WIFI DATA LEAKAGE DETECTION

[1]Era Kaushik,[2]Shweta Kumari,[3]Manan Jain , [4]Aaditya Katyan
[1]Student,[2]Student,[3]Student,[4]Student
[1]Department of Computer Science,
[1]Dr. Akhilesh Das Gupta Institute of Technology and Management, affiliated to GGSIPU
Delhi,India

Abstract: The paper demonstrates a mechanism for inferring the behavior of the user from encrypted wireless network activity. This mechanism also operates without any level of network access and without the need of breaking any encryptions. It also shows how an entirely passive, external observer can detect the information that is being transmitted over the network. The large number of applications that we install on our smartphones generates huge amount of network traffic patterns. The traffic that is generated by the user while using these applications also contains some characteristic traffic generated by various applications including their background activities or periodic updates or some specific information of particular applications. Although the encryption system present in various networks for transmitting the data prevents malicious intruders or eavesdropper from getting access to analyze the content of the data, the periodic traffic patterns generated by the applications leak side channel information such as data packet size, data transfer timing and the volume of the data. Since, wireless communications are broadcast in nature, various information that can be transmitted like data packet size, it's volume, frame size and the modulation scheme used get exposed. This information can be used and exploited by an intruder to passively attack or steal confidential user information. Such kind of problem cannot be avoided even though we encrypt the frame headers and the payloads of the data. With the long range wireless communications becoming more prevalent and increased commercial interest in tracking and analyzing publicly broadcast wireless data, this paper highlights the threat to users private activities.

Index Terms **- Wi-Fi; Mobile-Apps; Privacy; Security ;Data Protection; Information Inference; Data Leakage**

I. INTRODUCTION

A. Data Leakage

Accidental or unintentional distribution of private or sensitive data to an unauthorized entity is known as data leakage. When data stored in data centres is transmitted in an unauthorized and uncontrolled manner, data leakage occurs. Data leakage can lead to huge losses to the organization and also result in loss of trust of the people. Data leakage is said to have occurred when sensitive data is accessed by an unauthorized user. Sensitive data consists of intellectual property, financial data, personal information like contact information, personal credit-card data, and other confidential information. Data leakage is a very common and important issue these days. The number of incidents and the users affected by it are on a rise. The reason of exposure of confidential data or information might be malicious intent or an inadvertent mistake by an insider outsider, but it can seriously damage the reputation of an organization.

B. Scenario and Environment

WiFi data usage and communication is escalating and is now a basic need of an ever-increasing modern society; permeating through homes, business and almost everywhere. Sensitive data may include intellectual property (IP), financial information, patient information, personal credit-card data, and the information depending on the business and the industry. When these are leaked out it leaves the company unprotected and goes outside the jurisdiction of the corporation, this may put business in a vulnerable position. Once this data is no longer within the domain, then the company is at serious risk. Data leakage poses a serious issue for companies as the number of incidents and the cost to those experiencing them continue to increase.

A standard local area wireless network providing secure internet access for users of PCs and major handheld devices WiFi-enabled devices. With companies recognizing wireless broadcasts as an important source of data for the bigger firm, the commercial interest in wireless broadcasts has increased ever since making access points more common and popular. Our study concentrates on review of work done with the WiFi, the 'side-channel' information used to infer user activities is reviewed common to all modern IP networks so should generalize to 4G LTE or any other observable network communication.

