IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 27, 2017, accepted January 26, 2018, date of publication February 9, 2018, date of current version March 13, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2802783

A Covert Channel Over VoLTE via Adjusting Silence Periods

XIAOSONG ZHANG[1,2], YU-AN TAN[1], CHEN LIANG[1], YUANZHANG LI[1,3], AND JIN LI[4]

[1]School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China
[2]Department of Computer Science and Technology, Tangshan University, Tangshan 063000, China
[3]Research Center of Massive Language Information Processing and Cloud Computing Application, Beijing 100081, China
[4]School of Computer Science, Guangzhou University, Guangzhou 510006, China

Corresponding author: Jin Li (lijin@gzhu.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant U1636213.

ABSTRACT Covert channels represent unforeseen communication methods that exploit authorized overt communication as the carrier medium for covert messages. Covert channels can be a secure and effective means of transmitting confidential information hidden in overt traffic. For covert timing channel, the covert message is usually modulated into inter-packet delays (IPDs) of legitimate traffic, which is not suitable for voice over LTE (VoLTE) since the IPDs of VoLTE traffic are fixed to lose the possibility of being modulated. For this reason, we propose a covert channel via adjusting silence periods, which modulates covert message by the postponing or extending silence periods in VoLTE traffic. To keep the robustness, we employ the Gray code to encode the covert message to reduce the impact of packet loss. Moreover, the proposed covert channel enables the tradeoff between the robustness and voice quality which is an important performance indicator for VoLTE. The experiment results show that the proposed covert channel is undetectable by statistical tests and outperforms the other covert channels based on IPDs in terms of robustness.

INDEX TERMS Covert channel, VoLTE, voice activity detection, silence period.

I. INTRODUCTION

Covert channels are used for the secret transfer of information. Encryption only protects communications from unauthorized parties to decrypt data, while covert channels are designed to hide the presence of communications. Lampson [1] first originated covert channels in 1973 and views covert communication as the process of communicating data through a transferring channel that is neither designed nor intended. A covert channel is also defined as a communication channel where information is transferred in a manner that violates the system security policy [2]. To build an effective covert channel, numerous issue-specific solutions have been proposed [3]–[7]. An important class of covert channels is covert timing channel that can transfer information to a receiver by modulating the timing behavior of an entity, such as the inter-packet delays (IPDs) or the reordering packets of a packet stream.

Covert timing channels are designed to conceal the existence of the covert message by hiding the fact that the covert channel exists. They are mainly characterized by undetectability and robustness. Undetectability prevents the adversary from detecting the existence of the timing channel by distinguishing covert from legitimate traffic. This advantage of simplicity and effectiveness has given rise to many statistical detection methods such as Kullback-Leibler divergence (KLD) test [8], Kolmogorov-Smirnov (KS) test [9] and entropy test among others [10]–[13]. On the other hand, active adversaries may disrupt the covert channel by introducing additional noise, such as jamming, into the channel. Despite inherent network jitter and malicious noise, robustness enables the covert channel to function properly, so that it is of necessity to integrate robustness with designing covert channels to guarantee covert message to be successfully transmitted.

For the voice telephony, the existing works mainly build covert channels over voice over Internet protocol (VoIP). No research has been available on building covert channels over VoLTE (Voice over LTE) which provides high-definition (HD) voice calls. Apart from traditional VoIP whose services are provided over various types of wireless networks, VoLTE is a technology via which voice is transported over the LTE network only. When mobile terminal user makes a call, the voice is actually carried over the carriers through high-speed data network instead of the

Circuit Switching network. VoLTE provides a smooth transition path from mixed network voice service to ubiquitous all-LTE network voice evolution [14]. Even though the advent of Internet based voice services, such as Skype and WeChat pose new challenges to VoLTE in cost and portability, VoLTE is technically superior to these VoIP applications. For example, IP traffic loads in networks and devices affect the quality of VoIP services, but not in VoLTE. Also, VoLTE offers much better interoperability as it is based on standards. Moreover, mobile networks will evolve into their fifth generation (5G) and VoLTE is a natural fit within 5G. Therefore, VoLTE will serve as the foundation for telecom-grade voice and video calling services in future 5G networks. Covert channels can be a secure and effective means of transmitting confidential information over 5G networks.

However, the existing schemes based on IPDs cannot be applied to VoLTE because the IPDs are fixed and then difficult to modulate covert message into IPDs of the VoLTE traffic. Therefore, we proposed a novel covert channel over VoLTE, in which the covert message is hidden by postponing and extending silence periods of the overt traffic. The fixed IPDs will not affect the proposed covert channel since the silence periods do not involve IPDs. Moreover, our covert channel is robust enough since packets out of order and packets loss, caused by VoLTE traffic jitter, have litter effect on the silence periods.

Our contribution is that we provide a covert channel by adjusting the silence periods over VoLTE, which is statistically undetectable by KS and KLD tests, while being robust against disruptions caused by inherent network jitter and malicious noise from active adversaries. Covert message is thus modulated by adjusting silence periods of VoLTE traffic. We employ Gray code and silence periods grouping to achieve robustness against intended and unintended channel noise. Our design features tunable parameters that allow to trade-off the intended voice quality against the robustness. In terms of robustness, the proposed approach outperforms the other methods based on IPDs due to the full reliability guarantee of VoLTE.

The remainder of our paper is organized as follows: In section II, we review related work on existing covert channels. In section III, we present the preliminaries including VoLTE, RTP and VAD. Then, in section IV, we give an overview of the proposed covert channel over VoLTE, with discussions on the performance metrics including undetectability and robustness and voice quality. In section V, we show how to build a covert channel for composite performance requirements. We present the experimental results and analyses in section VI and evaluate the proposed covert channel in section VII. Finally, we conclude with discussion and future research directions in section VIII.

II. RELATED WORK

Covert channels are defined as a communication mechanism that can evade access control policies by using a medium that normally goes by unmonitored. More recently, focus has shifted toward covert channels in network protocols from those which are initially identified as a security threat on monolithic systems such as mainframes. Previous work on covert channels in network traffic can be divided according to the media used: packet payload, packet header, or packet timing behavior. Different researchers have focused on identifying the possible applications in public networks [15]–[19]. Here we highlight the methods of building covert timing channels.

A covert timing channel can transfer information based on the modulation of some timing behavior of an entity. A most common entity is the IPD of a traffic generated by a distributed application. Specifically, covert message is transmitted by modulating the IPDs of overt traffic. For example, an IPD-based covert channel proposed called TCPScript in [3] which embedded covert messages into the TCP bursts. It maintains normal burstiness patterns of TCP to ensure undetectability. The authors presented results for robustness against packet loss and packet reordering, without offering solution to improve them. Another IPD-based covert channel was proposed in [15] including two version. One is undetectable and with low capacity, and the other is apt to be detected but with high capacity. In [20] a time-replay covert channel was proposed where covert messages were transmitted by replaying a previously recorded sequence of timing intervals. Similar to the previous two works, it emphasizes undetectability but does not address robustness.

