2023 Annual Conference & Exposition
Baltimore Convention Center, MD | June 25 - 28, 2023
The Harbor of Engineering Education for 130 Years
Paper ID #39747

A network analysis of the Twitter-Rxiv ecosystem for purveyors of science misinformation in preprints on the COVID-19 pandemic

David C. Brown, University of North Carolina at Charlotte
Mr. Erfan Al-Hossami, University of North Carolina at Charlotte

Erfan Al-Hossami is a Ph.D student at UNC Charlotte. Erfan has been mentored in teaching CS1 since 2016 and then in CS education research. His work mainly focuses on predictive learning analytics. His research interests include Machine Learning, NLP, an

Zhuo Cheng
Alyssa Lasmarias Alameda
Tia Nicole Johnson
Dr. Mesbah Uddin, University of North Carolina at Charlotte

Dr Mesbah Uddin is a professor of Mechanical Engineering at UNC Charlotte and has a long track record of providing leadership to multi-disciplinary activities within the campus.

Daniel Andrew Janies

A network analysis of the Twitter-Rxiv ecosystem for purveyors of science misinformation in preprints on the COVID-19 pandemic

David Brown[1], Erfan Al-Hossami[2], Zhuo Cheng[2], Alyssa Alameda[2], Tia Johnson[3], Samira Shaikh[2], Mesbah Uddin[4] and Daniel Janies[1]

[1] Department of Bioinformatics and Genomics, UNC Charlotte

[2] Department of Computer Science, UNC Charlotte

[3] Department of Geography and Earth Sciences, UNC Charlotte

[4] Department of Mechanical Engineering, UNC Charlotte

A network analysis of the Twitter-Rxiv ecosystem for purveyors of science misinformation in preprints on the COVID-19 pandemic

Abstract

This paper illustrates the final research product resulting from a team of diverse students of many educational stages and backgrounds in cyber intelligence-based research. We chose a real-world dataset of discussion of scientific preprints on SARS-CoV-2 virus and COVID disease on Twitter ™. The selection of the real-world dataset was driven by: (a) misinformation regarding COVID-19 disease and SARS-CoV-2 virus is rampant and undermines our ability to recover from the pandemic, (b) unfounded and false health-related claims are spreading on social media, and (c) the rapid dissemination of health misinformation provides challenging competition with information broadcast by public health or government authorities such as the World Health Organization and the U.S. Centers for Disease Control and Prevention. Thus, we focused on the close symbiosis between preprints, preprint servers (like bioRxiv and medRxiv), Twitter, scientific researchers, journalists, and the public that developed in the early months of the pandemic (e.g. first six months of 2020). This symbiosis, the "Twitter-Rxiv ecosystem", led to the rapid dissemination of results before traditional processes of scientific peer-review before publication. While much of the work in preprints is well-intentioned, concerns have been raised that this symbiosis may be exploited to disseminate spurious results or intentionally incorrect information. In response we constructed networks to represent public discourse surrounding scientific preprint literature on Twitter and develop metrics to score users within these networks. One such metric, peer-review percentage score, is useful for calculating the network prominence (i.e. influence) of a user while weighting that user for the quality of information propagated by the user. Peer-review percentage score can be used to identify subject-matter experts who transmit evidence-based information online. We found that these subject-matter experts outcompeted public health authorities in online forums by transmitting scientific results. Subject-matter experts engaged with the public whereas public health authorities did not.

Keywords

Preprints; Twitter; Science misinformation; "Twitter-Rxiv ecosystem"; Percolation centrality

1. Introduction

In late 2019 the SARS-CoV-2 virus emerged in China, spreading to every part of the world and causing the COVID-19 disease pandemic with roughly 18 million deaths by the end of 2021 [1]. The COVID-19 pandemic has caused much human suffering and impacted economies, governments, education, public health, and private life worldwide over the past 36 months. Early in the pandemic, individuals and governments turned to public forums of communication and social media, to rapidly share information related to COVID-19 treatment and response. At the same time, many scientific researchers released experimental results online through "preprint servers", pre-empting the traditional processes of scientific peer-review. Members of the scientific community posted thousands of scientific preprints about COVID-19 disease, treatment, response, and the SARS-CoV-2 virus [2]. Preprint servers are freely accessible, allowing preprinted results to be amplified on social media.

The online social media space proved especially susceptible to the spread of false health claims and misinformation. Both researchers and public health institutions/governments (authorities) have pointed to social media and other non-traditional media formats as transmitters of poorly fact-checked information or outright misinformation [3]. However, social media sites like Twitter have become public forums for discussing scientific results and what these data mean for public health. The discussions often focused on how scientific publications and preprints inform official guidance, such as the predicted epidemiological trajectory of the COVID-19 pandemic, the efficacy of public health responses, and the safety of vaccines [4] [1].

By the end of October 2020, over 125,000 peer-review and preprint papers have been published worldwide regarding SARS-CoV-2 and COVID-19 disease with over 30,000 of these using preprinting [4]. This volume of research on the novel pandemic is impressive but difficult to contextualize for several reasons. First, the sheer number of manuscripts represents a significant amount of work for volunteer peer-reviewers. Second, the avoidance of peer-review opens the possibility for the circumvention of publication standards, allowing scientific misinformation to enter the literature. Scientific misinformation could include 1) bad and rushed science or 2) purposely ill-intentioned or fraudulent papers posted and amplified to suit a narrative instead of scientific inquiry.

One prominent example of scientific misinformation is the "Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag" [5] which was posted to bioRxiv on January 31, 2020 (Version 1) and subsequently withdrawn by the authors on February 2, 2020 (Version 2). The quick withdrawal by the authors is noteworthy, but public unfamiliarity with scientific publishing controls and the document's versioning led to the broad sharing and amplification of the preprint on social media and in the popular press [6]. These results were treated as "evidence," despite withdrawal by the authors, which were eventually used by some to support and popularize alternative origin theories for SARS-CoV-2. Rapid and widespread discussion of the "HIV inserts" preprint demonstrates the efficiency with which potential science misinformation (PSM) can be disseminated on social media. Our intention in this paper is to understand the nature of preprints from bioRxiv and medRxiv on Twitter. To this end, we developed the Twitter-Rxiv ecosystem to investigate the users who interact with PSM.

While there are many preprint servers [4], [7], our investigation focuses on two of the major servers (bioRxiv and medRxiv), referred to together in this manuscript as "Rxiv". These servers are well known and have formed a tight relationship with the online scientific community, by virtue of an automated process for publicly posting preprints on Twitter using the @biorxivpreprint and @medrxivpreprint Twitter handles. Manuscript URLs from the Rxiv servers can be Retweeted by users, allowing rapid coverage by journalists in the popular press, web, and broadcast media. Especially in the early days of the pandemic, the Rxiv servers and Twitter created an ecosystem of discussion for preprints and peer-reviewed literature on SARS-CoV-2. In this study we focus on discovering the properties of the "Twitter-Rxiv ecosystem".

To investigate PSM in the Twitter-Rxiv ecosystem, and overcome the dual problems of obtaining SMEs and achieving scalability, we formed three research questions:

1) What is the pattern of discourse involving preprints in the Twitter-Rxiv ecosystem when studied through network graph analysis?

2) Can the network of the Twitter-Rxiv ecosystem be improved with metadata to represent the range of quality of the scientific information being transmitted while revealing PSM?

3) Are there computational methods useful for identifying SMEs who positively influence the spread of evidence-based science information?

2. Materials and Methods

2.1. Data set

We referenced a data set of COVID-19 pandemic-related Tweets collected and updated throughout the pandemic [8]. A snapshot of this data set, from January to June of 2020, was captured at the beginning of 2021. Tweet IDs were downloaded and rehydrated using the Twitter API (v1.1). We filtered Tweets for manuscript hyperlinks hosted by the Rxiv servers. This filtered Twitter data set included 12,340 Tweets referencing 2,020 unique preprint papers (as identified by DOI) from the first six months of 2020. Preprint metadata was obtained via the Rxiv servers' API in April of 2021. The full spreadsheet of joined Twitter and Rxiv preprint metadata is available upon request of the authors. The Rxiv metadata included information about the publication status of each preprint, specifically, whether a preprint had been subsequently published in a traditional academic journal at the time of the API data access.

