Understanding Voice Quality in the Context of Emergency Calling

Prepared by

Dr Murray Milner

Milner Consulting Limited

Final Version, 17 July 2009

Executive Summary

This paper investigates voice call quality as experienced over any form of telecommunications network, based on the best available sources of research and analysis on this topic. The E model is used to calculate the transmission rating factor R, which has been calibrated against the best available objective and subjective voice quality measurements. The ETSI work on this topic relates the values of R to the common subjective survey results and the controlled laboratory assessments defined in terms of Mean Opinion Scores. More importantly, the ETSI work also relates the R values to measures of expected customer reaction when they experience poor quality voice transmission, such as early termination of calls, call retries and customer complaints. These latter measures are much simpler for the average person to understand and relate to. They should also be very poignant for service providers delivering voice services. Early termination of calls, multiple call retries and most importantly, customer complaints would normally all be avoided to the greatest extent possible by responsible service providers.

This work on call quality is then related to the situation encountered with voice calls to emergency service facilities. It is found that the values of R defined for fixed PSTN voice calls generally are actually very consistent with what is required for national voice calls to an emergency response centre. The referred design goal for satisfactory voice quality is to have R greater than or equal to 70. However, to avoid such a target potentially excluding emergency calls from mobile cellular phones under generally accepted network standards, the target has been relaxed to R greater than or equal to 65 for 95% of all national calls It is shown that this relaxation is acceptable, but is the lower limit that should be used. Most voice calls to an emergency services centre should achieve this target, to ensure that voice quality is adequate for quick and efficient communications between the parties, under the stressful environment encountered in emergency situations.

It is also shown that the use of networks which are designed for a lower R value, will not deliver an acceptable quality of voice for emergency services communications. This is not to say that such services should not be offered, as they can be very attractive to users for other purposes. However, for a voice service which does not achieve the R value target (as is generally the case with Skype voice service, for example), the service provider should clearly indicate to the customer, that the service should not be used for emergency communication purposes. It is also shown that defining a low target of R=50 is not particularly useful under any circumstances, as most usable voice services would need to compete with the many voice services, such as Skype, which are already available on the Internet for very low prices and deliver voice quality equivalent to an R value of better than 55 under normal conditions.

Finally, an accreditation process for service providers is outlined, which would enable service providers to demonstrate their ability to support voice services which achieve the defined R value target and hence are suitable for use by their customers to make voice calls to emergency service facilities.

Given all of the material contained in this report, it is strongly recommended that the network design requirement for accrediting a provider of voice services to a quality standard suitable for emergency calls at retail is that more than 95% of emergency voice calls achieve an R value of equal to or better than 65. It is also strongly recommended that service providers who offer voice services that are not accredited to this standard should be required to inform their customers that they should not use the voice service for making emergency calls.

1. Introduction

The provision of voice communication for emergency purposes is one of the most stringent requirements for voice quality in telecommunications today. This is because it is essential to ensure that the message being communicated is clearly understood by the emergency response person, with the minimum of repeats, in order to achieve the right response for the caller in the shortest possible time. The communication will also be often undertaken in situations where the calling party is under a high degree of stress, which makes communication difficult under the best of conditions. Hence it is essential that the quality of voice communication available from networks designated as suitable for emergency communication does in fact support a minimum voice quality.

This does not mean that there will be other forms of voice service which are offered in the market which are not designated as suitable for emergency communications, but will be entirely adequate for casual communications, particularly where there is a sensitivity to cost. In addition, it is certain that in future there will be voice services that are offered in the market which are greatly enhanced relative to the minimum required for satisfactory emergency communications, probably attracting a higher price from the user.

This quality based discrimination of voice services is already starting to appear within the market today. Skype is a voice service provider that offers both on-net and off-net voice services.[1] Skype charges nothing for on-net calls and a low price for off-net calls. However, because Skype offers no guarantee for the quality of its off-net calls, it provides a clear warning to the user that the Skype Out service is not intended for the making of emergency calls: “Note: Skype itself is not a replacement for a traditional phone and cannot be used for emergency calling”.

