

Narrative Coherence and the Detection of Enemy Activity in Social Media Data: Case Study from the 2020 Azerbaijan / Armenia Conflict

Dr. Hezekiah Akiva Bacovcin / Mr. Maxime Martineau / Mr. Zach Devereaux
CANADA

abacovcin@nexalogy.com / mmartineau@nexalogy.com / zdevereaux@gmail.com

ABSTRACT

Information warfare relies on being able to identify enemy action and possible sites for offensive information activities. In a world where millions of messages are being generated by millions of people globally, purely human efforts at surveillance of the information landscape are doomed. We show, using a case study on the Azerbaijan / Armenian conflict, that by focusing on the various narratives at play, it is possible for AI-assisted human analysis to narrow down the surveillance targets to a manageable size, and identify automation deployed by antagonists in an armed conflict.

1.0 INTRODUCTION

Information warfare is fundamentally about narrative shaping. This can take the form of either: (1) tactical narrative shaping, i.e., narratives about troop placement/numbers and casualties or (2) strategic narrative shaping, i.e., narratives about validity of the war/influence on troop/civilian morale. A critical problem in the information warfare space is to be able to rapidly identify and understand enemy information activity in order to accurately respond.

Given the centrality of narratives to information warfare, *narratives cannot lie*. Therefore, we should seek enemy activity by looking carefully at narratives that support the enemy's broader agenda. By seeking enemy activity through the narratives that they promote, the enemy cannot hide the activity without undermining their own effectiveness; this is the double-edged sword of information warfare, and also applies to public diplomacy around the same operational webspace.

In this paper, we suggest a specific methodology that uses this narrative property of information warfare to identify information warfare activity. In particular, we highlight a process in which one: (1) identifies relevant narratives and narrative signals, (2) collects data in a structured way based on narrative signals, (3) investigate actors and stories within enemy narratives. In the rest of this paper, we will walk through a case study that will demonstrate a specific use of this methodology and then we will highlight some on-going technological improvements that we are working on to improve the ease and efficiency of this methodology for field use. The goal of this paper is a working methodology for the exploitation of social media for intelligence in a ubiquitous information warfare environment.

2.0 METHODOLOGY AND CASE STUDY

2.1 Azerbaijan/Armenia Conflict as Case Study

In this section, we will present a case study of using our methodology to investigate the conflict between

Armenia and Azerbaijan in October of 2020. During this conflict, military action was taken over the disputed territory of Nagorno-Karabakh situated between Armenia and Azerbaijan. While internationally recognized as part of Azerbaijan territory, the population is ethnically majority Armenian and in an initial conflict in 1994 had achieved de facto independence from Azerbaijan.

In late September 2020, the conflict reignited with both the Armenian and Azerbaijani militaries becoming involved in the conflict. Ultimately (Nov. 10, 2020), Azerbaijan was able to regain control of most of the Nagorno-Karabakh region. Our case study was conducted during the middle of the conflict (Oct. 27th, 2020) and we were tasked as part of an industry-media collaboration, to determine if there was any suspicious information activity from either side that could be better understood through artificial intelligence.

2.2 Analysis, Step 1, Identify Relevant Narratives and Narrative Signals

The first step when attempting to analyse a dataset is to dig into the data and identify the narratives at play and the signals that can be used to distinguish between the narratives. This process is rarely unidirectional. One often has to gather some broad data, analyse it, and use that initial pass to find new terms that can be used to collect a new, more pertinent set of data. This back-and-forth iterative process is the only way to guarantee that an appropriate sample of the data has been analysed; it is a step important for veracity.

For the Azerbaijan-Armenian conflict, we started by collecting messages using the following two queries:

1. First query :
 - a. Must contain: Armenia
 - b. Contain one of: Nagorno-Karabakh Fuzuli Mkrtychyan "Benik Hakobyan" "Yuri Adamyan" "Sophie Zhang" "Nathaniel Gleicher" "Rally Forge" "FARA" "ANCA" "Azerbaijan Diplomatic Academy" Baku
2. Second query :
 - a. Must contain: Azerbaijan
 - b. Contain one of: Nagorno-Karabakh Fuzuli Mkrtychyan "Benik Hakobyan" "Yuri Adamyan" "Sophie Zhang" "Nathaniel Gleicher" "Rally Forge" "FARA" "ANCA" "Azerbaijan Diplomatic Academy" Baku

By requiring the messages to contain the name of one of the participants of the conflict, we sought to do an initial gross separation of the messages into Azerbaijani and Armenian groups.