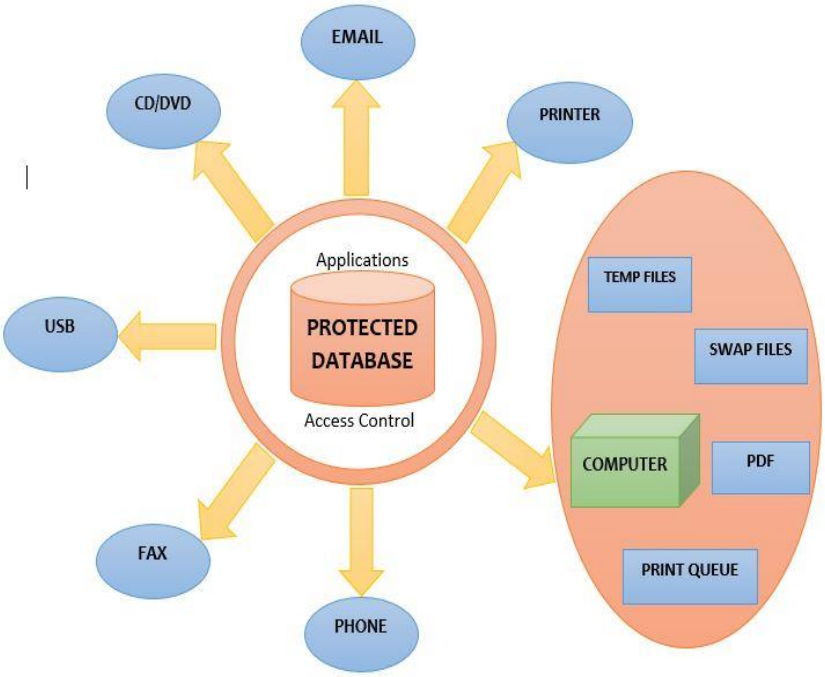

Figure1.1 Sensitive data storage

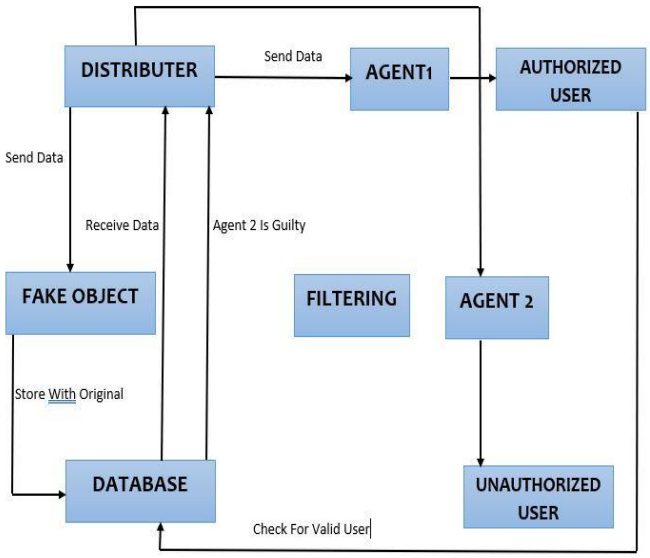

Figure1.2 Data Leakage

In order to show this, 34 popular and secure applications were selected, and the intended demographics of their users were found out. The next step was to collect the network data when users opened the applications. The usage of a particular application is determined by the network activity [1]. But because of encryption, only limited amount of information is available through side-channels. Histograms were constructed from frame size classification and inter-arrival time characteristics. These histograms detail the metrics distribution over a period of time.

This technique can recognize application which can be helpful in fingerprinting of different activities over the encrypted communication system. Mobile application is an appropriate and a vulnerable target due to their personal ties with other applications, availability in an openly ranked market with comparatively less diversity and the ease of collections. Although the methods shown in this paper use 802.11d WiFi standard, the processes should generalize to other protocols of wireless communications as well unless they are intentionally designed to withstand this type of analysis [1]. Especially for mobile applications, the measurements and methods used to carry out this analysis will also be present in long range protocols like 4G LTE network available in cell phones.

However, here we tried to review about how the combination of Direction, Frame size and Inter-arrival timing of the data packet makes user's personal and insightful information vulnerable. Any passive party within the wireless network range can cause information leakage. The observer here operates without any level of network access or authorization and does not attempt to break the encryptions in place. Skype traffic was also analyzed to infer user activities. Even when the Skype traffic is combined with different and more confusing traffic simultaneously, the detection ability still remains the same. This in turn makes it remote, passive, undetectable and inexpensive.

A closer look at Fig 1.3, shows a description so as how a mobile device gets connected to an Access Point (AP), a service provided by Internet Service Provider (ISP). The applications running on any device using internet may use the same connection to communicate with certain remote internet servers. The content of the application is provided through the information from these servers. One of the main advantages of being dependent on an internet connection is that its centralized architecture reduces the dependency over processing power and storage. These equipment's can thus be smaller, cheaper and also offer latest content or backups when connected. Also, even the applications that only offer unchanging limited subject matter use the internet connection to some extent when accessible. In our situation the access point offers an 802.11g Wi-Fi network which uses industry standard encryptions so that only the users who have access to the network access it.

