QOS AND QOE ASPECTS OF DIGITAL FINANCIAL SERVICES

ITU-T FOCUS GROUP ON DIGITAL FINANCIAL SERVICES

International Telecommunication Union

ITU-T

TELECOMMUNICATION STANDARDIZATION SECTOR OF ITU

FG-DFS

(05/2016)

ITU-T Focus Group Digital Financial Services

QoS and QoE Aspects of Digital Financial Services

Focus Group Technical Report

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The procedures for establishment of focus groups are defined in Recommendation ITU-T A.7. TSAG set up the ITU-T Focus Group Digital Financial Services (FG DFS) at its meeting in June 2014. TSAG is the parent group of FG DFS.

Deliverables of focus groups can take the form of technical reports, specifications, etc., and aim to provide material for consideration by the parent group in its standardization activities. Deliverables of focus groups are not ITU-T Recommendations.

QoS and QoE Aspects of Digital Financial Services

About this Report

This Technical Report was written by Joachim Pomy and Wolfgang Balzer.

Special thanks to Jan Holub and Peter Pocta for their helpful review and contribution.

If you would like to provide any additional information, please contact Vijay Mauree at tsbfgdfs@itu.int

CONTENTS

Page

List of Tables

List of figures:

Executive Summary

This Report summarizes the Quality of Service (QoS) and Quality of Experience (QoE) aspects of Digital Financial Services (DFS) as concluded by the ITU-T Focus Group on Digital Financial Services (FG DFS).

Guidance and suggestions are provided for stakeholders involved in DFS taking into account regulatory and consumer related aspects.

It analyses different use cases and the applicability of currently available standards.

The report details that persisting problems with the KPIs basic functionalities of a mobile network need to be resolved by the stakeholders in the interest of any mobile service and are therefore out of scope of QoS-for-DFS-considerations.

Since the number of technical KPI is overwhelming and target values cannot be set on a global level, the report provides a novel scheme, which enables stakeholders in any region or country to assess the fitness of networks, terminals, users, DFS implementations and society / government of the use of DFS implementations.

In addition, a motivation for future KPIs is discussed from a technology-agnostic point of view.

At various places in the report motivation for future standardization is included which is expected to be actively taken up by ITU-T Study Group 12.

1 Introduction

This Report summarizes the Quality of Service (QoS) and Quality of Experience (QoE) aspects of Digital Financial Services (DFS) as concluded by the Focus Group DFS.

Guidance and suggestions are provided for stakeholders involved in DFS taking into account regulatory and consumer related aspects.

The objective is to provide guidance mainly for Telecom Regulators but also to Service Providers of DFS. One main topic is the selection of Key Performance Indicators (KPIs) which should be focussed on.

Besides that, the report contains comments and notes which might not be appropriate for immediate guidance; this material is considered of importance for future work.

- Annex A discusses existing standards which are related to DFS.
- Annex B introduces underlying functionalities of DFS applications.
- Annex C summarizes a possible selection of a set of KPIs appropriate for DFS

1.1 Relationship of QoS and QoE

In addition to the term QoS, the term Quality of Experience (QoE) is often used nowadays in order to stress the purely subjective nature of quality assessments in telecommunications and its focus on the user's perspective of the overall value of the service provided.

The increased significance of the term QoE is related to the fact that in the past the term QoS was used mostly for only technical concepts focused on networks and networks elements. The definition of QoS, however, does include the degree of satisfaction of a user with a service. Thus, non-technical aspects are included, like e.g. the user's environment, his expectations, the nature of the content and its importance. But most service providers did use the QoS only in relation to the actual user-service interaction in order to cross-check whether the user requirements have been met by the service implementation of a provider (as perceived by the user). So there was a strong focus on the actual network performance and its immediate influence on user perceivable aspects while additional subjective and not directly service related aspects were omitted.

QoE is defined in in Appendix I of Recommendation ITU-T P.10 as the overall acceptability of an application or service, as perceived subjectively by the end-user. It includes the complete end-to-end system effects (client, terminal, network, services infrastructure, etc) and may be influenced by user expectations and context. Hence the QoE is measured subjectively by the end-user and may differ from one user to the other. However, it is often estimated by a combination of objective measurements and metrics describing subjective elements.

> NOTE: The definition of QoE and, in particular, the dividing line between QoS and QoE is, however, quite fuzzy, and up to today it does not appear that a globally accepted definition exists. For example, the Recommendation ITU-T E.800 does not use the term QoE at all; instead, it uses a 4-viewpoint model (similar to the one in Recommendation ITU-T G.1000) with terminology, like QoSE (E=experienced) or QoSP (P=perceived). In any case, the amount of energy put into the QoS/QoE discussion in the context of the FG DFS should be limited, since this is already on the agenda of ITU-T Study Group 12 and several other organizations.
> For working purposes, preferably the use of QoS can be limited to things which can be measured by machines or technical means (including e.g. speech quality metrics, like POLQA, Rec. ITU-T P.863, which already contain some perceptual considerations), and QoE should be used for items further down a “processing chain” where some kind of assessment has been applied. This assessment can be, for instance, some kind of usually nonlinear (clipping) function expressing limits where service quality is either “inacceptable” anyway, or so good that a further improvement will not have any practical consequences. It is important to note that such limits will be

strongly dependent on previous experience, i.e. will vary between regions or countries, and will also vary with time as people get accustomed to improvements. Therefore, the issue of "typical values" or "threshold values" is characteristic for the QoE domain.

Objective measurements deal with quantities which can usually be determined by technical measurements, such as information loss and delay. Subjective elements are components of human perception that may include emotions, linguistic background, attitude, motivation, etc. which determine the overall acceptability of the service by the end-user. An important part of subjectivity are expectations which usually are formed by previous experience of users for the same or similar types of service.

The following figure shows factors contributing to QoE. These factors are organized as those related to Quality of Service and those that can be classified as human components. QoE for voice and video is often measured via carefully controlled subjective tests where voice or video samples are played to viewers, who are asked to rate them on a scale. The ratings assigned to each case are averaged together to yield the mean opinion score (MOS).

Quality of service (QoS) is defined in Recommendation ITU-T E.800 as the collective effect of performance which determines the degree of satisfaction of a user of the service. In general, QoS is measured in an objective way.

In telecommunications, QoS is usually a measure of performance of services delivered by networks QoS mechanisms include any mechanism that contributes to improvement of the overall performance of the system and hence to improving the end-user experience. QoS mechanisms can be implemented at different levels.

EXAMPLE: At the network level, QoS mechanisms include traffic management mechanisms such as buffering and scheduling employed to differentiate between traffic belonging to different applications. Other QoS mechanisms at levels other than the transport include loss concealment, application Forward Error Correction (FEC), etc.

QoS parameters are used to describe the QoS observed. Similar to the QoS mechanisms, QoS parameters can be defined at different layers. Figure 1 below shows the factors that have an influence on QoS and QoE.

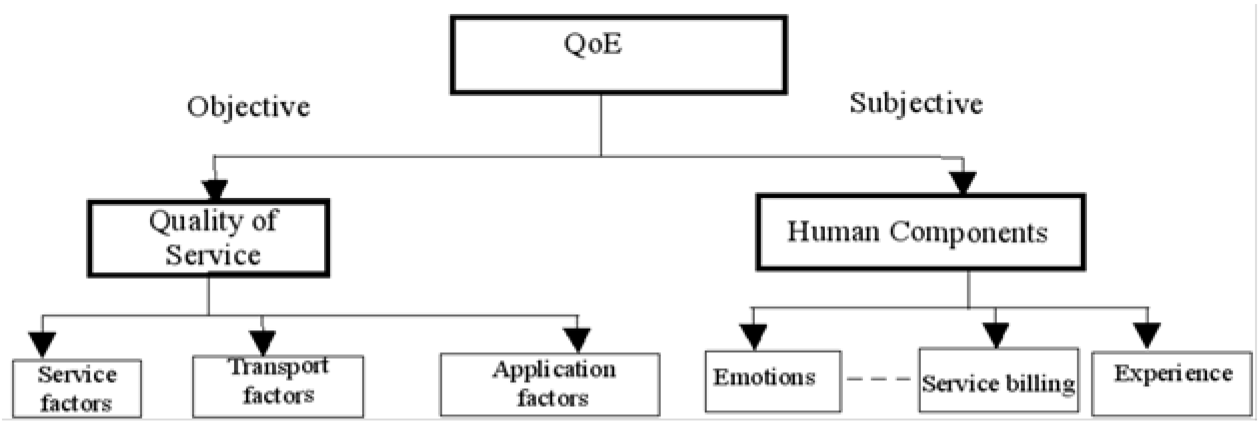

Figure 1 - Factors that have an influence on QoS and QoE

In general, there is a correlation between the subjective QoE as measured by the MOS and various objective parameters of Quality of Service.

Typically, there will be multiple service level performance (QoS) metrics that impact overall QoE. The relation between QoE and service performance (QoS) metrics is typically derived empirically. Having identified the QoE/QoS relationship, it can be used in two ways:

1) Given a QoS measurement, one could predict the expected QoE for a user.

2) Given a target QoE, one could deduce the net required service layer performance.

These prediction and deduction steps are built on assumptions and approximations.

Due to the complexity of services and the many factors which have an influence on QoS/QoE, there is not a close one-to-one relationship which would allow statements like "If the bandwidth is increased by 200 kbit/s, then the rating by the user will rise by 0.5 points".

To ensure that the appropriate service quality is delivered, QoE targets should be established for each service and be included early on in system design and engineering processes where they are translated into objective service level performance metrics.

Quality of Experience is an important factor in the marketplace success services and is a key differentiator with respect to competing service offerings. Subscribers to network services do not care how service quality is achieved. What matters to them is how well a service meets their expectations (e.g. in terms of price, effectiveness, operability, availability, and ease of use).

1.2 Services, Applications or "Popular Services"

Within the formal standardization community the term "Service" was always understood as a functionality for which all aspects are standardized (i.e. standardized service); the concept behind was that globally all networks would (be able and willing to) offer exactly the same – fully interoperable – harmonized service.

However, over time the terminology got corrupted in a sense that service today stands for any application. For example the Internet Engineering Task Force (IETF) refers to their standards which basically describe network functionalities as services.

Under end-user aspects the term service is used for any application offered in the networks; this makes it very difficult to standardize assessment methods and target values or requirements for related KPIs.

