NUDGING USERS TO SLOW DOWN THE SPREAD OF FAKE NEWS IN SOCIAL MEDIA

Christian von der Weth Jithin Vachery Mohan Kankanhalli

School of Computing, National University of Singapore
{chris| jithin|mohan}@comp.nus.edu.sg

ABSTRACT

The success of fake news spreading on social media is to a large extent because of normal users unknowingly parroting or sharing such content. Educational interventions such as information campaigns or revised curricula aim to raise users' awareness and critical thinking skills. However, these are long-term efforts with an uncertain outcome. In this paper, we present ShareAware, our prototype for nudging users into a more conscious posting and sharing behavior. ShareAware uses linguistic analyses to infer the factuality of content and credibility of sources before being posted or shared. The results provide users with immediate feedback to discourage the dissemination of questionable content. We demonstrate the benefits of our approach using a series of simulations.

Index Terms— fake news, social media, user nudging

1. INTRODUCTION

> *"Falsehood flies, and the truth comes limping after it"* (Jonathan Swift, 1710)

In recent years, the adverse effects of "fake news" came under public scrutiny, particularly within politics, health, and finances. While disinformation has long been used to shape people's thoughts and behavior, social media has amplified its adverse effects significantly. Firstly, never has it been so easy to access, publish, and share information, including the deployment of so-called social bots, i.e., software-controlled user accounts. Secondly, journalistic norms such as objectivity and balance are often forgotten, ignored, or purposefully dismissed. And lastly, besides bots, users often share content without fact-checking, especially when it contains controversial or emotionally charged content. Current countermeasures to fake news can be classified into three categories:

From a legal perspective, anti-fake news laws impose fines for fake news and enforce platform providers to curate or remove content. However, fake news and related notions are not well defined, making it very difficult to put such notions into a legal framework. It would also make governments the arbiters of truth. This form of hard paternalism raises legitimate concerns of censorship and misuse [1]. Moreover, manual fact-checking and re-actively taking down false information do not scale and typically take effect only after the damage is done.

From a technological perspective, a plethora of methods to tackle the publication and diffusion of disinformation have been proposed. The most common research directions concern the identification of social bots, automated credibility assessment, and fact-checking. Most efforts utilize state-of-the-art data mining and machine learning techniques to distinguish between human users and social bots or between false or truthful content; see Section 2. These tasks are inherently very challenging and can be viewed as a cat-and-mouse game – any improvements in detection will result in the development of improved bots or better fake content.

From a societal perspective, the success of disinformation can be attributed to two human "flaws". Firstly, disinformation often mobilizes the user's cognitive biases and heuristics, making it more likely for users to fall for it [2]. Secondly, disinformation is typically novel, controversial, emotionally charged, or partisan, making it more "interesting" and hence more likely to be shared [1]. Educational interventions such as public information campaigns or reformed school curricula aim to improve users' critical thinking skills as well as their digital literacy in general. However, these are either one-time or long-term efforts with uncertain outcomes [3] .

Summing up, legal and technical solutions focus on the "bad guys" – that is, social bots and human users ranging from individuals to state actors that intentionally generate and spread disinformation. However, the success of fake news going viral also strongly depends on normal users without any malicious intentions of sharing disinformation [1]. While information campaigns might (temporarily) raise users' awareness, such educational interventions typically happen outside of the context of users' everyday social media use. Posting and sharing content on social media is typically very situational actions that often do not reflect users' attitudes. This is particularly true due to the "infective" nature of fake news.

In this work, we motivate a hybrid approach using technological solutions to provide in-situ educational interventions to nudge the "good guys" into a more conscious social media use. Nudging is a form of soft paternalism to guide users by suggesting, instead of enforcing a certain behavior. To this end, we present ShareAware, which analyzes content before

978-1-7281-1485-9/20/$31.00

being posted or shared. If the analyses raise any concerns about posts' factuality or credibility of the sources the users get notified. We implement ShareAware as a browser extension to enable seamless integration into user's social media experience. We present and discuss the current features of our ShareAware platform, provide preliminary results towards its effectiveness, as well as outline our long-term research goals. We argue that our approach both combines and complements current legal, technological, and societal solutions.

