Computational and Mathematical Organization Theory (2019) 25:60–71
https://doi.org/10.1007/s10588-018-09280-3

S.I. : SBP-BRIMS 2018

FakeNewsTracker: a tool for fake news collection, detection, and visualization

Kai Shu[1] · Deepak Mahudeswaran[1] · Huan Liu[1]

Published online: 13 October 2018

Abstract
Nowadays social media is widely used as the source of information because of its low cost, easy to access nature. However, consuming news from social media is a double-edged sword because of the wide propagation of *fake news*, i.e., news with intentionally false information. Fake news is a serious problem because it has negative impacts on individuals as well as society large. In the social media the information is spread fast and hence detection mechanism should be able to predict news fast enough to stop the dissemination of fake news. Therefore, detecting fake news on social media is an extremely important and also a technically challenging problem. In this paper, we present FakeNewsTracker, a system for fake news understanding and detection. As we will show, FakeNewsTracker can automatically collect data for news pieces and social context, which benefits further research of understanding and predicting fake news with effective visualization techniques.

Keywords Fake news detection · Neural networks · Twitter visualization

1 Introduction

With people spending more time on the social media platforms, they are more prone to consume information from social media. Social media is free of cost, easy to access and help one to express opinions and hence it acts an excellent source for an individual to consume information from social media. But the quality of news on social media is generally lower than the traditional news organizations. It is because anyone can spread information they want in the social media

✉ Kai Shu
kai.shu@asu.edu

Deepak Mahudeswaran
dmahudes@asu.edu

Huan Liu
huan.liu@asu.edu

[1] Arizona State University, Tempe, AZ 85281, USA

and there is no regulating authority to control the information. *Fake news*, as a specific type of disinformation, means the false information that is spread deliberately to deceive people. Some individuals and organization use social media as a tool to spread disinformation for financial and political gains. It was approximated that over million tweets are related to fake news "Pizzagate" by the end of US presidential election. This consequence has adverse effects and the opinions of people are biased because of fake news. Thus, it is important to address this issue.

Fake news detection is an important and technically challenging problem. In an attempt to tackle the growing misinformation, several fact-checking websites have been deployed to expose the fake news. These websites play a crucial role in clarifying fake news, but they require expert analysis which is time-consuming. Numerous articles and blogs are written in order to distinguish fake news from the true news. However, they are not from authority sources and may be biased, which make the labels not fully reliable and convincing. Due to the volume and diversity of the social media, it is almost impossible to manually label the fake news and true news. Also, the information in social media is spread at an alarming rate and hence a framework is required to detect fake news in early stages and avoid dissemination. Thus, to solve these challenges, we present a system FakeNewsTracker, to facilitate the community for studying fake news. Figure 1 shows the various components in the FakeNewsTracker system. The major functionalities of FakeNewsTracker are as follows,

1. *Fake news collection* collecting news contents and social context automatically, which provides valuable datasets for the study of fake news;
2. *Fake news detection* extracting useful features and build various machine learning models to detect fake news;

