

NETWORK MONITORING AND ANALYSIS TOOL

Field of the Invention

- 5 The present invention relates to systems and methods for analysing the performance of a network. In particular although not exclusively the present invention relates to the provision of a network monitoring tool for mobile telecommunication systems.

10 Discussion of the Background Art

- Mobile networks involve a very large number of components ranging from Radio Network Controllers to Gateway GPRS Support Nodes (GGSNs) with numerous firewalls, routers and other devices. Data services through mobile networks, of
15 whatever underlying technology, remain a relatively recent development. Given this, mobile carriers are still developing procedures and tools to allow them to monitor, tune and upgrade these data networks.

- Typically carriers are able to implement a wide range of monitors and probes into the
20 infrastructure of the packet core to monitor its performance. However there is no current way to provide ongoing monitoring of the "user experience" where the data network is concerned. A user's experience of data services can vary widely according to location and a various numbers of other factors. For example, a network packet core could be performing extremely well, but if the cell towers are
25 aligned incorrectly then many users in a particular area could experience poor network performance. Providing a consistent user experience in this environment is problematic given the lack of visibility beyond the packet core.

- Clearly it would be advantageous to provide a monitoring tool that is capable of
30 providing data on the performance of a network beyond the packet core. It would also be advantageous to provide a monitoring tool which may be deployed as a hidden payload to a carrier supplied mobile messaging client, so as to permit a large volume of tests to be performed across the entire network.

SUMMARY OF THE INVENTION

Disclosure of the Invention

5

Accordingly in one aspect of the present invention there is provided a method for determining the performance of a communications network said method comprising the steps of:

transmitting a message from a mobile device to one or more servers
10 configured to direct the message back to the mobile device;

receiving at the mobile device the messages returned by the said one or more servers;

calculating a time differential between transmission of the message by said mobile device and receipt of the messages returned by the one or more servers by
15 said mobile device; and

forwarding the calculated time difference to a primary server selected from said one or more servers for storage.

Suitably the message sent by the mobile device to the one or more servers includes
20 the mobile device's identification number (MIN) which is registered with at least one of the servers. The message is kept sufficiently small to keep the process transparent from the user and to prevent unnecessary overhead that could adversely affect user experience. Suitably the message is more or less in the order of 100 bytes.

25

Preferably each of the servers apply a time stamp to the message when directing the message back to the mobile device to permit the monitoring tool to calculate the time for the outbound and inbound legs of the round trip for the message.

30

The step of transmitting the message may be performed at regular intervals. Suitably the interval between each transmission of the message is set to keep the power consumption required in the process minimal so as not to adversely affect battery life of the mobile device. For example the interval between each

transmission may be set to a period of 55 to 61 minutes. The primary server may instruct the mobile test client to vary the length of the interval between each transmission of the message based on network team test requirements or dynamically based on current network performance so as to obtain more detailed test data.

The step of calculating the time differential may further include the steps of calculating the time between the initial send and the subsequent response for each test message. Alternatively the step of calculating the time differential may include the step of calculating a combined/average value for all servers within the one or more servers. Utilizing multiple servers may eliminate issues specific to a single server and allow a network service provider to hone in on a network specific issue rather than a single anomaly within a given sector.

Suitably the step of forwarding the calculated time difference to the primary server further includes forwarding information on the type of mobile device, the application used to send the message, the mode of connection between the mobile device and the communications network, the User ID of the mobile subscriber, the Identification Number of the mobile device etc, cell identification and sector number in which the mobile is located.

The method may further include the step of formulating reports based on the time differential and the type of mobile device, the application used to send the message, the mode of connection between the mobile device and the communications network, the User ID of the mobile subscriber, the Identification Number of the mobile device etc. Preferably the reporting is in the form of an Excel extract of the raw data for 24 hours on a daily basis. Suitably the reports are provided in real-time or near real-time via web page via daily, SMS, email or the like. Preferably, the reports include information identifying performance issues in specific cell sites and cell site sectors to allow carriers to target scarce resources where they can provide most benefit to consumers.

