Forensic Analytics

Communications Data Experts

Briefing Paper

GPRS Billing:

Using CDR Data Evidentially

Version 3.1

23/01/2020

0058-BRF

Page left intentionally blank

Forensic Analytics has and will continue to take all reasonable efforts to ensure that the information contained in this material is accurate and up to date. Forensic Analytics Ltd’s responsibility for inaccurate or out of date information contained in this material is limited to the correction of such errors, Forensic Analytics Ltd will not be responsible for any loses (actual or consequential) that may result from such errors.

Contents

1. Introduction

The growing population of smartphone users and the increasing use of Internet-based services such as FaceTime and WhatsApp to supply basic communications functions have resulted in a decrease in the number of ‘normal’ communications events (voice calls, text messages, etc) recorded in disclosed CDRs (Call Detail Records). There has been an associated increase in the number of GPRS/mobile data events captured in billing data, which means that telecoms analysts and cell site experts are being forced to analyse and base their conclusions, in an increasing proportion, on less accurate GPRS/mobile data events.

1.1. Problem Statement

GPRS/mobile data connections work in a different way to traditional cellular voice or text connections and the associated billing data is consequently collected using different methods. These differences can lead to a degree of uncertainty being present in data CDRs in relation to the cell site that was in use at the time the CDR was opened.

The reasons for this uncertainty are bound up in the technical operation of cellular networks and are often not well understood by analysts, investigators or cell site experts.

1.2. Effects & Implications

The effect of the uncertainty that exists in relation to GPRS/mobile data CDRs, in particular in relation to the ‘start cell’ listed in a CDR and the correlation between that cell and the record’s timestamp, is that cell site conclusions based upon GPRS evidence often cannot be as definite as the conclusions that can be reached in relation to voice/text events.

Whereas, with a voice/SMS CDR, it is appropriate to reach conclusions along the lines ‘AT the start time of the record, the subject phone was within the coverage area of cell ID 12345’ – meaning that a definite correlation between the start cell and the start time is being drawn – it is less appropriate to draw the same conclusions based on some GPRS/mobile data events.

The implication of drawing incorrect or unsupported conclusions from GPRS CDR evidence is that analysts or experts might attribute an incorrect location to the user of a mobile device at a point in time, if that conclusion was based on an incorrect understanding of GPRS billing data.

A converse effect is that uncertainty around the applicability of GPRS data causes some analysts and experts to avoid using it at all or to fail to use it to its maximum advantage. This may result in perfectly valid evidence and conclusions being omitted from an investigation.

1.3. Objectives

The objectives of this briefing paper are to outline the technical operation of GPRS and other cellular data services; to highlight the ways in which GPRS billing data is collected; to explain the ways in which ‘Start Cell’ (in the case of EE, Vodafone and one type of Three data) or ‘End Cell’ (in the case of O2) details are captured in GPRS CDRs; and to explain the ways in which this data can be most effectively analysed and understood.

The expectation is that after reading this paper, telecoms analysts will be more confident in their understanding of the operation of GPRS, the collection of billing data and the reasons for the uncertainty in relation to GPRS data. The expected outcome is that telecoms analysts will correctly interpret GPRS data and apply their understanding in their analysis and conclusions.

2. Mobile Data Overview

2.1. GPRS Concept

GPRS (General Packet Radio Service) was an enhancement to 2G GSM networks and began to be deployed in the UK from approximately 1999 onwards.

The original GSM networks offered only a dial up or ‘Circuit Switched’ service that supported voice, fax and data calling. A Circuit Switched (CS) connection occupies the assigned connection capacity continuously for the entire call, meaning that a mobile network must assign an entire radio connection to each call irrespective of whether the people connected via it are talking or not. CS connections occupy a fixed amount of network capacity and offer a very predictable level of service (based on the network delay or latency experienced by traffic travelling over the connection), which makes them well suited to carrying voice calls, but they can also be very inefficient (in terms of their use of network resources).

GPRS offers a ‘Packet Switched’ data service. Packet Switched (PS) services typically only occupy connection capacity when they actually have something to send or receive; this can make them unpredictable (in terms of network delay or latency) but they can also be very efficient (in terms of network capacity), which makes them well suited to carrying data traffic.

Most data applications – web browsing, email, app usage etc – have data to exchange only intermittently; a user might cause some data to be exchanged when they browse to a new web page but no further data might need to be sent or received while they are reading the content of that new page.

Strictly speaking, the term ‘GPRS’ (and the associated term ‘EDGE’) refers only to the mobile data service provided via 2G networks. In 3G, 4G and 5G networks the facility is simply referred to as ‘mobile data’ or ‘packet switched data’, but the term ‘GPRS’ has stuck and is used, incorrectly, as a coverall term for all mobile data services. Unless specifically indicated, in the following text the terms ‘GPRS’ and ‘PS data networks’ will be used to refer to all generations of mobile packet data services, from 2G to 5G.

PS data networks, like the Internet, accommodate the intermittent nature of most data applications by transmitting information in discrete ‘packets’, each of which carries the address of the device to which it should be delivered. If an application has a lot of data to send then it will be carried in a lot of consecutive packets, if they only have a small amount of data to send, the app will only transmit a few packets.

The intermittent nature of packet-based data communications is incompatible with the fixed capacity allocations of a circuit switched network, as the ‘circuit’ would need to remain assigned to the user even when they had no data packets to transmit and this could prove to be prohibitively expensive for the networks to facilitate.

2G GPRS and the PS data services offered by 3G, 4G and 5G networks were designed to allow data services to be handled more efficiently and economically by cellular operators.

2.2. GPRS/Mobile Data Network Elements

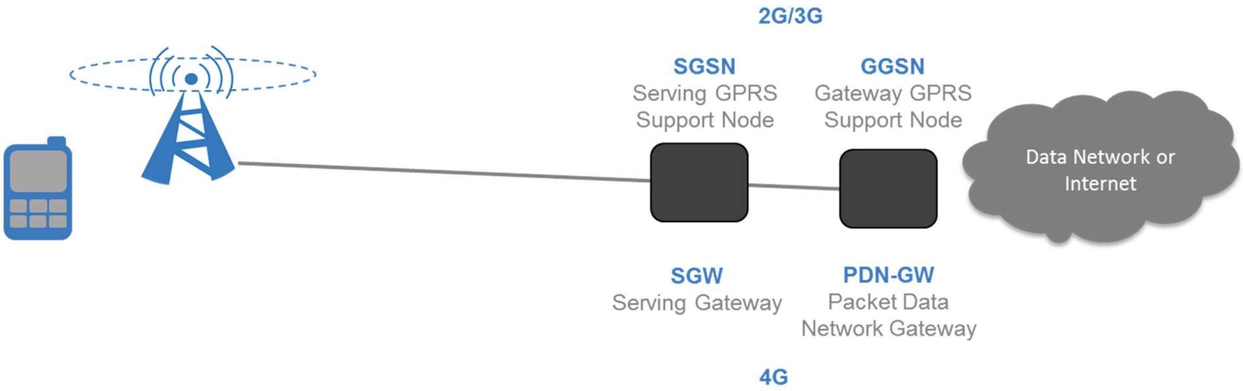
2G GPRS data services share the same radio channels, cells and access networks as 2G GSM voice services. GSM increases the carrying capacity of its radio channels by dividing each of them into a set of 8 ‘timeslots’ – up to 8 phones share the same radio channel concurrently, each periodically transmitting/receiving small amounts of data on their own timeslot in strict rotation.

In cells that support GPRS, some timeslots on some channels are given over to CS voice and others are used for PS data. The proportion of timeslots used for each service can vary depending upon demand.

The 2G BSC (Base Station Controller) handles traffic received on ‘voice’ timeslots and ‘data’ timeslots differently, forwarding voice traffic to the CS core network and data traffic to the PS core network.

3G/4G/5G radio channels do not employ timeslots but do provide shared access to the connection resources that they employ. The generic terms ‘connections’ or ‘sessions’ will be used to describe the PS data connectivity used by all generations of network.

The 2G/3G PS core network consists of two main elements known as the SGSN (Serving GPRS Support Node) and GGSN (Gateway GPRS Support Node).


2G/3G
SGSN
Serving GPRS Support Node
GGSN
Gateway GPRS Support Node
Data Network or Internet
SGW
Serving Gateway
PDN-GW
Packet Data Network Gateway
4G

The SGSN interfaces with the access network and manages the exchange of data packets with mobile devices. The SGSN receives data session set up requests from mobile devices and passes them on to the GGSN. Once a session is activated, the SGSN passes data packets in both upstream (from the mobile to the GGSN) and downstream (from the GGSN to the mobile) directions and keeps a count of the volume of data sent and received. The SGSN is also periodically kept informed of the cell that each active mobile device is currently using and takes note of any change of cell or Routing Area.

The GGSN is essentially a router that interfaces with external networks, like the Internet, and manages the flow of packets to and from those networks. The GGSN also maintains a count of the amount of data sent and received over each data session for billing purposes.

Each GGSN is able to connect to multiple external networks and incorporates a logical element known as an APN (Access Point Name) to handle the interfacing duties to each specific external network. The name assigned to the APN typically reflects the network or service to which it connects, so APNs of ‘Internet’, ‘MMS’ and ‘ims’ are typical.

The SGSN and GGSN generate CDRs (Call Detail Records) for each data session and pass them to the network’s billing system.

The 4G PS core network employs similar devices to the SGSN/GGSN, that are known as the S-GW (Serving Gateway) and the P-GW (PDN [Packet Data Network] Gateway) respectively. Connection control and subscriber management in 4G networks is handled by the MME (Mobility Management Entity). Billing data for 4G data sessions is generated in the S-GW and P-GW.

5G networks (that operate in Standalone Mode) employ similar methods for handling mobile data as 4G networks, although the names of the network nodes have changed again: instead of having different types of packet data handling nodes (e.g. SGSN/GGSN or S-GW/P-GW), the 5G core network has only one type of device known as the UPF (User Plane Function). Connection control and subscriber management are handled by a node known as the AMF (Access and Mobility Management Function). Billing data is generated in the UPFs.

2.3. GPRS/Mobile Data Operation

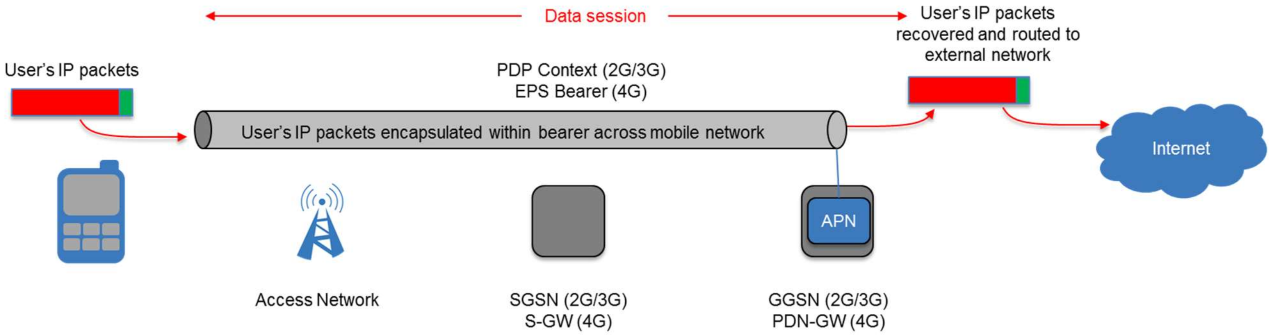
When a mobile device, like a smartphone or a tablet, needs to exchange data with a network or service, it asks the cellular network to set up a new data ‘session’.