This is a very reasonable warning, as the Skype voice quality can be highly variable, depending on the loading on the Internet associated with the path taken by the call media. Calls can also drop out from time to time due to severe overload conditions. This type of call quality is certainly not acceptable for emergency calling, but can be fully acceptable for many other forms of communication. I have personally used Skype for personal and business purposes, all around the world, due to the low cost of Skype calls, and I fully accept the intermittent quality degradation as a trade-off for the low cost.

2. PSTN Voice Quality

The traditional dominant application on telecommunication networks is voice. The fixed [2] Public Switched Telephone Network (PSTN) was specifically designed to carry the voice application, with a high degree of security and consistent high speech quality. There are very few complaints today about the quality and security of the voice which is carried over any fixed PSTN. Everyone takes it for granted that their call will get through to almost anywhere in the world, almost all of the time, and the resulting call quality will deliver excellent intelligibility. Furthermore, people expect that their conversation will not be overheard by others, so that their privacy is maintained.

The fact that these outcomes have been achieved for the voice application is the result of many decades of sound engineering practice and consistent regulatory policy across the globe. The quality and security is not only maintained within a single service provider’s PSTN, but rather, it can be maintained across multiple service provider PSTNs, even when these service providers are scattered across the globe. This has been achieved through the adoption of, and rigid adherence to, international standards for PSTN interconnection and end-to-end voice quality. Hence in the case of the voice application, the end user does get access to an application which uses multiple servers scattered around the world, owned by multiple service providers and the desired quality and security of the application is maintained end-to-end.

Furthermore, it has been proven over many years of actual use that the PSTN provides a voice quality that is suitable for communications involving emergency services. There are very few complaints about the call quality associated with calls received at emergency contact centres, when using the PSTN.

[1] The terms on-net and off-net are used here to refer, respectively, to a voice service where the calls are carried entirely on the Internet versus a voice service where the calls are carried in part only on the Internet.

[2] Fixed PSTN denotes a telephone access network hosting local line connections, conventionally copper cable, over which subscribers make and receive telephone calls.

The PSTN call quality is based on exhaustive research and analysis over 2-3 decades, in laboratories around the globe in an effort to quantify the requirements for "carrier grade" voice quality. The results of some of this work are summarized below.

3. Definition of Voice Quality

One of the basic human needs for telecommunications is to ensure that the speech leaving the lips of one party can be accurately understood through the ear of the receiving party. The fact that the communication channel between the sending and receiving party is digital and might use a TDM data stream or data packets is of little importance to the end user. From the end user perspective, the end-to-end communication channel is analogue between mouth and ear, and as such, intelligibility and recognition of the speaker's voice remain the most important attributes of a good telephone connection. These are followed closely by comfort factors, which include the level of the voice signal received, the level of the side-tone in the speaker's earpiece, the pickup of excessive room noise, echo on the connection, mouth-to-ear delay, circuit noise, etc.