The method may also include the step of forwarding a set time differential to the primary server, the set time interval being indicative of a timeout event wherein the return message from the servers have not been received by the mobile device within a preset period. Preferably the preset period is initially set to 10 minutes, although
5 such preset period may be varied to such period as may be optimal for its purpose. The primary server may vary the preset period in accordance with a variation of the predetermined period. Alternatively a timeout event may be indicated by way of a code in the transaction to the primary server.

- 10 Suitably the method may be implemented with no end user interaction. In such cases the mobile device includes a dedicated application which causes the mobile device to send messages to the servers at regular intervals. The application being in communication with the primary server and responsive to instructions transmitted from the server to alter the intervals at which it transmits messages to the servers.
15 Alternatively a method of monitoring may be manually activated by a user of the mobile device via one or more user screens.

- Where the method is implemented via a dedicated application installed on the mobile device the application may utilise one or more auxiliary applications on the
20 mobile device to send messages to the one or more servers. The primary server may deactivate the dedicated application when further monitoring is not required. In such instances the servers may set the transmission intervals to a large time frame, for example 9999 minutes. Alternatively the primary server may send a specific instruction to the mobile client application to stop monitoring. Where the test client
25 application is hidden within another application such as a mobile messaging application, the mobile network carrier must ensure no inconvenience results to the consumer. Suitably this would include ensuring the consumer is not charged for any data usage resulting from the monitoring tool. Further, test volumes must be kept to a low level to minimise battery consumption or any other impact on the consumer's
30 mobile experience.

Throughout the specification the term "comprising" shall be understood to have a broad meaning similar to the term "including" and will be understood to imply the

inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps. This definition also applies to variations on the term "comprising" such as "comprise" and "comprises".

5 BRIEF DETAILS OF THE DRAWINGS

In order that this invention may be more readily understood and put into practical effect, reference will now be made to the accompanying drawings, which illustrate preferred embodiments of the invention, and wherein:

10 FIG. 1 is a schematic diagram depicting the operation of a network monitoring tool according to one embodiment of the present invention;

FIG. 2 is a schematic diagram depicting the operation of a network monitoring tool according to a further embodiment of the present invention;

15 FIG. 3 is a schematic diagram depicting the transaction flow for the network monitoring tool according to one embodiment of the present invention;

FIG. 4 is a schematic diagram depicting of the operation of a network monitoring tool according to another embodiment of the present invention; and

FIGs. 5A to 5C are a series of users screens for use in the network monitoring tool according to one embodiment of the present invention.

20

DESCRIPTION OF EMBODIMENTS OF THE INVENTION

With reference to Fig 1 there is illustrated the operation of a network monitoring tool according to one embodiment of the present invention. In this example the
25 monitoring tool is deployed as an application on a mobile device roaming within a network. As shown the network monitoring tool causes the mobile device **101** to send a message **104** to a server **102** (known as the connect server) associated with the network. The message **104** is at least 100 bytes and contains the mobile device identification number (MIN) which is registered with the connect server **102**.
30 Transmitting a message of a reasonable size, such as 100 bytes, ensures that the monitoring tool can, with some degree of accuracy establish a base line for network performance which can be factored into later analysis.

On receipt of the message **104** the connect server **102** immediately forwards the message back **105** to the mobile device **101**. The network monitoring tool then calculates the Round Trip Time (RTT) of the message by comparing the time stamps applied by the mobile device to the sent and received messages to determine the total elapsed time. The round trip time information is then forwarded **106** to the connect server **102** for storage in database **107** for logging and reporting purposes.

Fig 2 depicts the operation of a network monitoring tool according to a further embodiment of the present invention. In this particular example the network monitoring tool causes the mobile device **101** to simultaneously send a message **104, 204** to the connect server **102** and a second server **202** known as the dedicated RTT server. On receipt of the message both the connect server **102** and the RTT server **202** forward the message directly back to the mobile device **105, 205**. On receipt of the return messages from both the connect and RTT servers **102, 202** the mobile device then proceeds to calculate the Round Trip Time.

While the calculated RTT in this instance may not be an entirely accurate view of the actual RTT, it will however, provide a measure of the general health of the network and once trends are established can be used to identify periods of good or poor performance relative to the baseline.