Therefore, if we speak about services, today, we can distinguish multiple dimensions:

a)	applications with global reach	vs.	b)	locally limited applications
c)	specifically named applications	vs.	d)	application class denominators

Typical examples are

a) Netflix or YouTube™

b) eGovernment application in country xyz

c) Netflix or YouTube™

d) Video streaming, IPTV

Since services in all these dimensions are not being standardized in their functionality a-priori, the communities involved in assessing QoS and QoE for such services have focussed on what is called "popular services". The concept behind is to provide assessment methods and targets at for such services which are used frequently by a huge number of users.

- Looking first at dimension a) with the examples given above, these are truly "popular services" – however the underlying technical aspects, such as carrier services may be changed from time to time.

- For dimension b) the main obstacle is the limitation itself. It is highly probable that there will not any international standard to measure the QoS or QoE of exactly one of that specific services.
- Dimension c) requires close cooperation between the stakeholder providing these services and the standardization experts.
- Proper dealing with dimension d) requires the standardization of new end-to-end mechanisms. Otherwise the existing carrier services will be confronted with more stringent targets for existing services.

1.3 Is DFS a "Popular Service"?

DFS is popular, yes – but DFS is only a class denominator.

> NOTE: At the time work on mobile QoS started (about 10 years ago), the experts considered "service" as something which has a direct impact to the customer's perception. Typical examples would be telephony or web browsing. A "service" in this view is understood as something connected to an end to end use case. However, many end to end use cases relate to "carrier services" (such as some type of packet data functionality having their own QoS metrics (KPI).
>
> In this context DFS can be considered as a classical example of such a user-related service, which can be realized in several ways, using "carrier services" such as SMS or packet data functionality of networks.
>
> DFS is not alone in this "top level service" view. Today's telephony is a prominent example. End users basically do not care if the function they are looking for (being able to orally communicate with another) is realized using legacy GSM or UMTS, VoLTE or some OTT VoIP technology. Their quality assessment is based on universal metrics such as setup time, call drop rate or speech quality, which are exactly those metrics which are at the core of standards such as Recommendation ITU-T E.804 or ETSI TS 102 250.
>
> The sometimes very detailed KPI definitions in these standards are owed to a "diagnostic" approach, but by no means not "the golden rule". Future developments will attempt to reveal true "end customer" related Key Quality Indicators (KQI).
>
> An additional example for this may be web browsing using HTTPS instead of HTTP. For the user, nothing seems to have changed, so top-level QoS KPI to assess user perception are the same - however, the networks are treating HTTPS and HTTP traffic in many cases differently, which will lead to a difference in usage of such KPI for diagnostic purposes.

If we want to technically assess the expected top-level QoS of a particular DFS offering using a carrier service point of view, we need to know the technical flow of data and signalization. This information is not normally available from service providers' websites or brochures.

> NOTE: Strictly speaking this is true for most of the other services offered by network operators. First of all, operators typically do not commit themselves (at least not towards end customers) to strict performance targets; in the case of mobile networks this is perfectly understandable as the local conditions vary in a wide range (e.g. from rooftop to cellar of a house even in the same geographical spot). Then, with networks going even more towards "content sensitive" behaviour for the sake of resource optimization, the performance cannot safely be predicted from just some general "bit pipe" properties, measured using simple end to end services such as web browsing. However, DFS can be - as will be shown later - made subject to objective measurement quite easily.

Ideally, this must be dealt with when licenses are negotiated between regulators and potential DFS service providers.

> NOTE: This is well known and understood for other services like for example, video streaming:

When "YouTube™" first became popular it was based on TCP streaming; with this information KPIs could be defined in standards, QoS could be assessed and QoE could be predicted. Today, for good reasons, the same service by the same entity is rendered as adaptive streaming using HTTPS. Consequently, new standards have been written with new KPIs in order to assess QoS for the "same service".
Strictly speaking, the KPI with respect to video quality are still the same; only the methods have changed (or were forced to change). Most importantly, KPI definitions using "low level" technical events as those from the IP level do not work anymore if encrypted connections such as HTTPS are used.

If we can identify categories of different DFS offerings, we could conclude, which of such categories constitute "popular services" (i.e. which are widespread and used by many customers) and start a more selective look into KPI definitions.

2 Problem statements

2.1 Different use cases

QoS aspects of DFS need to be assessed for two different use cases:

1) In use case #1 the targeted group of users of such service is limited to the use of (cheap) basic feature phones. This excludes for example browser-based DFS solutions.

2) In use case #2 the additional QoS aspects are assessed when the minimum requirements to the phones used for DFS are raised and basic smartphone functionality can be assumed.

2.2 Legal entities

Today, one can observe that the provision of a service offer ("service") is – as a general rule – independent from the physical operation of a telecommunication network.

Whereas for most service offers there is – beside the general legal framework – no specific regulation, DFS "services" are under the close control of the regulators of the banking sectors, whereas operators of telecommunication networks are under the control of the regulators of the telecom sectors.

Therefore, legal aspects (from a QoS perspective) need to assess two different legal cases:

a) In legal case #a: the provider of a DFS "service" and the operator of a physical telecommunication network are two distinct and different legal entities.

b) In legal case #b: the provider of a DFS "service" and the operator of a physical telecommunication network are the same and identical legal entity.

2.3 Mobile Network QoS affecting all services

The figure 2 (adapted from Recommendation ITU-T E.804 and ETSI TS 102 250) shows a model for quality of service parameters. This model has four layers.

The first layer is the Network Availability, which defines QoS rather from the viewpoint of the service provider than the service user. The second layer is the Network Access. From the service user's point of view this is the basic requirement for all the other QoS aspects and parameters. The third layer contains the other three QoS aspects Service Access, Service Integrity and Service Retainability. The different services are located in the fourth layer; the performance of these services is characterized by service specific QoS KPIs.

The first three layers (highlighted with green boxes) are common to ALL mobile services or applications.

They are characterized typically by the following parameters (KPIs):

- network availability
- network accessibility
- service accessibility
- service integrity
- service retainability

In cases where the KPIs in layers 1, 2 and 3 are not maintained at a stable high level it is useless to make attempts to assess the QoS of any kind of services because the statistical relevance of QoS figures received will be close to zero.

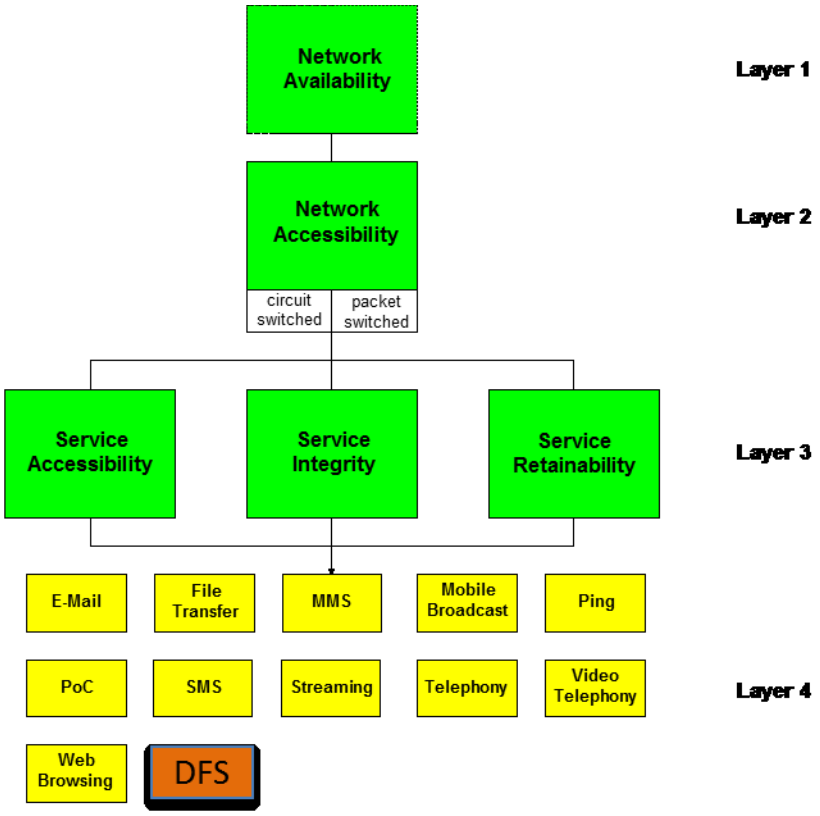

Figure 2 - Model for quality of service parameters

Persisting problems with the KPIs for layers 1, 2and 3 of a mobile network need to be resolved by the stakeholder in the interest of any mobile service and are therefore clearly out of scope of QoS-for-DFS-considerations.

NOTE: This diagram is in the process of being updated. First of all, layers 1 to 3 describe actually a kind of “pyramid of needs”, i.e. before one can start to think about service integrity (e.g. call drop rate in telephony), the service needs to be accessible first. Also, the “service” picture needs an overhaul. The “circuit/packet switched” division is legacy from 2G or 3G. Some of the “services" in Layer 4 actually depend on each other or belong to different groups. There are “carrier services” such as basic IP, and also combined services using one or more such carrier services, e.g. MMS relies on SMS (which is actually an end user related service as well) for notification, and uses packet data to actually transfer data. A “service” with the same effect for end users, e.g. some kind of OTT chat with attached files, uses only basic packet data.

In any case, there is no real "technology dependency" anymore. If an operator decides to suppress Skype, or prioritizes certain video streaming, this is not the result of some fundamental ability or inability, but just the effect of some "traffic shaping" elements.

2.4 Possible Solutions

Digital Financial Services are realized through utilization of basic services provided by a network. Assuming that the reliability of DFS has to be very high, there are two basic ways to ensure this reliability.

- By default, the QoS level for these basic services needs to be very high too. This is where we are with DFS today.
- The alternative is the use of robust end to end protocols which ensure the reliability of the actual service even in the presence of deficiencies in the underlying functionality. An analogous example would be the TCP protocol level which ensures lossless data transmission even in the presence of packet loss in lower layers. This needs to be developed and standardized.

Such robustness can be described by key criteria for DFS. Topmost is, for each transaction, a clear indication if it was successful or not, which needs to be consistent for both sides. Assume a transaction is composed of a number of steps, each step being the exchange of a data token. If the transfer of a data token has no clear "lost" criterion, but can take, in principle, indefinite time, a time-out needs to create a defined situation. The essential property of robustness is that, if a data token now arrives after its time-out, the protocol needs to ensure that this token is not causing any action any more.