2. RELATED WORK

Social bot detection. Most methods to identify social bots use supervised machine learning leveraging on the user, content, social network, temporal features, etc. (e.g. [4, 5]). The underlying assumption is that genuine human users and social bots exhibit sufficiently distinct behavior patterns. In contrast, unsupervised methods aim to detect social bots by finding accounts that share strong similarities with respect to social network and (potentially coordinated) posting/sharing behavior (e.g. [6, 7]). The challenge is that bots with malicious intent get better and better at blending into the population of genuine user accounts.

Manual and automated fact-checking. Fact-checking websites such as Snopes or Politifact validate or debunk popular claims including personal statements, rumors, urban legends, etc. Sites such as Media Bias/Fact Check (MBFC) evaluate news sites regarding their credibility and biases. Also, most social platforms hire dedicated staff to fact-check submitted content. However, manual fact-checking scales very poorly with the amount of information published online. Various solutions for automated fact-checking have been proposed (e.g., [8, 9]). However, fully automated fact-checking system are far from mature [10]. More common are hybrid solutions where humans are supported by automated systems.

User nudging in social media. What to post or share is often very situational and rushed, resulting in users often reporting regrets [11, 12]. A common cause for regret is the (unintentional) self-disclosure of sensitive information. Acquisti et al. evaluated series of privacy nudges [13]. For example, users see a subset of contacts who will be able to read a post to remind users of their audience. Another cause for regret is a blunder, which is concerned with mistakes and factuality issues [12]. Nekmat [14] conducted a series of user surveys to evaluate the effectiveness of fact-check alerts (e.g., the reputation of a news source). The results show that such alerts trigger users' skepticism, thus lowering the likelihood of sharing information from questionable sources.

3. ShareAware PLATFORM

Our ShareAware platform aims to nudge users into a more conscious posting and sharing behavior. Before a user posts or shares content, ShareAware analyzes it and makes subtle educational interventions if the content raises concerns about its level of certainty and/or credibly of its sources. In this section, we use the fictitious tweet *"Trump said that supermarkets might fail to restock food."* in the midst of the COVID-19 pandemic as a running example.

3.1. Existing User Account & Link Analysis

For Twitter, we can leverage on existing efforts towards social bot detection as a proof of concept to utilize such information as user nudges. More specifically, we use the Botometer API [4] returning a score between 0 (low) and 5 (high), representing the likelihood that a Twitter account is a bot. We adopt this score but color-code it for visualization: green (0-2), yellow (2-3), orange (3-4), red (4-5); see Figure 3.

ShareAware displays credibility information for linked content in a social media post. To this end, we collected data for 2.7k+ online news sites provided by Media Bias Fact Check (MBFC).[1] MBFC assigns each news site one of six factuality labels and one of nine bias or category labels. We provide an API that for a given URL returns the corresponding credibility information, for example (simplified):

```
{
  "unshortened": "https://www.breitbart.com/entertainment
       /2020/01/24/f-donald-trump-rapper-yg-arrested-...",
  "domain": "www.breitbart.com",
  "rating": {
    "ratings": {
      "fact": "mixed factual",
      "bias": "fake news"
    },
    "labels": {
      "fact": ["not always credible", ...],
      "bias": ["questionable source", "not credible", ...]
    }
  }
}
```

3.2. Triple Extraction

Textual content (e.g., social media posts) represents unstructured information. To support downstream tasks such as factuality degree annotation (see below), we convert text into a structured, triple-based representation. This refers to the task of Open Information Extraction, which is most commonly applied to statements from factual content (e.g., to enrich knowledge bases); see [15]. In contrast, we focus on posts that can express doubts, uncertainty, counterfactuals, etc.

Our triple extraction algorithm takes as input a dependency graph [16]; see Figure 1 for an example. From this graph, we derive all triples based on universal dependencies[2] (edge labels). To preserve the connection between statements within on sentence, we extract three types of triples:

Subject-Predicate-Object: spo-triples represent basic, standalone statements such as *"supermarkets restock food"*,

[1] https://mediabiasfactcheck.com/

[2] https://universaldependencies.org/

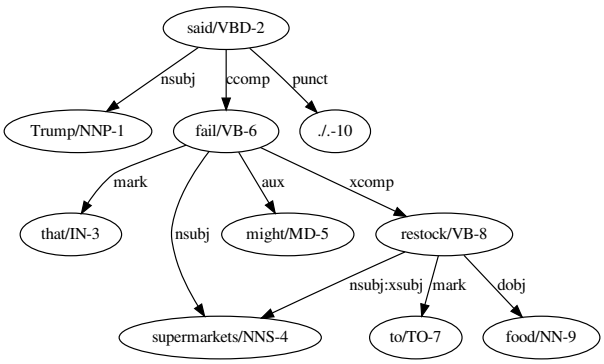


Fig. 1: Dependency graph for example sentence *"Trump said that supermarkets might fail to restock food."*