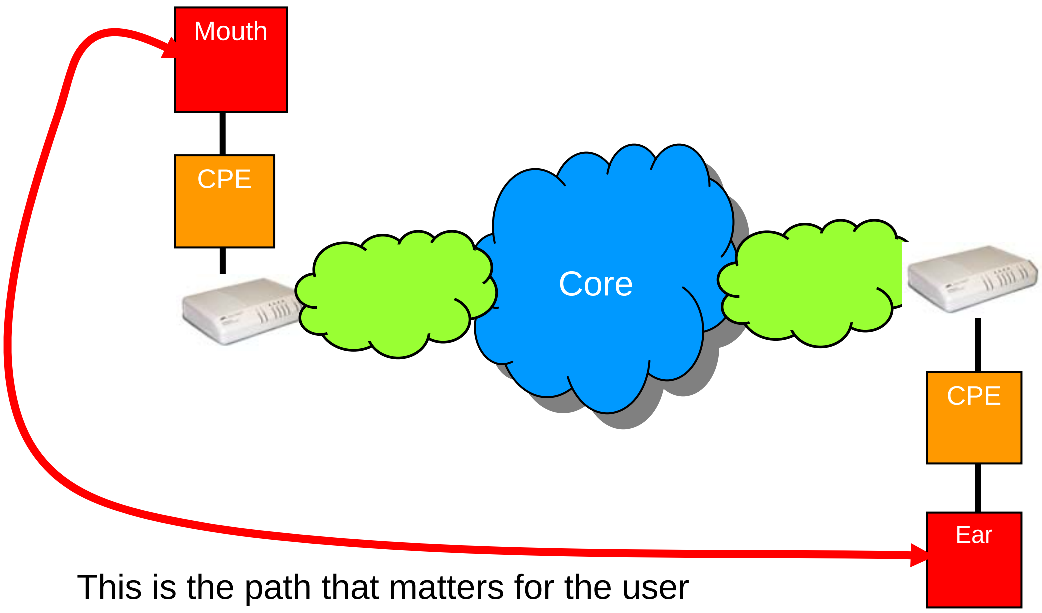

Figure 1: Reference path for voice communications.

The International Telecommunications Union (ITU) sets standards for international telephone connections and circuits, and these are as applicable for a mobile connection or a Voice over IP (VoIP) connection as they are in the PSTN world for which they were originally developed. The ITU Recommendation G.109 entitled "Definition of Categories of Speech Transmission Quality" provides guidance on speech transmission quality in terms of a "transmission rating factor" R, which takes into account all of the transmission parameters involved in any telephone connection. The R value is based on analytical calculation, but is calibrated against the best available subjective quality measures. Table 1 shows the relationship between the R-value and the more traditional, but subjective, Mean Opinion Scores (MOS) and the percentage of the population that would rate a call good or better (GOB) and poor or worse (POW) for various equivalent R-values. The R-values are preferred today, as they can be calculated in a much more objective manner, based on end-to-end connection impairments, using the E-model as defined in ITU Recommendation G.107.

With reference to Table 1, it is normally considered that for minimum acceptable voice quality, the end-to-end voice connection should have an R-value of no less than 50. This is considered to be the level below which nearly all users would be dissatisfied. This objective applies whether the call connection is local, national or international. In order to ensure that most telephony connections meet this objective, the typical design objectives for connections in different parts of the PSTN and mobile networks are as follows:

- Local calls: 95% of connections need to exhibit R >=70
- National calls: 95% of connections need to exhibit R >=65
- Cellular calls: 95% of connections need to exhibit R >=65
- International calls: 95% of connections need to exhibit R >=55

The above objectives, although originally defined for fixed PSTNs, have been extended to include Public Land Mobile networks (PLMNs) and can apply equally to VoIP call connections, through either a private IP network or the Internet, or in fact any other form of end-to-end telephony connection.

R-value	**MOS**	**%GOB**	**%POW**
100	4.50	99.4	0.03
90	4.34	97.0	0.25
80	4.03	89.4	1.4
70	3.60	73.4	5.9
60	3.10	50.0	17.4
50	2.58	26.6	37.7
40	2.06	10.6	62.3
30	1.61	3.04	82.6
20	1.25	0.62	94.1
10	1.04	0.09	98.6

Table 1: Relationship between subjective voice quality measures and R-values
MOS: Mean Opinion Scores

%GOB: % of population that would rate a call good or better
%POW: % pf population that would rate a call poor or worse

4. What Does All This Really Mean?

The above provides the formal definition of voice quality as defined by the ITU. It also provides the expectations for call quality for various types of connection. But what do these numbers mean in reality to the end user of the voice connection?