2.2. Network graphs

We combined the Twitter and Rxiv data to create network graphs using the Python package NetworkX. Users were assigned as nodes, while edges were defined by interactions connecting two users (Figure 2). On Twitter, two immediate types of user-user interactions are immediately apparent: mentions and Retweets. Mentions follow the schema "@handle" in the text of a Tweet, while Retweets are a reposting of a previous user's Tweet by a new user.

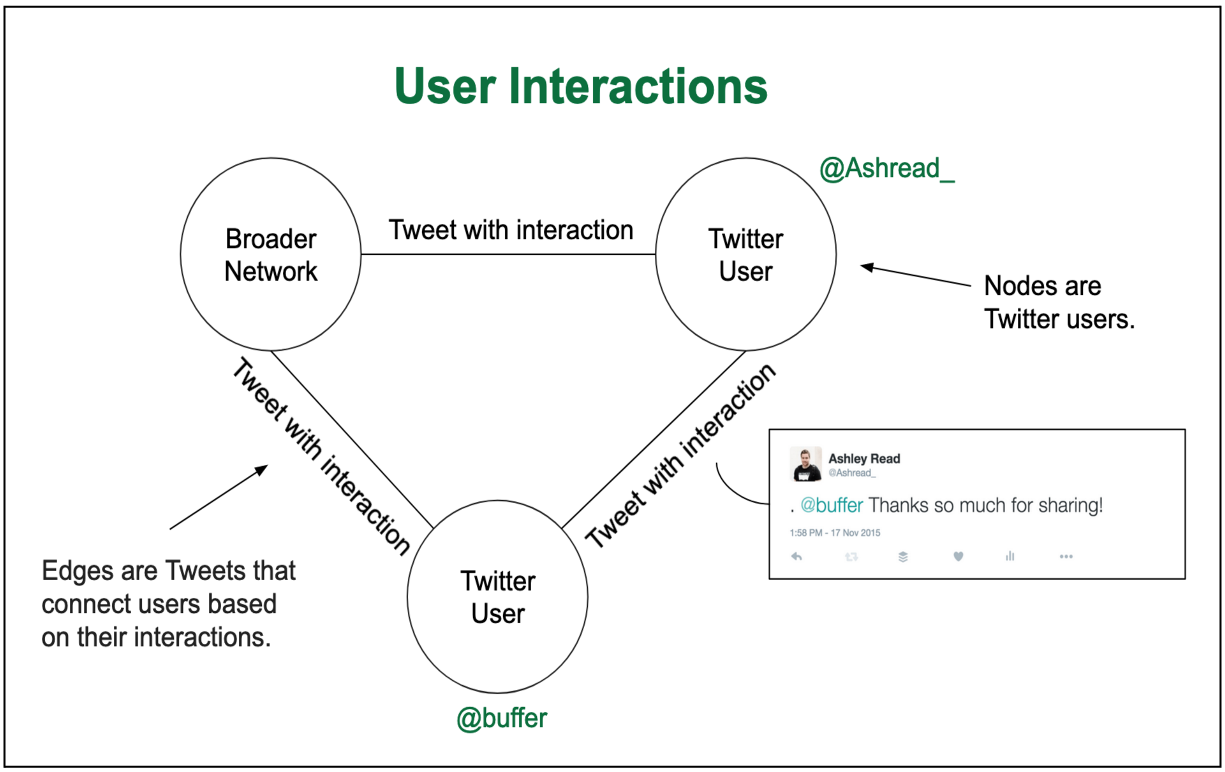

Figure 2. Users, mentions, and Retweets as interactions in a network graph.

2.3. Metrics for importance and competition

Centrality metrics are another tool for ranking the importance of nodes within complex networks. Various centrality methods were applied to the data in order to identify important users that could be potential SMEs within the LCC networks. This broader idea of importance was conceptualized as "competitiveness" within the bounds of the Twitter-Rxiv ecosystem. A competitive user would have two qualities: 1) higher relative importance to the network than their surrounding nodes, and 2) an ability to transmit scientific information as measured by originating Tweets including preprints. A specific node is therefore "outcompeted" by another node when it lacks some aspect of these qualities, with lowered importance and/or lessened transmission. Given the shape of the initial graph in Figure 4, we assumed that competitive users would often be found at the center of or between "communities", groups of users sharing at least one edge with the given user or "community leader".

Closeness centrality (Figure 5) is the calculation of the relative importance of a given node based on the shortest available path from that node to other nodes. Betweenness centrality (Figure 6) is the calculation of the relative importance of a given node based on its rate of presence on the shortest paths from any node to any other node. While both closeness and betweenness centrality can identify competitive nodes, neither metric worked well to identify potential SMEs.

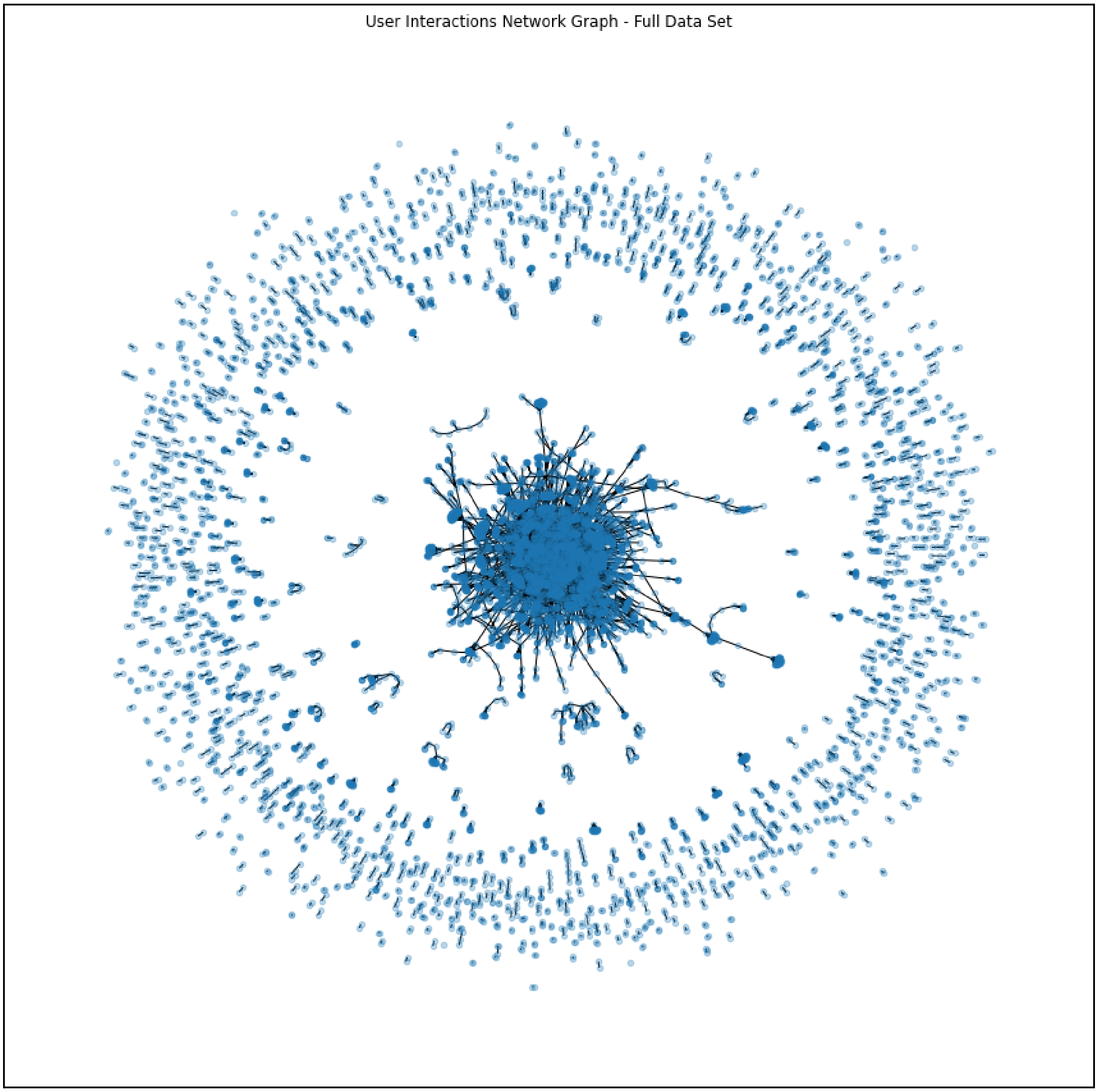

Figure 3. Network graph calculated by FR algorithm, LCC centered, for mentions and Retweets.