Using the data collected from these two queries, we looked at the top words/hashtags and most retweeted messages, in order to quickly get a sense of what the core of the discussion was, and what were the flashpoint narratives warranting further investigation. By engaging with this data, we identified the following pro-Armenia and pro-Azerbaijan narratives, along with hashtags that served as narrative signals:

- For pro-Armenia:
 - Azerbaijan is the aggressor (#stopazerbaijaniaggression #stopazerbaijanaggression)
 - Azerbaijan is engaging in war crimes/terrorism (#azeriwarcrimes #stopaliyevsterror)
 - All Armenia wants is peace (#peaceforarmenians)
- For pro-Azerbaijan:
 - Armenia broke the ceasefire (#armeniabrokeceasefire #armeniaviolatesceasefire)

- Karabakh is naturally part of Azerbaijan (#karabakhisazerbaijan #karabakhisazerbaijan #karabakhhisazerbaijan #freearabakh)
- The situation in Karabakh is generating a refugee problem for ethnic Azerbaijani (#1millionazerbaijanirefugees)
- Armenia is engaging in war crimes/terrorism (#armeniakillscivilians #stoparmenianterrorism #stopsupportingarmenianterror #stoparmenianterror #armeniakillschildren)
- Armenia is the aggressor (#stoparmenianaggression #stoparmenianoccupation #stoparmenianaggression #dontbelievearmenia #dontbelievearmenian #armenianoccupation #armeniaendtheoccupation)
- Azerbaijan has popular support for Karabakh (#westandwithaliyev #supportaliyev #azerbaijanisnotalone)

For Armenia, the goal was to show that Azerbaijan was illegitimately contesting the status quo (which benefited Armenia). Azerbaijan's goal was to show that conflict in Karabakh was justified based on Azerbaijani territorial and ethnic claims and that Armenia was the true source of the problem necessitating the conflict (which benefited Azerbaijan).

2.3 Analysis, Step 2, Collect Data

As shown above, each side has a number of narrative messages that are associated with their overall narrative strategy for shaping public understanding of the conflict. For each of these messages, we identified one or more hashtags associated with the message. At this stage, we took these sets of hashtags and used each set as a query to collect the sets of social media messages that contain those hashtags. Each set of documents can be associated with a particular narrative message, since we know that it contains the hashtags that serve as signals for that message. Thus, from the moment of collection, our data collection is sorted into different categories based on their narrative content, each with different information exploitation goals.

2.4 Analysis, Step 3, Investigate Actors and Stories within Enemy Narratives

The next step is to look at global summaries and the major posts within each side of the conflict (pro-Armenia/pro-Azerbaijan) to make sure that the correct narratives have been identified and that no major sub-narratives have been missed. Below is the automatically extracted AI summaries of the total dataset, which shows a balance between pro-Azerbaijan and pro-Armenia messages.

- @ani_mheryan Are the Armenians peaceful?
- Hear our voice, countries of the world.
- We have to enlighten #Armenian youth about the realities regarding history and legal status of #Karabakh
- #armenia is so insane that fired @euronews journalists before they were threatening French journalist @LiseronBoudoul #StopWarmeria #FactsAboutArmenia
- Armenia keeps killing civilians and damaging civil infrastructure
- Impose sanctions upon Turkey and Azerbaijan!

We then looked at the most retweeted messages from each side; another quick method for social media exploitation for intelligence. This was possible, because all of the messages were associated with one of the narratives identified earlier based on which query brought the data in. For pro-Azerbaijan, (as expected) we found that the messages (Fig. 1) focused on the legitimacy of Azerbaijan’s claims to Karabakh and the atrocities attributed to Armenia.

