The effect of encrypted traffic on the QoS mechanisms in cellular networks

Chunshan Xiong and Milan Patel – Huawei

1 Introduction

Through this workshop we aim to explore how to solve the conflict between the encrypted IP traffic and radio network management, and endeavour to optimize use of radio resources and provide the required QoS to encrypted IP traffic. QoS handling is the most important part of the 3GPP radio resource management. 3GPP networks have limited radio and transmission resources and need to strictly schedule the utilization of radio and transmit resources using different granularity of bearers to provide and ensure Quality of Service (QoS) for the IP traffic. Different bearers with different QoS parameters will provide different QoS handling for the IP flows on each bearer. Different IP flows can share the same bearer; IP flows on the same bearer will receive the same QoS handling of the 3GPP network. With this binding mechanism, the 3GPP network can provide any IP flow with its required QoS handling. Therefore, the 3GPP network firstly needs to know the IP flow information and its QoS requirements. If this information is unknown, possibly as a result of encryption applied to the IP flow, the 3GPP network will discard this IP flow or handle the IP flow with default QoS.

We propose a paper to provide a detailed description of the QoS mechanisms of the 3GPP network and why encrypted IP traffic makes current QoS management mechanisms almost useless. Finally, we propose some ideas to solve this conflict to allow QoS mechanisms to be applied to encrypted IP traffic whilst maintaining the confidentiality of the IP traffic.

In this paper, we present the following content:

1) Why QoS is important for the wireless (3GPP) network and describe the QoS related mechanisms defined in 3GPP
2) The mechanisms to determine IP flow/traffic information used in 3GPP networks
3) The potential effects of encryption of IP traffic on 3GPP QoS mechanisms.
4) The possible methods that can be used by 3GPP networks to preserve the current QoS mechanisms.
5) Thoughts on future cooperation between mobile device, mobile network operator and 3rd parties/OTT (Over the Top).

2 The importance of QoS management for wireless network operators

The use of the cellular network by people and machines continues to rise. More and more services are carried on the cellular network and it has become one of the essential national infrastructures such as roads, bridges, electricity and tap water.

Initially, the cellular network set out to replicate communication functions similar to fixed line Public Switched Telephone Network (PSTN); later, adding the Short Messaging Service, and data communications services based on IP. As a principal of the network architecture, Governments have explicit service and security requirements for the cellular network, such as telephony, tele-conference call, emergency call (911) [1] similar to what the PSTN provides. In addition, some countries also require operators to provide eCall (for car accidents) [2], Public Warning System (tsunami, earthquake early warning) [3], Mission Critical Push-To-Talk [4] (cluster services for police or fire officers). Cellular networks need to support various kinds of fundamental services and a variety of value-added services and data services [5].

The characteristics and requirements of services and the scarcity of radio resource can affect the behaviour of the cellular network and end-user experience. Implementation of access control, scheduling and guarantee of resources is determined in advance of session establishment by QoS rules. For example, consider a telephone call and a file download simultaneously initiated by two different users in the same cell. Due to the scarcity of radio resource, the file download will seize communication resources and potentially impact the success of completing the telephone call. The network needs to ensure both the priority of resource allocation and the priority of resource use is higher for voice than for download. Likewise, as voice services have the same priority, the network needs to ensure that the radio resources during the call are always independently guaranteed, not shared. As a result, radio resources must be guaranteed for some services. However, for the services such as file download, Email, communication resources can be shared between services and users, resulting in little impact to the user's QoE. Similarly, in the event of public emergency, if the radio resources are insufficient (such as if base stations or communications equipment are partly damaged), the emergency services expect to be able to use the radio resources, thus causing the resources of ordinary users to be redirected. This shows that there are services with higher priority than normal voice calls.

In addition, in a wireless system, as a result of the user’s movement, a user can switch from one base station to another base station. Due to the handover, the radio resources in the target base station may not be the same as in the original base station. Resource reservation should be done in advance according to the QoS of a user’s ongoing services. If resource reservation cannot be realized, when the User Equipment (UE) switches to the new base station, it will cause interruption to all or part of the ongoing services as there is not sufficient resources, which degrades user experience.

The diversity of radio link and user mobility can cause handover between different wireless access technology which have different capacity and QoS characteristics. For example, the user may be switched from 4G to 3G (or 3G to 4G). If, for example, a user is on a phone call when the UE switches from 3G to 4G, it can modify its voice call to video call, or it can adjust its video call from Standard Definition (SD) to High Definition (HD). Or conversely, when the user switches from 4G to 3G, it can adjust its video call from HD to SD or even remove the video part.