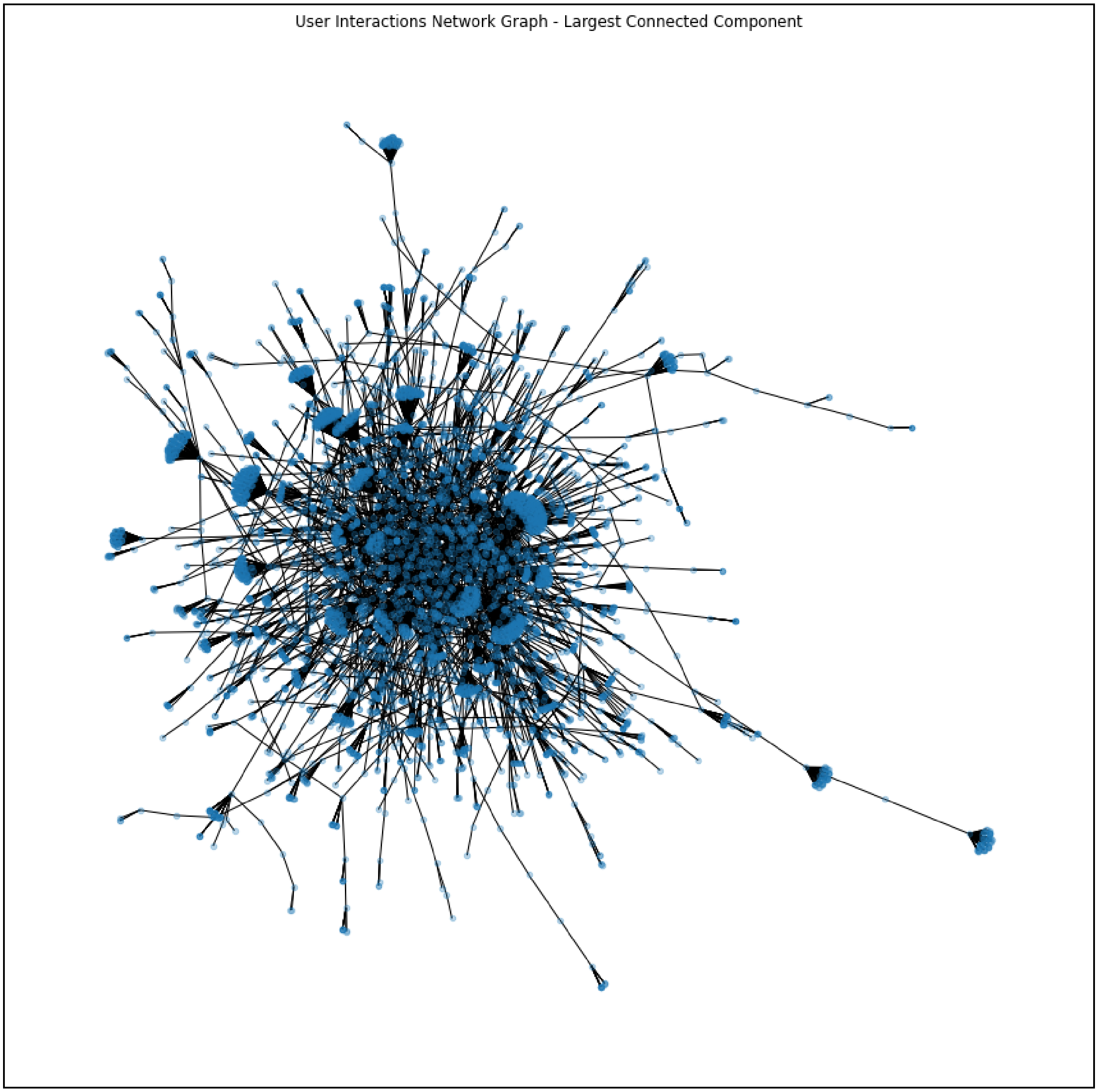

Figure 4. Magnified view of the LCC. Each node is a Twitter user and each edge represents a Tweet about a COVID-19 preprint connecting the users either by mention or Retweet. Note the existence of separated user clusters, often linked by only one or two users.

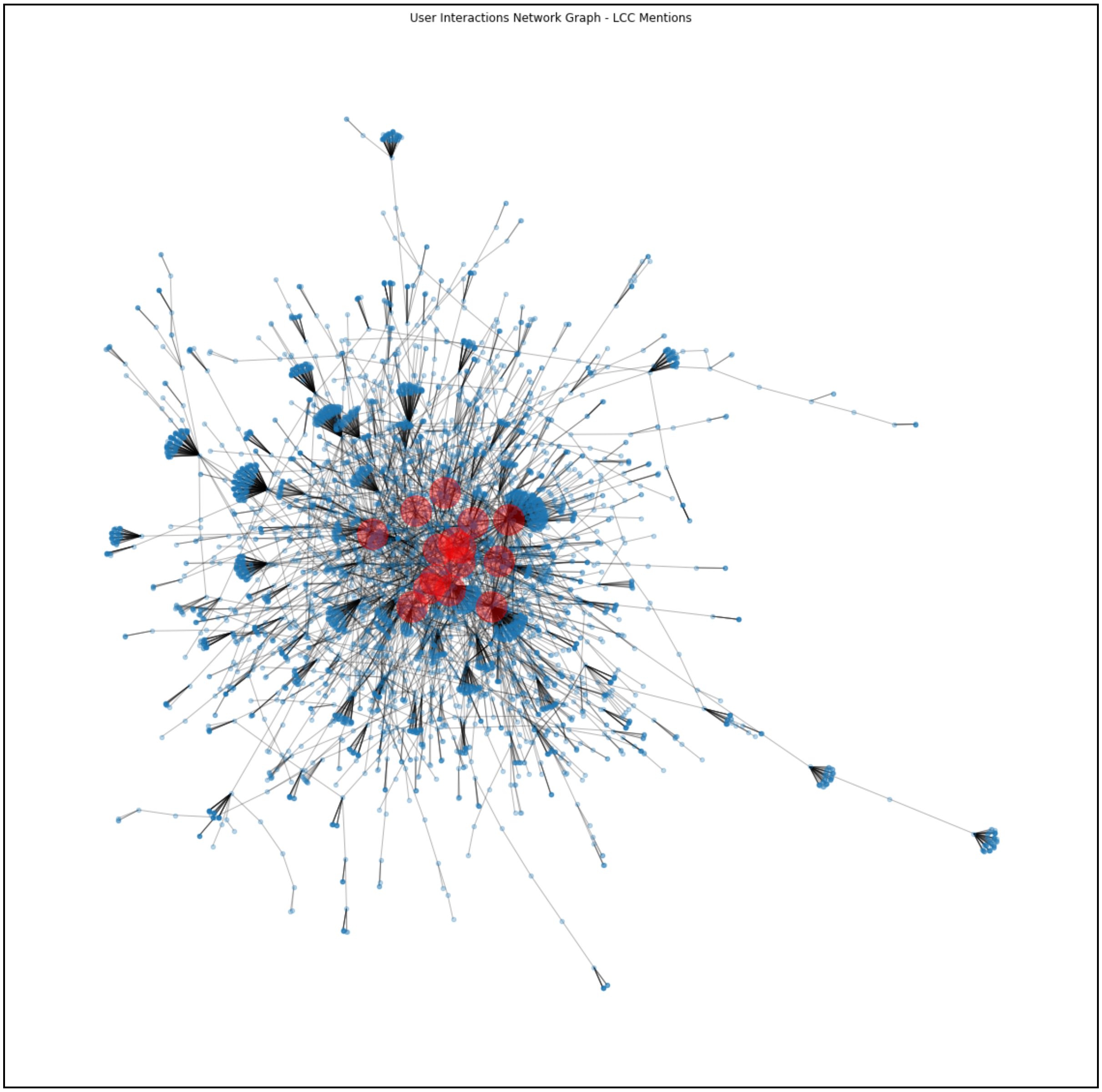

Figure 5. LCC of mentions only, with 15 users identified by highest closeness centrality (red).