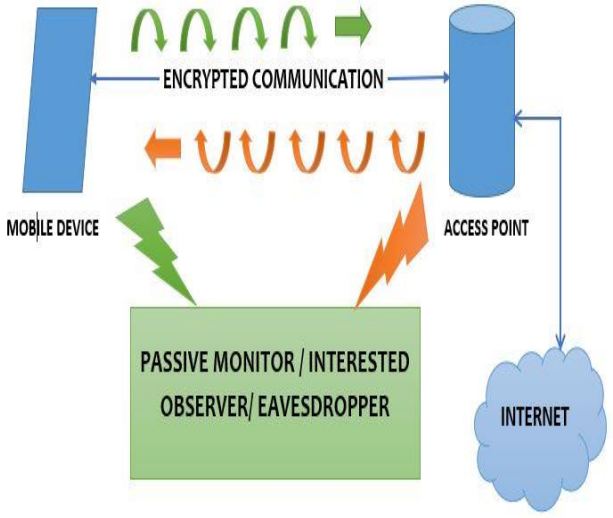

Fig 1.3 Observing encrypted mobile device WiFi traffic.

Since the process of observation is completely passive, the person observing it is completely unidentifiable by the user or the person operating a network. They are not essentially malicious or dangerous as they do not have any network authorization or security keys, nor do they ever try to break or find them.

II. RELATED WORK

Today, Wifi is widespread everywhere around the world. The broadcast nature of wireless networks utilizes different nearby applications to detect receiver's location and is a familiar phenomenon in various mobile devices. Though encryption technique is used for maintaining the confidentiality of data in wireless communications over wide area, the unbroken encryption can also affect the user's activities to large extent about which many people are still not aware. However great amount of effort is needed to carry out these kinds of analysis and therefore is not much prevalent.

In order to achieve compatibility, most of the network protocols follow OSI model. OSI model separates the protocols into different layers with respect to their responsibilities.[2] The data at lower layer protocols is encapsulated with the higher ones. The WiFi standard (IEEE 802.11) has been adapted from the wired Ethernet standard (IEEE 802.3) and redefines the Physical layer and Data Link layer of OSI model.[3]

To ensure confidentiality of data, WiFi networks must employ an encryption technique in order to prevent data from being read directly by an unauthorized user. The work below has not attempted to find or exploit flaws to break WiFi encryption.

A. PAPER I: What Do Your Mobile Apps Gossip About You

The paper discusses about how mobile device applications can inadvertently broadcast user's personal information via use of wireless networks despite encryption working perfectly as designed. It is demonstrated how usage of application can be tied to personal information through selection of personas. Personal information is assumed to be confidential by the users while using an encrypted network. However, here we illustrate how analysis of an encrypted traffic pattern can allow a distant observer to infer possibly sensitive data passively and untraceably without any network authorization.[4]

The availability of limited side channel data enables remote app detection. However, here we do not have the capability to read encoded WiFi data directly. Random Forest classifier is constructed using side-channel data measurements that are represented as histograms. It has the ability to precisely identify mobile applications from their generated encrypted traffic. This algorithm has a mean accuracy of ~99% within the training set for accurately identifying the applications. [5]

Finally, potential applications, methods to minimize such kind of data leaks and efforts required to demonstrate this phenomenon to users is discussed. This paper highlights a privacy vulnerability that is difficult to solve and cannot be minimized without making significant changes to the existing and subsequent generation of wireless communication protocols.

B. PAPER II: Inferring User Behavior, Encryption Irrelevant

In this paper we illustrate a method for inferring the behavior of the user from encrypted wireless network activity. Even without the necessity of breaking any encryption or having any level of network access this mechanism operates successfully. It demonstrates how a completely passive and remote observer can identify when a user is using Skype and when not. It is also illustrated that this detection capability remains even when Skype's traffic is combined with bewildering concurrent traffic such as Bittorrent.

The method demonstrated here challenges the presumption that secure cryptography means secure information and the approach of the mechanisms used may help guide in analyzing increasingly large volumes of encrypted data.

Based on the difference in the size of the frame and it's interarrival time, we were able to distinguish between the Skype and Bittorrent data packets. Bittorrent was chosen because of its high complexity i.e. to increase the overall difficulty.

C. PAPER III: Techniques and Countermeasures Of Website/Wireless Traffic Analysis And Fingerprinting

The behavior of a communication traffic can reveal some patterns such as packet size, packet direction, interarrival time that can expose user's identity, activities and private relations or connections. Even if the encryption algorithms are adopted precisely it is very difficult to conceal these kinds of information. Such situations are analyzed by traffic analyzers giving them an opportunity to infer characteristics of different users visiting a specific website or application that are running in the wireless networks. In order to preserve the privacy of the user, a defense mechanism and anonymous networks can be used to conceal the traffic patterns and features during communication.

