International Journal of Recent Trends in Engineering & Research
ISSN (ONLINE) : 2455 – 1457
IMPACT FACTOR : 4.101

Full Frame Encryption and Modulation Using Friendly CryptoJam Scheme

Mrs. G.Sasikala[1], Mrs. D. Kavitha[2]
[1]*Assitant Professor, Department of Computer Science, Adhiparasakthi college of Arts and Science, Kalavai*
[2]*Research Scholar, Department of Computer Science, Adhiparasakthi college of Arts and Science, Kalavai*

Abstract: The overall network traffic patterns generated by today's smart phones result from the typically large and diverse set of installed applications. In addition to the traffic generated by the user, most applications generate characteristic traffic from their background activities, such as periodic update requests or server synchronisation. Although the encryption of transmitted data in 4G networks prevents an eavesdropper from analysing the content, periodic traffic patterns leak side-channel information like timing and data volume. The broadcast nature of wireless communications exposes various transmission attributes, such as the packet size, inter-packet times, and the modulation scheme. These attributes can be exploited by an adversary to launch passive (e.g., traffic analysis) or selective jamming attacks. This security problem is present even when frame headers and payloads can be encrypted. For example, by determining the modulation scheme, the attacker can estimate the data rate, and hence the payload size. In this paper, we propose Friendly CryptoJam (FCJ), a scheme that decorrelates the payload's modulation scheme from other transmission attributes by embedding information symbols into the constellation map of the highest-order modulation scheme supported by the system (a concept we refer to as indistinguishable modulation unification). Such unification is done using the least-complex trellis-coded modulation schemes, which are combined with a secret pseudo-random sequence in FCJ to conceal the rate-dependent pattern imposed by the code. It also preserves the bit error rate performance of the payload's original modulation scheme. At the same time, modulated symbols are encrypted to hide PHY-/MAC layer fields. To identify the Tx and synchronously generate the secret sequence at the Tx and Rx, an efficient identifier embedding technique based on Barker sequences is proposed, which exploits the structure of the preamble and overlays a frame-specific identifier on it. We study the implications of the scheme on PHY-layer functions through simulations and test bed experiments.

I. INTRODUCTION

USING commodity radio, unauthorized parties can easily eavesdrop on wireless transmissions. Although advanced encryption algorithms like AES can be applied to ensure data confidentiality, parts of the frame (e.g., PHY/MAC headers) must be transmitted in the clear for correct protocol operation and device identification. For example, 802.11i, the primary security amendment of 802.11, provides confidentiality only for the MAC-layer payload. Even if we hypothetically encrypt the entire PHY frame, the transmission is not completely immune to eavesdropping. An adversary can still fingerprint encrypted traffic through analyzing its side-channel information (SCI). It refers to statistical traffic features, such as packet size distribution, traffic volume, and inter-packet time sequence. These statistical features can be obtained by estimating and correlating leaked transmission attributes, including frame duration, the modulation scheme, traffic directionality (uplink/downlink), and inter-packet times. Traffic fingerprints can be used to breach user privacy by tracking her or discerning her identity, activity, and interests. For example, by eavesdropping on 802.11 WLAN traffic for only 5 seconds, an adversary (Eve) can determine the type of user activities with 80% accuracy. The sizes (in bytes) and direction of packets exchanged between a mobile user and an access point may reveal what phrase the user is searching for in a search engine, and identify

the browsed page or the language used in an encrypted instant messaging application. SCI can also facilitate geographically tracking the user by identifying her particular smart phone among many possible devices. By analyzing transmission attributes, Eve can further learn the type or stage of a communication, and launch selective jamming attacks. For example, Noubir et al. demonstrated a reactive jammer that can significantly hammer the network throughput by intercepting the rate field in the header and accordingly decide whether to jam the rest of the frame. If a packet is not correctly decoded as a result of jamming, the transmitter (Alice) mistakenly assumes a poor channel and lowers the rate when retransmitting the same packet, wasting network resources.

To obtain transmission attributes, Eve can intercept unencrypted fields in the PHY and MAC headers. These fields include the source/destination MAC addresses, payload transmission rate and modulation scheme, frame length/duration, traffic directionality, number of MIMO streams, and others. Eve can also perform low-level RF analysis to obtain SCI even when PHY/MAC headers are encrypted, a threat that has not been well-studied in the literature. Consider, for example, the detection of the payload's modulation scheme of an entirely encrypted PHY frame. Using an off-the-shelf device such as a signal analyzer or a dedicated device equipped with an FPGA, one can detect the modulation scheme, and accordingly estimate the payload's data rate. The same device can also measure the frame duration and determine the packet size based on the estimated data rate.

