

Narrative Threads Extracted Through Topic Modeling and Sentiment Analysis

Chloe Quinto
Stevens Institute of Technology
662 Beers Street
Hazlet, NJ 07730
(+1) 732-673-2290
cquinto@stevens.edu

Aron Lindberg
Stevens Institute of Technology
1 Castle Point Terrace
Hoboken, NJ 07030
(+1) 201-216-3391
aron.lindberg@stevens.edu

1. ABSTRACT

EVE Online is a space-based massively multiplayer online role-playing game (MMORPG) that engages thousands of players. These players are part of a single online community, but also organize themselves in corporations and alliances. Understanding the narratives that unfold across such communities is crucial for being able to interact successfully with the customers of such games. By extracting, analyzing, and visualizing topics and their evolution across time within the forums of EVE Online community, we can reach such an understanding. This poster proposes a method for achieving this.

2. BACKGROUND

EVE Online was launched in 2003 by the Crowd Control Productions (CCP) Games. Players of EVE Online take on the role as a spaceship pilot and participate in activities such as mining, piracy, manufacturing, trading and combat. Due to the nature of EVE Online, players became heavily invested in the virtual world and its attendant online community has consequently flourished.

Traditionally, market researchers have tried to understand customers through surveys and focus groups of various kinds. Unfortunately, such methods have drawbacks. Consider surveys which tend to produce low participation and representation bias due to their voluntary nature.

As more and more products and services are being delivered as part of complex ecosystems hosted on online platforms, new sources of data are becoming available. One important form of such data is the conversations that occur in online forums.

We seek to further our understanding of such conversations through the use of a novel topic modeling and visualization approach that seeks to understand the evolution of “narrative threads” over time.

3. APPROACH AND UNIQUENESS

Understanding conversations found in online forums is a challenging task as they are often interwoven between

other speakers and other topics. A solution is to focus on the topics that are found within each conversation. In natural language processing, topic modeling is a type of statistical model for finding abstract themes that occur in documents. After running this analysis on a sub forum called *Player Features and Ideas*, the top 10 most occurring words within each topic are given as well as the probability of find these topics in the documents. By using an analyst’s intuition and experience in the detection process, topics that were not understandable were eliminated. Finally, the probabilities found in each topic were mapped over time. Trends were discovered and were found to be significant after comparing major events found in the game.

In a similar fashion, the sentiments found in the online forums were found and followed the same patterns in a few topic model visualizations.

This approach provides refined narrative threads of how conversations on the forum evolve over time, thus enabling us to understand the conversations that customers engage in from a new perspective.

4. DATA COLLECTION AND ANALYSIS

There is one main and popular site that hosts online forums for the EVE Online games [1]. Since there were multiple sub forums, the sub forum called *Player Features and Ideas* was chosen as it is a place to discuss, develop, and criticize features of the game. This is central to the proposal as critical events in EVE Online revolved around the implementation of new ideas.

A scraper was built to extract thread IDs, usernames, date, and posts. This was then written in a csv file for analysis. A topic modeling tool called MALLET was used to extract the probabilities of topics from the csv file. The MALLET package provides a fast and scalable implementation of Latent Dirichlet Allocation (LDA) which is one of the most common algorithms for topic modeling. After running this package, the probability of topic occurrences in the corpus as well as the top 10 words were given as a result.

From personal experience of the game and intuition, topics that were unclear were eliminated from the study.

The reason is that almost always, topics that are unclear or vague show almost no patterns over time.

Using the date of each posts with the topic probability, visualizations of topics over time were created and analyzed. Out of the 15 topics that were analyzed, only five topics were clear and showed significant patterns. The same methods were used for the sentiment analysis. Using a library in RStudio, analysis was conducted on the CSV and was afterwards mapped over time.

Major events of EVE Online such as the Incarna Crises and the NDA Leaks were noted on the visualizations to give a perspective of the timeline.

5. RESULTS

After running the MALLET package on the sub forum csv, the topic probabilities were given as well as the top 10 words associated with that topic. Table 1 shows the clear and well-defined topics that were extracted from the MALLET package.