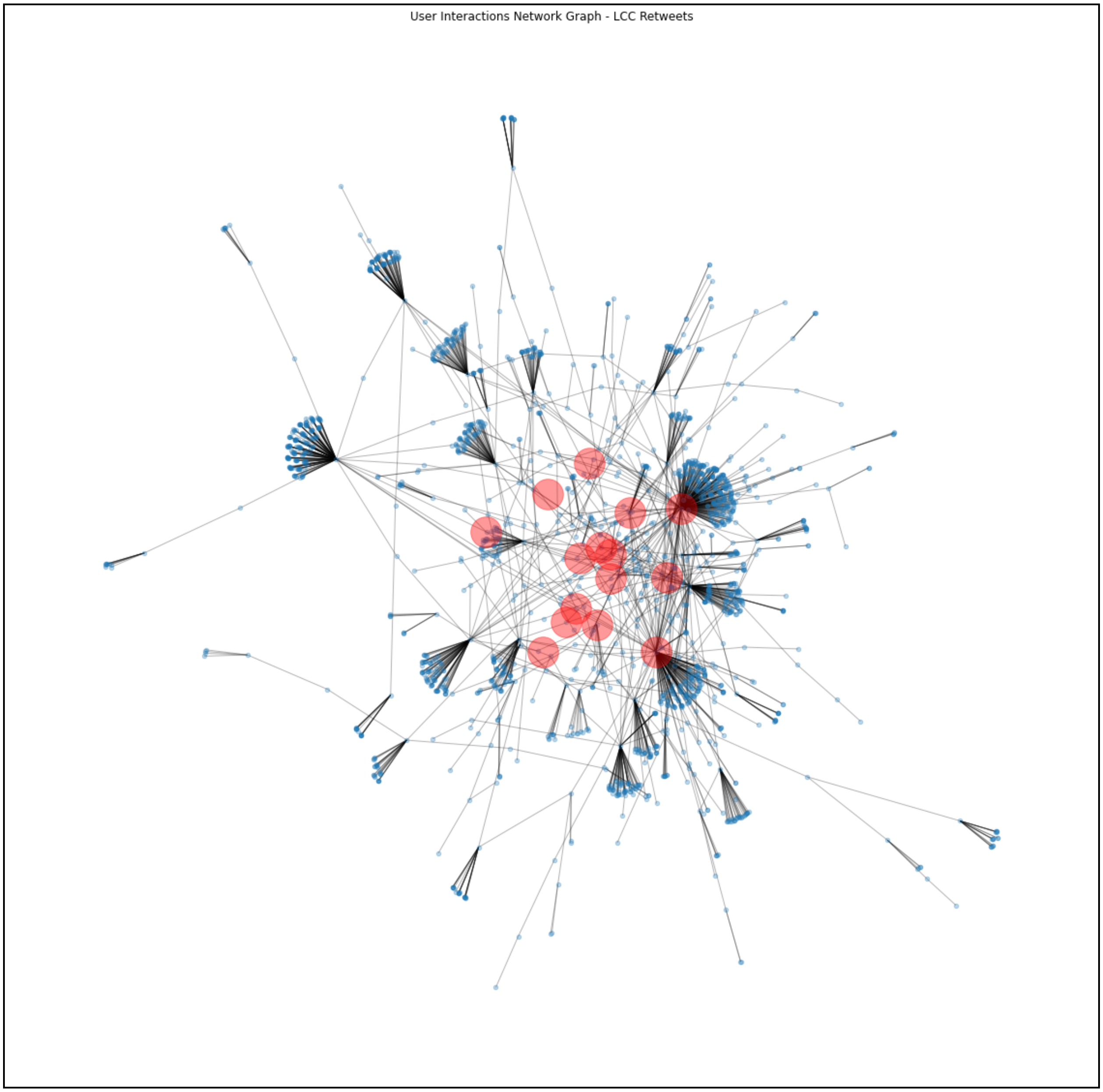

Figure 6. LCC, Retweets only, with 15 users identified by highest betweenness centrality (red).

3. Results

The data set included 10,053 unique users who originated Tweets and 12,340 unique Tweets. These Tweets were expressed in 39 languages, the vast majority being English (~84%). A summary of the descriptive statistics for Tweets per user and Tweet frequency per user can be seen in Table 1. These Twitter data represent 2,020 unique preprint manuscripts as determined by uniform resource locator (URL) and digital object identifier (DOI). At the time of our Rxiv application programming interface (API) data request, a minimum of six months had passed from the posting of the final Tweet in the data set with a maximum of eighteen (18) months from the posting of the first Tweet.

Table 1. Descriptive statistics for favorites and Retweets per Tweet in the data set.

	Favorites per Tweet	Retweets per Tweet
Maximum (count)	2430	1955
Median	0	1
Mean	4.91	28.64

3.1. Twitter-Rxiv Ecosystem Patterns

3.1.1. Government and Public Health Authority Accounts Failed to Amplify Preprints Shared on Twitter

Appendix Table B includes 25 Twitter handles (nodes) from our network, sorted in descending order by closeness centrality. Notably, the government and public health authority accounts have 0 (zero) out edges. With an out degree of 0, these Twitter handles never Tweeted a COVID-19 preprint, but the presence of in edges indicates that they were actively mentioned by other Twitter users. Government and public health authority accounts during the first six months of the COVID-19 pandemic were largely "dead ends" that did not amplify preprints. On Twitter, the accounts of government and public health authorities were often mentioned in conjunction with preprinted results, but those same authoritative accounts did not amplify preprints through Tweet origination or Retweeting. Preprints were shared with authorities but not by authorities.

3.1.2. Centrality Metrics Identify Uncompetitive Nodes, not SMEs

High betweenness and closeness centrality values can be seen for all the nodes in Appendix Table B, but those nodes never transmitted scientific preprints (see "Out Edges" column). The Twitter handles in Appendix Table B are important (in both degree and closeness centrality) to the network, but they do not transmit preprints. A lack of transmission invalidates our second criteria for competitiveness, meaning that users in Appendix Table B are outcompeted in terms of scientific preprint transmission by almost any other user in the network.

Notably, within Appendix Table B are the @biorxivpreprint and @medrxivpreprint automated accounts that post URLs and titles from preprints submitted to those respective servers. At this time, the automated behavior of these accounts does not include mentioning or Retweeting other users, so these two accounts are not full participants in the scientific discourse. These accounts originated Tweets including preprint URLs but did not link their posts to other users. Instead, these accounts are often referenced by mention or Retweet as other users discuss scientific information, explaining the high importance but lack of out degree for these two automated accounts, see Figure 1.