II. RELATED WORK

2.1 Risk Associated with Network Security

- Attackers can "eavesdrop" on unencrypted data traveling over a network, not only impacting privacy but potentially opening the potential to modify or substitute data as a way to stage more sophisticated attacks.
- Because industry mandates often require protection for data in motion, organizations that do not implement this protection risk fines, embarrassing data breach disclosure statements, and resulting damage to their reputation.
- Depending on the application, encryption capabilities embedded in routers and switches may not offer the combination of security and performance you need.

2.2 Network Encryption: Thales e-Security Solutions

Using standalone network encryption platforms from Thales e-Security, you can deploy proven solutions to maximize confidence that your sensitive, high-value data will not be compromised during transport. Datacryptor network encryption platforms offer increased levels of protection over both unencrypted data transport and basic encryption capabilities embedded in routers and switches.
The Datacryptor family of network encryption platforms is designed to offer the widest range of support for different network types, encryption protocols, and certification levels—while delivering state-of-the-art throughput and latency. This ideal combination of security, performance, and deployment flexibility is essential for organizations and service providers wishing to secure point-to-point and multipoint networks where latency, bandwidth utilization, and powerful separation of duties are of utmost importance.

- Institutions or organizations with geographically distributed offices interconnect by virtual private networks.
- Organizations with mirrored or replicated data centers using high-speed wide area network (WAN) connections.
- Organizations using microwave or radio based campus networks.
- Service providers wishing to provide premium, encrypted data networking services.
- Governments wishing to support national algorithms or key management practices for high assurance restricted networks.

1.3 Existing Methods and their limitations

Techniques to prevent SCI leakage can be divided into three categories: SCI obfuscation at upper layers, rate hiding in our initial work ,and in more recent scheme, and eavesdropper deafening at the PHY layer. Upper-layer SCI obfuscation techniques aim at invalidating SCI, usually at the cost of traffic overhead. For example, packet padding can be used to alter the traffic statistics. However, the overhead can be as high as 400% . Traffic reshaping is a MAC layer technique that involves configuring several virtual interfaces with different MAC addresses for the same device so as to create different traffic patterns on each interface. This prevents Eve from associating all the packets with the same sender. Similarly, the sender and receiver can agree on a set of confidential time-rolling MAC addresses. However, these identifier concealment techniques cannot hide certain attributes, including the modulation scheme.

To hide the payload's modulation scheme, Conceal and Boost Modulation (CBM) was proposed in , whereby convolutional codes based on a Generalization of Trellis Coded Modulation (GTCM) are used, combined with a cryptographic interleaving mechanism to conceal the rate information of the underlying code. GTCM directly encodes the symbols of any modulation scheme into the highest-order modulation scheme. A symmetric-key scheme was also proposed to encrypt the PHY-layer header. While CBM can achieve up to 8 dB asymptotic coding gain (in idealized simulation scenarios), it does not address the issue of sender identification and the decryption of the PHY-layer header. Moreover, the complexity of GTCM codes, interleaving, and expensive symmetric-key encryption result in a large decoding delay at Bob. Due to acute susceptibility of denser modulation schemes to phase offset, GTCM codes also suffer significantly from inaccurate FO estimation, reducing its coding gain.

PHY-layer eavesdropper deafening techniques include friendly jamming (FJ). In this method, Eve's channel is degraded without impacting the channel quality at Bob. This is done using (distributed) MIMO techniques to transmit a jamming signal that is harmless (friendly) to Bob. However, four fundamental issues limit the practicality of this approach. First, if Eve is equipped with multiple antennas too, she can cancel out a transmitter-based FJ signal. For example, Schulz et al. exploited a known part of Alice's signal (e.g., frame preamble) to estimate the precoding matrix used in generating the FJ signal and then eliminate it from the received signal at Eve. This matrix is supposed to be secret and unique, as it depends on the channel state.

