An Empirical Methodology for Detecting and Prioritizing Needs during Crisis Events

M. Janina Sarol, Ly Dinh, Rezvaneh Rezapour, Chieh-Li Chin, Pingjing Yang, Jana Diesner
University of Illinois at Urbana-Champaign, IL, USA
{mjsarol,dinh4,rezapou2,cchin6,py2,jdiesner}@illinois.edu

Abstract

In times of crisis, identifying essential needs is crucial to providing appropriate resources and services to affected entities. Social media platforms such as Twitter contain a vast amount of information about the general public's needs. However, the sparsity of information and the amount of noisy content present a challenge for practitioners to effectively identify relevant information on these platforms. This study proposes two novel methods for two needs detection tasks: 1) extracting a list of needed resources, such as masks and ventilators, and 2) detecting sentences that specify who-needs-what resources (e.g., we need testing). We evaluate our methods on a set of tweets about the COVID-19 crisis. For extracting a list of needs, we compare our results against two official lists of resources, achieving 0.64 precision. For detecting who-needs-what sentences, we compared our results against a set of 1,000 annotated tweets and achieved a 0.68 F1-score.

1 Introduction

During crises, substantial amounts of information are shared and discussed on social media (Palen and Anderson, 2016; Reuter et al., 2018). Some of these posts may contain relevant information about the needs of affected and at-risk populations (Basu et al., 2018; Dutt et al., 2019; Purohit et al., 2014). The recent COVID-19 virus outbreak is no exception; online platforms such as Twitter have been crucial means for sharing information about the impact of the outbreak (Singh et al., 2020), personal accounts from infected individuals (Jimenez-Sotomayor et al., 2020), and updates from medical professionals (Rosenberg et al., 2020). Crisis responders and practitioners have also turned to online platforms to obtain actionable information that could aid them in response planning (Vieweg et al., 2010; Zade et al., 2018). In particular, scholars in crisis informatics have provided solutions to detect relevant Twitter messages that express resource needs and availabilities related to crisis events, e.g., during the 2015 Nepal earthquake (Basu et al., 2017; Dutt et al., 2019) and the 2015 Chennai floods (Sarkar et al., 2019). This paper builds upon and extends prior literature by proposing two needs detection tasks and applying needs detection to data about the COVID-19 crisis. In particular, we (1) *extract a list of needs* by using word embeddings to identify closest terms to *needs* and *supplies* with respect to their cosine similarity, and (2) *detect who-needs-what* sentences to determine social entities who need particular resources.

This study makes two contributions. First, we propose a method for identifying and prioritizing resource needs during a crisis. Second, we present a set of heuristics to determine the social entities that need specific resources. Overall, our study provides a reliable set of methods that might help response professionals identify immediate types of needs in the general population quickly and make effective decisions accordingly.

2 Related Work

A large body of literature from the field of crisis informatics has used natural language processing and machine learning methods to extract relevant situational awareness content from large text corpora (Vieweg et al., 2010; Verma et al., 2011). One of several categories of situational awareness content is needs expressed by (affected) individuals and communities (Imran et al., 2016; Purohit et al., 2014; Varga et al., 2013; Temnikova et al., 2015). Imran, Mitra, and Castillo (2016) analyzed tweets about eight major natural disaster events and found that about 21.7% of all tweets contained crucial information about urgent needs for shelter, donations, and essential supplies, such as medical aid, clothing, food, and water. Varga and colleagues (Varga

Findings of the Association for Computational Linguistics: EMNLP 2020, pages 4102–4107
November 16 - 20, 2020.

et al., 2013) leveraged machine learning models to match tweets, indicating problems with aid being offered to minimize the waste of resources during a crisis. Similarly, Purohit and colleagues (2014) classified tweets based on requests and offers of resources, and further matched requests with offers using regular expressions. Temnikova, Castillo, and Vieweg (2015) developed a lexical resource that contained 23 categories of situational awareness, most of which are based on needs requested and resources available (e.g., clean water, shelter material), as well as services (e.g., rescue workers, relief work) to meet the needs. Basu and colleagues (2017; 2019) identified need and availability tweets, and matched the identified needs with availabilities (Basu et al., 2018). Our paper builds upon this prior work that has primarily focused on classifying need/non-need tweets. More specifically, we propose methods that identify a general overview of the needs and specify where and by whom these resources are needed.