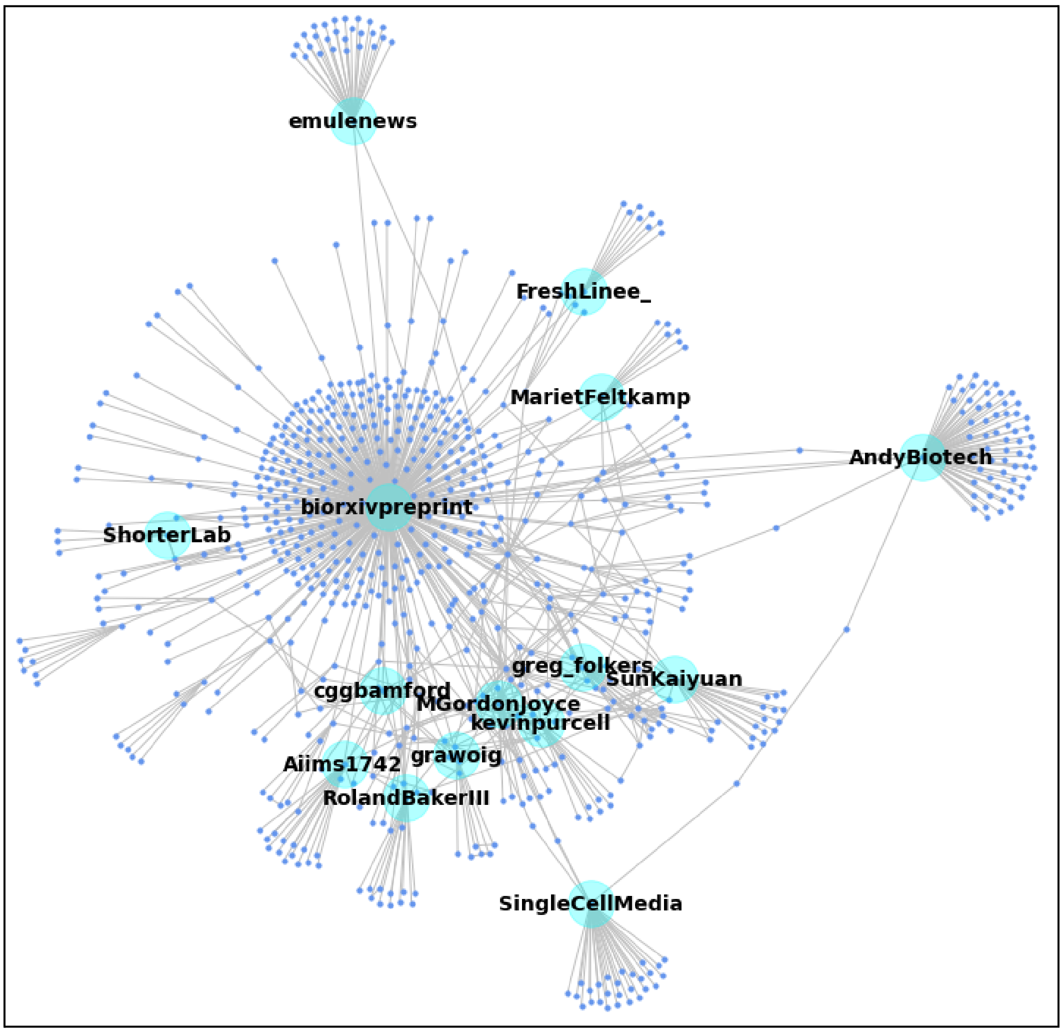

Figure 1. Simplified (ego) network graph of the @biorxivpreprint handle, identifying top 15 competitive handles via percolation centrality, employing PRPS as the percolation state for each node. Note that proximity with @biorxivpreprint, indicates that the users were interacting with the automated @biorxivpreprint account as a source of data.

3.2. Twitter-Rxiv Ecosystem Improvement with Metadata

3.2.1. Percolation Centrality

While both closeness and betweenness centrality can identify competitive nodes, neither metric worked well to identify potential SMEs. Path exclusive metrics on network graphs do not account for information, so other approaches are necessary [9], [10]. In response to this problem, Piraveenan et al., 2013 developed percolation centrality to quantify the role of users in spreading information in a network. For percolation centrality, a value (percolation state) is assigned to each user, representing their capacity to spread information along the network. When the percolation state of all nodes is equal, the percolation centrality algorithm simplifies to betweenness centrality. Percolation centrality enables differentiation among users based on their information quality, or conversely, their misinformation potential.

3.2.2. Peer-review Percentage Score to Represent Misinformation

To address the misinformation potential of a given user, we chose to represent the evidence-based nature of the Rxiv preprints with a novel metric, the "peer-review percentage score" (PRPS). The PRPS represents the ratio between the number of preprints that were eventually published in a traditional, peer-reviewed journal and the total number of preprints shared by a given Twitter user in the timeframe. That value was calculated as follows for each user:

$$PRPS = \frac{num.\ of\ preprints\ peer\ reviewed\ and\ published\ as\ of\ April\ 2021}{total\ num.\ of\ originated\ Tweets\ including\ a\ preprint}$$

PRPS is a retrospective metric applying the traditional process of peer-review to the information shared by each user. We calculated the PRPS for a given user by examining scientific preprints that had achieved peer-reviewed publication after at least six months. We chose six months [11], [12] to be an appropriate minimum span of time for authors to post the preprint on the Rxiv servers and follow through with subsequent publication in a traditional academic journal.

The PRPS measures a user's capacity for transmitting information that is consistent with community-accepted scientific publication standards, applying peer-review as a proxy for evidence-based information. Lower PRPS values represent the transmission of fewer peer-reviewed manuscripts (less evidence) and higher PRPS values represent the transmission of more peer-reviewed manuscripts (more evidence). For example, consider a user who posted two Tweets in our data set, where each Tweet links a different preprint. At the time of the Rxiv API request, only one of those preprints was published. This hypothetical user would have a PRPS of 0.50 (50%). A user who originated multiple Tweets, but only ever referenced a single preprint (that was later found to have been published) would have a PRPS of 100% or 1.00. Any users who never originated a Tweet, but were only mentioned by other users, would have a PRPS of "N/A". We designated the PRPS values as the percolation state for all users when calculating percolation centrality for the networks, enabling us to identify competitive users (potential

SMEs) that spread high-quality information in the form of preprints as benchmarked by eventual peer-reviewed publication.

3.3. Computational Identification of SMEs using the PRPS

Twitter handles who 1) held high PRPS and 2) were centrally important to the network-at-large could be considered as the most competitive nodes in the data set (Appendix Table C) using the percolation centrality algorithm. In order to shorten the list of all users to a smaller subset of prospective SMEs, we developed and applied filtering criteria (see Table 2) to the data.

Table 2. Criteria and rationale for filtering the data set. Useful as a proof-of-concept for implementing PRPS as a metric to identify SMEs by reducing human workload.

Criteria	Rationale
Top 25th percentile for PRPS	Allows for the comparison of users who interacted with different numbers of preprints.
Two or more unique preprints	SMEs should engage with multiple manuscripts.
Non-zero number of both in & out edges	SMEs should participate in dialogue with others about scientific literature.
Top 10th percentile for betweenness	SMEs should occur on the network's shortest paths.
Top 10th percentile for closeness	SMEs should have short paths to other users.

This filtering resulted in a list of 35 names, the top 15 are by percolation centrality are seen in Appendix Table 3. These 15 users are the Twitter handles most likely to share evidence-based scientific results from our data set. These 35 represent a 99.65% reduction in data from the original sample of 10,053 users. No government or public health authority accounts appear in Appendix Table 3. The users identified by these criteria, and especially PRPS, allow for the programmatic identification of SMEs who improve the scientific dialog online.

4. Discussion

We investigated purveyors of PSM in the Twitter-Rxiv ecosystem in three ways:

1) What is the network of public discourse involving preprints?
2) Can the network be augmented to represent the information quality transmitted by members of the network?
3) Are computational methods useful for identifying SMEs who spread evidence-based science information?

To answer these questions, we created a novel data set by merging publicly available information from the Twitter and Rxiv preprint server APIs. We calculated a novel metric (PRPS) to represent the quality of information spread by a user, based on the publication fate of the preprints that they referenced. The implementation of PRPS as a percolation state, revealed prospective SME users. These SMEs are acknowledged in retrospect as those who spread preprints that were eventually peer-reviewed and traditionally published.

4.1. Twitter-Rxiv Ecosystem Patterns

4.1.1. Failure of Authority Accounts in Amplifying Preprints Shared on Twitter

For this investigation, we treated Retweets the same as originating a Tweet, without assigning a penalty to users who spread the words of others. The purpose of this consideration was to account for public health authority handles that amplified scientific results (preprints) on Twitter early in the pandemic. Most surprisingly, none of the Twitter handles for government or public health authorities transmitted preprints. Appendix Table B demonstrates that for the top 25 nodes ranked by closeness centrality, several of which are government and public health authorities, none had outbound edges. In effect, discourse on Twitter was overwhelmingly imbalanced: the public shared scientific preprints with authorities, but not vice versa. The authorities did not pass on the information.

For example, the Rxiv handles are automatic (bot) accounts that Tweet when a preprint is released by the servers. Other bot accounts, like @cryoEM_Papers, certainly exist in the data set, so bot behaviors could account for some of the observed network characteristics. However, the noteworthy presence of presumably human-managed handles for authorities like @CDC, @CNN, @realDonaldTrump, and @WHO in the top 25 demonstrate that bot activity does not explain all Twitter interactions involving scientific preprints and COVID-19.