This paper categorizes the traffic analysis into two domains, website domain and wireless domain and also analyzes the present traffic analysis technique and its countermeasures. A combination of different set of layers that illustrates the various stages of analysis technique is highlighted with a help of integrated traffic analysis process model. After that, the factors that can affect the accuracy of the fingerprinting mechanisms are studied to demonstrate how the change of such factors can affect the success results of fingerprinting. At last, we illustrated various possible challenges that needs to be considered when we are planning to implement and deploy it into real-world traffic analysis systems.

D. PAPER IV: Full Frame Encryption and Modulation Using Friendly Crypto Jam Scheme

The large number of applications that we install on our smartphones generates huge amount of network traffic patterns. The traffic that is generated by the user while using these applications also contains some characteristic traffic generated by various applications including their background activities or periodic updates or some specific information of particular applications. Although the encryption system presents in various networks for transmitting the data prevents malicious intruders or eavesdropper from getting access to analyze the content of the data, the periodic traffic patterns generated by the applications leak side channel information such as data packet size, data transfer timing and the volume of the data. Since the wireless communications are broadcast in nature various information that can be transmitted like data packet size, it's volume, frame size and the modulation scheme used get exposed. This information can be used and exploited by an intruder to passively attack or steal confidential user information. Such kind of problem cannot be avoided even though we encrypt the frame headers and the payloads of the data.
To get access to various transmission attributes of the transmitted data the eavesdropper can cut off unencrypted fields in Physical and MAC headers of data packets. By measuring the frame duration, the intruder can determine the packet size of the frame by estimating the data rate of specific frame. These intruders can find information about the MAC addresses of the source and destination machine, modulation scheme and transmission rate of payloads, directions of traffic, length or duration of frame etc. [6]

E. PAPER V: Mining Application on Analyzing Users' Interests from Twitter

It is difficult to provide social media users with posts that are analyzed from their interests efficiently. It makes it difficult to provide good quality and variety of posts to the users based on their interest. The ever-increasing use and reach of smartphones with an internet connection have enabled to analyze users interests from Twitter. Twitter has a huge user base and is used by a large number of people to share posts on a variety of topics and interests as tweets. Mining user's interests from twitter can increase a number of efficacies like advertisements, trending topics that can be analyzed according to user's interests and recommendation of posts.

For the same, this paper provides an android application which incorporates Web Services, Jsoup, JSON, Firebase Real-time Database and MVC. The application helps to select the posts which include huge images and text that are shown to users as a training set. These personalized posts can later be analyzed by the users themselves using suffix, array ds and artificial neural network. Under ANN, here backpropagation methodology has also been used that fires neurons as posts. Kosarju algorithm and palette lib is also used to help remove redundant posts while also retaining relevant posts with specific hashtags more efficiently and accurately.

III. METHODOLOGY

A. *WiFinspect - Random Classifier Application*

The first and foremost task was to select 34 apps which could easily be found and will be used for monitoring passively and cover a wide range of demographics. We are not interested in personal data sets which can be used for identification purposes. Even though, we are not able to infer the personal information of an individual, but we can deduce the data such as age groups, country etc. Information categorized as sensitive is data where additional security is required. The inference generated are only generalized. For instance, most of the people using car trading applications are male while there are few females using those applications as well.

After selecting the applications for studying, there is an issue with measuring app activity like a perspective of external observer. In order to generate more accurate and synchronized results, an external capture mechanism was developed to capture various app activity. A 'WiFinspect' [7] app was used for this purpose along with some other readily available software in order to collect data in large quantity with the help of above software used. It is important to note that the test was done with network used WPA2(with PSK).

Developing a classifier from the data sets is available. A rather subtle technique called "Random Forest Construction "with bigRF package in R [8][9]. After the development of the app, the process can analyze the usage the accuracy of which can be predicted by the process as a measure of Out-Of-bag (OOB) error. The three-step process of Decision tree construction [10].

This process to deduce the generic information does not make any attempt to break Wi-Fi security measures which is considered as most secured. Thus, these measures only allow us to gather data which is present in the header of the frame and access point (AP). Along with side-channels we can calculate:

Size of the frame (read starting from header)

- Inter-arrival frame time
- Directions of the frame
- Size of the frame

With the above information the measurement can be represented as value distributions, Histograms in this case helped to obtain the desired results.

B. *Inferring User Behavior Without Breaking the Encryption*

Most of the communication are done through wireless medium. Operating with Linux drivers with a standard 802.11b/g wireless network card which is in 'monitor mode', Kismet (a wireless packet 'sniffer') is used to perform these observation tasks. This stores the observed transmissions as de facto standard PCAP (Packet Capture) files for analysis.
This stores the observed transmissions as de facto standard PCAP (Packet Capture) files for analysis. We specifically measure the size of the frame and its direction (Between the MAC addresses of the receiver and sender) can be read from 802.11. Additionally, no. of frames and frame time slot can be logged.