With respect to practical aspects of DFS implementations, this poses some fundamental differences. When the main goal is to introduce DFS in the near future, it needs to operate with the existing installed base of end-user devices. This will automatically limit the spectrum of applicable methods to those which can be supported by those devices. A possible drawback of this approach is, of course, the fact that if a technology has been deployed and is widely used, it will – as long as is working without major problems - be difficult to replace it even if the new technology is superior. This may be less an issue with respect to end user devices as the penetration of smartphones will be continue to increase strongly due to their manifold advantages. It may be the case that these retaining factors are more on the side of infrastructure, as introduction of new technologies requires new investment which may, at least in the first years of usage, not be balanced by likewise new opportunities to generate additional revenue.

3 Conclusions

The following conclusions are, with respect to the preceding clause, based on the assumption that necessary DFS performance is achieved by ensuring a sufficiently high performance of the basic services used to implement DFS. The case of using a robust end to end protocol is not treated here.

3.1 Conclusions for use case #1

Four different techniques are discussed in Annex B which might be used in conjunction with DFS offers for use case #1.

- SMS is a store and forward service. Even if the share of short transfer times may be high in typical cases, it cannot – without modifications – be used reliably for real-time transactions.
- DTMF has limited transfer capabilities and will most probably only be used to complement one of the other techniques.
- IVR typically requires reasonably high listening quality which might pose a problem with feature phones in environments with higher levels of background noise.

- USSD is a true real-time technique. However, the message transfer which could be used for DFS are not standardized.

3.2 Conclusions for use case #2

Seven different techniques are discussed in Annex B which might be used in conjunction with DFS offers for use case #2.

As per availability on smartphones, HTTPS based solutions appear to be the optimal carrier technology for DFS.

3.3 Conclusions related to the fitness for DFS

A successful introduction of DFS via a mobile network requires fitness of the whole environment used, which is

– Fitness of the mobile network, to provide a minimum level of availability and accessibility
– Fitness of the mobile network to provide the services required for realization of DFS
– Fitness of mobile devices used, to support the basic services used to realize DFS
– Fitness of the DFS service itself to provide useable interfaces
– Fitness of users to successfully use DFS. This may include the necessary skills to operate DFS on phones as well as basic understanding of properties of DFS in general, to protect users against exploitation of insufficient knowledge
– Fitness of the general society and the governmental institutions for DFS

The following subsections contain decision diagrams, figures 3 to 7 which are meant to facilitate the discussion between stakeholders in the different regions or countries. The diagrams do not contain any numbers or specific target values. This is by intention, because target values acceptable for all stakeholder will vary from region to region and from country to country.

The term "Major Events" used throughout the five diagrams refers to work in progress in ITU-T SG12, Question 12, aiming at QoS in mobile networks during major events, as for example, major sports events.

