Fake News Detection Based on Subjective Opinions

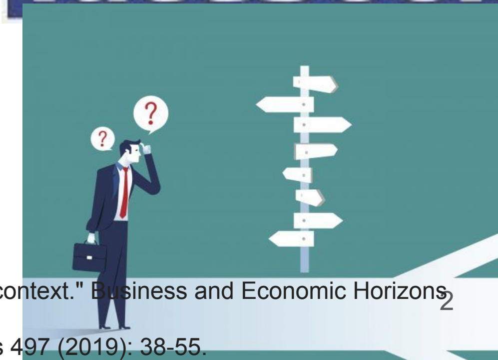
***Presenter:* Danchen Zhang**
***Author:* Danchen Zhang, Vladimir I. Zadorozhny**

School of Computing and Information, University of Pittsburgh
Pittsburgh, Pennsylvania, US

Fake News in social media

A survey shows 62% US adults get news from social media [1], which become a primary source for people to read the news.

Fake news could spread more quickly and widely in our society. More and more fake news detection studies appear [2].

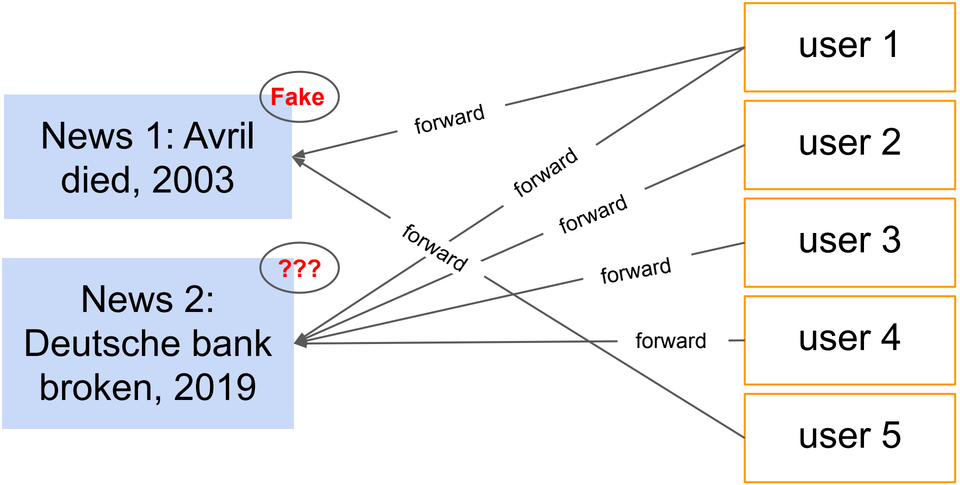


[1] Sihombing, Sabrina Oktaria. "Predicting intention to share news through social media: An empirical analysis in Indonesian youth context." Business and Economic Horizons 13.4 (2017): 468-477.
[2] Bondielli, Alessandro, and Francesco Marcelloni. "A survey on fake news and rumour detection techniques." Information Sciences 497 (2019): 38-55.

Fake news detection scenario

Social media users forward the news articles. Some news is real, and some is fake. A subset of news are labeled by experts.

Goal: predict unlabeled news veracity.



Reliability assessment in fake news detection

User reliability: the possibility of user spreading real news.

News reliability: the possibility of news being real.

Existing models (Castillo [3], Harmonic [4], HC-CB-3 [5], TriFN [6]) assess news and users reliability based on labeled data, and **do not fully consider the uncertainty lead by the unlabeled data**. We propose a **Subjective Opinions** based fake news detection model (SO_fnd).

[3] Castillo, Carlos, Marcelo Mendoza, and Barbara Poblete. "Information credibility on twitter." *Proceedings of the 20th international conference on World wide web*. 2011.

[4] E. Tacchini, G. Ballarin, M. L. Della Vedova, S. Moret, and L. de Alfaro, "Some Like it Hoax: Automated Fake News Detection in Social Networks," in Proceedings of the Second Workshop on Data Science for Social Good, vol. 1960. Skopje, Macedonia: CEUR-WS, 2017

[5] Della Vedova, Marco L., et al. "Automatic online fake news detection combining content and social signals." *2018 22nd Conference of Open Innovations Association (FRUCT)*. IEEE, 2018.

[6] Shu, Kai, Suhang Wang, and Huan Liu. "Exploiting tri-relationship for fake news detection." *arXiv preprint arXiv:1712.07709* (2017).

Subjective Opinions

Subjective Opinions could represent the **probability** affected by the degrees of **uncertainty**, and **Subjective Logic** is a calculus for subjective opinions [7].

