ETSI TR 102 529 V1.1.1 (2009-03)

Technical Report

Speech and Multimedia Transmission Quality (STQ);
SMS Testing Guidelines;
Measurement Methodologies and Quality Aspects

Reference

DTR/STQ-00098m

Keywords

SMS, testing, scheduling

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://webapp.etsi.org/IPR/home.asp).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Report (TR) has been produced by ETSI Technical Committee Speech and multimedia Transmission Quality (STQ).

1 Scope

The goal of the present document is to point out all the aspects impacting SMS service measurement results, providing many different approaches for SMS testing.

2 References

References are either specific (identified by date of publication and/or edition number or version number) or non-specific.

- For a specific reference, subsequent revisions do not apply.
- Non-specific reference may be made only to a complete document or a part thereof and only in the following cases:
  - if it is accepted that it will be possible to use all future changes of the referenced document for the purposes of the referring document;
  - for informative references.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication ETSI cannot guarantee their long term validity.

2.1 Normative references

The following referenced documents are indispensable for the application of the present document. For dated references, only the edition cited applies. For non-specific references, the latest edition of the referenced document (including any amendments) applies.

Not applicable.

2.2 Informative references

The following referenced documents are not essential to the use of the present document but they assist the user with regard to a particular subject area**.** For non-specific references, the latest version of the referenced document (including any amendments) applies.

[i.1] ETSI TS 102 250 (all parts): "Speech Processing, Transmission and Quality Aspects (STQ); QoS aspects for popular services in GSM and 3G networks".

[i.2] ETSI TS 123 040: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); Technical realization of Short Message Service (SMS) (3GPP TS 23.040)".

3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

HLR	Home Location Register
MNRF	Mobile Not Reachable Flag
MO	Mobile Originated
MS	Mobile Station
MWDL	Message Waiting Data List

QoS	Quality of Service
SM	Short Message
SMS	Short Message Service
SMSC	Short Message Service Center
UDH	User Data Header
UDHI	UDH Information
VLR	Visitor Location Register

4 General aspects of SMS functionality

4.1 Maximum character lengths

The SMS service is designed to send up to 140 bytes of user data with a single SM. From a customer's point of view, this leads to different amounts of available characters available to compose text messages, based on the encoding used.

Table 1 provides an overview of the different available character sets including the maximum possible message lengths for a single SM.

NOTE 1: Even if not all the available data is used from a customer perspective, e.g. if the actual text used for testing is less than 160 characters, the size of the user data element will stay constant.

NOTE 2: Content integrity of single SMs is ensured by mechanisms on lower protocol layers of GSM and UMTS networks. Thus, there is - from an E2E testing perspective - no need to implement content integrity checking mechanisms on top of the SMS service.

Table 1: Overview of the different available character sets including the maximum possible message lengths for a single SM

Character set	Encoding	Maximum message length
Default GSM alphabet	7-bit	160 characters
ANSI (e.g. Cyrillic)	8-bit	140 characters
Unicode (e.g. Arabic, Asian)	16-bit (UCS2)	70 characters

4.2 Concatenated short messages

In order to transfer text messages not fitting into a single SM, the text message can be split into multiple SMs containing a so called UDH (User-Data-Header), which will lower the amount of available user data. The UDH is in general used to transfer additional information related to the user data of the SM and will vary in size. Please refer to TS 123 040 [i.2] for a list of allowed UDH information elements (UDHI) defining possible ways to use the UDH. According to TS 123 040 [i.2], the UDH contains for each SM of a concatenated SM the indication that this single SM is part of a concatenated SM and will also provide information about the position of the SM within the long message. Thus, fewer characters can be transferred per single SM if this single SM is part of a concatenated SM. Table 2 provides an overview of the maximum available characters of a single SM when using UDH to concatenate SM in order to form and send longer messages.

Table 2: Overview of the maximum available characters of a single SM when using UDH to concatenate SM in order to form and send longer messages

Character set	Encoding	Maximum message length using UDH for concatenation of SM
Default GSM alphabet	7-bit	153 (160-7) characters
ANSI (e.g. Cyrillic)	8-bit	134 (140-6) characters
Unicode (e.g. Arabic, Asian)	16-bit (UCS2)	67 ((140-6)/2) characters

NOTE: For further details on SMS, please refer to TS 123 040 [i.2].

