Enhancing Mobile Network Efficiency - Recommendations for Devices

Version 1.0

16 July 2014

This is a Non-binding Permanent Reference Document of the GSMA

Security Classification: Non-confidential

Copyright Notice

Disclaimer

Antitrust Notice

Table of Contents

1 Introduction

1.1 Overview

Smartphone signalling can be very inefficient as standards leave room for different device implementations.

The continuous rise in smartphone and mobile data usage has also impacted the signalling load at the network level, independent of the volume of data traffic.

Issues experienced include Keepalives, Polling, Network Rejections and Infinite Signalling Loops. These issues, experienced across different operators' networks, can in turn negatively impact user experience, battery life, download time, network congestion etc.

This document addresses mobile network efficiency issues related to smartphones experienced by operators and manufacturers around the world, and makes a number of recommendations for device implementations to improve mobile network efficiency.

The main target audience for this document are device and chipset manufacturers. It will be up to each individual operator to adopt and enforce these requirements in their bilateral device requirements.

1.2 Scope

The document addresses issues experienced relating to:

- Signalling Loops and Aggressive Retries (chapter 2)
- Excessive Background Traffic (chapter 3)
- Other Inefficiencies (chapter 4)

The document outlines the problems experienced and the impacts they have on the network, device and user, and provides related recommendations for device implementations with the aim of enhancing the smartphone network efficiency. Consideration is also made for existing specifications and work on-going in other organisations such as 3GPP and OMA.

This document complements GSMA's published Smarter Apps Guidelines which gives background information and recommendations to application developers to improve apps connectivity, power consumption, user experience, security, and device battery life.

1.3 Definitions

Term	Description
ME Model	ME Model is defined as an ME which is different from other MEs (as per TS.06) a) By hardware design b) By transceiver c) By control software d) By frequency bands or e) As it is manufactured or assembled or commissioned by different ME Owners

1.4 Abbreviations

Term	Description
APN	Access Point Name
EMM	EPS Mobility Management
EPC	Evolved Packet Core
IMEI	International Mobile Equipment Identity
ME	Mobile Equipment
MNO	Mobile Network Operator
OMA AOI	Open Mobile Alliance Always Online Infrastructure
OS	Operating System
PDP	Packet Data Protocol
PDN	Packet Data Network
PRD	Permanent Reference Document
RAT	Radio Access Technology
SINE	Signalling Improvements for Network Efficiency
SNE	Smartphone Network Efficiency
TAC	Type Allocation Code
UE	User Equipment

1.5 References

Ref	Doc Number	Title
[1]	GSMA TS.06	IMEI Allocation and Approval Process v7.0 (31 October 2013)
[2]	3GPP TS25.331	Radio Resource Control
[3]	GSMA TS.18	Fast Dormancy Best Practices

2 Signalling Loops and Aggressive Retries

2.1 General problem statement

In many cases, devices are sending signalling traffic in a continuous loop, or aggressively re-sending signalling with no success or when it is not necessary. This can be due to open interpretation of 3GPP specifications, unexpected device response to network rejects, or applications making unnecessary requests to the OS.

2.1.1 Impacts on network efficiency

- Waste of network resources (reduces the real capacity of the network)
- Potentially creates signalling storm

2.1.2 Other impacts

- Unnecessary consumption of device battery energy
- Possibly reduced throughput or increased delay for other users

- Users may not have access to data services

Note: *GSMA SNE work in this area aims to change the relevant 3GPP specs e.g. TS 24.301 and thus this section is limited in this document. The SNE related work in 3GPP CT1 is under the "Signalling Improvements for Network Efficiency" (SINE) Work Item created in Q4/13 to be included in 3GPP Rel-12.*

2.2 Continuous device attempts to activate PDP context

The device continuously sends PDP/PDN attempts. This can be due to device response to reject codes sent by the network, no user plane or parameters, or a malicious application aiming to create a denial of service-like attack on the network.

ID	Recommendation
TS28_T1_1	If the device exceeds X (e.g. 5) PDN/PDP attempts within time period Y (e.g. 1 minute), the device SHALL cease to make any further PDN/PDP requests for the given APN for an increasing wait interval. This algorithm is to be implemented on the device.