Wireless operators charge for providing a variety of services through the cellular network [6]. As we all know, different quality versions of goods and services have different prices. Likewise, in cellular network, the price for services with different QoS requirements is different. In general, high QoS service charges higher in the transmission of a single bit compared with low QoS service. Previously, the charges for voice calls and SMS were higher due to their high priority and exclusive use of radio resources. Video download services were charged at a lower rate per bit, therefore, we often had to face the problems of interruption and poor quality when watching video streams. This requires that Wireless operators clearly know the QoS of user's service, duration of service using, location (e.g., roaming) and flow rate to charge different rates [7].

EPS QCI					Example Services
QCI value	Resource Type	Priority	Packet Delay Budget	Packet Error Loss Rate	
1	GBR	2	100 ms	10^{-2}	Conversational Voice
2		4	150 ms	10^{-3}	Conversational Video (Live Streaming)
3		5	300 ms	10^{-6}	Non-Conversational Video (Buffered Streaming)
4		3	50 ms	10^{-3}	Real Time Gaming
5	Non-GBR	1	100 ms	10^{-6}	IMS Signalling
6		7	100 ms	10^{-3}	Voice, Video (Live Streaming) Interactive Gaming
7		6	300 ms	10^{-6}	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)
8		8			
9		9			

Table 1. EPS defined QCI characteristics and corresponding Applications [7]

3 Policy and Charging Control (PCC) and QoS Management in Cellular network

In this section the Evolved packet System (EPS), the wireless system defined in 3GPP, and the QoS mechanisms used are described. For the sake of simplicity, the case for 3GPP S5/8 interface with GTP (GPRS Tunnelling Protocol) is discussed here as an example. The case for 3GPP S5/8 interface with PMIP is not discussed in this paper, but this does not affect the conclusion of the discussion.

3.1 IP Services, EPS bears and QoS

EPS provides different levels of QoS guarantee for IP services (see table 1). It is an indirect QoS guarantee mechanism. Any IP service can be identified by one or more Service Data Flows (SDFs) of the transfer data. A SDF can be identified by one or more IP Flow Filters, and a SDF is transferred through an EPS bearer. By implementing the QoS of EPS bearer, it can realize the QoS of SDF, and realize the QoS of IP services. The EPS bearer is one type of logical transport channel between the UE to Packet Gateway (PGW) (see figure 2).

Therefore, QoS guarantee of an IP service is converted into QoS guarantee of EPS Bearer. So, in the EPS, EPS Bearer becomes the minimum granularity of the QoS guarantee mechanism. A SDF is mapped to a specific QoS EPS Bearer, and SDFs associated with different IP services can be mapped to the same EPS Bearer with the same QoS parameters (namely QCI (QoS Class Identifier) and ARP(Allocation and retention priority)). That is, multiple SDFs mapped to the same EPS bearer will take the same levels of data processing. If two SDFs have different levels of QoS requirement, a separate EPS bearer needs to be established for each SDF. Different service data flows mapped to the same EPS must have the same QCI and ARP. The process to map a SDF with specific QoS to an EPS bearer is known as the binding process.

QoS parameters of EPS mainly include: QCI, ARP, GBR (Guaranteed Bit Rate) and MBR (Maximum Bit Rate). The mapping between QCI, ARP and IP service is given in table 1.

There are different service types, thus, according to service types that the bearer supports, the bearer can be divided into two categories: GBR bearer and non-GBR bearer. An EPS bearer is referred to as a GBR bearer if dedicated network resources related to a Guaranteed Bit Rate (GBR) value that is associated with the EPS bearer are permanently allocated (e.g., by an Admission Control Function in the eNodeB) at bearer establishment/modification. Otherwise, the An EPS bearer is referred to as a Non-GBR bearer. Each bearer, GBR bearer or Non-GBR bearer includes parameters: QCI and ARP

QCI is a scalar similar to the Differentiated Services Code Point (DSCP). It is a QoS parameter index in data forwarding of bearer level for specific access control node. The meaning of QoS parameters index is pre-configured to specific access nodes by operators. Its specific meaning is expressed by standardized characteristics (resource types, priority, packet delay and packet lost ratio, as per columns 2-4 of table 1). Bearer level data forwarding function includes scheduling weight, acceptance threshold, queue management threshold, protocol configuration of link layer, etc.