4.1.2. Centrality Metrics Identify Uncompetitive Nodes, not SMEs

Appendix Table B also highlights the ratio between in and out edges for the top 25. Centrality metrics applied to the Twitter-Rxiv ecosystem could identify important nodes, but these nodes were often information sinks instead of sources. Those accounts received but did not amplify scientific preprints. Thus, another method was needed to identify misinformation and/or SMEs.

4.2. Twitter-Rxiv Ecosystem Improvement with Metadata

4.2.1. Percolation Centrality

Percolation centrality, and other information reliant centrality measures could be seen as a way to measure "contagion" [13]. We chose to implement percolation as a measure of the viral amplification of scientific results through preprints. This approach allowed the retrospective application of traditional scientific publishing standards and the purposeful exclusion of "popularity" metrics like lift. The implementation of the PRPS and percolation centrality serve to provide a contrasting approach for scoring the capabilities of an individual user with regards to misinformation.

4.2.2. Peer-review Percentage Score to Represent Misinformation

For the formula in 2.2.2, users who never originated a Tweet including a preprint would cause a divide by zero (0) error. A PRPS cannot be calculated, so both "N/A" and "0" PRPS were grouped together as "0". As such, it is important to note that as implemented here, the percolation centrality metric using PRPS indicates the ability of a node to spread higher-quality scientific information, but the reverse conclusion is not immediately true. As PRPS values decrease, the potential for transmitting science misinformation increases. High PRPS scores indicate a relatively higher information quality, due to the increase in evidence-based results. However, low PRPS scores *do not guarantee* the presence of misinformation, merely its potential or lack of participation in transmitting preprints. With the above caveats and the strict criteria (found in Section 2.3 and Table 2) applied to filter the users, it is more feasible to identify the best users than the worst in terms of transmitting evidence-based information.

4.3. Computational Identification of SMEs using the PRPS

As shown in Appendix Table D, the use of PRPS allowed the automatic filtering of the initial data set of 10,053 users down to 35. The presence of a handle in this list is not an endorsement, as some of those accounts may have been subsequently banned on Twitter. The filtering served to identify active Twitter accounts who were important to the Twitter-Rxiv ecosystem and interacted with preprints that were eventually peer-reviewed. This study did not investigate the content, tone, or implied meanings of the language in users' Tweets when interacting with preprints. Appendix Table D represents users who met our listed criteria, but these filters do not cover a user's intention. Further human labor is necessary to identify the motivations for those who transmit scientific results. Is a user amplifying an evidence-based discovery, or are they interacting adversarially with scientific publications? Both the work of "fact-checkers" and the repeatability of scientific results are still of utmost importance. This work acts as a proof-of-concept for programmatically reducing the workload of content reviewers and allowing them to understand larger networks of discourse.

This use of the PRPS could become a scoring system for individuals who spread peer-reviewed, publishable scientific results. PRPS as percolation can be seen as akin to impact factor [14], where impact factor is to a journal, as PRPS is to a user. Computationally identified SMEs could prove useful for combating online misinformation, by recognizing individuals who bring evidence-based information into the public discourse.

4.4. Identifying purveyors of science misinformation

A challenge for the analyses and metrics proposed here is the implementation of effective solutions to combat online misinformation. The results published here are meant to discover how

Twitter users discuss preprints and users' roles in transmitting PSM. Such results are not meant to conclusively solve the problem of online misinformation, but instead to inform the development of improved community and editorial guidelines. Evidence-based information needs to outcompete both dis- and misinformation in order to benefit the public good through public health and societal norms.

The prospective SMEs identified here are more likely to transmit claims with substantial evidence-based scientific information in the Twitter-Rxiv ecosystem. However, those SMEs may not be as popular or famous on Twitter, so their contributions may be unable to consequentially impact the greater public discourse. Another challenge for the retrospective PRPS metric is the time involved for traditional publication practices, which can often take months [11], [12]. In the rapidly changing environment of scientific publication, a standard is useful to the scientific community as a means of ensuring the quality of research and discourse [4], [15], [16].

4.5. Future Work

Next steps for this research are the improvement of the PRPS metric by including more parameters for prospective SMEs. The inclusion of metrics for lift will further the development of PRPS as a potential "influencer score". The networks could also be improved by accounting for the relative rate of preprints' versus peer-reviewed publications' transmission speed through the network or how the information contained within a publication's title affects the online dissemination of the manuscript. These research ideas will lead to a more precise understanding of the observed differences in competitiveness for high and low-quality scientific information, potentially leading to the naive prediction of nascent echo chambers based solely on user network topological data and manuscript metadata.

This study introduces a novel, simple, retrospective metric in PRPS to measure the science-based information sharing of social media users on Twitter. The implementation of PRPS and the percolation centrality algorithm to analyze a user interaction network enabled the identification of potential SMEs who spread more evidence-based information.[2] PRPS and percolation centrality demonstrate how scientific results could outcompete misinformation online: through the promotion or lift of potential SMEs. Such a solution is aligned with the global, distributed nature of scientific research while also serving to address potential bias, confounding decisions, or conflicts of interest from paid expert fact-checkers [17]–[19].

Further, the use of PRPS provides a metric to identify the relative quality of information within an online social media community by serving as a measure of dissemination for evidence-based communications. Much evidence exists for misinformation susceptible communities on social media, especially pertaining to the safety and efficacy of vaccines [20]–[23]. An "anti-algorithm" or "anti-recommender" system based on a metric like PRPS could prove disruptive to echo chambers, by identifying topologically close SMEs that could be connected to individuals with low PRPS. The effects of such an approach can be investigated by future research.

Social media is a powerful tool for the rapid broadcast of scientific results and the dissemination of evidence-based policies. There is a trend towards the acceptance of preprints and social media as respected sources over traditional peer-review and broadcast and print media. Joining progressive and traditional practices together via a metric like PRPS represents a valuable compromise for high-quality scientific discourse.

4.6. Conclusions

This paper illustrates the final research product resulting from a team of diverse students of many educational stages and backgrounds in cyber intelligence-based research. We chose a real-world dataset of discussion of scientific preprints on SARS-CoV-2 virus and COVID disease on Twitter ™. The selection of the real-world dataset was driven by: (a) misinformation regarding COVID-19 disease and SARS-CoV-2 virus is rampant and undermines our ability to recover from the pandemic, (b) unfounded and false health-related claims are spreading on social media, and (c) the rapid dissemination of health misinformation provides challenging competition with information broadcast by public health or government authorities such as the World Health Organization and the U.S. Centers for Disease Control and Prevention. Thus, we focused on the close symbiosis between preprints, preprint servers (like bioRxiv and medRxiv), Twitter, scientific researchers, journalists, and the public that developed in the early months of the pandemic (e.g. first six months of 2020). This symbiosis, the "Twitter-Rxiv ecosystem", led to the rapid dissemination of results before traditional processes of scientific peer-review before publication. While much of the work in preprints is well-intentioned, concerns have been raised that this symbiosis may be exploited to disseminate spurious results or intentionally incorrect information. In response we constructed networks to represent public discourse surrounding scientific preprint literature on Twitter and develop metrics to score users within these networks. One such metric, peer-review percentage score, is useful for calculating the network prominence (i.e. influence) of a user while weighting that user for the quality of information propagated by the user. Peer-review percentage score can be used to identify subject-matter experts who transmit evidence-based information online. We found that these subject-matter experts outcompeted public health authorities in online forums by transmitting scientific results. Subject-matter experts engaged with the public whereas public health authorities did not.

Acknowledgements

The authors appreciate Dr. Michael Simpson and Dr. R. Marielle Hjelm for their support and guidance during the execution of the project. We are grateful for the ideas and comments provided by the other members of the program review board while advancing this investigation. A large thank you to Khyati Mahajan for her efforts in data collection, specifically, hydrating the Tweets used in this research, and answering our questions on the dataset. Much thanks to researchers Cameron Clark, Mya Fernandez, Alex Graffeo, Triston Kersenbrock, Ruben Ramirez Martinez, Nathaniel Shepherd, and Tarique Stephens who assisted with other aspects of this investigation. The authors appreciate all the support from the Office of OneIT, University Research Computing staff at the University of North Carolina at Charlotte who have assisted us in using high-performance computing resources for this project.

Funding: This work was supported by the Office of Naval Research (ONR) Navy Reserve Officers' Training Corps (ROTC) Cybersecurity Training Program (grant number N00014-20-1-2804) and the Truist Business Research and Innovation Program.