4.3 SMSC queuing per destination/user

The SMSC platform maintains a single queue for every destination, storing the SMs in its memory buffers. Consequently, a newly arrived SM is added to the queue of the destination user. This queuing method limits the SMSC to deliver SMs to the same destination based on arrival order.

This queuing method causes the following behaviour:

- Before accepting an incoming SM, the queue length is verified and the message is accepted only if the queue for the specific recipient has not reached the maximum allowed length.
- An incoming SM could trigger the delivery mechanism of all the queued SMs for the same destination.

4.4 SMSC retry mechanism

In order to increase SMS reliability, SMSC platforms usually implement a mechanism to autonomously trigger the delivery procedure of an SM. This feature is usually called "retry mechanism" and the applied temporal plan is defined as a "retry scheme". Usually, many delivery trials are foreseen within a retry scheme to achieve the foreseen goal of improving the perceived QoS of the Short Message Service, while reducing time between SM submission and delivery to its destination and increasing the delivery success rate.

The applied "retry scheme" is implementation-dependent and established based on the network operator's plans. The time interval between two subsequent delivery attempts can vary from minutes to many hours. To avoid wasting network resources while optimizing the SMS delivery process, the retry scheme usually starts with time intervals of few minutes for the first attempts, gradually increasing up to many hours for subsequent trials.

The retry mechanism also aims at ensuring SM delivery in case of fault of the following "alerting mechanism".

4.5 Alerting mechanism

If an SMS delivery is unsuccessful for a multitude of causes (terminal switched off, subscriber not reachable, etc.), the SM is stored in the SMSC for a subsequent delivery attempt.

In order to trigger the new delivery procedure, different mechanisms can be utilized. The first is the above-mentioned retry mechanism, while the alternative is an alerting mechanism started by the HLR.

In the following part of this clause, a detailed overview of the alerting mechanism will be provided.

While delivering an SM to the foreseen destination, considering an MS as the receiving party, a submission failure can happen due to unavailability of the destination. For example, the terminal of the receiving party could be switched off or could not be reachable due to poor radio signal power. In this case, the SMSC, as a consequence of the failed delivery attempt, informs the HLR that the requested user is not reachable, setting the "Mobile Not Reachable Flag" (MNRF) for that user. Furthermore, it adds its identifier to the Waiting Data List; this action, performed by the SMSC, has to be interpreted by the HLR as "there is at least one waiting SM for this user stored in the buffers of this SMSC".

The steps of the above-mentioned procedure are represented in the following scheme.

