IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 15, 2021, accepted November 1, 2021, date of publication November 2, 2021,
date of current version November 22, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3125128

Social Media and Steganography: Use, Risks and Current Status

R. GURUNATH[1], MOHAMMAD FADEL JAMIL KLAIB[2], DEBABRATA SAMANTA[1], (Member, IEEE), AND MOHAMMAD ZUBAIR KHAN[2]

[1]Department of Computer Science, CHRIST (Deemed to be University), Bengaluru, Karnataka 560029, India
[2]Department of Computer Science, Taibah University, Madinah Al Munawara 42353, Saudi Arabia

Corresponding authors: Mohammad Fadel Jamil Klaib (mklaib@taibahu.edu.sa), Debabrata Samanta (debabrata.samanta369@gmail.com), and Mohammad Zubair Khan (mkhanb@taibahu.edu.sa)

ABSTRACT Steganography or data hiding is used to protect the privacy of information in the transit; it has been observed that the information that flows through Online Social Networks (OSN) is very much unsafe. Therefore, people hesitate to communicate their sensitive data on social media.. Most of the information on the online social network is not useful to users and appears to disregard such details. People's actions provided a possibility for digital Steganography through the Internet.. TCPIP covert channels were used for steganography until the last decade. People began to utilize social media as a covert conduit to communicate hidden messages to targeted users as social media grew in popularity. There are numerous Online Social Networks accessible nowadays, ranging from Facebook to the more contemporary Twitter and Instagram. All of them may be utilized as covert channels without the general public noticing. The primary characteristic of steganography is the protection of information privacy; nonetheless, it has been utilized more for illicit message transmission, which is a source of concern. To make matters worse, adversaries are using steganalysis techniques to mess with the concealed data. In this article, we examine the different social media steganography techniques, such as those used on Facebook, WhatsApp, and Twitter, as well as the difficulties that these approaches raise. The positive and negative consequences of social media, as well as its current state, are discussed in this study. This paper discusses how the performances of Steganography methods may be assessed using the Entropy value of the Stego object. A look of the three features of steganography. It has been given with undetectability, robustness, and payload capacity. Finally, the paper's concept's future scope is explored.

INDEX TERMS Steganography, social media, cyber stalking, cyber bullying, steganalysis, OSN, Facebook, Twitter.

I. INTRODUCTION

Steganography is a very old method that dates back about 2000 years, and digital steganography has just been around over the last two decades. People utilized covert channels to hide text, image, audio, video, and network protocols like TCP/IP in the early phases of the digital era. In the recent decade, people began to use clandestine social media platforms. As the Internet grew in popularity, so did social media networks, which began to generate massive amounts of data on a regular basis. The majority of the data created is unimportant and unattended, and individuals prefer to ignore large chunks of social media data. People's attitudes influenced the development of social media steganography.

Social media has brought transformative changes in the individuals and it removed the barriers that were existed. The main idea of being the user of online social network is to stretch out to oneself to multiple communities of certain interest. Engaging in sharing the work with communities in social media, that facilitates connection with similar interest people, information, and discussions and so on [1]. Social media is described as forms of electronic Communication and it was initiated by WWW through a website, SIX DEGREES in 1997. In 2000, MySpace and LinkedIn two social media sites were released. Then in 2007, very famous Facebook and later hundreds of social media sites were released. Social

The associate editor coordinating the review of this manuscript and approving it for publication was Fabrizio Marozzo.

media and networks made our lives lucrative; it is a huge evolution that helped globalization of education, business, culture, social life, etc. Though there are plenty of positive aspects, the negativity factors namely, cyber bullying, illegal content, harassment, stalking, Child abuse, health impact, terror utilization, are equally dominant [2].

Social networking (SNSs) has become an important part of the medium for communicating inside and around interpersonal relationships. The strongest SNS is Facebook in the United States and beyond. One theory why? Facebook is the most popular social networking site. It is the selection of services that consumers are offered. Online connections allow users to quickly interact with members of the network. Facebook may also have a negative effect on relationships that are romantic. Studies have shown that Facebook can foster romantic jealousy [3], [4]. Given that each participant has a major negative relationship and psychological experience linked to Facebook, social media management and its role in our relationships should be an important part of education in media literacy.

It's nothing new to Steganography. It has got the history of usage in the malware as the first instance. This concept had used even before 440 BCE. Namely, the ancient royal people of Greece used to shave the slaves head, on their head tattooing some secret message, and then before sending them to the other destination; used to wait for hair to grow so that such an act could not be recognized by any enemies. This way, it has been transformed up to the age of digital Steganography. Steganography is a Greek word meaning ''covered writing'' [5]. The sender of the message encodes the secret on an innocent text, and whoever receives decodes the same to get back the secret.

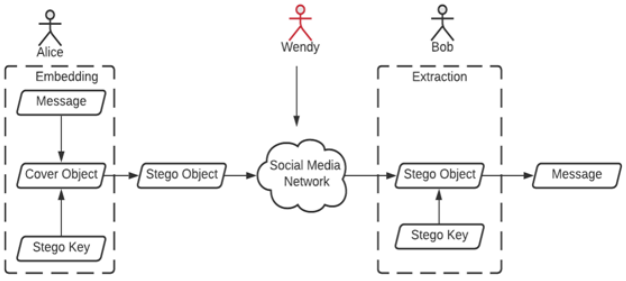

FIGURE 1. The process of data hiding in social media networks.

The process of data concealment and extraction is depicted in the Figure 1 below. Two users, Alice and Bob, function as transmitter and receiver, respectively. Alice and Bob need to communicate privately and invisibly via the Internet (Online Social Networks), and Steganography can enable them do so. Alice transmits the secret information to Bob by embedding it in a normal-looking cover file using a specific technique. Bob receives the Stego object and extracts the secret using the same technique. Between the source and the target, there would always be attackers known as 'Steganalysts' (in our case, Wendy), who would study the messages and then attack them. Recently, people of social media are using the image and text Steganography more to send their messages. Steganography alone cannot provide complete anonymity; sometimes, it reveals the existence of the secret. Even attackers can come to know and tries to decode it; the certain time they may succeed in their mission as well. It is very important for a sender to protect their message and the second level of security in the form of encryption might be employed. Google+ and Facebook today are recognized social media networks that are being used for secret sharing. Flickr is another similar kind of Social media for photo-sharing website [6].

