

Diminishing Spread of False Message in Twitter using Block chain and Machine Learning



Alagu Vignesh A, Harini N.

Abstract: A communication platform such as social media often works to change or influence opinions when it comes to political views. Microblogging platforms such as Facebook, Pinterest, Tumblr, Instagram enable users to contribute posts which are publicly visible. Twitter stood out as over-powering news equipment, giving information to the world. It is very common for users to log on to this site for immediate update on recent happenings like natural disasters, political events etc. Due to the enlarging reputation of Twitter, this aroused spammers to circulate false information. The proposed scheme seeks to reduce the spread of false information in Twitter using blockchain consensus algorithm, along with classification and machine learning algorithms. A "tweet" gets posted based on the validation provided by the nodes in the blockchain. The system works according to the sensitivity of the tweet.

Key Words: Blockchain, Consensus, Communication, False tweet, Microblogging, Politics, Tweet, Twitter.

I. INTRODUCTION

Social media are web-based applications which help in the creation and sharing of any sort of information, ideas, documents, and media such as photos, videos, music, etc. It allows people to share content faster in real-time.

Microblogging has become a major source of news these days [18], [19]. Statistics show that 2.4 billion (approx. 64.5%) online end users obtain news from applications like Facebook, YouTube, Tumblr, Pinterest and Twitter [1], [2]. With the exponential growth of the dependent users on social media to carry out their day to day activities it is undoubtedly evident that social applications like Fb, Insta etc. stand as most preferred applications.

Microblogging service has the highest political participation, it enables news stories to quickly go viral which can lead to mis-interpretations that can cause conflicts later. Considering the political arena Twitter has become authentic and incessantly used microblogging service in the political ring [6], [9].

The increased popularity of the social media "Twitter" is the reason behind spammers spreading false information. Therefore, Twitter has to be shielded from false news spreaders or spammers.

Revised Manuscript Received on October 30, 2019.

* Correspondence Author

Ridha Muldina Negara*, School of Electrical Engineering, Telkom University, Bandung Indonesia.

Rohmat Tulloh, School of Applied Science, Telkom University, Bandung, Indonesia.

Nandy Hadiansyah P.N., School of Applied Science, Telkom University, Bandung, Indonesia.

Rizka Triani Zahra, School of Applied Science, Telkom University, Bandung, Indonesia.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Companies strive to filter false tweets making twitter safer to use. For example, Trend Micro filters spam URLs using blacklisting service known as WRP (Web Reputation Technology).

Due to time lag, it couldn't blacklist URLs before it reaches the user. In order to overcome this, rule-based algorithms were used by researchers to filter spam tweet [7]. With industries relying on social media for popularizing their business the presence of spam tweets pose great security threat. This has initiated the need for researching on ways to minimize and filter spam. Fake news are messages created to deliberately misinform readers and today the rate of its spread is about 365% more

than it was during 2017 [8].

Fake news is so alluring because they tend to be more interesting than the true ones, and they gain more attention. For example, Spouse of Supreme Court Justice Clarence Thomas tweeted about the migrant caravan which was found to be false [4]. Research reports that the spread of fake news is approximately six times quicker than the spread of original information [3]. The likeliness of spread of false news is about 70% in twitter [8]. The depth analysis performed on retweet revealed that there were less than 10 retweets with original news and 19 with fake news. Generally, bots or trolls are used for disseminating fake news [8]. Although fake news bots have gained attention, humans are equally spreading false news across social networks. A study at MIT shows that even after the removal of fake news spreading bots from their dataset, the false news spread rate remained unchanged [5].

Therefore, in this paper, the focus is to reduce the spread of false information through Twitter using blockchain Proof of Work (PoW) consensus algorithm integrated with classification and machine learning algorithms.

II. LITERATURE SURVEY

1. Block Chain

Blockchain is a technology that facilitates decentralized storage of data and secure distributed computation of process using strict timestamp, digital signature, encryption/decryption schemes and consensus algorithms. The storage and data processing capability along with its strengths of sustainability, data sharing and inter connectivity of this technology has attracted a wide variety of applications to rely on it [10].

