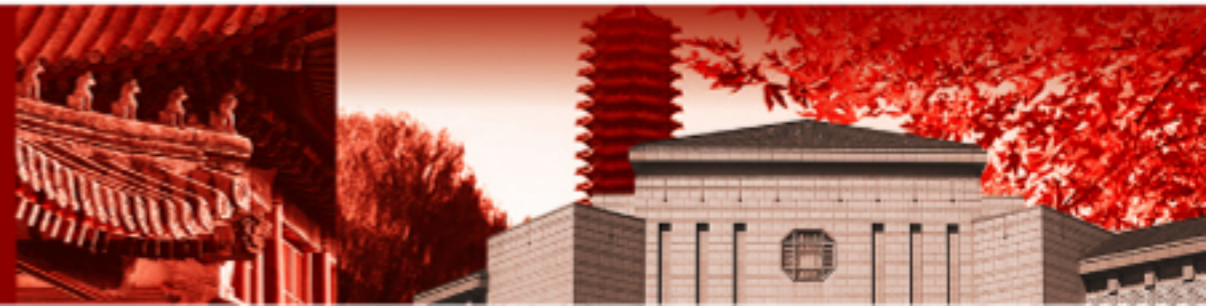


# How does Truth Evolve into Fake News? An Empirical Study of Fake News Evolution

Mingfei Guo, Xiuying Chen, Juntao Li, Dongyan Zhao, Rui Yan

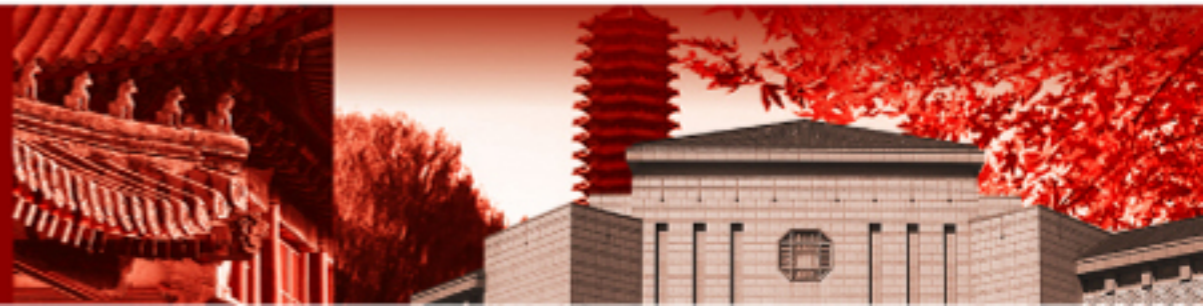
**Presenter: Mingfei Guo**

April 14, 2021



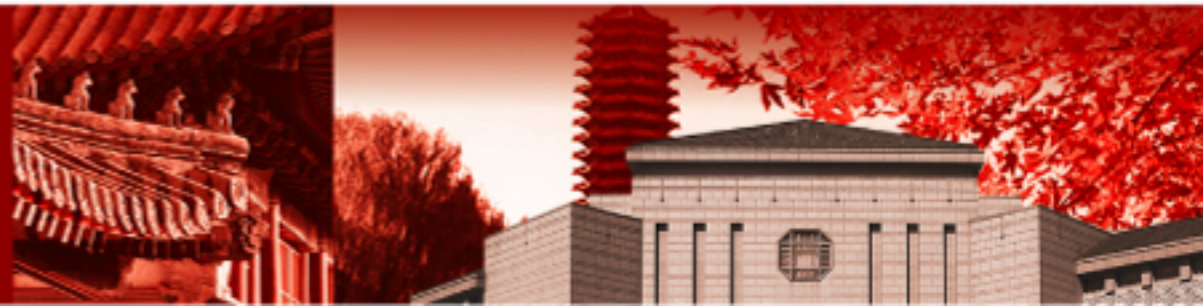
# Motivation

- In real social networks, the news evolves constantly in the spreading process, which grows more concise, more attractive, more easily grasped and told.
- We believe the fake news recognition models would obtain better results if they learn the deeper pattern of the evolution process.
- Hence, we introduce the Fake News Evolution (FNE) dataset and observe the evolution from facts to fake news.