In 2G and 3G networks the specific name given to a data session is ‘PDP Context’, in 4G networks it is termed an ‘EPS Bearer’ and in 5G networks it is a ‘PDU session’; but whichever terminology is used, the data session extends like a flexible pipe from the mobile device, through the currently serving base station via an SGSN/S-GW/UPF and terminates on a GGSN/P-GW/UPF, as shown in the diagram. These connections will simply be called ‘sessions’ in this document.

Each GGSN/P-GW incorporates a logical element known as an APN (Access Point Name) – the 5G UPF instead calls it a DNN (Data Network Name) – which acts as the gateway into the external network the session is being connected to. APN/DNNs are often given descriptive names, such as ‘internet’ and ‘ims’. A mobile device is permitted to have multiple concurrent data sessions established as long as they connect to different APN/DNNs – so a typical 4G phone might have one session established the ‘ims’ APN to carry VoLTE (4G voice) calls and a separate, but concurrent, session established the ‘internet’ APN to carry everything else.

Each session is assigned a temporary IP Address via the serving APN. These IP Addresses are typically assigned at random from a common pool of addresses, so the same IP address might over time be used by many different subscribers. However, all UK operators apply some form of ‘sticky’ IP address allocation, which ensures that a consistent IP address is assigned to a user over a short period of time, meaning that a subscriber might be assigned a random IP address when they first connect to the network on a given day, but will then end up using that same IP address over several sessions during the rest of that day, for example.

The ‘data session’ concept is employed for two reasons; firstly, it defines the path over which traffic for the mobile device should be routed. As mobile devices move around a network they will come under the control of a succession of base stations and core network nodes (although their data sessions will remain anchored on the same GGSN/P-GW) and an active data session will be re-routed to move with them, so that it realigns to deliver data traffic to the phone’s current cell. This ensures that data always travels from the anchor GGSN/P-GW, through the currently serving SGSN/S-GW, through the current BSC/RNC (for 2G/3G connections) and on to the current base station. The second role of the data session is to provide a logical connection between the mobile device and the GGSN/P-GW even when no physical radio connectivity is assigned.



Access to a physical radio connection in a cell is provided to a mobile device, to carry data for a session, if it requests a connection. The physical connection is released if the network decides that is has been unused for too long – this access network ‘idle timeout’ usually takes place if the data connection has been idle for a relatively short time of between 10 seconds (in 4G/5G) to a few minutes (in 2G/3G).

The logical data session is maintained even if the physical connectivity provided in a cell is released, so the logical session is independent of the physical radio connectivity provided by the cellular network.

In the same way that the physical connectivity in a cell is released if the mobile device has no further data to send, the logical session (the PDP Context, EPS Bearer or PDU Session) is released if the user has exchanged no data for an extended period – usually in the order of several hours. This process is controlled by an 'inactivity timer' or 'idle timeout' in the GGSN/P-GW which is started after each data packet is forwarded and which is stopped if a new data packet is exchanged. If no further data packets are exchanged by the time the session-level inactivity timer expires, the data session is released. This prevents unused sessions clogging up the network.

After a data session expires, if the mobile device finds that it has more data to send it simply requests the set-up of a new data session and the process begins again.

There are several levels of timer used to manage mobile data sessions:

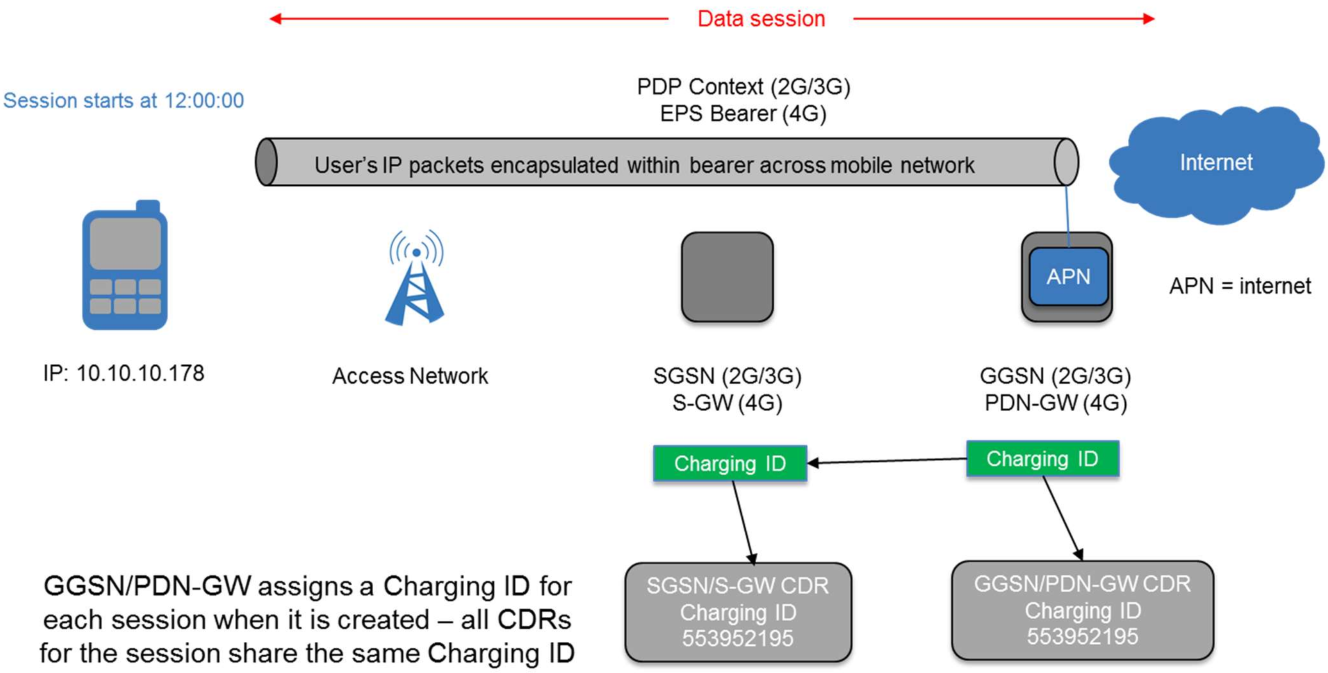
- The access network – via the BSC/RNC for 2G/3G services or the base station itself for 4G/5G connections – employs a timer to release radio resources when phones have been idle for too long. These timers typical operate on a scale of around 2-3 minutes for 2G/3G services and 10 seconds for 4G/5G services
- The CDR capture system in the traffic-handling nodes (SGSN/GGSN, SGW/PGW, UPF) employs a 'maximum open time' timer to control the length of time that each individual CDR file can be held open for; this is described in more detail below
- The data session management nodes (GGSN, PGW, UPF) employ 'session-level' inactivity timers which will release the entire session if it is unused for a period of time. Session timeout values employed by UK networks vary in duration from around 2 hours up to 6 or even 12 hours.
- The subscriber management nodes (SGSN/MME/AMF) employ 'periodic location update' timers. If idle an mobile device fails to check in by performing a location update when it is expected to, the network will perform an 'implicit detach', marking the device as unavailable and cancelling any active data sessions that it might have established. The CDRs associated with those data sessions get finalised at the same time.

It should be clear from the above that there are a number of factors that contribute to the number of CDRs that are generated in relation to mobile data connections and the period of time covered by each record can vary considerably.

2.4. GPRS Billing Data

Charging or billing data for a GPRS/PS Data session is captured at two points within the core network:

- The SGSN/S-GW, captures ‘open time’ and ‘close time’ timestamps, data volumes, serving (or at least, last known serving) Cell ID, LAC/RAC/TAC details and billing ‘event trigger’ information such as change of Routing Area
- The GGSN/P-GW, which also captures ‘open’ and ‘close’ times and data volumes, but which does not capture details of Cell IDs or access network events
- In 5G networks the UPF performs the roles of the SGSN/SGW and the GGSN/PGW, there can be expected to be two UPFs assigned to manage each data session – the UPF will be omitted from the following explanation to keep it simple, but it can be assumed that 5G billing works in much the same way as 2G-4G billing is described.



The request for a new data session or a new bearer within an existing session is always issued by the mobile device – the network might ‘page’ a device or issue some other form of alert that causes the device to request a session, but the session set up request always originates in the device. The request is passed through the network – via the SGSN in 2G/3G and the MME/S-GW in 4G networks – and is delivered to the GGSN/P-GW.

Authorisation for a new or amended data session is provided by either the GGSN/P-GW directly (this is the traditional method) or it is provided by the PCRF (Policy & Charging Rules Function) node in the network (this is the more modern method). Both methods employ a set of basic rules and permissions for data services decided by the operator.

When a new session is authorised, the GGSN/P-GW selects the parameters for the session and assigns it a Charging ID; all of which is then passed to the SGSN/S-GW and onwards to the mobile device to establish the connection. The fact that the GGS/P-GW and SGS/S-GW use the same Charging ID for each session ensures that CDRs from both devices for the same session can be correlated (or ‘mediated’) in the billing system.

2G SGSN CDRs and 4G S-GW CDRs capture the Cell ID of the current (or last reported) cell, whereas 3G SGSN CDRs capture the current (or last reported) SAC (Service Area Code). Most networks ensured that their 3G SACs were set to the same value as the corresponding Cell ID, meaning that the SAC in 3G CDRs pointed to the cell's Cell ID.

O2 originally made the decision to assign a SAC to each sector of each 3G site, each of which might be supporting two or 'stacked cells', and did not align these to the Cell IDs of the individual cells deployed on those sectors. This is why O2 3G CDRs traditionally did not identify the Cell ID on multi-cell sectors (instead listing the Cell ID as '-1' or 'unknown'). More recently they appear to have fallen in with general industry practice and have been assigning SAC values to individual cells that often match the cell's Cell ID.

2.5. CDR Files

A data session will remain established until either the mobile device explicitly releases it (e.g. because the user has closed that app that was using the session), the network explicitly releases it (e.g. because the mobile device has failed to perform an expected periodic update) or because the session is released because of 'session-level' inactivity.

The 'chatiness' of modern mobile apps, which keep up a regular flow of small message exchanges, means that, depending upon the mix of apps a mobile device is running and on the pattern of use by the device's user, explicit release and session-level idle timeout can be comparatively rare. This further means that some for some users, sessions can last for extended periods – several hours or even days – whereas for other users, with a different mix of apps, each session might only last for a matter of seconds or minutes.