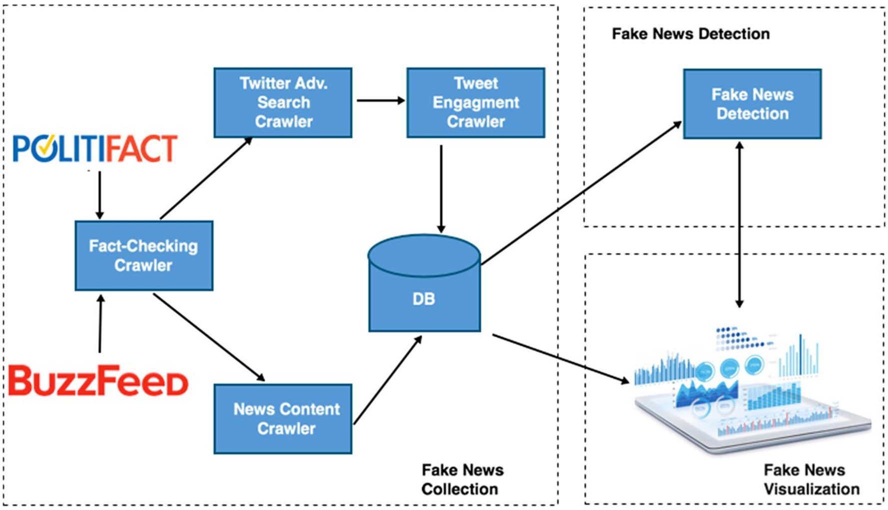

Fig. 1 The flowchart of FakeNewsTracker system

3. *Fake news visualization* presenting the characteristics of fake news dissemination through effective visualization techniques.

2 FakeNewsTracker

We propose an end to end framework for collecting data related to fake news, detecting fake news and visualization of the fake news data which provides insights on the nature of data. The data collected as part of this paper is comprehensive to an extent for fake news detection because it includes both news content and the social context. We use both this information in our classification task and provide a software interface for visualization of the data in different manners.

2.1 Collecting fake news data

Fake news is widely spread across various online platforms. We use some of the fact-checking websites like PolitiFact as a source for collecting fake news information. In these fact-checking sites, fake news information is provided by the trusted authors and relevant claims are made by the authors on why the mentioned news is not true (Wang 2017).

We propose a strategy for collecting fake news in a periodic manner to update the repository. First, we collect the verified fake news and true news from fact-checking websites like PolitiFact on daily basis. Then, using the Twitter's advanced search API, we gather the tweets related to the fake/real news that spread them in Twitter. Moreover, we gather social engagements of users such as replies of tweet, retweet, and favorites through Twitter APIs. Users who interact with these tweets that are related to fake news are more vulnerable to them. If the user likes the tweet related to fake news they are prone to be affected by the fake news. Based on the comments on the retweets we could infer whether the user is able to differentiate fake news or not. In social media, users form social groups and so people in the same group will also be affected by the fake news because of echo chamber effect. So, we also collect the followers and followees of the users who engage with fake news help characterize user features in the detection task. We have publihsed this dataset with detailed description and analysis at Shu et al. (2018c).

2.2 Detecting fake news

Detection of fake news is a difficult task as it is intentionally written to falsify information. We propose Social Article Fusion (SAF) model that uses the linguistic features of news content and features of social context to classify fake news. We formulate the fake news detection as a binary classification problem (Shu et al. 2017b).

2.3 News representation learning

We can use linguistic features likes news content to find the clues between fake news and real news. Although fake news is intentionally written to appear similar to fake news studies have shown that the language style used to falsify information and the topic content could be a factor for determining fake news. For using the news content in our classification, we use autoencoders to learn the news content in low dimensional latent feature space (Bowman et al. 2015).

Auto-encoders are widely used to represent the text content in lower dimensional space. So, we also use auto encoder to capture the text content of the news articles in the lower dimensional space z. Given a new article text document $\mathcal{X} = \{x^1, x^2, \dots, x^m\}$ where x^i is the ith word in the document we find the representation $z \in \mathbb{R}^{k\times1}$. The encoder E $: \mathcal{X} \rightarrow \mathcal{Z}$ learns a mapping function to map input document to lower dimensional space z. The decoder $D = \mathcal{Z} \rightarrow \mathcal{X}$ learns the mapping to reconstruct the original document from the representation z.

The encoder network takes an input word at each time stamp and produces an output and the hidden state of the last timestamp represents the content representation z.