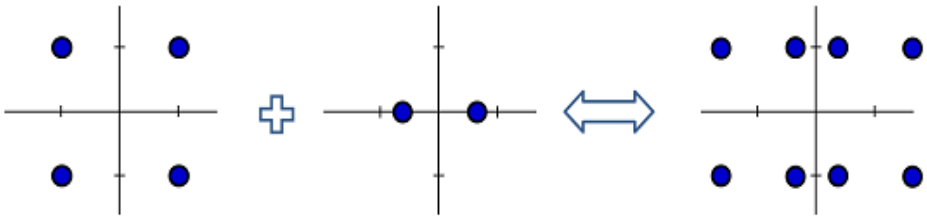


Figure 2.1 Combining QPSK-modulated and BPSK-modulated signals with different powers results in a 8-symbol constellation map.

Information (CSI) for the Alice-Bob channel, i.e., it represents a signature of the Alice-Bob channel. This known-plaintext attack can thwart any deafening scheme that relies on signal pre filtering (precoding) at Alice. Furthermore, the uniqueness of the Alice-Bob CSI has been shown not to be true in the presence of strong LOS component. Specifically, a few adversaries located several (~ 18) wavelengths away from Bob can cooperatively reconstruct Alice-Bob channel's signature.

Second, FJ requires additional transmission power and antenna(s), which come at the expense of throughput reduction for the information signal. The FJ power may need to be even higher than the

information signal power to achieve nonzero secrecy capacity. Moreover, Alice may not have sufficient number of antennas (degrees of freedom) to apply FJ.

Third, transmitter- and receiver-based FJ are still vulnerable to cross-correlation attacks on (unencrypted) semi-static header fields, the fields that can take one of a few valid values. Eve can detect the start of a frame, even if it is combined with a jamming signal. By knowing where each field is supposed to start in the underlying header format, Eve can pinpoint a targeted field in the received signal. Because of FJ, Eve may not be able to successfully decode the field value. However, she can correlate the sequence of modulated symbols of each possible value with the received signal and guess the true field value. In general, this cross-correlation attack can be formulated as a composite hypothesis testing.

Last but not least, FJ cannot effectively hide the modulation scheme and frame duration. If the jamming signal is random, Eve can employ detection techniques for low SNR to detect the modulation scheme. Even if the FJ signal takes the form of a digitally modulated signal (as opposed to random noise), Eve may still detect the modulation scheme of the payload by analyzing the order and constellation map of the received superposition. The superposition of the I and Q components of the complex symbols that belong to the two signals results in a modulation scheme whose order and con stellation depend on the original schemes and the respective received powers. For example, the constellation map resulting from the superposition of two signals, one modulated with QPSK and the other with BPSK, can disclose the constituent modulation schemes.

III. OVERVIEW OF FRIENDLY CRYPTOJAM

To address the aforementioned limitations, we propose Friendly CryptoJam (FCJ), a form of friendly jamming but with the information and jamming signals intermixed right after the digital modulation phase and before the frame is transmitted over the air. Our intermixing method makes FCJ a form of modulation-level encryption (for the whole frame) and also a form of modulation obfuscation (for the PHY-layer payload). To generate a secret FJ sequence, Alice exploits an unpredictable sender identifier as a seed, which is then embedded in the frame preamble (i.e., a PHY-layer identifier). This way, Bob can identify the sender for key lookup and synchronize with Alice in generating the same FJ sequence. Hereafter, we call this secret sequence as “FJ traffic”. This identifier is independent of the link features and is robust to known plaintext attacks. Compared to our initial proposal of FCJ, the modulation encryption in this paper preserves the Gray coding structure of the encrypted symbols on the original constellation map. In contrast to conventional (digital domain) encryption, the encryption in FCJ is modulation-aware.

Using parts of the same FJ traffic, encrypted symbols of the payload are then simultaneously coded and mapped (upgraded) to the constellation map of the highest-order (target) modulation scheme supported by the system. We develop a modulation coding that prevents the disclosure of the payload’s original modulation scheme, i.e., it provides indistinguishable modulation unification. In contrast to the uncoded modulation unification in the initial design and variable-rate coding for upgrading different modulation schemes to same target modulation scheme in CBM, the novel mapping proposed in this paper employs only two minimal trellis-coded modulation (TCM) codes with constraint length $\leq$ 2and constant rate (irrespective of the target modulation scheme). These codes are inseparably combined with the FJ traffic so as to continuously move the low-density coded symbols on the target constellation map while maintaining the BER. This way, we hide both the true modulation scheme and the rate- dependent structure imposed by the underlying TCM code without symbol interleaving. Compare to, FCJ also enjoys lower coding complexity, decoding delay, and susceptibility to FO, but at the expense of lower coding gain. We further provide an analytical study

of the impact of uncompensated FO. In contrast to classic FJ techniques, a single antenna is sufficient to transmit both the information and FJ signals.