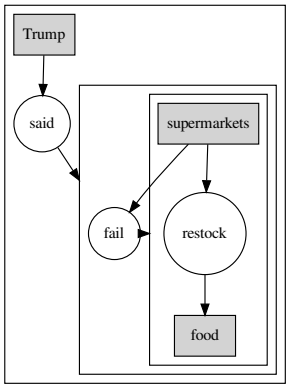


Fig. 2: Graph representation of extracted triples from example sentence *"Trump said that supermarkets might fail to restock food."* The nested boxes reflect the tree-like reference structure between all statements.

typically indicated by the presence of an object (e.g., `dobj`). Note that the object can be null (e.g., *"supermarkets closed."*)

Subject-Predicate-Reference: spr-triples represent statements about statements, i.e., the target of an *spr*-triple is a reference to another triple. Common cases are causal complements (`ccomp`) such as *"said that [...] failed"*), and open clausal complements (`xcomp`) such as *"failed to restock"*). *spr*-triples are of particularly importance when it comes to evaluate the factuality of their referenced statements.

Reference-Predicate-Reference: rpr-triples connect two statements. For example, adverbial clause modifiers (`adcvl`) such as *"Trump was blunt when Obama visited"* connect the statements *"Trump was blunt"* and *"Obama visited"*.

When applied to the example post, our triple extraction algorithm returns the following set of triples (simplified):

ID:1	Trump	→	said	→	ID:2
ID:2	supermarkets	→	fail	→	ID:3
ID:3	supermarkets	→	restock	→	food

In this example, ID:1 and ID:2 are *spr*-triples and ID:3 is an *spo*-triple. Figure 2 visualizes the triple set as a hierarchical statement graph. Note that for each non-reference node in a triple, we retain important information such as negation, adverbs, adjectives, and auxiliaries that can affect on the factuality of a statement. For example, we keep the information that *"might"* is an adverbial modifier of *"fail"*.

	positive	**negative**	**underspec.**
certain	`CT+`	`CT-`	`CTu`
probable	`PR+`	`PR-`	n/a
possible	`PS+`	`PS-`	n/a
underspec.	n/a	n/a	`Uu`

Table 1: Possible factuality degrees (n/a = impossible degree) [17]

3.3. Factuality Annotation

Our example post makes the following assertions regarding factuality: (a) Trump is not certain that supermarkets fail, (b) Trump is not certain that supermarkets did *not* restock food, and (c) the author herself does not assert the correctness of the statements about supermarkets but only (d) that Trump made those statements. To formalize these assertions, we adopt the framework proposed by Saurí and Pustejovsky [17].

Factuality degree. Linguistic literature commonly agrees on four discrete levels of certainty (*certain*, *probable*, *possible*, *underspecified*) and three polarity values (*positive*, *negative*, *underspecified*); Table 1 shows the resulting set of possible factuality degrees. For example, `PR-` and `PS-` is associated with statements that are not probable and not certain, respectively. The factuality degree of a statement can be modified by individual words and whole phrases. For example, given the statement *"supermarkets might fail to restock food"*, the core statement *"supermarkets restock food"* (`CT+`) gets modified twice: Firstly, it inherits the level of uncertainty from *"supermarkets might fail"*. (`PS+`). And secondly, *failed (to)* is a simple implicative verb that flips the polarity of its referenced statement. As a results, the factuality degree of *"supermarkets restock food"* evaluates to `PS-` (unlikely).