3.3.1 Fitness of a mobile network for DFS

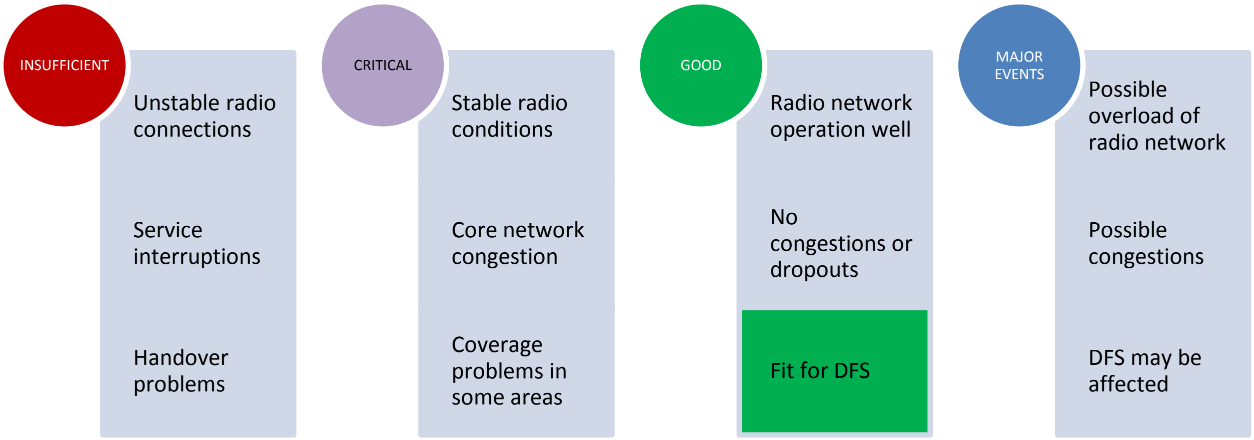

Figure 3 - Decision diagram for fitness of a mobile network for DFS

3.3.2 Fitness of mobile terminals for DFS

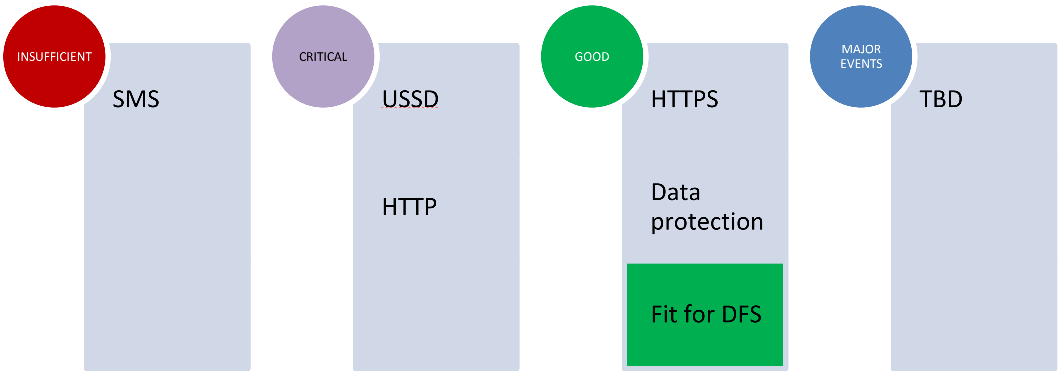

Figure 4 - Decision diagram for fitness of mobile terminals for DFS

3.3.3 Fitness of mobile services for DFS

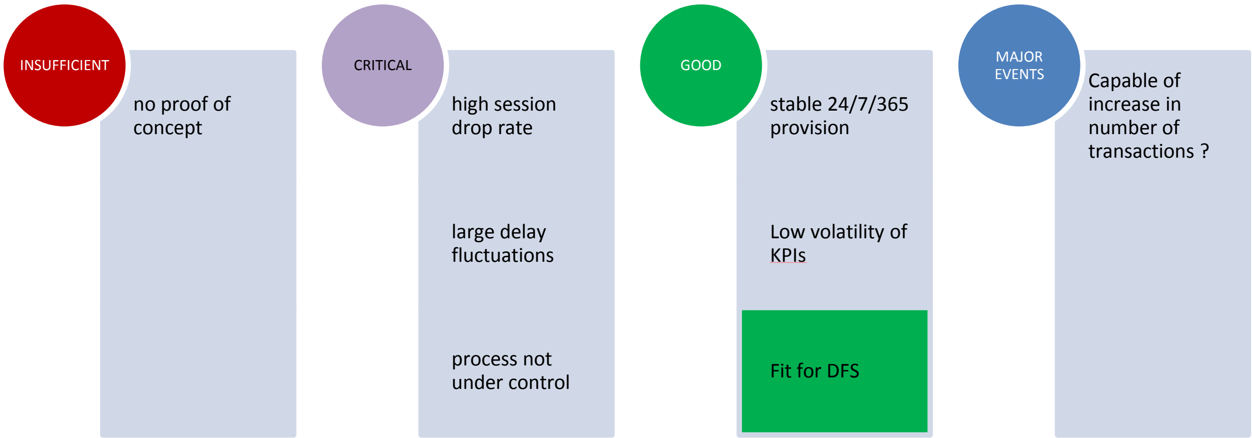

Figure 5 - Decision diagram for fitness of a mobile services for DFS

3.3.4 Fitness of mobile users for DFS

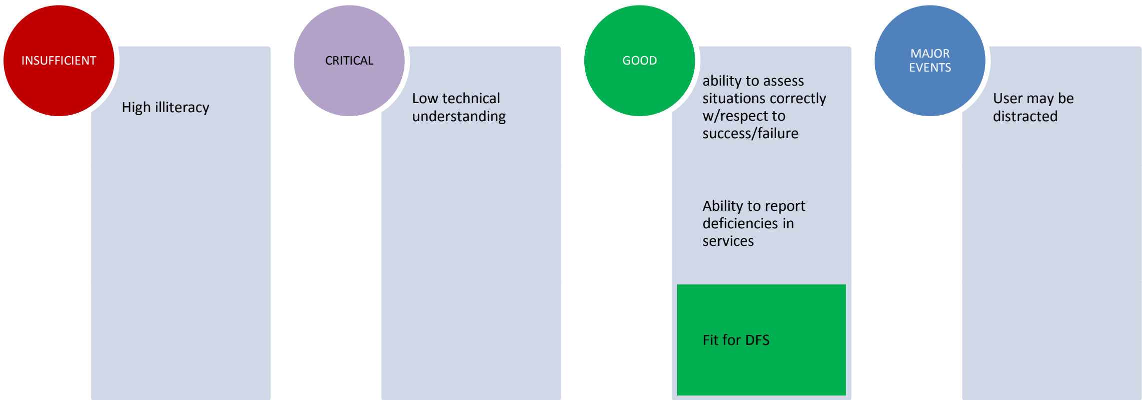

Figure 6 - Decision diagram for fitness of a mobile users for DFS

3.3.5 Fitness of society / government for DFS

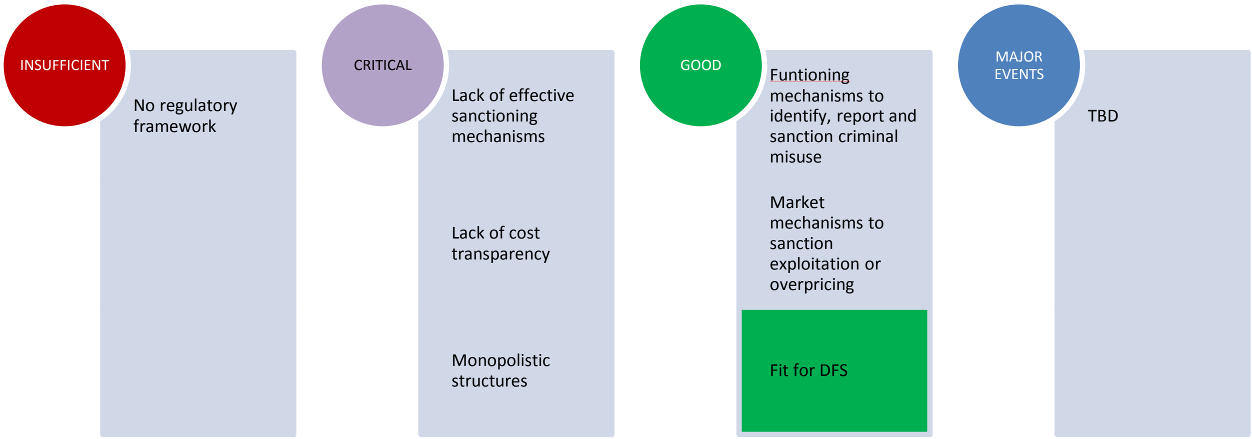

Figure 7 - Decision diagram for fitness of a society / government for DFS

3.4 Conclusions related to Digital Financial Services

It is of importance for any further work in the field of QoS / QoE for DFS to get access to more detailed information, such as descriptions of the various DFS offers to see on a technical level, which underlying services in the network are used and which are the technical parameters associated with them, e.g. timer values, timeout events, number of interactions involved in a single financial transaction.

Therefore, it is suggested that telecom regulators collect such information prior to the issue of licenses in order to make their own judgment of the quality of the planned DFS offering.

Such flowchart information should be submitted by regulators to ITU-T SG12, where the experts could start categorizing the different approaches and provide comments and guidance on such implementations.

There are even more issues remaining currently open, which will need further discussions:

- Mobile operators have increasing problems with the huge amount of data traffic in their networks. Therefore, if high speed fixed networks are available, there is a massive trend to use so-called **WIFI offloading**, where data traffic is redirected via WIFI accesses to the internet backbone core. The consequences for DFS seem to be quite unexplored, by now.
- The text displayed in the course of DFS interactions or the accentuation in spoken dialogue systems may be loaded with emotions, which could affect the users' experience of the service (QoE). Emotion detectors could be used to minimize any negative impact from this text and speech material. Currently, Requirements for **emotion detectors in telecommunications** are under study in an ETSI project (STF#504), which will provide a new Technical Specification (ETSI TS 103 296).

- A serious problem (mostly for regulators) are effects which cannot easily be allocated to one of the stakeholders in the DFS process. A prominent example are so-called **early timeouts in the DFS**, which anyone outside the DFS provider would interpret as dropped-calls, i.e. blame the network or blame the terminal or blame the user - in reality it turns out just to be a badly designed flow-of-actions: users still reading instructions on their screens before initiating the next step of a transaction are hit by an invisible timer's timeout action.

Because the field of DFS and its related QoS and QoE aspects is both of high importance and quite complex, **capacity building** is essential. Therefore, it is suggested that the ITU start the development of **online e-learning courses** in this area.

4 Guidance and suggestions

4.1 Use case #1

In this situation, where it is assumed that the end-user has phone with limited capabilities and uses that phone directly for DFS, it is important to make the use of USSD mandatory whenever possible at least for the initial part of the transaction.

Following-up communication with DFS user, like balance statements etc. may be done via SMS, but encryption should also be imposed as well as other mandatory features of SMS, like delivery attempts until confirmation.

- KPIs for USSD are under study.
- KPIs for SMS that should be monitored are the following:
  - SMS service non-accessibility [%]
  - SMS completion failure ratio [%]
  - SMS end-to-end delivery time [s]
  - SMS receive confirmation failure ratio [%]

4.2 Use case #2

In this situation, where it is assumed that either end-users or agents have access to smartphones it is suggested to mandate the use of HTTPS as the only protocol to be used for DFS.

KPIs for HTTPS are not easy to monitor, due to the use of the Secure Socket Layer (SSL) protocol:

- HTTPS Service non accessibility [%]
- HTTPS set-up time [s]
- HTTPS session failure ratio [%]
- HTTPS session time [s]
- HTTPS data transfer cut-off ratio [%]

NOTE: The above implies that the networks are a "uniform bit pipe" which can be characterized with one type of service (e.g. web browsing or upload/download) and the results being extrapolated to the behaviour of other services using the same "carrier (the http protocol level of packet data). This cannot be taken for granted in all cases.
Also, the fact has to be considered that a typical DFS use case needs a couple of round trips for an end to end completion of a transaction.
Of course there is the problem that actual end to end tests for DFS also mean that real money is transferred. However, to establish the necessary level of trust or safety, it needs to be seen how close to the "real thing" tests have actually to be. At least a kind of monitoring, tracking how accurate such extrapolations actually are is suggested.

4.3 Guidance related to mobile networks

This section discusses **possible ranges for target values for selected KPIs**.

NOTE: With an end to end DFS service description available, which would identify the "component services", a somewhat more "integrated" view could be provided. In such a view, target values would be integrated, coming from a "top level" view which can be related to requirements on lower levels.

4.3.1 USSD service non-accessibility [%]

The USSD service non-accessibility is the probability that the end-user cannot access the Unstructured Supplementary Service Data (USSD) when requested while it is offered by display of the network indicator on the UE.

- Target: 2% - 1% - 0.5% ?

NOTE: This KPI and its technical basis are currently not standardized and therefore cannot be assessed in a comparative manner.

4.3.2 USSD completion failure ratio [%]

Definition under study.

- Target: 1% - 0.5% - 0.1% ? Or even 0% ??

NOTE: This KPI and its technical basis are currently not standardized and therefore cannot be assessed in a comparative manner.

4.3.3 USSD end-to-end delivery time [s]

Definition under study.

- Target values:
  - 60 sec for 90%, 120 sec for 100%
  - 30 sec for 95%, 90 sec for 100%
  - 10 sec for 98%, 30 sec for 100% ?

NOTE: This KPI and its technical basis are currently not standardized and therefore cannot be assessed in a comparative manner.

4.3.4 USSD receive confirmation failure ratio [%]

Definition under study.

- Target: 1% - 0.5% - 0.1% ?

NOTE: This KPI and its technical basis are currently not standardized and therefore cannot be assessed in a comparative manner.

4.3.5 SMS service non-accessibility [%]

The SMS service non-accessibility is the probability that the end-user cannot access the Short Message Service (SMS) when requested while it is offered by display of the network indicator on the UE.

- Target: 2% - 1% - 0.5% ?

4.3.6 SMS completion failure ratio [%]

The SMS completion failure ratio is the ratio of unsuccessfully received and sent messages from one UE to another UE, excluding duplicate received and corrupted messages.

A corrupted SMS is an SMS with at least one bit error in its message part.

- Target: 1% - 0.5% - 0.1% ?

4.3.7 SMS end-to-end delivery time [s]

The SMS end-to-end delivery time is the time period between sending a short message to the network and receiving the very same short message at another UE.

- Target values:
  – 60 sec for 90%, 120 sec for 100%
  – 30 sec for 95%, 90 sec for 100%
  – 10 sec for 98%, 30 sec for 100% ?

4.3.8 SMS receive confirmation failure ratio [%]

The SMS receive confirmation failure ratio is the probability that the receive confirmation for a sent attempt is not received by the originating UE although requested.

- Target: 1% - 0.5% - 0.1% ?

4.3.9 HTTPS Service non accessibility [%]

The HTTPS service non-accessibility ratio is the probability that a subscriber cannot establish a PDP context and access the service successfully.

The packet data protocol (PDP) context is a data structure present in several parts of the mobile network which contains the subscriber's session information when the subscriber has an active session.

- Target: 2% - 1% - 0.5% ?

NOTE: This KPI and its technical basis are currently not standardized and therefore cannot be assessed in a comparative manner.

4.3.10 HTTPS set-up time [s]

The HTTPS set-up time is the time period needed to access the service successfully, from starting the connection to the point of time when the content is sent or received.

- Target values:
  – 30 sec for 90%, 60 sec for 100%
  – 15 sec for 95%, 30 sec for 100%
  – 8 sec for 98%, 20 sec for 100% ?

NOTE: This KPI and its technical basis are currently not standardized and therefore cannot be assessed in a comparative manner.

4.3.11 HTTPS session failure ratio [%]

The HTTPS IP-service access ratio is the probability that a subscriber would not be able to establish a TCP/IP connection to the server of a service successfully.

- Target: 2% - 1% - 0.5% ?

NOTE: This KPI and its technical basis are currently not standardized and therefore cannot be assessed in a comparative manner.

4.3.12 HTTPS session time [s]

The HTTP session time is the time period needed to successfully complete a packet switching data session. It is also called page loading time.

- Target values:
  – 30 sec for 90%, 60 sec for 100%

- 15 sec for 95%, 30 sec for 100%
- 8 sec for 98%, 20 sec for 100% ?

NOTE: This KPI and its technical basis are currently not standardized and therefore cannot be assessed in a comparative manner.

4.3.13 HTTPS data transfer cut-off ratio [%]

The HTTP data transfer cut-off ratio is the proportion of incomplete data transfers and data transfers that were started successfully.

- Target: 2% - 1% - 0.5% ?

NOTE: This KPI and its technical basis are currently not standardized and therefore cannot be assessed in a comparative manner.

4.3.14 Integrity of complaint resolution [%]

Ratio of the number of complete and professional resolutions of the contributory causes of a complaint, to the total number of user complaints accepted.

- Target: 2% - 1% - 0.5% ?

4.3.15 Complaint resolution time

Definition under study. Working days or calendar days?

- Target values:
  - 2 days for 90%, 6 days for 100%
  - 1 day for 95%, 3 days for 100%
  - 4 hours for 98%, 2 days for 100% ?

4.3.16 Mean Time to Restore (MTTR)

Definition under study.

- Target: Minutes ?, Hours ?, Days ???

4.4 Guidance related to specific Digital Financial Services implementations

As mentioned in the previous section guidance on specific DFS implementations requires the detailed technical knowledge on the components and technical factors and flow of action of each and every DFS implementation. Regulators are the key stakeholders who are in a position to mandate that such information finds its way into the standardization sphere of the ITU-T.