3 Data

We collected 665,667 tweets posted between February 28, 2020 and May 8, 2020, with a maximum of 10,000 samples for each day using Crimson Hexagon[1]. Each tweet contains at least one of the following hashtags: #COVID19, #COVID-19, #coronavirusoutbreak, #WuhanCoronavirus, #2019nCoV, #CCPvirus, #coronavirus, #CoronavirusPandemic, #SARS-CoV-2, #coronavírus, #wuhanflu, #kungflu, #chineseviruscorona, #ChinaVirus19, #chinesevirus. Our sample includes only tweets from users in the United States and tweets written in English.

4 Methodology

4.1 Extracting a List of Needs

For detecting needs, we trained an embedding model on the dataset and identified the terms closest to the seed terms *needs* and *supplies* with respect to their cosine similarity. Specifically, we performed the following steps:

1. Detect phrases using AutoPhrase (Shang et al., 2018), setting the threshold for salient phrases to 0.8, and annotate dataset with phrases.

2. Split tweets into sentences and tokens using the NLTK (Loper and Bird, 2002) sentence and tweet tokenizer, respectively.

3. Run word2vec (Mikolov et al., 2013) on the tokenized sentences.

4. Select the top 100 nouns closest to the word embeddings of *needs* and *supplies*. These nouns are representative of the needed resources.

To identify nouns, we ran the NLTK part-of-speech (POS) tagger on the tweets (before phrase annotation). We considered nouns as words whose most frequent POS tag is a noun, and a phrase as a noun if its final token is a noun (e.g., *testing-capacity* is a noun as *capacity* is a noun).

4.2 Detecting Who-Needs-What Sentences

We developed a rule-based method to identify who-needs-what sentences, where *who* is an entity (noun or pronoun) and *what* is a resource or an item (noun). We leveraged the grammatical structure of sentences for this purpose by using a dependency parser to identify sentences containing this triple. We developed two simple rules to identify these types of sentences.

The first rule considers the occurrence of the word *need* as a verb (as per its POS) in a sentence. This is a straightforward application of the who-needs-what format. We identified sentences where *who* is the subject and *what* is the direct object. After identifying that *need* (or its other word forms) is used as a verb, we selected sentences where the left child of *need* in the dependency parse tree is a nominal subject (nsubj), and the right child is a direct object (dobj). Figure 1 shows an example sentence that follows this rule and its dependency parse tree. The second rule considers the use of the word need as a noun (as per its POS). Our initial data exploration identified many sentences in the form of *X is in need of Y*, where, in the dependency parse tree, the *who* and *what* are not direct children of the term *need*. The *who* is a child of a copular verb (e.g., is), which is an ancestor of *need*. The term linking the copular verb and *need* is a preposition (i.e., the copular verb is the

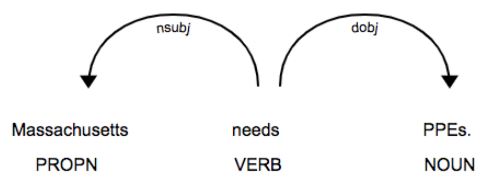

Figure 1: Rule considering need as a verb

[1]https://forsight.crimsonhexagon.com/

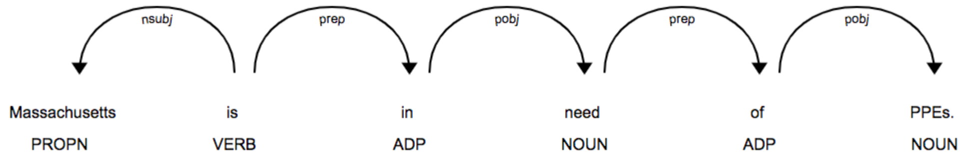

Figure 2: Rule considering need as a noun

term's parent and *need* is its prepositional object (pobj). The *what* is a descendant of need, also linked through a preposition. Figure 2 shows an example sentence that conforms to this rule and its dependency parse tree.

Similar to the first needs detection task, we used the NLTK sentence and tweet tokenizer to split the tweets into sentences and tokens, respectively. We used spaCy (Honnibal and Montani, 2017) to generate the dependency parse trees. Our source code is available on GitHub[2].

