Detection of displacement and needs related to the Russia-Ukraine conflict

Hamiz Anjum*, Arpandeep Khatua*, Alex Stratton*, Pingjing Yang, Jana Diesner

{manju2, akhatua2, als15, py2, jdiesner}@illinois.edu

Introduction

Goal:
Replicate a prior study on detecting needs during a crisis situation. We want to determine who needs what for the ongoing conflict in Ukraine using social media data.

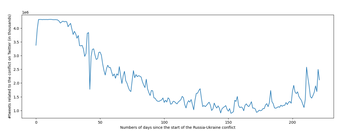
In this project, we answer 3 main questions:

- How can needs be matched to supplies? What kind of supply chain issues are revealed by the request of resources?
- How can we analyze needs over time? Does the quantity of different types of needs change as a crisis progresses? Can we predict this change based on the stage of the crisis?
- How do people react to different types of needs on Social Media? (e.g., supplies needs, financial needs, weapon needs, supply chain needs).

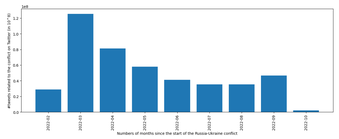
Methods

Data Collection:

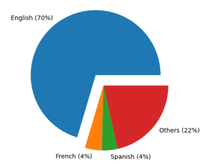
We used Twarc to rehydrate Tweets obtained from a data set containing Tweet ids. These Tweets all contained a word relating to the Russia-Ukraine conflict.



(a) Number of tweets by day



(b) Number of tweets by month



(a) Number of tweets by language

Keywords for filtering tweets:

```
'need',    'needs',
'needed',  'needing',
'want',    'wanting',
'wanted',  'requires',
'required','require'
```

(b) Keywords used for filtering

Needs detection:

In order to find texts which talked about particular needs, we extracted tweets posted geographically in and around Ukraine with keywords such as help, need, support, etc. We fed all versions of these "need" verbs in order to extract a list of types of needs that were tweeted by people in Ukraine and around the world.

References

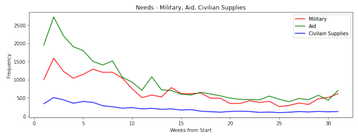
[1] Emily Chen and Emilio Ferrara. 2022. Tweets in Time of Conflict: A Public Dataset Tracking the Twitter Discourse on the War Between Ukraine and Russia. arXiv:cs.SI/2203.07488

[2] Sarol, M. J., Dinh, L., Rezapour, R., Chin, C.-L., Yang, P., amp; Diesner, J. (2020). An empirical methodology for detecting and prioritizing needs during crisis events. Findings of the Association for Computational Linguistics: EMNLP 2020. https://doi.org/10.18653/v1/2020.findings-emnlp.366

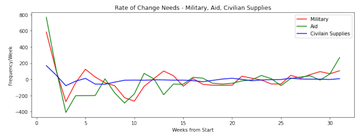
Matching Needs to Supplies

Supply Chain Issues:

From the extracted needs, observing a subset of them relating to requests for conflict aid, general support, and civilian supplies, the demand for different needs can be seen over the course of the war. Moreover, looking at how the rate of these requests changes, discrepancies in the supply chain line relating to these requests can be pinpointed. During periods of combat, for example, the request of military needs can be seen to be requested more. Major changes in the rate of these requests can signal the need to match supplies accordingly, shifting resource priority.



(a) Frequency of Needs

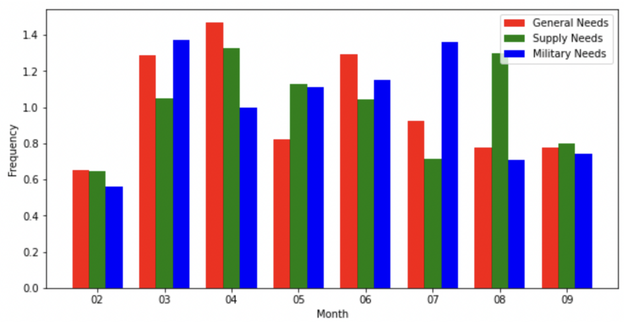


(b) Rate of Change of Needs Frequency

Temporal Analysis of Needs

Determining Frequency of Needs:

In order to obtain the quantity of each type of need, we start by part of speech tagging each word in each tweet. Then we look to see if a tweet contains a keyword, such as "needs, needing, need". If the tweet has a keyword, we find the earliest appearing noun after and return that as a need.



(a) Quantity of need type normalized by mean

We classified needs into three categories: general, supply, and military needs.

Predicting Needs:

During June, July, and August, there was heavy fighting in southeastern Ukraine. We came to the conclusion it was likely we would see an increase in military needs during this time. This was the case in July and can be seen in the graph above. Despite this, there is not a discernible pattern in the number of needs over time. Hence, we believe it is only possible to predict the number of needs based on current conflict news and not using statistical models.

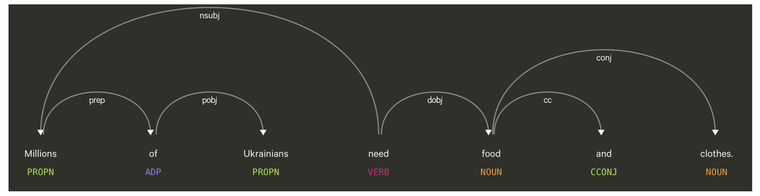
Reactions to Needs

Reaction-based needs gauging:

Let's consider a tweet $\mathcal{T}$ mentioning a set of needs that we detected in the previous step as $\mathcal{M} = m_1, m_2, \ldots, m_n$. Each tweet on Twitter is associated with some public engagement metrics; likes $\mathcal{T}(l)$ and retweets $\mathcal{T}(r)$. The importance score of each need is $\mathcal{S}(m)$.

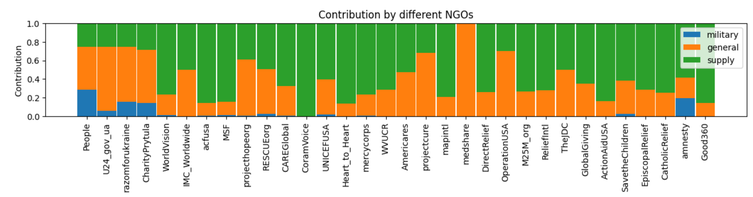
$$\mathcal{S}(m) = \sum_{t=1}^{\mathcal{T}} (0.5 * \mathcal{T}(l) + \mathcal{T}(r)) * 0.001$$

We consider the like count of a tweet to be half as important as a retweet, and we loop through every tweet in order to get the final score for each need. We add scaling to prevent overflow, and we pass it through a softmax.



(a) Tweet parsing dependency graph.

In order to find the types of needs of regular people and compare them against the contributions from various NGOs we represent this plot here. We split it into different types using these words: **Military:** guns, missiles, weapons. **General:** peace, security, support, freedom. **Supply:** money, food, fuel, camps.



(a) Contribution from different NGOs.

Future Work

Classifying Needs:

To classify needs, we created a framework to group needs based on what the corresponding tweets were about. Going forward, we would like to train and test classification models, which could possibly do this more rigorously and accurately.

Added Language Support:

At this time, the methods we used to identify quantity needs from Twitter data are unusable with languages other than English. Given the crisis, there may be significant relevant data in Russian and Ukrainian. Adding this functionality may supply us with more data and allow us to make new insights.

*Support for this project was provided by the Illinois Geometry Lab and the Department of Mathematics at the University of Illinois at Urbana-Champaign. *Denotes equal contribution.*