For very long-lived sessions, it isn't practical for the network to keep a single CDR file open and collating billing data for an extended period, so the common practice is to regularly close the individual files and open new ones during the lifetime of a data session.

One data session might therefore be described by a succession of contiguous CDR records, all combining to show the overall resource consumption over the whole period that the session was active, but each individual record showing a subset of that overall billing picture. This practice is sometimes referred to as 'part billing', as each individual CDR record carries billing data for only part of a longer session. All CDRs that belong to the same session carry the same Charging ID.

In relation to part-billing, the network (usually via the SGSN/SGW) can end a currently open billing file (and immediately open a new file) for a variety of reasons, including:

- Maximum open time reached – networks can set a maximum amount of time that a CDR file can be open for, when this time is reached (for example, 2 hours), the current file is closed and a new one is opened
- Maximum data volume reached – networks can set a maximum amount of data that can be billed for in a single CDR file (for example, 20MB). Once this total is reached the current file will be closed and a new one opened
- Change of SGSN/SGW – a handover to a cell that is managed by a different SGSN/SGW will cause the old CDR file (held by the old SGSN/SGW) to be closed and a new CDR file to be opened (by the new SGSN/SGW)
- Change of Technology – this is really the same as change of SGSN/S-GW, as the change from a 3G bearer to a 4G connection will cause the 'old' 3G SGSN to close its current CDR and the 'new' 4G S-GW to open a new record
- Change of RAC (Routing Area Code) – this would occur when the mobile device detects that it has moved to a new Routing Area and transmits a RAU (Routing Area Update). The SGSN/SGW *could* close the current CDR file and open a new one containing the new RAC in these circumstances, but real-world testing has

shown that networks aren't configured to do this, instead just noting the change of RAC and Cell ID for future reference and not closing the current CDR file

- Change of tariff – if the user's tariff (e.g. the amount they are charged per unit of resource consumption) changes at certain times during the day, a new CDR will be opened to reflect the changed amounts to be charged. This isn't very common in reality
- Change of timezone – in large countries a handover to a cell in a new timezone could cause the old CDR to be closed and a new one opened, but generally only if the network employs time-based tariffs

In all of the above examples, even though the current CDR file is closed, the overall data session continues, so a new CDR file is opened to continue to collate charging information.

2.6. Cell Change Updates

In theory, every time a mobile device with an active data session is handed over to a new cell, the core network should be informed of that change so that the current CDR file can be updated. Mobile data CDRs capture updated details of the 'last used' cell as a session progresses and this information is used as the 'end cell' in the finalized CDR.

In practice, the core network elements that are compiling the CDR files (SGSN/SGW/UPF) are not informed of every change of cell that occurs.

In 2G/3G networks, the connection between the base station and the core network passes through the access network controller (BSC in 2G, RNC in 3G); these nodes deal with intra-RAC handovers locally and do not necessarily pass the details on to the core network and consequently the 'last used cell' information held in the SGSN does not always get updated.

In 4G/5G networks, although the traffic flow from a base station is exchanged directly with the SGW/UPF (there is no equivalent to the BSC/RNC in newer network types), signalling traffic travels via a different node - in 4G networks signalling travels via the MME (Mobility Management Entity) and in 5G it travels via the AMF (Access & Mobility Management Function). These nodes are informed of the current serving cell for each phone in every signalling message sent on from a base station and are also involved in handovers between base stations, but the MME/AMF will only pass 'access network' details on to the nodes that are constructing the CDRs (SGW/UPF) if they have been instructed to. These instructions are contained in the charging settings used when a new data session is being set up and could be different for different types of session. If access network updates are requested for a data session, the MME/AMF will pass on details of changes in cell, RAC, access technology type and timezone, but if those updates aren't selected then the S-GW/UPF will be updated as to changes of cell less regularly.

The fact that the core network node that compiles the CDRs for a data session is not necessarily kept informed of changes in serving cell for the associated mobile device is at the heart of the uncertainty that surrounds mobile data CDRs

It will rightly be pointed out that disclosed mobile data CDRs don't usually contain an 'end cell', so the lack of certainty associated with that item of information shouldn't be too much of an issue. However, in circumstances where 'part billing' takes place, when an 'old' CDR file is closed and a new one is opened immediately for an ongoing session, the 'end cell' captured in the 'old' CDR is used as the 'start cell' for the new CDR. This means that for 'follow on' CDRs the start cell might not actually be the cell that was in use at that time, it might merely be the 'last reported cell' that the network was last informed of – this idea is explored in much more detail below.

3. Call Detail Records

3.1. Voice/Text CDRs

Voice and text events typically have very distinct temporal identities, in the sense that there is an absolute start time and end time for a voice call and an irrefutable transmission or delivery time for a text event.

Call records for voice calls typically capture details of the cell ID used to set the connection up (the 'start cell') and also the cell ID that was in use when the call ended and the connection was released (the 'end cell') – they do not, unfortunately, (in UK billing at least) provide details of any 'handover' cells used in between those times and so are less useful for long duration calls than for shorter calls.

Call records for SMS events identify the cell ID that was used to transfer the text message. SMS transfer typically takes just a few tens of milliseconds and there is no mechanism to allow a handover to take place during an SMS event, so SMS records typically detail just the 'start' cell ID.

In the case of both voice and SMS events, cell site analysts can be quite certain that a target phone was within the coverage area of the cells listed at the start time for an event and (in the case of voice records) at the end time of that event. It is therefore possible for cell site reports to conclude, in relation to a cell that serves at a significant location, that use of that cell is consistent with the possibility that the target phone could have been 'at' that location at that time (with caveats).

3.2. GPRS/PS CDRs

GPRS/PS data CDRs can be far less definite in relation to the correlation of their timestamps and their start cells and are therefore potentially less valuable in terms of the cell site evidence they can provide.

A new GPRS/PS data CDR is opened when a mobile device establishes a new logical session with a data network, such as the Internet, and each new session is assigned a different Charging ID. Separate CDRs are generated by SGSN/S-GW and GGSN/P-GW nodes.

PS data sessions are 'released' if the mobile device explicitly indicates that the session is no longer required (see 'Session Explicitly Terminated' below) and the associated CDR files will be finalised and sent to the billing system.

In other scenarios, mobile data sessions will be released in a more implicit fashion due to long term 'user inactivity' or due to the mobile device failing to check in with the network at the expected time.

Mobile devices will be in one of the following states when attached for PS services:

- Standby (2G) or Idle (3G/4G/5G) Mode, where the device is attached and may have activated PS data sessions but does not have radio resources assigned to actually carry any data
- Ready (2G) or Connected (3G/4G/5G) Mode, where the device does have radio resources assigned and can transmit and receive data traffic.
- 3G and 5G networks also have intermediate states that are somewhere between 'idle' and 'connected' and that allow devices to drop and restore physical radio connections more quickly during periods of intermittent activity without fully releasing the associated data sessions.

Devices that are in Ready/Connected Mode are required to perform a ‘cell update’ or re-selection or handover when they change cells while involved in a data session. The access network will therefore know the cell that each connected mobile device is currently using but may not pass these details on the core network, which will continue to use the ‘last reported’ cell as the ‘end cell’.

Devices that are in Standby/Idle Mode are only required to update the network periodically or if they roam into a new Routing or Tracking Area. In between updates, the network only knows a device’s location down to the current Routing or Tracking Area, each of which will consist of a number individual cells and the network will not necessarily know which of these cells the phone is camped on at any point in time.

The effects of these procedural aspects of mobile data activities can be summarised into two main issues:

Issue 1 – Delayed Update of Current Cell Location

Theoretically, the core network node managing a data session’s CDRs will have a continuously updating record of the cell that an active phone is currently using; in reality, cell updates are less frequent and the value that the core network stores as the ‘current cell’ should be regarded as being more like the ‘last reported’ cell and can be out of date by many minutes or even hours.

The effect of the slow update of the current cell for a given mobile device is one contributing factor to the problems associated with the accuracy of GPRS/mobile data CDRs.

Issue 2 – Part-billing

Another contributing factor is related to the ‘part-billing’ that takes place in some scenarios.

GPRS CDRs can be closed without the associated data session ending; these cases are known as ‘partial records’ and are generated mid-session. Typical reasons for creating a partial record include (but are not limited to):

- Data volume limit reached – if the maximum data volume for a record is set at, for example, 20MB, then the current record will be closed, and a new record opened when the data throughput count (uplink + downlink) reaches that value.
- Time (duration) limit – if the maximum ‘open time’ for a record is set at 2hrs, as an example, the current record will be closed and a new one opened when the open time reaches this value.

When a CDR file is closed due to a part-billing trigger, the network node collating the CDR does not contact the associated device and therefore does not ascertain the device’s current cell location, it simply closes the current billing file and opens a new one using the ‘last reported’ serving cell, and this is the second contributing factor to the issues commonly experienced with GPRS CDRs.

Combining these Issues

When a part-billing event is triggered – when the current CDR is closed and a new, ‘follow-on’ CDR is opened – the network node collating the billing data (the SGSN, S-GW or UPF) is required to add a ‘start cell’ to the new CDR.

Because the node does not contact the phone in relation to the closure of the current CDR it is forced to rely on its knowledge of the ‘current cell’ the phone was using at the time that the new CDR was opened.

As the update of the current cell is often delayed, it means that the ‘current cell’ information held by the network is really the ‘last reported’ cell and could therefore be out of date by minutes or hours.

Mitigating these Issues

Due to a combination of the issues outlined above, the ‘start cell’ included in a follow-on CDR may not be the cell that the phone was really using when the new CDR was opened, but may instead be that last cell that the phone used that the network was updated with before the new CDR was opened.

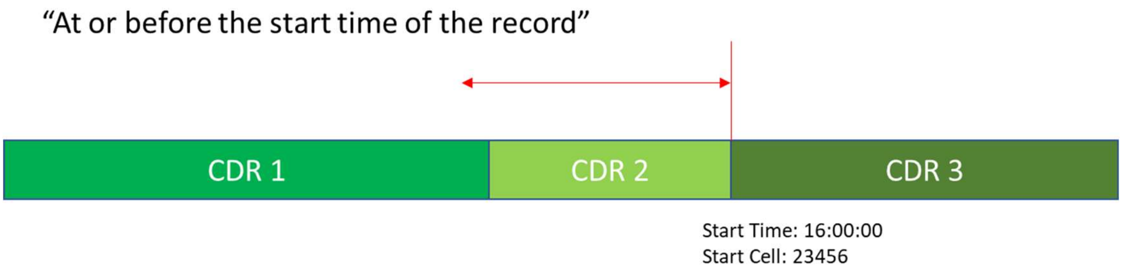
It is often not possible to tell if the ‘start cell’ in a follow-on CDR was the cell the phone was using ‘**at**’ the time that the CDR was opened or if it was a cell that was used sometime ‘**before**’ the CDR was opened.

Based on the issues outlined above, the UK College of Policing recommends that conclusions based on GPRS/mobile data events do not use the same terminology as used for voice/SMS events.