4.5 KPIs for non-utilization stages

This needs further discussions. There is a huge number of possible KPIs standardized, but the selection and the assessment methodology need to be defined.

4.6 Mystery shopping

Mystery shopping was standard practice by the early 1940s as a way to measure employee integrity. Tools used for mystery shopping assessments range from simple questionnaires to complete audio and video recordings. Mystery shopping can be used in any industry, with the most common venues being retail stores, hotels, movie theatres, restaurants, fast food chains, banks, gas stations, car dealerships, apartments, health clubs and health care facilities.

Since 2010, mystery shopping has become abundant in the medical tourism industry, with healthcare providers and medical facilities using the tool to assess and improve the customer service experience. In the UK mystery shopping is increasingly used to provide feedback on customer services provided by local authorities, and other non-profit organizations such as housing associations and churches.

Mystery shopping is increasingly used to evaluate user experience related to DFS. However, due to the complexity of DFS offerings, the results may be interesting, but also may lack statistical significance:

This needs further discussions and in this context a critical look into crowdsourcing and quality evaluation via social media has to be done.

4.7 Legal entities

If the provider of the DFS implementation and the operator of the physical network are the same legal entity it should be unproblematic for the telecom regulator to impose certain QoS requirements with respect to the DFS “service” offered by such entity.

However, if the provider of the DFS “service” and the physical network operator are distinct and separate legal entities, it might turn out problematic to impose any QoS requirements with regard to the DFS onto the network operator.

Therefore, it is suggested to predominantly only accept DFS “services” offered by the physical network operator.

> NOTE: However, the question may be raised whether this is realistic. Up to now, network operators were not even able to successfully establish higher-value services in the entertainment sector (such mobile music or video). On the other hand de-coupling a DFS "service" from the physical network makes it strictly speaking to an OTT "service" provided under best effort conditions, which by their nature withstand technical regulation.

5 Future Considerations: Top-level view

This section deals with an end-to-end model of DFS. It focuses on the essence for user-related functionality of DFS by providing a top-level view of (selected) DFS use cases.

The term “transaction” is used to describe a single instance of a complete use case from a customer point of view, in accordance to the usage of this term in other fields of QoS standardization[1]. It is noted that in this case the term is also part of the common expression “financial transaction”.

The use cases described serve as examples to explain the underlying framework. The underlying model can, however, be easily applied to other use cases which are identified to be relevant in the DFS context.

From the use cases, quality metrics are derived. The key point of the model is that it is, on its topmost level, “technology agnostic”. The actual implementation may be in manifold ways, with specific technical characteristics, strengths and weaknesses; these come in in lower levels of the model. The technology agnostic top level makes sure that no “technology-related” allowances are made (such as “discounts” for known technical weaknesses of particular implementations). Also, the model makes sure that new technical developments in realizing DFS do not disrupt existing QoS metrics.

The underlying general principle of the QoS metrics proposed is also to provide the smallest possible number of KPI, with each KPI having a clearly defined relation to user perception. This shall avoid the situation – which can be observed in some KPI sets – that single KPI overlap from their meaning, which can lead to unclear or even contradictory results.

An actual DFS implementation will be using different network- related “services” or functionality. The respective section shows how the use case related top-level view – and its KPI – can be mapped

[1] FOR INSTANCE, A TRANSACTION FOR THE SERVICE „TELEPHONY“ WOULD BE A CALL FROM AN A PARTY TO A B PARTY, FROM CALL SETUP TO A CALL USAGE PHASE TO THE CALL HANG-UP BY THE A PARTY.

to this technological level of currently existing “carrier services” with respective (mostly already existing) KPI.

The principle of having a small number of strong KPI does not exclude additional KPI with diagnostic or administrative function.

It is recognized that there are several stakeholders with different interests. The respective section – which is also to be seen as an expandable illustration of the underlying concept – describes this view in more detail.

The fact that different stakeholders have different interests also leads to the conclusion that not all of the KPI are of equal importance for all stakeholders. This aspect can provide guidance when it comes to the provision of a legal or regulatory framework to enable or support emergence of DFS.

The final section of this chapter deals with considerations about a practical monitoring of DFS service performance can be implemented. It differentiates between test and measurement in the introduction phase, and continuous quality monitoring in the operational phase of DFS.

5.1 Use cases and related top-level KPI

5.1.1 Transfer of money from A to B

Basic flow of activities

Party A decides to transfer amount X from his account to the account of B.

Key interests of this transfers are:

1 The transfer shall be made with a clear indication of success or failure on both sides within a reasonable time span

2 The success rate of a money transfer shall be high

3 The duration of a transaction shall be reasonably short

4 If the transaction fails, the situations needs to be completely reverted within a reasonably short time span (i.e. no money “lost in limbo”)

5 The transaction shall lead to a stable and correct end state for all participants in a reasonably short time span (i.e. all accounts have to be “up to date” as fast as possible)

6 There must be no losses or duplications of money during the transaction (i.e. money not deducted from A’s account but appearing on B’s account).

NOTE: Not all of these conditions are of equal importance to all stakeholders, e.g. the absence of “money duplications” may not be of interest to end users.

A further differentiation of the use case may come from the question if some kind of proof for the transaction is created, and if yes, in which way. This may be a crucial element if money is paid to serve some duties as e.g. the electricity bill. This may involve another data transaction towards, possibly, a third party to send such a proof, or access to respective services to produce this.

From these requirements, the following end to end KPI can be derived:

– Money Transfer completion rate
– Money Transfer completion time
– Money Transfer False Positive Rate
– Money Transfer False Negative Rate
– Money Transfer Failed Transaction Resolution Rate
– Money Transfer Account Stabilization Success Rate
– Money Transfer Account Stabilization Time

– Money Transfer Loss Rate
– Money Transfer Duplication Rate

NOTE: These KPIs and their technical basis are currently not standardized and therefore cannot be assessed in a comparative manner.

This list clearly contains elements which are not primarily related to mobile network behaviour or performance; they also relate to the performance of underlying banking processes and implementations. So, the list can probably be reduced to elements which are assumed to be primarily linked to mobile networks.

There is, however, a connection. If, for example, a connection loss occurs during a transaction consisting of a number of roundtrips estimated to complete a DFS transaction, this may have different results depending on a particular implementation of such banking processes. Therefore, it is assumed that the robustness and stability of such processes against failures which are typical to specific basic services of mobile networks will also have an effect on overall QoS of DFS.

5.2 Technological components of DFS

As outlined in other parts of this document, there are some services and functionalities within existing mobile networks which can be used – with a further selection by available features of mobile devices - to realize DFS.

From the concept of a "pyramid of needs" and assessment of the end to end KPI for DFS, a clear hierarchy of quality requirements can be derived.

The topmost requirement will be the integrity of a transaction. Integrity in DFS is the clear and reliable assessment if a transaction has been successful or not. This is seen as even more important as the overall success rate of an implementation. If a transaction is erroneously assessed as being successful or failed, the objective damage (e.g. to a person's financial condition) will be larger than a case where a transaction has to be repeated due to a detected failure. The same applies to a transaction which is erroneously assessed as unsuccessful, which would result in duplicate transfer due to a repetition of the process.

From a QoE point of view, the situation can be more complex. Assumed there are two implementations, one of them being stable and robust in the sense of low (ideally zero) probability of false positives or negatives, but slow; the other one faster but more sensitive to such errors. Unless the false-assessment error will be quite large, it is likely that in the customer perception, the latter will appears as the "better" one. It follows that in this area, considerations beyond a mere competition according to market rules need to be undertaken.

An end to end approach needs to be taken because the overall robustness of a particular implementation depends on several factors.

Assume that there are two alternatives, one of them requiring N1 roundtrips, each having a time duration of T1, and a success rate per roundtrip of S1; the other one characterized likewise by characteristics N2, T2 and S2. Clearly, there are several interactions with typical network properties. For instance, if the transaction is performed while the actor is moving (e.g. in a public transport vehicle or as a passenger in a car), the change of network conditions during a transaction influences the overall success rate. This links the time scale of motion-related impairments to transaction characteristics. If the typical overall duration of a DFS transaction (T1*N1 and T2*N2) is above the typical time during which network properties show degradations, the probability of failure increases. In a more general view, the overall success rate of a DFS transaction can be expressed as S1N1 and S2N2. So even if an individual success rate per roundtrip of a specific implementation (where the motion profile can be factored in) is lower, the resulting E2E success rate may be higher if the number of roundtrips in this implementation is sufficiently smaller.

The same linkage between characteristics includes the times involved. For instance, if a transaction fails (in a "proper" way, i.e. with correct assessment of the result), the negative impact on QoE will assumedly be smaller if this result is obtained in a shorter period of time, as a follow-up try can be started and completed faster.

5.3 Stakeholders

The following is not meant to be a complete analysis of stakeholder structure and their requirements. The point to be made here is that different stakeholder types exist, and that their concerns and main interests differ. This will have an impact on the relative weighting of particular QoS metrics and therefore on definition of QoE.

End customers

The main interest of end customers will be to have access to DFS at low cost (which also means without the need to spend additional on new mobile devices) and with a high degree of reliability, as financial losses due to service failures will be felt relatively strong, in particular in low-income segments. It is assumed that transaction speed considerations are (as long as transaction times are within certain reasonable limits) of less importance.

Businesses

With assumedly the same basic need for reliable and affordable transaction, at least larger enterprises will have an interest in DFS technologies which allow for efficient processing of recurring or larger scale transactions. It is further assumed that there may be interest in technologies which can be performed from fixed-network equipment (i.e. computers) without excessive cost. This will in turn affect market acceptance of solutions with different ways of interfacing. An example would be access to certain gateways or other network based functions as SMSC.

Network operators

As network operators are, usually, subject to regulation, relevant factors actually can be separated into two categories. The first category are general technical and commercial requirements, such as cost of operation of a particular technology in relation to profits which can be generated. The second category may include cost of noncompliance to legal or regulatory requirements (SLA), or linkages between e.g. licenses and obligations to provide certain services or service properties.

DFS operators

As far as DFS operators are not identical with network operators, they will basically underlie similar conditions as network operators with perhaps other governmental entities responsible to set and enforce the rules under which they operate. Commercially, their market power will probably be large enough as to impose quality standards (SLA) or other market forces to service providers (network operators).