ARP is used to determine whether bearer creation/modification requests can be accepted, or used to refuse the request when resources are constrained which mainly occurs in situations where radio resource capacity is limited for GBR bearer. ARP can also be used to decide whether the bearer can be discarded when there is an abnormal limited resource, for example, in handover cases. However, when the bearer is completely established, the ARP parameter has no effect on data processing. ARP contains three pieces of information: priority level, pre-emption capability, pre-emption vulnerability. Priority level defines the relative importance of the resource request priority, which determines if the bearer creation/modification requests can be accepted or rejected. EPS defines 15 priorities, 1 as the highest priority, 15 as the lowest priority.

Each GBR bearer includes GBR and MBR. GBR indicates the bit rate the bearer can provide. MBR indicates the maximum bit rate the bearer can provide. For example, when performing the rate shaping function, additional data will be discarded. For a GBR bearer, MBR is greater than or equal to the GBR.

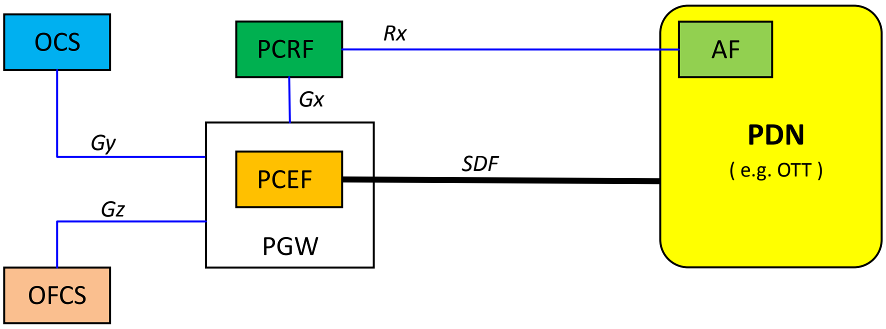
In accordance with the demand of the EPS, it is needed to provide "always online" service for users. So the bearer types are divided into two categories: the default bearer and exclusive bearer. Default bearer is an IP connection kept in a certain life cycle between the UE and Packet Data Network (PDN) once EPS bearer between the UE and the PDN is established. The other bearers which connect to the same PDN (e.g. UE uses the same IP address) as the default bearer are called exclusive bearers. QoS parameters (QCI and ARP) of each bearer under the same PDN must not be the same.

According to services the EPS bearer can support, the service can be divided into GBR service (such as voice service) and non-GBR service (such as Web page download), shown in the last column of table 1. Since the default bearer needs to always keep an IP connection between UE and a PDN, the default bearer is a non-GBR bearer, which is always a non real-time service, and has the characteristics of best effort. Exclusive bearer can be either GBR bearer or non-GBR bearer.

3.2 Bearer binding mechanism

As mentioned above, QoS control is strongly associated with charging. PCC is the integration of QOS management and control and accounting management and control architectures. On-path PCC architecture model in EPS is the policy control model that is used in the S5/S8 interface with GTP transport protocol (EPS also defines an Off-path PCC architecture, which is not described here in detail).

The PCC architecture is shown in figure 1 below:



AF:	Application Function
PCRF:	Policy and Charging Rule Function
PCEF:	Policy and Charging Enforcement Function
OCS:	Online Charging System
OFCS:	Offline charging system
PDN :	Packet Data Network
PGW :	PDN Gateway

Figure 1. Simple On-Path PCC logical Architecture

The AF generates dynamic session information of the application layer and receives specific information and notice of IP–CAN (IP Connectivity Access Network). AF is an entity in the service domain (such as OTT). Before launching a session, there should be signalling interaction between UE and AF in service domain. AF has to determine the session information in advance, namely the SDF of session media (i.e., one or more of the IP Flows) and media type of SDF and session information such as bandwidth requirements; the AF transfers the session information to the PCRF via the Rx interface.

PCRF is the strategy control function and charging control function based on flow and subscription data management function. PCRF is the policy center of PCC architecture. When PCRF receives session information sent by AF, PCRF decides PCC Rules for the service SDF according to the UE’s subscription information in Subscription Profile Repository (SPR), the operator's preconfigured strategy and the current wireless access technology which the UE is connected by. The PCC Rules include IP Flow Filters that identify SDF, QoS Rules and Charging Rules. PCRF sends PCC Rules to PCEF/PGW.