2. Proof of Work

This procedure is adopted to validate the transaction before it gets appended to the blockchain. In this method a ledger is maintained to record each of the exchanges that happens.

A block is then created for this information and appended to the chain by the miner once he/she solves the hash for that block[11]. Bitcoin based cryptography procedures are difficult to be compromised for reasons that it has built in capability to prevent double-spend attack due to its requirement in terms of huge computing power to solve the hash value[12].

3. Text Classification

Text classification is a process of categorizing text data into groups by using Natural Language Processing (NLP) [14], [15].

The proposed system uses Paralleldots API which provides machine learning and classification solutions. Figure 1 represents the procedure for classifying the text

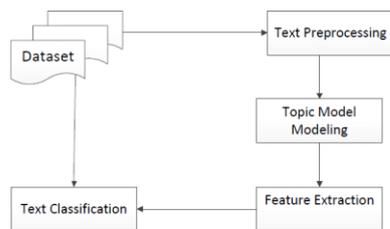


Fig. 1. Classification Process

4. Findings for Literature survey

Most of the works existing in the literature use machine learning and Deep learning methods for text classification and prevention of spread of fake news [13], [16], [17]. The focus of the existing methods was more on improving the accuracy of classification rather than providing a distributed methodology for approving the content of the message. To the best of our knowledge we have not found any existing work that is based on block chain that would diminish the spread of false messages in social media. The integration of blockchain and machine learning based approval scheme is the motivation for the proposed system.

III. PROPOSED SYSTEM

1. Overall Working

When a user posts a tweet, on reaching the twitter server, the tweet is put through a machine learning module component which segregates the information into political and non-political messages. Tweets segregated as non-political are passed without any further processing and gets posted in Twitter. If a message is tagged as political-related information it is classified based on the sensitivity and it is tagged with a threshold. In the case of a time-insensitive tweet, the threshold is decided by taking the number of active nodes N in the blockchain and is divided by 2 i.e., if 50% of the nodes agree the tweet is true then it is considered valid. The threshold is set to 25% if it is a time-sensitive tweet. For example, consider a tweet related to the decease of a person. The tweet then gets transferred to the block-chain management node with a threshold tagged from the previous step. The management node broadcasts the tweet to all the nodes which are in consensus. The nodes then validate and mark the tweet as either valid or invalid and then puts the tweet in the blockchain with their digital signature. The nodes mentioned here are either a person or an AI application that verifies the viability of the

information in real-time. After a time-interval, the chain is processed by the management node and if the validity count meets the threshold the tweet gets posted. If the validity count is less than the threshold, the tweet gets dropped without getting posted and the user's activity to spread false information is reported to twitter. The overall functioning of the proposed model is shown as an architecture diagram in Figure 2.

2. Obtaining and classifying tweets

Tweets collected for the experimentation process were collected using Python scripts using the twitter library. The tweets are segregated as political and non-political using machine learning algorithms, which were achieved through Paralleldots custom classification API. The data set is prepared as a text document and the same is depicted in Figure 3. The contents of the text document will further undergo classification based on the sensitivity of the tweet and then gets transferred to the block-chain for approval. In this application, the sensitivity of a tweet is calculated by using word matching strategy, which searches sensitive words like decease, funeral, death, etc. from a word list and returns true if any of the words in the list are found in the tweet.

3. Construction of Blockchain

The node has been constructed using Node JS which uses express, crypto-js, body-parser, ws (WebSocket) and file system libraries for implementing the blockchain. The node initializes HTTP service and P2P service in two different port numbers. The HTTP service is used to communicate between the user and the node, while P2P service is used to communicate between nodes. Initially, when the server starts, an initial block will be placed into the blockchain with index, previous hash, timestamp, data, hash and digital signature as parameters. The application sends all the tweets to a common management node which is chosen at random. The nodes in the blockchain network use web socket to communicate with each other. When a tweet is received by the management node, it first validates the tweet and then constructs a block with the validated data and the node's digital signature. This block is then added to the blockchain. Figure 4 shows the tweet added to the blockchain after getting validated by node

A. The block is then broadcast to all the connected nodes. Each node can retrieve the block and can re-validate the message. Figure 5 shows the tweets broadcast to node B from node A. After a time interval, the management node processes the chain and sends back the tweet with its validity to the twitter server.