Export Retweets

Top Retweets: 10

Top Retweeted Users: 10

Format: **JSON** DOWNLOAD

Count	Tweet
231	RT @azerbaijanmfa: urgent! missile attack on #barda region by #armenia's occupation forces killed 3 civilians & seriously injured 10 people. including women and kids. another brutal v...
212	RT @shedeivrplus: thank you ilham aliyev @azpresident for all your efforts to bring peace to south caucasus and ensure the territorial integrity of the azerbaijan republic! #westandwit...
158	RT @arzujaeed: going to post this all over again until everyone gets it. #azerbaijan is fighting on its own land and we are liberating occupied historical lands. this year i'm coming back ...
145	RT @shedeivrplus: we are restoring historical justice on the battlefield! #armeniakillscivilians #prayforganjacity #karabakhisazerbaijan #westandwithaliyev https://t.co/amjevmbwfa
138	RT @shedeivrplus: #westandwithaliyev because he defends our country, our home, we are grateful @azpresident #armeniakillscivilians #karabakhisazerbaijan #stoparmenianaggressi...
128	RT @wwwmodgovaz: armenian armed forces fire rockets at #goranboy #fartar and #barda regions. #stoparmenianerror
111	RT @azgomruk: armenia vandalizes even the monuments to world war II heroes #stoparmenianaggression #stoparmenianoccupation #karabakhisazerbaijan #dontbelievearmenia #ar...
110	RT @daily_karabakh: even back in 1920 #armenia lied about the number of #muslim population in #karabakh, while #karabakh was home to 415,000 muslims they stated only 80,000...
110	RT @arzujaeed: listen to our stories and name my country. #1millionazerbajaniarefugees https://twitter.com/jahangiryoussif/status/1319293842433011713
107	RT @arzujaeed: greatest nation greater than roman empire greater than all the empires vandalizes its own monuments, why? #dontbelievearmenia #karabakhisazerbjan https://t.co/hy...
Count	Tweeter
1186	@arzujaeed
525	@shedeivrplus
463	@daily_karabakh
298	@wwwmodgovaz
266	@azerbaijanmfa
242	@azgomruk
172	@maharramshalala
138	@nazulkin
137	@emin274
123	@leman95ibra

Figure 1: Main Retweets in Pro-Azerbaijan Dataset

For pro-Armenia, (as expected) we found that the messages (Fig. 2) focused on the fact that Armenia wants peace and that the Azerbaijani are the aggressors. As noted above, this fits well with Armenia’s strategic position, given that the status quo benefits Armenia.

Export Retweets

Tip: Retweets: 10

Top Retweeted Users: 10

Format: **JSON** DOWNLOAD

Count	Tweet
1096	RT @henrikhmkh: peace for armenians #openyoureyes #world #stonsaying #starfactio #artsakhstrong #armeniatrong #peaceforarmenians https://t.co/ybioj5axw
519	RT @mfankr: on oct 26, #artsakh's fm masis maylian meets with a delegation of french mps. mps presented to minister a copy of draft resolution on recognition of the republic of artsa...
144	RT @adamsmithmd: this is mete turksoy, representative of the azerbaijani republican alternative (real) party, a political activist. in his entry, he noted that the entire civilian population of...
128	RT @angelasaralyan: we don't need war ... we are a peaceful people ... all we want is peace! let there be no wars!! how do you say war in english an/lan? "war," ... "vor?" vor che, war ...
117	RT @anca_dc: dear secretary of state pompeo @secpompeo and national security adviser o'brien @whnsc: how are you holding #azerbaijan accountable for breaking the us-negotiat...
92	RT @hraqarabaghian: beautiful sayings about turks from different nations!♥ #stopaliyev #peaceforarmenians https://t.co/euerrmwn3b
86	RT @garmirsar: i am a @nyualumni(us) and i call on #nyu to cease their business relationship with #bgrgroup, who actively continues to support the genocidal government of #azerbaij...
84	RT @adamsmithmd: @ladygaga lady gaga: as you are making your voting decision, armenians are thinking only about the candidate that will help its people to stay alive!! a word of en...
78	RT @adamsmithmd: @frankpallone @secpompeo i am surprised that our state-of-the-art tech & sats have not been able to observe which side has violated it! @deptofdefense!? @cia...
76	RT @adamsmithmd: @secpompeo azerbaijani & turkish political/military elite is solely responsible for deliberate, jihadi terrorist employment using illegal weapons against civilians killin...
Count	Tweeter
1100	@henrikhmkh
729	@adamsmithmd
523	@mfankr
339	@anca_dc
193	@annabarseghyan2
139	@agbu
128	@angelasaralyan
119	@anzhela_yan
111	@syriahuy
99	@hraqarabaghian