PCEF is located in PGW. PCEF is the execution node of PCC Rules. It has three important functions to perform Bearer Binding, SDF Detection and charging based on the flow. When PCEF receives PCC Rules from PCRF, it detects under the same PDN connection whether there is an

EPS Bearer that has the same QCI and ARP according to the QoS parameters (QCI + ARP) of the QoS Rules. If EPS Bearer with the same QCI and ARP exists, then the Traffic Flow Template (TFT) of the EPS Bearer can be modified. IP Flow Filters in the received PCC Rules should be added to the EPS Bearer TFT (if the EPS bearer is a GBR Bearer, the GBR, MBR parameters should also be modified). If EPS Bearer with the same QCI and ARP does not existed, the PCEF initiates to create a new EPS Bearer with the QCI and ARP, and adds IP Flow Filters in PCC Rules to the TFT of the new EPS Bearer.

3.3 SDF Detection

Each PCC rule contains a SDF, which defines the data detection parameters; each SDF of service data flow contains a variable number of IP Filters. The Filters are the combination of the various parameters of the following:

[1] Remote Address and Subnet Mask.
[2] Protocol Number (IPv4) / Next Header (IPv6).
[3] Local Address and Mask.
[4] Local Port Range.
[5] Remote Port Range.
[6] IPSec Security Parameter Index (SPI).
[7] Type of Service (TOS) (IPv4) / Traffic class (IPv6) and Mask.
[8] Flow Label (IPv6).

It is important to note that the each parameter above allows using wildcards.

A TFT in EPS Bearer is the collection of the IP Flow Filters of the installed SDF. Because there may be many SDFs and multiple EPS Bearers, each IP Flow Filter should have a priority. Therefore, when PGW/PCEF receives a downlink packet, it determines which SDF the received IP packet belongs to by matching the information of IP packets and IP Flow Filter in turn according to the priority. If the Gating state in PCC Rules corresponding to the SDF is open, IP packets are charged in SDF Charging Rules, and at the same time IP packets are transferred though determined EPS Bearer in binding process. Otherwise, if the status of Gating is closed, then the downlink data packet will be discarded. If received IP packets cannot be matched with any one IP Flow Filter, the packet will be discarded. SDF is single direction detection; therefore, it is direction independent. The downlink is performed in PGW/PCEF, while uplink is in UE. The IP Filter Rules of PCC Rules on UE is passed to UE by PGW.

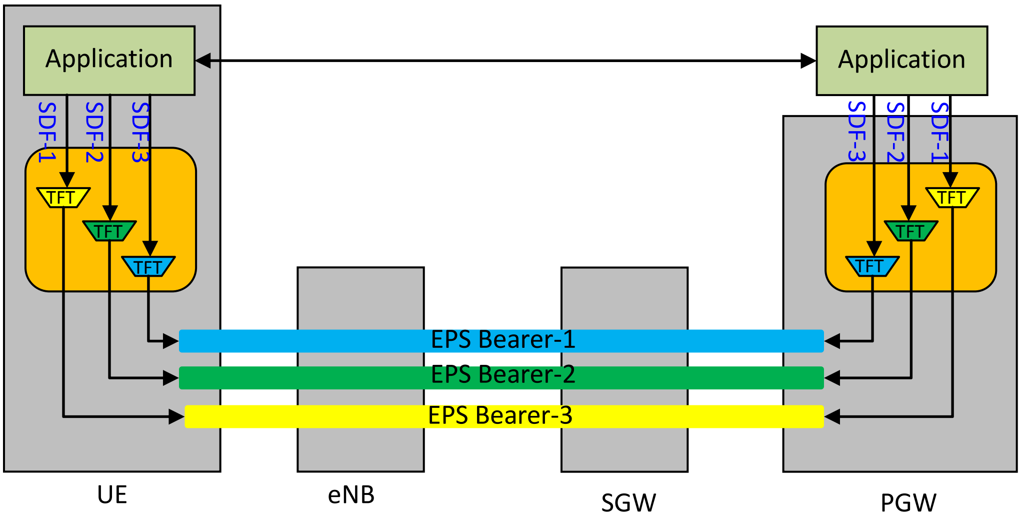


Figure 2. UL & DL SDF Detection and Binding to EPS Bearers

A point to note is the bearer with corresponding QoS should be established prior to the start of the session, and the establishment of the bearer requires signalling interaction (such as SIP signalling, or WebRTC signalling) related to session QoS provided at the beginning of the session. However, many Internet services (e.g., HTTP) have no such related QoS signalling interaction at the beginning of the session. This problem could be addressed by:

- Solution 1 is the use of default bearer mentioned above. Default bearer is created when the UE attaches to the cellular network.
- Solution 2 is to configure some static rules on the PCEF. These rules can be used to define the Filter based on parameters of the application layer (such as HTTP request or data types of HTTP). One or more exclusive bearers may be created according to these static rules, which are usually created after the default bearer is established. This requires a function similar to Traffic Detection on PCEF.
- Solution 3 is to using TDF/ADC (TDF: Traffic Detection Function, ADC: Application Detection and Control). The TDF can be independent of the PCEF and form a standalone node. If the node is not independent, the Traffic Detection function on the PCEF is called ADC.

3.4 Charging

PCC charging model supports the following models:
[1] Charging model based on volume;
[2] Charging model based on time;
[3] Charging model based on the combination of time and volume;
[4] Charging model based on event;
[5] No charging model.

The PCRF defines corresponding charging key, charging method and measurement method in SDF according to Charging Rules in PCC Rules defined by operators. Different charging keys reflect a specific service execution rate for the operator. The charging method indicates whether online billing or offline billing is used. Measurement method embodies the specified measurement methods adopted by the operator, such as based on volume, based on time, or both. PCEF completes the billing of users’ SDF according to the received Charging Rules and interaction with OCS or OFCS (Online and Offline Charging Systems). Charging control is based on the SDF granularity.

4 The influence of encryption on cellular network QoS management

After the detailed mechanism of cellular network QoS is well understood, it is easier to identify the encryption effects on cellular QoS mechanism.

In general, if the cellular network knows the SDF's IP Flow information, but doesn't know the content type of the transmission data and its QoS requirements, the SDF is usually mapped to the Default Bearer with the Default QoS to handle, or is mapped to a dedicated EPS Bearer with poor ARP with default QCI. If the network cannot know the SDF of one IP service in advance, the SDF of the IP service is also usually mapped to the Default Bearer with the Default QoS or is mapped to a poor ARP dedicated EPS Bearer with default QCI or is discarded because of the unknown service information of the SDF based on the predefined operators rules.

4.1 General impact on the cellular network QoS management from the Encryption

The cellular network supports all categories of services and at the same time, the cellular network is a radio resource-constrained network, unlike fixed broadband which always providing very high-rate access. In the case the wireless resources are abundant, the services transmission and QoS is good even without the cellular network's perception of the transmission content type. But when the wireless resource is in shortage, it is mandatory to execute the access control for some services and enforce the corresponding QoS scheduling. When SSL/TLS encryption is used， the IP 5-tuple information still can be known by the cellular network. If a combination service session has voice and video media at the same time, such as a video call, normally the voice and video media use different SDF, respectively; however, if encryption is used, the cellular network cannot perceive the different data transferred for voice and video if no especial heuristic tools are used and the voice and video media are normally mapped to the Default Bearer or a general dedicated bearer with poor ARP without mapping the different audio and video SDFs to different EPS Bearers. Generally, the cellular network will firstly ensure voice call's resources and scheduling in case of radio resource shortage; now the audio and video cannot achieve different ARP handling and QCI scheduling, which affects the voice part scheduling and further affects the quality of the voice call. At the same time, the network cannot provide the quality guarantee for the video part.

Likewise, when a Server provides different services (such as Web Email, or video/MP3 download and play) with the same IP address and same port number (e.g., 80) , the cellular network cannot learn the content type of the transmitted data and the transmitted SDF can only be mapped to the Default Bearer. Since most web pages are relatively small in size, the TCP with the Default QoS ensures an acceptable or good end user experience. However, with the same TCP connection of Default QoS (i.e., the same SDF) to transmit large volume of a video stream, it will result in a degraded end user experience, e.g. the video play is often stalled. If the cellular network can distinguish the transmitted content as large volume video instead of small size web pages, the base station can correspondingly perform some accelerated transmission in order to improve the user's video experience.