Footnotes

1. The traditional method of publication in scientific journals requires the solicitation of peer-reviewers and subsequent editorial approval prior to a manuscript's acceptance for publication. This process involves a significant amount of time and overwhelmingly volunteer effort from multiple parties to preserve high levels of integrity for the journal specifically and scientific literature as a whole. A preprint is simply a manuscript posted to an internet server (e.g., aRxiv.org, bioRxiv.org, medRxiv.org or other venues) by the authors without any peer-review. Pre-publication can occur in addition to the traditional process as an "early release" manuscript, depending on a specific journal's editorial policies or the authors' preferences. Thus preprints are often seen as a rapid route to the release of research results, and in combination with Twitter, preprints can be amplified within hours of posting. This amplification comes with a significant caveat, the lack of peer and editorial review.

2. Information contagion can be an NP-hard problem [24]. Approximation methods are important for quick responses to infection on social networks [25], so identifying evidence-based metrics like PRPS could prove useful as an approximate solution. Bad actors leverage traditional metrics for network importance, e.g. bots, fake viral lift, sock puppets, etc. [26]–[30]. Attempts have also been made to subvert evidence and cloak science through an incomplete mimicry of the scientific process [31]–[34]. However, it is the belief of the authors that traditional academic publishing controls have not yet failed, and so far have proven resilient to attack. As such, the incorporation of metrics based on traditional standards for academic research could improve the evidence-based transmission online through the identification of potential SMEs.

References

[1] H. Wang *et al.*, "Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21," *Lancet*, Mar. 2022, doi: 10.1016/S0140-6736(21)02796-3.

[2] H. Else, "How a torrent of COVID science changed research publishing - in seven charts," *Nature*, vol. 588, no. 7839, p. 553, Dec. 2020, doi: 10.1038/d41586-020-03564-y.

[3] V. H. Murthy, "Confronting Health Misinformation: The U.S. Surgeon General's Advisory on Building a Healthy Information Environment [Internet]," US Department of Health and Human Services, 2021. [Online]. Available: https://www.hhs.gov/sites/default/files/surgeon-general-misinformation-advisory.pdf

[4] N. Fraser *et al.*, "The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape," *PLoS Biol.*, vol. 19, no. 4, p. e3000959, Apr. 2021, doi: 10.1371/journal.pbio.3000959.

[5] P. Pradhan *et al.*, "Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag," *bioRxiv*, p. 2020.01.30.927871, Jan. 31, 2020. doi: 10.1101/2020.01.30.927871.

[6] "Report for: Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag." https://biorxiv.altmetric.com/details/74957328 (accessed Apr. 21, 2023).

[7] R. Sever *et al.*, "bioRxiv: the preprint server for biology," *bioRxiv*, p. 833400, Nov. 06, 2019. doi: 10.1101/833400.

[8] E. Chen, K. Lerman, and E. Ferrara, "Tracking Social Media Discourse About the COVID-19 Pandemic: Development of a Public Coronavirus Twitter Data Set," *JMIR Public Health Surveill*, vol. 6, no. 2, p. e19273, May 2020, doi: 10.2196/19273.

[9] A. Landherr, B. Friedl, and J. Heidemann, "A Critical Review of Centrality Measures in Social Networks," *Business & Information Systems Engineering*, vol. 2, no. 6, pp. 371–385, Dec. 2010, doi: 10.1007/s12599-010-0127-3.

[10] K. Das, S. Samanta, and M. Pal, "Study on centrality measures in social networks: a survey," *Social Network Analysis and Mining*, vol. 8, no. 1, p. 13, Feb. 2018, doi: 10.1007/s13278-018-0493-2.

[11] R. J. Abdill and R. Blekhman, "Tracking the popularity and outcomes of all bioRxiv preprints," *Elife*, vol. 8, Apr. 2019, doi: 10.7554/eLife.45133.

[12] B.-C. Björk and D. Solomon, "The publishing delay in scholarly peer-reviewed journals," *J. Informetr.*, vol. 7, no. 4, pp. 914–923, Oct. 2013, doi: 10.1016/j.joi.2013.09.001.

[13] R. S. Mehta and N. A. Rosenberg, "Modelling anti-vaccine sentiment as a cultural pathogen," *Evolutionary Human Sciences*, vol. 2, 2020, doi: 10.1017/ehs.2020.17.

[14] "Journal Citation Reports," *Clarivate*, Oct. 21, 2022.

https://clarivate.com/products/scientific-and-academic-research/research-analytics-evaluation-and-management-solutions/journal-citation-reports/ (accessed Apr. 21, 2023).

[15] C. Watson, "Rise of the preprint: how rapid data sharing during COVID-19 has changed science forever," *Nat. Med.*, vol. 28, no. 1, pp. 2–5, Jan. 2022, doi: 10.1038/s41591-021-01654-6.

[16] F.-X. Coudert, "The rise of preprints in chemistry," *Nat. Chem.*, vol. 12, no. 6, pp. 499–502, Jun. 2020, doi: 10.1038/s41557-020-0477-5.

[17] S. Jiang, R. E. Robertson, and C. Wilson, "Bias Misperceived:The Role of Partisanship and Misinformation in YouTube Comment Moderation," *ICWSM*, vol. 13, pp. 278–289, Jul. 2019, Accessed: Mar. 30, 2022. [Online]. Available: https://ojs.aaai.org/index.php/ICWSM/article/view/3229

[18] C. Lim, "Checking how fact-checkers check," *Research & Politics*, vol. 5, no. 3, p. 2053168018786848, Jul. 2018, doi: 10.1177/2053168018786848.

[19] E. Vincent and M. Mestre, "Crowdsourced Measure of News Articles Bias: Assessing Contributors' Reliability," in *SAD/CrowdBias@ HCOMP*, 2018, pp. 1–10. [Online]. Available: http://ceur-ws.org/Vol-2276/paper1.pdf

[20] A. S. Bradshaw *et al.*, "Propagandizing anti-vaccination: Analysis of Vaccines Revealed documentary series," *Vaccine*, vol. 38, no. 8, pp. 2058–2069, Feb. 2020, doi: 10.1016/j.vaccine.2019.12.027.

[21] J. M. Sharfstein, "Vaccines and the Trump Administration," *JAMA*, vol. 317, no. 13, pp. 1305–1306, Apr. 2017, doi: 10.1001/jama.2017.2311.

[22] C. B. Joseph, "Anti/Vax: Reframing the Vaccination Controversy," *J. Med. Libr. Assoc.*, vol. 108, no. 1, p. 147, Jan. 2020, doi: 10.5195/jmla.2020.826.

[23] J. Roozenbeek *et al.*, "Susceptibility to misinformation about COVID-19 around the world," *R Soc Open Sci*, vol. 7, no. 10, p. 201199, Oct. 2020, doi: 10.1098/rsos.201199.

[24] G. Alexe, S. Alexe, Y. Crama, S. Foldes, P. L. Hammer, and B. Simeone, "Consensus algorithms for the generation of all maximal bicliques," *Discrete Appl. Math.*, vol. 145, no. 1, pp. 11–21, Dec. 2004, doi: 10.1016/j.dam.2003.09.004.

[25] M. Shapiro and E. Delgado-Eckert, "Finding the probability of infection in an SIR network is NP-Hard," *Math. Biosci.*, vol. 240, no. 2, pp. 77–84, Dec. 2012, doi: 10.1016/j.mbs.2012.07.002.

[26] S. Kumar and N. Shah, "False Information on Web and Social Media: A Survey," *arXiv [cs.SI]*, Apr. 23, 2018. doi: 10.1145/nnnnnnn.nnnnnnn.

[27] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, "The rise of social bots," *Commun. ACM*, vol. 59, no. 7, pp. 96–104, Jun. 2016, doi: 10.1145/2818717.

[28] S. Kumar, J. Cheng, J. Leskovec, and V. S. Subrahmanian, "An Army of Me: Sockpuppets

in Online Discussion Communities," *arXiv [cs.SI]*, Mar. 21, 2017. [Online]. Available: http://arxiv.org/abs/1703.07355

[29] N. Shah, H. Lamba, A. Beutel, and C. Faloutsos, "The many faces of link fraud," in *2017 IEEE International Conference on Data Mining (ICDM)*, New Orleans, LA, Nov. 2017. doi: 10.1109/icdm.2017.140.