Researchers developed various types of coding schemes to improve the robustness of covert channels. Liu *et al.* [21] presented an IPD-based covert channel using spreading codes which made the channel robust but with low capacity. Wu *et al.* [22] proposed a covert channel where the covert message was compressed by Huffman coding and gave the experimental results of undetectability and capacity. In CoCo [23], four different types of error correcting coding methods were used to improve robustness of the scheme. Besides robustness, undetectability and capacity were also considered in this work.

Gianvecchio *et al.* [24] proposed a model-based covert channel to mimic the statistical properties of legitimate traffic but did not consider the robustness requirement. However, all requirements of an efficient covert channel including undetectability and robustness were considered in [25] which proposed a covert channel with distribution matching. Archibald and Ghosal [26] proposed a model-based covert channel using fountain codes which was the first robust timing channel scheme based on rateless codes. The work implements a guard band strategy at the sender or the receiver to further improve the robustness. A more efficient turbo covert channel was proposed in [27], which also employed precoding and guard band strategy. Liu *et al.* [28] proposed model-fitting covert channels using analog fountain codes that not only met the traditional goals of undetectability and robustness but were also model-adaptive, however, the proposed covert channels may not perform well for particular type of network traffic.

Another example of entities of covert timing channels is the packet ordering in a network. The covert channel based on packet reordering is implemented by encoding the covert messages in the ordering of packets in overt traffic. Ahsan and Kundur [29] proposed a covert channel based on packet ordering within the IPSec framework. Another covert channel presented in [30] using polynomial-time optimal encoding and decoding algorithms to obtain high capacity. However, there is no practical solution proposed to improve undetectability and robustness in these two works. Different from [29] and [30], El-Atawy and Al-Shaer [31] proposed a similar covert channel which used specific permutations of consecutive packets to improve robustness and mimics real traffic distribution to improve undetectability.

While these works have performed well in terms of undetectability and robustness, they are mainly based on computer networks rather than VoLTE. Since the fixed IPDs of VoLTE traffic result in some existing methods unfeasible, we will employ new examples as the carrier to carry the covert message. Consequently, in this paper, inspired by the two examples of covert channel entities including the IPD and the packet ordering, we design a new covert channel by adjusting silence period which transmits covert message through adjusting the silence periods, considering both undetectability and robustness and voice quality.

III. PRELIMINARIES

A. COVERT CHANNELS FOR VoLTE

Although some existing technologies today also provide voice services over LTE, VoLTE has gained prominence and evolved as a viable solution because it has the advantages of higher spectral efficiency, low cost and compatibility with rich multimedia voice experience over the competitors or counterparts [14]. To provide guaranteed QoS to VoLTE service, mobile network operators logically differentiate the VoLTE traffic and non-VoLTE traffic using different bearer type. Bearer is a logical link through which a certain type of service traffic can be carried on. VoLTE service is delivered over dedicated bearer, which has the highest priority, while non-VoLTE service including VoIP service is delivered over default bearer which only preserves best effort QoS. Therefore, the building of covert channel over VoLTE cannot significantly reduce the voice quality.

All of the VoIP traffic over LTE network is generated by third party mobile applications, while VoLTE traffic can only be generated by the phone dial application which is typically pre-installed by mobile network operators. It is for this reason that capturing and processing VoLTE traffic is more difficult than VoIP traffic.

While some work has been done on building covert channel over VoIP, the existing research has not involved the construction of covert channels over VoLTE. These methods based on IPDs cannot apply to VoLTE application because the IPDs of VoLTE traffic are fixed. Therefore, we propose a new solution building the covert channel over VoLTE by adjusting the silence periods of VoLTE traffic.

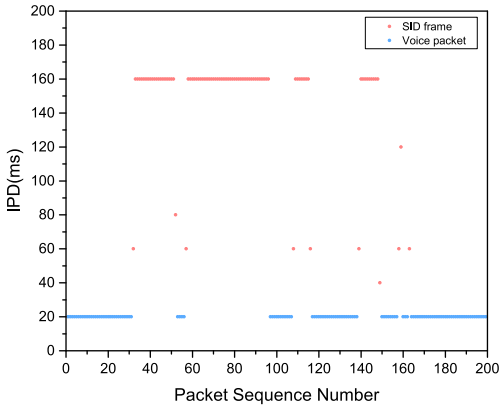


FIGURE 1. IPDs of RTP packets of overt VoLTE traffic.

B. SILENCE PERIODS IN VoLTE VOICE TRAFFIC

Real-time Transport Protocol (RTP) is an application layer protocol which was developed to transfer real-time data, including video and audio on the multicast and unicast network. VoLTE use RTP to transfer voice and video traffic in real-time manner on top of UDP.

VoLTE uses AMR-WB as a vocoder. AMR-WB codec is a speech codec which has been applied in the 3GPP LTE network for voice compression and decompression. It is fully described in [32]. AMR-WB codec uses a sampling rate of 16 kHz, which covers 50-7000 Hz audio bandwidth. It has 9 different codec modes (from mode 0 to mode 8) corresponding to 9 source bit rates in range of 6.6-23.85 Kb/s. Each of them generates an encoded 20 ms voice frame and switches among them every 20 ms. The bits in the encoded voice frame are ordered according to their subjective importance. AMR-WB packet size depends on the bit rate.

For voice communication, bandwidth optimizations are needed. One of the most effective and popular is voice activity detection (VAD), which allows the sender side to stop the transmission during speech pauses. This can provide major bandwidth saving, as typically VoIP calls are characterized by 35% to 70% of silence periods [33]. A talk spurt is a continuous segment of speech between silent intervals where only background noise can be heard. Segmenting speech streams into talk spurts allows bandwidth to be conserved by not sending excess data in silent intervals, and also allows synchronization, buffering and other parameters of the communications system to be readjusted in the intervals between talk spurts. The AMR-WB codec also utilizes an integrated VAD. The function of the VAD algorithm is to indicate whether each 20 ms frame contains signals that should be transmitted.