One important challenge in designing FCJ is how to modify the FJ traffic on a per-frame basis. Not changing the FJ traffic during a session opens the door for a dictionary attack against semi-static header fields. Furthermore, relying on a preshared secret sequence for the FJ traffic makes the design prone to synchronization errors. To ensure consistency in the generation of FJ traffic at Alice and Bob, Alice conveys a frame-specific seed (e.g., frame and sender ID) whose modulated value is superposed onto the known frame preamble. Together with the session key, this seed is fed into an appropriately selected pseudo-random number generator (PRNG) to generate the secret FJ traffic. The seed is also used for sender identification at PHY layer. Superimposing the seed with the preamble, however, may degrade the preamble's crucial functions (including frame detection). To mitigate that, we exploit the low cross correlation property of cyclically rotated Barker sequences to construct a seed-bearing signal in 802.11b systems.

TABLE I

DSSS SIGNAL SPREADING BASED ON AN 11-CHIP BARKER SEQUENCE FOR DBPSK MODULATION (802.11b)

Input	Sequence
0	+1, −1, +1, +1, −1, +1, +1, +1, −1, −1, −1
1	−1, +1, −1, −1, +1, −1, −1, −1, +1, +1, +1

3.1 Frame Detection and FO Estimation

Each PHY header is preceded by a preamble, which is used for frame detection, FO and CSI estimation. 802.11b systems exploit a scrambled version of a 128-bit all-one preamble that is spread using an 11-chip Barker sequence (see Table I). For a Barker sequence of length N, its autocorrelation function at lag k, denoted by $A(k)$, is very low at non-zero lags (orthogonality property). This can be exploited for frame detection and timing. Formally

$A(k) = \left|\sum_{j=1}^{N-k} b_j b_{j+k}\right| \leq 1, 1 \leq k < N$ ---(1)

where $b = \{b_1 b_2 \dots b_N\}$ is a Barker sequence. The receiver correlates this known sequence with the received sam ple sequence $r = \{r_1 r_2 \dots\}$ and computes the square of the cross- correlation value, denoted by $R(b, n)$:

$R(b,n) = \left|\sum_{j=1}^{N} b_j^* r_{j+n-1}\right|^2$ ----------(2)

R(b, n) is expected to peak when the nth sample of r marks the beginning of one of the transmitted Barker sequences. To improve the detection accuracy, b is replaced with a series of identical Barker sequences, one sequence per preamble bit.

The preamble consists of several repetitions of a publicly known pattern. FO estimation involves detecting the arrival of at least two identical portions of the preamble. 1 An FO in the amount of $\delta f Hz$ creates a time-varying phase displacement $(t) = 2\pi\delta f t$. To decode a frame, Bob estimates δ f by taking one of the repetitions in the received signal as a reference and comparing it with another repetition that is T seconds away. Specifically, Bob subtracts the phases of any pair of identical samples to find $\phi(T)$. Because of noise, usually there will be a

TABLE II
LIST OF KEY NOTATIONS

Notation	Definition
$\mathcal{M}_i$	Payload modulation scheme $i, i = 1, \ldots, M$
δ_f	The amount of frequency offset
$\varphi(t)$	The phase offset corresponding to δ_f after t seconds
$\mathbf{j}$	FJ traffic (j: Decimal value of $\log_2 \lvert\mathcal{M}_M\rvert/\lvert\mathcal{M}_i\rvert$ bits)
$\mathcal{U}_j$	Set of $\lvert\mathcal{M}_i\rvert$ elements of $\mathcal{M}_M$ corresponding to FJ bits j
$\mathcal{F}_{\mathbf{j}}(\mathcal{M}_i)$	The static mapping from $\mathcal{M}_i$ into $\mathcal{M}_M$ based on $\mathbf{j}$
$\mathcal{E}_{\mathbf{j}}(\mathcal{M}_i)$	Modulation encryption function based on $\mathbf{j}$
$\gamma_i(M)$	The gain of uncoded mapping from $\mathcal{M}_i$ into $\mathcal{M}_M$
$\gamma_i^{(q)}(M)$	The asymptotic coding gain when embedding $\mathcal{M}_i$ in $\mathcal{M}_M$ using the minimal q-state TCM code

residual FO estimation error even after averaging over several of such identical pairs. Depending on the frame duration, the residual FO may move a received symbol to a wrong region on the constellation map, causing a demodulation error. After compensating for FO, Bob compares the known pattern in the preamble with its received value to estimate the CSI.