Figure 2: Main Retweets in Pro-Armenian Dataset

We then looked at a network visualization (Fig. 3) that grouped each of the most prolific actors in the dataset by which of the various narrative queries their messages are associated with. As can be seen in the figure, the community is divided into two groups. The group on the left is associated with the pro-Armenian narratives and the group on the right are associated with the pro-Azerbaijan narratives. This is also the first sign of some suspicious activity as there are two individuals who share both pro-Armenian and pro-Azerbaijani narratives (@Aliyeva25648403 and @TuralHasanov_) an action which is counter-intuitive given the goals of the narratives deployed on either side.



Figure 3: Network Visualization by Query

In both cases, this is driven by a peculiar error where they both include an anti-Azerbaijan hashtag in otherwise solidly anti-Armenian messages (see Fig. 4). The fact that they both make a similar error with the exact same hashtag (#StopAzerbaijaniAggression) suggests that they may be part of a shared information campaign. It also suggests that they may be using the same automated posting system that mixed up the hashtags. As such, rudimentary automation is a dangerous tool to deploy.

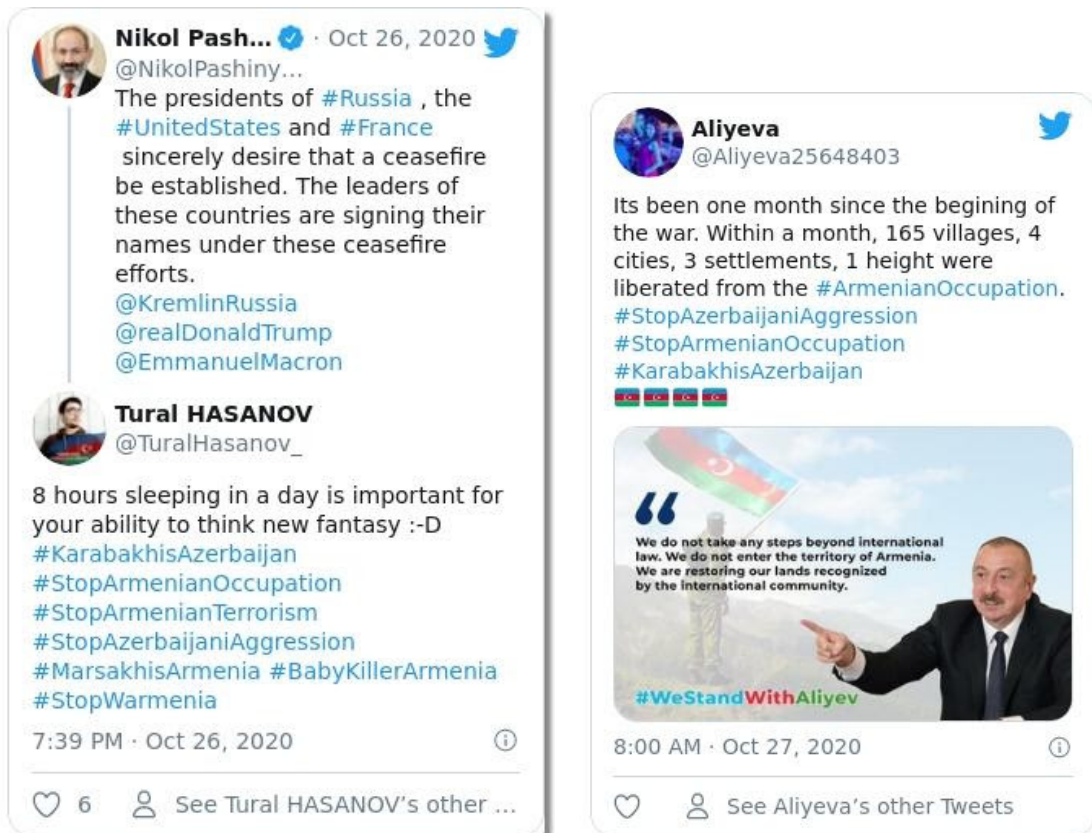


Figure 4: Example Tweets from Suspicious Accounts

Finally, in order to identify suspicious activity, we examined the most prolific agents on both sides and searched for suspicious agents. In this case, we focused particularly on agents who had just joined Twitter and had usernames that ended in a long string of numbers, as these traits are in keeping with our earlier methodology to identify automation [1]. We found that there were nine pro-Azerbaijani actors that fit this profile (@gunay93689808, @aliriaz49205348, @ramil78050199, @sakina76340523, @mn77795967625, @sahile57910429, @joel62114369, @gunel50811681, @melek18635643), all of whom produce a number of tweets with identical texts. In particular, we collected a total of 589 posts by these users, but only 16 distinct message strings. Which indicates that they are reposting the same messages over and over again. Here are an example of some of these messages:

- Despite the ceasefire, they still attack civilians. #prayforBarda #StopArmenianAggression #StopArmenianTerror #ArmeniaKillsCivilians
- We dont kill civilians but they kill #DontBelieveArmenia #Armenia IgnoresUNresolutions #stoparmenianagression

For the pro-Armenian side, there are 4 users (@jacquel73189742, @peaceon74411255, @bellaam27153020, @arthur83922116) who share 112 posts with only 8 distinct texts and again joined Twitter in the few months before the conflict.

In both cases, the identified actors are suspicious, but not definitely proven to be agents of information campaigns. However, in the course of the few hours that this methodology took, we narrowed the field from hundreds of thousands of posts by thousands of actors to a small set of posts by a limited number of suspicious actors. This targeted set is of a size that it is feasible to do more focused and human-intensive investigation and surveillance. This is in keeping with the previously published methodology to identify authentic chatter [2].