STEPS IN THE WORKFLOW

- Step 1: Retrieve the tweets using twitter python library.
- Step 2: Send the tweets to Parallel dots Custom classifier to classify the tweets as Political and Non-Political.
- Step 3: Post the tweet directly if it is classified as Non-Political using tweepy python library.

- Step 4: Send the tweet to a sensitivity analysis algorithm and tag the tweets with a threshold based on the sensitivity.
- Step 5: Send the tweets to the blockchain nodes.
- Step 6: After time-out, blockchain nodes return the validity of the tweet as shown in Figure 7.
- Step 7: Post the tweet if it is said to be valid by the nodes using tweepy python library, this is shown in Figure 8.
- Step 8: Drop the tweet if it is said to be false by the nodes, report the user to Twitter.

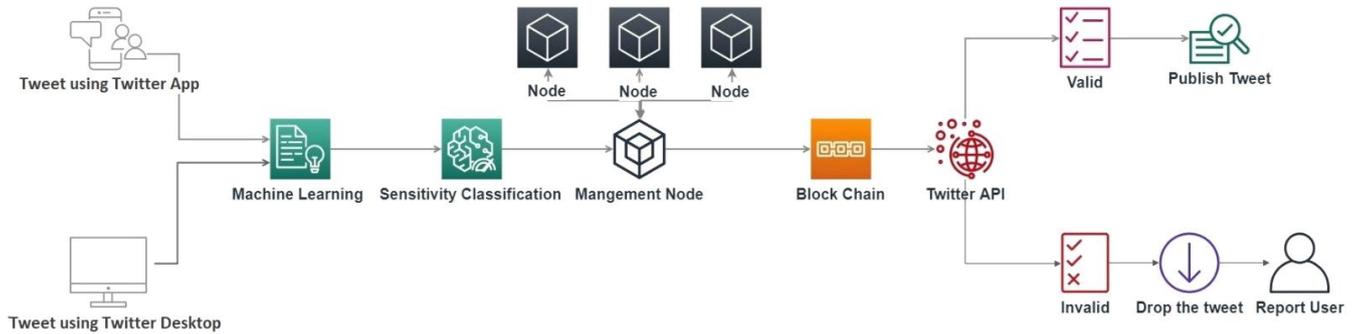


Fig. 2. Architecture diagram of the proposed system

```

Prime Minister Narendra Modi tweeted a link to the speech Human Resource Development Minister Smriti Irani made in the Lok Sabha during the bate on the ongoing JNU row and the suicide of Dalit scholar Rohith Vemula at the Hyderabad Central University.

Diwakar Raote, senior Shiv Sena leader and minister in Maharashtra government had recently said the alliance will break if the Sena doesn't get to contest 50 per cent of the seats.

Nitish Kumar lashed out at his detractors, claiming that people 'lacking political acumen' were trying to gain publicity by launching personal attacks on him.

Ms Jayalalithaa, 68, died last night after three months in hospital

Priyanka Gandhi accused the BJP government of delay, and also added that it was the people's ire that ultimately forced BJP to take action against its own leader.

J Jayalalithaa's funeral took place with full state honours in Tamil Nadu.

Prime Minister Narendra Modi, who flew to Chennai today to pay his last respects to Ms Jayalalithaa, expressed deep sadness over her death in a series of tweets.

BJD MP Slaps Party Worker at Public Event, Calls it 'Friendly Gesture' After Backlash.
    
```