4.3 Evaluation

There is no single comprehensive list of resources needed by people in the U.S. for the COVID-19 crisis that could serve as ground truth for evaluation. We found two sets of sources that we deemed as proxies for such a list. First, the World Health Organization's (WHO) essential resource planning guidelines (2020) provide a set of forecasting tools and documents for calculating the required manpower, supplies, and equipment needed to respond to the virus adequately. Second, the U.S. Department of Health and Human Services (HHS) Office of Inspector General published the results of a survey conducted about hospitals' experiences in responding to the pandemic (Grimm, 2020). To evaluate our results for the first needs detection task, we counted the number of matches between the list we had generated and the resources mentioned in the WHO and HHS documents. This helps to capture precision. We report our results as precision@k, with k ranging from 10 to 100.

For the who-needs-what detection task, two annotators identified who-needs-what sentences from a random set of 1,000 sentences that contained any word form of *need* (i.e., need, needs, needing, and needed). Each annotator was assigned 600 sentences, where 200 sentences also appeared in the other annotator's list. Cohen's kappa was 0.91.

We report our results for the who-needs-what detection task using precision, recall, and F1-score. We compare our work to the needs detection algorithm proposed by Basu and colleagues (Basu et al., 2017), who classified need vs. non-need tweets by ranking tweets based on their cosine similarities to the embeddings of the stemmed terms *need* and *requir*. We set the cut-off value of the need-related tweets to 250 and performed the same pre-processing steps outlined in (Basu et al., 2017). While their work is focused on identifying all need tweets, it is still the closest prior work to our task.

5 Results

Table 1 shows the top 10 resources generated by our first needs detection method. The full set of results is shown in Appendix A. Comparing them to the WHO guidelines, precision@10 is 0.8, and comparing them to the HHS survey, it is 0.9. When both WHO and HHS documents are considered, the precision@10 is 1. The top 13 terms (and 19 of the top 20 terms) appear in at least either one of the WHO or HHS documents. Overall, 41 of the top 100 terms appear in the WHO guidelines, 57 in the HHS survey, and 64 in at least one document.

Figure 3 shows the precision@k, where k is in increments of 10. There is a steep drop-off in the results when the cut-off is relaxed from 20 to 30, but the precision@k decreases at a more controlled rate after this drop-off. This indicates that the resources

Resource	WHO	HHS
medical-equipment	✓	✓
equipment	✓	✓
medical-supplies	✓	✓
protective-gear	✓	✓
stockpile	✗	✓
protective-equipment	✓	✓
ppe	✓	✓
manufacturing	✗	✓
personal-protective-equipment	✓	✓
medicines	✓	✗

Table 1: Resources generated for COVID-19

[2]https://github.com/janinaj/needs_detection

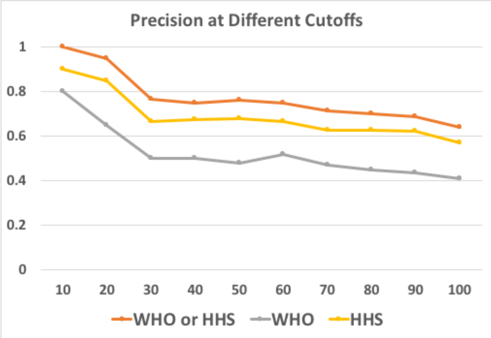

Figure 3: Precision at different cutoffs

needed still appear lower in the list. High precision scores for lower k values suggest that our proposed method can identify resources needed and produce a rigorous ranking of needs.

For the who-needs-what detection task, our method produced a precision of 0.66, recall of 0.70, and F1-score of 0.68. Sentences that were incorrectly predicted as positive examples include those of the form *if you need x, then..*, while false negatives include more complex sentences. Only using the first rule produces a precision of 0.66, recall of 0.68, and an F1-score of 0.67, indicating that most who-needs-what sentences follow this rule, where the *who* is the subject of the sentence or clause and the *what* is the direct object. Our baseline method, inspired by the work by (Basu et al., 2017), performed poorly, achieving only 0.28 precision, 0.26 recall, and 0.27 F1-score.

6 Discussion

The first needs detection results vary in terms of specificity (e.g., equipment vs. medical equipment, personal protective equipment vs. respirators, funding vs. federal funding). Several retrieved terms that are not on the WHO and HHS lists are general terms such as goods, aid, efforts, programs, and assets. In addition, several terms are synonymous (e.g., personal protective equipment and PPE). These results suggest that clustering the terms may lead to a more distinct set of results.