To measure the information a series of code were written using sikuli. To imitate actions of various application under study. This behavior is then logged using the data traffic observer which was described previously. At first, behaviors were observed using a private network. A computer was connected to a mobile device (3G). This connection between the computer and the mobile device was made using 802.11g in order to collect Skype [11] traffic. Various other tools such as Hyenae and target devices were used and retained for analysis.

With the measurement and collection of the information we needed to characterize the presence of Skype voice traffic. Only some of the data have the relevant frames and Wireshark Suite can be used to access important packets. Once all the data is filtered. Rest of the process is as follows:

- To generate sliding time windows that is to group different frames received within a given time duration. These windows are used to generate sliding windows at an interval less than the capture time duration of the grouped frames. Thus, a single frame will be present inside some of the overlapping windows. This prevents the loss of data which can occur if separate windows were used
- To generate metric distributions for each of these windows, generating a distribution characterizing each metric over the given time duration.

Statistical picture measures were done to form Distribution over 5s (50 ms distribution of time) Window period. Each of the distribution includes data packets from Skype traffic and is assigned a metric's score for given time window (Sw) calculated as follows. The metric score is calculated by taking pre-calculated expected distribution (ei) multiplying with the numbers logged over that time duration by the number actually observed over that time window (oiw) [12].

$$Sw = \sum_{i=1}^{n} \quad ei\ oiw$$

The numeric scores that would be generated is used to measure Skype-like activity showing the metric scores over time. A dull background indicates when a Skype voice call was in progress. Inter-arrival scores are scaled by a factor of 100 for illustration only. FSize is considered as the strongest indicator of Skype activity. Metric Thresholds can also be determined during a Skype voice call by comparing the plots for each metric.

C. Frame Encryption by Using Friendly CryptoJam Scheme

Friendly Crypto Jam(FCJ) can be defined as when the information along with jamming signals are intermixed after processing digital modulation phase but before pushing it to the channel. The encryption for modulation used for the process preserves the reflected binary (RB) also known just as the Gray Code of the encrypted symbols on the original mapping. In contrast to conventional (digital domain) encryption, the encryption in FCJ is modulation-aware [14].

Frame detection, FO and CSI detection can be done with the help of physical header since each physical header comes before the preamble. Several repetitions of a publicly known pattern can be obtained with the help of preamble. Detecting the receival of minimum of two portions of the preamble is required for the estimation of FO. By considering repetitions in the arrived signal as a reference and comparing it with another repetition that is T seconds away, a frame can be decoded. CSI estimation can be performed by comparing the known patterns in preamble with its received value.

An adversary can intercept the preamble, physical and MAC headers since all of the headers are sent during the transfer of information. Usually, the preamble and the Physical layer header are transmitted at the minimum supported rate2 while the transmission rate for the frame payload (including MAC header) is adjusted based upon the conditions of the channel. It results in variation of frame time (in seconds) for similar payload. In 802.11n, the 'Modulation and Coding Scheme' field represents both the coding rate and the modulation scheme, like the 'rate' field in 802.11a. All 802.11 variants specify a 'length' field, which represents the payload size in octets (for 11a/n) or in milliseconds (for 11b) [15].

Any kind of rate-based SCI classification can be prevented if the scheme which is for the modulation purpose of different frame payloads always investigate it. This can be achieved by embedding the payload's original modulation symbols in the constellation map of the highest-order modulation scheme.

D. Website Fingerprinting Technique

A number of experiments have been conducted in different papers to show that an improvement is still needed in existing traffic analysis countermeasures. One of the first attempts in attacking website traffic to recognize users' identities was introduced by Wagner and Schneier.[16] The random padding method used in SSL protocol works in block cipher modes only, which allows cipher text to reveal plain text lengths. This shortcoming has been used effectively by Wagner and Schneier to infer the identities of the visited website.[16] As an outcome, it was suggested to adhere to random length padding to all cipher modes to resolve the problems SSL.

Sun et al also concluded that a considerable amount of information can be revealed from the encrypted communication.[17] Another significant work was done by Bissias et al [18] , they kept on collecting the data packets for a year with the use of Firefox linked with OpenSSH tunnel. After selecting two features of each HTTP trace, namely the inter-arrival time and the size of packets, the similarity of the two traces of different time gap packets was recognized and a cross-correlation metric was generated. The obtained results were quite good despite of low success rates.