The first segment of the paper covers, literature survey on Social Media Steganography, for that Facebook, Twitter, and WhatsApp are chosen. The second segment deals with the use of entropy value of the Stego object to assess the performance of a Steganographic method, followed by Bright side of Social media, and Dark side of Social media. The last segment provides the idea of current trends and status of Social media and future scope.

II. MOTIVATION OF RESEARCH

Steganography is a method of concealing a hidden message within regular files. Due to their innocence, these files serve as hidden communication carriers. The major design difficulty of Steganography is undetectability, which implies that when the secret is hidden, no one will suspect it. We utilized social media data as the bearer of a secret message in this article since massive amounts of social media data are generated on a regular basis. The hidden sharing is not visible to users of social media. Aside from social media data, there will always be the element of suspicion. As a result, it is motivated to use social media as a hidden message carrier.

III. OBJECTIVE OF WORK

The objective of this article is to examine the benefits and drawbacks of social media and steganography, as evidenced by the following arguments. Nowadays, there are a plethora of Online Social Networks to choose from, ranging from Facebook to the more modern Twitter and Instagram. All of them might be used as clandestine channels without the public's knowledge. Steganography main feature is that it protects information privacy; nonetheless, it has been used increasingly for unlawful message transmission, which is a cause of worry. To make matters worse, attackers are tampering with the hidden data using steganalysis techniques. It's concerning that cyber thieves are increasingly employing steganography to steal genuine data from businesses.

IV. CONTRIBUTION OF THE RESEARCH

The following are some of the article's key features:

- A steganography process diagram employing Facebook networks is proposed.
- Proposed steganography process diagram employing Twitter networks.

- Implementations of word shift/extended line Steganography are presented as examples (Figure 3).
- Suggested the Shannon entropy for determining steganography. imperceptibility characteristic.
- Benefits and drawbacks of using hidden channels on social media.
- Current social media trends and status.

V. LITERATURE SURVEY ON SOCIAL MEDIA STEGANOGRAPHY

A piece of social media information can be used as a cover text to carry secret messages to destinations. The secret messages are made invisible for the normal view, only the destination users able to retrieve it again. The researchers use the concept extended line to embed to the secret information. To send the secret the authors experimented on the FACEBOOK data as the cover object [7]. The process Figure 2 below shows how to hide and extract messages utilizing social media networks. Word shift/extended line steganography is used in this approach. M-bits and cover text are fed through the aforementioned data concealing algorithms to generate Stego text, which is then sent via social media networks (Facebook). Basically, the Stego Text is sent to anybody who is a member of the Facebook network. The concealed information can only be retrieved /extracted by the designated receiver [8].

Cover text (a) and Stego text (b), respectively, are shown in the accompanying (Figure 3) texts. We need to insert the phrase ''NTUEE'' into the Cover text (a), with the effect given in (b). The chosen word is converted to a stream of binary bits as follows:

100 111 001 010 100 010 101 010 100 010 101 000 101

$\rightarrow$ *NTUEE*

The bits are divided into three groups, and the embedding is done as follows: The line length of the cover text is set to a predetermined length (in this case it is 50). If the first bit is 1, an extra white space character will be added between any two words in the line at random.

If the second bit is 1, it will read one more word from the cover text, allowing the line length to exceed L. If the third bit is 1, one white space will be added to the end of the line. If all 000 is true, then there will be no change.

The claim of the proposed algorithm is to hide more bits using, Facebook information. The same algorithm can be extended to other social media applications as well. The drawback of this method is the use of any word processing tools, that might remove the white spaces embedded and in turn, defeat the purpose. WhatsApp is third biggest social media app used to send and receive any kind of data instantaneously. About 90% of the WhatsApp messages are normally useless. This kind of WhatsApp information can be used as innocent carrier to send the secret information. This paper gives an idea of embedding the secret data on a WhatsApp image file [9], [10]. First the message is hidden under an image using the Steganography software to make the image

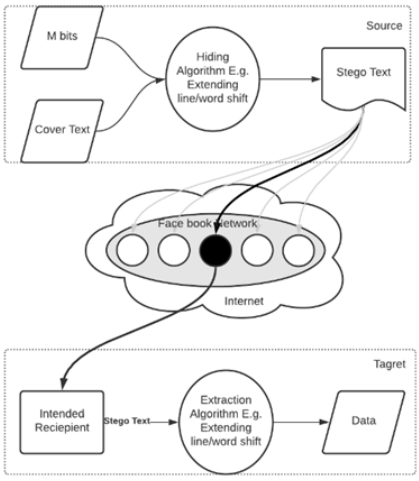

FIGURE 2. Visualization of Steganography using Facebook networks.

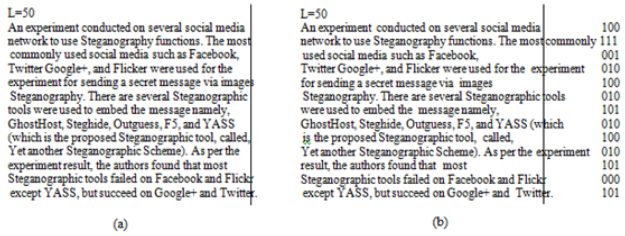

FIGURE 3. A sample message shown as cover text (a) and (b) a Stego text containing hidden bits using extended Steganography.

as stegano container. The processed image is sent via the WhatsApp to the destination address. Recipient retrieves the hidden message from the image using initially agreed upon Steganography tools and authentication. There are several of the Steganography tools have been used to check the performances such as SteganPEG, OpenStego, Quickstego, JP hide, and etc. The Method is devised for android mobile phones. As per the authors, tools such as Stegais and Steganos showed promising results. Therefore, many Steganographic tools are available; some are very specific to Windows OS and some for Android OS. This paper discusses the application of Steganography on Social media data especially the Twitter. The algorithm uses a linguistic Steganography which reduces Steganalysis attacks. The method to embed data into Twitter data is called 'Cover Tweet (Figure 4)'. A particular tweet is selected, to embed the payload (message).