Governments/Regulators

Assumed that the main objective of governments is economic development, their task is to find a balance between "carrot and stick", i.e. a level of rules and regulations which enable technical evolution, leave DFS operators enough room to run a profitable service, and make sure that cost of DFS services are in an affordable range. For this stakeholder group, assumption is that the main objectives are stable, reliable services in combination with a technology which gives the target segment of the population a sufficiently barrier-free access to DFS.

Furthermore, there are different ways how each of these stakeholder groups have influence on other stakeholders, for instance in rewarding or sanctioning market offerings or, more general, decisions. The crucial point to be made here is that beyond the directly visible first-order effects, second-order interactions exist which do not necessarily have to be weaker, but may work in a "cybernetic" way, i.e. with longer time constants but with likewise or even stronger effects than first-order dependencies.

5.4 QoS Monitoring

In order to secure the necessary quality level of DFS, respective regulatory guidance and comprehensive performance targets need to be established. Basically it would be possible to refer to basic performance measurements of respective carrier services (such as SMS, telephony (for DMTF or IVR) or packet data. Due to the nature of services implementation this will, however, be a surrogate with considerable risk of predicting actual DFS performance incorrectly.

It is therefore - owing to the importance of DFS - assumed that a better way of monitoring needs to be established. This monitoring should – while being fully aware of practical issues in definition and implementation - use actual use cases, i.e. actual money transfer.

The monitoring is proposed to have multiple forms which cover all of the stages of the technical life cycle of any DFS implementation.

Assessment and Roll-out phase:

End to end performance measurements as professional done by dedicated systems, e.g. under control of regulatory authorities.

Operational phase:

Regular End to end performance measurements as professional done by dedicated systems, e.g. under control of regulatory authorities.

“Test Panel” performance measurements, integrated in selected end user’s devices/apps:

For this kind of measurement, a group of end users, selected as to be representative for the general usership, would be recruited and equipped with specially designed DFS clients. This group would, along with doing their “real life” DFS usage, also file additional reports. These reports would then allow responsible entities to constantly assess the performance of DFS in the field.

“Crowdsourced” performance measurements, integrated in end user’s devices/apps:

This would be a simple and non-intrusive way to obtain information on DFS performance on a broad scale. Professional systems used would be equipped with functionality to not only measure E2E performance, but also collect diagnostic information allowing to track root causes for poor performance or malfunction of services.

Of course using real use cases creates additional cost. This cost needs to be assessed against the benefits of obtaining real data instead of surrogate data which only can estimates actual service performance. Moreover, it is possible, with little additional effort in planning and implementation, to design processes which optimize such additional cost, such as re-transferring money which has been moved by a DFS usage.

It is therefore proposed to add respective concepts to a DFS implementation strategy. To increase the effectiveness of such concepts, it is recommended to design a pilot phase which shall give insight into practical aspects and provide information to optimize respective operations.

Annex A

Overview of existing standards which are related to DFS

Recommendation ITU-T Y.1541 "Network performance objectives for IP-based services" specifies network IP performance values between User Network Interfaces (UNI) for each of the performance parameters defined in [ITU-T Y.1540]. The specific performance values vary, depending on the network QoS class. This Recommendation defines eight network QoS classes. It applies to international IP network paths (UNI-UNI). The network QoS classes defined here are intended to be the basis of agreements between end-users and network service providers, and between service providers. The classes should continue to be used when static agreements give way to dynamic requests supported by QoS specification protocols.

However, Rec. Y.1541 does only apply to fixed networks.

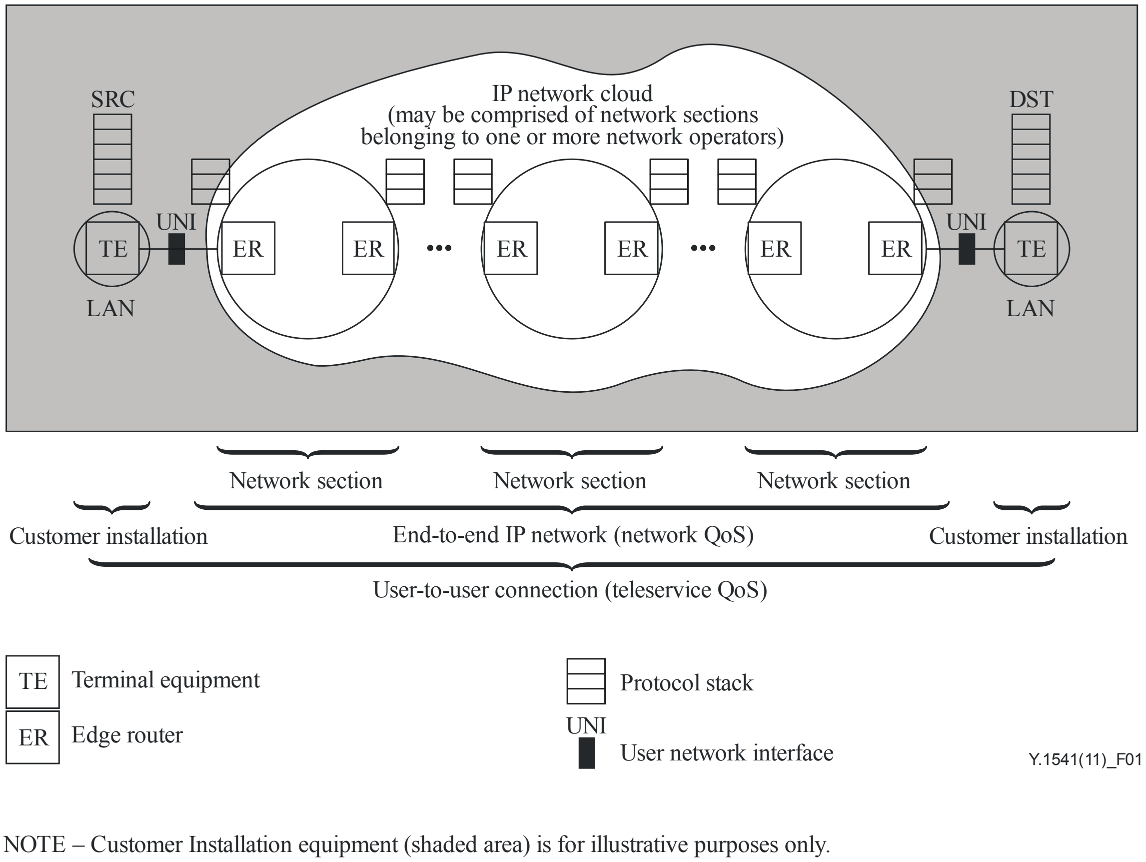

Figure 8 - Reference path from Recommendation ITU-T Y.1541

Its counterpart in the mobile environment is 3GPP TS 23107 "Technical Specification Group Services and System Aspects; Quality of Service (QoS) concept and architecture".

When defining the UMTS QoS classes, also referred to as traffic classes, the restrictions and limitations of the air interface have to be taken into account. It is not reasonable to define complex mechanisms as have been in fixed networks due to different error characteristics of the air interface. The QoS mechanisms provided in the cellular network have to be robust and capable of providing reasonable QoS resolution.

There are four different QoS classes in UMTS:

– conversational class;
– streaming class;
– interactive class; and

– background class.

The main distinguishing factor between these QoS classes is how delay sensitive the traffic is.

Conversational class

The most well-known use of this scheme is telephony (e.g. over GSM). But with Internet and multimedia a number of new applications will require this scheme, for example voice over IP and video conferencing tools. Real time conversation is always performed between peers (or groups) of live (human) end-users. This is the only scheme where the required characteristics are strictly given by human perception.

Real time conversation scheme is characterised by that the transfer time shall be low because of the conversational nature of the scheme and at the same time that the time relation (variation) between information entities of the stream shall be preserved in the same way as for real time streams. The maximum transfer delay is given by the human perception of video and audio conversation. Therefore the limit for acceptable transfer delay is very strict, as failure to provide low enough transfer delay will result in unacceptable lack of quality. The transfer delay requirement is therefore both significantly lower and more stringent than the round trip delay of the interactive traffic case.

Real time conversation - fundamental characteristics for QoS:

– preserve time relation (variation) between information entities of the stream;
– Conversational pattern (stringent and low delay).

Streaming class

When the user is looking at (listening to) real time video (audio) the scheme of real time streams applies. The real time data flow is always aiming at a live (human) destination. It is a one way transport. This scheme is characterised by the fact that the time relations (variation) between information entities (i.e. samples, packets) within a flow shall be preserved, although it does not have any requirements on low transfer delay.

The delay variation of the end-to-end flow shall be limited, to preserve the time relation (variation) between information entities of the stream. But as the stream normally is time aligned at the receiving end (in the user equipment), the highest acceptable delay variation over the transmission media is given by the capability of the time alignment function of the application. Acceptable delay variation is thus much greater than the delay variation given by the limits of human perception.

Real time streams - fundamental characteristics for QoS:

– preserve time relation (variation) between information entities of the stream.

NOTE: This shall also be true for data communication even if not in the real time class. In packet data, higher protocol levels (TCP and upwards) guarantees this time relation. Preservation of order is not directly linked to low latency.

Interactive class

When the end-user, that is either a machine or a human, is on line requesting data from remote equipment (e.g. a server), this scheme applies. Examples of human interaction with the remote equipment are: web browsing, data base retrieval, server access. Examples of machines Interaction with remote equipment are: polling for measurement records and automatic data base enquiries (tele-machines).

Interactive traffic is the other classical data communication scheme that on an overall level is characterised by the request response pattern of the end-user. At the message destination there is an entity expecting the message (response) within a certain time. Round trip delay time is therefore one of the key attributes. Another characteristic is that the content of the packets shall be transparently transferred.

Interactive traffic - fundamental characteristics for QoS:

– request response pattern;

– preserve payload content.

Background class

When the end-user, that typically is a computer, sends and receives data-files in the background, this scheme applies. Examples are background delivery of E-mails, SMS, and download of databases and reception of measurement records.

Background traffic is one of the classical data communication schemes that on an overall level is characterised by that the destination is not expecting the data within a certain time. The scheme is thus more or less delivery time insensitive. Another characteristic is that the content of the packets shall be transparently transferred (with low bit error rate).

Background traffic - fundamental characteristics for QoS:

– the destination is not expecting the data within a certain time;

– preserve payload content.