Liberatore and Levine [19] introduced two new techniques for fingerprinting websites traffic using Naive Bayes (NB) classifier along with a density estimation. The experiments done by them showed that there is a need to secure OpenSSH protocol in order to block the threats from attackers so that they cannot analyze the traffic passing through it.Shi et al. [20] proposed a technique for website fingerprinting which can be used to analyze traffic communication over Tor. In their attack, they divided the different direction of data packets into several intervals and then converted them into vectors. Using, the Cosine Similarity formula, similarities are calculated between observed vectors and well-known fingerprints. The generated results using above technique was then evaluated theoretically and practically in order to show how effective their technique is in generating vulnerability to user's anonymity as compared to Tor.

Cai et al. [21] proposed a new fingerprinting method for attacking websites in which data was collected by capturing packets generated from visiting websites through Firefox linked with Tor using Tshark. Each of the website was visited with a number of times in order to collect a large number of trace packets. The Damerau-Levenshtein Distance algorithm was deployed for identifying the visited web pages. They were able to find the sizes, ordering, directions and other useful information of the transmitted packets with the help of the classifier. In addition, Hidden Markov Model was also used to identify. These experiments demonstrated there is a need to improve the current defense schemes against traffic analysis over Tor. As an outcome, Congestion-Sensitive BuFLO which was considered an improved defense scheme was introduced by them. As shown in their work, it provides better security systems than its predecessor.

Based on the attack proposed by Cai et al., Wang and Goldberg introduced Combined Optimal String Alignment Distance (OSAD) attack which wan an improved version of their so-called Cai's fingerprinting attack. This new metric was proposed to identify the similarities between two packets of traffic data collected from Firefox with the help of Tor. By eliminating the SENDME packets from Tor cells, higher accuracy can be achieved. These experiments were carried over real world accessible websites, like

Panchenko[22], and could result in better outcome.By combining the attacks proposed by Cai et al and Wang and Goldberg a new fingerprinting attack was introduced. The resulting technique got a higher success rate than previous ones in terms of accuracy and processing time. These experiments were conducted in real world scenarios over large open world data gathered through Tor.

E. Wireless Fingerprinting Technique

Ever since the introduction of wireless communication, WLANs are an ever-present part of our society. With the increase in the use of WLAN, it has become easy for adversaries to snoop and analyze the user traffic over WiFi links even though the encryption is deployed perfectly in the communication system. By analyzing and observing the specific patterns in various packets about the characteristics and behavior of the transmitted packet the intruders can infer user's online and local activities. It is difficult to extract traffic features when a user is running more than one application simultaneously due to interference of applications packet between each other.

The encoded VoIP calls were identified by a new method which was introduced by Wright et al. [23]The phrases spoken during call conversations can be identified through the packet sizes of the encrypted VoIP signals. This method adopted a mechanism called Hidden Markov Models for the purpose of recognizing 122 target sentences that are trained using the TIMIT training data. An accuracy of more than 90% was achieved using this method.

A hybrid mechanism was proposed for classifying network traffic by Tavallaee et al.[24]Few machine learning techniques and signature-based method was applied in this mechanism. A hierarchical classification based system was introduced by Zhang et al. [25] to infer online activities of various users. In order to analyze the identity of users that uses encrypted wireless communication, a technique was proposed by Atkinson et al.[26]The experiments were carried out by running users' apps and then gathering data by simulating various user's actions. These packets are first gathered and then identified by a specific set of metrices such as the size of the packet and inter-arrival time. At last, the traffic is extracted by applying aggregate normalized distributions over each metric. Multiple accuracies were gained based on the number of packets collected.

.

F. Neural Networks Back Propagation

Backward Propagation of Errors abbreviated as backpropagation, is a common method of training Artificial Neural Networks. It is used in combination with an optimization method. It is used for calculating the gradient of a loss function with respect to the weights present in the network. Attempting to minimize the loss function, a gradient is used with the optimization method which further updates the weights. Backward Propagation requires a desired input known with the method for each input value to compute the loss function gradient. Therefore, it is usually considered a supervised learning method. It can also be used in some unsupervised methods like auto-encoders. Backpropagation is used for proper arrangement of user's posts.