Therefore, VoLTE voice services exist in two states: talk spurts and silence period. In talk spurts, the sending interval of the voice packet is 20 ms and the voice packet size depends on the currently used coding rate. In silence period, the sending interval of SID (Silence Insertion Descriptor) frame is usually 160 ms. The difference between talk spurts and

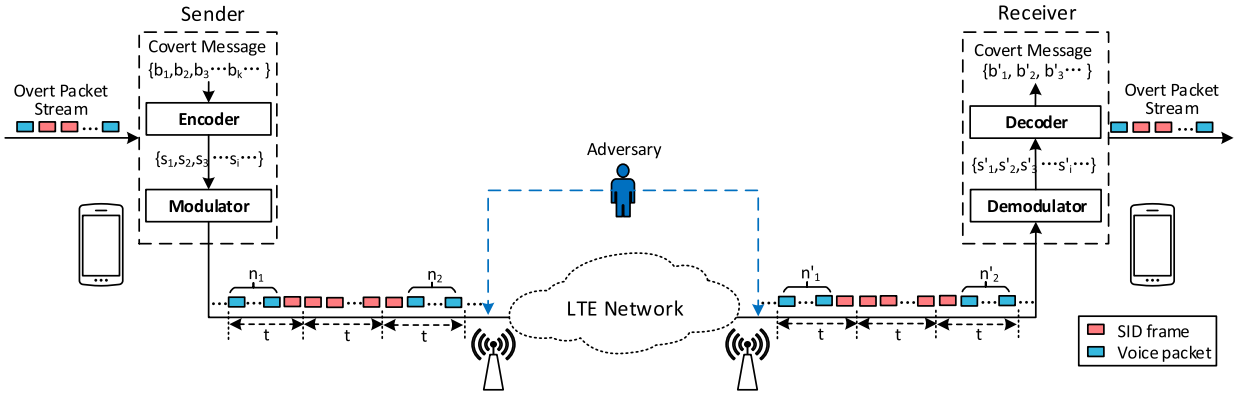


FIGURE 2. System model of the proposed covert channel.

silence period is that the size of voice packet is larger than that of SID frame, and there is a significant difference between the time intervals of the adjacent voice packets and the time intervals of the adjacent SID frames. The RTP header for the SID frame should be constructed as if the comfort noise were an independent codec [34]. Thus, the RTP timestamp designates the beginning of the silence period. Fig.1 shows that the IPDs of RTP packets of VoLTE traffic calculated by the timestamps from RTP packets. All the IPDs of the voice packets are 20ms and most IPDs of the SID frames are 160ms. Through further detailed observation, there are four silence periods in fig.1, and the IPDs of the first SID frames of the silence periods are all 60ms and the IPDs of the last SID frames of the silence periods are all in the range of (20ms-160ms) which are multiples of 20ms. Based on the characteristics of the silent period of VoLTE traffic, we can adjust the silent periods to build a covert channel over VoLTE.

IV. DESIGN OF COVERT CHANNELS OVER VoLTE

The general communication model considered in this work is composed by a sender and a receiver wanting to communicate through the LTE network in a covert manner. To this aim, we establish a robust and undetectable covert channel by exploiting silence periods of VoLTE traffic. We will use covert communication to refer to a communication with embedded covert channel and overt communication without embedded covert channel. These two different channels are associated to the traffic types by covert traffic and legitimate traffic. Figure 2 displays system model of the overall covert timing channel that shows the data flow from the sender to the receiver. The main notations and symbols are shown in Table 1.

A. SYSTEM MODEL

The sender encodes the covert message to distinct symbols that are modulated by postponing and extending silence periods of legitimate traffic, and then transmit the message covertly to the receiver over VoLTE network. The receiver receives the adjusted packets and demodulate them into symbols that are decoded back to be read. We define the sender and the receiver entities as the two ends in a covert communication. The covert message is $\{b_1, b_2, b_3, \ldots, b_k, \ldots\}$ where b_k is the k-th information bit. The covert information bits are encoded to symbols, which are finally modulated in the numbers of adjusted packets $\{n_1, n_2, n_3, \ldots, n_i, \ldots\}$ of the packet stream that is sent from the sender to the receiver. To decode the covert message accurately, the custom parameters related to the covert channel are assumed to be shared between the sender and the receiver prior to the covert transmission.

TABLE 1. The main notations and symbols.

Notation	Description
b_k	The k-th information bit of the covert message
b'_k	The k-th information bit decoded at the receiver
s_i	The i-th code symbol encoded with Gray code
s'_i	The i-th code symbol demodulated at the receiver
n_i	The i-th adjustment amount transmitted from the sender
n'_i	The i-th adjustment amount received at the receiver
bl	The length of covert information bits hidden each time
N_{sp}	Number of Silence periods
N_p	Total number of packets
T_C	Transmission Time of covert message
R_l	Packet loss rate
C	Capacity of the covert timing channel (bit/s)
P_e	Bit error rate (BER)

B. ADVERSARY MODEL

An adversary is an intrusion detection system or a channel jammer that can monitor or manipulate the transmission between the sender and the receiver. We assume that an adversary has access to both the legitimate and the covert traffic, but not simultaneously, since the covert traffic is generated by modifying the legitimate input traffic. At any time only a single traffic type, either legitimate or covert, can exist in the transmission channel. The adversary can easily derive some characteristics, such as the distribution of the numbers of SID frames in silence periods, and has the knowledge of the modulation algorithm of the covert channel. However, if he cannot detect the traffic abnormality, it is difficult for him to identify the covert channel. In particular, regarding undetectability it is adequate to restrict analysis to an attacker

that passively listens to the network communication. Since the hidden communication lies within the VoLTE channel, the adversary cannot alter the traffic too much actively as it will result in the poor call voice quality. Thus, the real-time nature of VoLTE HD voice represents a sort of protection against aggressive network-based countermeasures, such as deep packet inspection tools or attackers limiting the performance of the covert channel through traffic normalization approaches. Nonetheless, the covert channel needs to be robust against network noise which is injected by the network or intentionally by an adversary trying to disrupt the covert channel.

C. DESIGN CRITERIA

We design the covert channel for VoLTE according to the following three goals, undetectability and robustness and speech quality, which are also the performance metrics of the proposed covert channel.

1) UNDETECTABILITY

On a high level, undetectability means that no efficient algorithm can distinguish between the numbers of SID frames (NoSFs) of legitimate traffic and covert traffic. A covert channel is Polynomial Undetectable with respect to a security parameter δ if there exists a negligible function $f(\delta)$ such that $|T(n)-T(n')| \leq f(\delta)$ for some probabilistic polynomial-time statistical test T, referring to the definition of [10], where n and n' are arbitrary N samples of the numbers of voice packets (NoVPs) or NoSFs of the legitimate traffic and covert traffic. We assume that N is a positive integer. Polynomial time statistical tests include KS test, KLD test among others. In this paper, KS test and KLD test are employed to evaluate the undetectability of the proposed covert channels. Moreover, we also use the average packet number of silent period and ratio of silent packets for the overt traffics and covert traffics to appraise the undetectability.