Fig. 3. Sample Data set

```

root@focsociety:~/Block_Chain# python get_blocks.py
Message Validity
Prime Minister Narendra Modi tweeted a link to the speech Human Resource Development Minister Smriti Irani made in the Lok Sabha during the bate on the ongoing JNU row and the suicide of Dalit scholar Rohith Vemula at the Hyderabad Central University.
:false info
Diwakar Raote, senior Shiv Sena leader and minister in Maharashtra government had recently said the alliance will break if the Sena doesn't get to contest 50 per cent of the seats.
:false info
Nitish Kumar lashed out at his detractors, claiming that people 'lacking political acumen' were trying to gain publicity by launching personal attacks on him.
:false info
Ms Jayalalithaa, 68, died last night after three months in hospital
:false info
Priyanka Gandhi accused the BJP government of delay, and also added that it was the people's ire that ultimately forced BJP to take action against its own leader.
:false info
J Jayalalithaa's funeral took place with full state honours in Tamil Nadu.
:valid
Prime Minister Narendra Modi, who flew to Chennai today to pay his last respects to Ms Jayalalithaa, expressed deep sadness over her death in a series of tweets.:valid
Digital Signatures
4162ac6d26ab264c16fc2cf488965a84d6d797e4d5bc97f5940ed43e9969a61b715061309d51fa6b569bf2ba988fc5c4547c1b4ce44f06aad886fd01910fbbab47f06c7d4ec1f498501d614b64a16a10ebc323e6ee446e41cf85b890cb07ce9c375cb56a4992f64c7b25fe3961ac73bc5fc911ae95d8110a2743f6be8a1c162aa55225c7cd1098304a62b45870db04c8ef7cc09c0e9f4fd2ef709ceb91b881e0cf66f28661bb0934cebbdf9f903a764319e1735be7d97618039aaebc729954862901bd21d51ee33cabd7f0b49e2bc0244bb6948ce2cc04349c98d808c38c2534ae8552da7a54bc8a9676c1f839872e6bc3a8b806076cbcedb8689f3be852d,true
7ecc082e38503d9445d20f8aae458d9a46a831ed48219ddddd33142505c46cc63c87c86b5e9879b36b26e59a7344710d8a2eb6907cfc8100127e5370f2fd488ba43f7ad3aa55806344e515d89823e2e8021812d3cfff179b45fe18be8ed81ef070217a95a89bd99b049042c9449940bdd346e4b94a2f7bb985be35ba1b5970e27164b55e8261593159edf120415d15d9e667b86085102a326a72b53c1884dc30a9c8ae17025e9695848a5844568085e1dc4a2db9c50f0e841fab21f49f8c45275259741d330a131f6adc3789ef48f7dbfa70cc6e9a73c5fa45605dae5182d10dbb33608a8d66a16ac194abcd135640cd5e0aff49d2af8cdcaed47e65eaff80,true
    
```