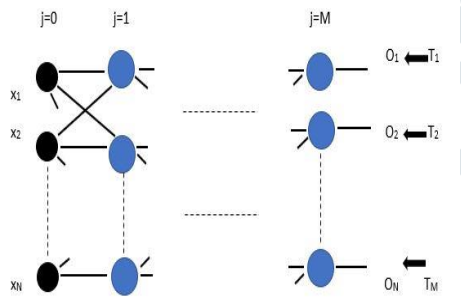

Fig 1.4 Neural Network

G. Kosaraju Algorithm

It is a depth first search algorithm. A graph is said to be a strongly connected directed graph only if there is a path between all the pairs of vertices. Strongly Connected Component of a directed graph is a maximal strongly connected sub-graph.[27] DFS is done twice here. A single tree is produced by a DFS of a graph if all vertices are reachable from the DFS starting point. In the next step, the graph is reversed. Here the algorithm is used for removal of redundant posts. The vertices are marked as not visited and are filled in a stack according to their finish time in the first DFS. In second DFS, the vertices are marked as non-visited. All the vertices are processed in an order as defined in the stack. Vertices are then popped from the stack and the SCC[28] of the popped vertex is printed out. In the next step, the direction of the pointing to SCCs are reversed.

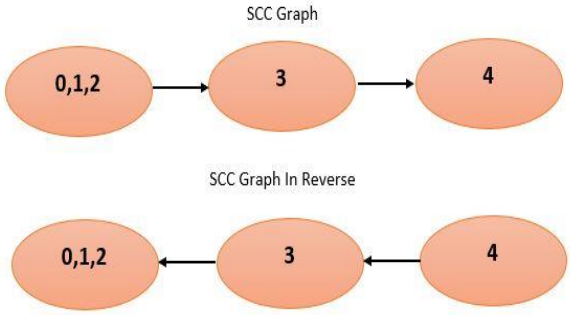

Fig 1.4 SCC Graph

IV. CONCLUSION

This paper highlights that even though various Wifi encryption techniques works perfect as intended, sensitive and personal information of the users are still at the risk of getting broadcasted whether intentionally or unintentionally. A distant and untraceable detection mechanism is deployed by the intruders in order to gather user's private information by observing the network activity of the encrypted app.

It is possible to infer and detect user behavior to an extent by an external and entirely passive user, despite providing only limited data and using correct encryption techniques. It is difficult to design efficient protocols that resist the analysis. Side effects of efficient networking such as variable frame size and interarrival times cannot be ignored.

The methods and the corrective measures used here for analyzing traffic and fingerprinting have also been reviewed here. They have been divided into two domains namely websites and wireless. In here, we have shown how existing defense schemes, encryption protocols, and anonymity networks do their best to conceal user's personal and private information. Here, it has also been demonstrated how malicious users utilize all possibilities to analyze and study user's application traffic for the sake of disclosing their identities.

The prevention of transmission attributes like unencrypted PHY/MAC header fields and the payload's modulation scheme is a difficult task. Friendly CryptoJam (FCJ) effectively protect the confidentiality of lower-layer fields and prevent SCI- based traffic classification, rate-adaptation, plaintext, dictionary, modulation detection, and device-based tracking attacks [6]. With wireless communication becoming increasingly common, and commercial companies becoming keener in tracing and analyzing publicly transmitted wireless data, this paper therefore presents an important and demonstrable threat to user's privacy.

V. ACKNOWLEDGEMENT

This research was supported by Mr. Pranav Shrivastava, Assistant Professor (Department of CSE, ADGITM). We thank him for his constant support who provided insight and expertise that greatly assisted the research, although he may not agree with all of the interpretations/conclusions of this paper.

REFERENCES

[1] Your WiFi is leaking By:JohnS.Atkinson, John E.Mitchell, Miguel Rio, George Matich
[2] H. Zimmermann, OSI reference model-the ISO model of architecture for open systems interconnection, IEEE Trans. Commun. 28 (4) (1980) 425–432.
[3] IEEE-SA, Wireless LAN Medium Access Control, MAC, and Physical Layer, PHY, Specification, IEEE Standards Authority, 2007.
[4] N. Lawson, Side-channel attacks on cryptographic software, IEEE Secur. Privacy Mag. 7 (6) (2009) 65–68.
[5] A. Lim, L. Breiman, A. Cutler, bigrf: Big Random Forests: Classification and Regression Forests for Large Data Sets, 2013.
[6] Mrs. G. Sasikala, Mrs. D.Kavitha Full Frame Encryption and Modulation Using Friendly CryptoJam Scheme
[7] A. Hadjittofis, WiFinespect Play Store Page, 2014. https://play.google.com/store/apps/details?id=uk.co.optician cms.wifiprobe.
[8] L. Breiman, Random Forests, Mach. Learn. 45 (1) (2001) 5–32.
[9] L. Breiman, Random Forests, 2013. http://www.stat.berkeley.edu/~breiman/Random Forests/cc_ home.htm
[10] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2013.
[11] S. A. Baset and H. G. Schulzrinne, "An analysis of the skype peer-to-peer internet telephony protocol," in INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings, Apr. 2006, pp. 1– 11.
[12] F. Zhang, W. He, X. Liu, and P. G. Bridges, "Inferring users' online activities through traffic analysis," in Proceedings of the fourth ACM conference on Wireless network security, ser. WiSec '11. New York, NY, USA: ACM, 2011, pp. 59–70.