It is not surprising that more of the terms we detected appeared in the HHS than in the WHO document because we collected our tweet data from the U.S., and the HHS document is from a survey of U.S. hospitals, while the WHO list is for a global audience. Overall, our results suggest two findings: 1) our needs detection method works, and 2) most COVID-19 needs mentioned on Twitter are either of medical or financial nature (see Appendix A).

Our who-needs-what detection results show that a simple rule-based method can retrieve sentences that mention entities needing resources and the resources needed (0.68 F1-score). This is an interesting finding with several implications. We can produce a simple white-box method for identifying who-needs-what sentences. While deep learning may increase the scores, our method requires no training data. Another implication of our findings is that mentioning needs on Twitter often follows a specific, uniform format, which could be due to the limited characters available per tweet. Testing the generalizability of this method on other crisis events is part of our future work.

While social media has been shown to be a valuable source of information during crises, finding useful information is still akin to finding a needle in a haystack. For our who-needs-what detection task, we only found 262 positive examples. Overall, our first needs detection method can generate a ranked set of needs for 600,000+ tweets in less than 30 minutes. Running steps such as phrase detection and POS tagging in parallel may even improve this time. For the who-needs-what detection task, our method can classify 1,000 sentences in 8 seconds.

7 Conclusions and Future Work

In this paper, we presented two needs detection methods: one for extracting a list of needed resources during a crisis, and another one for detecting the who-needs-what sentences. We believe that these two methods are helpful in capturing the broad range of needs that emerges during crisis events. Specific to the COVID-19 crisis, our results suggest that the essential needs are protective equipment and financial assistance. Our methods can help detect the essential needs of the general population and affected stakeholders so they can properly plan and respond effectively.

In future work, we aim to expand our methodology to identify the availability of needs, if they have been met, and social entities who address them. In addition, we plan to differentiate between a more comprehensive set of requests, including *hopes*, *wants*, and *wishes* during a crisis.

Acknowledgments

This work was supported in part by the U.S. Department of Homeland Security under Grant Award Number 2015-ST-061-CIRC01. The views and conclusions contained in this document are those

of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security.

References

Moumita Basu, Kripabandhu Ghosh, Somenath Das, Ratnadeep Dey, Somprakash Bandyopadhyay, and Saptarshi Ghosh. 2017. Identifying post-disaster resource needs and availabilities from microblogs. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*, pages 427–430.

Moumita Basu, Anurag Shandilya, Kripabandhu Ghosh, and Saptarshi Ghosh. 2018. Automatic matching of resource needs and availabilities in microblogs for post-disaster relief. In *Companion Proceedings of the The Web Conference 2018*, pages 25–26.

Moumita Basu, Anurag Shandilya, Prannay Khosla, Kripabandhu Ghosh, and Saptarshi Ghosh. 2019. Extracting resource needs and availabilities from microblogs for aiding post-disaster relief operations. *IEEE Transactions on Computational Social Systems*, 6(3):604–618.

Ritam Dutt, Moumita Basu, Kripabandhu Ghosh, and Saptarshi Ghosh. 2019. Utilizing microblogs for assisting post-disaster relief operations via matching resource needs and availabilities. *Information Processing & Management*, 56(5):1680–1697.

Christi A Grimm. 2020. Hospital experiences responding to the covid-19 pandemic: Results of a national pulse survey march 23–27, 2020. Washington DC: Office of the Inspector General; April 3, 2020. Report no. OEI-06-20-00300.

Matthew Honnibal and Ines Montani. 2017. spaCy 2: Natural language understanding with bloom embeddings, convolutional neural networks and incremental parsing.

Muhammad Imran, Prasenjit Mitra, and Carlos Castillo. 2016. Twitter as a lifeline: Human-annotated twitter corpora for nlp of crisis-related messages. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, pages 1638–1643, Paris, France. European Language Resources Association (ELRA).

Maria Renee Jimenez-Sotomayor, Carolina Gomez-Moreno, and Enrique Soto-Perez-de Celis. 2020. Coronavirus, ageism, and twitter: An evaluation of tweets about older adults and covid-19. *Journal of the American Geriatrics Society*, 68(8):1661–1665.

Edward Loper and Steven Bird. 2002. NLTK: the natural language toolkit. *arXiv preprint cs/0205028*.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2*, pages 3111–3119.

World Health Organization. 2020. Coronavirus disease (COVID-19) technical guidance: Essential resource planning. Available at https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/covid-19-critical-items.

