2018 International Conference on Advances in Computing and Communication Engineering (ICACCE-2018)
Paris, France 22-23 June 2018

Social Media Security

Analysis of Threats and Security Measures

[1]Sanyuj Singh Gupta,[2]Abha Thakral,[3]Tanupriya Choudhury
[1,2]Amity University Uttar Pradesh,Noida,India,[3]University of Petroleum and Energy Studies,Dehradun,India
[1]sanyujgupta@gmail.com,[2]abhareads@gmail.com,[3]tanupriya1986@gmail.com

***Abstract*— 20th century has connected us in ways that we were never connected in before, our privacy, our fate are in our hands like never before. The stakes are at its highest. We must take care of our online footprint such that we must not be filled with regret and/or be fooled by people online, it's our own responsibility of what happens in the online realm. This paper analyzes the different threats to the ever growing network of social media, to its clients and users alike. Then tries to give solutions for the same.**

Keywords: Security; Social Media; Facebook; Gmail.

I. Threats to Social Media Networks

A. Media content threats

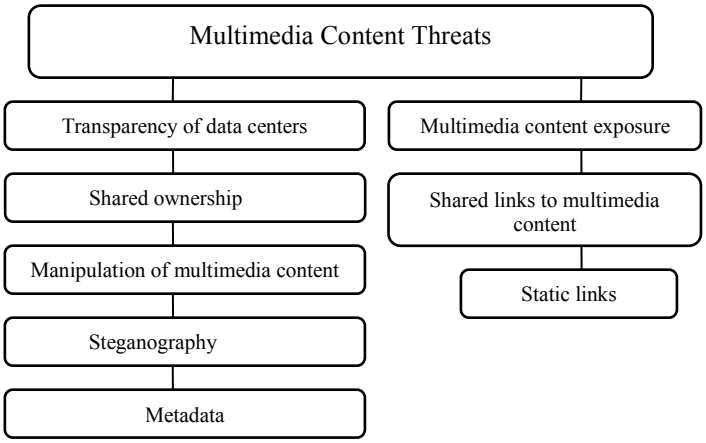

Fig. 1. Multimedia Content Threats([3],[5])

A.1 Multimedia content exposure

Shared mixed media information on social networking services can straightforwardly uncover huge measure of client's touchy data, for example, client's street number, late activities. Impacts of these types of threats are Information revelation, Reputation misfortune, Location spillage, Cyber provocation, Profiling and loss of Safety.[2]

A.2 Shared ownership

Multimedia information partook in social networking services may identify with various clients and just a single client can choose the favored protection settings[2] for the mixed media data. Main Impact being the proprietorship loss on the Content

A.3 Control of interactive media content

In social networking services, a pernicious client can alter the individual pictures of real clients to damage or derision them.[2] Impacts of this ownership breach are loss of reputation, extortion/Blackmail and Cyber-Bullying.

A.4 Steganography

A pernicious client can share malevolent data by covering it inside sight and sound information, for example, a picture of a cat may have a [8] malicious code that deletes the system files once opened, or it may steal personal information etc. This results in loss and misuse of data, and reputation, and makes one's safety compromised.

A.5 Metadata

Multimedia substance go about as a metadata in light of the fact that these substance may uncover other profitable information, for example, IDs, location. This results in Information exposure,[13] Location spillage, Reputation misfortune, Cyber stalking, Profiling & Safety compromise.

A.6 Shared links to interactive media content

Social networking services give a component in which clients can share sight and sound substance in[5] unsupported organization. This results in Reputation misfortune, Information divulgence, Account loss.

A.7 Static links

48.6% of the clients[7] in social networking services utilize static connects to share the interactive media information. This results in Multimedia information exposure, Data possession loss.

A.8 Video/Audio conferencing

Noxious client can capture the communicate video stream by misusing the conceivable vulnerabilities in hidden correspondence architecture. This results in Reputation misfortune, Information revelation, Blackmailing, Cyber-bullying, and Cyber stalking.

A.9 Tagging

Tagging may connect the general population with social networking services who are not the individuals from any social networking services and would prefer not to share any of their own information. This results in Multimedia information leakage, Attackers have access to location, Loss of ones reputation, Cyber-Crime.