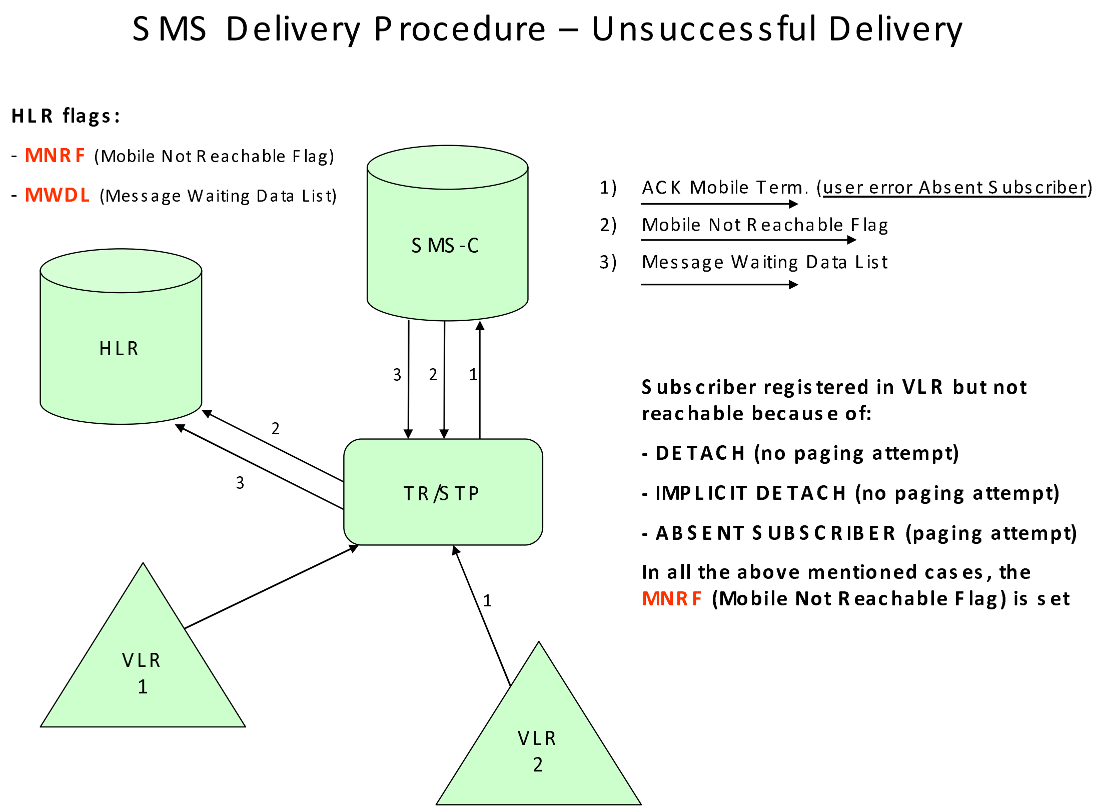

Figure 1: SMS Delivery Procedure - Unsuccessful Delivery

The above-mentioned MNRF flag is also implemented in the VLR. As a consequence, when the user becomes unreachable for one of the above-mentioned reasons, the VLR autonomously sets the MNRF flag.

The duplicate implementation of the MNRF aims at increasing the efficiency of the alerting mechanism. In the following, the details and motivations of this double information are explained.

4.6 Triggering of the delivery procedure caused by the alerting mechanism

In the following, the alerting method is explained by means of two different examples.

EXAMPLE 1: MS Procedure involving HLR:
Considering a terminal being initially switched off. After switching it on, the registering procedure will inform the HLR that the user became reachable. If, for the respective user, the HLR detects an MNRF flag set to 1, it will change this value to 0, taking into account the list of SMSCs storing SMs for that user, reported in the "Message Waiting Data List" (MWDL). Subsequently, the HLR will inform each SMSC in the above-mentioned MWDL, sending an alerting message using the MAP primitive "MAP-ALERT-SERVICE-CENTRE". As a consequence, each alerted SMSC starts delivering the stored SMs towards the foreseen destination.

EXAMPLE 2: MS Procedure involving only the VLR:
If a mobile user is temporarily unreachable, for example being located in an area with no or very poor radio signal conditions, and an SM delivery attempt fails during this period, the MNRF is set in the VLR. If the same user tries to make a call after returning into a covered area, the VLR deducts that the user is actually reachable again. Taking into account the state of the MNRF flag, the VLR will inform the HLR about the new state of the user. Thus, the HLR starts the triggering procedure towards the list of SMSCs in the MWDL. Finally, each alerted SMSC starts delivering the stored SMs towards the foreseen destination.

4.7 SMSC behaviour at arrival of a new SM directed to the same destination/user

While storing a previous SM at the arrival of a new SM, the SMSC should start a new delivery attempt for all stored SMs, usually using a first-in-first-out-approach in treating the messages. As a consequence, the delivery of an old message can be triggered by the arrival of a subsequent SM towards the same user. If this approach is not adopted, the delivery of a stored SM can be triggered only by the alerting mechanism or by the retry scheme.

The behaviour of an incoming, SMS-driven delivery mechanism could vary based on the specific implementation and operator settings. When establishing and defining the testing method to assess the QoS of SMS services, all the mechanisms acting on the delivery of an SM and impacting on perceived quality should be taken into account. Following this general criteria, the suggested testing method should prevent correlation between subsequent tests. Correlation between following tests would impact the validity of the measured values, as it could happen if we send a new SM before establishing if the previous one has expired.

5 SMS Testing guidelines

5.1 Settings for SMS test and correlation between subsequent tests

Taking into account all previously correlated mechanisms, particular attention should be paid to scheduling SMS tests. This means that a subsequent SM towards the same destination can affect the result of a previous SMS test if the second SM is sent before stating the result of the first one.