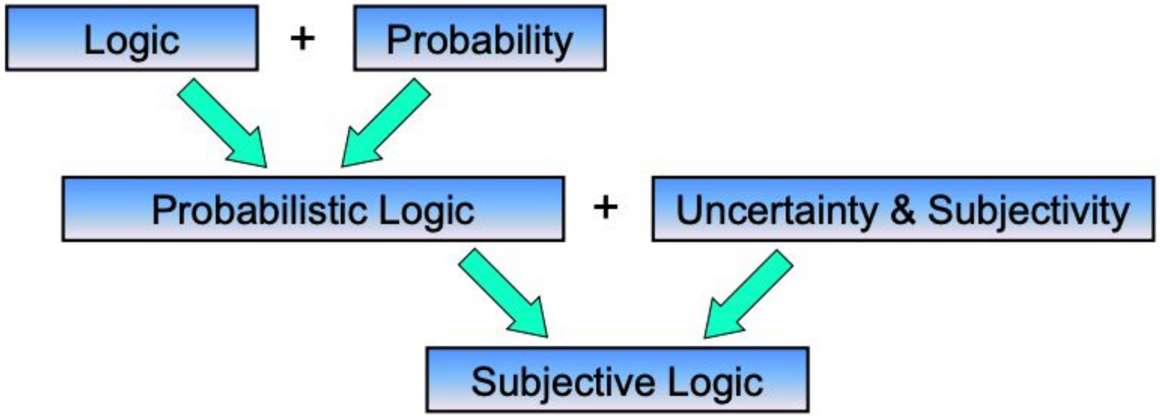


Figure from Josang & Kaplan Tutorial at FUSION 2019 (http://folk.uio.no/josang/sl/SL-Fusion2019.pdf)

[7] Jøsang, Audun. *Subjective logic*. Heidelberg: Springer, 2016.

Subjective Opinions

Source **A**'s opinion towards a statement **p** can be described as [7]:

$$\omega_p^A = \{trust_p^A, distrust_p^A, uncertainty_p^A\}$$

- Trust + distrust + uncertainty = 1

	Subjective Opinions	Probabilistic Logic
A believe p is true with <u>higher</u> uncertainty.	{0.4, 0.2, 0.4}	0.6
A believe p is true with <u>lower</u> uncertainty.	{0.5, 0.3, 0.2}	0.6

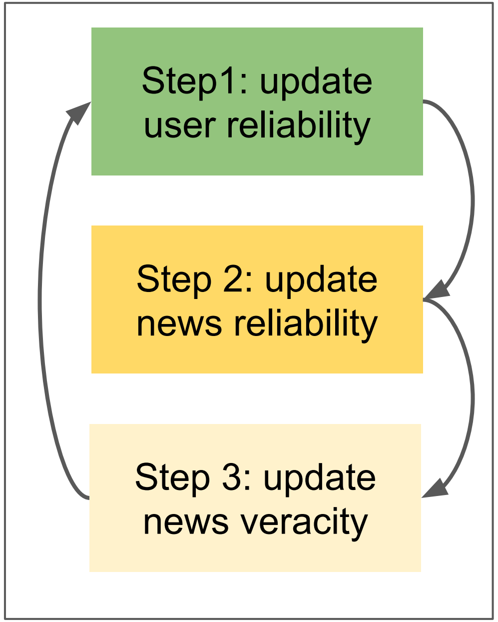
[7] Jøsang, Audun. *Subjective logic*. Heidelberg: Springer, 2016.

Two proposed fake news detection models

Probability (Prob_fnd) & **Subjective Opinions (SO_fnd)**
based fake news detection

Table 2: Comparison of Prob_fnd and SO_fnd

Steps	Prob_fnd	SO_fnd
Step 1: Update *user_reliability*.	Use **Probability** to describe our belief in a user sharing real news. **Ignore unknown cases**.	Use **Subjective Opinions** to describe our belief in a user sharing real news. **Record unknown cases with uncertainty**.
Step 2: Update *news_reliability*.	Use **Average** to fuse related users' reliability.	Use **Consensus** to fuse related users' reliability.
Step 3: Update news veracity.	Predict with *news_reliability*.	Predict with *news_reliability*.



Prob_fnd reliability assessment

$$user\_reliability(user_i) = \frac{\#(r\_n)}{\#(r\_n) + \#(f\_n)}$$

- #(r_n) is count of **real news** user forwards; #(f_n) is count of **fake news**.

$$news\_reliability(news_j) = Average(user\_reliability(related\ users))$$

- SVM classifier predicts news veracity with {news_reliability} as the feature.