The cellular network normally maps the encrypted SDF to the Default Bearer or to a poor ARP dedicated EPS Bearer with default QCI. If the radio resource of the base station is plentiful, this mapping maybe will not introduce a big problem. In the case the UE handovers to a target base station, the target base station may lack sufficient radio resources, and the target base station will schedule the SDF with Default QCI; the radio resources do not match and the QoS requirements of the SDF are not met. Worse still, if the encrypted SDF is mapped to a poor ARP dedicated EPS Bearer, the SDF corresponding to dedicated EPS Bearer in the target base station is released during the handover procedure because the target base station cannot allocate the resources to the dedicated EPS bearer with the Poor ARP in case of insufficient radio resources. In this case, the connection for the SDF is broken. When the cellular network maps the SDF to a suitable dedicated EPS Bearer according to the services characteristics of the SDF (e.g. by the content type to SDF), the target base station can carry on the corresponding radio resource allocation based on the dedicated EPS bearer; if the radio resource of the target base station is insufficient, the dedicated EPS Bearer corresponding to less important SDF is released in order to ensure the EPS Bearer corresponding to the important SDF gets radio resource allocation. For example, for a video call during handover to the target base station with shortage of radio resource, the target base station usually releases the dedicated EPS Bearer corresponding to the video SDF and ensures allocation of radio resources to the dedicated EPS Bearer corresponding to the voice SDF.

Furthermore, consider when the UE moves to a new radio access technology. Different radio access technologies have different radio capacity and QoS characteristics, especially when the UE moves from a high bandwidth radio access technology to a low bandwidth of radio access technology, e.g. from 4G to 3G/2G; if the target radio capacity is limited, the cellular network must maintain the continuity of the critical service in the target side, and reduce the radio bandwidth of non-critical services, or even release the non-critical services. If the cellular network cannot perceive the type of the SDF, it may for example, release the user's key services and still maintain the non-critical services, e.g. closing the phone (VoIP) call service and keeping the file download service. When the UE moves from a low bandwidth into a high bandwidth radio access technology, usually all the critical and non-critical service can be maintained in the target side.

When the cellular network cannot perceive the encrypted content type, it cannot performance the optimization of transmission based on the content type and QoS guarantee. Thus, all services are handled with the unified Default QoS/QCI scheduling and guarantee, but since different services essentially have different network transmission requirements, this will damage services with strict QoS requirements, reducing the overall user experience. Additionally, as all the services have the same Default QoS/QCI, the operator will charge the user with the unified rate for all of the services. The single charging rule and single QoS for all the service will eventually hurt the user, operators and OTT service providers.

4.2 IPsec/VPN Tunnel-based IP layer encryption effect on QoS of network management

In this case, the internal real port number is invisible to cellular network and the tunnel-based IP traffic is usually mapped to the Default Bearer with Default QoS or to a dedicated EPS bearer with poor ARP and the same default QCI. If the VPN is from a big customer, the special tunnel-based IP traffics are mapped to a special dedicated EPS bearer with special QoS according the predefined rules and SLA. This might result in more dedicated EPS bearers with different QoS used to transport the different tunnelled-IP traffic with different QoS requirements.

4.3 WebRTC/IMS/SIP session type services encryption effect on QoS of network management

The cellular network can beforehand obtain the IP 5-tuple information of SDF of the voice, video and data (only for WebRTC) parts and the content type of each SDF during the Offer/Answer signalling interaction if the signalling connection between the WebRTC/IMS/SIP UA and WebRTC/IMS/SIP server is plaintext without encryption. Alternatively, the WebRTC/IMS/SIP Server or the AF in the server can actively tell the cellular network via the Rx interface to the PCRF all the voice, video and data SDF information even when the signalling connection is encrypted. Even if the transmission of voice, video media above the transport layer is encrypted, such as using SRTP, the cellular network can realize SDF detection and further can guarantee the SDF with the correct ARP and QoS control because the IP Flow Information is known by the cellular network beforehand.

If the cellular network cannot obtain prior SDF information on the voice, video and data part of the session because the signalling connection is encrypted and the server/AF does not provide the SDF information, if the voice and video use different IP flows, the cellular network still can identify the SDF type through using intelligent heuristic algorithms which can identify the difference content type by the transmission span of two successive packets, packet size and other information. After the cellular network identifies the SDF information of voice, video and other (data) parts, the cellular network can realize the corresponding QoS control and ARP and ensure the whole session's QoS.