The decoder takes the output from time stamp $t-1$ and previous timestamp's hidden state for predicting the word at time stamp t. The output of the decoder is given as

$$s_t = f_{dec}(s_{t-1}, \hat{x}_t)$$

where s_{t-1} is the hidden state of the decoder at time t and $s_0 = z$ and $\hat{x}_t$ is the output symbol from the timestamp $t-1$ which becomes to the decoder at time t. The output of the decoder is fed to a fully connected dense layer for mapping it to one of the words in the vocabulary and it produces an output $y_t \in \mathbb{R}^{|v|\times1}$ at each timestamp t.

The autoencoder tries to minimize the reconstruction error from representation z to original document $\mathcal{X}$. The reconstruction error is given by

$$\mathcal{L}_{AE} = L(x, D(E(x))$$

We use cross entropy as the loss function. The cross-entropy error for the reconstruction of the document is given as

$$\mathcal{L}_{AE} = \sum_{i}^{m} \left(y^i\left(\log\left(y_t^i\right)\right) - \left(1 - y_i\right)\left(\log\left(1 - y_t^i\right)\right)\right)$$

2.3.1 Social engagement learning

Social engagements could be an another major feature for the fake news detection task (Shu et al. 2018a). Social context provides valuable information about how users interact with the fake news and the true news. The interactions of the users on the social media change over the period of time. To capture the temporal engagements (Ruchansky et al. 2017) of the users with the fake news we have used Recurrent Neural Networks (RNN). Studies have shown that recurrent neural network performs efficiently for capturing the temporal relationships and hence we use this for our problem. Social engagements like tweets and their replies are embedded in a certain format by the embedding layer before giving it to the network. The output of the RNN is considered as the social context feature for the classification. Long Short-Term Memory (LSTM) has been used in our experiments as it solves the long-range dependencies and vanishing gradient problem.

Social engagements involve users interacting with the news article over a period of time. In order to capture the user information who engaged with the news article we have performed SVD on the user-news engagement matrix $E \in \mathbb{R}^{u \times a}$ where $E_{ij} = 1$ if there is engagement by user i on the article j and $E_{ij} = 0$ when there is no engagement by the user i on the article j. We decompose the engagement matrix using SVD as follows

$$E = U\Sigma V$$

here U is the latent user feature that captures the user article engagement matrix. The text content in the social engagement provide useful information as user may provide their viewpoint on the news articles. The text content is represented using the doc2vec (Le and Mikolov 2014) algorithm implemented in genism.

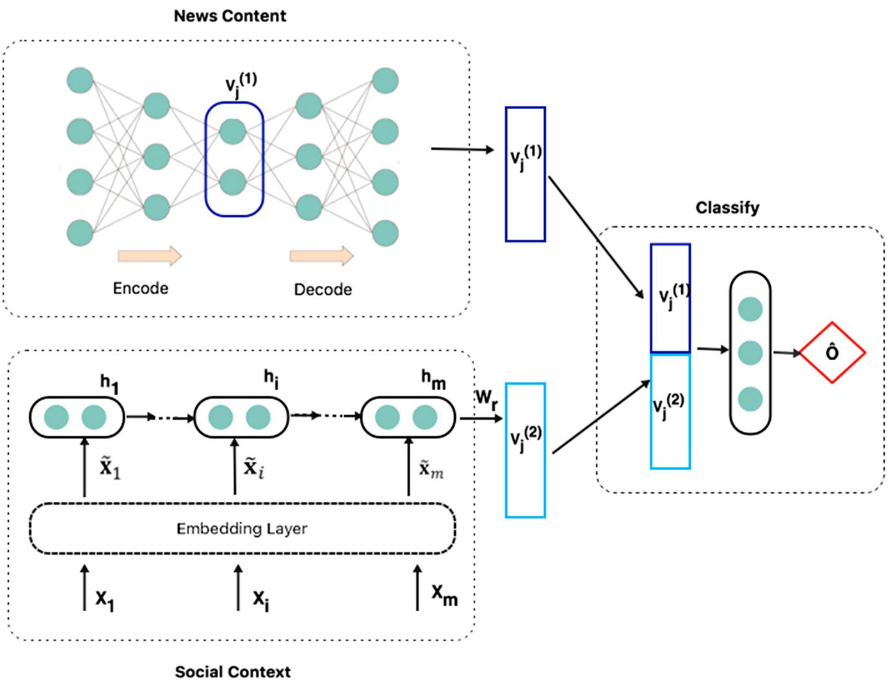

Fig. 2 Fake news detection model. It integrates both news content and social context into a coherent model to learn news feature for prediction