Source. By default, the source of a statement in a sentence is the writer of the sentence, denoted by s_0. However, a sentence might contain *source-introducing predicates* (*SIPs*) that introduce new sources in the discourse. Common SIPs are predicates of report (e.g., *say*, *tell*), predicates of knowledge (e.g., *know*, *forget*), predicates of belief and opinion (e.g., *think*, *consider*), and others. For example, *"Trump said [...]"* introduces *Trump* as a new source which is asserting the factuality of the following statement(s). Note that this assertion made by source s_{Trump} is itself an assertion made by the author, denoted with s_{trump_0}. In principle, sources can be arbitrarily nested. For example, the sentence *"Trump said that Obama claimed that Hillary did X"* yields the source $s_{hillary\_obama\_trump\_0}$ (among s_0, s_{trump_0}, etc.). This nesting of sources also means that a statement has a factuality degree for each "outer" source. In the previous example, for instance, the writer s_0 is uncommitted regarding the factuality of statement *X* (this is true for most SIPs since they push any following assertions to the new source) .

Factuality calculation (by example). With these two notions of source and factuality degree, we systematically assign each statement (here: triple) one or more factuality labels depending on the mentioned sources and linguistic information

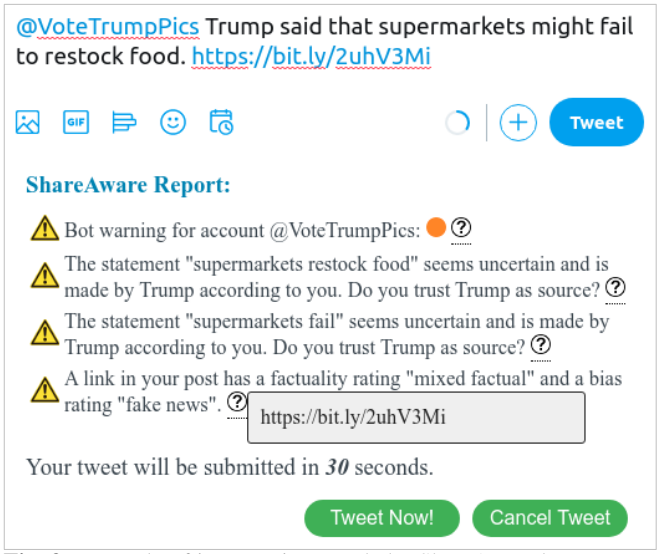


Fig. 3: Example of interventions made by ShareAware browser extension when submitting a new tweet.

(adverbs, implicative verbs, SIPs, etc.). An individual factuality label is a tuple $\langle s, fd \rangle$ of a source s and a factuality degree fd. A detailed discussion of the algorithm is beyond the scope of the paper, and we point the interested reader to [17]. When applied to our example post, the algorithm returns the following factuality labels for the three triples:

ID:1	Trump	$\rightarrow$	said	$\rightarrow$ ID:2
	$\{\langle s_0, CT+\rangle\}$			
ID:2	supermarkets	$\rightarrow$	fail	$\rightarrow$ ID:3
	$\{\langle s_0, Uu\rangle, \langle s_{trump\_0}, PS+\rangle\}$			
ID:3	supermarkets	$\rightarrow$	restock	$\rightarrow$ food
	$\{\langle s_0, Uu\rangle, \langle s_{trump\_0}, PS-\rangle\}$			

The annotations reflect our intuitive assertions we made at the beginning of this section. For example, in the label $\{\langle s_0, Uu\rangle, \langle s_{trump\_0}, PS-\rangle\}$ for ID:3, $\langle s_{trump\_0}, PS-\rangle$ reflects assertions (b/c) and $\langle s_0, Uu\rangle$ reflects assertion (d).

3.4. Educational Interventions

We implemented the frontend of ShareAware as a browser extension for the seamless, integration of factuality and credibility information into users' normal social media experience. The browser extension intercepts post and share requests and injects factuality and credibility information in terms of warning messages. Using the results of the user account and link analysis is straightforward (cf. Section 3.1). Factuality labels, on the other hand, can be interpreted in various ways; and how to display them is an on-going research question. For the time being, we use the following heuristics – we warn users if: (a) a statement made by the author is considered uncertain (factuality is neither `CT+` nor `CT-`), (b) a statement is made by another source instead of the author; in this case, we ask the author to reconsider the source's credibility. Figure 3 shows the resulting warning messages for our example post when a user would submit it as a new tweet. Given these warnings, the user can change the post accordingly (e.g., link to a more credible news source), immediately submit it "as is", or wait for the post the be automatically submitted after a delay.