3.2 Detection of Lower-Layer Fields

The preamble, PHY, and MAC headers are all transmitted in the clear, allowing an adversary to intercept them. Typically, the preamble and the PHY header are transmitted at the lowest supported rate2 while the transmission rate for the frame payload (including MAC header) is adjusted based on channel conditions, resulting in different frame durations (in seconds) for the same payload. Many standards, including 802.11 variants, specify the frame length and payload's transmission rate in the PHY header. For example, in 802.11b/g, the data rate and the modulation scheme are specified in the 'Signal' and 'Service' fields, respectively. In 802.11n, the 'Modulation and Coding Scheme' field represents both the coding rate and the modulation scheme, similar to 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).

The payload's size and transmission rate may also be determined by detecting the payload's modulation scheme and combining that with the frame duration to compute the payload size. A modulation scheme is usually associated with two or three data rates of different code rates. For example, in 802.11a, 16-QAM is used for data rates 24 and 36 Mbps. Hence, by determining the modulation scheme, it is rather easy for Eve to correctly guess the data rate.

3.3 Modulation Unification

In this section, we introduce a method for indistinguishably unifying different modulation schemes using FJ traffic. For now, we assume that the FJ sequence is already available at both Alice and Bob and Bob can decrypt the PHY header and obtain the true modulation scheme.

3.4 Uncoded Modulation Unification

To prevent any rate-based SCI classification, the modulation scheme used for different frame payloads should always look the same to Eve. We achieve that by embedding the payload's original modulation symbols in the constellation map of the highest-order modulation scheme supported by

the underlying system (denoted by $\mathcal{M}_m$). At the same time, we need to preserve the original demodulation performance at Bob.

To unify various payload modulation schemes, denoted by $Mi, i = 1, 2, \ldots, M$, each modulated symbol of Alice's payload is combined with one modulated FJ traffic, producing one point in the constellation map of $\mathcal{M}_m$. As long as the distribution of these points in the target constellation map is uniform, similar to the distribution of the points of a random $\mathcal{M}_m$-modulated information signal, and a given symbol is independent of the previous and next symbols (from Eve's perspective), Eve cannot determine if $M_i = \boldsymbol{\mathcal{M}_m}$.
In general, a higher-order modulation scheme is more susceptible to demodulation errors. The minimum Euclidean distance between the symbols in the constellation of Mi , denoted by d_{min}, i, specifies the probability of a demodulation error (hence, the BER) at a given SNR value. This d_{min}, i generally decreases with .

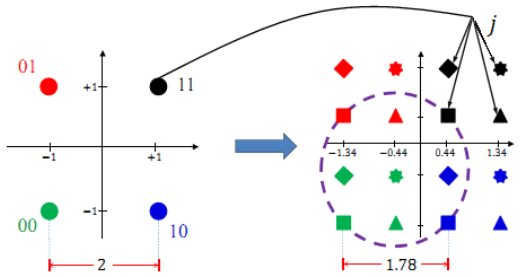


Figure-3.1 Optimal mapping from QPSK to 16-QAM.

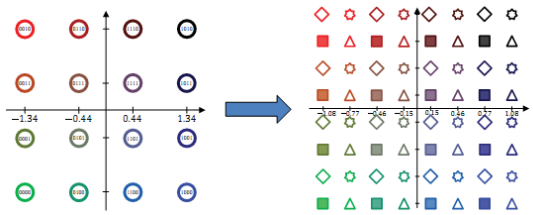
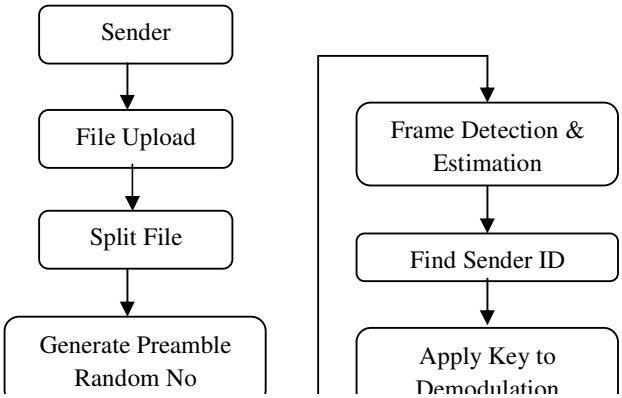