2.3.2 Social article fusion

Social Article Fusion model combines the features generated by the auto-encoder (Bowman et al. 2015) and social context recurrent neural network as shown in Fig. 2 and concatenates them together to form a single concatenated feature vector V_j for the classification. Then weights are learned to map the combined features to classified labels. The softmax layer is used to classify the news articles in the output layer. The output of the SAF network is given by

$$y_o = softmax(WV_j)$$

We train both the feature learning and classification tasks together so that the features are learned relative to detection task rather than capturing plain linguistic differences and social engagements. We use cross entropy as loss function for output prediction and it is given by

$$\mathcal{L}_{pred} = \left(y^i_{label}\left(\log\left(y^i_o\right)\right) - \left(1 - y^i_{label}\right)\left(\log\left(1 - y^i_o\right)\right)\right)$$

For our experiment, we optimize the following loss function with regularization on the concatenated features in order to avoid overfitting.

$$Loss = \mathcal{L}_{pred} + \mathcal{L}_{AE} + \sum_{i=1}^{n} V_i^2$$

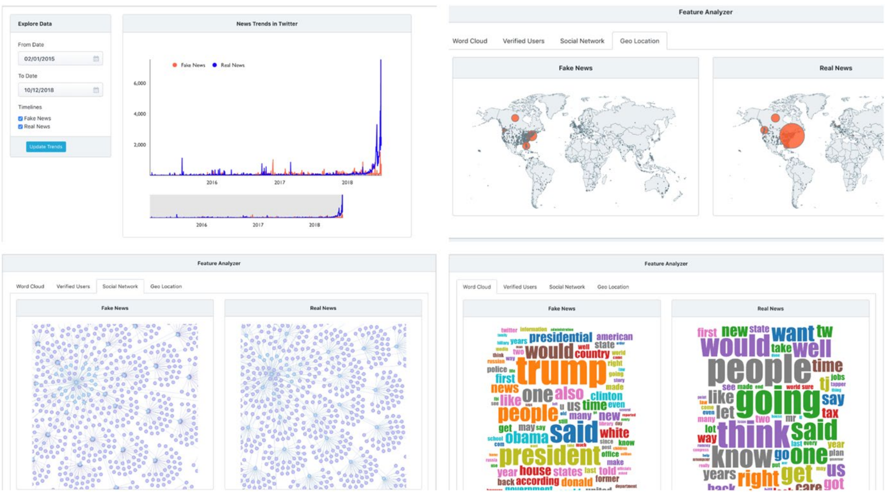

Fig. 3 FakeNewsTracker fake news visualization interface

2.4 Visualizing fake news

We have developed a fake news visualization as shown in Fig. 3 for the developing insights (Morstatter et al. 2013) on the data. We have developed various interfaces for visualizing the data from our dataset. For identifying the differences in the news content of the true news and the fake news we have used word cloud representation of the words for the textual data (Fig. 4). From Fig. 3 we can search for fake news within a time frame and identify the relevant data. Also, we have provided the comparison of feature significance and model performance as part of this dashboard.

Using the geolocated fake tweets (Luke and Morgan 2015) as shown in Fig. 5 we could identify how the fake news is spread around certain areas of United States as the news pieces collected were mostly related to US politics. Using the network of the users (Shu et al. 2018b; Castillo et al. 2011) as shown in Fig. 6, we visualize the social network to identify the differences between the users who interact with the fake news and the true news. Using these visualizations, we see differences between the user characteristics such as the social network and geolocation.