978-1-5386-4485-0/18/$31.00 ©2018 IEEE

A.10 Data Breach

The Social Media Client may Publically post pictures which are meant for the sole use of the individual. This is a great violation of the individuals privacy, and the from the past experiences like "The Fappening"[15]. This results in Reputation misfortune, Information exposure, Location spillage, Content proprietorship loss, Theft of identity, Extortion/Blackmailing, Cyber stalking,

B. Traditional Threats

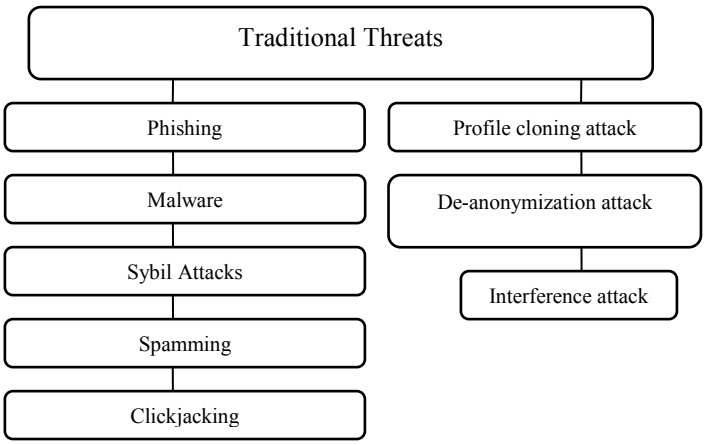

Fig. 2. Traditional Threats([3],[5])

B.1 Phishing

Phishing is the process used to acquire delicate media or information, such as, passwords, usernames,charge card subtle elements (and cash) for mallicious reasons, that is done via camouflaging as a dependable service, which is known not to be fraudulent or mallicious in an electronic correspondence. Phishing[2] is derived from the fishing, which is basically "baiting" people. Phishing is generally done via email mocking or texting, it generally guides the clients to enter private or sensitive data via a fraudulent web-site, the asthetics of which are not distinguishable to the real website and generally the main contrast is the URL of the site in concern. Correspondences implying to be from social sites, sell off destinations, online installment directors or IT chairmen, banks, are frequently used to bait-in casualties. Phishing media generally spam to web-services that are known to spread malware content.

B.2 Malware

The malevolent URLs[3] can divert the client to counterfeit sites, and later, transmit malware to client's PC for taking secret information of user. This leads to Revelation Confidential data, Account misfortune, Data possession misfortune.

B.3 Sybil Assault and Phony profile

Malicious clients can oversee and handle a few phony characters in social networking services. By working these phony personalities,[3] they can outvote the honest to good users. This results in Reputation misfortune, Corrupt client's data, Extortion/Blackmailing, Pornography, Cyber harassment.

B.4 Spam

Using social networking services, Attackers may send spontaneous messages called as/known as spam[2] in a mass add up to the social networking service users. This may result in Reputation misfortune and Account loss.

B.5 Clickjacking

Attackers cover up malignant applications behind complex UIs' or catches to take the snaps of clients and utilize them for the vindictive purposes.[2] This may result in Data disclosure, Click stolen, Reputation loss, Blackmailing, Decrease user's experience, theft of users' online identities.

B.6 De-Anonymization attack

Attackers utilize the techniques, for example, client gather enrollments, arrange topologies, following treats to uncover the client's genuine character in social networking services[5]. This results in Identity divulgence, Relationship exposure, Reputation misfortune, Profiling.

B.7 Surmising attack

An aggressor derives client's private data by misusing other distributed data about the client on social networking services.[7] Private data spillage, Location spillage, Identity revelation, Relationship divulgence, Reputation misfortune, Profiling.

B.8 Profile Cloning attacks

Attackers may clone an effectively existing client's profile to accumulate touchy confidential data about the client's companions or to submit a few kinds of web scam.[3] This results in Cyber stalking, Sensitive data spillage, Cyber-bullying, Extortion or Blackmailing Cyber harassment, Account misfortune, Reputation misfortune.