The tweet is subjected to paraphrase database (PPDB). PPDB tokenize the tweet input, and generates feasible paraphrases, and it is subjected again to hash function to generate a unique 4-bit hash code. Using Bayes rule probability, rank of all the paraphrases generated. The human operator then selects a specific rank (hash code) and i.e. going to be our message. The paraphrase selected is sent to the recipient through twitter. Recipient uses same procedure to retrieve the message. This method resists the Steganography attacks and avoids detection of the message [11], [12]. PPDB-The

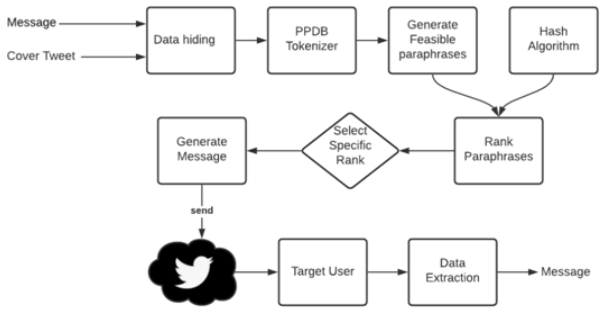

FIGURE 4. The process steps of Steganography using twitter network.

PPDB data set is the world's biggest accessible collection of paraphrases, with over 22 crore paraphrase pairings, 7.3 crore phrasal paraphrases, and 8 million lexical paraphrases for syntactical transformations [14]. the inventors write social media networks are frequently targeted in two ways: actively and passively. In all circumstances, a single piece of information about the legal user may be used to disclose their identity [13]. Active social network attack aimed at creating user profiles for a small number of people whose personal information it needed to obtain. Phishing allows attackers to get users' personal information and begin targeting them. Passive attacks, on the other hand, do not create any false accounts; instead, they try to find themselves on the network and gather other genuine facts about individuals. The researcher gives the idea of data hiding mechanism on Facebook messages. The steganography is done on the user's profile pictures of Facebook, which serves as a covert channel for secret transmission. The channel utilized for this purpose is to convey both control and data information. For further protection, data can be encrypted with keys and delivered through email to the intended users for secret extraction [10]. Facebook, on the other hand, utilizes a re-sampling method on photos, which may cause the secret to be corrupted. The data concealing technique can endure the harm if it is sturdy enough. Steganography used in criminal activities for the first time during 2010, a group of Russians encoded the messages in online images. This was used for hiding a malicious data. During 2012, Al-Qaeda used Steganography for transporting secret messages, when a German security official held a Pakistani Al-Qaeda person with video of pornography embedded plans of terror attacks in Europe [11]. Social bots are the programs written to control online social network accounts instead of human user. These social bots automated and can work similar to human beings, and it is hard to notice. Social bots can be good and malicious one. Attackers usually adopt these malicious social bots to collect the account details from the social network users illegally. These Botnets can penetrate into legitimate user's domain and steal the authentication details [14]. Adversaries employ such social bots huge in number and perform the functions like sending messages, connection request, etc. The social bots have devised using the artificial Intelligence algorithms. Therefore, these threats are no longer from a human being rather from a software program. Like social bot, StegBot is another kind of bot, used for transferring illicit secrets concealed into a normal text, image, audio, or video. The embedded secret can infect computer, communicate stolen data, passwords, Account information, etc.

This paper [15] throws light on which particular online social networks are Stego-friendly or not; among selected OSN's. An experiment conducted on several social media network to use Steganography functions. The most commonly used social media such as Facebook, Twitter Google+, and Flicker were used for the experiment for sending a secret message via images Steganography [16]. There are several Steganographic tools were used to embed the message namely, GhostHost, Steghide, Outguess, F5, and YASS (which is the proposed Steganographic tool, called, Yet another Steganographic Scheme). As per the experiment result, the authors found that most Steganographic tools failed on Facebook and Flickr except YASS, but succeed on Google+ and Twitter. Therefore, Google+ is best OSN for hiding message and the next is Twitter. GhostHost fails on Twitter, but other platforms are capable of conceiving secret information efficiently, however, Facebook and Flickr are tough enough. Google+ and Twitter protect the dignity of the images to a great degree. Popular methods for Steganography can be used directly on photographs posted to these pages. Facebook and Flickr process submitted. Specifically, the properties of the data fields (metadata), the image elements (pixels) or the quantitative form of the 3-D signal elements (DCT coefficients) of the image are susceptible to manipulation by these sites [17]. Failure is due to the use of the LSB solution to protect an image message that is more likely to shift during processing. The authors then thought that instead of using the standard LSB type, a mixture of LSB+2-LSB (second least significant bit) would be safer with caution. This reduces the chance of Steganalysis. And this approach can be used efficiently on Facebook and Flickr.

VI. USAGE OF ENTROPY IN STEGANOGRAPHY

Text hiding scheme or Steganography [22] in which the cover message and the Stego text are statistically imperceptible, meaning; the cover message and the Stego text have the same distribution of probability. The sender is involved in sharing information through public channels with the recipient with an expectation of the presence of attackers. The sender transmits secrets to the recipient via cover message and the attacker always does not notice it. Let's say that the intruder is involved in finding some secret message in the text being sent instead of interrupting the transmission. The question arises; what strategy is the attacker going to use? The usual solution to that is the process of Shannon entropy.

Shannon originally developed entropy as result of his communication theory, for which a data communication device is made up of three components: a data source, a transmission medium, and a recipient. The ''fundamental problem of communication'' in Shannon's theory, as articulated by

Shannon, is for the receiver to be able to recognize what data was produced by the source, based on the signal it collects via the stream. In general, data entropy is the average amount of information transmitted by an event when all possible outcomes are considered [18]. Entropy is the direct measure of the randomness of information, which is regarded as a measure of the distribution of symbols in the information. A simple lack of continuity or predictability of events is the randomness of any information; it has no structure and does not follow an intelligible pattern. Such events are unpredictable. But it can be predictable by the distribution of probabilities.

The entropy value is a number, showing the low or high entropy status of the information. Predictability is lower for greater entropy and higher predictability for lower entropy [19]. By the following formula, entropy can be determined.

$$E(D) = \sum_{i=0}^{n} -P_i \log_2 (P_i) \tag{1}$$

where E, denotes entropy or randomness of an information and D, denotes data, $\sum$ denotes the sum of the possible symbol values of the variable and log is the logarithm. A unit of bits is given by Base 2, the potential results, i=1. . .n occur with the likelihood of $p_1, p_2, \ldots, p_n$. The following graph shows entropy vs. probability for two class variables. A symmetric curve is used to graphically represent the equation (1), as illustrated below. The probability of two-class variables is on the x-axis, while the randomness of data, indicated by E (D), is on the y-axis. When there are two potential outcomes, $log_2 pi$ is used in equation (1). Either 1 or 0.5 are the probability (Pi). As a result,

$$log_2(1) = 0 \tag{2}$$

for Pi=1

$$log_2(0.5) = 1 \tag{3}$$

for Pi=0.5

The entropy is low where the likelihood of one result is close to 0 or 1; if the probability of both results is 0.5, the entropy is optimum shown in Figure 5.