# Contribution

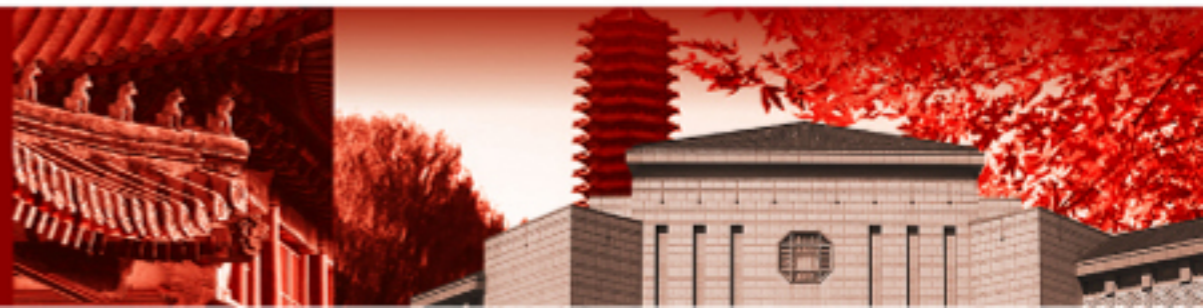
- We propose a dataset that tracks the fake news evolution process. Each piece of data includes three articles, the truth, the fake news, and the evolved fake news, and all of them are derived from the same particular event.
- We look into the data properties to study the features during the evolution process. Here, we observe the statistics of the disinformation techniques, text similarity, top 10 keywords, classification accuracy, sentiment properties, and parts of speech of words.



# Overview

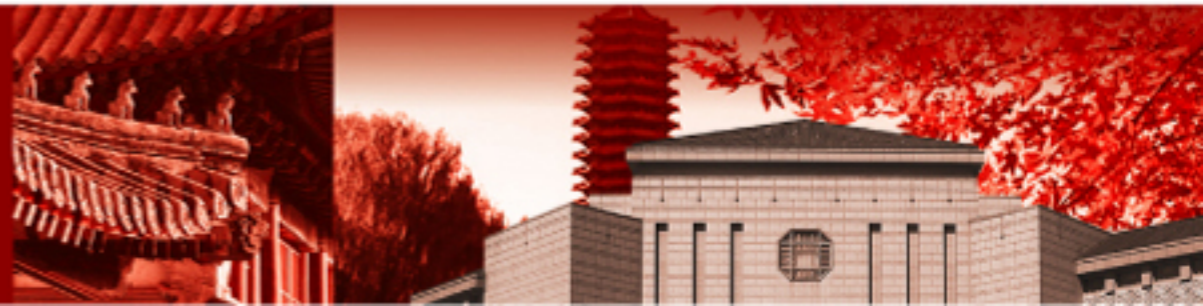
<p><b>Truth:</b> The laws restrict people from various military-like activities that result in violence and civil disorder. The amendment to the Virginia law, specifically, prohibits marching with weapons or explosives for the purpose of intimidation.</p>
<p><b>Fake News:</b> Proposed virginia law would outlaw krav maga, brazilian jiu jitsu, kickboxing, tai chi, firearms instruction and self-defense training, the story reads: The law would instantly transform all martial arts instructors into criminal felons.</p>
<p><b>Evolved Fake News:</b> The state of virginia, now entirely run by truly insane democrats who support infanticide and child murder, is proposing a new 2020 law known as sb64 (see link here) which will be taken up by the democrat-run senate beginning january 8, 2020.</p>

**Table 1: An example of the dataset. We highlight the original truth information by using yellow and the fake news and the evolved fake news by using pink.**