2) ROBUSTNESS

Due to channel unreliability during transmission, the NoVPs and NoSFs generated at the sender will be perturbed at the receiver. In addition, a channel jammer can be introduced to reduce the channel capacity of covert communication, which causes further deviation of the numbers of packets from their designated values. To resist these unintended and malicious disruption, covert communication must be robust. Specifically, the robustness can be measured as the capability to achieve a decoding bit error rate (BER) $P_e \leq \epsilon$ under a given robustness requirement $\epsilon \in R^+$ [10]. Different parameters can be selected according to the given robustness requirement.

3) VOICE QUALITY

As discussed, preserving the voice quality is important to not reveal the presence of the covert channel. This is especially true for the case of VoLTE voice call, which deal with human-to-human HD voice communication. In addition, the availability of a rich literature on multimedia analysis for detecting artifacts requires to completely preserve the overall quality of the conversation. Therefore, to assess the impact of the covert message within the VoLTE conversation, we perform voice quality tests and discuss the appropriate length of covert information bits.

D. PERFORMANCE CALCULATION

The capacity of our covert timing channel is given by

$$C = \frac{N_C bl}{T_C} \tag{1}$$

where N_C denotes the number of locations where the covert message can be hidden. When bl is different, N_C is different too. This is because not all silent periods can be delayed and expanded, for example, when the silence period is too short or the adjacent voice period is too short, the covert message cannot be embedded.

The BER of our covert channel is given by

$$P_e = \frac{N_{sp} R_l}{N_C bl} \tag{2}$$

where R_l denotes the packet loss rate of VoLTE traffic, NspRpl is the number of lost packets at the end of silence periods since the error bit is mainly caused by the inability to determine whether the lost packet is voice packet or SID frame at this location.

The average number of SID frames in silence periods is given by

$$N_{average} = \frac{\sum_{i=1}^{N_{sp}} N_{S_i}}{N_{sp}} \tag{3}$$

where N_{S_i} denotes the number of SID frames in the i-th silence period.

The ratio of SID frames to total packets is given by

$$R_{sf} = \frac{\sum_{i=1}^{N_{sp}} N_{S_i}}{N_p} \tag{4}$$

The average number of SID frames in silence periods and ratio of SID frames to total packets, along with KLD and KS test, are used to evaluate undetectability of our covert channel.

V. IMPLEMENTATION

The sender and receiver have to agree on an overt channel to host their covert communication. This includes determining the encoding scheme to be used in encoding covert information bits into code symbols to be modulated by adjusting silence periods. Both parties have to agree on their choice according to network conditions and users requirements to successfully transmit the covert information.

While the covert message is modulated into the RTP traffic of VoLTE telephony, to avoid detection, the proposed covert channel must not disrupt the stream or alter the packets in an aggressive manner. The adversary could find incorrect

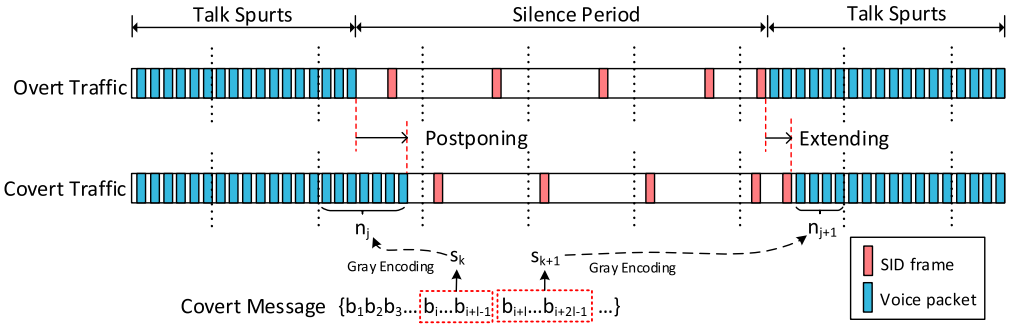


FIGURE 3. Modulation of covert message by postponing and extending the silence periods.

details about the amount of exchanged data or the number of dropped packets. Hence, the sender must adjust RTCP reports accordingly. The postponing and extension of silence periods could affect the timing of the stream and reducing the voice quality of the conversation. Hence, ''time warps'' should be properly addressed as to meet the HD voice quality requirements for VoLTE.

A. ENCODING

We employ Gray code to mitigate channel noise since Gray code characterized by two successive values different in only one bit results in being robust to packets reordering and packets loss. The covert message is encoded by Gray code according to the length bl of covert information bits that the sender and receiver have determined. The covert information bits are encoded in selected length bl into code symbols, where each silence period carries two pieces of information bits of length bl. The greater the bit length bl of symbol is, the higher the transmission efficiency of covert message is. Although the bit error rate is relatively lower than many other IPD-based methods, the bit length has a certain impact on the bit error rate. The smaller the bit length of code, the higher the bit error rate.

B. MODULATION

Since the transformation of a symbol s_i to the n_i must be invertible, we consider a one-to-one mapping: $n_i := G(s_i), i = 1, 2, \ldots,$ where $G(\cdot)$ is an invertible function. The demodulation at the receiver is done by: $n'_i := G^{-1}(s'_i), i = 1, 2, \ldots,$ where n'_i and s'_i are the received NoVPs and the symbols which the NoVPs demodulated into, respectively. In the following, we show how to obtain $G(\cdot)$ and $G^{-1}(\cdot)$, which correspond to modulation and demodulation, respectively. The modulation of covert message by postponing and extending the silence periods is shown in Fig.3.

The first packet of silence period functions as a synchronization identifier which marks the start of covert channel. The sender clocks at the beginning of the first silence period and cycles through 2^{bl} voice packet time. At the end of the silence period, it calculates the remaining time in the current time segment to determine how many voice packets can be transmitted. The number of voice packets must be in the range of $(0 \sim 2^{bl} - 1)$. At this moment, it gets the symbol value encoded from covert information bits according to the selected code length by Gray code. If the number of voice packets is not equal to the symbol value, the sender will extend the current silence period and shorten the adjacent talk spurts to make the two values equal. Then, at the beginning of next silence period, the sender computes the number of voice packets in the current time segment and gets the symbol value encoded from covert information bits again at the same time. Comparing the number of voice packets and the symbol value again, if not equal, the sender will extend the current talk spurts and postpone the adjacent silence period to make the two values equal. The modulation process is also demonstrated in Algorithm 1.