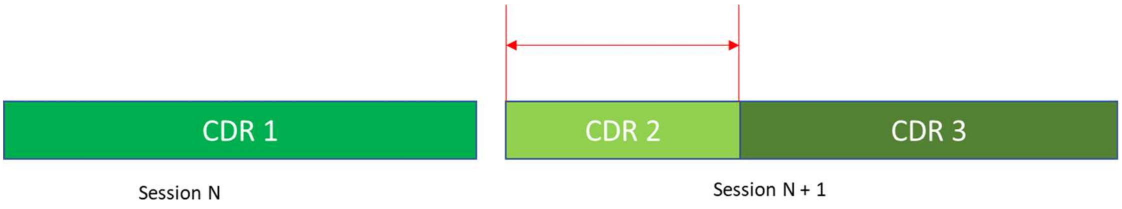
For voice/SMS events it is acceptable to draw conclusions that say: ‘the subject device was in the coverage of cell ID 12345 **AT** the CDR start time’.

For GPRS/mobile data events it is recommended that conclusions are more circumspect and say: ‘the subject device was in the coverage area of cell ID 12345 **AT OR BEFORE** the start time of the CDR’.

This construction of words covers the equally likely possibilities that the target phone could have been using the cell in question ‘at’ the start time of the CDR (i.e. that the ‘start cell’ was actually the current cell at that time) and that the ‘start cell’ had been used sometime ‘before’ the start time of the CDR (i.e. that the ‘start cell’ was actually just the ‘last reported’ or ‘last notified’ cell).



It must be stressed at this point that this issue of ‘at or before’ really relates to some scenarios of ‘follow-on’ records, where a new GPRS record follows on immediately after a previous one and where, crucially, both records belong to the same session and have the same Charging ID.



For the first CDR in a new session (with a new Charging ID) and some other scenarios of ‘follow-on’ record, it may be perfectly acceptable to trust the correlation between the start cell and start time and consequently use an ‘at’ conclusion. These scenarios will be explored in more detail below.

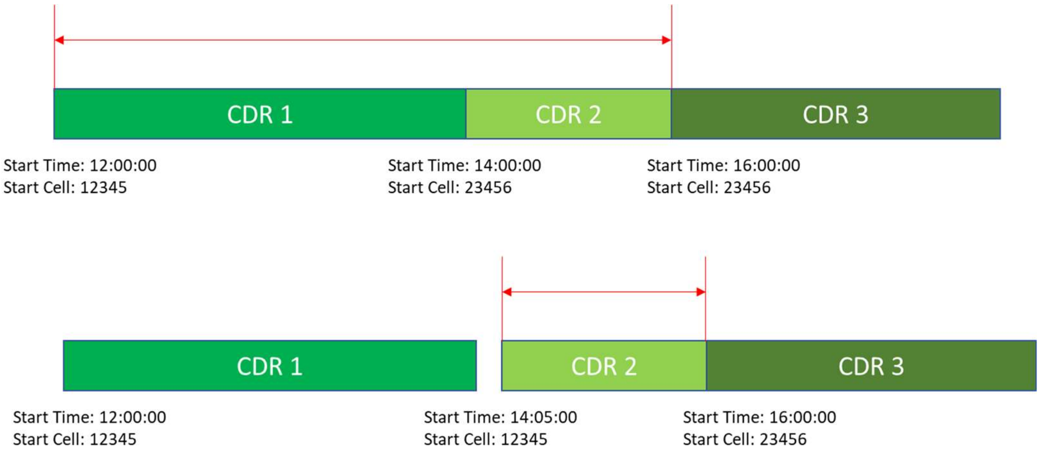
Constraining the 'Before' time

The phrase 'at or before the start time' creates the inevitable question of 'how long before?' – meaning 'how long before the start time of the CDR could the phone have been using that cell?'.

The answer to this is related to the nature of follow-on records; without wanting to be patronizingly simplistic about this, 'follow-on' records follow on from a previous record belonging to the same ongoing data session.

This means that if the 'start cell' for the record in question was used 'at or before' that start time of that record, it must have been used sometime during a previous consecutive record that covers the same data session. The furthest in time we can conceivably go back is to the start of the first record in that particular data session, which could be hours or days earlier. Realistically, however, the furthest we can go back is to the start of last, previous consecutive record that had a different start cell ID – the phone must have switched to a new cell sometime between the start of that record and the start of the record in question.

This means that we can modify the standard GPRS-related conclusion to add a constraint for how far back the conclusion stretches – the modified wording is: 'the subject device was in the coverage area of cell ID 12345 **AT OR BEFORE** the start time of the CDR *and after the start time of the last consecutive record that had a different cell ID'*.



Can GPRS events ever have an 'at' conclusion?

Before embarking on this next topic, it should be stressed that our recommendation is to stick with the 'at or before' conclusion when using GPRS/mobile data evidence if at all possible – it is far simpler and more consistent to apply the same wording to similar events.

However, there are times when an 'at' conclusion *could* be used for GPRS/mobile data events – the main example of this is the first record for a new session. To start a new session, the mobile device must have contacted the network to make the session request; the network knows which cell the device used to send the request and therefore captures an accurate start cell for the new CDR.

CDR records that show the start of a new session can be determined in two ways – for CDR formats that explicitly include the Charging ID (in the UK, this is Vodafone and O2, who include it by default, and also EE, where it was an additional field that could be requested, although there is now some doubt as to whether EE still offer this option), the first record with a new Charging ID will be the first record in a new session. For formats that don't explicitly show a Charging ID, the first record in a new session can be inferred from there being a gap between the start of the new record and the end of the previous data session – to cover situations where timestamps produced by different network nodes can be a little out of sync, we'd suggest that a reasonable gap between the end of one record and the start of a new one would be 10 seconds or more.

There are other scenarios in which the start cell of a record can be expected to correlate with the start time, these include change of SGSN/S-GW and change of technology – which are scenarios in which the device can be expected to have contacted the network to trigger the part-billing to take place.

How long do sessions typically last for?

Mobile data sessions can vary in length based on a number of factors, but the two main ones are:

- Whether the apps being used trigger explicit device-side session release
- Whether the device is ever inactive long enough to trigger network-side session release

The popularity of modern smartphones, with 'chatty' apps that send and receive small amounts of data almost continuously, means that many data sessions rarely trigger the 'user inactivity' process to cause session release.

From a forensic point of view, the chattiness of smartphone apps is potentially good news as more frequent contact with base stations means that the cell location data appended to the CDRs is more likely to be 'fresh', this is contrasted however with the fact that the duration associated with a GPRS CDR is likely to be longer for sessions that never trigger inactivity and that don't trigger part-billing (i.e. the device stays in one place and is in use very often) except for maximum open time/maximum data volume.

There are some circumstances in which 'chatty' apps – possibly WhatsApp or Facebook – are sending regular bursts of data but are also regularly informing the host smartphone that the connection is no longer required.

In these circumstances, the associated CDRs will show regular, short duration, standalone data sessions, each of no more than a few seconds or minutes in duration, with a few minutes of disconnection between them. These CDRs can often be almost as good as voice/SMS CDRs in the sense that, for non-follow-on records with short durations, the correlation between the start time and the start cell can be expected to be valid and usable. The problems associated with the interpretation and use of mobile data CDRs really related to follow-on records in sessions with longer durations.

Part-billing & follow-on records

If a data session remains active for an extended period, its current CDR file will eventually be closed due to the maximum data volume or maximum open time values being reached. When an existing CDR file is closed as a partial record, a new record will be opened immediately.

A CDR file could potentially also be closed mid-session (and a new record opened) if the mobile device roams into a new Routing Area, or hands over to a different access network type (e.g. 3G to 4G handover) or any one of several other reasons.

In the case of a Routing Area change, although the network specifications state that 'RAC Change' should be a trigger to close the current CDR file, testing of real world networks has shown that this doesn't happen in practice; that following a RAC change, the current CDR file stays open until a different trigger event (such as maximum open time) is detected. The details of the cell used for the RAC/TAC change will be captured, however, and might still be the 'last reported' cell reference retained by the network when the part-billing trigger occurs, so the cell used at the time of the RAC/TAC change might end up being used as the 'start cell' for a follow-on CDR sometime later.

Special Handling of O2 GPRS/mobile data event records

O2 GPRS records need to be handled slightly differently to those of other networks and the wording of the standard conclusion needs to be altered as well.

O2 admitted several years ago that, apparently due to an error in the way their CDR data is extracted from their billing system, instead of capturing the 'start cell' for each GPRS record they in fact capture the 'end cell'. This means that whereas each CDR purports to show the 'start time & start cell' it in reality shows the 'start time & end cell'.

The recommended wording for conclusions to be used for O2 data events is therefore changed to: 'the subject device was in the coverage area of cell ID 12345 **AT OR BEFORE** the **END** time of the CDR *and after the start time of the last consecutive record that had **the same** cell ID'*.

Special Handling of Three/H3 mobile date event records

Three disclose two different formats of mobile data record – IP Locations and LTE Locations.

IP Locations disclosures provide details of mobile data connections per target phone that are grouped into 15-minute periods. The billing system creates a record for each 15-minute period in which the target phone was connected for mobile data services. The record shows the cell ID used by the target phone during that period.

If the phone uses more than one cell in a period there will be multiple records, one for each used cell ID. If a phone continues to use the same cell in the following period there will be no record created, they are only created for the first period of continuous use in which the phone uses a particular cell. If the phone doesn't use mobile data during a period there will be no record created.

The evidence provided in an H3 IP Locations disclosure is therefore able to indicate the cells used by a phone to carry mobile data sessions, however, the fact that periods of continued usage of a cell aren't reported on means that it is not possible to tell the difference between periods when a phone is connected but hasn't changed cells and periods when the phone isn't connected at all. Testing by H3 has shown that the usage of a particular cell might happen slightly before the start time of the record.

The recommended wording for H3 IP Locations events is: 'the subject device was in the coverage area of cell ID 12345 **AT OR BEFORE** the **START** time and before the end time of the CDR*'*.

LTE Locations disclosures provide a more traditional summary of CDRs for mobile data sessions, with a record per session or a succession of follow-on records for a session.

The individual session records show start time and start cell, similar to other networks' GPRS CDRs, but they don't provide an end time, a duration or a Charging ID, so it's not possible to reliably distinguish between standalone and follow-on records, which in turn means that it is not possible to constrain the 'before' component as we can't determine contiguous records.

The recommended wording for H3 LTE Locations events is: 'the subject device was in the coverage area of cell ID 12345 **AT OR BEFORE** the **START** time of the CDR*'*.