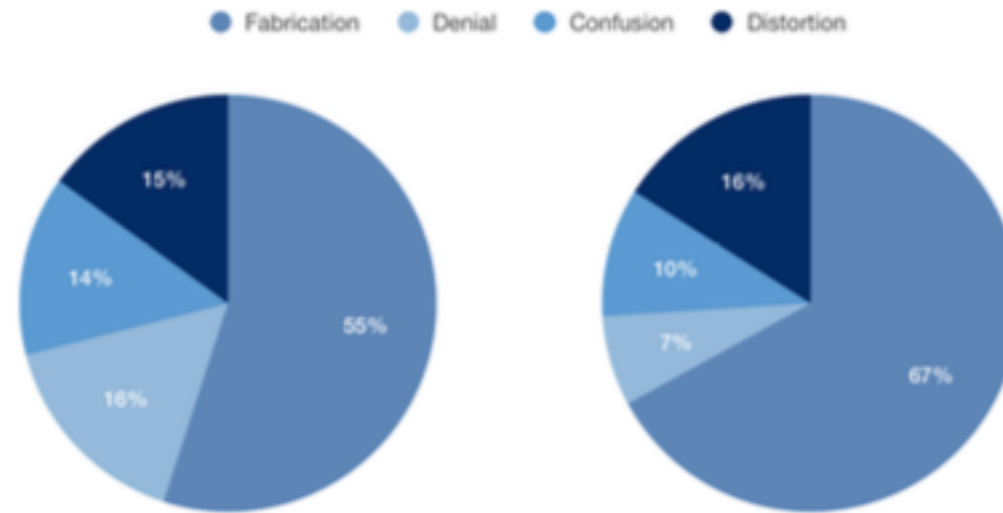


# Dataset

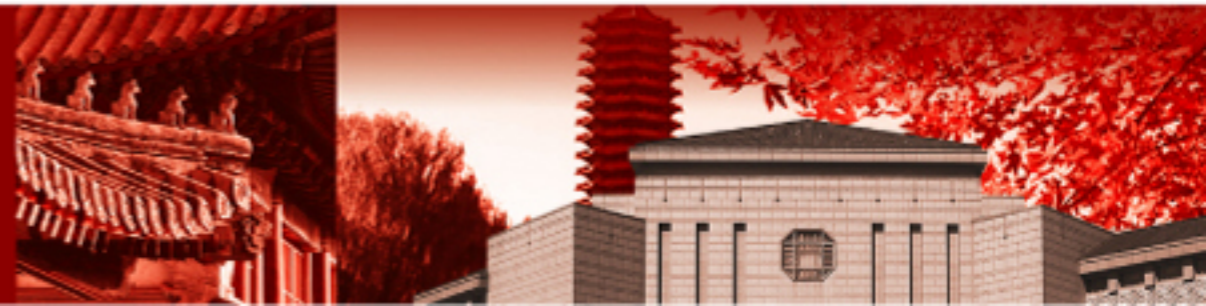
- First, we crawl a fact-checking website Snopes and collect true articles because it is transparent about examination sources.
- Secondly, we collect fake news from the citation of the content.
- Thirdly, social media platforms are likely to regulate questionable content, so it is more preferable to gather the evolved fake news on a snapshot website. We regard the records on the archive website as evolved fake news.



# Observation



**Figure 1: The left one represents the proportion of disinformation techniques from truth to fake news, while the right one represents that from fake news to evolved fake news.**



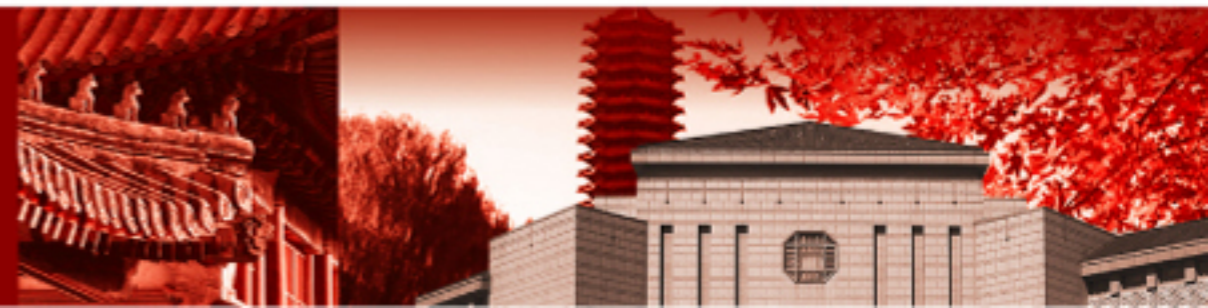
# Observation

truth	trump	said	news	president	people
	law	one	state	states	also
fake news	trump	president	clinton	obama	school
	police	law	state	bill	health
evolved fake news	trump	obama	people	said	president
	clinton	ago	like	new	one

**Table 2: The top 10 keywords of the truth, the fake news and the evolved fake news.**

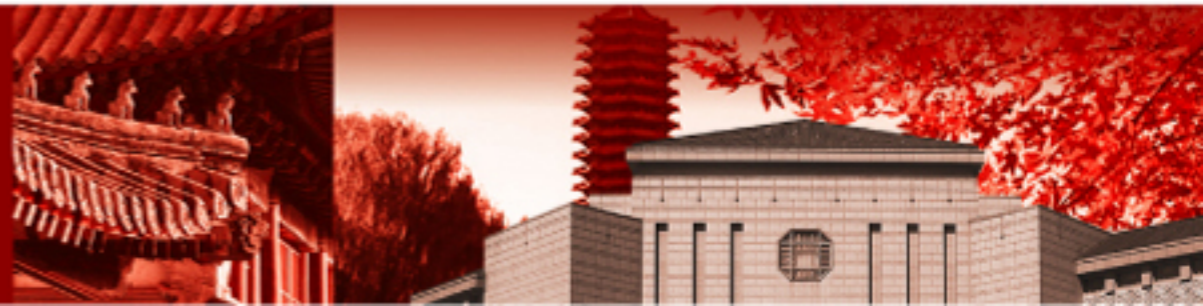
	Truth vs Fake	Truth vs Evolved	Fake vs Evolved
Polarity	0.1735	7.2598	7.0863
Subjectivity	0.2696	6.8796	7.1492