In order to obtain a more representative statistic, an alternative way to increase the number of tests is sending the SMs towards different destinations. Using this approach, the queuing per destination mechanism is bypassed.

A high frequency of SMS submission attempts is typically used to increase the number of measurement samples aiming both at increasing the statistical validity of the measured values and at efficiently using the expensive "drive test time". On the contrary, by reducing the above-mentioned frequency (in order to exclude forced delivery attempts caused by incoming SMs), testing time would not be used efficiently. This last approach, even if producing valid measurements from a statistical point of view, could cause low testing activity, with long inactive periods for SMS tests and the measurement system.

To define an SMS testing approach, the general validity of indicators and of relevant timeouts should be taken into account.

Even if alternative solutions could be analysed and proposed, the objective of the defined approach is to identify a testing method that is able to verify the QoS of the SMS service according to user experience, analysing End-To-End behaviour of the service without breaking the delivery chain.

For example, during a drive test, it could be possible to increase the frequency of SMS sending attempts to primarily assess the ability of the network to transport SMS from the originating party to the SMSC, without analysing the delivery part of the SMS service process. Following this approach, to test the delivery-side part of the SMS service and measuring the related indicators, the introduction of a traffic generator as a source of SMs to be delivered to receiving parties is needed. This traffic generator could be represented by a stationary SM sender or an SM generator directly connected to the SMSC platform (intrusive testing method).

On the contrary, using testing equipment simulating a real user, it is also possible to benchmark the performance provided by different operators for the same service without knowledge of the specific implementation of the service (like in the Thevenin/Norton theorem approach in circuit analysis).

NOTE: Using the last approach, for example, it is not possible to obtain the "SMS delivery time" as a primary indicator but only the "SMS End-To-End Delivery Time", as established by the current version of TS 102 250-2 [i.1].

The SMS performance can also vary in different hours of the day, according to specifically submitted traffic and relating to the network condition. To assess the perceived quality, a high temporal density of tests could better represent the real situation. As a consequence, the foreseen testing approach could take into account this aspect and consider that the highest possible SMS sending frequency would be the "best" one. The chosen testing method should also avoid an ameliorative estimation of the performance for SMS service.

So there would be the following practical proposals, to be read as alternatives:

- Use a test case design with long intervals between SMS postings, filled up with other testing activities. Of course this requires good control over the process of test case design. Even if, as suggested in the rest of this clause, the introduction of another test during the above-mentioned intervals could impact the results of the QoS measurement.
- Send SMs to a variety of destinations. If we have M drive test systems and N receiving units, each of the drive test systems could direct their SMs to the N destinations in a round-robin scheme. Prerequisite is that N is well above 1.
- As stated, once an SM is in the SMSC, it will not matter anymore how it got there. Use the drive test system to create SMs as before, with the highest possible frequency, and complement measurements by low-frequency stationary tests. So the drive test measures the ability of the UE to deliver SMs to the SMSC, while the stationary tests probe the probability of an SM being delivered in time.

5.2 Exemplary testing model - using one receiving terminal

The following model shows a testing setup using one terminal as SMS receiver from a single sender terminal.