3.0 ONGOING RESEARCH

One of the most difficult tasks in this methodology is the identification of narratives and detection of narrative signals. In order to make this process more efficient and easier for people without specialized training, we are engaged in research to develop novel technologies to assist people in making these determinations. Here, we'll present a recently developed technology (clustering and summarization) that is designed to help agents quickly identify potential narrative targets for further analysis. We will then discuss future research into automated recognition of narrative signals by determining the in-cluster vs. out-cluster properties of relevant narrative targets.

In order to assist in identifying narrative clusters, we have developed new AI technology involving narrative clustering. This process clusters messages together based on their content. By applying the process iteratively, we are able to get a small number of high-level narrative groupings as well as a small number of sub-groupings. This clustering is paired with automatic summarization technology, in order to quickly label the various narrative groups based on the shared content of the associated messages.

We are currently working on using ML techniques to extract from sets of narrative clusters the signals necessary to distinguish between the messages from one narrative and the messages from another narrative. In our case study, there were a number of hashtags that clearly distinguished the different narratives, but it is not always this easy to find markers of a given narrative. By automating the process of creating narrative-aligned data collection, we will allow more subtle narratives to be tracked more efficiently and reliably.

4.0 CONCLUSIONS

Current state-of-the-art AI techniques and soon-to-be-developed features provide a means of quickly ascertaining the narratives at play in a particular context and rapidly investigating the major actors and messages in each narrative space. These include term and agent ranking, network visualization, narrative clustering and summarization and automated narrative signal detection.

Information warfare relies on being able to identify enemy action and possible sites for offensive information activities. In a world where millions of messages are being generated by millions of people globally, purely human efforts at surveillance of the information landscape are doomed. We have shown that by focusing on the various narratives at play in a particular conflict, it is possible for AI-assisted human analysis to narrow down the surveillance targets to a manageable size, and identify automation deployed by antagonists in an armed conflict.

REFERENCES

- [1] Forrester, Bruce, et al. "Propaganda Filters: Tracking Malign Foreign Interventions on Social Media," *IST-178 Specialists meeting on Big data challenges: situational awareness and decision support*, Budapest, October 2019.
- [2] Forrester, Bruce. "Authentic chatter." *Computational and Mathematical Organization Theory* 26.4 (2020): 382-411.