Note*: Ideally, X, Y and "increasing wait interval" should be defined by the Mobile Operator in their Mobile Device Requirements issued to their device vendors.*

2.3 PDP context activation rejected due to failed user authentication

2.3.1 Context

PDP context activation initiated by the device and rejected by the network with cause value "User authentication failed". This happens occasionally in the existing networks when the end-user misspells his/her username and password, or makes other similar mistakes.

No clear device behaviour is defined after Reject Code #29 in the 3GPP specifications.

2.3.2 Impacts

The device may retransmit the PDP context, before any credentials are changed, therefore wasting network resources because the procedure will just fail again.

2.3.3 Solution

If the user is informed about the failed user authentication, he/she can try to input the correct username and password for the service, or otherwise achieve successful authentication.

ID	Recommendation
TS28_T1_2	Device SHALL inform the user about rejected PDP Context Activation due to failed user authentication.

2.4 LTE/UMTS Continuous Loop

2.4.1 Context

After receiving five Attach or TAU Reject messages with Cause #17 (Network Failure), the device moves to UMTS, however it will immediately retry connecting on LTE. The device is in a continuous loop, bouncing between LTE and UMTS.

The same device behaviour will occur for a device operating in CS/PS Mode 1 with “IMS voice not available” upon receiving five Attach or TAU Accept EPS Only with cause #16 (MSC temporarily not reachable).

2.4.2 Impacts

In addition to wasting its battery life, the device could be missing pages and will not have service while attempting on LTE.

ID	Recommendation
TS28_T1_4	After receiving five Attach or TAU Reject with cause #17 (Network Failure) and E-UTRA capability has been disabled, the device SHALL start a timer with a recommended minimum value of 12 minutes for enabling E-UTRA.
TS28_T1_5	After receiving five Attach or TAU Accept EPS Only with cause #16 (MSC temporarily not reachable) and E-UTRA capability has been disabled, the device operating in CS/PS mode 1 of operation with “IMS voice not available” SHALL start a timer with a recommended minimum value of 12 minutes for enabling E-UTRA.
TS28_T1_6	In addition to TS28_T1_4 and/or TS28_T1_5, an optional back-off algorithm MAY be implemented at the discretion of a carrier or a UE vendor where the back-off timer is extended if one or more subsequent consecutive attempt cycles also fail. The network operator shall be able to disable the optional back-off algorithm if provided.

3 Excessive Background Traffic

3.1 General problem statement

Several long-term system capacity issues have been related to background data traffic from smartphones. This traffic is seen 24 hours per day, 7 days per week.

> ***Note**: In this context, "background" refers to actions which the phone performs when the user is not actively using the device, e.g. when the screen is off, e.g. keepalive pings, checking for updates, updating social media status, etc.*

Reduction of background traffic can substantially improve the battery life on the device without impacting the user experience.

3.1.1 Impacts on network efficiency

Unnecessary signalling, waste of network resources (reduces the real capacity of the network), signalling congestion, and potential network overload.

3.1.2 Other impacts

Unnecessary consumption of device battery energy, reduction of data throughput, increased delay for access by other users, and consumption of subscriber's data usage plan.

3.1.3 Solutions

Some solutions to these issues are described below. However, other possible solutions also exist.

3.2 Each application has its own Keep Alive

When several always-online applications run in the device at the same time, each application will keep sending a heartbeat message in accordance with its own defined interval time.

The interval time between heartbeat messages may vary between different applications. The device may send heartbeat messages more frequently than a single always-online application scenario.

The use of a Push Notification Server (PNS) in the operator network would remove the need for Keep Alive messages and the need for each device application to keep any TCP connection open as the server will “wake up” the device whenever a notification for it is received.

ID	Recommendation
TS28_T2_1	The device SHOULD use a Push Notification Server if available.
TS28_T2_2	The device MAY use an existing Push service (OMA Push, W3C Push API for Web Apps, or platform-specific push) combined with developer program promotion of Push methods.
TS28_T2_3	The device MAY use OMA AOI (Always Online Infrastructure) enabler: The AOI Client SHOULD have only one associated AOI Server at a time; • A UE compatible with OMA AOI Enabler SHOULD NOT allow Client Application(s) from associating with multiple AOI Clients at the same time. • An AOI Client MAY be integrated into the Operating System on the device, or integrated with Client Application, or installed independently, depending on the deployment policy.

3.3 Keep Alive interval is too short

The device sends frequent keep alive messages to OS push servers by using a fixed polling interval.

ID	Recommendation
TS28_T2_4	The device SHOULD automatically detect the operators’ TCP_IDLE value (NAT timers) when using push services.