In order to have a specific service transported in the appropriate QoS class, it has to be recognized by the protocol instances to which class it belongs. This is of special importance in cases where new services demand for close to real-time transmission and make use of existing services.

The best example in this context is a financial service which makes use of the SMS service. Without any additional measures taken, the network does not recognize the financial service but only the SMS service and will transmit it in the background class. In consequence the financial service is not being provided with the necessary real-time transmission.

Recommendation ITU-T E.804 "Quality of service aspects for popular services in mobile networks" provides sets of quality of service (QoS) parameters from an end-user's perspective for the operational aspects of mobile communication. As services per se are not standardized, it focuses on popular services, which means commonly or widely used services.

This does not preclude applying the definitions in this Recommendation for other (not widely used) services, if feasible.

It provides QoS parameter (KPI) definitions for mobile services and related trigger points. Furthermore, it discusses all aspects of practical application thereof, including field testing and statistical considerations.

Currently, DFS as a specifically defined end to end service is not included in Recommendation ITU-T E.804. Services on which actual DFS implementations are based – such as SMS or http – are however treated in in broad detail.

Note: SMS is a store and forward service which – without modifications – cannot be used for real-time transactions which will be required for certain transaction types of DFS

Recommendation ITU-T G.1040 "Network contribution to transaction time" provides the definition, description, and examples of the network contribution to transaction time (NCTT) performance metric for short data transactions with relevance to network providers and users. This is a metric derived primarily from the performance characteristics of the user-network interface to user-network interface (UNI-UNI) path, although it also uses limited configuration information from clients and hosts.

This performance metric is intended to be applied in situations where packet network communications are used to complete repetitive data transactions, such as credit card authorization for purchase, and where measurements of the supporting network's performance are available.

The NCTT metric is derived from packet transfer delays and packet loss ratios from client to host and host to client, effectively a round-trip across the network. Measurements will usually supply the needed network characterization.

A typical data transaction takes the form of a packet conversation, where the client identifies itself to a remote host and submits some request for processing on behalf of a user. The host, after assuring the identities and authorization of the client device and user, performs the request and communicates the result. In the case of "short" transactions considered here, the result is a simple confirmation of the request to exchange funds, or an account balance.

The reference path and reference transaction (illustrating a transaction with eight round-trip exchanges) are described in the figures below.

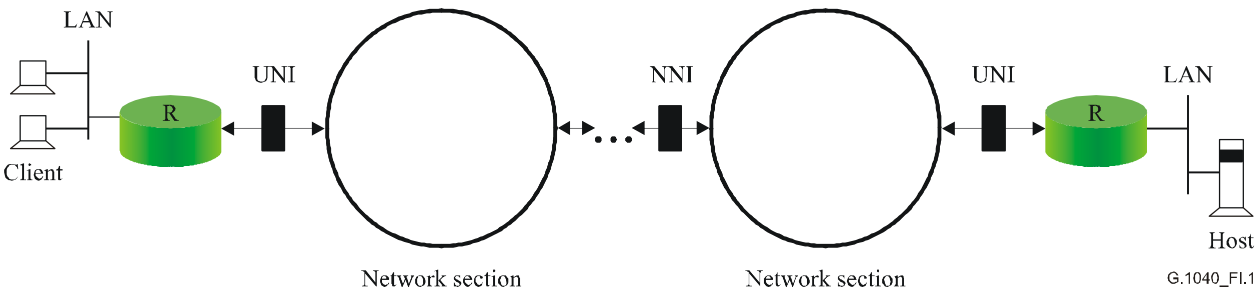

Figure 9 - Reference path from Recommendation G.1040

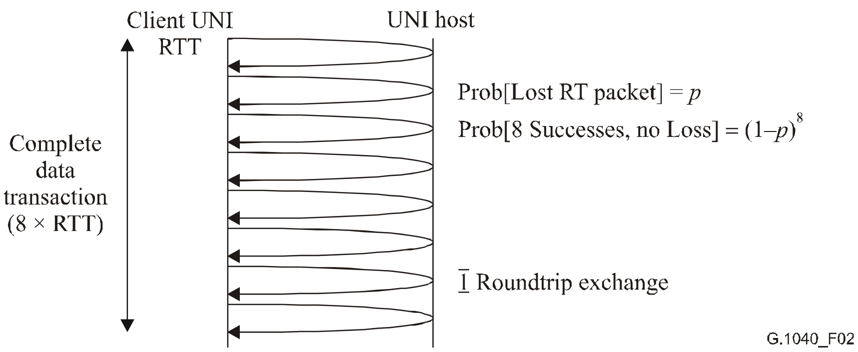

Figure 10 - Reference transaction from Recommendation G.1040

Supplement 9 to the ITU-T E.800-series Recommendations "Guidelines on regulatory aspects of QoS" focuses on end-to-end QoS as perceived by the user when using modern mobile and broadband services. The intent here is to assist regulators or administrations who need to achieve desired levels of QoS for one or more information and communications technology (ICT) services under their jurisdiction.

Recommendation ITU-T E.803 "Quality of service parameters for supporting service aspects" lists 88 generic parameters over the product life cycle of ICT services which will enable a regulator, stakeholder or any interested party to select a pertinent number of parameters about the Service Provider (SP) that provide performance data. Performance data on the non-utilization stages of services, in addition to the service specific performance usually dealing with in-use performance, are necessary to enable customers to choose a service provider (SP) most suited to meet their specific quality of service (QoS) requirements.

QoS performance on non-utilization stages can benefit customers, regulators, stakeholders and service providers (SPs) to monitor performance levels for the benefit of the customers and ICT industry. The essential information to be obtained for measurement and reporting of performance levels is illustrated on a selection of parameters. Guidance on presentation of performance results is also provided. Service providers reporting of delivered performance to a recommended procedure will enable comparability among providers.

Annex B

Underlying functionalities of DFS applications

Table B.1: Summary of technologies for use case #1

Technique	Main features	Disadvantages	Advantages
SMS	Store-and-forward alphanumerical messages	Not real-time	Globally available Interconnection ok
IVR	Interaction with user by artificial or recorded voice, voice recognition and/or DTMF	Requires good speech quality transmission	Real-time
DTMF	Simple keypad operation	Limited character set	Real-time
USSD	Alphanumerical messages	Requires USSD Gateways	Real-time

Table B.2: Summary of technologies for use case #2

Technique	Main features	Disadvantages	Advantages
SMS	Store-and-forward alphanumerical messages	Not real-time	Globally available Interconnection ok
IVR	Interaction with user by artificial or recorded voice, voice recognition and/or DTMF	Requires good speech quality transmission	Real-time
DTMF	Simple keypad operation	Limited character set	Real-time
USSD	Alphanumerical messages	Requires USSD Gateways	Real-time
WAP	Simple web browser	Limited set of functions	Available on some phones even if they do not support http
HTTP	Standard web browser	Unsecure	Internet-like access
HTTPS	Safe web browser	Complex	Encrypted, not even subject to traffic shaping

B.1 Use Case #1

From a pragmatic point of view it is assumed that the Focus Group DFS (FG-DFS) focusses on DFS applications that can be run using simple mobile feature phones (low-end mobile phones which are limited in capabilities in contrast to modern smartphones). Therefore we assume in the following that financial services requiring ftp, http or browser based transactions can be safely excluded from the discussion in this section.

B.1.1 Short Message Service (SMS)

SMS is used to send text messages to and from mobile phones, fax machines and /or IP addresses. The messages can typically be up to 160 characters in length, though some services use 5-bit mode, which supports 224 characters. SMS was originally created for phones that use GSM (Global System for Mobile) communication, but now all the major cell phone systems support it. Once a message is sent, it is received by a Short Message Service Center (SMSC), which must then get it to the appropriate mobile device.

To do this, the SMSC sends a SMS Request to the home location register (HLR) to find the roaming customer. Once the HLR receives the request, it will respond to the SMSC with the subscriber's status: 1) inactive or active 2) where subscriber is roaming.

If the response is "inactive", then the SMSC will hold onto the message for a period of time. When the subscriber accesses his device, the HLR sends a SMS Notification to the SMSC, and the SMSC will attempt delivery.

The SMSC transfers the message in a Short Message Delivery Point to Point format to the serving system. The system pages the device, and if it responds, the message gets delivered.

The SMSC receives verification that the message was received by the end user, then categorizes the message as "sent" and will not attempt to send again.

SMS falls into the group of the so-called store-and-forward services and is normally being transported in the background class according to 3GPP TS 23107. As a consequence, parameters like SMS delivery time or SMS response time depend very much on the traffic load of the mobile network and cannot be guaranteed.

B.1.2 Interactive Voice Response (IVR)

Interactive voice response (IVR) is a technology that allows a computer to interact with human users through the use of voice and DTMF tones input via keypad.

In telecommunications, IVR allows customers to interact with a company’s host system via a telephone keypad or by speech recognition, after which they can service their own inquiries by following the IVR dialogue. IVR systems can respond with pre-recorded or dynamically generated audio to further direct users on how to proceed. IVR applications can be used to control almost any function where the interface can be broken down into a series of simple interactions.

B.1.3 Dual Tone Multi Frequency (DTMF) signalling

The DTMF system uses a set of eight audio frequencies transmitted in pairs to represent 16 signals, represented by the ten digits, the letters A to D, and the symbols # and * as described in Recommendation ITU-T Q.23. Detailed requirements for DTMF are specified in ETSI ES 201 235. As the signals are audible tones in the voice frequency range, they can be transmitted like speech signals. Originally used to dial the number of the remote terminal, it became a common method to transmit small amounts of data.

In packet based networks there are 3 common ways of sending DTMF:

- SIP INFO packets as described in IETF RFC 2976
- As specially marked events in the RTP stream – as described in IETF RFC 2833
- Inband as normal audio tones in the RTP stream with no special coding or markers

For mobile networks 3GPP TS23014 describes how DTMF signals are supported. A message based signalling system is used across the 3GPP system air interface. Inband transmission is not possible.

That means that in mobile communication the originating mobile terminal is directly creating the relevant messages when the keys are pressed by the user during a call.

B.1.4 Unstructured Supplementary Service Data (USSD) – both push and pull services

Unstructured Supplementary Service Data (USSD) is a protocol used by mobile terminals to communicate with the network of the mobile operator.