4. GPRS Billing Scenarios

The following examples show the CDR data that is likely to be captured in a range of connectivity scenarios. Each example contains a diagram showing the movement of the subject phone, the cells it connects to and the times of significant events. There will also be examples of the information that would be captured in the resulting CDRs, which includes:

- Start Cell - which will be populated with either the cell the phone was connected to at the start of the CDR period or, in the case of some ‘follow-on’ CDR scenarios, the stored ‘last reported cell’ details for that phone
- Start Time – the time the CDR was opened
- Last Used Cell – the last known/last reported connected cell for the phone, as stored in the SGSN/S-GW
- Closing Time – the time this particular CDR file was finalised and closed
- Duration – the length of the ‘open time’ for the CDR file
- Volume – the aggregate amount of upstream and downstream data that flowed across the session during the CDR open time
- Cause – the reason why the CDR was closed – in the UK, only Vodafone routinely provide details of the ‘closure cause’ values in their CDRs

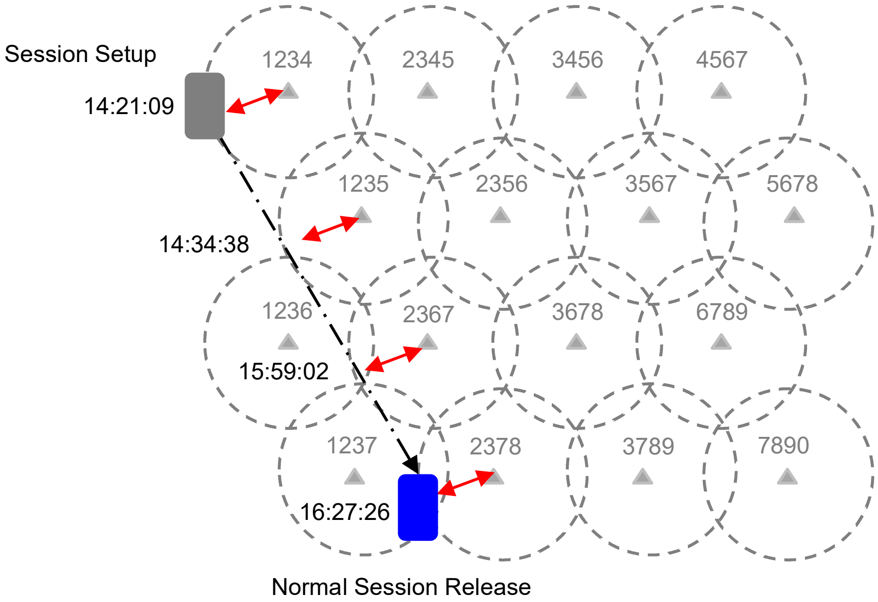
Real CDRs will show a much wider range of details, such as Charging ID, APN, LAC, RAC/TAC, etc. The examples show the minimum amount of information necessary to allow us to draw cell site conclusions from the events.

In scenarios in which a ‘follow-on’ CDR would be generated the examples will show CDR1 (an initial record in a new session) and CDR2 (a follow-on record for the same session) details.

The examples will also provide an indication of the cell site conclusions that are safe to draw from the CDR evidence provided.

4.1. Session Explicitly Terminated

If the mobile device decides (or is informed by an app) that it no longer requires a currently-active data session, it will send a ‘session release’ request to the network. The currently open SGSN/S-GW and GGSN/P-GW CDRs will be ‘finalised’, closed and transmitted to the billing system.



Previous data CDR’s End Time:
12:45:17

CDR	
Start Cell:	1234
Opening Time:	14:21:09
Last Used Cell:	2378
Closing Time:	16:27:26
Duration:	02:06:18
Volume:	2.78MB
Cause:	Normal

Start cell details are accurate
Last used cell details are accurate (although not included in the CDR)

In this example:

This session can be determined to be a ‘new’ session as there were no other GPRS/PS Data sessions established that closed immediately prior to the start time at 14:21:09hrs; the previous GPRS/PS Data CDR closed at 12:45:17hrs.

The session started at 14:21:09hrs when the phone sent a ‘session setup request’ via Cell ID 1234. The ID of the cell that was used to send the setup request is captured and the Start Cell details are therefore accurate, in the sense that the subject phone must have been within the coverage of the Start Cell ID shown in the CDR at the Start Time of that CDR to have sent the setup request.

The last time the phone connected to the network was at 16:27:26hrs, using Cell ID 2378, when it sent a ‘session release’ request to the network. Cell ID 2378 will be stored as the ‘last used cell’ in the SGSN/S-GW.

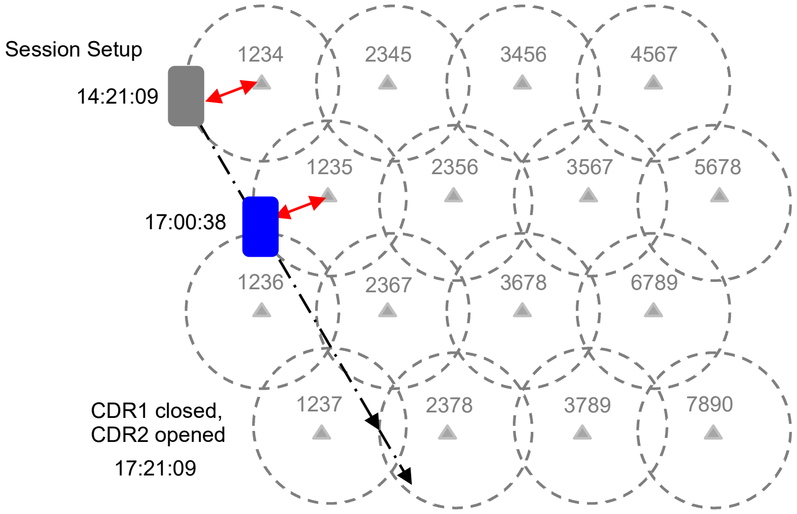
There will be no ‘follow-on’ CDR generated for this session, so the accuracy of the last used cell details are irrelevant.

Cell Site Conclusion: it is safe to draw the conclusion that the subject phone was within the coverage area of the Start Cell (Cell ID 1234) **AT** the start of the CDR period (14:21:09), although we’d still recommend using the standard ‘at or before’ conclusion if possible.

There was no ‘follow-on’ CDR opened as the session terminated, so the stored ‘last used cell’ details are irrelevant.

4.2. Time Limit Reached

If the network has a configured ‘CDR open time’ limit in the SGSN/S-GW then the currently-open CDR will be closed and a follow-on CDR will be opened when the current CDR ‘open time’ duration reaches that limit. In this example the time limit is 3hrs.



Previous data CDR’s End Time: 12:45:17

CDR1	
Start Cell:	1234
Opening Time:	14:21:09
Last Used Cell:	1235
Closing Time:	17:21:09
Duration:	**03:00:00**
Volume:	2.67MB
Cause:	Time Limit

CDR2	
Start Cell:	1235
Opening Time:	17:21:09
Last Used Cell:	open
Closing Time:	open
Duration:	open
Volume:	0.00MB
Cause:	open

Start Cell details for the follow-on CDR2 could be inaccurate if the mobile device has moved since connecting with Cell ID 1235 at 17:00:38

In this example:

CDR1 is the first record for a new session, so it’s start cell can be regarded as being accurate; CDR2 is a follow-on record and the start cell needs to be treated with caution.

The ‘time limit’ trigger operates independently of data transmission across the session, so the subject phone could be in either idle or connected mode when the limit is reached. It is therefore not possible, based only on the data contained in the disclosed CDR, to determine whether the Start Cell ID captured for CDR2 is the currently used cell or is a stored ‘last used cell’.

The Cell ID captured for the start of CDR2, Cell ID 1235 at 17:21:09, may or may not be accurate and depends on whether the phone moved after its cell location (cell ID 1235) was last reported to the network at 17:00:38.

As the session is a continuing one, CDR2 is shown in its ‘open’ state immediately after being opened, so no further relevant information is shown.

Cell Site Conclusion: it is safe to draw the conclusion that the subject phone was within the coverage area of the Start Cell (Cell ID 1234) **AT** the start of the CDR1 period (14:21:09).

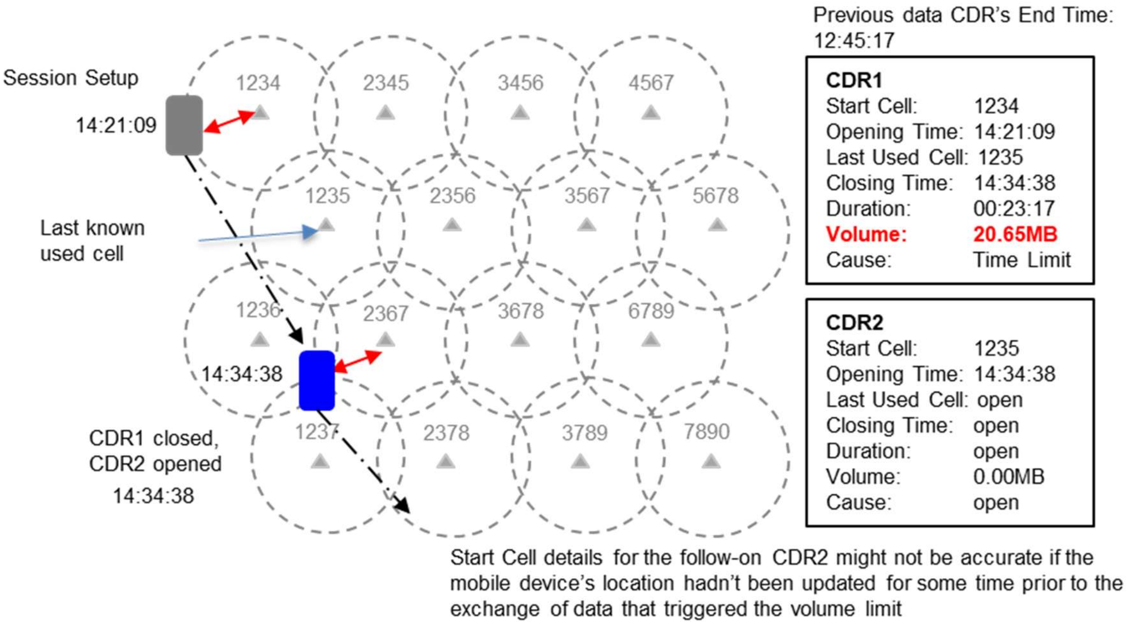
It is NOT safe to draw the conclusion that the subject phone was within the coverage area of the Start Cell (Cell ID 1235) at the start of the follow-on CDR2 period (17:21:09), as there is no totally reliable way of determining whether that was a ‘current’ cell ID or simply a ‘last used’ cell ID at that time. The generally accepted ‘safe’ conclusion in this scenario is to say that the subject phone was within the coverage area of Cell ID 1235 ‘**AT OR BEFORE**’ the CDR time of 17:21:09hrs).

4.3. Volume Limited Reached

If the network has a configured 'data volume' limit (in the SGSN/S-GW and/or GGSN/P-GW) then the currently open CDR will be closed and a follow-on CDR will be opened when the data count reaches that limit. In this example the data volume limit is 20MB.

User data flows are separate from the flow of administrative or 'signalling' information for mobile data sessions, so the SGSN/S-GW can be actively exchanging data packets with a mobile device without knowing exactly which cell that device is currently being served by.