SO_fnd reliability assessment

$$user\_reliability = \{user\_trust, user\_distrust, user\_uncertainty\}$$

$$\begin{cases} user\_trust = \frac{\#(r\_n)}{\#(r\_n)+\#(f\_n)+\#(u\_n)} * (1-\alpha), \\ user\_distrust = \frac{\#(f\_n)}{\#(r\_n)+\#(f\_n)+\#(u\_n)} * (1-\alpha), \\ user\_uncertainty = \frac{\#(u\_n)}{\#(r\_n)+\#(f\_n)+\#(u\_n)} * (1-\alpha) + \alpha. \end{cases}$$

- #(r_n) is count of **real news** user forwards; #(f_n) is count of **fake news;** #(u_n) is count of **unlabeled/unpredicted news**.

$$news\_reliability(news_j) = user_1\_reliability \oplus user_2\_reliability \oplus \ldots \oplus user_k\_reliability,$$

- SVM classifier predicts news veracity with {news_distrust} as the feature.

Experimental study

Dataset 1: PolitiFact

- 120 real news & 120 fake news
- 23865 users

Dataset 2: BuzzFeed

- 90 real news & 90 fake news
- 15257 users

Baselines:

- Harmonic [5], HC-CB-3 [6], TriFN [7]

Evaluation metrics:

$$Recall = \frac{TP}{TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$F1 = 2\frac{Precision * Recall}{Precision + Recall}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Experiments Results

Table 3: Repeated 5-fold cross-validation results on two real-world datasets.

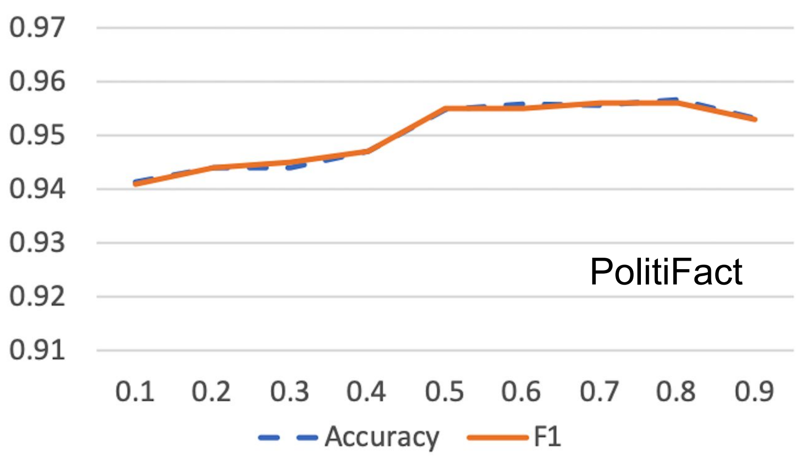
		Prob_fnd	SO_fnd	HC-CB-3	Harmonic	TriFN
BuzzFeed	accuracy	$.852 \pm .055$	$\mathbf{.871 \pm .051}^{**}$	$.856 \pm .052$	$.854 \pm .052$	$.864 \pm .026^{*}$
	precision	$.788 \pm .086$	$.816 \pm .079^{*}$	$.791 \pm .076$	$.782 \pm .075$	$\mathbf{.849 \pm .040}^{**}$
	recall	$.969 \pm .043^{*}$	$.960 \pm .004$	$.966 \pm .045$	$\mathbf{.983 \pm .041}^{**}$	$.893 \pm .013$
	F1	$.866 \pm .052$	$\mathbf{.880 \pm .050}^{**}$	$.867 \pm .050$	$.869 \pm .050$	$.870 \pm .019^{*}$
PolitiFact	accuracy	$.922 \pm .036$	$\mathbf{.953 \pm .029}^{**}$	$.938 \pm .029^{*}$	$.916 \pm .042$	$.878 \pm .020$
	precision	$.887 \pm .056$	$\mathbf{.941 \pm .048}^{**}$	$.899 \pm .057^{*}$	$.876 \pm .074$	$.867 \pm .034$
	recall	$.967 \pm .034^{*}$	$.967 \pm .034^{*}$	$.948 \pm .046$	$\mathbf{.970 \pm .030}^{**}$	$.893 \pm .023$
	F1	$.924 \pm .035^{*}$	$\mathbf{.953 \pm .030}^{**}$	$.921 \pm .041$	$.919 \pm .044$	$.880 \pm .017$