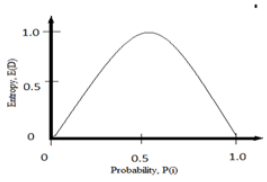

FIGURE 5. Entropy Vs probability for a two-class variable.

For Example, '*AAAAAAAA*' has one class and probability 1, then entropy can be:

$$E(D) = -1 \log_2(1) = 0 \tag{4}$$

Similarly, '*AAABBCD*' has 4 classes A, B, C, and D, the corresponding probabilities are 4/8, 2/8,1/8 and 1/8, then entropy would be:

$$E(D) = -\frac{4}{8}\log_2\frac{4}{8} - \frac{2}{8}\log_2\frac{2}{8} - \frac{1}{8}\log_2\frac{1}{8} - \frac{1}{8}\log_2\frac{1}{8} = 1.75 \tag{5}$$

for, '*AABBCCDD*', the entropy would be 2. The results for the three cases are 0, 1.75 and 2. It happened that the corresponding entropy was low, medium, and high.

Shannon entropy can be used in data hiding and even discovering the secret in the cover file. The world's largest data exchangers of any data around the world are social media networks. Intruders play a negative role in attempting to gather information to locate and harm users on Social media networks. Intruders often use Shannon entropy for Steganalysis; it's also used to shield hidden data from the intruder.

The concepts of information entropy and information imperceptibility are inextricably linked. The formula above (see equation no. 1) is used to determine whether or not there is any uncertainty or variation in the probability distributions of the given texts. Different kinds of steganography employ various strategies to conceal data. Some may alter the cover text's structure, while others may not. Information entropy is used by attackers to analyze concealed messages within Stego text. If the Stego text is well-known, steganalysis using the entropy formula is simple. The entropy value is a measure of uncertainty that was determined for a paragraph with varying entropy values as shown in the texts above (Figure 6). The formula produces entropy values of 4.06 and 4.02 when applied to regular Sentence case text (a) and Upper case text (b). The third example (c) uses the same text but adds a newline and a space to encode a secret with entropy of 4.09. The high entropy indicates that the data has a lot of variation and so contains a lot of data and/or noise. Noise is the worthless information in the data, yet they both contribute to the entropy, making the text suspect. The entropy of the selected image can be extracted and an entropy matrix can be defined. And then, according to some algorithms, the random hidden message is mapped on images to create a full library of features; therefore, establish hidden communication without altering the original images [20]. The entropy of data will be altered because after embedding another test message into cover images, the number of near color pixel pairs increased and in turn entropy changes can effectively represent the hiding ratio by examining the information characteristic of the cover image and Stego-image [21]. The ability to exploit the entropy function of at least four-letter words in a regular text file by using their constituent characters' alphabetic ordering to hide information. In the sense that embedding capability is very appreciable, the experimental results have shown encouraging features, however there is no significant difference in file sizes between the cover file and the Stego file.

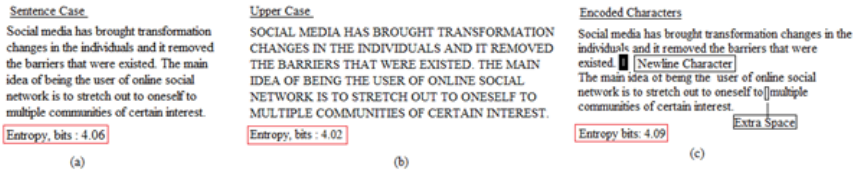

FIGURE 6. Information entropy of a paragraph, showing varying value when subjected to entropy calculation of (a) Sentence case, (b) Upper case (c) after applying word-shit and extended line embedding algorithm.

VII. EVALUATION OF STEGANOGRAPHIC SYSTEMS

Undetectability, robustness, and payload capacity are the three essential features of data concealing technologies. Undetectability is linked to the Steganographic system's security or the difficulty of detecting the existence of concealed data and it is a primary requirement [30]. The MSE (Mean Square Error) and PSNR (Signal to Noise Ratio) formulae are used to calculate the performance of Steganography approaches. MSE is a statistic that is used to assess the quality of an image in general. A peak error is shown by PSNR. The higher the PSNR, the better the image quality; for MSE, the number should be as low as feasible to indicate a good system. The pixel value difference between the cover image and the Stego image can be used to assess the Steganography method's undetectability. This is determined using the MSE formula below; the higher the MSE number, the more suspicious the Stego image is.

$$MSE = \sum_{a=1}^{n} \frac{(p_a - q_a)^2}{N} \tag{6}$$

where, $a = 1, 2, 3, , \ldots, n$ indicates first to last pixel in an image. p_a, is $a^t h$ iteration value of the Cover object. q_a, is $a^t h$ iteration value of Stego object. N, is the total number of pixels in a chosen image. PSNR is expressed in decibels and computed using MSE as the basis.

$$PSNR = 10 * \log\left(\frac{(\text{ Peak Signal Value })^2}{MSE}\right) \tag{7}$$

where, MSE is value of the previous equation (4) and Peak Value Signal of cover image, which is calculated as maximum of $(p_a - q_a)$. In general, the PSNR value decreases as the amount of concealed data rise. Thus, the payload capacity and PSNR are reciprocal to each other [31].

If an attacker suspects the Stego text based on what they see or observe, statistical approaches can be employed to uncover the secret, provided the attacker is competent. The probability distributions of the Cover and Stego media must be same, and the Stego medium should not be distorted once the embedding technique is used. If any image processing, or any kind of transformations is done after the concealment of message, a Steganographic system is said to be robust since the concealed data and cover media are not affected. Accidental alterations and human-intentional changes are two types of attacks that might occur on an image. If the secret is intact within the Stego image in both situations, it is said to be ''robust'' [22]. The most significant characteristic of steganography systems is payload capacity. It refers to the quantity of information concealed within the cover item. The concealed text must be less than the volume of the cover. The payload capacity is determined by both the method and the type of cover employed. Normally, an image cover file may store a significant amount of hidden data, but a text cover file has a restricted amount of hidden data, because the image includes more repeating bits than the text, and vice versa.

