Networking 2013 1569701137

Characterizing High-frequency Subscriber Sessions in Cellular Data Networks

Jingtao Li
Fudan University
Email: lijt@fudan.edu.cn

Wengang Pei
Fudan University
Email: 08302010024@fudan.edu.cn

Zhen Cao
Fudan University
Email: fdsscz@gmail.com

***Abstract*—Cellular systems operate under restrictive constraints of resources including radio channel capacity and network processing capability. The tremendous growth in the cellular data network usage brings operators with unprecedented signaling overloads and threatens the stability of the network. Understanding the characteristics of the subscribers who are resource inefficient has an important significance of capacity planning and optimal allocation of resources.In this paper, we perform the first large-scale investigation of session characteristics of a particular category of mobile subscribers, called high-frequency subscribers who access network frequently based on anonymized traces of an city-wide operational 3G network. We fund that high-frequency subscribers can be extremely signaling resource-inefficient. For subscribers who activate more than 10 sessions per hour, they only account for 0.5% of the total subscribers and generate about 1.73% data traffic but roughly 12.6% of the signaling resources consumption.**

We also propose a novel approach for discovering their session patterns and study the applications that generate such session patterns from semantic level. We observed that the frequency of subscribers' session activation shows positive correlation with the periodicity of session intervals. We also found periodic session activations have a certain correlation with abnormal behaviors, to our best knowledge, we first report that although TCP SYN flooding is detected and blocked by security measures, attackers may still send requests for context activation continuously, which adds unwanted signaling loads. We demonstrate that our findings have significant implications on network optimization.

I. Introduction

Data cellular networks have undergone tremendous growth in recent years due to the rapid increase in subscriber base size, cellular communication bandwidth, cellular device capability and mobile devices such as smartphones and tablets.

Cellular networks are originally designed mainly for voice communication. They assume dumb terminals, depending on the network for complex control ability. In order to send a packet, a typical data network (e.g. Internet router) simply forwards individual packets as they arrive, while a cellular data network interprets the first packet in a flow as an indicator of more traffic to come. Rather than simply forwarding that packet to its final destination, the network dedicates significant processing and bandwidth resources (connection establishment, etc.) to ensure that the end device is ready to receive data [26]. Such traffic patterns are unfriendly to the signaling resource control mechanism of cellular network carriers, and have put unprecedented pressure on signaling resources [18]. Due to an android VoIP application which generate huge levels of signaling traffic, DoCoMo suffered a major outage in its 3G network on January, 2012 [1].

This paper presents the first city-wide 3G operational network measurement of a specific category of mobile subscriber sessions, whom we named as "high-frequency subscriber sessions", Our study is motivated by the following two key observations.

First, data transfers produced by high-frequency subscribers can be extremely signaling resource-inefficient as they activate data sessions with high frequency but transfer few data in each session. In our collected trace, some sessions are activated with only one UDP packet transferred, which is extremely inefficient in the term of signaling consumption. To investigate the distribution of aggregate traffic, we observed that the subscribers, who activate more than 10 sessions per hour, only account for 0.5% of the total subscribers and generate about 1.73% data traffic but create 13.6% of the total sessions, and consume 12.6% of the signaling resources.

Second, high-frequency subscribers cause unfairness in charging. We observed that high-frequency subscribers produce few data traffic but have disproportionately high signaling overhead. However, network carriers generally charge their subscribers based on data traffic volume. High signaling resource consumption with few data transfers are unfair to other subscribers, which puts high signaling pressure on network operators but produces low fees.

In recent year, problems concerning signaling overheads of IP traffic in cellular networks become gradually concerned by researchers. Mao et al. characterized periodic-transferred IP packets in each data session and clarify that optimizing or batching these transfers can effectively reduce the consumption on network resources in [18]. They thought that mobile applications lack a thorough understanding of the RRC(Radio Resource Control) control mechanism, and as a result, intermittent packet transfers will cause lots of RRC state promotions and demotions, which is resource inefficient. For example, tail time of demotions will cause lots of device energy consumption [19]. Qian et al. estimated signaling loads based the RRC state transitions by measure the intervals of IP packet arrival collected from data plane in [11], and based on this estimation they calculate signaling overheads of common network applications. However, no characterization has been done for high-frequency subscriber sessions.

In this paper, we perform an in-depth and comprehensive study to quantitatively understand the following important characteristics from the session level:

- The regularity of subscriber's session activation;
- The correlation between the frequency of session activation and the periodicity of activation intervals;
- The impact on signaling resources of high frequency subscriber sessions for commercial cellular networks;
- The application-level semantics of high frequency subscriber sessions.

None of the four aspects was addressed by previous work, and we investigate all of them in this measurement study. They are important to study because the findings will provide insights on how to fundamentally eliminate resource inefficiency caused by high-frequency subscribers. We detail our key contributions as follows.

1. Data session level. Prior works analyzed based on packet or flow level, and they used an idle time (e.g. 5 minutes) to approximate the termination of a session. We extract the session accurately using the information collected from the control plane. We reconstructed complete data sessions by aggregating all Packet Data Protocol (PDP) control messages and IP packets in their own sessions. Then we studied session patterns in the consideration that data session maintenance is the basic goal of network signaling and can be used to capture the resource usage behaviors of mobile subscribers.

2. The frequency of subscribers' session activation shows positive correlation with the periodicity of session intervals. By clustering session activation interval of subscribers with DBSCAN algorithm [4] and classifying based on their cluster series, we observed that the intervals of sessions activated by high-frequency subscribers show some periodicity, and the more frequently subscribers activate their sessions, the more possible that their session intervals are periodic. However, by taking all the subscribers into consideration, we find that in most time, the session activation intervals (SAI) are in accordance to Poisson Distribution.