The ITU relates the R values back to the Mean Opinion Scores, as determined under a statistical sampling of individual perceptions of quality. For example, an R value of 70 relates to a MOS of 3.60 as indicated in table 1, which implies that:

- 73.4% of people communicating over a connection exhibiting this R value would consider the voice quality to be Good or Better, while,
- 5.9% of people communicating over a connection exhibiting this R value would consider the voice quality to be Poor or Worse.

The problem with this interpretation is that it is not clear what the reaction of the calling party will be if this type of call quality is encountered. In order to better understand this issue, it is useful to return to some of the source documents associated with this voice quality analysis. One of the key source documents is that prepared by the European Telecommunication Standards Institute (ETSI) referenced as ETSI Technical Report 250 and entitled “Transmission and Multiplexing (TM); Speech communication quality from mouth to ear for 3,1 kHz handset telephony across networks”, which was prepared in 1996.

This document provides considerably more detail on the work undertaken to define and calibrate the R factor, including its relationship with MOS results and the interpretation of what these figures really mean in terms of expected customer reaction.

ETSI Technical Report 250 identifies that the users' reactions to the voice transmission quality of a connection can be assessed in many ways, some of which have been discussed above and all of which are rather costly to undertake:

a) By user surveys, i.e. actual network users are interviewed and asked for their opinions. Results may be given as percentage of users finding the connection "good or better", GOB, or "poor or worse", POW;
b) By observing users' calling behaviors such as the percentages who:
   (i) Terminate their calls unusually early (TME),
   (ii) Terminate the call and then Re-dial (RETR) or even ,
   (iii) Actually complain to the network operator (CMPL);
c) By subjective tests, involving test teams in controlled laboratory experiments. The results most often are given as "Mean Opinion Scores", MOS. (Sometimes also "percentage having difficulties" are recorded).

The aim of computation models is to give estimates of how such user assessments would turn out, provided the assessments really were made. As identified above, from the ITU standards, the transmission parameters are combined into a "Transmission Rating Factor" R using the ETSI or E model.

The values for R can lie in the range from 0 to 100 or even higher. R = 0 represents an extremely bad and R = 100 a very high transmission quality for the connection. These values for R are then calibrated against the above subjective customer perception and customer reaction measures.

ETSI show that there is a quantified relationship between the various customer perception assessments of voice quality and the analytical transmission rating factor R. Some of the key relationships are shown in Figure 2 below. The ETSI computation model is fundamentally based on knowledge gained from a number of subjective tests made in the past by various organizations, as well as experiences derived from actual network operations and quality surveys.