C. Social Threats

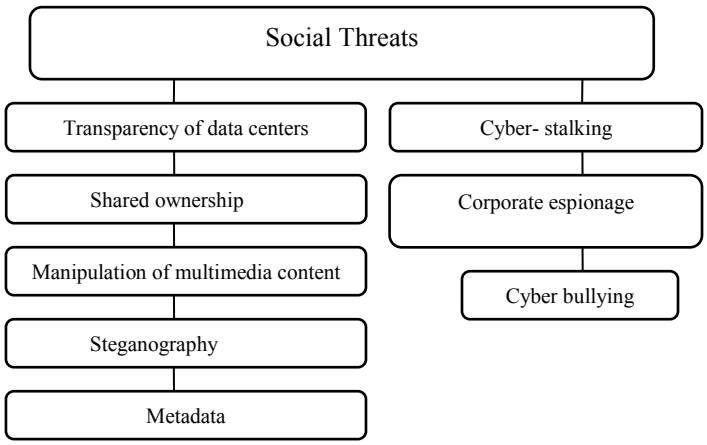

Fig. 3. Social Threats([3],[5])

C.1 Grooming and Cyber Bullying

Grooming is when some legal adult(i.e over 18 years of age) try or are successfully able to lure minors using the internet for the purpose of sexual assualt.[3] Children Scared for life. Many of these acts are Pedophilic and are illegal all around the world. Also Impersonating someone online, Giving wrong information online are all illegal. Teen depression, Child Pornography, are all as consequence to this.

C.2 Corporate Espionage

A social specialist can assemble valuable data, for example, representative's situation inside the organization, full name, email addresses, and numerous more about organization workers by utilizing social networking services and can penetrate the company.[1] This Affects the organization's[3] and representative's notoriety, Information spillage, Disclosure of organization approaches, Profiling.

C.3 Digital stalking

Stalkers online can get the victims home address, and then harm the person. This results in Reputation misfortune, Data revelation, Blackmail[7], Cyber badgering, Safety misfortune, Location spillage.

III. SECURITY ALGORITHMS

A. Private Key/Symmetric Algorithms

Private key algorithms use a single mystery key for encoding of the major measurement of the information and have quick preparing speed. These calculations utilize one mystery key that is known by the receiver and the sender. 3DES, Blowfish, RC6,are some best cases of the algorithm.[1]

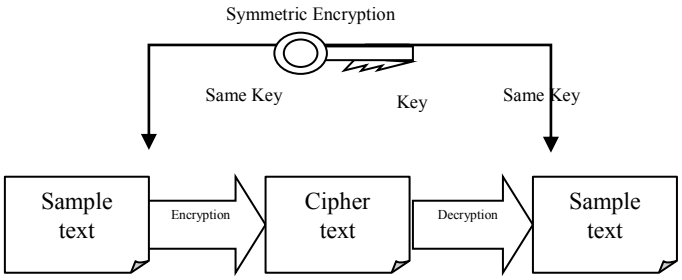

Fig. 4. Symmetric Encryption

B. Public Key/Asymmetric Algorithms

These algorithms use a two key encryption and decryption method. The public key of the receiver is used to encrypt data, by the sender, and the decrypted using the private of the receiver. RSA and Diffie-Hellman(DH)[18] are some examples of open key calculations.

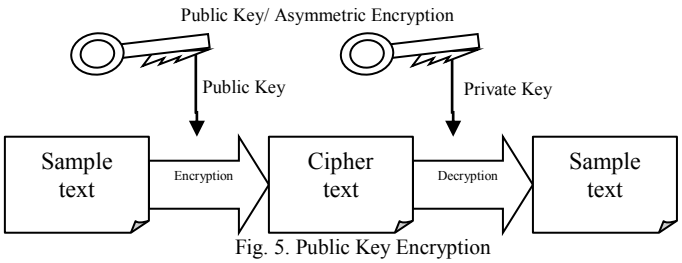

Fig. 5. Public Key Encryption

C. Signature Algorithms

These algorithms are mainly utilized in order to sign and validate media that may be single key based. Examples include: DH, RSA. That hash algorithm, when it is used as first step of a signature generation or verification algorithm, will be called "signature hash algorithm".[3] When we say something like "RSA/SHA-256"[5], we mean "RSA signature, with SHA-256 as accompanying hash function