Fig. 4. Message broadcast by Node A after validation

```

Kavis-MacBook-Air:Block_Chain kavidharshini$ python get_blocks.py
Message Validity
Prime Minister Narendra Modi tweeted a link to the speech Human Resource Development Minister Smriti Irani made in the Lok Sabha during the bate on the ongoing JNU row and the suicide of Dalit scholar Rohith Vemula at the Hyderabad Central University.
:false info
Diwakar Raote, senior Shiv Sena leader and minister in Maharashtra government had recently said the alliance will break if the Sena doesn't get to contest 50 per cent of the seats.
:false info
Nitish Kumar lashed out at his detractors, claiming that people 'lacking political acumen' were trying to gain publicity by launching personal attacks on him.
:false info
Ms Jayalalithaa, 68, died last night after three months in hospital
:false info
Priyanka Gandhi accused the BJP government of delay, and also added that it was the people's ire that ultimately forced BJP to take action against its own leader.
:false info
J Jayalalithaa's funeral took place with full state honours in Tamil Nadu.
:valid
Prime Minister Narendra Modi, who flew to Chennai today to pay his last respects to Ms Jayalalithaa, expressed deep sadness over her death in a series of tweets.:valid
Digital Signatures
4162ac6d26ab264c16fc2cf488965a84d6d797e4d5bc97f5940ed43e9969a61b715061309d51fa6b569bf2ba988fc5c4547c1b4ce44f06aad886fd01910fbbab47f06c7d4ec1f498501d614b64a16a10ebc323e6ee446e41cf85b890cb07ce9c375cb56a4992f64c7b25fe3961ac73bc5fc911ae95d8110a2743f6be8a1c162aa55225c7cd1098304a62b45870db04c8ef7cc09c0e9f4fd2ef709ceb91b881e0cf66f28661bb0934cebbdf9f903a764319e1735be7d97618039aaebc729954862901bd21d51ee33cabd7f0b49e2bc0244bb6948ce2cc04349c98d808c38c2534ae8552da7a54bc8a9676c1f839872e6bc3a8b806076cbcedb8689f3be852d,true
7ecc082e38503d9445d20f8aae458d9a46a831ed48219ddddd33142505c46cc63c87c86b5e9879b36b26e59a7344710d8a2eb6907cfc8100127e5370f2fd488ba43f7ad3aa55806344e515d89823e2e8021812d3cfff179b45fe18be8ed81ef070217a95a89bd99b049042c9449940bdd346e4b94a2f7bb985be35ba1b5970e27164b55e8261593159edf120415d15d9e667b86085102a326a72b53c1884dc30a9c8ae17025e9695848a5844568085e1dc4a2db9c50f0e841fab21f49f8c45275259741d330a131f6adc3789ef48f7dbfa70cc6e9a73c5fa45605dae5182d10dbb33608a8d66a16ac194abcd135640cd5e0aff49d2af8cdcaed47e65eaff80,true
    
```

Fig. 5. Messages retrieved by Node B

Diminishing Spread of False Message in Twitter using Block chain and Machine Learning

```
Kavis-MacBook-Air:Block_Chain kavidharshini$ python fire_module.py
Prime Minister Narendra Modi tweeted a link to the speech Human Resource Development Minister Smriti Irani made in the Lok Sabha during the bat
e on the ongoing JNU row and the suicide of Dalit scholar Rohith Vemula at the Hyderabad Central University.

False Message, Reporting User
Diwakar Raote, senior Shiv Sena leader and minister in Maharashtra government had recently said the alliance will break if the Sena doesn't get
to contest 50 per cent of the seats.

Message is Valid posting tweet
Nitish Kumar lashed out at his detractors, claiming that people 'lacking political acumen' were trying to gain publicity by launching personal
attacks on him.

Message is Valid posting tweet
Ms Jayalalithaa, 68, died last night after three months in hospital

Message is Valid posting tweet
Priyanka Gandhi accused the BJP government of delay, and also added that it was the people's ire that ultimately forced BJP to take action agai
nst its own leader.

Message is Valid posting tweet
J Jayalalithaa's funeral took place with full state honours in Tamil Nadu.

Message is Valid posting tweet
Prime Minister Narendra Modi, who flew to Chennai today to pay his last respects to Ms Jayalalithaa, expressed deep sadness over her death in a
series of tweets.

Message is Valid posting tweet
Kavis-MacBook-Air:Block_Chain kavidharshini$
```