	• This can be achieved by increasing the polling interval until a timeout occurs and then operating just below the timeout value. • The device should adapt to the new value as opposed to hard coding a polling interval on the device.
TS28_T2_5	Dynamic polling interval (ref. TS28_T2_4) is preferred. However, if a fixed polling interval is used, the device SHALL use 29 minutes as the polling interval.

3.4 Each application event results in a RRC transaction

Almost all applications running in the background send pulling notification messages and upgrade software versions in an asynchronous manner, waking up the modem each time. This generates excessive RRC signalling.

- Note: In this context, “background” is defined as scenarios and/or time periods when the end user is not expecting certain application data traffic activities.

ID	Recommendation
TS28_T2_6	The device MAY use an algorithm to aggregate and synchronize the traffic from background applications in order to reduce network signalling and device power consumption
TS28_T2_7	The device SHALL provide the end-user with the ability to activate and deactivate the algorithm (ref. TS28_T2_6) on a per application basis.

3.5 Excessive background traffic when user has not configured anything

Most smartphones send background traffic frequently. This occurs even if a user does not install any applications or does not change any configurations on the device.

ID	Recommendation
TS28_T2_8	Prior to operator customizations, the amount of data usage in KBs/day in the default configuration SHALL be documented by the OEM and provided to operator. (Note: beyond OEM’s basic configuration, the data usage by Operator apps is the Operator’s business decision)
TS28_T2_9	When a user powers on the device for the first time, the device SHOULD notify the user that any data usage and related costs incurred after device activation are the responsibility of the user and they should select their applications and configure their device carefully. Additional notifications may be required by the operator.
TS28_T2_10	Devices SHOULD be shipped with a system-level (OEM or Operator provided) app which can provide users with real-time and historical data usage per installed app and system service. If accurate battery life impacts can also be provided, the data should include that, per app.

4 Other Inefficiencies

4.1 LTE device not allowed LTE access, but still attempts to attach to LTE network

4.1.1 Context

1. LTE capable smartphone with a 2G/3G USIM card (no LTE subscription).
2. LTE capable smartphone when LTE subscription is not active.

At power on, the LTE device tries to attach to an LTE network, but the procedure fails with EMM 3GPP cause #15 (No Suitable cell in Tracking area). The device will re-attempt LTE if the cell belongs to a different Tracking Area.

This procedure takes around 20-30 seconds and in this interval the user is unreachable.

4.1.2 Impacts

The user is unreachable for 20-30 seconds while the user's device is re-attempting LTE resulting in a bad user experience.

ID	Recommendation
TS28_T3_1	Device SHALL include a Radio Access Technology (RAT) option on the Network Selection menu of the LTE device's User Interface (UI) which – when selected – allows the device to automatically select between 3G and 2G networks, but not LTE networks.

4.2 LTE device reports LTE capabilities even if it is not allowed to LTE RAN

4.2.1 Context

1. LTE capable smartphone with a 2G/3G USIM card (no LTE subscription).
2. LTE capable smartphone when LTE subscription is not active.

At Attach, the LTE device sends its capabilities to the network including LTE features supported (Information Element: "*EPC capability = 1"*) even if the subscriber is not allowed access to LTE RAT.

Consequently, if return to LTE feature is enabled, the device will be sent to LTE, even if it is not allowed, rendering the user unreachable for 20-30 seconds.

4.2.2 Impacts

Bad user experience.

The device will also be connected to EPC network, occupying Network Resources allocated to only LTE subscribers e.g. IP address, HSS licenses, traffic loading. This may also mislead traffic reports.

ID	Recommendation

TS28_T3_2	LTE capable device SHALL NOT send E-UTRAN and EPC capabilities to the network when the “Automatic 3G/2G RAT” option (ref. TS28_T3_1) is selected via the device’s user interface.
TS28_T3_3	When E-UTRAN and EPC capability is changed (e.g. when a user changes RAT selection from LTE to 2G/3G, or vice versa), the device SHALL perform Attach in order to ensure that network resources are assigned appropriately and to prevent issues with Inter-RAT transfers.

4.3 Fast Dormancy trigger

4.2.3 Context

The algorithm to trigger the Fast Dormancy [2] [3] procedure is not standardized and sometimes does not work efficiently in some devices.