USSD messages are up to 182 alphanumeric characters in length. USSD messages create a real-time connection during a USSD session. The connection remains open, allowing a two-way exchange of a sequence of data. This makes USSD more responsive than services that use SMS.

> Messages sent over USSD are not standardized:
>
> Normally, USSD is used in the format *nnn# as part of configuring the phone on the network. In order to transfer text messages via USSD to another mobile network, a special USSD gateway is required which mobile operators not normally provide.

USSD is sometimes used in conjunction with SMS. The user sends a request to the network via USSD, and the network replies within the same USSD session with an acknowledgement of receipt.

Subsequently, one or more mobile terminated SMS messages communicate the status and/or results of the initial request. In such cases, SMS is used to "push" a reply or updates to the handset when the network is ready to send them. In contrast, USSD is used for command-and-control only.

All mobile phones of phase II or later have USSD capability.

USSD is generally associated with real-time or instant messaging services. There is no store-and-forward capability, as is typical of other short-message protocols like SMS.

USSD is specified in GSM 02.90 and in GSM 03.90.

USSD Modes:

- Mobile-initiated: USSD/ PULL or USSD/ P2P
  when the user dials a code from mobile terminal
- Network-initiated: USSD/ PUSH or USSD/A2P
  when the user receives a push message from the network

USSD can be used e.g. for prepaid callback service, mobile-money services, location-based content services, menu-based information services, and as part of configuring the phone on the network.

B.2 Use Case #2

In addition to use case #1, the following underlying techniques can be taken into account. Even basic smart phones will provide services based on these techniques.

B.2.1 WAP

Wireless Application Protocol (WAP) is a technical standard for accessing information over a mobile wireless network. A WAP browser is a web browser for mobile devices such as mobile phones that uses the protocol.

WAPs that use displays and access the Internet run what are called microbrowsers -browsers with small file sizes that can accommodate the low memory constraints of handheld devices and the low-bandwidthconstraints of a wireless-handheld network.

Although WAP supports HTML and XML, the WML language (an XML application) is specifically devised for small screens and one-hand navigation without a keyboard. WML is scalable from two-line text displays up through graphic screens found on items such as smart phones and communicators. WAP also supports WMLScript. It is similar to JavaScript, but makes minimal demands on memory and CPU power because it does not contain many of the unnecessary functions found in other scripting languages.

B.2.2 HTTP

The Hypertext Transfer Protocol (HTTP) is an application protocol for distributed, collaborative, hypermedia information systems. HTTP is the foundation of data communication for the World Wide Web. Hypertext is structured text that uses logical links (hyperlinks) between nodes containing text. HTTP is the protocol to exchange or transfer hypertext.

HTTP functions as a request-response protocol in the client-server computing model. A web browser, for example, may be the client and an application running on a computer hosting a web site may be the server. The client submits an HTTP request message to the server. The server, which provides resources such as HTML files and other content, or performs other functions on behalf of the client, returns a response message to the client. The response contains completion status information about the request and may also contain requested content in its message body.

B.2.3 HTTPS

HTTPS (also called HTTP over TLS, [1] [2] HTTP over SSL, [3] and HTTP Secure [4] [5]) is a protocol for secure communication over a computer network which is widely used on the Internet. HTTPS consists of communication over Hypertext Transfer Protocol (HTTP) within a connection encrypted by Transport Layer Security or its predecessor, Secure Sockets Layer. The main motivation for HTTPS is authentication of the visited website and to protect the privacy and integrity of the exchanged data.

In its popular deployment on the internet, HTTPS provides authentication of the website and associated web server with which one is communicating, which protects against man-in-the-middle attacks. Additionally, it provides bidirectional encryption of communications between a client and server, which protects against eavesdropping and tampering with and/or forging the contents of the communication.

Annex C

Selection of a set of KPIs appropriate for DFS

Traditionally, the ITU does neither specify a specific set of KPI nor does the ITU specify target values. The technical and economic conditions are too different in different regions or even in different countries of the same region to make it all the same.

Therefore, it is an important task for the stakeholders involved in DFS (regulators, service providers and user organizations) to enter into a process of the selecting an appropriate set of KPIs that fits the local situation.

This selection could reflect the local market characteristics, customer's preferences and requirements. The number of parameters may be chosen to be manageable both for reporting and for practical application. Where local market characteristics require different sets of parameters for different customer sectors this may be reflected in the choice of parameters.

Guidance on the process can be found in Supplement 9 to the ITU-T E.800-series Recommendations.

This section discusses several areas from which KPIs might be selected.

C.1 KPIs for non-utilization stages

For digital financial services it is crucial to set KPI for interactions between the user and the service provider outside the actual usage of the service. Due to the nature of DFS it would obviously be beneficial if in addition to the telecom regulator (and the DFS SP) the related regulator for the banking sector is involved in the selection of these KPIs. The following stages need to be taken into account:

- Preliminary information on ICT services
- Contractual matters between ICT service providers and customers
- Provision of services
- Service alteration
- Technical upgrade of ICT services
- Documentation of services (operational instructions)
- Technical support provided by service provider
- Commercial support provided by service provider
- Complaint management
- Repair services
- Charging and billing
- Network/Service management by customer
- Cessation of service

Further details and guidance can be found in Recommendation ITU-T E.803.

C.2 Technical KPIs

This section points to technical KPIs that can be used against the four techniques discussed in this paper: SMS, IVR, DTMF and USSD. It is important to understand that a specific DFS service offer may make use of a combination of these four techniques. In such a case KPIs for each of the techniques should be taken into account.

Example: A DFS transaction could be initiated using USSD/PULL and afterwards be concluded with SMS status report etc.

NOTE: For the following subclauses, a cross-reference table would be helpful. However, it requires detailed information on the way a DFS service is using these underlying services. Also a description of the role of the respective basic technology in a DFS transaction is missing (or a reference to a new section, as suggested). Example: A DFS transaction is made of a couple of usages of some basic services of functions. For those functions, KPI definitions may exist. With the current structure, there is no information how far a DFS implementation is covered. For example: DTMF. A DFS transaction based on DTMF needs a connection set-up (telephony? Currently not covered; telephony is again treated by existing standards e.g. TS 102 250/E.804).

C.2.1 SMS

The following is a non-exhaustive list of technical KPIs which may be applied for SMS:

- Recommendation ITU-T E.804 (subclause 7.4.4):
  - SMS service non-accessibility [%]
  - SMS access delay [s]
  - SMS completion failure ratio [%]
  - SMS end-to-end delivery time [s]
  - SMS receive confirmation failure ratio [%]
  - SMS receive confirmation time [s]
  - SMS consumed confirmation failure ratio [%]
  - SMS consumed confirmation time [s]
- Recommendation ITU-T G.1040
  - Network contribution to total transaction time (in case the network is not packet based, the principles laid out in G.1040 can be applied in analogy)
- 3GPP TS 23107
  - Traffic class

C.2.2 IVR

The following is a non-exhaustive list of technical KPIs which may be applied for IVR:

- Recommendation ITU-T G.1040
  - Network contribution to total transaction time (in case the network is not packet based, the principles laid out in G.1040 can be applied in analogy)
- Recommendation ITU-T P.863
  - Perceptual objective listening quality assessment
- No Recommendation currently available
  - Intellegibility, voice recognition

C.2.3 DTMF

The following is a non-exhaustive list of technical KPIs which may be applied for DTMF:

- ES 201 235
  - DTMF characteristics
- Recommendation ITU-T G.1040
  - Network contribution to total transaction time (in case the network is not packet based, the principles laid out in G.1040 can be applied in analogy)
- 3GPP TS23014
  - DTMF transport over the radio network

C.2.4 USSD

The following is a non-exhaustive list of technical KPIs which may be applied for USSD:

NOTE: These KPIs and their technical basis are currently not standardized and therefore cannot be assessed in a comparative manner.

- Recommendation ITU-T G.1040
  - Network contribution to total transaction time (in case the network is not packet based, the principles laid out in G.1040 can be applied in analogy)

C.2.5 WAP

The following is a non-exhaustive list of technical KPIs which may be applied for WAP:

- Recommendation ITU-T E.804 (subclause 7.3.11):
  - WAP activation failure ratio
  - WAP activation time
  - WAP (page) IP access failure ratio [%]
  - WAP (page) IP access set-up time [s]
  - WAP (page) session failure ratio [%]
  - WAP (page) session time [s]
  - WAP (page) request failure ratio [%]
  - WAP (page) request time [s]
  - WAP (page) mean data rate [kbit/s]
  - WAP (page) data transfer cut-off ratio [%]
  - WAP (page) data transfer time [s]
- Recommendation ITU-T G.1040
  - Network contribution to total transaction time (in case the network is not packet based, the principles laid out in G.1040 can be applied in analogy)
- 3GPP TS 23107
  - Traffic class

C.2.6 HTTP

The following is a non-exhaustive list of technical KPIs which may be applied for HTTP:

- Recommendation ITU-T E.804 (subclause 7.3.8):
  - HTTP Service non-accessibility [%]
  - HTTP set-up time [s]
  - HTTP IP-service access failure ratio [%]
  - HTTP IP-service set-up time [s]
  - HTTP session failure ratio [%]
  - HTTP session time [s]
  - HTTP mean data rate [kbit/s]
  - HTTP data transfer cut-off ratio [%]
- Recommendation ITU-T G.1040
  - Network contribution to total transaction time (in case the network is not packet based, the principles laid out in G.1040 can be applied in analogy)
- 3GPP TS 23107

  - Traffic class

C.2.7 HTTPS

To be defined.

For the current purposes the same KPI as for HTTP apply (as in user perception there is no difference). However, it should be noted that HTTPS is treated (and routed) in a different way then HTTP by many network operators.

Basically, the test cases for HTTPS can be the same as for HTTP (upload, download, web browsing). Therefore, the set of KPI as defined in Recommendation ITU-T E.804 can be used. However, the technical events on IP level, used there as the primary example are not applicable as these may not be accessible due to encryption. Therefore, equivalent events on higher protocol levels (up to the application level) need to be used, which may require additional validation.

NOTE: These KPIs and their technical basis are currently not standardized and therefore cannot be assessed in a comparative manner.

- Recommendation ITU-T G.1040
  - Network contribution to total transaction time (in case the network is not packet based, the principles laid out in G.1040 can be applied in analogy)
- 3GPP TS 23107
  - Traffic class

____