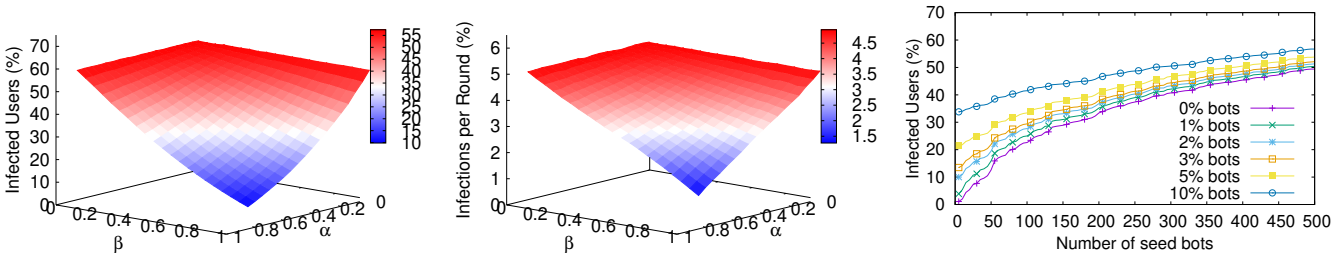
4. EVALUATION

To get some first insights into the effectiveness of ShareAware, we leverage on existing surveys showing that users respond to factuality nudges [14] and conduct a series of simulations to evaluate the effect of such nudges on the information diffusion with a social network.

Dissemination Model & Methodology. We model a social network as a directed graph with N nodes representing the users, and *Independent Cascade Model* [18] captures the propagation of fake news. We use four graphs: one graph derived from a real-world Twitter follower network (Higgs dataset) [19] and three random graphs (scale-free graph, Erdős-Rényi graph, and a clustered graph simulating partisan networks). Each graph contains bN (randomly chosen) bots and $(1-b)N$ non-bots, i.e., regular users. Each bot forwards a fake news message with probability $p_{bot} = 1.0$. Without nudging, the probability p_{user} that a user forwards a fake news message is proportional to the outdegree of message sender. This reflects that users are more likely to share messages from influential users. We simulate information dissemination in discrete time steps. At each time step t_i, a user or bot may receive a message and forwards it with probability p_{user} or p_{bot}. An attack, i.e., the initial submission of a fake news message, is started by s seed bots at time t_0.

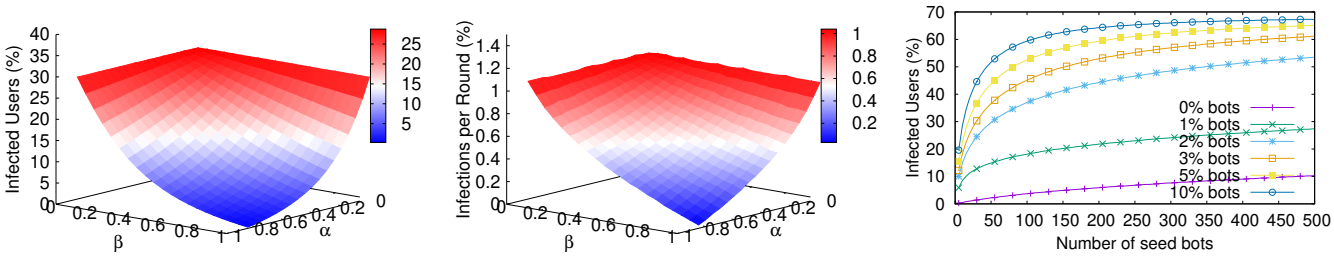
We model the usage of ShareAware using two parameters: α, the fraction of users using our platform, and β, the fraction to which the forwarding probability p_{user} of a user (which uses ShareAware) is decreased when presented with a nudge. For example, with $\beta = 0$ users will not forward the message; with $\beta = 0.5$ the forwarding probability is $0.5p_{user}$. We use two metrics to measure the effectiveness of nudges: (1) the number of infected nodes in the graph, i.e., the number of users that have received and forwarded the fake news message; (2) the number of effected nodes per time step as an estimator for the dissemination time.