3. Exploration of application-level origins of high-frequency subscribers. Some subscribers activate data sessions at a fixed SAI repeatedly, which seem to be generated automatically by applications. By further analyzing IP packets of these sessions, we find they have certain correlation with abnormal behavior, such as continuously PDP context activation, TCP SYN flooding, periodical UDP packets, personal information uploading and fraud billing. Some other subscribers access the network at irregular SAIs but activate lots of data sessions. We find that a large fraction of them use always-on applications like instant messaging (IM) and social networking services (SNS), which ask the client devices to send heartbeat messages to the server every several minutes in order to remain online and thus bring a heavy signaling load.

4. Blocked SYN flooding still can bring unwanted signaling loads. To the best of our knowledge, we first report that although SYN flooding is detected and blocked by security measures, attackers may still activate PDP contexts continuously, which adds unwanted signaling loads.

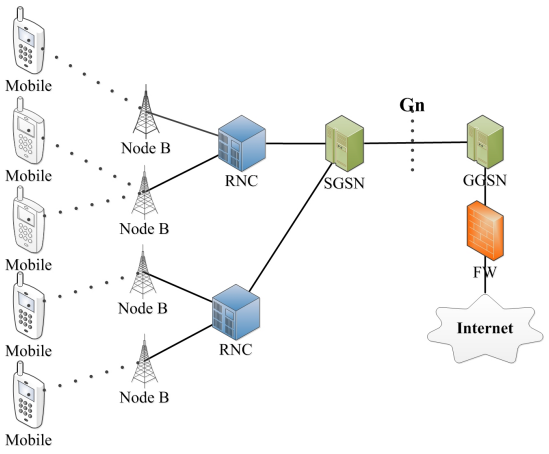

Fig. 1. UMTS network architecture

II. Datasets and Methodology

A. Datasets

Our traces are collected from the Data transmission link between SGSN and GGSN, namely the Gn interface, in a cellular operator's core UMTS network with its architecture as Figure 1, which services a large metropolitan with a population of more than one millions in China. Both PDP control messages between SGSN and GGSN for the initiation, termination and updating of the session (a period of continuous activity) and Tunneled IP packets between mobile terminals and the GGSN are collected. Our collection lasts for three days in January, 2010.

In our work, each data session is represented as a record which is a summary report of activities during one particular data session by one subscriber. Each record in the data set is indexed by a time stamp of session activation and an anonymized subscriber identifier (IMSI). One session record contains several control messages which can be used to determinate the start and end of the session. It also contains a set of IP packets during the holding time of the session and by extracting from those packets, we get some key fields, such as IP address, ports and compressed payload. Different applications are identified using a combination of port information, payload signatures, and other heuristics. More details about application identification are provided in [21].

Our dataset has the following limitations: 1) our datasets were collected in 2010. Given the dramatic expansion of the smart-phone market worldwide [22], we expect that there is a dramatic growth of the volume of data/control-plane traffic in 3G networks as of today. Also, there have been continuing version upgrades for smartphone operating systems (e.g., in Android and iPhone/iPad), and such upgrades may change the underlying data transmission behaviors. Nevertheless, since our methodology is based on the standard 3G specifications, it remains applicable for today's 3G networks in general. 2) Our collection lasted only 3 days. But our traces contain more than 910,000 sessions and 260,000 subscribers, which to our best knowledge is so far the relatively large cellular packet trace. As recently reported, traffic volume and behaviors of most application categories remains stable [15] and data traffic patterns are similar in each day during one week [17]. We

therefore expect our trace provides a largely representative snapshot of data traffic patterns. Also note that the trace duration is much longer than the session activation intervals under investigation.

B. Methodology

We primarily apply two methods in our study.

1) Methodology for Extracting Sessions: In this paper, we measure a data session from the successful activation of PDP context to the deletion of that PDP context. In order to accurately extract one session, we **restitute sessions by correlating the datasets that cover both data-plane and control-plane information**. The goal of our correlation is to identify the data/control traffic for each data session. We elaborate the details as follows:

1 We first extract the GTP(GPRS Tunnelling Protocol) signaling messages that are later used for reconstituting data packets in one session. For example, we will record TEID_Control (Tunnel Endpoint Identifier for control plane massages) in PDP Context Create and Update messages to collect GTP signaling messages in the same session and record TEID_Data for collecting data packets in the session later

2 Then we join the output record in the previous step with the data packets. The correlation is based on the TEID field, which identifies the data communication tunnel.

3 For one particular subscriber, the signaling messages in his sessions have the same anonymized IMSI. IMSI is identifier that identify the unique data sessions activated by the same subscriber. We will order all the sessions of the same subscribers according to the recorded time.

The methods proposed in [11], [18] extract sessions based on the same private IP address and a threshold of idle time to approximate the termination of a session. They mainly focus on how to estimate the signaling overheads brought by RRC state promotions and demotions in a single session, which are caused by the intervals between continuous IP packets. However, we extract sessions based on the TEID and IMSI, and by this method we can accurately identify the beginning and the end of a session. In addition, we can connect all the sessions belonging to one subscriber together, which will be used to analyze subscribers' session-level behaviors. Note that private IP addresses are dynamically allocated and normally one subscriber would get different IP addresses when connecting to the network twice, so methods based on IP addresses cannot be used to connect sessions.

2) Detect Periodicity in Session Activations: To detect the regularity of subscribers' session activation, we use DBSCAN algorithm and a novel classification method to analyze subscribers' data sessions from a time perspective. A se U of size high-frequency subscribers) can be represented as $U = \{u_1, u_2, ..., u_n\}$, and the i^{th} element of U means one of the high-frequency subscribers that we focus on. Here we use u_k as an example to illustrate our methods.

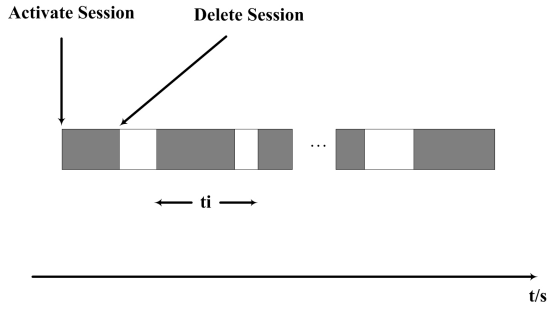

Fig. 2. Model of user session sequence

Reconstruct all data sessions initiated by u_k in chronological order and model a sequence of sessions as shown in Figure 2.