**Table 4:  $Q$  statistic of the Tukey's test among the three samples, the truth, the fake news and the evolved fake news.**



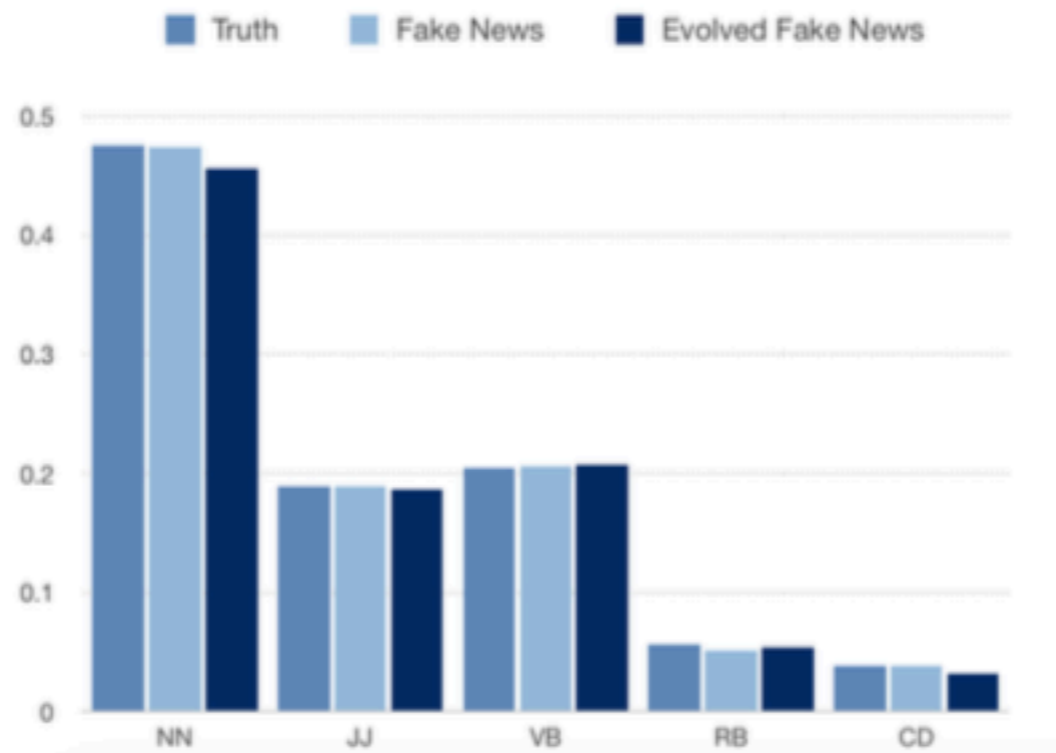
# Observation

- We calculate the similarity score between the truth and the fake news  $X_1$ , and that between the fake news and the evolved fake news  $X_2$ . Compared with the change between fake news and evolved fake news, the transformation between truth and fake news is more drastic and extensive.
- We train several standard machine learning models like Naive Bayes, Random Forest, and Adaptive Boosting. The accuracy of the evolved fake news is smaller than that of the fake news, which indicates that detecting the evolved fake news is more intractable.

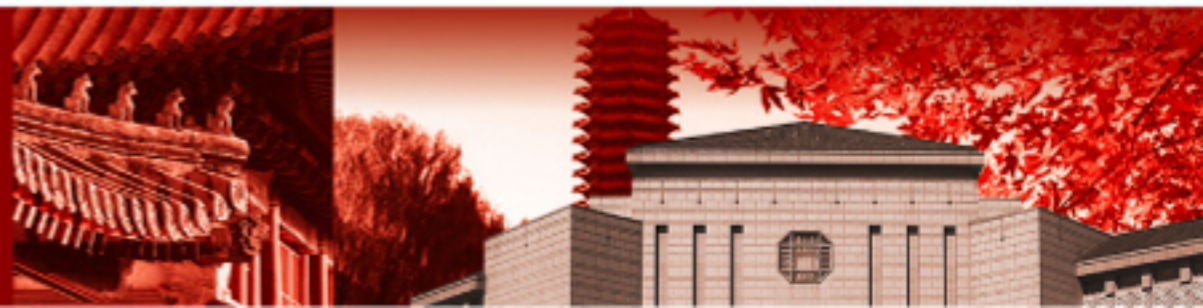




# Observation



**Figure 2: Proportion of different parts of speech of words during the evolution process.**



Thank you!



北京大學