Algorithm 1 Covert Message Encoding and Modulation

```
Get the start of the first Silence Period
TimeCounting(pow(2, bl))
while !eof(CovertMessageFile) do
  if End(SilencePeriod) then
    num ← Count adjustment amount
    grayvalue ← GraytoDec(GetMessage(bl))
    if num ≠ grayvalue then
      Extend(CurrentSilencePeriod)
      Shorten(AdjacentTalkSpurts)
    end if
  end if
  if Start(SilencePeriod) then
    num ← Count adjustment amount
    grayvalue ← GraytoDec(GetMessage(bl))
    if num ≠ grayvalue then
      Extend(CurrentTalkSpurts)
      Shorten(AdjacentSilencePeriod)
    end if
  end if
end while
```

The extension and postponing of the silence period is achieved by processing the SID frames. For example, when the sender need to prolong the silence period, it first check whether the last SID frame in the silence period has reached the maximum silence interval 160ms, and if not, it extends

the interval, and if yes, it inserts a new SID frame. The way to postpone the silence period is the opposite. Similarly, it can be inferred that the extension and shortening of talk spurts is attained by processing the voice packets. The sender just inserts and deletes the voice packet simply to achieve the goal. When a voice packet is inserted, the payload of the packet employs the voice data of the last packet of the current voice period. The reason for this is that the last voice packet is generally the noise at the end of a short conversation and has a very small impact on voice quality of VoLTE traffic.

C. DECODING AND DEMODULATION

The decoding and demodulation process is the reverse process of encoding and modulation. The receiver starts to clock when it receives the first SID frame, and calculates the NoVP n'_i on each edge of silence period including the end and the start. The symbol s'_i can be obtained with the NoVP n'_i. The covert information bits of selected length l can be available by converting the symbol s'_i to a binary reflected Gray code.

D. SILENCE PERIODS GROUPING

Network jitter can cause recovery data inaccurate and large packet loss rate will produce a chain reaction. To solve this problem, we employ silence periods grouping to achieve synchronization within the group and no effect each other between groups. The beginning of the first silence period in each group functions as synchronization identifier. Silence periods grouping can reduce the bit error rate which the packet loss causes since packet loss in current group do not affect other groups. Additionally, silence periods grouping will affect the capacity of the proposed covert channel since the beginning of the first silence period in each group cannot be used as modulation locations. The smaller grouping size gl which depends on the degree of the network jitter will reduce the capacity of the covert channel even more. Therefore, the grouping size is a tunable parameter that enables trade-off between the required robustness and capacity.

VI. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we mainly benchmark the proposed covert channel by examining the undetectability and robustness. Moreover, the capacity of covert channel and voice quality are analyzed.

A. EXPERIMENTAL SETUP

To analyze the performance of the proposed covert channel, we used VoLTE traffic between two mobile phones. Our schemes can be successfully implemented on different models of mobile devices and different versions of android, as long as the devices and the system support VoLTE. We select two Coolpad 8722V mobile phones, where the Android version is 5.1 and kernel version is 3.10.65, as the sender and the receiver to test our schemes. We capture traffic both at the sender and the receiver in our experiment. Since the existing software cannot capture VoLTE voice packets processed by baseband program, we developed a capturing program based on Android kernel. For overt traffic, we used our crawler software to simultaneously capture packets at both the sender and the receiver. Covert traffic is generated by encoding and modulating public traffic according to our scheme. In order to test voice quality, we also developed a program to extract voice information from the packets. We mainly test the proposed covert channel under different length of covert information bits, and compare and analyze the experimental results.

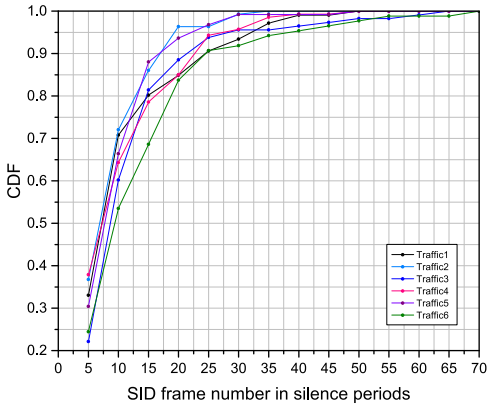


FIGURE 4. CDF of NoSFs of silence period of overt traffics.

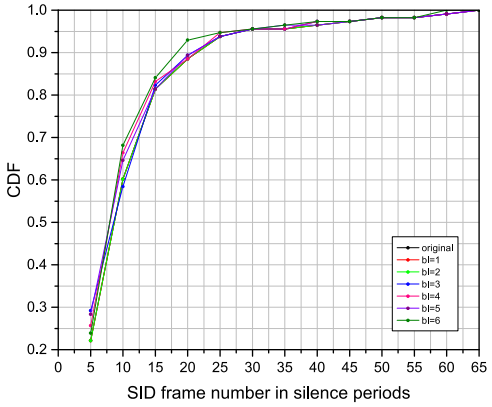


FIGURE 5. CDF of NoSFs of silence period of covert traffics.

B. UNDETECTABILITY

We analyze the undetectability from the experimental results of three aspects. First of all, we calculate the cumulative distribution functions (CDFs) of the NoSFs for the legitimate traffics and the covert traffics. Fig.4 demonstrates the CDFs of the NoSFs for the six legitimate traffics. The results are shown in Fig. 5 which depict the CDFs of the NoSFs for original overt traffic and the corresponding covert traffics with different lengths bl. Fig. 5 indicates that all of the covert traffics mimic the legitimate traffic well. There is a relatively large difference between the CDF of covert traffic ($bl = 6$) and that of original traffic. This difference arises as the silence periods extend and postpone greatly. In comparison, the difference of the CDFs between overt and covert traffic

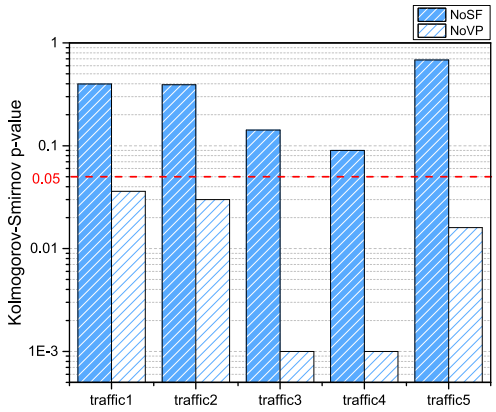


FIGURE 6. KS test for NoSFs of silence period of overt traffics.

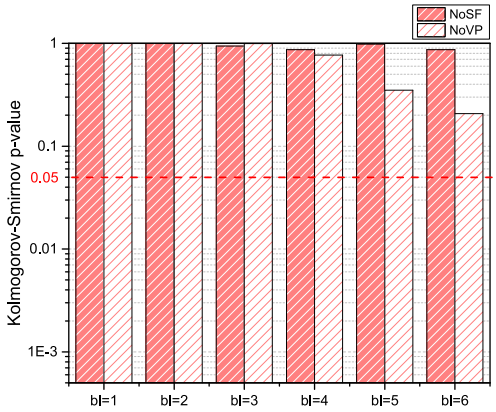


FIGURE 7. KS test for NoSFs of silence period of covert traffics.