Definition 1 SAI: the abbreviation for session activation interval, which represents the interval between two adjacent session activation requests initiated by the same subscriber. Note that an SAI is different from a session interval which means the interval between the end of the last data session and the beginning of the next one.

Definition 2 T_k: A sequence of numbers $T_k = \{t_1, t_2, ..., t_n\}$ and the subscript k represents the subscriber u_k . The i^{th} element of T_k means the SAI between i^{th} and $i+1^{th}$ of data sessions which we defined as initiated by u_k . It can clearly represent the behavior of session activation.

Secondly, apply DBSCAN algorithm to cluster SAIs in T_k based on their values, and then map from the lowercase alphabet to these clusters in proper order as in **Fig.4**. Thus, we get a cluster symbol series $C_k = c_1...c_m$, and c_i represents the symbol of the i^{th} cluster of u_k, i.e., $c_1 = a$ and $c_2 = b$.

Definition 3. TC_k: the transition of T_k and sequence $TC_k = \{S(t_1), S(t_2), ..., S(t_n)\}$. Here the method S means the transition from t_i to the symbol of the cluster which t_i belong to.

Definition 4. Pattern: a particular repeated subsequence of TC_k. Different from other subsequences of TC_k, a pattern appears more times in TC_k and accounts for a higher proportion. It is also worth noting that a subscriber may have more than one pattern.

Thirdly, extract all subsequences of TC_k to find his patterns. i.e., if his subsequence is **"dbcb"**, then we retrieve all occurrences of **"dbcb"** in TC_k and calculate its ratio. If it's high enough to represent the TC_k (here, we set it to be 65%), then his pattern is **"dbcb"**, else consider his other subsequences.

Finally, in order to capture the regularity in each pattern (that is, we don't care about what the symbols are), we map from the capital alphabet to different cluster symbols of a particular pattern in a proper order. For example, suppose we have two patterns **"abca"** and **"dbcd"**, in this case, they are both mapped to **"ABCA"**, which indicates that they have the same regularity.

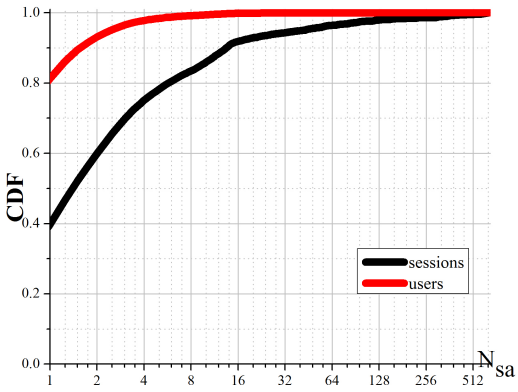

Fig. 3. CDF of users and sessions over N_{sa}

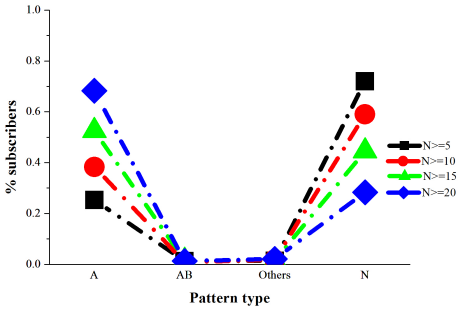

Fig. 4. Distribution of users over pattern types

III. Characterizing High-frequency Subscriber Sessions

A. Session Activation Frequency

In this section, we focus on the session characteristics of the subscribers with different session activation frequency. To characterize the session activation frequency for every subscriber, here we have done a classification operation. Let N_{sa} be the average session-activate times per hour for every subscriber, and subscribers with the same N_{sa} belong to the same category. Figure 3 plots the CDF of sessions and subscribers over N_{sa}.

We observed that the major subscribers (93%) have a N_{sa} less than 3. However, we still find that those subscribers just accounts for roughly 70% of the total sessions, suggesting that the other 7% subscribers have a high-frequency behavior to activate sessions, counting for approximately 30%. Furthermore, We observed that only 1% of the subscribers create nearly 20% of the total data sessions, suggesting that a few subscribers activate far more sessions than others. This shows a significant imbalance of network usage among subscribers with few subscribers hogging the much of the network resource, resulting in the unfairness of resources sharing.

B. Identify Session Activation Patterns

In this section, we study the session characteristics of the high-frequency subscribers. Using the method in section 2 to detect periodicity in session activations, we get tens session activation patterns. Here we choose some of them, which hold the largest proportion, and illustrate them in Figure 4:

We use "Pattern X" to represent the regularity of session activations extracted by our periodicity detection method. Session activation behavior of a subscriber follow Pattern A means more than 50% of his session intervals are in accordance to one periodicity, and Pattern AB means more than 50% of the session intervals are in accordance to two periodicities. "Others" means there exist more than two periodicities in these sessions. Some other subscribers follow no pattern and tend to frequently activate sessions irregularly, which is represented as Pattern N.

We observed that there is a correlation between the frequency of subscribers' session activation and the periodicity of session intervals. Higher frequency corresponds to a lower percentage of none-pattern traffic, which means the more frequently subscribers activate their sessions, the more possible that their session intervals are periodic. More than 40% of the subscribers with an N_{sa} value larger than 10 are in accordance to Pattern A,and for those with N_{sa} larger than 20, almost 70 % of them are Pattern A.

However, previous works on access behaviors of Dial-in Internet users propose that the access intervals are in accordance to Poisson Distribution [31]. Therefore, the session activation intervals of all the subscribers may also be in accordance to a Poisson Distribution.

To verify this assumption, We denote X(t)as the number of session activations observed in a duration of t, and we repeat such observations for m times. Let $A_i = X(t) = i, i = 0, 1, 2, ..., n$. Suppose the frequencies of event $A_0, A_1, ..., A_n$ in the m observations are $m_0, m_1, ..., m_n$ respectively, then we have $\sum_{i=0}^{n} m_i = m$.