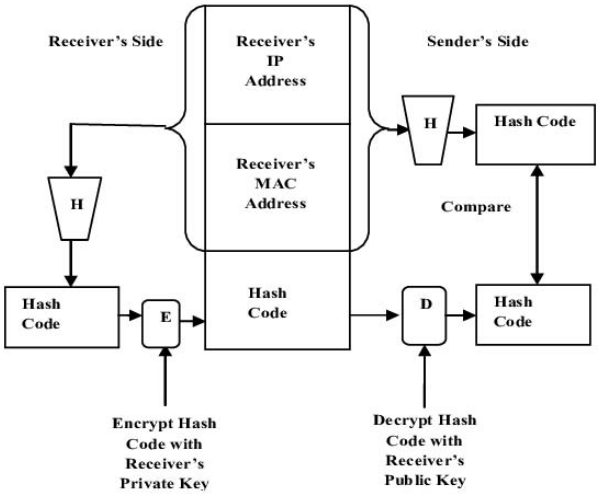

Fig. 6. Public Key Encryption

D. Hash Algorithms

A hash function is a function which takes input of any length and then returns a unique 160 bit Character string(depending on the algorithm). This unique identifier can identify the data uniquely. This is called a key. With the key, one may verify the integrity of the data. Hash Esteems are the quantities returned by the hash work. They are also called as hash codes, digests, or just hashes. One utilize is an information structure called a hash table, generally utilized as a part of PC programming for quick information query. Hash capacities quicken table or database query by distinguishing copied records in a vast document. Some examples include the MD(Message Digest) family like MD5#, SHA(Secured Hash Algorithm) family like SHA1,SHA2.

IV. PROBLEMS THAT ARRISE WITH SYMMETRIC ALGORITHMS

A. Exchanging of the Shared Secret Key via unsecure internet.

Mystery keys required by the sender and recipient amid encryption or decoding process shares Symmetric-key calculations. In the event that a third individual accesses the protected mystery key, figure instant messages can without much of a stretch be decoded. The reality of having one single mystery key calculation is the most basic issue looked by Cloud specialist organizations when managing end clients who convey over unsecure web. The main choice is to have that mystery key be changed regularly or kept as secure as conceivable amid the conveyance stage.

B. Problem confirming if substance is changed or really sent by the guaranteed sender.

In the event that a programmer has the mystery key, unscrambling figure media, changing the media that is to be sent using the same key and send to the collector. Since the solitary key can be included with our crypto procedure, either side of the exchanges that can be diminished. This information respectability[1] and non-renouncement issues ergo have to add utilization of the Hashing capacities like MD5.[2]

V. ANALYSIS OF THE AVAILABLE SOLUTIONS

As shown earlier, the threats to Social Media Networks are more apparent now than ever, if not managed the internet would become a place where no person's data is secure, privacy would be non-existent, and there would be no safety for minors and adults alike, who would be privy to stalking, murder, rape, and all the cyber-crime[15] that comes with this. However, security algorithms are being implemented everywhere. On websites, servers, and also by organizations, individuals, and governments. Main security features implemented include;