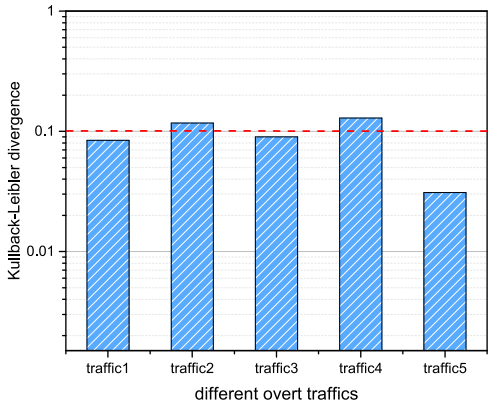


FIGURE 8. KLD for NoSFs of silence period of overt traffics.

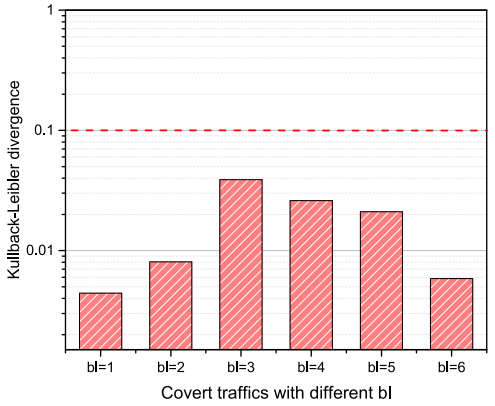


FIGURE 9. KLD for NoSFs of silence period of covert traffics.

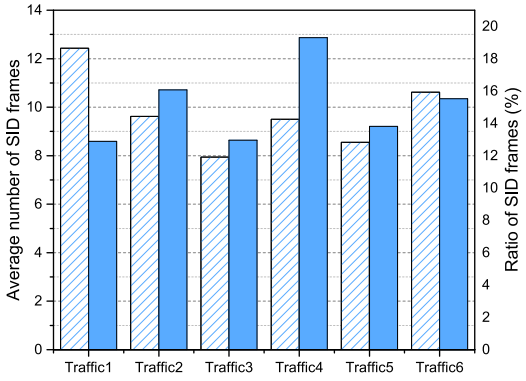


FIGURE 10. Average NoSFs of silence period and ratio of SID frames of overt traffics.

in Fig.5 is much smaller than the difference of the CDFs of overt traffics in Fig.4. Secondly, we use two standard statistical tests including KS test and KLD to visualize and verify undetectability. In Fig.6, the KS p-values are all greater than 0.05 for NoSFs of the five legitimate traffic, which represents that the NoSFs of these traffics fit the same distribution. On the other hand, the KS p-values are all smaller than 0.05 for NoVPs of the five legitimate traffic, which indicates that the NoVPs of these traffics are not distributed regularly. Fig.7 shows that KS p-values are all greater than 0.8 for NoSPs of the six covert traffics which can evade the KS test. Similarly, KS test is not valid in detecting the NoVPs of all the covert traffics. From the KS test results, we should focus on the undetectability for the SID frames because the SID frames have the same distribution in the overt traffics but the distribution is different for the voice packets.

Further, from Fig.8 and Fig.9, the KLDs of NoSFs for all covert traffics are smaller than the average KLD, even the smallest one, of the overt traffics. This performance is similar to that of KS test and the covert traffics can maintain good undetectability for both K-S test and KLD.

Finally, we calculate the average packet number of silence periods and ratio of SID frames to total packets for the overt traffics and covert traffics. From Fig.10 and Fig.11, the average packet numbers of silence period is similar to the original traffic and within the scope of the values of overt traffic, and the same holds for the ratio of SID frames.

C. ROBUSTNESS

Before we discuss the bit error rate, we first discuss capacity because the bit error rate is closely related to the capacity for our schemes. Fig.12 shows that the capacity generally increases with increasing length of the covert information bits and maximum reaches about 3 bit/s.

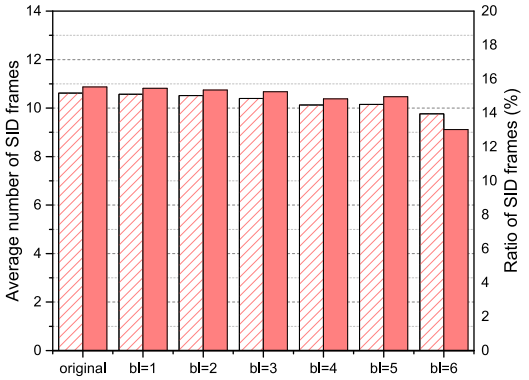


FIGURE 11. Average NoSFs of silence period and ratio of SID frames of covert traffics.

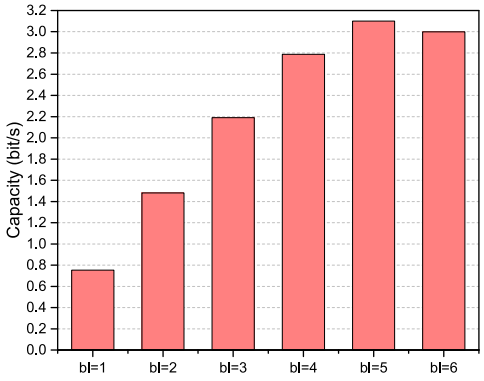


FIGURE 12. Capacity of our covert channel.

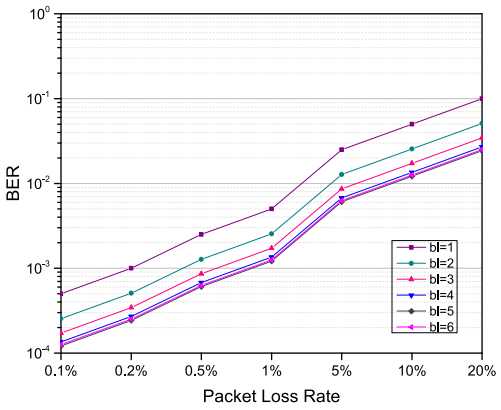


FIGURE 13. BER of our covert channel.

Without loss of generality, we measure the robustness of covert channels with BER of the decoded covert message. In Fig.13, the BER of the proposed covert channel under any bit length reaches 10^{-4} when Rr is less than 0.2%, whereas the BER increases to 10^{-2} when Rr is greater than 10%. The sort of performance for all bit lengths are similar in different network conditions. The sort of the bit error rate and the capacity for all bit lengths are exactly the same. In summary, the proposed covert channel remains valid even under high network jitter. Specifically, the covert channel with larger length *bl* can achieve better robustness.

Meanwhile, we compare the proposed silence-period-based covert channel (SPCC) with other five covert channels.