Suppose the data are in accordance to Poisson Distribution, then we have

$$H_0 : P_i = P(A_i) = \frac{\lambda^i}{i!} e^{-\lambda}, i = 0, 1, 2, ..., n-1$$

$$P_n = 1 - \sum_{i=0}^{n-1} P_i$$

We apply the maximum likelihood estimation to the parameter λ

$$\lambda^* = \sum_{i=0}^{n} i * m_i$$

We can get an estimated value of P_i by using λ^*. Then we do a chi-square test to verify H_0.

$$X^2 = \sum_{i=0}^{n} \frac{(m_i - m*P_i)^2}{m*P_i} = \sum_{i=0}^{n} \frac{m_i^2}{m*P_i} - m$$

Here we use periods of 300 seconds, and the maximum number of session activations in 300 seconds is 112. Thus, the degree of freedom in our chi-square test is 111, and $X^2(111)$ is 136. The verification result is showed in Figure 5

We can see from the figure that most of values are smaller than 136, which means the periods are in accordance to Poisson Distribution. And for those periods not passing the tests, they are among the typical peak time of network traffic.

Then we examine which periodicity values are commonly used. Figure 6 plots the CDF of session activation intervals. The key observation is that several particular values dominate the intervals. We notice small clusters of 1 minute, 5 minutes,

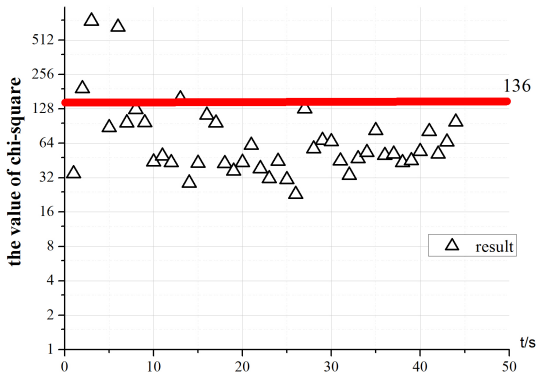

Fig. 5. Verification of Poisson Distribution

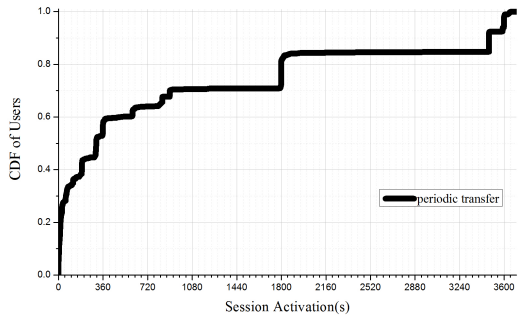

Fig. 6. CDF of session activation intervals

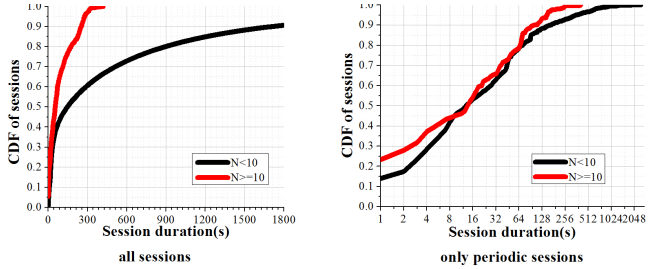

Fig. 7. Session Duration Distibution

30 minutes and 60 minutes. Such values are likely to be set by mobile application developers in an ad-hoc manner.

C. Session Duration

Fig 7(a) gives the CDF of session duration of high-frequency subscribers ($N_{sa} \geq 10$) and other subscribers ($N_{sa} < 10$). In (b), we only consider the periodic sessions. We find that high-frequency subscribers have relatively short session durations. For periodic sessions only, high-frequency subscribers still have shorter sessions durations, though the difference between high and low frequency subscribers is not so obvious as shown in Fig 7(a).

IV. Signaling Resource Impact

In this section, we use the large entire dataset described in Section 2 to study the signaling resource impact of high-frequency subscribers measured by N_{sa}. Here we totally study five sets listed in table I. U_0 is the all subscribers in the entire dataset; and U_1 to U_4 correspond to the subscribers with N_{sa} larger than 5, 10, 15, and 20 respectively.

TABLE I. Impact of high-frequency subscriber sessions.

Study scope	The contribution of high-frequency sessions			
	$\Delta V(Volume)$		$\Delta S(Signaling\ Overhead)$	
	All	Periodicity	All	Periodicity
$U_0 : all\ sessions$	100.00%	0.93%	-100.00%	-10.85%
$U_1 : N \geq 5$	3.82%	0.55%	-19.46%	-8.61%
$U_2 : N \geq 10$	1.66%	0.38%	-12.76%	-7.39%
$U_3 : N \geq 15$	0.52%	0.25%	-8.28%	-5.96%
$U_4 : N \geq 20$	0.30%	0.20%	-6.69%	-5.42%

We use the metric S to estimate the signaling overhead. It is quantified by the total signaling messages involved in creating and deleting a session. The signaling overheads consist of two parts: one part is the signaling overheads for radio resource control, and the number of signaling messages are estimated based on the signaling exchanges by the RRC state transitions [11]; the other part is the signaling overheads of PDP context control, and the number of messages in this part are counted directly from our dataset (and this part of signaling overheads are not taken into consideration by the previous works [11], [18]. Our estimation method has some limitations, because it neglects the signaling resources consumed in in the session holding time. For example, during session holding time, RRC state transition may still occur (i.e. $DCH \rightarrow FACH$), which brings a few round-trips of signaling messages. However, our method take into account the signaling resources consumed in RRC connection setting up (i.e., $IDLE \rightarrow DCH$) and the RRC connection release(i.e., $DCH \rightarrow IDLE$ or $FACH \rightarrow IDLE$). The recent cellular network measurement study [11] has demonstrated that the signaling messages of RRC connection setting up and release account for more than 60% of the total ones.

For subscriber set U_i, we define $V(U_i)$ as the total traffic volume consumed by these subscribers, and $S(U_i)$ the total signaling overheads brought by these subscribers' sessions.

