Computational Treatise on Urban Legends

Coen D. Needell

Introduction

We have to focus our attention though and draw boundaries between legend, fake news, conspiracy theory, conjecture, and so on. We are all fully aware that the systems that make the Internet a breeding ground for urban legend, also make it a fantastic substrate for the other types of rumor and hearsay (Blank and McNeille 2018). For our purposes, we will consider the genre of "Urban Legend" to be defined by the fact that the people who spread the legend know that it isn't true. Conspiracy theories tend to be spread by the true believers, and fake news tends not to be spread by people, but by robots and organizations. **(This isn't entirely true, there are a lot of people who talk about conspiracy theories without believing them)**. In addition, the Urban Legend tends to take place in a fuzzy time-frame, the near past **(elaborate)**. Conspiracy theories either exist in a sharp time frame, like the Kennedy Assassination, or they exist in the present **(This seems like an important distinction, make it more clear)**. Fake News on the other hand always exists in the Now, and often includes a Call-To-Action which forces it to be mixed up in current events. Fake News usually falls into the wider category of Hoaxes, being stories that are written and disseminated with the express purpose of deceiving the public. In addition, the star characters in Fake News and Conspiracy Theory are celebrities, senators, presidents. (Frank 2011) Urban Legend stars the everyman, the friend of a friend, the weird uncle, but it also stars some hard-to-understand entity, like a fairy, an alien, slenderman, or a large company's incomprehensible legal policies.

Some previous work on urban legends has defined them as being related to rumors. (Dunn and Allen 2005) This stance usually maintains that both rumors and legends are told with the intention of being believed, told as if they are true, and are difficult to verify. The distinction between the two being that rumors are generally smaller, simpler stories, and are not told in a narrative fashion. Consider a rumor like "I heard that John in HR is secretly dating Cassie in accounting" or even something with larger stakes like "Justin Trudeau is having an affair with Angela Merkel". These stories don't contain a narrative arc in the same way that the quintessential urban legends do.

Generally "Urban" or "Contemporary Legends" describes stories that few people truly believe, but are still told and retold time and time again. These stories often have common features, they leverage the unintuitiveness of modern life, a sense of distant credibility ("This happened to a friend of a friend"), and in some cases a sociopolitical call-to-action. Prototypical Urban Legends include the story about a man meeting a woman in a hotel bar and waking up the next day missing a kidney in a bathtub full of ice, or the story where your uncle's friend went down into the sewers and encountered a bask of exotic crocodiles that had been flushed down by irresponsible pet owners.

But let's focus for a moment on that third millennium (and also the decade leading up to it). The human experiment (although mostly the western countries and Japan), had just been linked to one another through the Internet. Before Google created the "clearnet", before Facebook acquired every social forum, we had specified forums, on separate sites, operated by separate people. We had Usenet, a social media system built out of email and simple text-based servers.

Usenet was a system of email mailing lists that were connected and chained together in a manner such that anyone could post on them, and anyone could read the history. They functioned somewhere between bulletin board systems and email newsletters, and filled the niche that would later be overrun by forum software. The main difference being that usenet was decentralized.

Usenet hosts would pass messages along to people, and receive messages from posters, often through a chain system, to ensure that a post which was sent to more than one newsgroup got to all of the requested recipients. (Mark R. Horton 1987)

Usenet was divided into newsgroups, which were in turn divided into subgroups. Generally these would be addressed by their hierarchy, separated by a period. For example, `sci.agriculture.beekeeping` would be a board about beekeeping, but it also fit under the subcategory agriculture, which was in the science hierarchy. There were a number of hierarchies which were controlled by specific groups. The major groups are referred to as the Big 8. Originally, in 1980 when Usenet was created there was only one major newsgroup, which was called `net.*`. In 1987, Usenet went through a major restructuring, creating the Big 7. These were intended to be an encompassing set of categories, and they were: `comp.*`, for discussion of computer related topics, `news.*`, for discussion of Usenet itself, `rec*`, designed for recreational activities like games and hobbies, `sci*`, discussing science, `soc.*`, for both socializing and for discussing social issues, `talk.*` which was targeted at discussion of contentious issues like religion and politics, and finally `misc.*`, things that don't fit in the other Big 7. This restructuring led to a controversy about what was allowed on usenet. Even `misc.*` was heavily moderated. The rules for each hierarchy were set up as to prohibit, from any of the big 7, discussions about recreational drug use, sex, or sharing recipes. So the self-styled Usenet Cabal set up an eighth major newsgroup, `alt.*`. Designed to keep the debauchery of the internet contained. `alt.*` is not considered an official member of the Big 8, and the real eighth member was added in the 90s, `humanities.*` for discussion of the humanities, in a similar spirit to `sci.*`. The `alt.*` newsgroup however, became a free-flowing, unmoderated message board, with all of the issues that includes. The Usenet community would often jokingly refer to `alt.*` users as "Anarchists, Lunatics, and Terrorists."