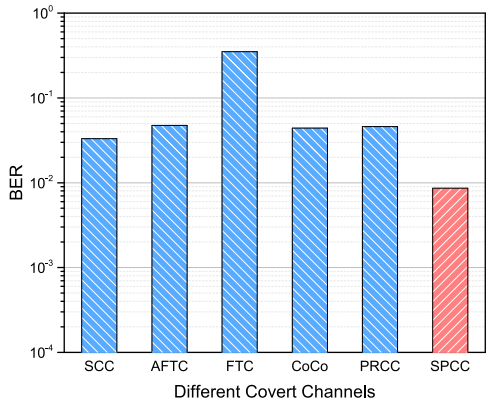


FIGURE 14. Comparison of BER between our CTC and other CTCs.

These schemes include spreading covert channel (SCC) [21], coding-based covert channels [23], the Fountain Timing Channel (FTC) [26], analog fountain timing channel (AFTC) [28] and packet reordering covert channel (PRCC) [31]. For each covert channel, we take the average BER in its best case. As shown in Fig.14, the BER of FTC is about 10^{-1}, whereas the BER of our covert channel reaches 10^{-3}, and that of the other four covert channels achieves 10^{-2}. The average BER of the proposed covert channel under normal conditions is observably lower than that of the other five schemes because only the special SID frames loss will affect the silence period.

D. VOICE QUALITY

To assess the quality of the conversation, we compared the original voice with the covert ones. The resulting quality has been evaluated by using mean opinion score (MOS). We considered seven different scenarios containing the origin traffic and covert traffics where the length of covert information bits is from 1 to 6. In all cases, the sender and the receiver used the AMR-WB codec which is applied in VoLTE and provides improved speech quality due to a wider speech bandwidth. Table 1 reports the collected results. Unsurprisingly, the scenario of origin voice has the highest MOS value. The MOS value of the scenario of $bl = 1$ reaches 4.231 and is very close to the value of the origin voice. The scenario of $bl = 6$ reaches the lowest MOS value 2.692, which indicates the voice quality is poor and barely accepted. Similar results, all in the $(3.5 \sim 4)$ range, have been obtained in other scenarios which is due to the fact that our covert channel does not greatly alter packets containing the voice, instead, it only produces negligible alterations to the stream due to additional jitter caused by postponing and extension of silence periods.

VII. EVALUATION

In the previous section, we give the experimental results in terms of undetectability and robustness and voice quality. By analyzing the experimental results, we have shown that increasing the length of covert information bits can significantly reduce the BER, which ensures that the covert message can be decoded out. However, greater length bl also results in

TABLE 2. MOS values of original voice and the covert ones.

Parameter setting	Origin	bl=1	bl=2	bl=3	bl=4	bl=5	bl=6
MOS value	4.308	4.231	3.923	3.846	3.769	3.550	2.692
Standard Deviation	0.480	0.439	0.277	0.376	0.439	0.510	0.480

deteriorating voice quality, which is a very important performance indicator for VoLTE. In addition, the capacity is proportional to bl. Therefore, such trade-off exists between the robustness and voice quality as different lengths *bl* yield different voice quality, given required robustness performance. To balance the robustness and voice quality, we can choose the moderate size of *bl*. In summary, we find that the length of covert information bits *bl* is a key factor, because larger length makes the covert channel liable to be exposed due to the poor voice quality, whereas smaller length results in higher bit error rate so that it is difficult to restore covert message.

Our covert channel has two limitations. One is that it can only be used to transmit small amount of covert message since the limited number of silent periods during normal conversations and the conversation time is often not long. It is inefficient to transfer big amount of data in this covert channel. The other is that voice quality gradually declines as the length of covert information bits increases due to that Excessive prolongation and shortening of the silence period will affect voice quality. However, VoLTE needs to ensure HD voice communications, thus the covert channel with too large length *bl* is not suitable for VoLTE. We plan to further improve this in future research.

VIII. CONCLUSION AND FURTHER WORK

In this paper, we pursue the possibility of building covert channels over VoLTE via adjusting silence periods, where the covert message is modulated by postponing or extension of silence periods. Silence period is a normal phenomenon in the voice communication and it is unlikely to raise suspicions. We employ Gray code and silence periods grouping to enhance robustness against intended and unintended channel noise. Using experiments with VoLTE traffic we have demonstrated the effectiveness of this covert channels with respect to undetectability and robustness and voice quality. Our approach outperforms the other IPDs-based methods in terms of robustness against channel jitter. These results may lead to several interesting future directions for this work. First, we can employ silence periods to convert covert message into covert symbols by different modulation algorithm to increase the capacity of covert channels. Second, we can design a valid covert channel for video telephony over VoLTE, inspired by the proposed approach for voice telephony. Third, we can design more efficient and effective protection techniques as well as detection mechanisms for VoLTE.

REFERENCES

[1] B. W. Lampson, "A note on the confinement problem," *Commun. ACM*, vol. 16, no. 10, pp. 613–615, 1973.

[2] *Trusted Computer System Evaluation Criteria*, document DoD 5200.28-STD, United States Department of Defense, 1985.

[3] X. Luo, E. W. W. Chan, and R. K. C. Chang, "TCP covert timing channels: Design and detection," in *Proc. IEEE Int. Conf. Depend. Syst. Netw. FTCS DCC (DSN)*, Jun. 2008, pp. 420–429.

[4] J. Li, Y. Zhang, X. Chen, and Y. Xiang, "Secure attribute-based data sharing for resource-limited users in cloud computing," *Comput. Secur.*, vol. 72, pp. 1–12, Jan. 2018.

[5] P. Li, J. Li, Z. Huang, C. Z. Gao, W. B. Chen, and K. Chen, "Privacy-preserving outsourced classification in cloud computing," in *Cluster Computing*. New York, NY, USA: Springer, 2017, pp. 1–10, doi: s10586-017-0849-9.

[6] J. Li, X. Chen, M. Li, J. Li, P. P. C. Lee, and W. Lou, "Secure deduplication with efficient and reliable convergent key management," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 6, pp. 1615–1625, Jun. 2014.

[7] J. Li, X. Huang, J. Li, X. Chen, and Y. Xiang, "Securely outsourcing attribute-based encryption with checkability," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 8, pp. 2201–2210, Aug. 2014.

[8] C. Cachin, *An Information-Theoretic Model for Steganography*. San Diego, CA, USA: Academic, 2004.

[9] S. Cabuk, C. E. Brodley, and C. Shields, "IP covert timing channels: Design and detection," in *Proc. 11th ACM Conf. Comput. Commun. Secur.*, 2004, pp. 178–187.

[10] F. Rezaei, M. Hempel, and H. Sharif, "Towards a reliable detection of covert timing channels over real-time network traffic," *IEEE Trans. Depend. Secure Comput.*, vol. 14, no. 3, pp. 249–264, May 2017.