For each set U_i, we first calculated its user subset UP_i detected in the real trace, which is the subscribers in U_i who have periodic session activation behaviors. Then we calculated the ratio of the traffic volume of U_i over that of U0 and the ratio of the traffic volume of UP_i over that of U_0 as:

$$\Delta V(all) = \frac{V(U_i)}{V(U_0)} \text{ and } \Delta V(Periodicity) = \frac{V(UP_i)}{V(U_0)}$$

The ratio of the signaling overheads brought by sessions of U_i over those brought by sessions of U0 and the ratio of the signaling overheads brought by sessions of UP_i over those brought by sessions of U_0 are also calculated as:

$$\Delta S(all) = \frac{S(U_i)}{S(U_0)} \text{ and } \Delta S(Periodicity) = \frac{S(UP_i)}{S(U_0)}.$$

We observed that as N grows larger, the traffic volume of high-frequency subscribers accounts for less proportion. But by removing these subscribers, the signaling overheads will reduce a lot. This trend is more obvious when applying it to periodic activated sessions. Clearly, there exists tremendous disparity between the traffic volume and the resource consumption of high frequent or periodical session activations,

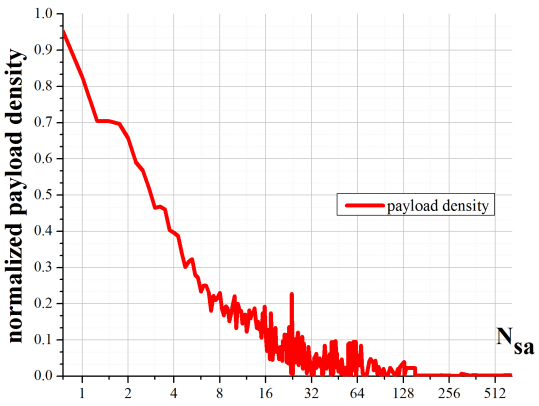

Fig. 8. Payload density distribution

indicating that high frequent or periodical session activations are extremely resource inefficient.

As an example, high-frequency subscriber sessions are responsible for only 0.55% of U_2, but their signaling overhead (ΔS) impact are 8 times higher. For periodic sessions in U_2, the impact are nearly 20 times higher.

To quantify disparity between the traffic volume and the resource consumption of high-frequency subscribers, we introduce the metric "payload density". Let C_{bytes} be the average payload size per session of a subscriber and $C_{signaling}$ be the total number of singling messages per session of the subscriber. We then compute the payload density, defined by $C_{bytes}/C_{signaling}$. Payload density is essentially one metric for measuring the effective data transfers per signaling message.Figure 8 plots the normalized payload density distributions of subscribers with different session activation frequency. We use the maximum payload density as the basis for normalizing. That is, normalized payload density of the subscriber $i = \frac{i's\ payload\ density}{maximum\ payload\ density}$. We observed that some active subscribers have an extremely high session-activate frequency, but a smaller payload density. The session activation frequency shows a negative correlation with the payload density.

From the operator's perspective, they charge only based on the traffic that subscribers have generated, and they prefer to the situation of lower signaling cost but higher traffic volume, however, those active subscribers generate just little traffic and cause significant signaling load. From the perspective of other subscribers, in the process of generating the same amount of data traffic, those subscribers consume more resources, showing significant unfairness of resource consumption.

V. Application-level Semantics

In this section, we will go further to study the application-level origins of high-frequency subscribers. To carry out a deep investigation, we use not only the categories of network traffic characterized by port number but also application layer headers to distinguish the traffic from different applications [21]. However, we find that large amounts of high-frequency subscribers' traffic use non-standard and custom application protocols, which makes it difficult to identify the application types by using common methods. That's why the categories of applications in our work are quite different from those in previous works. For those that tunnel their data through HTTP protocol (they hold the largest proportion), we also use the field in HTTP headers, such as HOST, UserAgent and URI, to identify application types. For those that cannot be characterized by a common port number, we check their packets manually.

TABLE II. Apps or behaviors in Pattern A

	users(%)	behaviors	users(%)
Periodical PDP context activation	16.0%	abnormal or suspected	40.35%
TCP SYN flooding	5.39%		
Periodical UDP packets	8.35%		
Personal info uploading	8.70%		
Fraud billing	1.91%		
SNS, IM, gaming, etc.	39.13%	normal but unwanted	53.04%
sync messages	13.91%		
unknown/other	6.61%		6.61%

The application behaviors mentioned below are extracted from the sessions of high-frequency subscribers ($N_{sa} \geq 10$) measured in section3 on the operational UMTS network. We respectively study the applications that trigger periodic session activations and the apps that generate non-periodic sessions (Pattern N) in these high-frequency subscribers.

A. Apps Generating Periodic Session Activations

We find a certain correlation between the subscribers periodic session activations and abnormal behaviors. The overview of these apps or behaviors is illustrated in table II.

1) Periodical PDP Context Activation: Some subscribers periodically initiate PDP context creation with a fixed time interval (from 1 second to 3 minutes), and then delete the PDP context soon. There is no User Plane data transmission (no IP packets), and the durations of these sessions are very short (within 3 seconds). Obviously, this would only cause the network to exchange signals continuously, wasting a lot of signaling resources without any actual utility (effective data transmission). In addition, the subscribers won't be charged with any fees based on data traffic accounting in this case.

2) TCP SYN Flooding: Some compromised mobile devices continuously send TCP SYN requests to the service ports of target hosts. As shown in Figure 9, we observe that these data flows have been successfully recognized by the Firewall or the intrusion detection systems between the GGSN and the Internet where all traffic is composed of IP packets, which causes the GGSN initiate deletion of PDP contexts and terminates the connections. Although these methods prevent the target hosts or the core network from suffering SYN flooding, but the over-consumption of signaling resources cannot be avoided. Attackers will continuously reactivate the PDP context and send SYN packets, while GGSN will delete these PDP contexts immediately. Back and forth like this, a lot of signaling resources will be consumed.