VIII. BRIGHT SIDE OF SOCIAL MEDIA

It has been more than 10 years social media in the arena and offering a multitude of opportunities to the entire world. However, Social media has a destructive hidden face as well. It is a powerful tool to share, consume and create information in a very little time. There is a vast difference that conventional way of sharing thoughts, and Social media. Social media has been playing a key part in democratizing voiceless people.

Secret material may be safely shared using Steganography via pictures on Google+ Twitter; with Facebook being the most used OSN nowadays. Similarly, Flickr is now the most popular photo-sharing site. Of course, LSB embedding in images is always subject to steganalysis, whereas DCT-based image embedding on social media networks is secure against steganalysis. On OSN, text steganography using generated Stego text is more invisible.

Social networking is one of the top technologies used by billions of people on daily basis across various platforms of hardware and software and network. It is a technology constantly offering new ideas and innovations, helping business, Science and Research, Education, social relationships, and so on. Social media generates enormous amounts of information that is useful for every organization to measure the customer behaviors for promoting their products. To learn, unlearn and relearn new skills and information, learners have an ocean of data on social media. Social networking can be considered a category of digital Internet technologies for the sharing of digital content by users.

Facebook, Twitter, Instagram, WhatsApp, etc., as shown in are the major platforms of social media today, where people do a lot of creating, sharing, and consuming of information [23]. The people who are in constant touch with Social media for over the past several years might recall the olden day social media platforms such as MySpace and Friendster eventually transformed into Facebook, and others.

Politics, autocratic regimes and the activity of social media worldwide are critical trend in the organization of protests. Social media is bringing reforms in the sectors of economic, cultural grievances. Major reform has been taken place in conventional mediums such as newspapers and television due

to the advent of Social media. Those who had the wisdom and being pushed to corners, due to the bad politics, now they can voice through social media and bring changes in the society [24]. Social media has the power to unite the people, and there is a tendency for teaching a lesson to tyrannical democratic forces. This will enable political figures to work on their limitations.

Most of the politicians today use a social media platform for analyzing the new future amendments. Indian present Prime Minister, Narendra Modi is a frequent user of social media; for the election campaign he used the social media and ultimately, he won the elections. He is known for reaching out the masses through social media, as per the 2019 statistics prime minister has over 5,00,00,000 followers.

Social media is an interactive multitasking network, facilitates in finding useful people on the Internet and creates an environment to share their views, opinions, suggestions on most of the topics. Social media provides users with a great virtual resource for solving day to day issues [39]. Social media provides situational and reliable information during the bad times of natural disaster, war, societal infighting, etc., and conveys warning messages and awareness about the situations. It facilitates users with information revelation through tweets and enables them to manage such situations.

Social media-based digital surveillance data are very much crucial in diagnosing the deceased status, Twitter for that matter extensively used social media platform [25]. Content categorization of Twitter information is possible through a keyword search or NLP tools to properly identify the health conditions due to epidemics, particularly COVID-19 kind of issues. The tweets made by patients, doctors, relatives, and others help to build models of for such pandemics.

Social media in the research is another vital domain for scientists, researchers, scientific bloggers, etc. for publishing their work. Here only the true work can be published; skeptical researchers hesitate to publish because the reactions are instantaneous. Therefore, Social media research platform is becoming a trusted medium. The medium improves the confidence level of true researchers; they get instant suggestions, critics, and opinions that would help the researcher in a greater extent. Research Gate is one of the academic social media outlets. The work that is published here can reach thousands of researchers and engage in discussions with eminent researchers and build a researcher network.

IX. DARK SIDE OF SOCIAL MEDIA

Social media is very well incorporated into the younger generation's lives. Approximately 83% of teenagers have smart phones and spend more than 20 hours a week on social media, triggering results, low self-esteem, anxiety, and depression [26]. Specifically, young women users are high in such platforms. An Australian Social media star, Essena O'Neill in 2015 quits from social media, though she earned millions of dollars out of it. On Instagram, she had 5000 followers, 250,000 on YouTube and 60,000 on Snap chat. Social media is not a real world in her view and it did not offer her satisfaction, and eventually, she left all her social media accounts and said "I don't want to spend hours and hours of my time scrolling, watching and comparing myself to others, serving no real purpose other than self-promotion" [44]. Social media always sweet and vibrant, gorgeous, colorful; however, in reality it is very hard and not so pretty.

Social media users often say that they are more connected compare to earlier days. One can gain popularity by posting pictures with minimum clothing and videos and assume that they are celebrities, in a way they are living in the false world. The same picture later can turn into pornographic photo and being harassed to an extent of suicide [27]. In this day and age, teenage boys and girls engage an excessive time on social media, posting unnecessary information such as videos, photos, texts, selfie without worrying that they are on public platforms. Most of the girls often believe that the act committed by them on social media is, their right. Even feminists claim that wearing minimum clothing is their right.

Children are the worst affected due to undesirable content on Social media. They are the ones who easily become prey to cybercriminals. The harmful information such as provoking images, videos, bullying, hatred, horror, and so on make them upset and harm mentally. Particularly, at a very young age, children develop some sort of depression and anxiety tends to lose interest and enthusiasm in their life. Certain times, they post their personal information on social media and place themselves into trouble. Such information might be used by criminals to attract children through fake profiles and start cyber bullying and intentionally upset them. Sometimes perpetrators call children to an isolated place as well for ransom.

Steganography, a technology that is used along with cryptography to protect the privacy of any sensitive information transmitted using an innocent cover text. The same technology along with Social networks is often cursed by the people, because it has a dark side application; it is being used by the terrorists to send their code words, illicit messages, to recruit terrorists at remote places and to maintain their loyalty, so on. As we all know Social media is generating huge amount data, in that relatively insignificant number of this information are being used for terrorist's activities to conceal their unlawful messages, and to create fear, tension, panic in the minds of public. This kind of Steganography messages is too hard to detect.