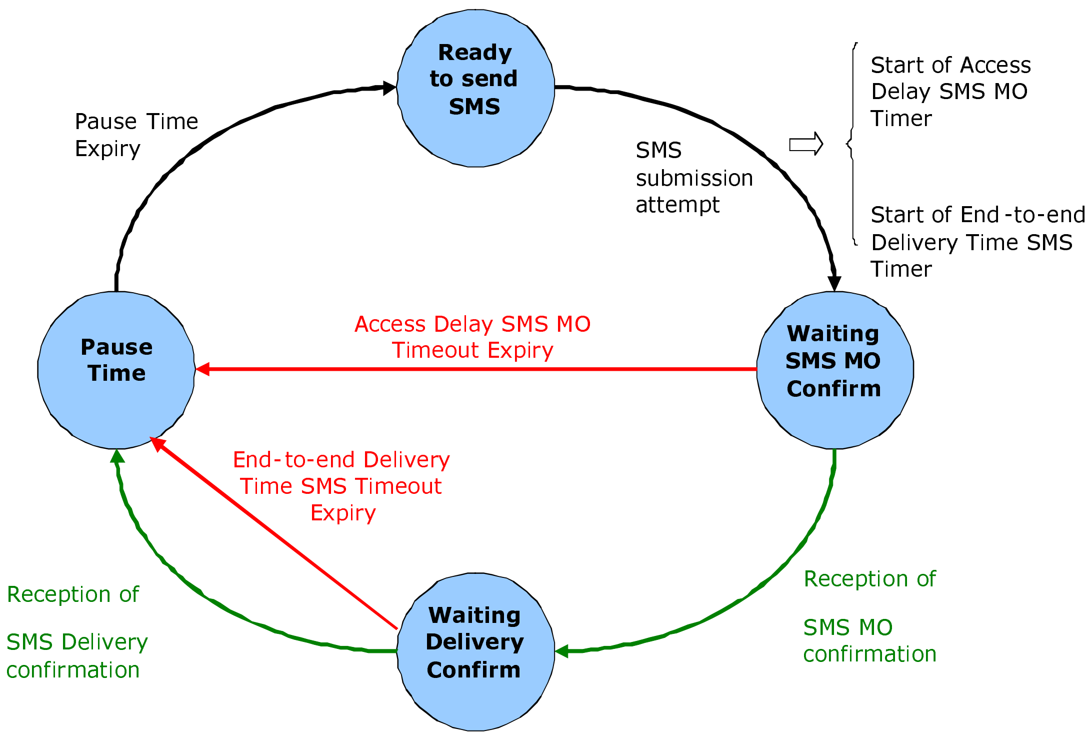

Figure 2: Exemplary testing model - using one receiving terminal

5.2.1 Effects of timeouts

Instead of using fixed measurement windows defined as the sum of the End-to-end Delivery Time SMS Timeout and a configurable pause, the method described in the previous clause uses a *best effort* approach in order to increase the number of measurement samples per test session.

Instead of waiting for the start of the next measurement window, the new test cycle would start with a configurable pause right after the delivery confirmation of the previously sent SM was received. In case of an SM delivery failure, the full "End-to-end Delivery Time SMS Timeout" would take effect between two subsequent SMs leading to the same length of a measurement cycle as if using the fixed measurement window approach described above.

As an example, assume an "End-to-end Delivery Time SMS Timeout" of 360 seconds and a pause time of 10 seconds. The two methods would produce the following results in one hour:

- Using fixed measurement windows, each measurement cycle would take the whole timeout plus the pause (360 + 10). Thus, in one hour, 9 SMs would be sent (3 600/370).
- Applying the proposed testing approach with an exemplary success rate of 100 % and assuming an average "End-to-end SMS Delivery Time" of 30 seconds plus the pause (30 + 10), 90 SMs would be sent per hour (3 600/40).

With respect to drive test measurement campaigns covering geographical regions with different coverage characteristics, including areas where "End-to-end Delivery Time SMS Timeout" timeouts are likely to happen, the proposed approach will lead to different sampling rates for successful and unsuccessful measurements, respectively.

NOTE: This is due to the fact that for unsuccessful samples, the respective measurement window is larger. This effect has to be accounted for in the statistical analysis of the campaign.

In general, fixed measurement windows should be used in such scenarios in order to achieve a representative distribution of samples with respect to the tested geographical regions. Furthermore, such measurements should be executed using a sufficient number of different receivers per sending party, addressing the receiving parties in a round-robin scheme without using delivery confirmation SMs in order to minimize effects described in clause 5.1.

5.2.2 Measurement flows

A measurement flow with varying short and successful test cycles and an unsuccessful measurement flow are shown below.