[13] A. White, A. Matthews, K. Snow, and F. Monrose, "Phonotactic reconstruction of encrypted VoIP conversations: Hookt on fon-iks," in Security and Privacy (SP), 2011 IEEE Symposium on, may 2011, pp. 3 –18.
[14] C. Cardoso, A. R. Castro, and A. Klautau, "An efficient FPGA IP core for automatic modulation classification," IEEE Embedded Syst. Lett., vol. 5, no. 3, pp. 42–45, Sep. 2013.
[15] IEEE-SA, Wireless LAN Medium Access Control, MAC, and Physical Layer, PHY, Specification, IEEE Standards Authority, 2007.
[16] Wagner, D., Schneier, B.: Analysis of the ssl 3.0 protocol. In: The Second USENIX Workshop on Electronic Commerce Proceedings, pp. 29–40 (1996)
[17] Sun, Q., Simon, D.R., Wang, Y.M., Russell, W., Padmanabhan, V.N., Qiu, L.: Statistical identification of en- crypted web browsing traffic. In: Security and Privacy, 2002. Proceedings. 2002 IEEE Symposium on, pp. 19–30. IEEE (2002) Sun, Q., Simon, D.R., Wang, Y.M., Russell, W., Padmanabhan, V.N., Qiu, L.: Statistical identification of encrypted web browsing traffic. In: Security and Privacy, 2002. Proceedings. 2002 IEEE Symposium on, pp. 19–30. IEEE (2002).
[18] Bissias, G.D., Liberatore, M., Jensen, D., Levine, B.N.: Privacy vulnerabilities in encrypted http streams. In: Privacy Enhancing Technologies, pp. 1–11. Springer (2006)
[19] Liberatore, M., Levine, B.N.: Inferring the source of encrypted http connections. In: Proceedings of the 13th ACM conference on Computer and communications security, pp. 255–263. ACM (2006)
[20] Shi, Y., Matsuura, K.: Fingerprinting attack on the tor anonymity system. In: Information and Communications Security, pp. 425–438. Springer (2009)
[21] Cai, X., Zhang, X.C., Joshi, B., Johnson, R.: Touching from a distance: Website fingerprinting attacks and defenses. In: Proceedings of the 2012 ACM conference on Computer and communications security, pp. 605–616. ACM (2012)
[22] Panchenko, A., Niessen, L., Zinnen, A., Engel, T.: Website fingerprinting in onion routing based anonymization networks. In: Proceedings of the 10th annual ACM workshop on Privacy in the electronic society, pp. 103–114. ACM (2011)
[23] Wright, C.V., Ballard, L., Coull, S.E., Monrose, F., Masson, G.M.: Spot me if you can: Uncovering spoken phrases in encrypted voip conversations. In: Security and Privacy, 2008. SP 2008. IEEE Symposium on, pp. 35–49. IEEE (2008)
[24] Tavallaee, M., Lu, W., Ghorbani, A.A.: Online classification of network flows. In: Communication Networks and Services Research Conference, CNSR'09. Seventh Annual, pp. 78–85. IEEE (2009)
[25] Zhang, F., He, W., Liu, X., Bridges, P.G.: Inferring users'online activities through traffic analysis. In: Proceedings of the fourth ACM conference on Wireless network security, pp. 59–70. ACM (2011)
[26] Atkinson, J., Adetoye, O., Rio, M., Mitchell, J., Matich, G.: Your wifi is leaking: Inferring user behaviour, encryption irrelevant. In: Wireless Communications and Networking Conference, pp. 1097–1102. IEEE (2013)
[27] Kim, J., & Hastak, M. (2018). Social network analysis. International Journal of Information Management: The Journal for Information Professionals, 38(1), 86-96.
[28] Alshomrani, S., & Iqbal, G. (2012). Analysis of Strongly Connected Components (SCC) Using Dynamic Graph Representation. International Journal of Computer Science Issues, 9(4)