3) Periodical UDP Packets: After creating the PDP sessions, the subscriber sends UDP packets to the target host in a certain period (from 1s to more than 10s). The payloads of these UDP packets are mostly fewer than dozens of bytes and this type of sessions often has a relatively longer duration. As predicted in [13], periodical UDP packets may provide a way

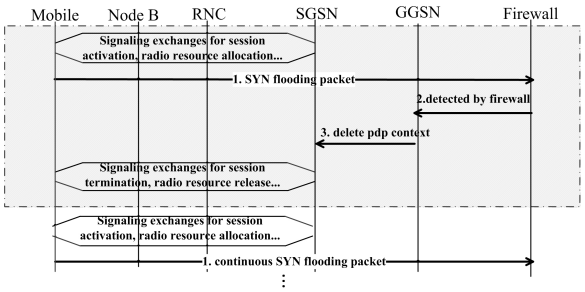

Fig. 9. Unwanted signaling loads of blocked SYN flooding

for the attackers to drain the battery power of subscribers' mobile devices by exploiting PDP context retention and the paging channel.

4) Privacy Information Uploading: The subscribers continuously initiate PDP contexts and upload privacy information, such as IMEI, phone numbers, etc. After finishing uploading, they disconnect themselves and after a while repeat these operations. These behaviors may threaten the privacy security and at the same time waste signaling resources.

5) Fraud Billing: A subscriber in this category continuously creates sessions, visits the same site and downloads large files. After further analyzing these sites, we find that their domain names are among those that are notorious for malicious consumption of user traffic, which brings users with lots of fees unwittingly.

To further clarify the aforementioned five categories of traffic behaviors, we still need more information and evidences to determine whether these session sequences are generated by intended DoS attacks aiming at overloading signaling resources (Although we are sure that some of them, e.g. SYN flooding, are definitely from compromised phones). However, according to recent studies on low-rate targeted DoS attack models [13], [26]–[28], these behaviors have many similar characteristics with these models (such as packet type, intervals between packets), and they do consume a lot of signaling resources.

6) SNS, IM, Gaming, etc: These types of applications hold the largest proportion of normal data traffic, most of which tunnel their data through HTTP protocol, and they are mainly always-on applications. These applications, such as instant messaging (IM) and social networking services (SNS), ask the client devices to send heartbeat messages to the server every several minutes in order to remain online. This also generates a heavy signaling load.

7) Sync Messages: Some mobile applications request clients to synchronize data with servers periodically (i.e. 5s), such as applications for time display or weather report. This makes subscribers create data sessions in a certain period, which consume lots of signaling resources.

In addition, the aforementioned apps or behaviors of a subscriber are partly relevant to the interval and duration of his session sequences, which is shown in Figure 10 in details.

B. Apps Generating Pattern N

Sessions that follow Pattern N can be generated by applications and network behaviors in various scenarios. An overview

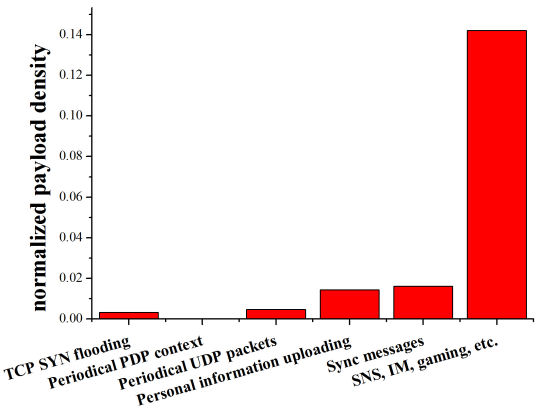

Fig. 10. Payload density of application behaviors

TABLE III. APPS OR BEHAVIORS IN PATTERN N

	users(%)
SNS, IM, gaming, etc.	61.09%
Blogging, online reading, serach,etc.	31.64%
Imei, uploading, etc	1.49%
unknown/other	5.78%

of these applications is given in Table III. In general, these subscribers' data sessions have longer duration than those of Pattern A. Also, more packets are transmitted in each session. We give two primary types of applications and user behaviors which belong to Pattern N .

1) SNS, IM, Gaming, etc.: One of the primary types is always-on applications. The reasons for repeated session activation have been analyzed above.

2) Blogging, Online Reading, etc.: This type of application- s waits a relatively long time before users' next operations on mobile devices. For example, in online-reading applications, two consecutive pages may be separated into two sessions, because users need some time to finish the current page. However, if there is no data transmission within a certain time (usually 3 to 10 seconds), smartphones or the network will automatically start dormancy mechanism to abort the wireless connection, which may cause the data session deactivated. The connection will be reestablished when the data transmission is needed. All these generate large amounts of signaling traffic.

VI. DISUSSION

Based on the previous characterization of high-frequency subscriber sessions, we examine its implication on cellular network operators to manage the signaling load from the perspectives of session activation frequency monitoring and abnormal traffic detection.

A. Subscribers Session Activation Monitoring

Session activation profiles for subscribers are necessary. Based on our study on the real trace from an operational 3G network, we found that high-frequency subscribers can be extremely signaling resource-inefficient as they activate data sessions with high frequency but transfer few data in

each session. As a result, it is necessary for cellular network operators to record subscribers' session activation profiles to distinct their session activation behaviors. We believe that this kind of profile will be one of the critical metrics for cellular service providers to improve the performance of resource allocation and security measurement. Those subscribers with a high N_{sa} value and low payload density value should be paid close attention to.

B. Abnormal Traffic Detection

In addition, **the enforcement of SGSN or RNC** is required for operators to act actively to recognize abnormal data traffic early and try to block them, which may reduce much unnecessary signaling consumption. Based on our studies mentioned in Section 5, abnormal data traffic have some relevance to the periodicity of session activations. The session activation profiles mentions above and the application-layer signatures can be used to detect these abnormal data traffic, which can contribute to the secure enforcement of network components such as RNC and SGSN. For example, we may collocate the detection elements with an RNC/SGSN for monitoring the session activations in each subscriber profile. When the N_{sa} of a subscriber becomes larger than the predefined threshold or the payload density becomes smaller than the threshold, countermeasures will be automatically taken to eliminate abnormal or repeated activations. How to determine to threshold value of N_{sa} and payload density to reduce false positive and false negative in detecting abnormal traffic is one of our future work.