Fig. 4 Word cloud of fake news content and true news content

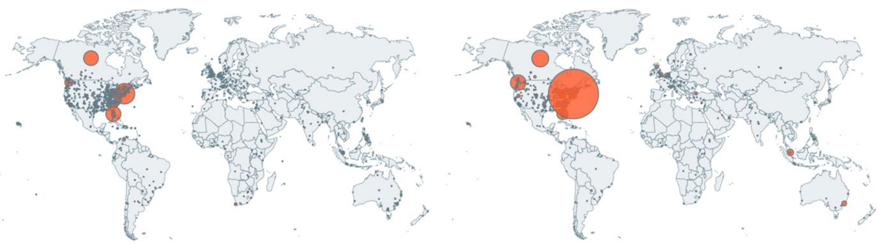

Fig. 5 Geo-visualization of the tweets of the fake news and the true news

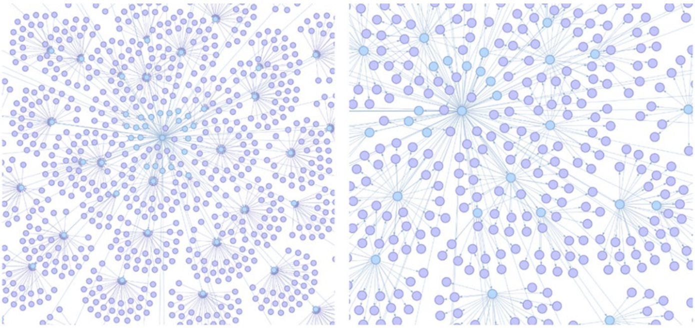

Fig. 6 The social network of the users who share fake news and true news

Table 1 Statistics of the dataset

Platform	PolitiFact	BuzzFeed
# candidate news	394	174
# true news	197	87
# fake news	197	87
# users	18,013	6118
# social engagements	23,106	9740

3 Experiments

In this section, we present the experimental results to evaluate the effectiveness of the Social Article Fusion model. Also, we try to answer the following questions through our evaluation.

(1) Can social media data be used for the fake news detection tasks?
(2) Can we learn representation of the text content and social engagements that could be effectively used for the fake news detection tasks?

3.1 Data

The dataset for the experiment is collected using the approach mentioned in the Sect. 2.1. We have collected the data from the fake checking websites in streaming manner and search in Twitter for any social engagements related to the news pieces. Also, seconds order engagements including favorites, retweets and replies are collected for recent tweets on periodic basis to get comprehensive temporal engagements on the previously collected tweets. From the Table 1 we can see that there is rich social information available from the dataset collected which could be further explored in different manners.

Table 2 Performance comparison for fake news detection

Dataset	Metric	SAF/S	SAF/A	SAF
PolitiFact	Accuracy	0.633	0.620	**0.670**
	Precision	**0.625**	0.595	0.600
	Recall	0.541	0.594	**0.891**
	F1	0.578	0.595	**0.717**
BuzzFeed	Accuracy	0.623	0.571	**0.742**
	Precision	0.526	0.444	**0.777**
	Recall	**0.714**	0.286	0.500
	F1	0.606	0.348	**0.608**