Fig 3 depicts the transaction flow for the RTT calculation and reporting process according to one embodiment of the present invention. As shown the mobile device firstly opens via a mobile client a HTTP connection **500** to the network **501**, if one is not already open. The network **501** creates the required Packet Data Protocol (PDP) context for the new HTTP connection **502**. Once the HTTP connection is established the mobile client sends **503** RTT test data, in the form of messages **104, 204**, to the connect server **102** and RTT server **202** using the TCP protocol. The mobile client then waits **513** for the reply for both the connect and RTT servers **102, 202**.

On receipt of the RTT test data **504** the connect server **102** sends an RTT reply **505** in the form of message **105**. Likewise the RTT server **202** on receipt of the RTT test

data **506** the RTT server **202** sends an RTT reply **507** in the form of message **205**. On receipt **508** of the messages **105**, **205** the mobile client then calculates the RTT times **509**. Once the mobile client has calculated the RTT times it proceeds to send an RTT report **510** to the RTT server **202** via message **106**. The RTT server **202**
5 then saves the report **511** to a database. Once this is complete the RTT server **202** may then proceed to send a reply message **512** to the mobile client via message **108**. The message **108** may contain RTT summary data which the mobile client can then display **514** on the mobile device's screen.

- 10 One example of the network monitoring tool in use is shown in Fig 4. In this case the network monitoring tool is as an application deployed on a plurality of mobile handsets roaming within a mobile communications network. For clarity of description the following discussion will focus on the calculation of the Round Trip Time (RTT) for a single cell with the network servicing plurality of mobile telephones. As shown
15 the cell **300** is composed of three sectors **301**, **302**, **303**. The cell being coupled through the network core to the connect server **102** and the RTT server **202**.

In this example the first slot **301** of the cell **300** is servicing two handsets **101a**, **101b** while the second and third slots **302**, **303** each service a single handset **101c**, **101d**.
20 As shown mobile device **101a**, **101b**, **101c** each send a message **104a**, **104b**, **104c** to the connect server **102** the connect server **102** then forwards each message **105a**, **105b**, **105c** directly back to the respective mobile devices **101a**, **101b**, **101c**. As with the case of the above examples each of the mobile devices running the mobile application may simultaneously send a message **204a**, **204b**, **204c** to the
25 RTT server **202**, the RTT server **202** immediately forwards each message back to the relevant mobile devices **205a**, **205b**, **205c**.

Alternatively the network monitoring application may wait for a timer to complete before opening a connection to the RTT server **202**. The timer is set to a
30 predetermined time interval, the length of the predetermined time interval is based on a number of factors such as the number of devices being sampled within the network, power consumption of the devices etc. The timer in effect controls the sampling rate for RTT data from the mobile device i.e. dictates send times of the

message to facilitate the RTT calculation. In either case the sampling time is kept large given the potential volume of samples that may be taken and to reduce the impact on the mobile device. In the example illustrated in Fig 4 the predetermined time interval is initially set to a period of 61 minutes. It is proposed that the RTT server may dynamically increase the sampling rate by reducing the predetermined time interval, particular in areas where it can be seen as potential network trouble spots in order to increase the data available on these areas for further analysis.

Since the objective of the monitor is to identify anomalous Round Trip Times, as well as to monitor regular network performance, the timeout for the message to be returned from the RTT server should be set as long as possible given network communication constraints but no longer than the time between tests (i.e. no longer than predetermined period). To ensure "Lag" situations are captured a suitable timeout value would be 10 minutes. A 10 minute timeout period is long enough to capture a typical "lag event" where data transmission is interrupted for a period (for example due to cell congestion). A longer timeout is more likely to result in the user logging out of the device so that the sample is lost. The timeout value should be configurable from the RTT server, however, the timeout set by the RTT server may be overridden by the mobile device if it is longer than the current predetermined interval time (sampling interval). Should the timeout be reached then this data sample should be reported as the timeout time. For example, if the timeout is 10 minutes and no response is received in that time interval then the RTT report should show that sample as 10 minutes for the Round Trip Time.