Topic 1	Topic 7	Topic 8	Topic 11	Topic 14
supported	plex	vote	alliance	evemail
signed	Incarna	issue	signed	spam
insurance	supported	jade	corp	implants
bonus	cosmetic	signed	pos	stop
scanner	bots	council	skill	kungutsumen
local	bridges	meeting	mission	bounty
drones	super	nano	macro	nullsec
ship	jump	support	rules	null
minmatar	nullsec	forum	exploit	fleet
amarr	trolling	discussion	queue	wormhole

Table 1: Topic models extracted from MALLET package

The topics can be labelled to the following:

- Topic 1:** Feature Requests
- Topic 7:** Microtransactions
- Topic 8:** Council Issues
- Topic 11:** Ideas Infrastructure
- Topic 14:** Communications

The probability of these topics occurring on the corpus was mapped from 2008 to 2016 on figures 1 to 6.

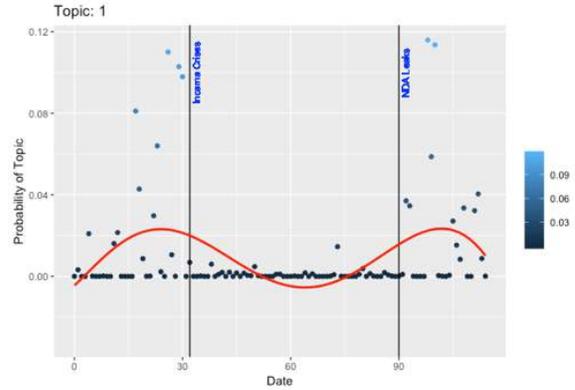


Figure 1: Feature Request topics over time

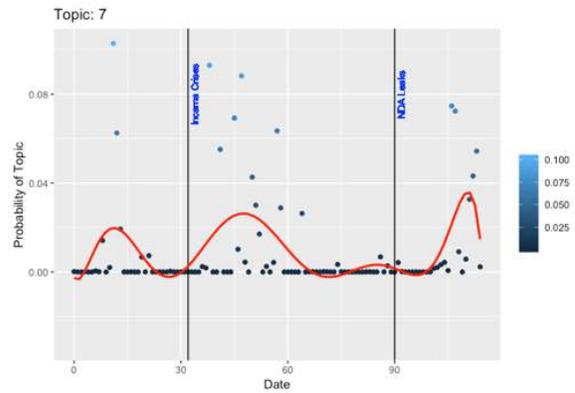


Figure 2: Microtransaction topics over time

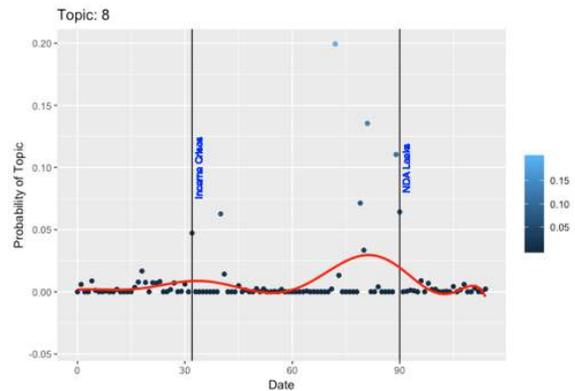


Figure 3: Council Issue topics over time

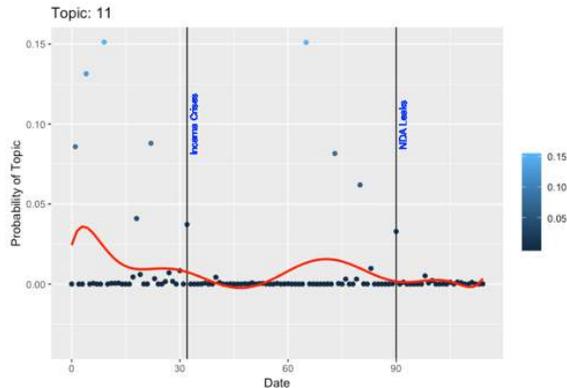


Figure 4: Ideas Infrastructure

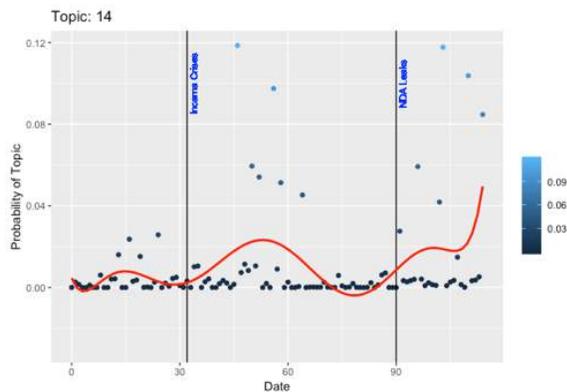


Figure 5: Communication topics over time

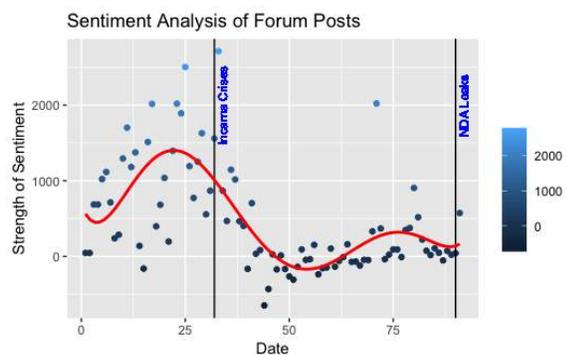


Figure 6: Sentiment Analysis over time

Two major events in EVE Online was outlined. Specifically, they were the Incarna Crises and the NDA leaks. The Incarna Crises was an event in 2011 in which an internal newsletter discussing the implementation of microtransactions were leaked to the general public [2]. This resulted in a massive uproar from the general public and a significant loss of economic capital for the CCP. The NDA leaks was a period in 2015 in which sensitive information was released to the public that diminished the credibility of the CCP. This also resulted in a large uproar and another significant loss in economic capital.

Figure 1 displays that feature request posts and Figure 4 shows ideas infrastructure posts. These two figures have the

same pattern in which topics were most predominant before the Incarna Crises and after the NDA leaks. This is intuitively a sound result because players would be interested about what new features should be implemented when the company is doing well. This would be the opposite when the company is facing crises and is reflected as such.