Fig. 7. System validation result

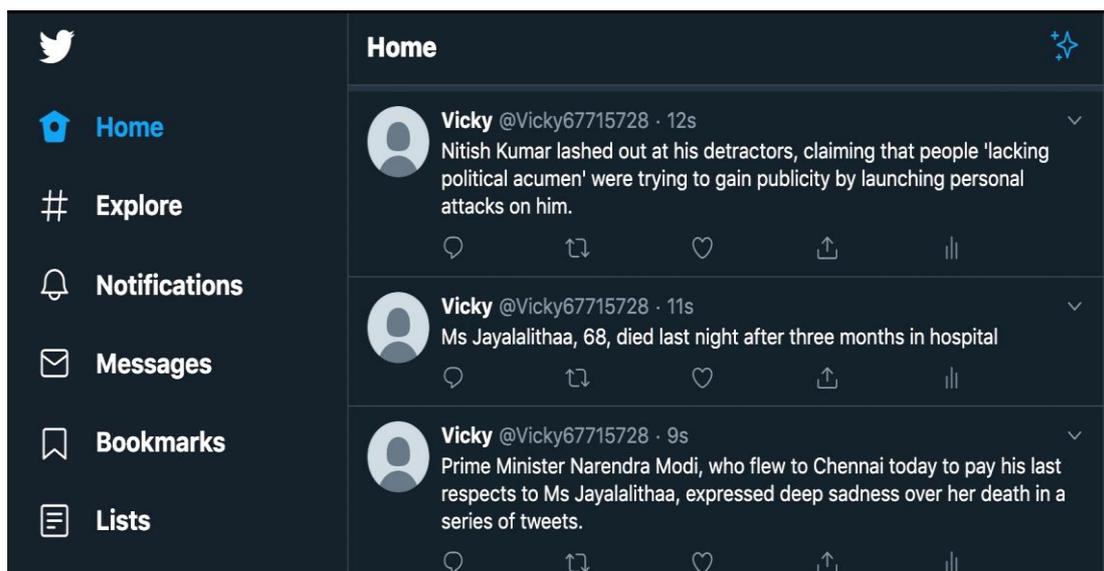


Fig. 8. Tweet posted after validation by the server

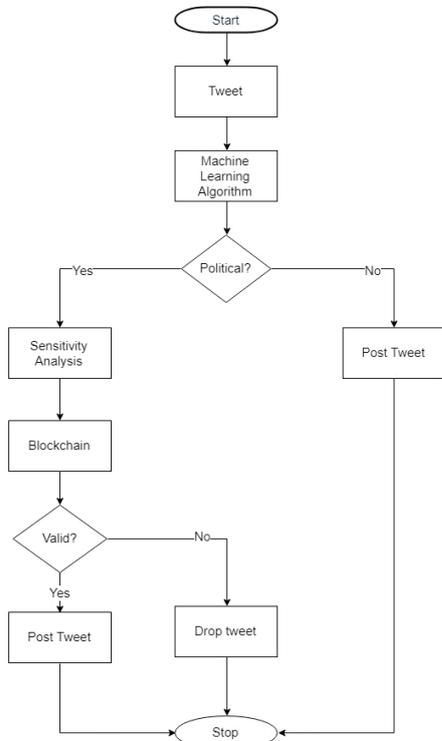


Fig. 6. Working of the proposed system

IV. CONCLUSION

In this paper, we have presented how the spread of false messages through Twitter can be reduced using Proof of Work (PoW) blockchain consensus algorithm along with classification and machine learning algorithms. The thorough experimentation process using the collected dataset from the twitter revealed the increased ability of the system to prevent spread of false messages in comparison to the methods discussed in the literature. This improved result is due to the integration of Blockchain based consensus protocol for tweet approvals.

REFERENCES

- [Online] Available: <https://www.forbes.com/sites/nicolemartin/2018/11/30/how-social-media-has-changed-how-we-consume-news/#61d6a7353c3c>
- [Online] Available: <https://www.journalism.org/2018/09/10/news-use-across-social-media-platforms-2018/>
- [Online] Available: <https://www.marketwatch.com/story/HYPERLINK> "http://www.marketwatch.com/story/fake-news-spread-fakeHYPERLINK" "http://www.marketwatch.com/story/fake-news-spread-fakeHYPERLINK" "http://www.marketwatch.com/story/fake-news-spread-fakeHYPERLINK"
- [Online] Available: <https://www.marketwatch.com/story/wife-of-supreme-court-justice-spreads-fake-news-about-the-migrant-caravan-2018-10-24>
- [Online] Available: <http://news.mit.edu/2018/study-twitter-false-news-travels-faster-true-stories-0308>
- Iyapparaja M, Tiwari. M, Security policy speculation of user uploaded images on content sharing sites , IOP Conf. Series: Materials Science and Engineering 263 (2017) 042018 doi:10.1088/1757-899X/263/4/042019,pp-1-8
- Rutuja Katpatal ; Aparna Junnarkar , "An Efficient Approach of Spam Detection in Twitter" , 2018 International Conference on Inventive Research in Computing Applications (ICIRCA).
- Meenakshisundaram, Iyapparaja, and Sureshkumar Sreedharan. "Intelligent Risk Analysis Model for Mining Adaptable Reusable Component." International Arab Journal of Information Technology (IAJIT) 12 (2015).