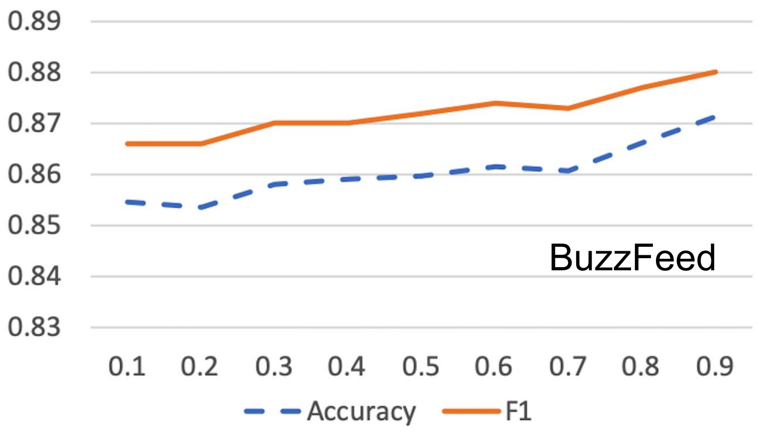
x^{**}: the run with the best performance.
x^{*} : the run with the second best performance.

SO_fnd significantly outperforms Prob_fnd (p-value<0.01).

SO_fnd is robust with parameter α.

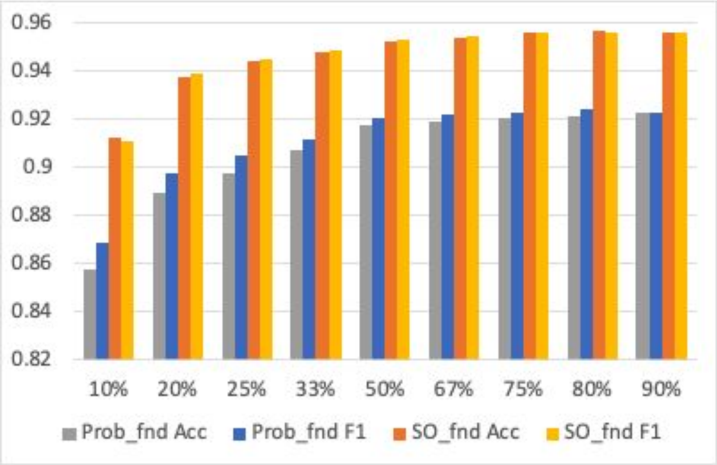
SO_fnd could get similar performance with different α.



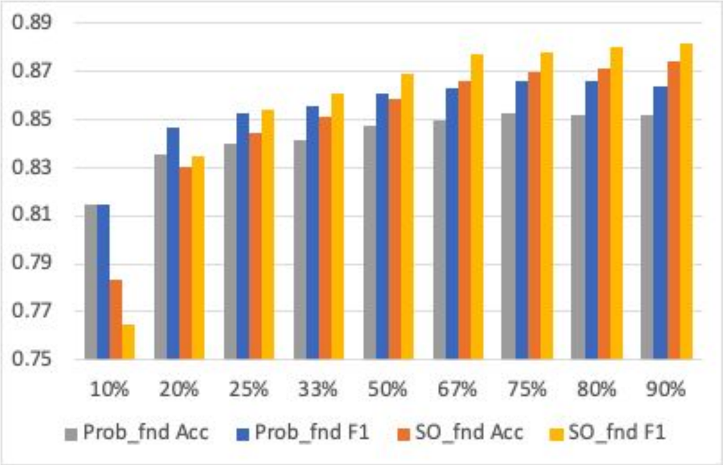


Effect from labeled training data size

Compared to TriFN, our proposed two models are able to achieve a similar or even better performance with much less labeled training data.