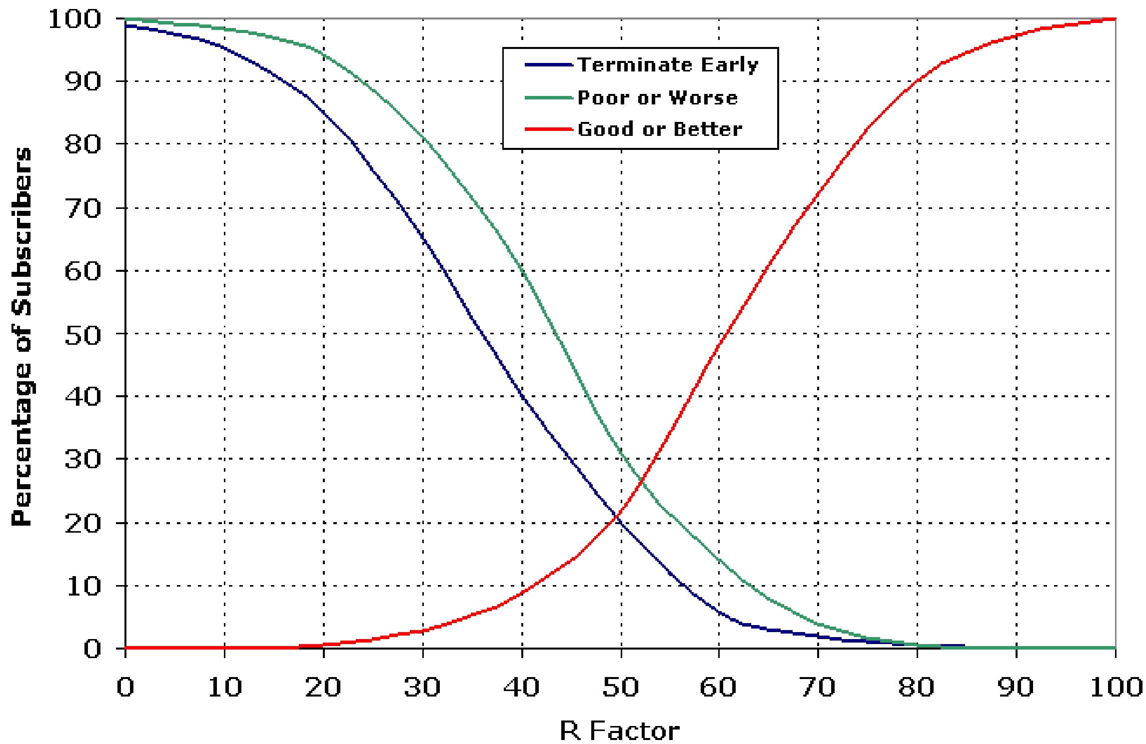

Figure 2: Chart showing the relationship between R factor and customer opinion measures.

One must note, however, that users' opinions vary with time and circumstances, so that there is an inherent uncertainty in the relationships shown in Figure 2. A user's perception of transmission quality depends partly on what kind of performance he (or she) has been used to and partly on what new systems, hopefully improved, he is being exposed to. Thus, a computational model cannot be expected to emulate each and every particular subjective test or opinion survey that is made. Rather, the opinion predictions as computed by means of the ETSI model need to be interpreted as "nominal" for reference purposes. There is in principle a limit to the "precision" of any model. Therefore, the predictions for GOB, POW, TME and MOS obtained by the ETSI model should be considered as nominal customer reactions for guidance in transmission planning. The advantage of using a model is that transmission planning and comparisons of transmission conditions can be done in a systematic and consistent manner. This type of approach is essential for network design.

To evaluate the global transmission quality of a network, it is useful to compute the traffic-weighted mean and standard deviation of the GOB for the connections. In addition, "limiting cases" should be investigated so that "hopeless" transmission situations are avoided. When comparing different transmission cases by means of the ETSI model one must remember the basic uncertainty of customer opinion evaluations. This means for instance that if the difference in GOB between two equipment schemes is only a few percent, this difference cannot be used as a guide to choose the optimum solution.

As shown in Figure 2, the ETSI model produces values of the rating factor R and predictions for the subjective quantities of Mean Opinion Scores (MOS), percentages of customers finding the quality Good Or Better (GOB), Poor Or Worse(POW) and the potential for Terminating calls Early (TME). Of these, non-experts often find GOB, POW or TME easier to interpret than MOS. For evaluating more or less "normal" connections in transmission planning, GOB appears to be an appropriate measure, while POW and TME can be used to characterize "limiting" (i.e. from a quality point of view) connections. In the transmission planning process itself the R-factors can be used directly. If one wishes to have more detailed quality information, when say mobile communication links are involved, R-ratings can be stated both with and without the expectation factor A included (see ITU Recommendation G.113 for an explanation of the A factor – extract contained in Appendix A).