[11] X. Du, M. Guizani, Y. Xiao, and H. Chen, "Transactions papers a routing-driven elliptic curve cryptography based key management scheme for heterogeneous sensor networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1223–1229, Mar. 2009.

[12] R. Zhu, Y. Tan, Q. Zhang, Y. Li, and J. Zheng, "Determining image base of firmware for arm devices by matching literal pools," *Digit. Invest.*, vol. 16, pp. 19–28, Mar. 2016.

[13] X. Zhang *et al.*, "Cryptographic key protection against FROST for mobile devices," *Cluster Comput.*, vol. 20, no. 3, pp. 2393–2402, 2017.

[14] Y. Jouihri, Z. Guennoun, Y. Chagh, and D. Zahi, "Towards successful VoLTE and VoWiFi deployment: Network function virtualization solutions' benefits and challenges," *Telecommun. Syst.*, vol. 64, no. 3, pp. 467–478, 2017.

[15] S. H. Sellke, C. Wang, S. Bagchi, and N. Shroff, "TCP/IP timing channels: Theory to implementation," in *Proc. INFOCOM*, Apr. 2009, pp. 2204–2212.

[16] X. Zhang, Y. Tan, C. Zhang, Y. Xue, Y. Li, and J. Zheng, "A code protection scheme by process memory relocation for Android devices," in *Multimedia Tools and Applications*. New York, NY, USA: Springer, 2017, doi: 10.1007/s11042-017-5363-9.

[17] J. Zheng, Y. Tan, Q. Zhang, X. Zhang, L. Zhu, and Q. Zhang, "Cross-cluster asymmetric group key agreement for wireless sensor networks," *Sci. China Inf. Sci.*, vol. 61, no. 4, pp. 048103:1–048103:3, 2018.

[18] H. Zhu, Y. Tan, X. Zhang, L. Zhu, C. Zhang, and J. Zheng, "A round-optimal lattice-based blind signature scheme for cloud services," *Future Generat. Comput. Syst.*, vol. 73, pp. 106–114, Aug. 2017.

[19] Y. Xue, Y. Tan, C. Liang, C. Zhang, and J. Zheng, "An optimized data hiding scheme for deflate codes," in *Soft Computing*. Berlin, Germany: Springer, 2017, doi: 10.1007/s00500-017-2651-2.

[20] S. Cabuk, "Network covert channels: Design, analysis, detection, and elimination," Ph.D. dissertation, Eng., Electron. Elect., Purdue Univ., West Lafayette, IN, USA, 2006.

[21] Y. Liu, D. Ghosal, F. Armknecht, A.-R. Sadeghi, S. Schulz, and S. Katzenbeisser, "Robust and undetectable steganographic timing channels for i.i.d. traffic," in *Information Hiding*. Berlin, Germany: Springer, 2010, pp. 193–207.

[22] J. Wu, Y. Wang, L. Ding, and X. Liao, "Improving performance of network covert timing channel through huffman coding," *Math. Comput. Model.*, vol. 55, nos. 1–2, pp. 69–79, 2012.

[23] A. Houmansadr and N. Borisov, "CoCo: Coding-based covert timing channels for network flows," in *Information Hiding*. Berlin, Germany: Springer, 2011, pp. 314–328.

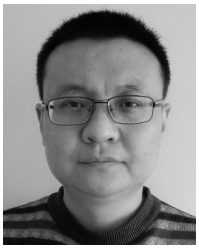
[24] S. Gianvecchio, H. Wang, D. Wijesekera, and S. Jajodia, "Model-based covert timing channels: Automated modeling and evasion," in *Recent Advances in Intrusion Detection*, vol. 8. Berlin, Germany: Springer, 2008, pp. 211–230.

[25] G. Liu, J. Zhai, Y. Dai, and Z. Wang, "Covert timing channel with distribution matching," in *Proc. Int. Conf. Multimedia Inf. Netw. Secur. (MINES)*, vol. 1. 2009, pp. 565–568.
[26] R. Archibald and D. Ghosal, "A covert timing channel based on fountain codes," in *Proc. IEEE 11th Int. Conf. Trust, Secur. Privacy Comput. Commun. (TrustCom)*, Jun. 2012, pp. 970–977.
[27] S. A. Ahmadzadeh and G. Agnew, "Turbo covert channel: An iterative framework for covert communication over data networks," in *Proc. INFOCOM*, Apr. 2013, pp. 2031–2039.
[28] W. Liu, G. Liu, J. Zhai, Y. Dai, and D. Ghosal, "Designing analog fountain timing channels: Undetectability, robustness, and model-adaptation," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 4, pp. 677–690, Apr. 2016.
[29] K. Ahsan and D. Kundur, "Practical data hiding in TCP/IP," in *Proc. Workshop Multimedia Secur.*, 2002, vol. 2. no. 7, pp. 1–8.
[30] R. Chakinala *et al.*, "Steganographic communication in ordered channels," in *Information Hiding*. Berlin, Germany: Springer, 2006, pp. 42–57.
[31] A. El-Atawy and E. Al-Shaer, "Building covert channels over the packet reordering phenomenon," in *Proc. INFOCOM*, Apr. 2009, pp. 2186–2194.
[32] *Wideband Coding of Speech at Around 16 kbit/s Using Adaptive Multi-Rate Wideband (AMR-WB)*, Standard G.722.2, 2003.
[33] J. Berger *et al.*, "Estimation of 'quality per call' in modelled telephone conversations," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar./Apr. 2008, pp. 4809–4812.
[34] R. Zopf, *Real-Time Transport Protocol (RTP) Payload for Comfort Noise (CN)*, RFC 3389, 2002.

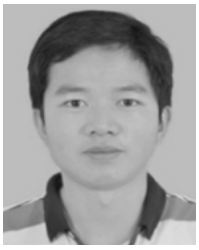
XIAOSONG ZHANG is currently pursuing the Ph.D. degree with the Beijing Institute of Technology. His main research interests include information security and mobile computing.

YU-AN TAN is currently a Professor and Ph.D. Supervisor with the Beijing Institute of Technology, and a Senior Member of the China Computer Federation. His main research interests include network storage, information security, and embedded system.

CHEN LIANG received the B.S. degree in EE from the Beijing Institute of Technology in 2014, where he is currently pursuing the Ph.D. degree. His main research interests include information security and coding theory.



YUANZHANG LI is currently a Lecturer with the Beijing Institute of Technology. His main research interests include mobile computing.



JIN LI received the B.S. degree in mathematics from Southwest University in 2002, and the Ph.D. degree in information security from Sun Yat-sen University in 2007. He is currently a Professor with Guangzhou University. He has been selected as one of science and technology new stars in Guangdong province. He has published over 80 research papers in refereed international conferences and journals. His research interests include applied cryptography and security in cloud computing. He was the program chair or program committee member in many international conferences.