- Vladimir P. Miletskiy ; Dmitry N. Cherezov ; Elena V. Strogetskaya, "Transformations of Professional Political Communications in the Digital Society (by the Example of the Fake News Communication Strategy)" , 2019 Communication Strategies in Digital Society Workshop (ComSDS).
- Hiroki Takikawa ; Kikuko Nagayoshi , "Political polarization in social media: Analysis of the 'Twitter political field' in Japan" , 2017 IEEE International Conference on Big Data (Big Data).
- Shitang Yu ; Kun Lv ; Zhou Shao ; Yingcheng Guo ; Jun Zou ; Bo Zhang, "A High Performance Blockchain Platform for Intelligent Devices" , 2018 1st IEEE International Conference on Hot Information-Centric Networking (HotICN).
- Shihab Shahriar Hazari ; Qusay H. Mahmoud , "A Parallel Proof of Work to Improve Transaction Speed and Scalability in Blockchain Systems" 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC).
- Iyapparaja, M., and S. Sureshkumar. "Coupling and cohesion metrics in Java for adaptive reusability risk reduction." (2012): 52-57.
- Iyapparaja M, Bhanupriya Sharma, Augmenting SCA project management and automation Framework, IOP Conf. Series: Materials Science and Engineering 263 (2017) 042018 doi:10.1088/1757-899X/263/4/042018,pp-1-8
- Shital Anil Phand ; Jeevan Anil Phand , "Twitter sentiment classification using stanford NLP" 2017 1st International Conference on Intelligent Systems and Information Management (ICISIM).
- Thayakorn Dangkesee ; Suthera Puntheeranurak , "Adaptive Classification for Spam Detection on Twitter with Specific Data" , 2017 21st International Computer Science and Engineering Conference (ICSEC).
- Sahar A. El Rahman ; Feddah Alhumaidi AlOtaibi ; Wejdan Abdullah AlShehri , "Sentiment Analysis of Twitter Data" , 2019 International Conference on Computer and Information Sciences (ICCIS).
- Umid Suleymanov ; Samir Rustamov ; Murad Zulfugarov ; Orkhan Orujov ; Nadir Musayev ; Azar Alizade , "Empirical Study of Online News Classification Using Machine Learning Approaches", 2018 IEEE 12th International Conference on Application of Information and Communication Technologies (AICT).
- Surya Prabha K; Harini N, "Chat-guard A system that ensures safe posting in social networking sites." , 2016 International Journal of Engineering and Technology.
- Smruthi M, Harini N, "A hybrid scheme for detecting fake accounts in facebook." , 2019 International Journal of Recent Technology and Engineering.

AUTHORS PROFILE



Alagu Vignesh A. is a B.Tech student at Department of Computer Science and Engineering, School of Engineering, Coimbatore India. He is actively involved in research in the areas of Cyber Security and Networks. He is a Certified Ethical Hacker (CEH) and CISO Certified Network Associate(CCNA R&S).



Dr. Harini N. currently serves as Assistant Professor at Department of Computer Science and Engineering, School of Engineering, Coimbatore Campus. Her areas of research include Security. She has published a conference paper Proceedings of International conference on Adaptive Technologies for Sustainable Growth (ICATS-2014) She has written a book named Cryptography and Security, Wiley India , 2011.