[30] V. S. Subrahmanian *et al.*, "The DARPA Twitter Bot Challenge," *arXiv [cs.SI]*, Jan. 20, 2016. [Online]. Available: http://arxiv.org/abs/1601.05140

[31] L.-M. Yan, S. Kang, and S. Hu, *Unusual Features of the SARS-CoV-2 Genome Suggesting Sophisticated Laboratory Modification Rather Than Natural Evolution and Delineation of Its Probable Synthetic Route*. 2020. doi: 10.5281/zenodo.4028830.

[32] J. Donovan and J. Nilsen, "Cloaked Science: The Yan Reports," Jan. 03, 2021. https://mediamanipulation.org/case-studies/cloaked-science-yan-reports (accessed Mar. 30, 2022).

[33] A. Harris, "New Peer Reviews: Yan Report's Claims that SARS-CoV-2 Was Created in a Chinese Lab Are Misleading and Unethical," *Rapid Reviews COVID-19*, Oct. 2020, Accessed: Mar. 30, 2022. [Online]. Available: https://rapidreviewscovid19.mitpress.mit.edu/pub/r94z275c/download/pdf

[34] K. L. Warmbrod, R. M. West, N. D. Connell, and A. G. K. Gronvall, "In Response: Yan et al Preprint Examinations of the Origin of SARS-CoV-2," Johns Hopkins Bloomberg School of Public Health, Sep. 2020. doi: 10.1002/jmv.25817.

Appendix A

List of Abbreviations

API	application programming interface
DOI	digital object identifier
ONR	Office of Naval Research
PRPS	peer-review percentage score
PSM	potential science misinformation
ROTC	Reserve Officers' Training Corps (United States Navy)
SME	subject-matter experts
URL	uniform resource locator

Appendix B

Appendix Table B. Top 25 handles in the Twitter-Rxiv ecosystem network, sorted by decreasing closeness centrality. Although the highlighted government and public health authorities are central, those users are "dead ends" (with no outgoing edges).

User Handle	Closeness	Degree	In Edges	Out Edges	Percolation	PRPS	Number Preprints
biorxivpreprint	0.2467	0.0654	365	0	0.3940	0.50	130
Alexis_Verger	0.2338	0.0018	1	9	0.0844	0.43	7
medrxivpreprint	0.2328	0.0383	214	0	0.1835	0.43	379
realDonaldTrump	0.2303	0.0106	59	0	0.1334	0.00	0
Aiims1742	0.2283	0.0034	10	9	0.0563	0.14	7
WHO	0.2264	0.0079	44	0	0.0861	0.00	0
RolandBakerIII	0.2229	0.0029	12	4	0.0171	0.31	16
TheSeeker268	0.2220	0.0090	1	49	0.0495	0.00	1
PandemicCovid20	0.2209	0.0236	124	8	0.1195	0.39	271
kevinpurcell	0.2175	0.0030	2	15	0.0390	0.50	32
Harvard2H	0.2171	0.0009	2	3	0.0185	1.00	4
Delana30183939	0.2132	0.0005	0	3	0.0028	0.44	16
SunKaiyuan	0.2131	0.0047	23	3	0.0219	0.33	3
CDCgov	0.2130	0.0050	28	0	0.0671	0.00	0
greg_folkers	0.2124	0.0030	12	5	0.0130	0.47	34
jflier	0.2119	0.0013	1	6	0.0179	0.00	2
jdm0004	0.2116	0.0004	1	1	0.0027	0.00	2
Micro_BSMT	0.2116	0.0005	0	3	0.0078	0.18	11
svscarpino	0.2115	0.0011	0	6	0.0063	0.33	3
Freitas_DRJ	0.2114	0.0007	0	4	0.0102	0.33	6
NIH	0.2113	0.0020	11	0	0.0121	0.00	0
reddykishore25	0.2112	0.0007	0	4	0.0179	0.00	1
neil_ferguson	0.2106	0.0034	7	12	0.0191	0.00	1
NIAIDNews	0.2105	0.0009	5	0	0.0360	0.00	0
epsilon3141	0.2104	0.0029	1	15	0.0374	0.67	9

Appendix C

Appendix Table C. Top 25 handles in the Twitter-Rxiv ecosystem network, sorted by decreasing percolation centrality. Highlighted government and public health authorities are still central to the network, despite a PRPS weight of "0.00".

User Handle	Closeness	Degree	In Edges	Out Edges	Percolation	PRPS	Number Preprints
biorxivpreprint	0.2467	0.0654	365	0	0.3940	0.50	130
medrxivpreprint	0.2328	0.0383	214	0	0.1835	0.43	379
realDonaldTrump	0.2303	0.0106	59	0	0.1334	0.00	0
SharonM57345162	0.2085	0.0244	0	136	0.1199	0.50	2
PandemicCovid20	0.2209	0.0236	124	8	0.1195	0.39	271
WHO	0.2264	0.0079	44	0	0.0861	0.00	0
Alexis_Verger	0.2338	0.0018	1	9	0.0844	0.43	7
trvrb	0.2084	0.0161	89	1	0.0762	0.50	2
CDCgov	0.2130	0.0050	28	0	0.0671	0.00	0
MackayIM	0.2044	0.0113	60	3	0.0664	0.55	11
krengnath	0.2014	0.0227	0	127	0.0622	0.00	1
cryoEM_Papers	0.1763	0.0120	67	0	0.0612	0.75	4
Aiims1742	0.2283	0.0034	10	9	0.0563	0.14	7
NAChristakis	0.1734	0.0120	66	1	0.0519	0.00	2
TheSeeker268	0.2220	0.0090	1	49	0.0495	0.00	1
WhiteBlabbit	0.1941	0.0165	5	87	0.0478	1.00	3
ferrisjabr	0.1907	0.0322	176	4	0.0454	0.50	2
AndyBiotech	0.2016	0.0118	65	1	0.0448	0.60	5
profvrr	0.1977	0.0086	48	0	0.0441	1.00	1
kevinpurcell	0.2175	0.0030	2	15	0.0390	0.50	32
epsilon3141	0.2104	0.0029	1	15	0.0374	0.67	9
NIAIDNews	0.2105	0.0009	5	0	0.0360	0.00	0
lwj70	0.1676	0.0213	0	119	0.0348	0.00	2
V2019N	0.2089	0.0057	28	5	0.0330	0.27	11
t2438	0.2003	0.0007	0	4	0.0324	0.57	7

Appendix D

Appendix Table D. Top 15 most competitive Twitter handles as identified by PRPS and filtering criteria (Table 2), sorted by descending percolation centrality. Note the lack of government and public health authority accounts. In Edges and Out Edges are abbreviated with (Edg.)

User Handle	Percolation	PRPS	Rank PRPS	Number Preprints	In Edg.	Out Edg.	Rank Betweenness	Rank Closeness
trvrb	0.0762	0.50	0.7508	2	89	1	0.9970	0.9923
MackayIM	0.0664	0.55	0.7680	11	60	3	0.9975	0.9900
WhiteBlabbit	0.0478	1.00	0.8903	3	5	87	0.9980	0.9234
ferrisjabr	0.0454	0.50	0.7508	2	176	4	0.9984	0.9178
AndyBiotech	0.0448	0.60	0.7703	5	65	1	0.9959	0.9868
kevinpurcell	0.0390	0.50	0.7508	32	2	15	0.9950	0.9984
epsilon3141	0.0374	0.67	0.7746	9	1	15	0.9971	0.9957
alykhansatchu	0.0279	0.50	0.7508	6	16	22	0.9953	0.9245
DrEricDing	0.0256	0.50	0.7508	4	26	1	0.9941	0.9907
BillyBostickson	0.0231	0.63	0.7713	8	2	9	0.9919	0.9862
PeterHotez	0.0200	1.00	0.8903	2	4	1	0.9966	0.9842
florian_krammer	0.0194	0.75	0.7790	4	25	4	0.9869	0.9275
thelonevirologi	0.0188	0.55	0.7683	20	32	1	0.9837	0.9885
Harvard2H	0.0185	1.00	0.8903	4	2	3	0.9902	0.9982
GermHunterMD	0.0178	0.67	0.7746	3	3	1	0.9912	0.9166