However, **coordinating the RNCs/SGSNs with the Firewalls or intrusion detection systems** between the GGSN and the Internet is necessary for designing a more accurate detection mechanism. For example, the Firewall can notify the detection module in an enhanced SGSN the suspected subscriber who is carrying out a TCP SYN Flooding attack. When the activation attempts exceed the certain threshold, the SGSN will reject the subscriber's activation requests to eliminate the abnormal activations at an early time, which avoids allocating resources repeatedly for SYN flooding. In fact, it is not easy since the Firewall can only see the subscriber's IP address, but the SGSN does not work at the IP layer, leading to extra efforts for extracting IP packets from encapsulated packets. Practical solutions will be further studied in our future work.

VII. Related Work

The characteristics of mobile data traffic and its impact on network capacity, signaling cost, and data transmission have attracted attention in the industrial circle. For example, HUAWEI thought that it is the shortcomings in the design of smartphones or the characteristics of IP flows (bursty traffic, numerous address scans, etc.) that cause signaling storms and a differentiated resource and capacity management solution is needed in order to analyze network and discover resource bottlenecks in a timely manner [2]. SignalsResearch found that as networks become congested with mobile data traffic and the underlying signaling traffic inherent to the "chattiness" of numerous smartphone applications, user experiences can degrade to unacceptable levels [6].

To the best of our knowledge, our work is the first to study session characteristics of high-frequency subscribers based on the real traces of an operational 3G network. We demonstrate that they waste lots of signaling resources but with little effective data transmission and propose a novel approach for discovering their session patterns and study the applications that generate such session patterns. Our findings have significant implications on network optimization.

Complex and heavy signaling procedures render internet-connected cellular networks vulnerable to a variety of low-rate targeted DoS attacks [25]–[28]. The goal of these studies is to design or model attacks, such as exploiting target specific components of the expensive connection setup and teardown procedures to prevent legitimate use of data services [26], [28]. However, they are still hypothetical. We study the characteristics of high-frequency subscriber sessions based on data collected in operational networks, and we observe that some behaviors of high-frequency subscribers possess the similar characteristics of low-rate targeted DoS attacks. In the last two years, measurement studies of data traffic in cellular networks have created significant avenues for exploitations. Previous studies can be classified approximately into several categories.

One is from ISP's view point of modeling network traffic or architecture. Shafiq et al. modeled the traffic dynamics using flow-level data in [29] and provided a fine-grained characterization of the geospatial dynamics of application usage in a 3G cellular data network [24]. S. Das et al. analyzed subscriber mobility, temporal activity patterns and their relation to traffic volume, and described their implications in pricing, protocol design and resource and spectrum management [5]. Mao et al. characterized the cellular data network infrastructure of four major cellular carriers within the U.S and found that the current routing of cellular data traffic was quite restricted [12]. They also invent a tool that unveils carriers' NAT and firewall policies by conducting intelligent measurement [20].

Another category of works emphasizes on the user behavior or usage of cellular networks. In their another work, Mao et al. investigated the diverse usage patterns of Smartphone apps [15]. Keralapura et al. formulated the user behavior profiling problem as a co-clustering one to study behavior patterns" in [23]. Some works mainly analyze particular applications, such as the measurement of video applications [9], [10], while some focus on one characteristic of user behavior in cellular networks, such as the mobility, the relationship that exists between people's application interests and mobility properties. i.e., [7] and [8]. There are also several subscriber behavior studies based on deploying a custom logger on smartphones [16], [30]. Other examples include a study of the interaction between the wireless channels and applications [14], and performance of TCP/IP over 3G wireless with rate and delay variation [3].

There also some other studies focus on signaling overheads of IP traffic in cellular networks, which are the closest to our work. Mao et al. characterized periodic-transferred IP packets in each data session and clarify that optimizing these transfers can effectively reduce the consumption on network resources in [18]. They thought that mobile applications lack a thorough understanding of the RRC control mechanism, and as a result, intermittent packet transfers will cause lots of RRC state promotions and demotions, which is resource inefficient. For example, tail time of demotions will cause lots of device energy consumption [19]. Qian et al. estimated signaling loads

based the RRC state transitions by measure the intervals of IP packet arrival collected from data plane in [11], and based on this estimation they calculate signaling overheads of common network applications. However, no characterization has been done for high-frequency subscriber sessions.

Note that our work fills an important void in the space of measurement study of cellular data network by focusing on the characteristics and session patterns on one specific group of high-frequency subscriber sessions with high signaling consumption.

VIII. Conclusion

In this paper, we comprehensively characterized the session patterns of high-frequency subscribers of an operational cellular network in China. They consume much more signaling resources but have a lower utilization.

We unveiled several different session activation patterns, and one of the most surprising findings is that the frequency of subscribers' session activation shows positive correlation with the periodicity of session intervals. High-frequency session activations are generated for various reasons, among which always-on applications are the main cause. We also found that the periodic session activations have a certain correlation with abnormal application behaviors, such as periodical PDP context activation, TCP SYN flooding, and privacy information uploading, and the payload density of these applications is extremely low.

We believe our findings in characterizing the session patterns of high-frequency subscribers in cellular data networks directly have important implications on session activation frequency monitoring and abnormal traffic detection.

References

[1] *DoCoMo demands Google's help with signaling storm.* CAROLINE GABRIEL. January,2012.http://www.rethink-wireless.com/2012/01/30/docomo-demands-googles-signalling-storm.htm.

[2] Chen Yang. *Weather the signaling storm.* Huawei Communicate, 2011.http://www.huawei.com/au/static/hw-094153.pdf

[3] P. Benko, G. Malicsko, and A. Veres. *A large-scale, passive analysis of end-to-end TCP performance over GPRS.* In INFOCOM, 2004.