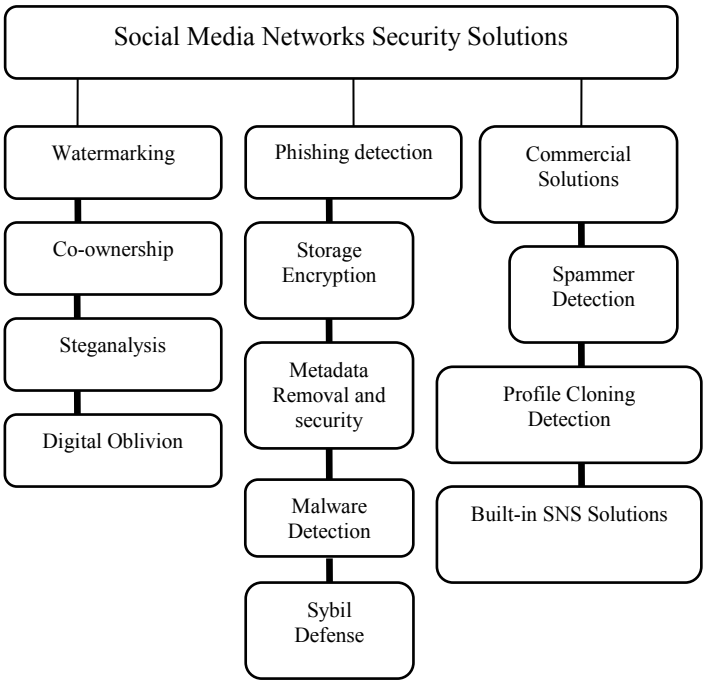

Fig. 7. Social Media Security solutions

A.1 Watermarking

This is a process in which there has been embedded a code, piece of audio, or even just a name, which identifies ownership to the content it is embedded in it. It may be Visible in the Media, like say the legal owners logo on one side of the media content. Or it is not visible when consuming the media. They may be robust, semi-delicate, or delicate. In robust watermarking, information can be recouped after pernicious assault or flag preparing is completed. Delicate watermarks can't be recouped or confirmed after basic flag preparing is done . Semi-delicate watermarks are a half and half of powerful and delicate watermarks. The nearness of watermarks in the interactive media document enables a client to track numerous exercises, for example, if different clients are re-transferring their mixed media record or adjusting it.[9]

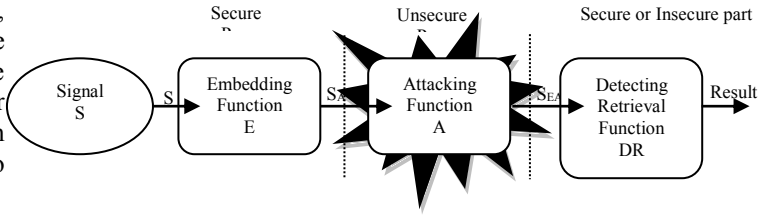

Fig. 8. Watermarking

A.2 Steganography

The method of disguising a message, video, record or picture inside another message, video, document or picture is called Steganography. The word steganography comprises of a Greek word steganos (στεγανός), [13] means "anchored, guaranteed, or disguised,". All the popular Social Media Networks provide the user with the functionality to upload high resolution images, this may result in users using this data for spreading malicious code with Steganography[13]. There exist many steganalysis methods which can detect malicious pictures and media. Steganalysis is done primarily using ML techniques in which a large data-set is used to train the machine in order to get it to learn about regular images and this then filters out the malicious images and doesn't let this spread.

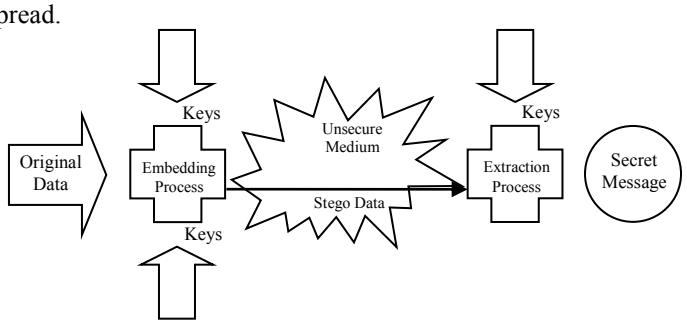

Fig. 9. Steganography[10][13]

A.3 The Co-Ownership Model

Co-Ownership model lets the users or (Co-owners) to use their privacy settings on shared content like media and mulitmedia. Squicciarini utilized the Clarke-Tax instrument, which depends on way separates. Later on, Squicciarini executed the Collaborative Privacy Management (CoPE) device as an application to guarantee the protection of pictures shared and created by social networking services clients. Adapt gives a cooperative estimation of security arrangements in light of the alternative with the most astounding number of votes.[6] They likewise exhibited a client investigation of this inside Facebook, which demonstrated that clients like the idea of community security administration and that it is valuable for ensuring the protection of their mutual interactive media information.

A.4 Digital Oblivion

This technique fundamentally an termination time for the media content. This means that the data is visible and usable for a finite period of time and after that time period it is not accessible. This method is very helpful in conservation of privacy of large amounts of data, since Social Media Networks are growing rapidly. Some tools like X-Pire![17] have been introduced which adds a timer to the image after

which access is revoked. Tools like this can improve privacy a lot. This method along with watermarking and digital signature verification are useful tools in protecting the privacy of the Social Media Network users.