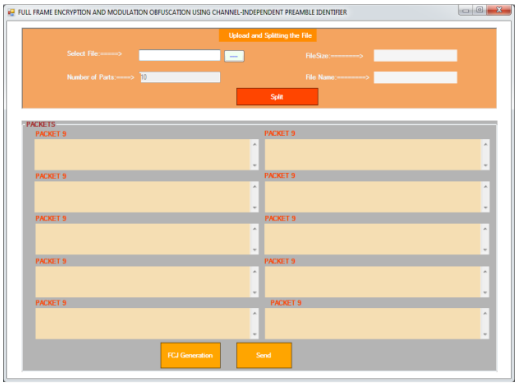
Figure 3.2 Optimal mapping from 16-QAM to 64-QAM.

DATA FLOW DIAGRAM:

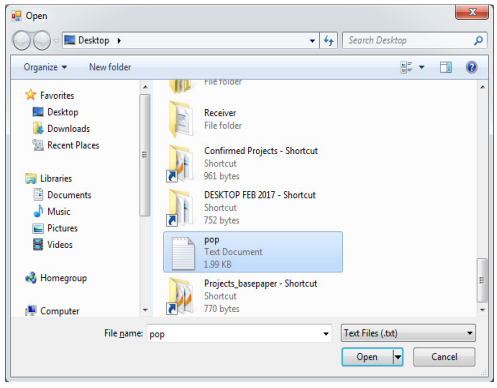
Full Frame Encryption



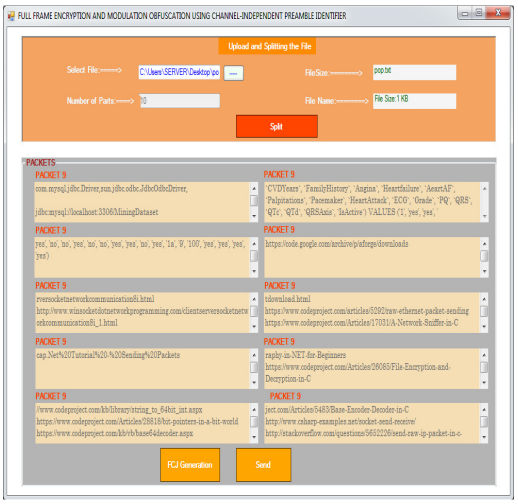
SENDER NODE



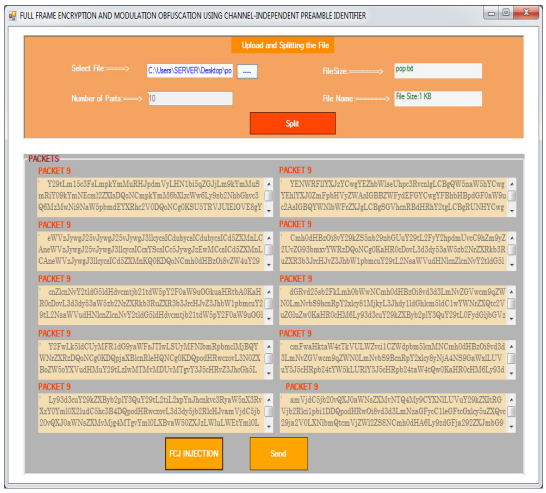
FILE UPLOAD



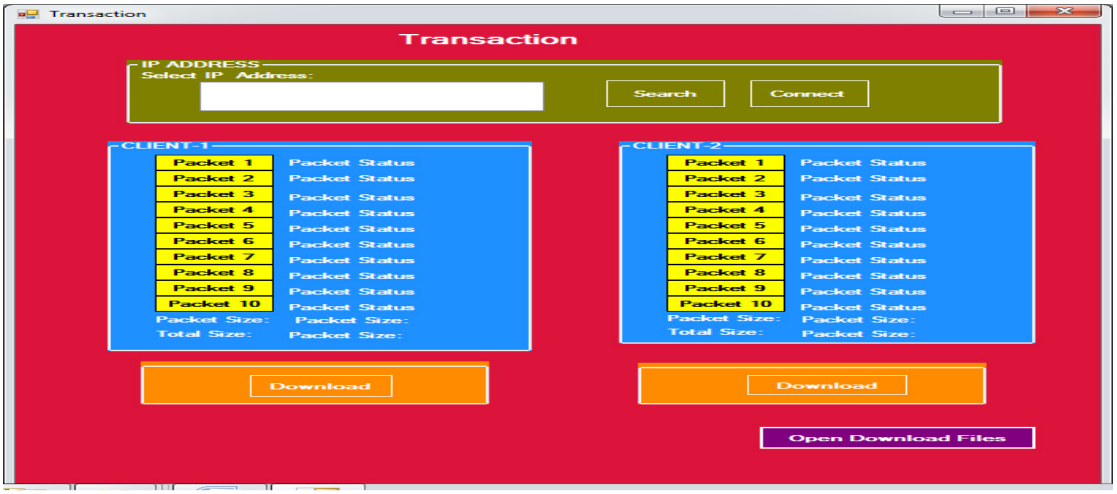
SPLIT FILE



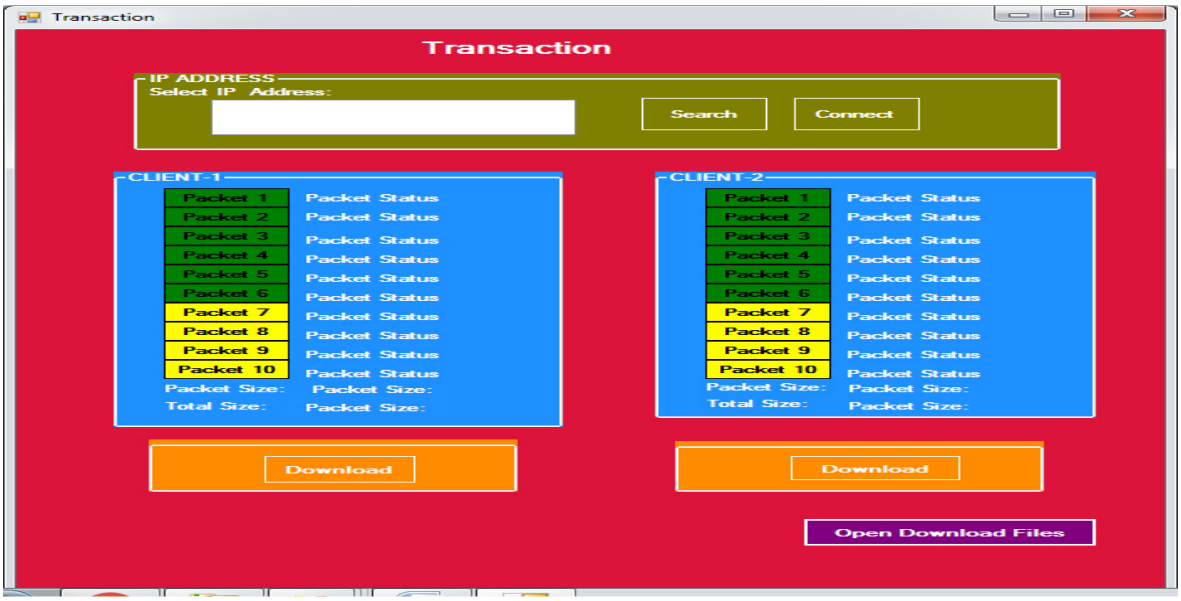
ENDOCE PACKET WITH FCJ



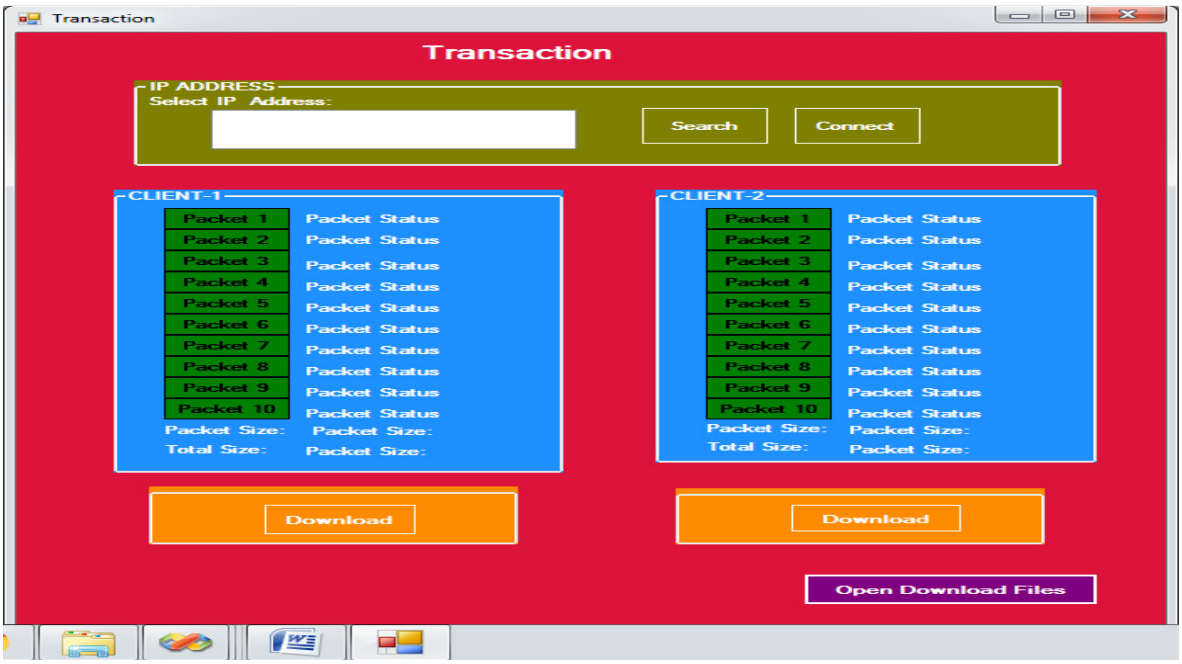
TRANSACTION



PARTIAL PACKET RECEIVED



DOWNLOAD RECEIVED PACKETS



IV. CONCLUSION

Preventing the leakage of transmission attributes, including unencrypted PHY/MAC header fields and the payload's modulation scheme, is challenging. In this paper, we proposed *Friendly CryptoJam* (FCJ) to 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. FCJ employs three main techniques. First, a message embedding technique was developed to overlay a frame-specific PHY-layer sender identifier on the frame preamble, obviating the need for MAC address and facilitating synchronous lightweight keystream generation and key lookup at PHY layer.

V. FUTURE ENHANCEMENT