X. INSTANCES OF STEGANOGRAPHY ON SOCIAL MEDIA

Some of instances and approaches of steganography utilizing social media networks are presented below:

- Vawtrak is a financial virus that was originally discovered as an information stealer on Social media networks in August 2013. The goal is to conceal the URL for downloading the configuration from the victim's computer. To conceal the URL, the LSB technique was employed.
- zbot is a Trojan virus designed by Hamza Bendellladj that attacks computers in order to steal sensitive

financial data. This is based on image steganography, which is used to hide the configuration file at the end of a JPEG file and distribute it over social media channels in order to steal sensitive data from genuine users.

- Some of the largest malvertising efforts have been linked to the AdGholas organisation. This virus conceals harmful JavaScript in picture, text, and HTML files in order to carry out destructive online social activities.
- Since September 2009, the FAKEM RAT (Remote Access Trojans) appears to have been actively utilised in cyberattacks. MSN and Yahoo Messenger traffic, as well as HTTP Conversation traffic. This was used to conceal commands and manage traffic.
- In 2017, SyncCrypt, a new malware, was found. This employs a form of image steganography to conceal the ransomware's basic components.

XI. CURRENT TRENDS AND STATUS OF SOCIAL MEDIA

There have been enormous developments in the world of information technology and social media. Knowledge exchange has become instantaneous with the advent of mobile technology [28].

Twenty years have now passed for social media. Social media, which has seen a lot of change since then, it is the breath of the people [29]. Mobile invention is making the most of the demand for social media day by day. People spend most of their time in mobile and social media. With this technology, information can be sent in a moment, whenever and wherever you want.

Why can individuals get so hooked to social media? This is because they are amused by mutual knowledge. This will improve the relationships in terms of business, family, community, social wellbeing, etc. 2020 has witnessed the use of Internet and social media a phenomenal increase compares to previous year, this is because the effect COVID-19. According to the survey conducted in 2020 by Global snapshot, about 4.57 billion users are on the net, out of that 346 million people have been added in 2020 itself. The Figure 7 gives some more statistics of Social Media.

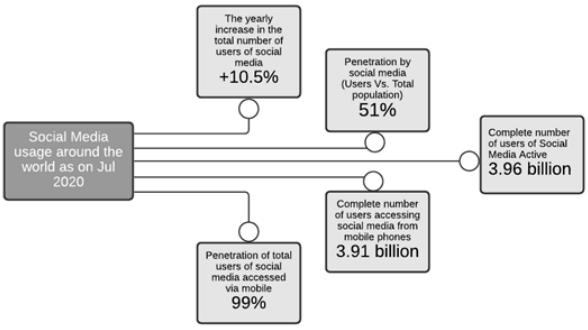

FIGURE 7. Social media usage around the world as on July 2020.

As per the Global digital and social media usage summary July 2020, Facebook at present leading with 2603 million users compare to rest of the Social Media Apps [30]. WhatsApp and YouTube stand second with 2000 million users. Others are no way near to these three Applications. The following Table 1 compares the active users as on July 2020, with social networks channels.

TABLE 1. Statistics showing the status of social media active users as on July 2020.

Social Media Channel	Social Media Active Users in Millions as on July 2020
Facebook	2603
YouTube	2000
WhatsApp	2000
Instagram	1082
TikTok	800
Twitter	326

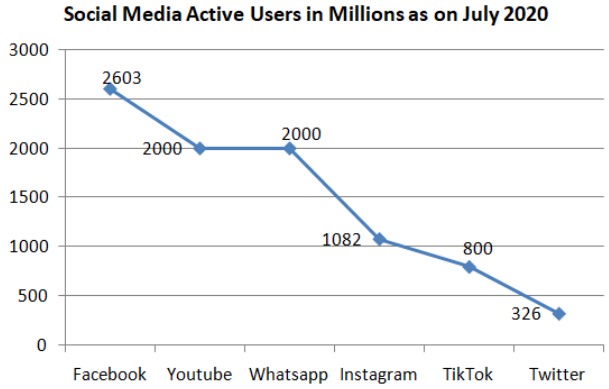

FIGURE 8. Comparison of individual social media users in millions against the total users of social media.

The corresponding line graph (Figure 8) is shown and Facebook stood top with 2603 million users against the total users of Social Media [31], [32]. The statistic indicates the providers and social media networks that are actually being used by the entire planet. The most famous are briefed out of these as under.

A. FACEBOOK

Mark Zuckerberg a psychology student at Harvard University in 2004 launched Facebook. At his young age, 23, he was into the development of social-networking websites for his friends namely, Coursematch, Facemash. It became Facebook.com in 2005, till such time it was used for the purpose of education, and later extended to the general public [33]. He made Facebook service free and made a profit through advertisement. In 2006 Yahoo and Google showed interest to buy it, however, Mark refused to do so and rest was the history.

B. YouTube

YouTube was developed by Chad Hurley, Steve Chen, and Jawed Karim. They were recruited by PayPal. The inspiration for making YouTube was two videos, one of which was Janet Jackson's wardrobe malfunction and the other, the Indian

Ocean tsunami. In 2005, the first-ever Me at the Zoo video was streamed. In 2019, revenue generated by YouTube was $46 billion and profit was $10.7 billion.

C. TWITTER

Twitter is an online micro blogging service, developed using Ruby on rails, a web-based application framework, by Evan Williams and Biz Stone in the year 2006. Earlier to launch these two were employees of Google. Basically, Twitter is short messaging service unlike SMS, the messages posted by the user goes to the Twitter servers and thereby it is broadcasted to the followers of the Twitter account. Therefore, messages became tweets. Presently, the Twitter INC. has revenue of 3.46 billion US dollar, about 4600 employees, headquartered at California, San Francisco, US, and the number of active users 326 million as on 2019 February.

D. INSTAGRAM

Kevin Systrom, a graduate of Stanford University and an employee of Nextstop, a start-up company, founded Instagram in 2009 [34]. Systrom was a computer illiterate, studied the program codes in weekends and developed a web app called Burbn that allowed users to enter figure their daily plan and photo sharing. In 2010 by the virtue of venture capitalists he raised a seed fund of $500,000. After, the major modification to the Burbn, and it was re-launched as Instagram app under iOS. In 2012 Facebook purchased for $1 billion. As on 2020 every day 500 million people are using the Instagram service.

XII. CONCLUSION

The future of social media will still be promising, but the situation at present is in the midst of good and bad. With leaps and limits, social media users can continuously increase. There will be no end to creativity in the areas of science and research, medicine, health care, education and human relationships, politics, etc. in social media. New services will be provided to the public by today's online social media network giants. New trends can enter the world of social media and good old social media services can disappear. There will be even better benefits for all aspects of society.