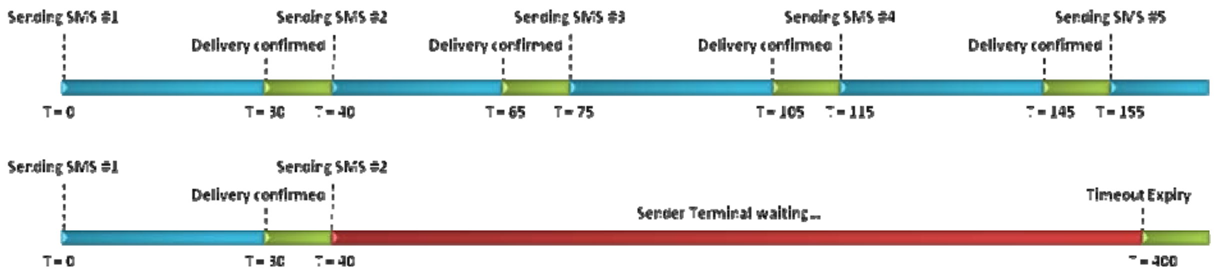

Figure 3: Measurement flow with varying short and successful test cycles and an unsuccessful measurement flow

Figure 4 compares different possible scenarios for subsequent test cycles.

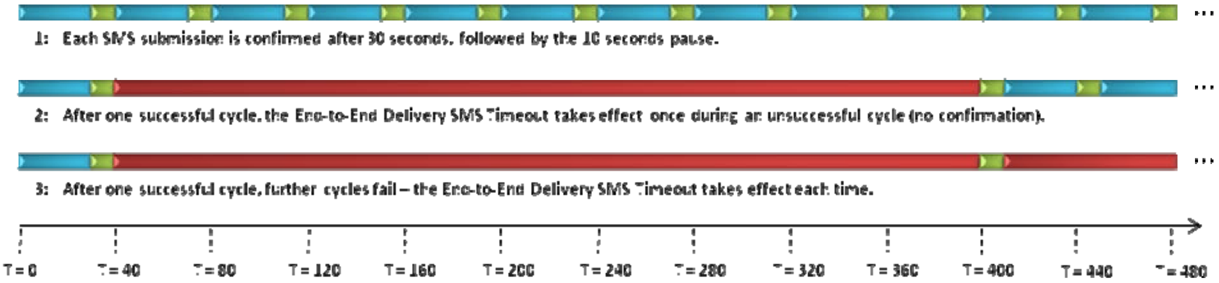

Figure 4: Comparison of different possible scenarios for subsequent test cycles

5.3 Exemplary testing model - using two or more receiving terminals

Sampling frequency can be further increased by utilizing two or more terminals as SMS receivers from a single sending terminal. After sending an SM towards terminal M and after submission has been confirmed by some terminal M, but before receiving delivery confirmation, the sender will already send another SM towards N.

In this approach, the pause after a confirmed delivery would be applied separately for each receiving device. The following example shows this method with two receiving mobiles M and N.