In future several upper-layer techniques, such as padding, traffic morphing, and packet features masking at the application layer, have to be proposed to prevent the leakage of SCI by altering the true traffic statistics. These techniques, however, trade off higher traffic overhead for increased privacy. To reduce the overhead, traffic reshaping at the MAC layer is used to dynamically distribute the traffic among several virtual MAC interfaces; hence reshaping the statistical traffic profile of each of the interfaces.

REFERENCES

1. F. Zhang, W. He, X. Liu, and P. G. Bridges, "Inferring users' online activities through traffic analysis," in *Proc. 4th ACM WiSec Conf.*, Hamburg, Germany, 2011, pp. 59–70.
2. S. Chen, R. Wang, X. Wang, and K. Zhang, "Side-channel leaks in Web applications: A reality today, a challenge tomorrow," in Proc. IEEE *Symp. Secur. Privacy (SP)*, May 2010, pp. 191–206.
3. B. Miller, L. Huang, A. D. Joseph, and J. D. Tygar, "I know why you went to the clinic: Risks and realization of HTTPS traffic analysis," in *Proc. 14th Int. Symp. Privacy Enhancing Technol. Symp. (PETS)*, Amsterdam, The Netherlands, Jul. 2014, pp. 143–163.
4. T. Stöber, M. Frank, J. Schmitt, and I. Martinovic, "Who do you sync you are?: Smartphone fingerprinting via application behaviour," in *Proc. 6th ACM WiSec Conf.*, Budapest, Hungary, 2013, pp. 7–12.
5. J. S. Atkinson, J. E. Mitchell, M. Rio, and G. Matich, "Your WiFi is leaking: What do your mobile apps gossip about you?" *Future Generat. Comput. Syst.*, in press, 2016.

6. 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.
7. J. Freudiger, "How talkative is your mobile device?: An experimental study of Wi-Fi probe requests," in *Proc. 8th ACM WiSec Conf.*, New York, NY, USA, Jun. 2015, Art. no. 8.
8. F. Zhang *et al.*, "Thwarting Wi-Fi side-channel analysis through traffic demultiplexing," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 86–98, Jan. 2014.
9. K. P. Dyer, S. E. Coull, T. Ristenpart, and T. Shrimpton, "Peeka-boo, I still see you: Why efficient traffic analysis countermeasures fail," in *Proc. IEEE Symp. Secur. Privacy (SP)*, May 2012, pp. 332–346.
10. A. Iacovazzi and A. Baiocchi, "Internet traffic privacy enhancement with masking: Optimization and tradeoffs," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 2, pp. 353–362, Feb. 2014.