On receipt of the relevant return messages from the servers **102**, **202** each of the mobile devices **101a**, **101b**, **101c** proceeds to calculate the RTT. Once each mobile device **101a**, **101b**, **101c** has calculated the relevant Round Trip Time (RTT) they proceed to forward messages **106a**, **106b**, **106c** containing the relevant RTT information to the connect server **102**. In addition to the RTT information, the messages **106a**, **106b**, **106c** also includes the User ID of the mobile subscriber, the Mobile Identification Number (MIN) of the device doing the test, the connection details for the mobile device (GSM/UMTS/etc and Application used), phone model etc. The message may further include information concerning outbound time from

the server to the mobile device, inbound time from the mobile device to the server and packet loss i.e. number of packets lost that have to be re-transmitted. A short listing of the desirable information fields for the RTT reporting message is shown in Table 1 below.

5

Initial reporting will be an excel extract of the raw data for 24 hours on a daily basis from 12:00 midnight to midnight. Server results should be available as follows: Via real time web page; via daily reports (see layout below); Via excel format extracts (see layout and options below); via real time alert SMS to a defined list of mobile numbers when thresholds are breached. The reports indicate the spread of RTT times by cell site/sector. Thus it is possible to identify sites or sectors that are experiencing higher than acceptable RTT times. These sites should then be targeted for more detailed analysis by the network team.

10

Report Field	Description/Purpose
Time/date	Time/date report was sent from client (this will be shortly after the ping however is more consistent than using the time of the mobile device).
User ID	User logged on to this client
MIN	Mobile Identification Number
APN	Application used to send RTT message (could be Smart1, Smart Internet or any other carriers Application)
Mode (2G/3G/3.5G)	Ensure system can cater for future e.g. LTE, WiFi, Other?
RTT to Connect server	In ms
RTT to RTT server	In ms
Phone Model	Only Symbian s60/3 rd at present

15

Table 1: RTT Reporting Message Fields

The server has the option to turn off the RTT testing from either for all or selected user ID's. In particular it should be possible to only have the tool turned on for

specific users. This can be simply achieved by setting the RTT timer interval to 9999 minutes.

5 In one iteration of the tool, the monitoring client software is incorporated into a messaging application which, in a Symbian s60 phone is designed to operate as "always on" and connected to the mobile internet. The messaging application provides a framework that supports sending and receiving messages as well as optionally providing RTT reporting back to the mobile user by way of EM type messages. EM stands for Extended Messaging and is a proprietary form of Internet
10 based instant messaging.

In the production version of the tool, no RTT information is to be visible to the mobile client user. In test versions, however, a user is to be able to select Round Trip time, from the Settings menu which will result in a transaction being sent to the server
15 requesting current RTT information. This information is provided in an EM type response from the RTT Bot. Fig 5A to 5C depict a series of user screens **400** of a test version of the network monitoring tool according to one embodiment of the present invention.

20 As illustrated in Fig 5A the user has activated an options menu **401** within the display screen **400**. To commence testing the user highlights the option 'Round Trip Time' from the listing **402** and the presses the select button **403**. Once the Round Trip Time is selected, a message **405** is displayed on the user screen informing the user that the RTT request has been sent as shown in Fig 5B. In order to dismiss the
25 message, the user then selects the OK button **406**.

After the monitoring tool has calculated the round trip time from the information obtained from the return messages, the relevant information is then displayed to the user **407** (see Fig 5C). As illustrated in the example of Fig 5C the RTT reporting
30 message in this instance includes the date and time at which the sample was taken, the phone model, the date and time of the previous test, the RTT for the current test (connect RTT and Reg RTT), the application used, type of network access (in this

case GSM) and the average RTT for the day (average connect RTT and average Reg RTT).

As mentioned above the analysis tool may be utilised to identify lag events within the network. In such instances each message transaction to another mobile is treated somewhat like the RTT test data discussed above.

Under lag testing conditions the sending client records the time when it sent each message (ts) and then records the time when the acknowledgment is received. The client is then able to calculate the duration (td) since original message was sent to the other client. This RTT time may be impacted by any kind of network issues including lag. The impact could be either the sending device or the receiving device.

