Misinformation on telegram with GNN

Contents

[Meetings 1](#_Toc178870767)

[Meeting 11.07.2024 1](#_Toc178870768)

[Meeting 18.07.2024 1](#_Toc178870769)

[Colloquium 1 2](#_Toc178870770)

[Meeting 04.10.2024 2](#_Toc178870771)

[Literature Review 2](#_Toc178870772)

[High level of correspondence across different news domain quality rating sets 2](#_Toc178870773)

[Detection and Discovery of Misinformation Sources Using Attributed Webgraphs 2](#_Toc178870774)

[Computational analysis of US Congressional speeches reveals a shift from evidence to intuition 2](#_Toc178870775)

# Meetings

## Meeting 11.07.2024

* GNN, Telegram, Misinformation classification
* URL graph network
* Scientific rumor, >>message propagation speed, time intervals between interactions, and temporal user activity patterns >>fast misinformation detection?
* >>hypothesis?
* Just a GNN with URL links? (appear often in the same chats/groups)
* How to extract links from telegram
* How to extract titles from URL’s
* What is misinformation? How to classify
* Score of misinformation - evidence minus intuition score
* Scrape attributes from websites
* Research team in göttingen – Data?

## Meeting 18.07.2024

* which time period? 2020-2023
* Power law distribution of domains, threshold, classification of domains (commercial, social media (Roman has a manual exclusion)
* Also some user data, more uniform behaviour, chats are more power law distributed (but it is more easy to filter chats than user)
* Lasser data as groundtruth
* Features (number of chats, number of tokens, emotional reactions, total number of user, channel shares, user shares, how many times it got forwarded, duration, velocity, timestamp, text of chatmessage, longest message with that URL …)
* Posting behavior of chats (small or big groups)
* How to get from domain to title of article?
  + 50% getting text from URL
  + 50% crawling text tokens with visiting websites from html
  + screenshot and extract header
* Not reachable indicates no credibility
* For each domain, text tokens of the group in which the domains got posted
* Start with English and then go on with German?

## Colloquium 1

* selenium web scraper
* log corpus
* conspiracy cluster as feature

## Meeting 04.10.2024

* embeddings out of scraped text?

## Meeting 15.10.2024

* url-title extraction
  + selenium works but takes too long
  + extraction from url directly is not that exact 🡪detecting if there is an actual world
* first MLP – bad performance
* Roman
  + Will send me speed indicator
  + And first 3 forwarded messages
* [giordano-demarzo](https://github.com/giordano-demarzo)
* Send Giordano a link to meeting and git
* Domain based so not all titles and text just a few per domain
* Giordano key: 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
* Chatgpt4 mini: 1 Million
* Lama3, replicate: r8\_AKSq3xdIIdYC1EDhSsgZ6PAH9GAqkJg45FVha
* <https://replicate.com/>

## Meeting 16.10.2024

* Write Expose for David!
* Forward metric might be better than the views (active engagement)
  + 3 most forwarded & 3 most viewed text messages with the URL
  + Just for most shared URL
  + Virality indicator, Euclidian normality, shares in a time normalized between 0 and 1
  + Max shares in 3 days (days is varible, but 3 is good)#
  + Timeseries, highest peak
  + Views is just scrolling
* 20GB
* Published forward network
* Channel can moderate emojys, mask quantity, reactions in total, polarized emojys 🡪misinformation
* 60 000 channels, 100 random messages to get the topic
* 30 of 250
* Pytorch network

# Literature Review

## High level of correspondence across different news domain quality rating sets

* <https://doi.org/10.1093/pnasnexus/pgad286>
* “A limitation of using domain ratings is that it assumes each domain produces content of identical quality, which is rarely (and perhaps never) the case. For example, the New York Times has a relatively high PC1 rating of 0.86, but its content can sometimes be fairly left-leaning or biased, even though factually accurate (26). Conversely, lower-quality domains often reproduce content from other higher-quality domains (such as the Associated Press), and therefore, their content is certainly not always inaccurate” page 6
* Can serve as groundtruth
* Domains classified as not fake could be used for comparison

## Detection and Discovery of Misinformation Sources Using Attributed Webgraphs

* <https://doi.org/10.1184/R1/25174193.v1>
* Domain detection with webgraph and Search Engine Optimization (SEO)
* Advantages: non-English settings, not reliant on API access to social media data, content agnostic system will not be impacted by increasing use of generative AI in SEO settings
* avoids common pitfalls of the content-based approach.
  + What are common pitfalls of the content-based approach?
* 4206 domains
* Media Bias Fact Check (MBFC)
* “we limit the webgraph to a single hop” ?

## Computational analysis of US Congressional speeches reveals a shift from evidence to intuition

* Evidence minus intuition score Segun et al.
* We constructed dictionaries to capture evidence-based and intuition-based language styles that underlie the two conceptions of truth (e.g. “fact” and “proof” in the evidence-based dictionary and “guess” and “believe” in the intuition-based dictionary, see Supplementary materials for the full dictionaries).
* 49 keywords for evidence-based language and 35 keywords for intuition-based language
* English

A set of globe icons with cursors

Description automatically generated

A set of globe icons with cursors

Description automatically generated

A set of globe icons with cursors

Description automatically generatedA set of globe icons with cursors

Description automatically generatedA set of globe icons with cursors

Description automatically generatedA set of globe icons with cursors

Description automatically generated