The other side of social media is going to be less evil than before, because the darker side of it is now more conscious of it. Innovations in security would undoubtedly safeguard the rights of consumers. The highest importance of child protection will be provided. Using social media networks as covert channels to hide users' sensitive data, Steganography attack detection will be given prominence.

FUTURE SCOPE

Extended line and Word shift techniques are both text-based and, as a result, are sensitive to OCR detection. Because the structure of the Stego file has been altered to some extent while concealing the data, it raises suspicion, prompting attackers to attempt to reveal the secret. In the case of image data concealing, Face book's data analysis sampling method may damage the embedded data to some extent. Due to the tradeoffs between image quality and embedding capacity, the distortion component in image concealment is highly matters. As the payload capacity grows, so does the distortion of the image. Because of its ease of use, utilizing WhatsApp covert channels to communicate hidden data in text and photos is more popular than using other social media networks. a large number of Internet users use WhatsApp to share large amounts of data, making them more vulnerable to vulnerability. The security flaw in WhatsApp is due to attackers' use of steganography for malicious purposes, resulting in data loss. The traditional steganography methods, such as word shift, line shift, and related approaches, alter the structure of the cover text and make it suspect. Furthermore, when scanners and OCR tools are employed on these changed structures, the concealed data is lost. The payload capacity of these traditional approaches is quite limited, thus the answer is to use generative methods based on neural networks. Covert social media platforms are currently more popular for conveying camouflaged info to targets than for using it for good; assailants are more likely to utilize them for destructive reasons. Controlling such an act is extremely challenging owing to the massive amount of data produced every day [35]. The majority of terrorist actions are carried out utilizing steganography on social media networks. Organizations, which run social media networks are concentrating their efforts on growing their networks by offering more user friendliness and simple access for the sole purpose of attracting more members, while ignoring user privacy concerns. More study into these areas will lead to a brighter tomorrow, at least until the adoption of a fully functional semantic web or web 3.0.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES

[1] S. Pramanik, D. Samanta, S. Dutta, R. Ghosh, M. Ghonge, and D. Pandey, "Steganography using improved LSB approach and asymmetric cryptography," in *Proc. IEEE Int. Conf. Advent Trends Multidisciplinary Res. Innov. (ICATMRI)*, Dec. 2020, pp. 1–5.
[2] S. R. Yaghobi and H. Sajedi, "Text steganography in webometrics," *Int. J. Inf. Technol.*, vol. 13, no. 2, pp. 621–635, Apr. 2021.
[3] M. Haenlein and A. Kaplan, "A brief history of artificial intelligence: On the past, present, and future of artificial intelligence," *California Manage. Rev.*, vol. 61, no. 4, pp. 5–14, Aug. 2019.
[4] R. Abduljabbar, H. Dia, S. Liyanage, and S. A. Bagloee, "Applications of artificial intelligence in transport: An overview," *Sustainability*, vol. 11, no. 1, p. 189, Jan. 2019.
[5] X. Du-Harpur, F. M. Watt, N. M. Luscombe, and M. D. Lynch, "What is AI? Applications of artificial intelligence to dermatology," *Brit. J. Dermatol.*, vol. 183, no. 3, pp. 423–430, Sep. 2020.
[6] M. N. Abdali and M. Z. Hussain, "Reference-free detection of LSB steganography using histogram analysis," in *Proc. 30th Int. Telecommun. Netw. Appl. Conf.*, Nov. 2020, pp. 1–7.
[7] N. Shahid, T. Rappon, and W. Berta, "Applications of artificial neural networks in health care organizational decision-making: A scoping review," *PLoS ONE*, vol. 14, no. 2, Feb. 2019, Art. no. e0212356.
[8] L. Y. Por and B. Delina, "Information hiding: A new approach in text," in *Proc. 7th WSEAS Int. Conf. Appl. Comput. Appl. Comput. Sci. (ACACOS)*, Hangzhou, China, Apr. 2008.