To optimize the amount of data captured by such a testing approach the client compares the duration td with a reporting threshold value (for example 3 minutes). If td exceeds the threshold, then the sending client transmits a notification to the RTT server to be stored and collated. Suitably the lag threshold, as with all similar parameter data, should be easily modifiable and may be set by the RTT server. The report may include such information as, the userid of sender and receiver, td (effectively RTT to the other client), cell site of reporting client (if known), phone model of reporting client and the MIN of sender and receiver.

It will be appreciated by those of skill in the art that the lag testing may also be performed at the connect server. Initiating testing in this manner would allow the server to calculate response times for every transaction as well as to identify which device of a pair is actually experiencing lag.

It is to be understood that the above embodiments have been provided only by way of exemplification of this invention, and that further modifications and improvements thereto, as would be apparent to persons skilled in the relevant art, are deemed to fall within the broad scope and ambit of the present invention described herein.

CLAIMS

1. A method for determining the performance of a communications network said method comprising the steps of:

5 transmitting a message from a mobile device to at least one server, each server within the at least one server being configured to direct the message back to the mobile device;

receiving at the mobile device the messages returned by each server within the at least one server;

10 calculating a time differential between transmission of the message by said mobile device and receipt of the messages returned by each server within the at least one server, by said mobile device; and

forwarding the calculated time differential to a primary server selected from the at least one server for storage.

15

2. The method of claim 1 wherein the message sent by the mobile device to each server within the at least one server includes the mobile device's identification number (MIN).

20 3. The method of claim 2 wherein the mobile device's identification number is registered with the primary server.

4. The method of any one of claims 1 to 3 wherein each server within the at least one server is configured to apply a time stamp to the message when directing
25 the message back to the mobile device.

5. The method of claim 4 wherein the step of calculating further includes the step of calculating a transit time for the message to reach each server within the at least one server.

30

6. The method of claim 4 or 5 wherein the step of calculating further includes the step of calculating transit time for the message to reach the mobile device from each server within the at least one server.

7. The method of any one of claims 1 to 6 wherein the time differential is indicative of the total round trip time for the message.

5 8. The method of any one of claims 1 to 7 wherein the step of transmitting the message to the at least one server is performed at regular intervals.

9. The method of claim 8 wherein the interval between each transmission is set to a predetermined time period.

10

10. The method of claim 9 wherein a secondary server selected from the at least one server is configured to dynamically vary the length of the predetermined time period based on variations in the calculated time differential.

15 11. The method of any one of claims 1 to 10 wherein the step of forwarding the calculated time differential to the primary server further includes forwarding information on the type of mobile device, the application used to send the message, the mode of connection between the mobile device and the communications network, the User ID of the mobile subscriber, the Identification Number of the
20 mobile device, cell identification and sector number in which the mobile is located.

12. The method of claim 11 further including the step of formulating reports for one or more cells within the network based on the time differential and the information on the type of mobile device, the application used to send the message,
25 the mode of connection between the mobile device and the communications network, the User ID of the mobile subscriber, the Identification Number of the mobile device.

13. The method of any one of claims 1 to 12 further including the step of
30 forwarding a set time differential to the primary server on the expiry of a preset period for receipt by the mobile device of the return messages from the at least one server.

14. The method of claim 12 or 13 wherein the secondary server is further configured to vary the preset period in accordance with a variation of the predetermined period.

5 15. A mobile device said device including a dedicated application for performing the method of any one of claims 1 to 14.

10 16. The mobile device of claim 15 wherein the dedicated application utilises one or more auxiliary applications on the mobile device to send the message to the servers.

15 17. The mobile device of claims 15 or 16 wherein said dedicated application is incorporated as an embedded application in a host application installed on the mobile device.

18. The mobile device of claim 17 wherein the host application includes messaging capabilities and the dedicated application is configured to use the messaging capabilities of the host application to provide feedback to the user regarding calculated time differential.

20 19. The mobile device of any one of claims 15 to 18, said dedicated application can be deactivated by each server within the at least one server.

25

30

35