4.4 HTTP encryption effect on QoS of network management

Currently HTTP 1.1 is the most widely used service/application protocol and it is expected to be widely replaced by HTTP 2 in the near future. HTTP supports transport of various types of data in a single TCP connection. Due to a single TCP connection corresponding to a single SDF, and different types of data and services are transmitted on the same TCP connection, the result is traditional SDF-based mapping SDFs transmitting different types of content/data to different EPS Bearers with different QoS and ARP no longer works well or is applicable for the cellular network. Instead, cellular network operators evolve and adopt new types of QoS-related acceleration technologies to realize and improve the user's experience. Therefore, Mobile CDN technology, Mobile Video Optimization technology, Mobile Web Optimization, Anti-Virus, Anti-Spoofing, Parent Control technology and all kinds of value-added technologies emerge and are widely used. These technologies can reduce the transport cost of cellular network and at the same time can greatly improve mobile user video and web browsing experience.

When HTTP2 and HTTP1.1 use TLS to encrypt the TCP connection, the widely used Web acceleration and value-added technologies no longer work well. The usual result is the HTTPS connection is mapped to the Default Bearer with Default QoS or dedicated EPS bearer with default QCI and poor ARP. Therefore, there is no guarantee for the different services provided by HTTPS websites. One exception is if there is a SLA/cooperation agreement, then the cellular network can map the TCP connection of the HTTPS website to a dedicated EPS bearer with special QoS, then the QoS for the HTTS website may be improve respectively with the special dedicated EPS Bearer and the specific QoS.

5 Possible approaches of perception the encryption contents

In Paper [14] and [19], many possible methods of perception of the content type of encrypted data are listed and analyzed.

Presented below are some key possible methods of inferring the content type of encrypted data without detailed technical description. An additional method is also proposed here. Further information can be found in the related reference documents.

5.1 No OTT participation methods

- A special certificate, as mentioned in paper [16], is issued to the cellular network operator in order to determine the content of encrypted HTTP URI without deducing the content of the encrypted HTTPS URI. This method has a lot of potential security issues.
- Certificate Spoofing method as mentioned in [17]. This method is similar to a Man-in-the-Middle Attack, in theory can work very well, but it introduces a lot of law/regulation issues and also can be used unlawfully if not carefully controlled. Thus, this method can only be used for limited controlled environment else introduces a big security risk, and requires the additional complexity of users to install special (Root) CA Certificate.
- The Captive Portal method as mentioned in [18], is also a type of Man-in-the-Middle Attack. Firstly, it needs to capture the user connection to the portal e.g. by DNS direction to the Captive Portal. If the original web page contains running codes (e.g. JavaScript), this approach will usually fail without changing the running codes. Modifying the running code to work well in Captive Portal mode is a great challenge.
- Intelligent heuristic (algorithm) method. By collection and convergence of the information of packet interval, packet size, port number, protocol type etc, the intelligent heuristic algorithm can guess correctly some the types of the content of the packet transmission as mentioned in previous chapter of WebRTC/IMS/SIP session type communication.

5.2 Methods requiring OTT participation

- DSCP method requires OTTs to set the right DSCP field of outer IP packet corresponding to different content types in the encrypted TLS connection. But it seems that OTTs won't always use it and furthermore, routers from the OTT to the cellular network may modify the DSCP settings.
- A new TCP option to identify the encrypted content type. This method has certain feasibility since it can normally pass through a lot of Middle-boxes, but OTTs are not willing to add this new TCP option.
- Operators open the PCRF and Rx interface to OTT and OTT provide the off-path content type information of the encrypted connection. This method is also unlikely to get OTT support.

5.3 Methods of industrial cooperation

- In this paper, the authors put forward a new method which modifies the up to down packet encapsulation of HTTP, TLS and TCP packet as described in figure 3. This method can be implemented in the mostly widely deployed Apache and or nginx HTTP Server package without destroying any current protocols. This method requires the OTT to deploy the modified Apache/nginx HTTP Server and an intelligent heuristic algorithm running in the cellular network to identify the dynamically changed content type of the encrypted HTTPS connection.

