: https://www.technologyreview.com/2020/03/17/905279/facebook-twitter-social-media-infodemic-misinformation/
[84] M. Charquero-Ballester, J. G. Walter, I. A. Nissen, and A. Bechmann, "Different types of COVID-19 misinformation have different emotional valence on twitter," *Big Data Soc.*, vol. 8, no. 2, 2021. [Online]. Available: https://doi.org/10.1177/20539517211041279
[85] R. Kumari, N. Ashok, T. Ghosal, and A. Ekbal, "Misinformation detection using multitask learning with mutual learning for novelty detection and emotion recognition," *Inf. Process. Manage.*, vol. 58, no. 5, 2021, Art. no. 102631.
[86] C. Guo, J. Cao, X. Zhang, K. Shu, and M. Yu, "Exploiting emotions for fake news detection on social media," 2019, *arXiv:1903.01728*.
[87] A. Kamplean, "Influence of emotion on fake news sharing behavior: The case study from Thailand," in *Proc. Int. Telecommun. Soc.*, 2020. [Online]. Available: https://EconPapers.repec.org/RePEc:zbw:itso20:224861
[88] X. Zhang, J. Cao, X. Li, Q. Sheng, L. Zhong, and K. Shu, "Mining dual emotion for fake news detection," in *Proc. Web Conf.*, 2021, pp. 3465–3476.

Mingyu Wan is a lecturer with the School of Continuing Education, Hong Kong Baptist University. Prior to the current position, she worked with a dual role of being a visiting scholar with the School of Foreign Languages, Peking University and a postdoc with the Department of Chinese and Bilingual Studies, Hong Kong Polytechnic University. Her research interests are mainly in computational linguistics, including metaphor detection, emotion analysis and fake news detection.

Yin Zhong received the MA degree in Teaching Chinese as a Foreign Language, in 2015, and the doctorate degree in applied language sciences from the Hong Kong Polytechnic University, in 2020. She is a lecturer with the Center for Language Education of the Hong Kong University of Science and Technology. Her research mainly integrates the corpus-based method and behaviour experiments to examine the relationship between sensorimotor information and lexical representations in Chinese.

Xuefeng Gao was the PhD candidate with the Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University. His research focused on emotion expressions and conceptual metaphor using corpus-based methods. He had published two important papers in emotion studies prior to his decease, in 2020. His work was partly shaped on the basis of his theoretical framework.

Yat Mei Lee received the PhD degree in linguistics from The Hong Kong Polytechnic University, Hong Kong, in 2010. She is an associate professor with the Department of Chinese & Bilingual Studies, The Hong Kong Polytechnic University. Her current research interests include Chinese linguistics, natural language processing, and emotion analysis.

Chu-Ren Huang received the PhD degree from Cornell and honorary doctorate from Aix-Marseille. He is a chair professor of Applied Chinese Language Studies, Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University since 2008. He is widely published in Chinese, computational and corpus linguistics. He is interested in how to extract knowledge about human and social interactions from language data.