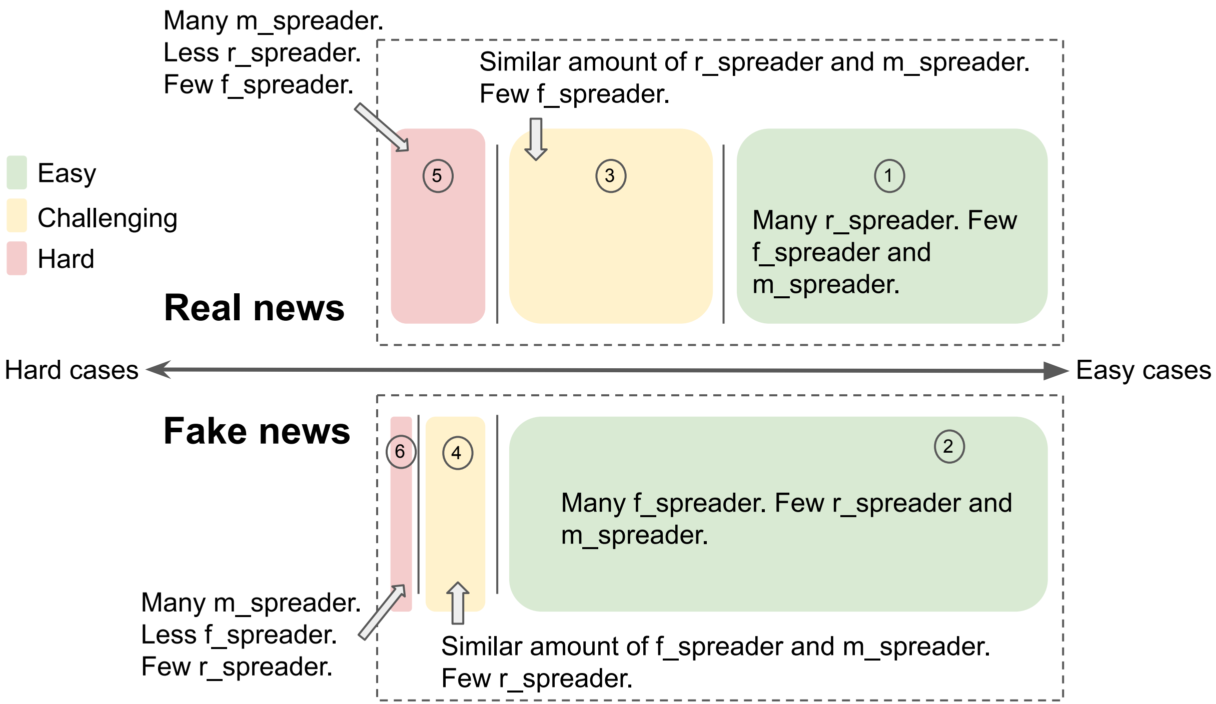


PolitiFact



BuzzFeed

When does Prob_fnd & SO_fnd win and lose?



Real news spreader (r_spreader), user_reliablity >80%; Fake news spreader (f_spreader), user_reliablity <20%; Mixed news spreader (m_spreader), 20%<user_reliablity<80%.

Conclusion and future works

We proposed two new models to evaluate users and news reliability in fake news detection area. Extensive experiments on two real-world datasets, validate our proposed methods' effectiveness. Especially, SO_fnd shows higher performance than state-of-art models.

Next step, we plan to

- introduce **news content** into news reliability assessment;
- Introduce **user profile**, e.g., age and gender, into user reliability assessment.

Thanks

Q&A

Probability & Subjective Opinions based reliability representation

Statement labels	User 1	User 2	User 3
News 1 is real	yes	yes	yes
News 2 is real	yes	yes	yes
News 3 is real	yes		yes
News 4 is **fake**		yes	
News 5 is **fake**		yes	
News 6 unknown			yes
News 7 unknown			yes
News 8 unknown			yes

⇨

User 1 reliability	100%
User 2 reliability	50%
User 3 reliability	100%? 50%? or 75%?

Prob_fnd [1], Harmonic [2], HC-CB-3 [3]

User 1 reliability	{trust = 1, distrust = 0, uncertainty = 0}
User 2 reliability	{trust = 0.5, distrust = 0.5, uncertainty = 0}
User 3 reliability	**{trust = 0.5, distrust = 0, uncertainty = 0.5}**

[1] Zhang, Danchen, et al. "Fake News Detection Based on Subjective Opinions." *European Conference on ADBIS,* Springer, Cham, 2020.
[2] E. Tacchini, et al., “Some Like it Hoax: Automated Fake News Detection in Social Networks,” 2nd Workshop on Data Science for Social Good, 2017
[3] Della Vedova, Marco L., et al. "Automatic online fake news detection combining content and social signals." *2018 22nd FRUCT*. IEEE, 2018.

Subjective Logic consensus operator

Consensus operator: cumulatively fuse subjective opinions.

$$\omega_s^{A,B} = \omega_s^A \oplus \omega_s^B$$

$$\begin{cases} t_S^{A,B} = (t_S^A u_S^B + t_S^B u_S^A)/(u_S^A + u_S^B - u_S^A u_S^B) \\ d_S^{A,B} = (d_S^A u_S^B + d_S^B u_S^A)/(u_S^A + u_S^B - u_S^A u_S^B) \\ u_S^{A,B} = (u_S^A u_S^B)/(u_S^A + u_S^B - u_S^A u_S^B) \end{cases}$$