Experiments & Results. For each network, we ran 2,000 simulations attacks (i.e., seeding a new fake news message) and measured the number of infected nodes and the number of infections per round. In the case of the clustered network, we assume that all seed nodes and bots are in the same cluster. If not stated otherwise, we assume that 10% of all nodes in the network are bots ($b = 0.1$), and attack is started by 200 seed nodes ($s = 200$). In all the following plots, we show the mean over all attacks. The standard deviation is consistently minimal, and we, therefore, omit their visualization using error bars. Due to space constraints, we only show the results for the Higgs network (Fig. 4) and the clustered network (Fig. 5);



(a) overall dissemination (b) rate of dissemination (c) Effects of #seeds and #bots

Fig. 4: **Higgs network**, #Nodes: 456k, #Edges: 14M



(a) overall dissemination (b) rate of dissemination (c) Effects of #seeds and #bots

Fig. 5: **Clustered Network**, #Nodes: 100k, #Edges: 1M, #Clusters: 4

the results for the other random graphs show similar trends. For the clustered network, we show our two metrics with respect to the nodes that are not part of the attacking cluster.

Unsurprisingly, for all networks, the more users use ShareAware (larger values for α), and the more they are influenced by it (larger values for β), the more contained and slower is a fake news attack. Note that even in the optimal setting where no fake news messages are forwarded by regular users ($\alpha = 1.0$ and $\beta = 1.0$), both the overall number of infected nodes and the infections per round is still noticeable. This is due to two reasons. Firstly, apart from the seed bots, the message is always forwarded by any other bot in the network. Moreover, secondly, we make the worst-case assumption by picking the bots from all nodes with the highest outdegrees (i.e., users with the most followers). In short, successfully nudging users can significantly limit both the spread and the speed that fake news travels across social media, thus limiting its potential consequences. In the context of our running example, this is more likely to reduce the number of people engaging in panic buying. It will also give supermarkets more time to fill their stocks, as well as give governments more time for interventions (e.g., publishing corrective messages or fact checks on both social media and trusted websites). In practice, the main challenges will be to promote the use of platforms such as ShareAware (maximizing α) as well as to provide effective and trusted nudges (maximizing β).

For the sake of completeness, Fig. 4(c) and Fig. 5(c) show the effects of the number of seed bots s and the number of bots b in networks; we report the overall number of infected nodes. In both cases, we set $\alpha = 0.05$ and $\beta = 0.5$, which are arguably more realistic values compared to a very high number of nudged users. As expected, the number of infected nodes increases with increasing values for s and b; however, the increases are sublinear for both parameters.

5. LIMITATIONS & ROADMAP

ShareAware is an early prototype towards user nudging in social media. In this section, we briefly discuss current limitations and provide a roadmap for our ongoing efforts.

Multimodal content analysis. Apart from considering available credibility data about information sources (links or users), we have proposed a linguistic approach to evaluate factuality information of written statements. A natural next step is to extend these efforts to multimodal content, particularly images and videos. This includes digital forensic techniques to spot tampered photos or content retrieval techniques to find related images and videos from credible sources.

Downstream task support. We convert unstructured text into structured triple-based representation to evaluate the factuality labels of each statement. Such a structured representation improves and enables a wide range of downstream tasks such as indexing, searching, and (entity) linking. Effectively and efficiently performing such tasks can further advance fake news detection, e.g., by finding and tracing related statements about the same topic but from different sources.

Deeper semantic understanding. Given two posts *"Trump said A"* and *"Trump said B"*, we treat statements *A* and *B* equally even if the former is a trivial or funny statement while the latter is of great political importance. Arguably, not all fake news is equal with respect to their potential gravitas. To improve factuality nudges – for example, through ranking

or filtering by scoring statements – we need to understand the semantic meaning of a statement better and define meaningful metrics to compare different statements.

UX/UI design. In this paper, we focused on the backend architecture for the analysis of social media content in terms of factuality and credibility information. However, when and how to present this information to users significantly impacts the effectiveness of factuality nudges. An inherent trade-off is the level of detail and ease of understanding. For our current prototype, we mainly use simple templates to generate a warning message based on our analysis results.

Factuality nudges beyond social media. Our in-situ approach using a browser extension makes ShareAware directly applicable to all online platforms beyond social media. For example, we can inject the credibility information of a link source into any website, including search result pages, online newspapers, online forums, etc. Such a more platform-agnostic solution poses additional challenges towards UX/UI design to enable helpful but also smooth user experience.

Towards privacy nudges. Similar to existing works (e.g. [13]), we aim to support privacy nudges to also warn users in case of potentially harmful self-disclosure. In this context, ShareAware analyzes a post with respect to sensitive information that can be extracted or inferred from the content.