In a time before Google, or even Snopes, you couldn't fact check that troll's rant, and those rants spread like wildfire.(Dunn and Allen 2005) (Donovan 2004) Even in the later times of transition, we saw legends like the story of slenderman, video game related legends like

"Herobrine", and so on. Obviously we see the scarier legends show up on forums dedicated to horror stories and urban legends, we see video game legends sprout up out of those games' communities, and we see less intense legends appear in more general use communities(Blank 2007). These are things like the story about the Neiman Marcus cookie recipe. This brings us to the question: How do Urban Legends spread through online communities, and how do those communities effect the nature of the legend itself?

Methods

Data Sources

Much of the early internet is made available to us. Groups like the Internet Archive host repositories full of early internet interactions. Many of the forums that people used to discuss broad topics are still existing, and keep their old posts up for posterity. That being said, this is far from complete, IRC (Internet Relay Chat) logs will be mostly wiped from the record. IRC servers were often private, and hosted by private entities. In addition, the Usenet archive that will be used for this discussion is not complete, it is however very comprehensive for the boards that it exists for, and this includes the largest boards. By compiling usenet records, we see a very large section of the picture. We will be able to see links between individuals, and links between communities (archive.org 2020). I will focus on a couple of usenet newsgroups, `talk.*`, `soc.*`, `rec.*`, and `sci.*`. I have also procured data from select `alt.*` newsgroups, including `alt.folklore.*`.

Using text classification methods, we can sift through these posts, and pick out a selection of urban legends that we will use to study their dynamics overall. Then, we can find posts that reference those legends. Conceptually, we can imagine that posts fit into one of three categories, irrelevant, which consists of posts that do not reference the urban legend at all, carrier, which consists of posts that reference the story in passing, but don't tell the story, and spreader, which consists of posts that tell the legend, thus spreading it to other people in the community. We can imagine that a single user can write posts that belong to all of these categories, but each individual post can only belong to one. **(Flesh out the theory a little more, cite common**

characteristics in post types.)

Analysis

Using those three categories above, we would construct a temporal model of spread through the network that will let us understand how the legend spreads overall. This is a contagion model of ideas that is based off of the idea that the individuals in the community are exposed to a legend, and then spread it for some amount of time, before they get bored of it and stop. In addition, we can use topic modeling to create a set of features for all of the spreader posts for all of the urban legends that we're studying, and see how the spreader posts change their telling of the story based on the community that they're talking to. For example, we'd expect the story of slenderman to take on a different tone in an anthropology focused community than it would in a parenting focused community. **(Be more comprehensive about how that contagion model works)**

Ideally the topics that we deconstruct from the spreader post serve as analogues for the "genes" in the Darwinian model of idea spread (Dawkins 1989) (Kronfeldner 2014). Under this theory, we would expect that the topics in a spreader post will change as the legend interacts with a new substrate, a specific type of community **(Citation Needed)**. Previous work on the Darwinian model of ideas focused on the concept of emotional selection. Ideas that employ emotional techniques, pathos argumentation, or evocative imagery tend to spread faster than those that don't.

We would also expect that some stories are less prone to mutation than others. For example, there is a certain subgenre of scary story that appeared on the early internet called "creepy pasta" which is a mutation of "copy pasta" which is in turn a mutation of "copy paste". These are stories which mostly appeared in the same format, implying that a person spreading the story would be copying the story from the place that they first saw it, and pasting it to a new post. While "Creepy Pasta" have died down over the years, and now the term is synonymous with internet scary stories, "copy pasta" is still a core part of how internet communities interact, with many of the famous copy pasta becoming memes of their own, often mutating in such a way that they

change topics entirely, but keep the cadence and tone of the original. **(Citation Needed, possibly irrelevant information.)**

If the Darwinian model is accurate, we should expect the topics that make up an Urban Legend to change when the story is introduced to a new environment. Ideally the topics that change will be in some way connected to the new substrate's nature. For example we would expect the version of an urban legend on an academic community to make more reference to complex ideas and theory, whereas in it's original substrate, the language would likely be more general. **(Citation Needed)**

Results

Discussion

Conclusion

———————————

References

archive.org. 2020. "Internet Archive: Digital Library of Free & Borrowable Books, Movies, Music & Wayback Machine." https://archive.org/.

Blank, Trevor J. 2007. "Examining the Transmission of Urban Legends: Making the Case for Folkore Fieldwork on the Internet."

Blank, Trevor, and S McNeille Lynn. 2018. *Slender Man Is Coming: Creepypasta and Contemporary Legends on the Internet.* Utah State University Press.

Dawkins, Richard. 1989. *The Selfish Gene.* New ed. Oxford University Press.

Donovan, Pamela. 2004. *No Way of Knowing: Crime, Urban Legends and the Internet.* Routledge.

Dunn, Henry B, and Charlotte A Allen. 2005. "Rumors, Urban Legends and Internet Hoaxes." In *Proceedings of the Annual Meeting of the Association of Collegiate Marketing Educators*, 85.

Frank, Russell. 2011. *Newslore: Contemporary Folklore on the Internet.* University Press of Mississippi. https://doi.org/10.14325/mississippi/9781604739282.001.0001.

Kronfeldner, Maria E. 2014. *Darwinian Creativity and Memetics*. Routledge.

Mark R. Horton, Rick Adams. 1987. "RFC 1036: Standard for Interchange of Usenet Messages." https://www.hjp.at/doc/rfc/rfc1036.html/.