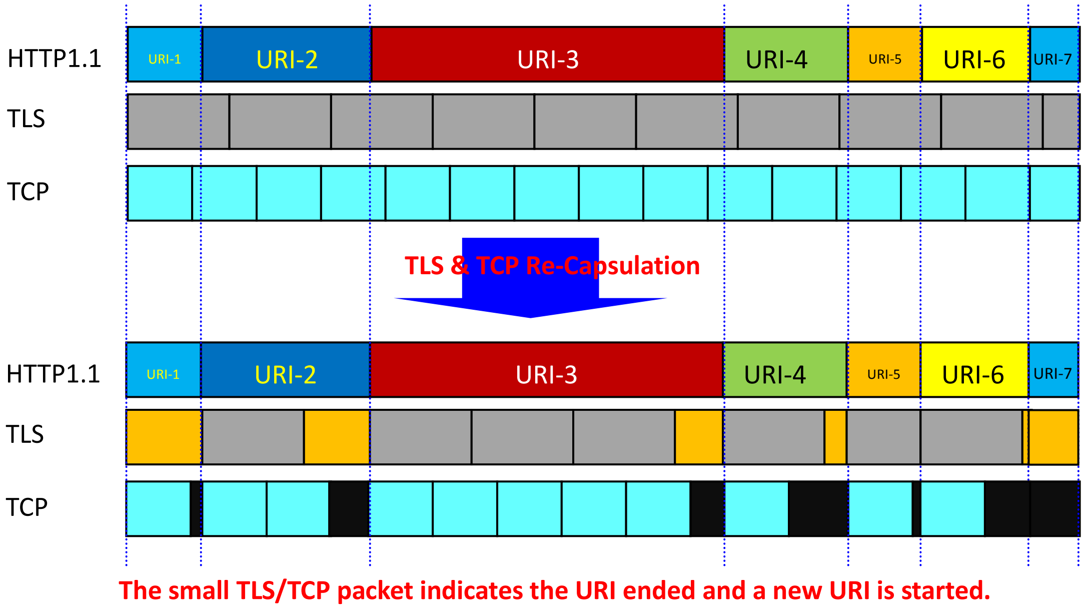


Figure 3. TLS & TCP Re-Capsulation for Identify the URI Content in the HTTP

- Operators cooperate with OTT, similar to the CDN Service Provider cooperating with OTT; this provides better performance for the OTT services and at the same time the operator can also determine some information about the encrypted content so that it can be managed correctly within the cellular network.

5.4 New encryption rules

- Similar to the MCIC (Multiparty Content Integrity and Confidentiality), OLC, SPUD (Substrate Protocol for UDP Datagrams), etc., let users, operators, OTT all benefit from the new encryption rules.

6 Conclusion

In this paper the importance of QoS in the cellular network service is discussed and the basic QoS management concept of the QCI, ARP, EPS Bearer, PCC, Bearer Binding, and SDF Detection in the EPS system is described. The influence of SDF encryption to the cellular network's QoS is discussed. At last, in the case of the encryption, some key content perception methods are listed and a new HTTP-TLS-TCP Re-capsulation method is provided to help the cellular networks to intelligently identify the content type of the encrypted data.

7 Reference

[1] 3GPP TS 22.101, "Service aspects; Service principles"

[2] 3GPP TS 26.267, "eCall data transfer; In-band modem solution; General description"

[3] 3GPP TS 22.268, "Public Warning System (PWS) requirements"

[4] 3GPP TS 22.179, "Mission Critical Push to Talk (MCPTT) over LTE; Stage 1"

[5] 3GPP TS 22.105, "Services and service capabilities"

[6] 3GPP TS 22.115, "Service aspects; Charging and billing"

[7] 3GPP TS 23.203, "Policy and charging control architecture"

[8] 3GPP TS 29.214, "Policy and charging control over Rx reference point"

[9] 3GPP TS 29.213, "Policy and charging control signalling flows and Quality of Service (QoS) parameter mapping"

[10] 3GPP TS 23.401, "General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access"

[11] IETF RFC 5246, "The Transport Layer Security (TLS) Protocol Version 1.2"

[12] IETF RFC 7230, "Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing"

[13] IETF RFC 7540, "Hypertext Transfer Protocol Version 2 (HTTP/2)"

[14] "Network Management of Encrypted Traffic, Version 1.0," 28 February 2015, GSM Association. Non-confidential Official Document WWG.04 - Network Management of Encrypted Traffic

[15] "A Rationale for Fine-grained Intermediary-aware End-to-End Protocols," draft-reschke-objsec-01,

[16] "Explicitly Authenticated Proxy in HTTP/2.0," draft-loreto-httpbis-explicitly-auth-proxy-01

[17] "SSL/TLS Interception Proxies and Transitive Trust," http://www.secureworks.com/cyber-threat-intelligence/threats/transitive-trust/

[18] "Captive Portal," https://en.wikipedia.org/wiki/Captive_portal

[19] "Network management of encrypted traffic," draft-smith-encrypted-traffic-management-02