[4] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu (1996-). *"A density-based algorithm for discovering clusters in large spatial databases with noise".* In Evangelos Simoudis, Jiawei Han, Usama M. Fayyad. Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96). AAAI Press. pp. 226-231. ISBN 1-57735-004-9.

[5] Utpal Paul, Anand Prabhu Subramanian, Milind Madhav Buddhikot, Samir R. Das. *Understanding Traffic Dynamics in Cellular Data Networks.* INFOCOM, 2011.

[6] *The impact of mobile computers and smartphones on cdma2000 networks.* Signals Research Group, January 2011. http://www.cdg.org/resources/files/white_papers/mobile_computers_smartphones_on_CDMA2000_networks.pdf

[7] I.Trestian, S.Ranjan, A.Kuzmanovic, and A.Nucci. *Measuring serendipity: Connecting people, locations and interests in a mobile 3G network.* In IMC, 2009.

[8] Emir Halepovic and Carey. *Characterizing and Modeling User Mobility in a Cellular Data Network.* Williamson PE-WASUN 2005.

[9] Jeffrey Erman, Alexandre Gerber, K.K. Ramakrishnan, Subhabrata Sen and Oliver Spatscheck. *Over The Top Video: the Gorilla in Cellular Networks.* IMC 2011.

[10] Yuheng Li, Yiping Zhang and Ruixi Yuan. *Measurement and Analysis of a Large Scale Commercial Mobile Internet TV System.* IMC 2011.

[11] Li Qian, Edmond W.W. Chan,Patrick P.C. Lee, and Cheng He. *Characterization of 3G control-plane signaling overhead from a data-plane perspective* In MSWiM, 2012.

[12] Qiang Xu, Junxian Huang, Zhaoguang Wang, Feng Qian, Alexandre Gerber and Zhuoqing Morley Mao. *Cellular Data Network Infrastructure Characterization and Implication on Mobile Content Placement*. SIGMETRICS 2011.

[13] Radmilo Racic, Denys Ma, Hao Chen. *Exploiting MMS Vulnerabilities to Stealthily Exhaust Mobile Phone's Battery.* Securecomm and Workshops, 2006.

[14] Xin Liu, Ashwin Sridharan, Sridhar Machiraju, Mukund Seshadri and Hui Zang. *Experiences in a 3G Network: Interplay between the Wireless Channel and Applications.* Mobicom 2008.

[15] Qiang Xu,Jeffrey Erman,Alexandre Gerber,Zhuoqing Mao,Jeffrey Pang and Shobha Venkataraman. *Identifying Diverse Usage Behaviors of Smartphone Apps.* IMC 2011.

[16] Hossein Falaki, Dimitrios Lymberopoulos, Ratul Mahajan. *A First Look at Traffic on Smartphones.* IMC. 2010.

[17] Xiuqiang He,Patrick P. C. Lee,Lujia Pan,Cheng He and John C. S. Lui. *A panoramic view of 3g data/control-plane traffic: mobile device perspective* In IFIP, 2012.

[18] Feng Qian,Zhaoguang Wang, Yudong Gao, Junxian Huang, Alexandre Gerber, Zhuoqing Mao, Subhabrata Sen and Oliver Spatscheck. *Periodic transfers in mobile applications: network-wide origin, impact, and optimization* In WWW, 2012.

[19] F. Qian, Z. Wang, A. Gerber, Z. M. Mao, S. Sen, and O. Spatscheck. *TOP: Tail Optimization Protocol for Cellular Radio Resource Allocation.* In ICNP, 2010.

[20] Zhaoguang Wang, Zhiyun Qian, Qiang Xu, Z. Morley Mao, Ming Zhang. *An Untold Story of Middleboxes in Cellular Networks.* SIGCOMM, 2011.

[21] J. Erman, A. Gerber, M. T. Hajiaghayi, D. Pei, and O. Spatscheck. *Network-aware forward caching.* In Proceedings of the 18th international conference on World wide web, pages 291-300. ACM, 2009.

[22] IDC. *Worldwide Smartphone Market Expected to Grow 55% in 2011 and Approach Ship-ments of One Billion in 2015, According to IDC* http://www.idc.com/getdoc.jsp\?containerId=prUS22871611 ,Jun2011

[23] Ram Keralapura, Antonio Nucci, Zhi-Li Zhang, Lixin Gao. *Profiling Users in a 3G Network Using Hourglass Co-Clustering.* Mobicom 2010.

[24] Shafiq, M.Z., Lusheng Ji ; Liu, A.X. ; Pang, J. ; Jia Wang. *Characterizing geospatial dynamics of application usage in a 3G cellular data network.* INFOCOM 2012.

[25] P. Traynor, M. Lin, M. Ongtang, V. Rao, T. Jaeger, P. McDaniel, and T. La Porta. *On Cellular Botnets: Measuring the Impact of Malicious Devices on a Cellular Network Core.* In Proc. of CCS, 2009.

[26] P. Traynor, P. McDaniel, and T. La Porta. *On Attack Causality in Internet-connected Cellular Networks.* In Proc. of USENIX Security, 2007.

[27] J. Serror, H. Zang, and J. C. Bolot. *Impact of Paging Channel Overloads or Attacks on a Cellular Network.* In WiSe, 2006.

[28] P. P. C. Lee, T. Bu, and T. Woo. *On the Detection of Signaling DoS Attacks on 3G Wireless Networks.* In Proc. of INFOCOM, 2007.

[29] M. Zubair Shafiq,Lusheng Ji,Alex X. Liu,Jia Wang. *Characterizing and Modeling Internet Traffic Dynamics of Cellular Devices.* SIGMETRICS, 2011.

[30] H. Falaki, R. Mahajan, S. Kandula, D. Lymberopoulos, R. Govindan, and D. Estrin. *Diversity in smartphone usage.* In MobiSys, 2010.

[31] Ron Hutchins, Ellen Zegura, Oleg Kolesnikov and Phil Enslow.*Usage Characteristics of Dial-in Internet Users: A National Study.* 2001.