4.2.4 Impacts

Some tests conducted have shown that the same application installed and run in different devices may generate a lot of network signalling due to the different Fast Dormancy Trigger algorithm. This has an impact on the battery life.

ID	Recommendation
TS28_T3_4	The Fast Dormancy algorithm SHOULD be triggered based on device data inactivity following suggested time parameters: For SCREEN OFF state: • 3 to 5 seconds (the specific value in range is to be defined by operator) for networks with PCH RRC State support (URA-PCH or Cell PCH) • 12 seconds for networks without PCH RRC State support (URA-PCH or Cell PCH) For SCREEN ON state: • 5 to 10 (the specific value in range is to be defined by operator) seconds for networks with PCH RRC State support (URA-PCH or Cell PCH) • Trigger disabled for networks without PCH RRC State support (URA-PCH or Cell PCH) The device should ensure that background IP or IMS data flows would not be suspended by the SCRI

4.4 Low rate of device software upgrades by user

4.2.5 Context

When a device is launched in a market, initial non-optimal features or software bugs may exist which impact network efficiency. These issues may be resolved by a new standards release and/or a new device software release. However, without the support of Firmware Over-The-Air (FOTA), the rate of software upgrades by users will remain low.

4.2.6 Impacts

Technical improvements do not always reach all devices since users do not have an easy method to upgrade device software.

ID	Recommendation
TS28_T3_4	The device SHALL support FOTA over: 1) WLAN 2) Cellular (configurable option on the device, could be zero-rated by the operator i.e. zero cost to the user)

4.3 Type Allocation Code (TAC)

4.3.1 Context

An operator's mobile network traffic analysis carried out for troubleshooting purposes showed a number of mobile devices with different ME Models having the same TAC code. This violates the requirements in GSMA TS.06 "IMEI Allocation and Approval Process" [1].

Based on TS.06, TAC and ME Model identify the features and relative behavior towards networks. Thus, any issue or inefficiency that comes up in Device-Network Interworking should be related to an ME Model and its TAC, provided that such identification is unique.

Provided that the TAC code identifies an ME Model uniquely, MNO can:

- Recognize the platform model and market name of the device
- Disable network features for devices which do not behave properly.

Below are some examples of TAC related issues found in live networks:

- Issue #1: Some devices do not work correctly during switching from radio state CELL-DCH to URA/CELL-PCH, impacting the user experience.

  ⇨ Operators can switch off relevant features only for those devices recognised by TAC in order to guarantee service continuity to the user.

- Issue #2: Some M2M devices do not work correctly with E-GSM despite what their product specification declares.

  ⇨ Based on TAC, operator can assign frequency carrier on GSM band permitting the M2M device to work.

- Issue #3: Some devices do not work correctly with A5/3 ciphering algorithm.

  ⇨ Based on TAC, Operator can choose to negotiate A5/1 for such devices

Provided that *The TAC code shall uniquely identify an ME Model,* MNO can:

- Make reports differentiating device type, features and relative traffic volumes
- Based on traffic report, the MNO can move towards strategic decision to handle its resources (IP pool Assignment, APN steering…)

4.3.2 Impact on network efficiency

From an operator's point of view, the duplicated use of TAC codes leads to wrongful use of network functions and inaccurate Network Traffic reports which rely on correct TAC codes.

In order to manage its network efficiently, the operator should be able to analyse network traffic correctly based on different device types, features and relative traffic volumes. If devices with different hardware platforms use the same TAC code, it is more difficult for the operator to manage network resources (e.g. IP Pool Assignment, APN Steering) optimally which leads to decreased network efficiency. For example, the operator cannot reliably disable network features for devices which behave improperly or switch off a feature for a specific ME Model avoiding a particular type of network inefficiency.

ID	Recommendation
TS28_T3_5	Every Device SHALL be associated with and identified by a unique Type Allocation Code (TAC) according to GSMA TS.06 [1].

Annex A Document Management

A.1 Document History

Version	Date	Brief Description of Change	Approval Authority	Editor / Company
1.0	28 May 2014	First published version of TS.28	PSMC/TSG	Lauri Söderbacka / Huawei

Other Information

Type	Description
Document Owner	Terminal Steering Group
Editor / Company	Lauri Söderbacka / Huawei Technologies (UK) Co., Ltd.

It is our intention to provide a quality product for your use. If you find any errors or omissions, please contact us with your comments. You may notify us at prd@gsma.com

Your comments or suggestions & questions are always welcome.