4G S-GWs have a better understanding of a device's current cell location as they connect directly to the 4G base station (but they may not necessarily know which specific cell on a base station the device is currently using). 2G/3G SGSNs connect to the access network controller (BSC or RNC) and not directly to the base station, and so only get updated cell location information when the BSC/RNC decide to pass it on.



Previous data CDR's End Time: 12:45:17

CDR1	
Start Cell:	1234
Opening Time:	14:21:09
Last Used Cell:	1235
Closing Time:	14:34:38
Duration:	00:23:17
Volume:	**20.65MB**
Cause:	Time Limit

CDR2	
Start Cell:	1235
Opening Time:	14:34:38
Last Used Cell:	open
Closing Time:	open
Duration:	open
Volume:	0.00MB
Cause:	open

Start Cell details for the follow-on CDR2 might not be accurate if the mobile device's location hadn't been updated for some time prior to the exchange of data that triggered the volume limit

In this example:

The details and conclusion related to the start of CDR1 match those in the previous examples.

A packet or sequence of packets sent via the data session triggers the SGSN to detect that the 'volume limit' has been reached or passed.

The Closing Time for CDR1 and Opening Time for CDR2 (14:34:38hrs) reflects the time at which the data volume limit was reached.

The Cell ID captured for the start of CDR2, Cell ID 1235, may or may not be accurate depending upon whether the SGSN had been informed of the cell change. In the example in the diagram, the SGSN was last updated with cell locations when the mobile device was connected to cell ID 1235, but it wasn't informed that the device had moved on the cell ID 2367, so the 'last used' cell in CDR1 and the 'start cell' in CDR2 are inaccurate.

As the session is a continuing one, CDR2 is shown in its 'open' state immediately after being opened, so no further relevant information is shown.

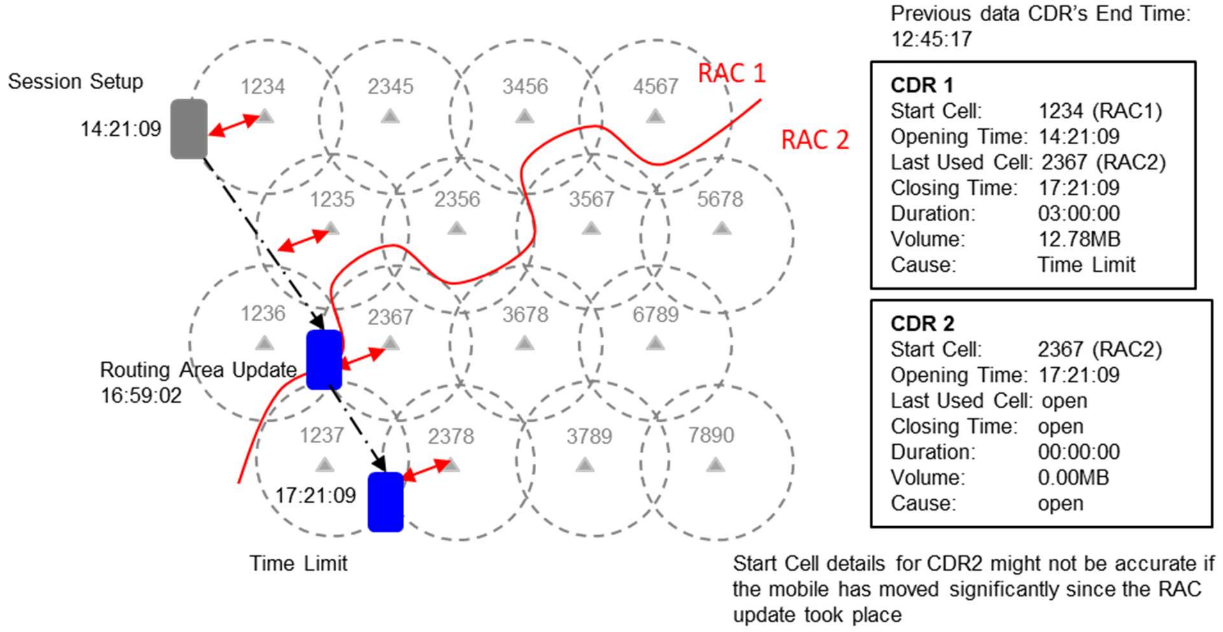
Cell Site Conclusion: it is safe to draw the conclusion that the subject phone was within the coverage area of the Start Cell (Cell ID 1234) at the start of the CDR1 period (14:21:09).

It is **NOT** safe to draw the conclusion that the subject phone was within the coverage area of the Start Cell (Cell ID 1235) at the start of the follow-on CDR2 period (14:34:38), as there is no totally reliable way of determining whether that was the current 'used' cell ID or simply a 'last reported' cell ID. The generally accepted 'safe' conclusion in this scenario is to say that the subject phone was within the coverage area of Cell ID 1235 '**AT OR BEFORE**' the CDR time of 14:34:38hrs).

4.4. RAC/TAC Change

If a 2G/3G connected mobile device that is engaged in an active data session reselects to a cell in a new Routing Area it will detect this based on the changed RAC being broadcast by the new cell. In this scenario the phone must transmit a RAU (Routing Area Update) to the serving SGSN to update its location. This requires the phone to connect to the new cell to enable the RAU to be sent. The SGSN picks up the phone's new cell as a consequence of the RAU being transmitted. This also applies to 4G-connected devices, but the network nodes involved are different – the MME receives the location update and the CDR is maintained by the S-GW – and the trigger will be a change of Tracking Area.

Theoretically, the currently open SGSN/S-GW CDR should be closed (with a cause code of 'RAC Change') and a follow-on CDR should be opened that uses the reported 'RAC Change cell' as the start cell. Real world testing has shown, however, that if the RAU doesn't lead to a change of serving SGSN/S-GW (e.g. the mobile moves between RACs controlled by the same SGSN/S-GW) networks don't accept RAC change as trigger for part-billing, so although the SGSN logs the reported change of 'last used' cell, it doesn't trigger the closure/reopening of the CDR.



Previous data CDR's End Time:
12:45:17

CDR 1
Start Cell: 1234 (RAC1)
Opening Time: 14:21:09
Last Used Cell: 2367 (RAC2)
Closing Time: 17:21:09
Duration: 03:00:00
Volume: 12.78MB
Cause: Time Limit

CDR 2
Start Cell: 2367 (RAC2)
Opening Time: 17:21:09
Last Used Cell: open
Closing Time: open
Duration: 00:00:00
Volume: 0.00MB
Cause: open

Start Cell details for CDR2 might not be accurate if the mobile has moved significantly since the RAC update took place

In this example:

The details and conclusion related to the start of CDR1 match those in the previous examples.

The phone connected to the network via Cell ID 2367 in RAC 2 to send its RAU at 16:59:02. Cell ID 2367 would be captured as the 'last used cell' (at 16:59:02hrs) but CDR1 was not closed at that time.

As data exchanged continues and the phone continues to move, the change to cell 2378 was not passed on to the SGSN, so the 'last used' cell stays as 2367.

CDR1 reached its ‘time limit’ at 17:21:09 and was closed and CDR2 was opened.

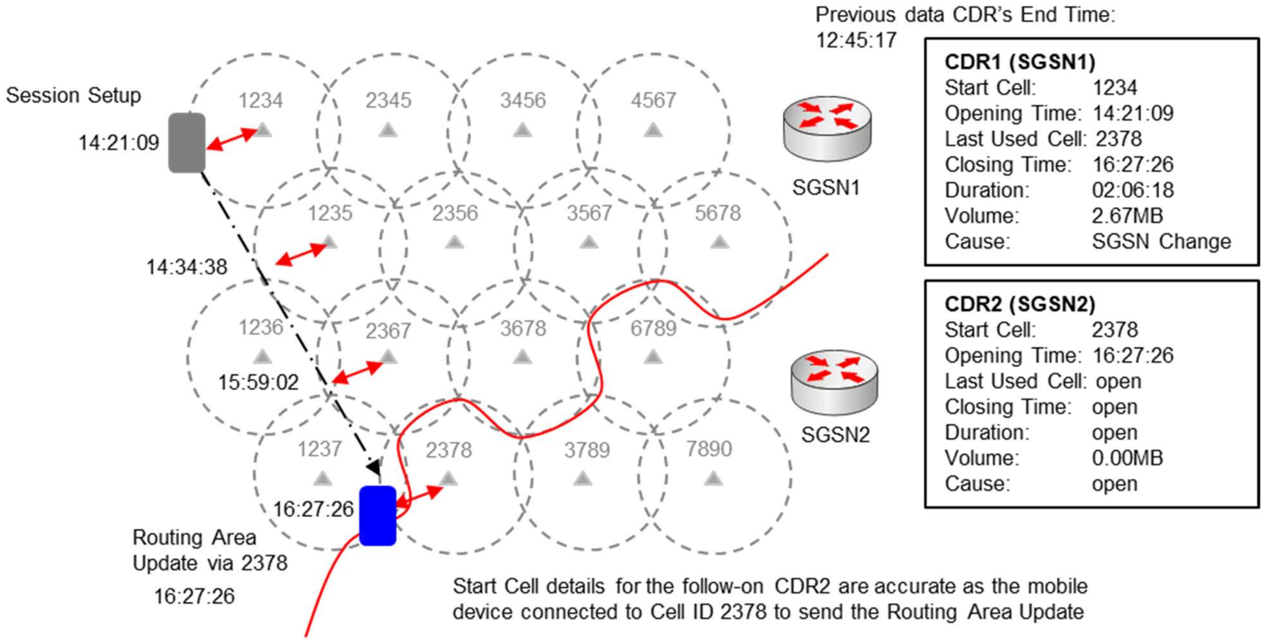
Cell ID 2367 was captured as the Start Cell of CDR2 with an opening time of 17:21:09hrs, even though by that time the device was being served by cell 2378.

Cell Site Conclusion: it is safe to draw the conclusion that the subject phone was within the coverage area of the Start Cell (Cell ID 1234) at the start of the CDR1 period (14:21:09).

It is **NOT** safe to draw the conclusion that the subject phone was within the coverage area of the Start Cell (Cell ID 2367) at the start of the follow-on CDR2 period (17:21:09), as although the change of RAC was reported to the SGSN (at 15:59:02) it did not trigger a change of CDR and the mobile may have moved by the time CDR2 was opened.

4.5. SGSN/S-GW Change

SGSNs and S-GWs are typically configured to serve only a subset of a network’s cells and are associated with specific RACs or TACs. If a mobile device reselects to a new cell that is under the control of a different SGSN/S-GW it will transmit a RAU (Routing Area Update - 2G/3G) or TAU (Tracking Area Update - 4G)to a new SGSN/S-GW, causing the current CDR to be closed at the ‘old’ SGSN/S-GW and a follow-on CDR to be opened in the ‘new’ SGSN/S-GW. The GGSN/P-GW CDR will not need to be closed as a consequence of SGSN change.



