Comparison of ISIS and Al-Qaeda Internal Communication Proposal

With this project, Tolu, Matthew and I intend to explore the inner workings of two prominent terrorist groups, ISIS and Al-Qaeda, and to compare their internal-facing communications. Specifically, we want to examine the letters exchanged. While we will all have this theme in mind as we create the different models and visuals, my interest specifically is in how the two organizations compare when responding to the same events. This will reveal a little bit more information about character or each group so to speak. We have found two corpuses to accomplish this, as will be discussed below. Depending on the actual content within those data, different analyses are possible. If there is little overlap in the actual discussion of a specific event between the two corpuses that we found, then I can analyze how the two groups communicate in a common time period, relating the time period to major events that may have preceded these discussions. This would still address how the groups reacted to an event, but would be less direct. If there are fortunate overlaps in the actual discussion of events, meaning the same events appear in both sets of letters directly, then even more specific analysis can be conducted. A little bit of research could help determine who was responsible for the events. Then, perhaps I could compare how each of the two groups talks about attacks they are responsible for, versus how they talk about the attacks attributed to the other group we are examining (ISIS or Al-Qaeda), versus how they respond to events that are attributable to other groups (i.e. not Al-Qaeda nor ISIS).

|  |  |  |
| --- | --- | --- |
|  | Group for Analysis of Internal Communications | |
| Attributed Attack Responsibility | ISIS | Al-Qaeda |
| The group itself | How ISIS communicates when they conduct an attack | How AQ communicates when they conduct an attack |
| The Other Organization {ISIS, AQ} | How ISIS communicates when AQ conducts an attack | How AQ communicates when ISIS conducts an attack |
| External (neither ISIS nor AQ) | How ISIS communicates when an external group conducts an attack | How AQ communicates when an external group conducts an attack |

Ultimately, my question is: How do ISIS and Al-Qaeda differ in their internal communication in response to attacks put on by their own group compared to other groups?

The raw data consist of approximately 600 documents originating from ISIS and 250 documents associated with Al-Qaeda. We found the ISIS documents from the Harmony Project at the Combating Terrorism Center (CTC) at West Point and we found the Al-Qaeda documents from the office of the Director of National Intelligence (DNI), where they had been translated from the documents found at Osama bin Laden’s Abbottabad Compound. Since we want to analyze letters only, we will filter these out before hand, using some automation and, if needed, manual inspection. Some of the letters have dates, while other do not, so dates could be helpful on a more detailed level, but there will not be enough to conduct an analysis on a broader scale. These documents are all pdf files that will be converted to text with the help of a Python package called TextMiner. We will then use the paragraphs as the documents for comparison. Since each letter has quite a few paragraphs, this should give us a decent amount of documents as a result. These corpuses are inward facing documents, which hopefully will provide a little more insight into how these organization operate and shine light on what they value. The only thing that may be a problem here is the amount of data that happens to overlap between the two groups. If there are not enough documents in the end, then a change in course would need to occur. However, until I get to that point, this is my current plan.

To address this question, there will be several methods that our group will use. Common among all of our projects will, of course, be the cleaning and introduction of the data into our framework. Some visualizations will also be nice to reference from other group members, such as Tolu’s organizational hierarchy that will be represented as a radial graph. The central methods that I intend to use will be clustering, sentiment analysis, and, most likely, force directed word networks to compare the different documents. The clustering will allow for the identification of similarity between paragraphs. This will allow me to identify paragraphs with similar subject matter. This is where I will be able to determine if there will be enough overlapping data to construct the matrix above. It will allow me to tease out if there are any trends between certain time periods as well, if I analyze the dates (if they exist for a given document) associated with a letter. This could be a neat comparison for documents within a cluster to see why the two are so closely related, if they are indeed at different time periods. The sentiment analysis will allow for a closer look at how these groups respond beside the frequency or inverse document frequency. Is one group more worried than other? Is Al-Qaeda more level-headed than ISIS? These questions can be addressed by examining the sentiments involved with the documents. Lastly, I will use the force directed word networks to examine the opposition of the language used within each group and to compare the different circumstances in which they occur.

The final result of these analyses will be a two-page paper detailing the insights gleaned on each of these groups as a whole, as well as a comparison of the reactions between the two groups to both internal and external events. Visualizations will also include word networks to represent the opposition, as mentioned in the methods, visual representations of the sentiment analyses, and, perhaps, I can tie people to an overall sentiment given a situation to show how they respond to events. This would illuminate the dynamic of the organization and refine our understanding of the roles they played or continue to play.