[9] H. Mohammad and N. Tayarani, "Applications of artificial intelligence in battling against COVID-19: A literature review," *Chaos, Solitons, Fractals*, vol. 142, Jan. 2021, Art. no. 110338.
[10] O. Zawacki-Richter, I. V. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education—Where are the educators?" *Int. J. Educ. Technol. Higher Educ.*, vol. 16, no. 1, p. 39, Oct. 2019.
[11] Z. Yang, L. Xiang, S. Zhang, X. Sun, and Y. Huang, "Linguistic generative steganography with enhanced cognitive-imperceptibility," *IEEE Signal Process. Lett.*, vol. 28, pp. 409–413, 2021.
[12] B. Mandal, A. Pradhan, and G. Swain, "Adaptive LSB substitution steganography technique based on PVD," in *Proc. 3rd Int. Conf. Trends Electron. Inform.*, Apr. 2019, pp. 459–464.
[13] K. Tiwari and J. Sahil Gangurde, "LSB steganography using pixel locator sequence with AES," in *Proc. ICSCCC*, 2021, pp. 302–307.
[14] Y.-H. Wu, S. Zhuang, and Q. Sun, "A steganography algorithm based on GM model of optimized parameters," in *Proc. ICCEA*, 2020, pp. 384–387.
[15] N. Wu, P. Shang, J. Fan, Z. Yang, W. Ma, and Z. Liu, "Research on coverless text steganography based on single bit rules," *J. Phys., Conf. Ser.*, vol. 1237, Jun. 2019, Art. no. 022077.
[16] R. Gurunath, H. Ahmed Alahmadi, D. Samanta, M. Z. Khan, and A. Alahmadi, "A novel approach for linguistic steganography evaluation based on artificial neural networks," *IEEE Access*, vol. 9, pp. 120869–120879, 2021.
[17] J. Deng, M. Tang, Y. Wang, and Z. Wang, "LSB color image embedding steganography based on cyclic chaos," in *Proc. ICCC*, 2019, pp. 1798–1802.
[18] C. Chang and S. Clark, "Practical linguistic steganography using contextual synonym substitution and a novel vertex coding method," *Comput. Linguistics*, vol. 40, no. 2, pp. 403–448, Jun. 2014.
[19] L. Xiang, S. Yang, Y. Liu, Q. Li, and C. Zhu, "Novel linguistic steganography based on character-level text generation," *Mathematics*, vol. 8, no. 9, p. 1558, Sep. 2020.
[20] R. Indrayani, "Human perception evaluation toward end of file steganography method's implementation using multimedia file (image, audio, and video)," in *Proc. ICITISEE*, Nov. 2019, pp. 200–204.
[21] M. Grosvald and C. O. Orgun, "Free from the cover text: A human-generated natural language approach to text-based steganography," *J. Inf. Hiding Multim. Signal Process.*, vol. 2, no. 2, pp. 133–141, Apr. 2011.
[22] Z. Yang, X. Guo, Z. Chen, Y. Huang, and Y. Zhang, "RNN-stega: Linguistic steganography based on recurrent neural networks," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 5, pp. 1280–1295, May 2019.
[23] A. Darbani, M. M. AlyanNezhadi, and M. Forghani, "A new steganography method for embedding message in JPEG images," in *Proc. KBEI*, Feb. 2019, pp. 617–621.
[24] R. Gurunath and D. Samanta, "Studies on encrypted secret data storage techniques analogous to steganography," *Int. J. Adv. Sci. Technol.*, vol. 29, pp. 3705–3711, Dec. 2020.
[25] M. Sharifzadeh, M. Aloraini, and D. Schonfeld, "Adaptive batch size image merging steganography and quantized Gaussian image steganography," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 867–879, 2020.
[26] springerprofes-sional.de. *Text Coverless Information Hiding Based on Word2vec*. Accessed: 2018. [Online]. Available: https://springerprofessional.de
[27] R. Gurunath, M. Agarwal, A. Nandi, and D. Samanta, "An overview: Security issue in IoT network," in *Proc. I-SMAC*, Aug. 2018, pp. 104–107.
[28] S. K. A. K. D. Samanta and M. Paul, "Message encryption using text inversion plus n count: In cryptology," *Int. J. Inf. Sci. Intell. Syst.*, vol. 3, no. 2, pp. 71–74, 2014.
[29] V. Dhanush, A. R. Mahendra, M. V. Kumudavalli, and D. Samanta, "Application of deep learning technique for automatic data exchange with air-gapped systems and its security concerns," in *Proc. ICCMC*, Jul. 2017, pp. 324–328.
[30] D. Watni and S. Chawla, "A comparative evaluation of jpeg steganography," in *Proc. ISPCC*, 2019, pp. 36–40.
[31] T. Fang, M. Jaggi, and K. Argyraki, "Generating steganographic text with LSTMs," 2017, *arXiv:1705.10742*.
[32] A. G. Benedict, "Improved file security system using multiple image steganography," in *Proc. IconDSC*, Mar. 2019, pp. 1–5.
[33] Y. Li, J. Zhang, Z. Yang, and R. Zhang, "Topic-aware neural linguistic steganography based on knowledge graphs," *ACM/IMS Trans. Data Sci.*, vol. 2, no. 2, pp. 1–13, Apr. 2021.
[34] D. Samanta, "A novel approach for semantic web application in online education based on steganography," *Int. J. Web-Based Learn. Teach. Technol.*, vol. 17, no. 4, pp. 1–13, Sep. 2022.
[35] N. Manohar and P. V. Kumar, "Data encryption & decryption using steganography," in *Proc. ICICCS*, May 2020, pp. 697–702.

R. GURUNATH is currently an Assistant Professor with the Dayananda Sagar College of Engineering, Bengaluru, India, and also a Research Scholar with the Department of Computer Science, CHRIST (Deemed to be University), India. His research interest includes text steganography.

MOHAMMAD FADEL JAMIL KLAIB received the B.Eng. degree in electronic engineering from Al-Quds University, Jerusalem, in 2002, the M.Sc. degree in computer engineering from Near East University, Cyprus, in 2004, and the Ph.D. degree in software testing from the Electrical and Electronic School, University Sains Malaysia, in 2010. He is currently working as an Assistant Professor with Taibah University, Saudi Arabia. His research interests include software engineering, testing automation, parallel computing, algorithm design, AI, and big data.

DEBABRATA SAMANTA (Member, IEEE) received the Ph.D. degree in computer science and engineering from the National Institute of Technology, Durgapur, India, in the area of SAR image processing. He is currently working as an Assistant Professor with the Department of Computer Science, CHRIST (Deemed to be University), Bengaluru, India. He is keenly interested in interdisciplinary research and development and has experience spanning fields of SAR image analysis, video surveillance, heuristic algorithm for image classification, deep learning framework for detection and classification, blockchain, statistical modeling, wireless ad hoc networks, natural language processing, and V2I communication. He has successfully completed six consultancy projects. He has received funding under International Travel Support Scheme in 2019 for attending conference in Thailand. He has received Travel Grant for speaker in Conference and Seminar for a period of two years, since July 2019. He is the owner of 18 patents (two design Indian patent and two Australian patent granted, and 14 Indian patent published) and two copyright. He has authored or coauthored over 154 research papers in international journals (SCI/SCIE/ESCI/Scopus) and conferences, including IEEE, Springer, and Elsevier conference proceeding. He is the coauthor of ten books and the co-editor of five books, available for sale on Amazon and Flipkart. He has presented various papers at international conferences and received best paper awards. He has authored or coauthored of 20 book chapters. He has received "Scholastic Award" at 2nd International conference on Computer Science and IT application, CSIT-2011, Delhi, India. He also serves as an Acquisition Editor for Springer, Wiley, CRC, Scrivener Publishing LLC, Beverly, USA, and Elsevier. He is a convener, keynote speaker, session chair, co-chair, publicity chair, publication chair, advisory board, and technical program committee members in many prestigious international and national conferences. He was invited speaker at several institutions.

MOHAMMAD ZUBAIR KHAN received the Master of Technology degree in computer science and engineering from U. P. Technical University, Lucknow, India, in 2006, and the Ph.D. degree in computer science and information technology from the Faculty of Engineering, M. J. P. Rohilkhand University, Bareilly, India. He was the Head and an Associate Professor with the Department of Computer Science and Engineering, Invertis University, Bareilly. He has more than 15 years of teaching and research experience. He is currently an Associate Professor with the Department of Computer Science, Taibah University. He has published more than 60 journal articles and conference papers. His current research interests include data mining, big data, parallel and distributed computing, theory of computations, and computer networks. Since 2004, he has been a member of the Computer Society of India.