Previous data CDR’s End Time:
12:45:17

CDR1 (SGSN1)

Start Cell:	1234
Opening Time:	14:21:09
Last Used Cell:	2378
Closing Time:	16:27:26
Duration:	02:06:18
Volume:	2.67MB
Cause:	SGSN Change

CDR2 (SGSN2)

Start Cell:	2378
Opening Time:	16:27:26
Last Used Cell:	open
Closing Time:	open
Duration:	open
Volume:	0.00MB
Cause:	open

Start Cell details for the follow-on CDR2 are accurate as the mobile device connected to Cell ID 2378 to send the Routing Area Update

In this example:

The details and conclusion related to the start of CDR1 match those in the previous examples.

The phone connected to the network via Cell ID 2378 in RAC 2 to send its RAU to SGSN2.

SGSN2 will contact SGSN1 and inform it that the phone has passed into its control and will ask SGSN1 to provide the administrative details for the phone.

SGSN1 will finalise and close CDR1 at 16:27:26 and send it to the billing system.

SGSN2 will open CDR2 and will capture Cell ID 2378 as the Start Cell (at 16:27:26hrs).

As the session is a continuing one, CDR2 is shown in its ‘open’ state immediately after being opened, so no further relevant information is shown.

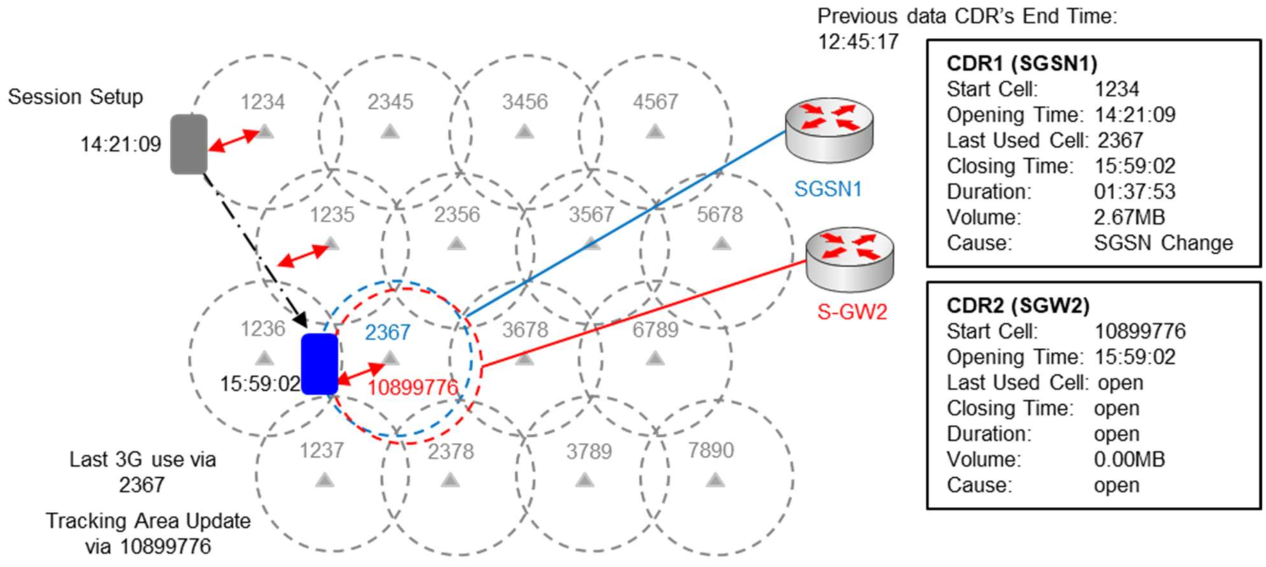
Cell Site Conclusion: it is safe to draw the conclusion that the subject phone was within the coverage area of the Start Cell (Cell ID 1234) at the start of the CDR1 period (14:21:09).

It is safe to draw the conclusion that the subject phone was within the coverage area of the Start Cell (Cell ID 2378) at the start of the follow-on CDR2 period (16:27:26), as the phone must have used that cell to transmit the RAU message and there was a change of SGSN.

'SGSN Change' is only detectable is some CDR formats (e.g. Vodafone), it might be inferred in O2 and EE data, for example from a change of LAC, but not reliably. So, although 'SGSN Change' is a valid reason for using an 'at' conclusion for GPRS events from a technical point of view, it may not be possible to reliably use this in most practical cases.

4.6. Change of Technology

SGSNs manage user traffic via 2G/3G cells and S-GWs manage user traffic via 4G cells (and eventually, UPFs will handle 5G cells). A mobile device might decide to reselect between, say, 3G coverage and 4G coverage without necessarily changing location, if the 4G coverage suddenly becomes more attractive than the current 3G coverage. This entails the device's session being transferred from a 2G/3G SGSN to a 4G SGW and the transfer of the CDRs associated with that session as well. 'Change of technology' is therefore another valid cause for part-billing.



Previous data CDR's End Time:
12:45:17

CDR1 (SGSN1)

Start Cell:	1234
Opening Time:	14:21:09
Last Used Cell:	2367
Closing Time:	15:59:02
Duration:	01:37:53
Volume:	2.67MB
Cause:	SGSN Change

CDR2 (SGW2)

Start Cell:	10899776
Opening Time:	15:59:02
Last Used Cell:	open
Closing Time:	open
Duration:	open
Volume:	0.00MB
Cause:	open

15:59:02

Start Cell details for the follow-on CDR2 are accurate as the mobile device connected to Cell ID 10899776 to send the Tracking Area Update

In this example:

The details and conclusion related to the start of CDR1 match those in the previous examples.

The phone was connected to the network via 3G Cell ID 2367 (connected to SGSN1) but decided to reselect to 4G cell ID 10899776 (connected to SGW2) at 15:59:02.

SGW2 will contact SGSN1 and inform it that the phone has passed into its control and will ask SGSN1 to provide the administrative details for the phone.

SGSN1 will finalise and close CDR1 at 15:59:02 (using Cell ID 2367 as the end cell if those details were reported to it) and send it to the billing system.

SGW2 will open CDR2 and will capture Cell ID 10899776 as the Start Cell (at 15:59:02hrs).

As the session is a continuing one, CDR2 is shown in its ‘open’ state immediately after being opened, so no further relevant information is shown.

Cell Site Conclusion: it is safe to draw the conclusion that the subject phone was within the coverage area of the Start Cell (Cell ID 1234) at the start of the CDR1 period (14:21:09).

It is safe to draw the conclusion that the subject phone was within the coverage area of the Start Cell (Cell ID 10899776) at the start of the follow-on CDR2 period (15:59:02), as the phone must have used that cell to transmit the TAU message.

4.7. Anomalies

Scenarios as complex as this can generate anomalies – GPRS billing data can be misinterpreted to show that a phone was active when in fact it had recently been switched off.

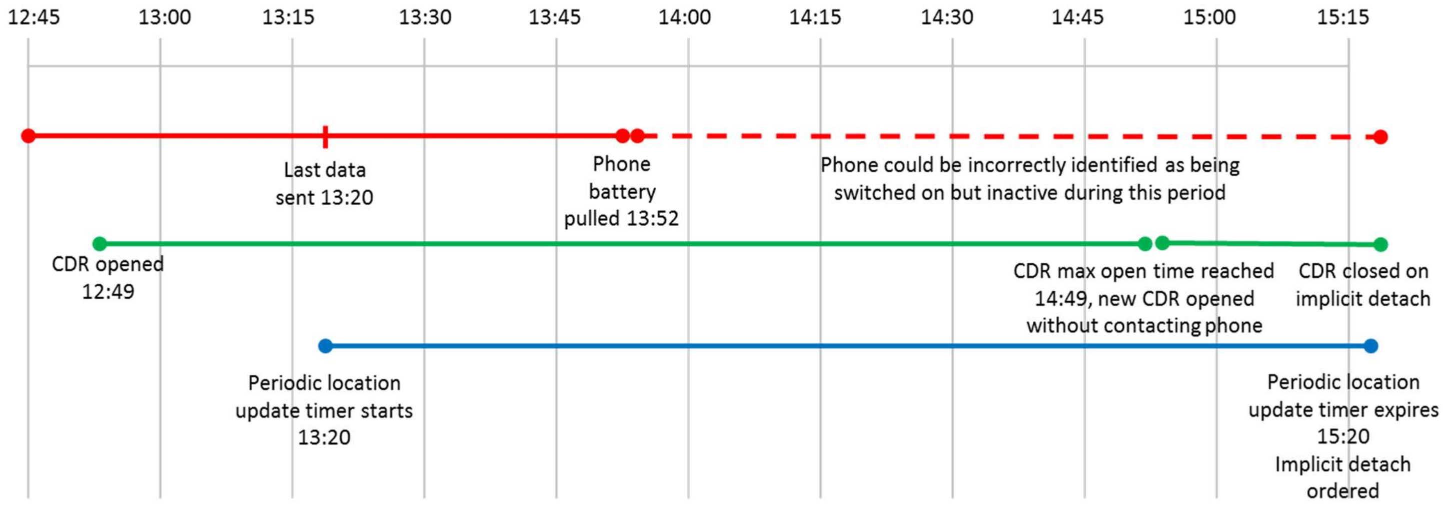
The circumstances in which a CDR record can continue after a phone has been switched off are related to whether the phone performed an 'orderly' shutdown or not

If a phone is shutdown in an orderly fashion (pressing power off button or because the battery is nearly exhausted), it sends a 'detach' message to the network that logs it off and closes any open billing records.

If a phone powers off in a disorderly fashion (battery pulled, dropped in water, placed in Faraday bag, etc) it doesn't have a chance to send a detach message, so the network doesn't know that it's not connected any more. Data transmission during GPRS/data sessions can be intermittent, the phone has no responsibility to contact the network unless it needs service, at least until its periodic location update timer expires, but if the phone fails to contact the network for a periodic LAU when expected, the network performs an implicit detach and logs it off.

In some circumstances, the interplay between the billing timers and the LAU timer can create unusual billing records.

For example: imagine that a network has a ‘maximum open time’ value for GPRS CDRs of 2hrs and a periodic location update timer of 2hrs.


12:45
13:00
13:15
13:30
13:45
14:00
14:15
14:30
14:45
15:00
15:15
Last data sent 13:20
Phone battery pulled 13:52
Phone could be incorrectly identified as being switched on but inactive during this period
CDR opened 12:49
CDR max open time reached 14:49, new CDR opened without contacting phone
CDR closed on implicit detach
Periodic location update timer starts 13:20
Periodic location update timer expires 15:20 Implicit detach ordered

Let's say, for a given phone, that the current GPRS CDR was opened at 12:49, so barring other causes for closure, it will be closed for 'maximum open time reached' at 14:49

Imagine that the phone has an active data session up until 13:20, when it sends the last data in that session and then sends no further data and exchanges no calls or texts. The data session remains in place but is in an idle state.

At this point (13:20) the 'periodic location area update timer' starts and, barring any other contact with the network, the phone will be expected to do a periodic LAU at 15:20.