Bold values indicate better results compared to other models

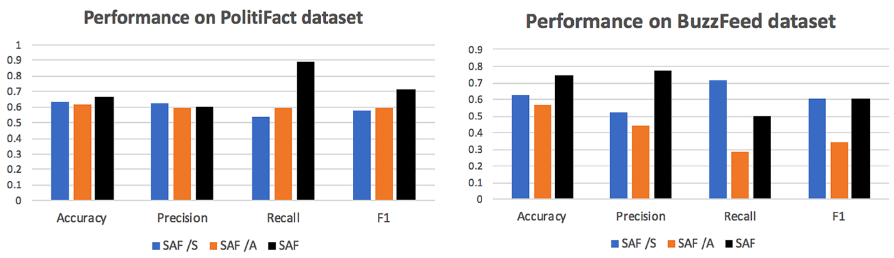

Fig. 7 Performance comparison of fake news detection on PolitiFact and BuzzFeed

3.2 Experimental design

In our experiment setting, for training the news article content we have set a maximum threshold on the sequence length as a hyperparameter with values at 100 because of long range dependencies problem in the RNN. In the encoder, we have used deep LSTMs with 2 layers with 100 cells at each layer and 200-dimensional word embeddings with an input vocabulary of 5000 words. Similarly, for the decoder we have 2 layers of LSTMs with 100 cells in each layer. The word embeddings are randomly initialized and they are learned along with the network.

Similarly, for the social engagements we have set a threshold of 100 engagements and selected the engagements based on a priority. For the twitter engagements, we have given first priority to the second order engagements like replies because they provide more useful information to identify fake news as users are more likely to provide their opinions on the news article that first level engagement where a user usually shares an article link. For social context network, we have used similar network architecture used for the encoder.

Table 3 Performance of using semi-supervised learned features from SAF for detection task

Dataset	Metric	SVM	LR	GNB
PolitiFact	Accuracy	0.684	0.683	0.620
	Precision	0.607	0.606	0.574
	Recall	0.919	0.910	0.729
	F1	0.731	0.728	0.643
BuzzFeed	Accuracy	0.543	0.542	0.486
	Precision	0.455	0.453	0.409
	Recall	0.714	0.786	0.643
	F1	0.555	0.579	0.500

3.3 Fake news detection performance

- To understand the importance of the social and news context in the fake news detection task we have performed experiments with some variations of our framework, detailed as follows.

  - *SAF/S* which uses only news article content.
  - *SAF/A* which utilizes only social context.
  - *SAF* which exploits both news article contents and social engagements.

From Table 2 and Fig. 7 we interpret the importance of each feature in the detection task. It is observed there is an improvement in performance in both the datasets when both the news article content and social context are used. From this we can answer our first question that the social media engagements could be used to detect fake news.

- To answer second question, the features are learned in a semi supervised manner in the SAF setting. The concatenated feature vector V_j from SAF network is taken as the learned feature for the classification task. To test the quality of the learned features we try the classification task using the learned feature with standard machine learning algorithms like Support Vector Machine, Logistic Regression and Naïve Bayes. From Table 3, we observe that the feature learned are a good representation with respect to the classification of the fake news articles. Thus, we can learn features from the SAF network for the classification task.

4 Related work

Fake news detection has attracted the interest of researchers in recent years and several approaches (Shu et al. 2017b) have been proposed. Recently there are works (Shu et al. 2017a, b; Wang 2017) on using text content of the news for the detection task. Wang (Wang 2017) uses CNN for the classification of the fake news content.

Shu et al. (2017a) uses the latent content embedding of the document as one of the features for the detection task. There are several other works which makes use of text content.

Ruchansky et al. (2017) uses the social engagements at post level to capture the differences in temporal engagement patterns between fake and real news. Since people express their emotions towards news through social media post and so it is reasonable to use social media posts as a potential feature for feature detection. Shu et al. (2018b) uses various features of the user engaging with the news articles to identify the fake news.

5 Conclusion and future work

In this paper, we provide a system FakeNewsTracker, which provides general solutions for data collection, interactive visualization, and analytical modeling towards fake news detection. Fake news data collection strategy is provided to collect data and a deep learning-based solution is provided to detect fake news. We have used linguistic and social engagements features for the detection task. Also, the software interface developed is useful for visualization to interpret result and identify new pattern in misinformation in social media.

There are several interesting options for future work. One is to make use of other features available in the dataset like favorites, retweets and social network and learn features for the fake news detection. Also, our proposed framework could be extended to detect fake news in real time as it is implemented in streaming manner.