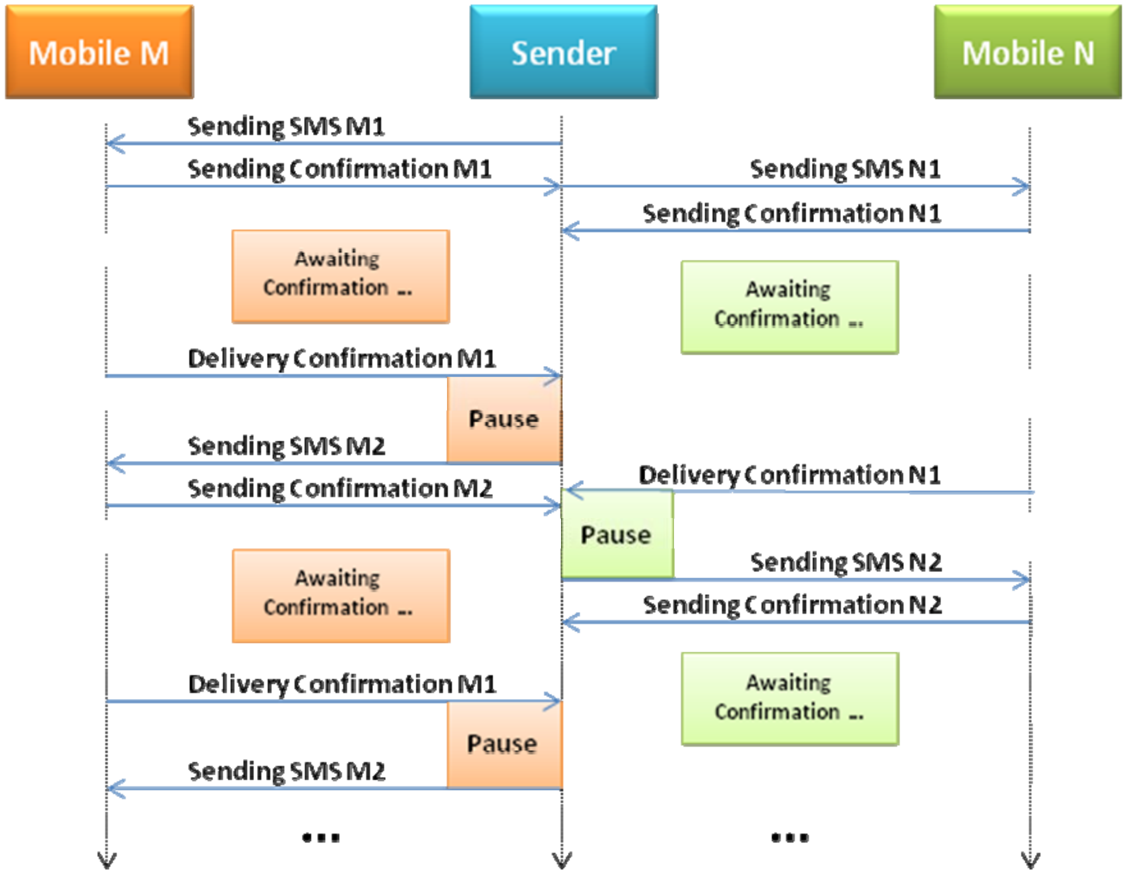

Figure 5: Exemplary testing model - using two or more receiving terminals

One drawback of this approach is that bypassing queuing problems on the receiver's side by using confirmation SMs can lead to queuing problems on the sender's side. Usually, the receiving parties do not know if some other receiving party has just sent a confirmation SM to the initial sender. Furthermore, the receivers also do not know if their confirmation SM has been successfully delivered to the initial sender. Thus, scenarios are possible where many receivers sent their confirmation SMs back to the sender at more or less the same time, which could then cause queuing problems on the side of the initial sender.

The possibility for such effects will increase in scenarios where it is likely that also the sender is not always reachable, e.g. when executing test measurement campaigns covering geographical regions with different coverage characteristics. Please also refer to clause 5.2.1 (Effects of timeouts).

Another drawback when using more than one receiving terminal is the possibility that the inactivity period desired before sending a new SM would vary or not even occur due to incoming confirmation SMs. Such scenarios can be compared to a single-sender scenario with a very small or no pause at all between single sending attempts. Thus, measurements with multiple receivers could lead to significantly shorter end-to-end delivery times and are thus not recommended if it is desired to measure end-to-end delivery times.

NOTE: Similar effects can be observed when measuring with a single terminal used as sender and also as receiver. If the terminal sends an SM to itself, the transfer time is usually much lower compared to measurements performed with dedicated sending and receiving terminals.

5.4 Testing methods for concatenated SMs

Depending on the used character set, the amount of charters fitting a single SM will vary. Thus, it will not always be possible to fit the same number of characters into a single SM. Especially with respect to more consumptive character sets, the likelihood for the end user to use concatenated SMs will increase. For this reason, testing of concatenated SMs should be taken into account.

In order to test concatenated SMs, a similar approach as when testing non-concatenated SMs can be used. In such a scenario, all SMs defining a single message should be sent to the receiving party one after another and in correct order. Afterwards, the sender should wait until all SMs have been successfully delivered before pausing and starting over.

From a technical point of view, each SM should be sent according to the testing model described in clause 5.2. Here, the attempt of sending the $(n+1)^{th}$ SM should be delayed until the sender has received the positive sending confirmation of the n^{th} SM, which corresponds for the n^{th} SM to entering the "Waiting Delivery Confirm" state with respect to the state machine described in the referred section.

In case that any of the SMs was not successfully sent, the whole attempt of sending a concatenated SM should be counted as a failure. In any case, the next attempt of sending a concatenated SM should be delayed until all SMs being sent in the current attempt have entered the "Ready to send SMS" state.

History

Document history		
V1.1.1	March 2009	Publication