The phone user pulls the battery at 13:52 - the phone does not send a Detach message, the network does not know that the phone is no longer available.

The current CDR is closed at 14:49 as ‘maximum open time’ is reached. The network does not need to contact the phone to manage this function, so the network still has no idea that the phone is unavailable. A new CDR is opened at 14:49 - which makes it appear as if the phone is active.

It gets to 15:20 with no contact from the phone and no periodic LAU is received, so the network initiates Implicit Detach, ending the new CDR session and logging the phone off from the network

From viewing the GPRS CDRs it could therefore appear that the phone was switched on and available between 12:49 to 15:20 and specifically that it was still active between 14:49 and 15:20, whereas it was actually unavailable from 13:52 onwards.

Situations like this are very difficult to detect with any certainty, but they can be inferred from the data with some types of CDR - especially those (VF, O2, EE) that show uplink/downlink data amounts. The last CDR (14:49-15:20) should have a zero data volume (0kb), as the phone wasn't active to use any data. With CDR types that don't show data volume (H3) the inference is more difficult to draw, but in all cases where this scenario pertains there would be a period with no data sessions or voice/text events after the end of the significant data session, as the phone would have been switched off.

This is just one example of the complex set of inferences that can be drawn from GPRS billing data in unusual scenarios; most billing records are comparatively straight forward and do not require such complex explanations.

4.8. Real-World Testing

Testing of a variety of GPRS/mobile data scenarios has been undertaken by Will Metters of the NPCC DCG Futures Group.

Each test followed a specific pattern of data upload/download activities using an RF test phone (e.g. a NEMO Handy), with event logging enabled and then comparing the activities undertaken with the logs captured by the test phone against the CDRs provided by the network.

This meant that, from the handset logs, it was possible to see when cell/RAC/technology changes occurred and from the CDRs it was possible to see when or if those changes were reported to the network and if they were reflected in the start cells on follow-on records.

The outcome of the testing was encouraging – in most cases the cell IDs shown in the CDRs reflected those that the phone reported itself to be using at that time, but there were differences.

The upshot of this testing was that in the majority of cases the ‘start cell’ listed in a GPRS/mobile data record, even for follow-on records, appeared to be accurate, but the fact that there were occasions when there were differences meant that the accuracy could not be guaranteed. This means that the recommendation to use the standard ‘at or before’ wording for conclusions based on GPRS data remains in place, unless there is additional evidence available to show that an ‘at’ conclusion could be used instead.

5. Summary & Conclusions

GPRS and PS data CDRs capture (or appear to capture, in the case of O2) details of the Start Time and Start Cell used by the subject phone.

For ‘standalone’ session records – those with no session records immediately before or after them – and for the first record in a new session, the Start Cell can, on the balance of probability, be trusted as being the cell that the subject phone was using at the time shown in the CDR’s Start Time field.

In the case of partial or follow-on records the correlation between Start Time and Start Cell can be considered to be less trustworthy, or alternatively that it may be more difficult to prove that the correlation is accurate.

The reason for the correlation difficulty is due partly to the intermittent nature of PS data, meaning that the mobile device isn’t required to remain in constant contact with the network, and also due to procedures that ensure that the core network isn’t updated every time the device changes cells. These issues combine to ensure that the network’s understanding of the ‘current’ cell being used by a device is more often just the ‘last reported’ cell and may not have been updated for minutes or even hours.

There are scenarios in which it is possible to be more definite about the correlation between start time and start cell and real-world testing has shown that in most cases the correlation is actually correct.

However, it may be difficult in many cases to determine between ‘standalone’ and ‘partial’ records and may be even more difficult to distinguish between the various forms of partial record event.

For this reason the generally accepted phrasing to be used for cell site conclusions drawn in relation to GPRS/PS Data events is that the subject phone was within the coverage area of the Start Cell ‘**at or before’** the time shown as the event timestamp.

More experienced practitioners may wish to employ more definite ‘**at**’ terminology in relation to events that have clear cut cell/timestamp parameters.

5.1. College of Policing Advice

The College of Police constantly update their advice in line with network changes should they affect conclusions that may be drawn from this type of data. Their current advice in light of O2 indicating that they only show “End Cells” in their GPRS records resulted in the following advice from the college;

As should be common knowledge the national guidance around GPRS data has been to describe the cell listed in a record to have been connected “at or before the start time” listed on that record to allow for the way Mobile Networks were understood to work. Recent work conducted by O2 has identified that on their network the guidance needs to be changed so as to describe the cell listed in a record to have been connected “at or before the end time” listed on that record. ***As O2 do not list an end time they advise that the end time should be calculated by adding the duration to the start time.***

As a result of this change of advice, work is on-going with O2 in order to ensure their GPRS (MDE) data is fully understood and can be correctly described whilst at the same time work has begun with all the other Mobile Network Operators to ensure that the current advice is still correct or whether it needs updating. Therefore whilst this work is being undertaken the following descriptions should be used:

[Note: the wording of these suggested conclusions has been edited by us to reflect recent re-evaluation of the wording for O2 and H3 records and the following does not exactly match College guidance]

- EE: “The cell shown was connected to at or before the start time of the record and after the start time of the earliest previous contiguous record with a different cell ID”
- Vodafone: “The cell shown was connected to at or before the start time of the record and after the start time of the earliest previous contiguous record with a different cell ID”
- O2: **“**The cell shown was connected to at or before the end time of the record and after the start time of the earliest previous contiguous record with the <u>same</u> cell ID**”**
- Three
  - IP Locations: “The cell shown was connected to at or before the start time and before the end time of the record”
  - LTE Locations: “The cell shown was connected to at or before the start time of the record”

5.2. Feedback, Comments & Questions

The information provided in this briefing paper is based mainly on the relevant 3GPP standards, on examination of examples of UK GPRS CDR formats and on personal experience.

It is also based on the results of real-world testing undertaken by Will Metters of NPCC DCG Futures Groups, which he has kindly allowed us to mention.

Other input, either directly to this paper or indirectly via discussion, has been provided by: Martin Griffiths and Dave Cutts of Forensic Analytics; James Matthews and Matt Lashley of Cyfor; Matt Tart and Iain Brodie of CCL.

There has been no specific input from the UK CSPs. We will be more than happy to discuss specific details with the CSPs or with practitioners who have had those conversations with CSPs. We would be happy to correct any errors or inaccuracies that are highlighted if CSPs indicate that their CDR formats should be subject to alternative interpretation.

Questions, comments and feedback can be sent to: enquiries@forensicanalytics.co.uk

6. Further Reading & Other Resources

The primary source for information related to mobile technologies is 3GPP, as they are responsible for maintaining the specifications that describe those networks.

3GPP specifications are available for free download from their website at www.3gpp.org/specifications.

The specifications that are particularly relevant to this document are:

- 32.251 – Charging Management: PS Domain Charging
- 23.060 – GPRS Service Description (2G/3G)
- 23.401 – GPRS Enhancements for EUTRAN Access Networks (4G)

There are also a number of educational resources at http://www.3gpp.org/technologies

Useful reference books for cellular technologies include:

- GSM Made Simple – George Lamb, Cordero Consulting 1997
- From GSM to LTE – Martin Saunter, Wiley 2010
- Essentials of UMTS – Christopher Cox, Cambridge University Press 2008
- An Introduction to LTE – Christopher Cox, Wiley 2012
- Forensic Radio Survey Techniques for Cell Site Analysis, Joseph Hoy, Wiley 2015

This topic is discussed at length in our Next Generation Communications Data course (FA017).

Forensic Analytics offers a wide range of training and support services and has links with many organisations that offer similar services – contact us to discuss your training needs.

7. Glossary

2G	Second Generation mobile network e.g. GSM
3G	Third generation mobile network e.g. UMTS
3GPP	Third Generation Partnership Project
4G	Fourth generation mobile network e.g. LTE
5G	Fifth generation mobile network e.g. NR
Active	a 3G cell currently selected to serve a mobile device’s Connected Mode connections
ARFCN	Absolute Radio Frequency Channel Number in 2G
BSC	Base Station Controller (in 2G)
CDR	Call Detail Record
CELL_DCH	Cell Dedicated Channel state (in 3G)
CELL_FACH	Cell Forward Access Channel state (in 3G)
CELL_PCH	Cell Paging Channel state (in 3G)
Connected Mode	The state a mobile device is in when a connection has been established to a base station and traffic flow is possible
CS	Circuit Switched e.g. traditional voice telephony service
CSG	Closed Subscriber Group (for 3G/4G femtocells)
Dedicated Mode	Original term for Connected Mode used in GSM
EARFCN	Evolved Absolute Radio Frequency Channel Number (in 4G)
EDGE	Enhanced Data Rates for Global Evolution, PS data for 2G networks
eNB	Evolved Node B – 4G base station
Femtocell	A small-scale cell/base station designed to be deployed at a user’s home or office, which provides a small bubble of network service
GGSN	Gateway GPRS Support Node
GPRS	General Packet Radio Service, PS data for 2G networks
GSM	Global System for Mobile, 2G network type
Handover	The process of passing the active connections for a mobile device in Connected Mode from one cell/base station to another
HSPA	High Speed Packet Access, fast PS data for 3G networks
Idle Mode	The state where a mobile device is powered on and attached to a network but has no active control or traffic connections
LA	Location Area (in 2G and 3G)
LAC	Location Area Code (in 2G and 3G)
LTE	Long Term Evolution, as 4G network type
MCC	Mobile Country Code e.g. 234 for the UK
MNC	Mobile Network Code e.g. 10 for O2 UK
MS	Mobile Station, a 2G mobile device
NR	New Radio (5G radio type)

P-GW	PDN (Packet Data Network) Gateway
PLMN	Public Land Mobile Network
PS	Packet Switched e.g. the data transmission mechanism used by data networks like the Internet
PSC	Primary Scrambling Code (in 3G)
RA	Routing Area (in 2G and 3G)
RAC	Routing Area Code (in 2G and 3G)
RAU	Routing Area Update (in 2G and 3G)
RAN	Radio Access Network
Reselection	In Idle Mode, the process by which a mobile device selects the serving cell that it will camp on
RNC	Radio Network Controller (in 3G)
Serving	Term applied to the cell that an Idle Mode device is currently camped on or that a Connected Mode device is connected to
SGSN	Serving GPRS Support Node
S-GW	Serving Gateway
SIM	Subscriber Identity Module
SMS	Short Message Service
TA	Tracking Area (in 4G)
TAC	Tracking Area Code (in 4G)
TAU	Tracking Area Update (in 4G)
UE	User Equipment
UMTS	Universal Mobile Telecommunications System, a 3G network type
URA_PCH	UTRAN Registration Area Paging Channel (in 3G)

CSAS

CDAN
Nexus

CSAS
Decypher

Cell Site
Analysis

Training
Services

Forensic Analytics Ltd
Pixmore Business Centre
Pixmore Avenue
Letchworth
SG6 1JG

www.forensicanalytics.co.uk