6. CONCLUSIONS

Nudging users to slow down the spread of fake news in social media complements existing automated approaches that focus on the "bad guys". In this work, we focused on concerns regarding factuality in posts and the credibility of the sources. Our simulations show that guiding users towards a more conscious posting and sharing behavior can significantly reduce the overall reach and dissemination speed of fake news. Based on these promising first results, in the future, we will investigate how to improve on ShareAware to consider different modalities, e.g., to identify and communicate factuality concerns found in images and videos.

Acknowledgements. This research is supported by the National Research Foundation, Singapore under its Strategic Capability Research Centres Funding Initiative. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore.

7. REFERENCES

[1] S. Vosoughi, D. Roy, and S. Aral, "The Spread of True and False News Online," *Science*, vol. 359, 2018.

[2] G. Pennycook and D.G. Rand, "Who Falls for Fake News? The Roles of Bullshit Receptivity, Overclaiming, Familiarity, and Analytic Thinking," *Journal of Personality*, 2019.

[3] D.M.J. Lazer, M.A. Baum, Y. Benkler, A.J. Berinsky, K.M. Greenhill, F. Menczer, M.J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, M. Schudson, C.R. Sunstein, E.A. Thorson, D.J. Watts, and J.L. Zittrain, "The Science of Fake News," *Science*, vol. 359, 2018.

[4] K.C. Yang, O. Varol, P.M. Hui, and F. Menczer, "Scalable and Generalizable Social Bot Detection through Data Selection," in *AAAI)*, 2020.

[5] Sneha K. and Emilio F., "Deep Neural Networks for Bot Detection," *Information Sciences*, vol. 467, 2018.

[6] N. Chavoshi, H. Hamooni, and A. Mueen, "DeBot: Twitter Bot Detection via Warped Correlation," in *ICDM*. 2016, IEEE.

[7] M. Mazza, S. Cresci, M. Avvenuti, and W. Quattrociocchi, "RTbust: Exploiting Temporal Patterns for Botnet Detection on Twitter," in *WebSci*. 2019, ACM.

[8] G. Karadzhov, P. Nakov, L. Màrquez, A. Barrón-Cedeño, and I. Koychev, "Fully Automated Fact Checking Using External Sources," in *RANLP*, 2017.

[9] N. Hassan, F. Arslan, C. Li, and M. Tremayne, "Toward Automated Fact-Checking: Detecting Check-Worthy Factual Claims by ClaimBuster," in *KDD*. 2017, ACL.

[10] L. Graves, "Understanding the Promise and Limits of Automated Fact-Checking," Tech. Rep., 2018.

[11] L. Zhou, W. Wang, and K. Chen, "Tweet Properly: Analyzing Deleted Tweets to Understand and Identify Regrettable Ones," in *WWW*, 2016.

[12] "I Read My Twitter the Next Morning and Was Astonished: A Conversational Perspective on Twitter Regrets," in *CHI*. 2013, ACM.

[13] A. Acquisti, I. Adjerid, R. Balebako, L. Brandimarte, L.F. Cranor, S. Komanduri, P.G. Leon, N. Sadeh, F. Schaub, M. Sleeper, Y. Wang, and S. Wilson, "Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online," *ACM Comput. Surv.*, vol. 50, no. 3, 2017.

[14] E. Nekmat, "Nudge Effect of Fact-Check Alerts: Source Influence and Media Skepticism on Sharing of News Misinformation in Social Media," *IEEE Transactions on Dependable and Secure Computing*, 2020.

[15] C. Niklaus, M. Cetto, A. Freitas, and S. Handschuh, "A Survey on Open Information Extraction," in *COLING*. 2018, ACL.

[16] D. Chen and C. Manning, "A Fast & Accurate Dependency Parser using Neural Networks," in *EMNLP*, 2014.

[17] R. Saurí and J. Pustejovsky, "Are You Sure That This Happened? Assessing the Factuality Degree of Events in Text," *Comput. Linguist.*, vol. 38, 2012.

[18] David Kempe, Jon Kleinberg, and Éva Tardos, "Maximizing the Spread of Influence through a Social Network," in *SIGKDD*. 2003, ACM.

[19] M. De Domenico, A. Lima, P. Mougel, and M. Musolesi, "The Anatomy of a Scientific Rumor," *Scientific Reports*, vol. 3, 2013.