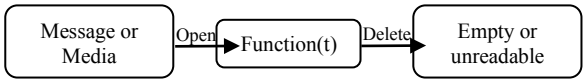

Fig. 10. Digital Oblivion

A.5 Storage Encryption

Most Social Media Networks do not own data centres. Most of the users data content is stored on third party-servers. These third-party data centers may share the users data with data distributers and or political/ malicious organizations without the users knowledge or approval.[15] These third party data centers are also used by many medical and health social media networks like WebMD,[15] practo.com , zocdoc and many more, these Networks may contain sensitive information which can be exploited and cause serious financial, mental and physical harm.

A.6 Removal of Metadata

Metadata can contain sensitive information about the user. like the physical location, credentials, government ID's and much more.[13] This information in the hands of malicious users can wreak havoc on the life of the user. They can be targeted and preyed upon by hackers, scammers, and people looking to cause harm for enjoyment or entertainment purposes. Metadata removal tool also known as Metadata scrubber[13] is a type of privacy software which protects the privacy of its users by deleting the potentially privacy-leaking metadata from files before they are shared with others. cleanDocs[13] by DocsCorp is one of the leading tools in comparison with VeExif[14]

A.7 Malware Detection

Malware detection can be done in several ways. Malware, as we know is a very serious threat and it is very common as well. One may not even know that they have been affected until it is too late. As of late, Zhu[12] proposed a Malware Propagation and Prevention Model (MPPM)[12] for Social Networking Services. This model characterizes the connections between malware recognition, clients' propensities, and malware spread in Social Networking Services. They depicted distinctive conditions of social networking services hubs in light of the qualities of social networking services and a few malware avoidance leads by utilizing dynamic development conditions. They likewise acquainted the discovery factor with control the spread of malware.

A.8 Defence against Sybil Attacks

Fake profile detection is also one of the key aspects of social media security as impersonating someone online is illegal. This is also means to extortion, crime, murder, loss of reputation and more. The Sybil assault in PC security is an assault wherein a notoriety framework is subverted by fashioning personalities in distributed systems. It is named after the subject of the book Sybil, a contextual investigation of a lady determined to have dissociative character issue. The name was recommended in or before 2002 by Brian Zill at Microsoft Research. Wei presented SybilDefender[10], which utilizes organize topology for shielding against Sybil assaults. It is adaptable and productive for extensive social networking services, and depends on playing out a constrained measure of self-assertive strolls inside the social diagrams. They tried their component on genuine topologies and guaranteed that it is powerful and effective in identifying the Sybil hubs in social networking services.[11]

A.9 Detection of Phishing and Phony Websites

Thusly, a considerable measure of against phishing strategies have been created to distinguish and counteract phishing assaults. The majority of these strategies depend on the machine-learning procedure in which highlights of sites are utilized to recognize the phishing sites.[8] With the huge development of phishing assaults on social networking services, a few analysts have proposed answers for identifying phishing assaults on social networking services. The phishAri[8] procedure for the ongoing ID of phishing that happens on Twitter[8]. This system can group tweets posted with URLs into the two classes of phishing or genuine by utilizing the tweet's substance and some particular highlights, for example, notices, hashtags, tweet length, number of devotees, number of tweets, and to what extent the record has existed. Another program, WarningBird,[9] which identifies the malicious Links that are posted via Twitter. They may counteract phishing assaults that shroud themselves by misusing the contingent redirection of URLs. The creators utilized related URL chains in various tweets to distinguish phishing. Lately it has been proposed that unsupervised procedure for recognizing phishing assaults inside Twitter. This method utilizes a two-stage unsupervised learning calculation. Additionally, It has been distinguished sending based pernicious URLs in social networking services by utilizing the three capabilities of sending based highlights, ordinary URLs highlights, and diagram based highlights.