Figure 2 shows microtransaction topics and Figure 5 displays communication over time. Both of these topics shows the same pattern. Topics are high after the Incarna Crises as well as the NDA leaks. This again is intuitively a sound result because microtransactions were at the core of the Incarna Crises. Communication between players and CCP were also questioned during the Incarna Crises and the NDA Leaks. The word Kungutsumen was found in the topic for figure 5 and this user was banned during the NDA leaks.

Figure 3 presents council issue topics over time. In the figure you see a period right before the NDA leaks where council issues were the highest. This coincides with what eventually occurred in the NDA leaks period. The online forum was talking about voting, meetings, and council right before the period where the credibility of the company was in question.

From what was described above, the sentiment analysis in Figure 6 is not a surprise. Up until the Incarna Crises, the sentiment of the online forums was extremely positive. Then there was a significant decrease in sentiment after the crises. This plateaued for only a short amount of time. Right before the NDA crises, the online forums users experienced a positive sentiment. Then, once the NDA leaks period occurred, however, the sentiments decreased again. The pattern that Figure 6 exhibits

Overall, these visualizations have given a new perspective of how the online forums evolve over time. There are narrative threads such as feature requests and ideas infrastructure that occur at the same time in online forums and coincide with the positive sentiment. Threads like communication and council issues coincide with the negative sentiment. What is innovative about this method is that complex ecosystems can be dissected and be given a different perspective. Topic modeling over time uncovers new conversations that are not present in other data analysis.

6. CONTRIBUTIONS

The main contributions of this work are:

- Present a novel approach to understanding topics talked about by online communities over time.
- Utilized Latent Dirichlet Allocation (LDA) algorithms to extract topics over time and map them to known periods of EVE Online. Unlike previous efforts, these topics provide a better insight to a customer's needs that other data collection methods do not meet.
- Validated the approach through an experiment on a sub forum of EVE Online Forum data called "Player Features and Ideas"

7. REFERENCES

- [1] "EVE Online Forums." *EVE ONLINE FORUMS*, forums.eveonline.com
- [2] Lindberg, A. 2018. *EVE Online: Knowledge Integration in a Virtual World*.