Acknowledgements This material is based upon work supported by, or in part by, the ONR Grant N00014-16-1-2257, ARO (W911NF-15-1-0328) and ONR N000141310835.

References

Bowman SR, et al (2015) Generating sentences from a continuous space. arXiv:1511.06349

Castillo C, Mendoza M, Poblete B (2011) Information credibility on twitter. In: Proceedings of the 20th international conference on World wide web. ACM

Le Q, Mikolov T (2014) Distributed representations of sentences and documents. International conference on machine learning, pp 1188–1196

Luke S, Morgan J (2015) Who tweets with their location? Understanding the relationship between demographic characteristics and the use of geoservices and geotagging on Twitter. PLoS ONE 10(11):e0142209

Morstatter F, et al (2013) Understanding twitter data with tweetxplorer. In: Proceedings of the 19th ACM SIGKDD international conference on knowledge discovery and data mining. ACM

Ruchansky, N, Seo S, Liu Y (2017) CSI: a hybrid deep model for fake news detection. In: Proceedings of the 2017 ACM on conference on information and knowledge management. ACM

Shu, K, Wang S, Liu H (2017a) Exploiting tri-relationship for fake news detection. arXiv:1712.07709

Shu K et al (2017b) Fake news detection on social media: a data mining perspective. ACM SIGKDD Explor Newslett 19(1):22–36

Shu, K, Bernard HR, Liu H (2018a) Studying fake news via network analysis: detection and mitigation. arXiv:1804.10233

Shu, K, Wang S, Liu H (2018b) Understanding user profiles on social media for fake news detection. In: IEEE conference on multimedia information processing and retrieval (MIPR), pp 430–435

Shu K, Mahudeswaran D, Wang S, Lee D, Liu H (2018c) FakeNewsNet: A Data Repository with News Content, Social Context and Dynamic Information for Studying Fake News on Social Media. arXiv:1809.01286

Wang WY (2017) "liar, liar pants on fire": a new benchmark dataset for fake news detection. arXiv:1705.00648

Kai Shu received his BS degree in computer science from Chongqing University (CQU), Chongqing, China, in 2012 and MS degree in computer science from Chongqing University in 2015. He is currently working toward the Ph.D. degree in CSE at Arizona State University (ASU), Tempe, AZ. He also works as a research assistant at the Data Mining and Machine Learning Lab (DMML) at ASU under the supervision of Dr. Huan Liu. His research interests are in data mining, machine learning and fake news detection. He has published innovative works in top conference proceedings such as WWW, AAAI, IJCAI, CIKM, and WSDM. He also worked as a research intern at Yahoo Research in 2018.

Deepak Mahudeswaran received his Bachelor of Engineering degree in Computer Science from Anna University, Tamil Nadu, India, in 2015. He is currently working toward the Master's degree in CS at Arizona State University (ASU), Tempe, AZ. He also works as Graduate Services Assistant at the Data Mining and Machine Learning Lab (DMML) at ASU under the supervision of Dr. Huan Liu. He worked as a Software Engineer for 2 years in MicroFocus, India from 2015 to 2017. He also worked as a Software Engineer Intern at Amazon India in 2015.

Huan Liu received the BE degree in computer science and electrical engineering from SJTU, and the PhD degree in computer science from the University of Southern California, Los Angeles, CA. He is currently a professor of CSE at ASU. He was recognized for excellence in teaching and research in Computer Science and Engineering at ASU. His research interests are in data mining, machine learning, social computing, and artificial intelligence, investigating problems that arise in many real-world, data intensive applications with high-dimensional data of disparate forms such as social media. His well-cited publications include books, book chapters, encyclopedia entries as well as conference and journal papers. He serves on journal editorial boards and numerous conference program committees and is a founding organizer of the International Conference Series on Social Computing, Behavioral-Cultural Modeling, and Prediction (http://sbp.asu.edu/).