Evaluation metric	**Naive Bayes**	**Decision Tree**	**Random Forest**
Accuracy	87.02%	89.28%	92.52%
Precision (phishing)	89.21%	88.05%	95.24%
Precision (safe)	92.12%	94.15%	97.23%
Recall (phishing)	68.32%	74.51%	92.21%
Recall (safe)	85.67%	89.20%	95.54%

Table. 1. Aanalysis of Naive Bayes,DecisionTree, and Random Forest [8]

A.10 Detection of spam

Spamming can be an effective tool for marketing teams but it is nothing more than a waste of time for regular people Proposed was graph and content-based features used to detect spam profiles via Twitter. After collecting a real dataset from publicly available information on Twitter and using it as an input for machine-learning classifiers to distinguish spam postings from the genuine posts. Maskara10 is one such smart spammer detection tool that is based on the model of Latent

Dirichlet Allocation (LDA)[7].

$$p(\mathbf{w}\,|\,\alpha,\beta) = \int p(\theta\,|\,\alpha)\left(\prod_{n=1}^{N} p(w_n\,|\,\theta,\beta)\right) d\theta,$$

Fig. 11. Latent Dirichlet Allocation[7]

"where p(wn |θ,β) are combination components and p(θ|α) are combination of weights. Figure 11 shows the interpretation of LDA. It depicts the distribution on p(w|θ,β) which is induced from a particular instance of an LDA model."[7]
This identifies spammers on the following topic features: 1) Global Outlier Standard Score (GOSS) which will expose the involvement of users to the global topic, and 2) Local Outlier Standard Score (LOSS) that shows users' involvement given a local topic. The major advantage approaching this way is, that it detects those smart spammers who may be posing as genuine users' profiles.

CONCLUSION

After analysis, it was found that 78.9% of the threats that actuate the social media networks are not actually security faults rather they prey on the users lack of knowledge[3] on handling computer systems and traversing the internet. By deceiving such users and making them click on malicious links and making them download malicious code they breach privacy of these users. Although companies like Google and Facebook use complex machine learning algorithms for spam detection, they also have a about 5% chance of not being able to detect spam.[9] By educating users of spam mails and phishing, we can actually negate out 70% of threats being faced.

REFERENCES

[1] AkashdeepBhardwajaG.V.B.SubrahmanyambVinayAvasthicHanumatSastryd Security Algorithms for Cloud Computing.

[2] ShailendraRathoreaPradip KumarSharmaaVincenzoLoiabYoung-SikJeongcJong HyukParka Social network security: Issues, challenges, threats, and solutions

[3] Barinka, Bad Day for Newsweek, Delta Amid Social-Media Hackings, Online; accessed 04 April 2017.

[4] C. Squicciarini, H. et al CoPE: enabling collaborative privacy management in online social networks J. Am. Soc. Inf. Sci. Technol., 62 (3) (2011).

[5] El Asam et al "Cyberbullying and the law: a review of psychological and legal challenges" Comput. Hum. Behav., 65 (2016), pp. 127-141.

[6] Kamilaris et al "The practice of online social networking of the physical world" Int. J. Space-Based Situated Comput., 2 (4) (2012), pp. 240-252.

[7] David M. Blei et al "Allocation Journal of Machine Learning Research 3" (2003) 993-1022.

[8] Anupama Aggarwal et al PhishAri: Automatic Realtime Phishing Detection on Twitter.

[9] Sangho Lee et al WARNINGBIRD: Detecting Suspicious URLs in Twitter Stream.

[10] Wei Wei et al Defend against sybil attacks in large social networks.

[11] Akshay Ambekar1et al India Sybil Identity Defender in Social Networks IARJ S, Engineering and Technology Vol. 2, Issue 4, April 2015

[12] Hui Zhu et al State Key Lab. of Integrated Service Networks, Xidian Univ MPPM: Malware propagation and prevention model in online SNS

[13] Hassan, Nihad et al Digital Privacy and Security Using Windows: A Practical Guide. Apress, 2017, pp. 56-59.

[14] Dennis O'Reilly. Remove metadata from Office files, PDFs, and images, CNET, May 16, 2014

[15] Alfonso, F (2014a) How Unidan went from being Reddit's most beloved user to its most disgraced (accessed 6 August 2015). Google Scholar

[16] Julian Backes, Mi– A digital expiration date for images in social networksarXiv:1112.2649v1 [cs.CR] 12 Dec 2011 Lorenz