Leysia Palen and Kenneth M Anderson. 2016. Crisis informatics—new data for extraordinary times. *Science*, 353(6296):224–225.

Hemant Purohit, Carlos Castillo, Fernando Diaz, Amit Sheth, and Patrick Meier. 2014. Emergency-relief coordination on social media: Automatically matching resource requests and offers. *First Monday*, 19(1).

Christian Reuter, Amanda Lee Hughes, and Marc-André Kaufhold. 2018. Social media in crisis management: An evaluation and analysis of crisis informatics research. *International Journal of Human–Computer Interaction*, 34(4):280–294.

Hans Rosenberg, Shahbaz Syed, and Salim Rezaie. 2020. The twitter pandemic: The critical role of twitter in the dissemination of medical information and misinformation during the covid-19 pandemic. *Canadian Journal of Emergency Medicine*, 22(4):418–421.

Abhinav Sarkar, Swagata Roy, and Moumita Basu. 2019. Curating resource needs and availabilities from microblog during a natural disaster: A case study on the 2015 chennai floods. In *Proceedings of the India Joint International Conference on Data Science and Management of Data*, pages 338–341.

Jingbo Shang, Jialu Liu, Meng Jiang, Xiang Ren, Clare R Voss, and Jiawei Han. 2018. Automated phrase mining from massive text corpora. *IEEE Transactions on Knowledge and Data Engineering*, 30(10):1825–1837.

Lisa Singh, Shweta Bansal, Leticia Bode, Ceren Budak, Guangqing Chi, Kornraphop Kawintiranon, Colton Padden, Rebecca Vanarsdall, Emily Vraga, and Yanchen Wang. 2020. A first look at covid-19 information and misinformation sharing on twitter. *arXiv preprint arXiv:2003.13907*.

Irina P Temnikova, Carlos Castillo, and Sarah Vieweg. 2015. Emterms 1.0: A terminological resource for crisis tweets. In *Proceedings of the ISCRAM 2015 Conference*, pages 134–146.

István Varga, Motoki Sano, Kentaro Torisawa, Chikara Hashimoto, Kiyonori Ohtake, Takao Kawai, Jong-Hoon Oh, and Stijn De Saeger. 2013. Aid is out there: Looking for help from tweets during a large scale disaster. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, pages 1619–1629, Sofia, Bulgaria. Association for Computational Linguistics.

Sudha Verma, Sarah Vieweg, William J Corvey, Leysia Palen, James H Martin, Martha Palmer, Aaron Schram, and Kenneth M Anderson. 2011. Natural language processing to the rescue? extracting "situational awareness" tweets during mass emergency. In *Fifth International AAAI Conference on Weblogs and Social Media*.

Sarah Vieweg, Amanda L Hughes, Kate Starbird, and Leysia Palen. 2010. Microblogging during two natural hazards events: what twitter may contribute to situational awareness. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1079–1088.

Himanshu Zade, Kushal Shah, Vaibhavi Rangarajan, Priyanka Kshirsagar, Muhammad Imran, and Kate Starbird. 2018. From situational awareness to actionability: Towards improving the utility of social media data for crisis response. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW).

A Appendix: Resources generated for COVID-19

medical-equipment	materials	hand-sanitizer	grants
equipment	access	face-masks	relief
medical-supplies	demand	gloves	essential-workers
protective-gear	essential-goods	local-hospitals	capability
stockpile	production	respirators	groceries
protective-equipment	face-shields	healthcare-workers	devices
ppe	personnel	recipients	pharmacies
manufacturing	federal-funding	refused	flexibility
personal-protective-equipment	reagents	essential-supplies	masks
medicines	federal-assistance	barriers	living-wage
#ppe	ventilators	demands	national-stockpile
supply	systems	repairs	medical-facilities
distribution	assets	relief-funds	assistance
goods	capacity	food-banks	packages
manufacturers	programs	utilities	trace
funds	aid	meds	dpa
plans	economic-relief	testing-capacity	purchases
essentials	kits	defense-production-act	handouts
essential-items	gowns	childcare	machines
financial-relief	food	ability	deliveries
needing	funding	services	local-governments
necessities	efforts	providers	paid-sick-leave
critical-supplies	medication	requirements	shortages
clean-water	supply-chain	surgical-masks	failed
resources	facilities	expenses	hospitals

Table A1: Resources generated for COVID-19