In the context of this discussion, one of the most useful parameters is TME. The percentage of users who terminate their calls early due to bad speech transmission is a measure which most people can recognize as a definitive action by users. If this percentage becomes too high, then obviously the voice quality is unacceptable. Any telecommunications service provider who operates a service for which customers are paying good money that allows a percentage of calls to be terminated early should be highly concerned about the service that they are offering to their customers. If cost is low and the early call termination is not too high, then it is likely that customers will not be too concerned. This is the case for instance with Skype today. Early call termination due to poor voice quality occurs for several calls out of 100 made, based on personal experience. This is considered to be acceptable by most users of Skype (including myself) as the cost of either on-net or off-net calls is low. On the other hand, if even one or two calls in 100 had to be terminated early due to poor voice quality, when using my primary line voice service provider, I would be highly annoyed and probably complain. The results presented by ETSI and other researchers have quantified the user reaction as identified above.

ETSI take the relationship between customer behavior and the transmission rating factor one step further. As well as the early call termination parameter TME, they have derived two further estimates of customer behavior as follows:

- RETR: Percentage of customers who will implement call retrial due to poor voice quality,
- CMPL: Percentage of customers driven to complaint about poor voice quality.

The three customer behavior attributes of TME, RETR and CMPL verses the transmission rating factor R are presented in Table 2, based on the ETSI analysis.

Based on Table 2, for national calls, where the absolute one way transmission delay is less than 100ms, at R=65, less than 1% of calls will result in a customer complaint due to voice quality. However, even at this level, up to 3% of calls could be terminated early due to call quality. It is these customer behavior attributes which determine the minimum call quality for the PSTN as identified in section 3 above. For national voice calls an R value less than 65 is expected to create too much customer dissatisfaction as illustrated by the data contained in Table 2.

R	TME%	RETR%	CMPL%	Comments
90	-	-	-	Very Good
85	-	-	-	Very good
80	-	-	-	Good
75	<1	-	-	Good
70	1.5	<1	-	Adequate
65	3	1	<1	Marginal early call termination
60	6	3	1	Unacceptable early call termination
55	12	5	2.5	Unacceptable call retrial
50	18	9	4	Unacceptable customer complaints
45	27	15	8	Unusable
40	40	25	14	Unusable
35	50	35	22	Unusable

Table 2: Customer behavior attributes verses transmission rating factor.
TME = Terminate early, RETR = Retrial, CMPL = Complaints

For international calls, a lower value of R is acceptable due to customer expectations being lower for international calls. In addition, international calls often have to contend with inherent absolute transmission delays of greater than 100ms, which also results in lower voice quality, but is accepted by customers as an inherent attribute of international calls.

Domestic satellite voice connections are treated in a manner similar to international calls. The absolute one way transmission delay is extended to around 270ms, which makes satellite voice calls typically of lower quality than that for national terrestrial voice calls. On the other hand, if the satellite link is well engineered, the delay variation can be minimal and the error rate can be very low. These two factors mean that a single hop satellite voice connection can be engineered to deliver an R factor of around 65 or better for 95% of all calls, which makes it satisfactory for national telephony use. On the other hand, double hop satellite connections with an absolute delay of over 500ms present a more challenging problem for voice quality. The R value for such connections is typically less than 50 and not surprisingly, most end users consider this type of voice connection to be of unacceptable quality. There is a small group of users though, who have no alternative means of voice communications and are aware of this fact. These users learn to persist with the voice quality of a double hop satellite voice connection. This is fully expected in the computation of R through the application of a high value for A in the analysis.

5. Impact on Emergency Services Voice Calls

Calls which are being made by customers to emergency service centres are normally national calls (ie emergency calls are either terminated in the local area in which they originate or are terminated at a location within the national borders of the originating country to be handed over to emergency service centres). It is very unusual to outsource such call centres to off-shore locations. Hence the national calling situation, with absolute transmission delay of less than 100ms, normally applies to these calls. The domestic satellite situation is discussed separately above.

Based on the data contained in Table 2 above, the preferred transmission rating factor R would be greater than or equal to 70. Under these circumstances only around 1% of customers are likely to terminate calls early due to poor voice quality and in emergency situations, due to the urgency of the communication, this percentage is likely to be reduced to near zero. At this same R value, there will be almost zero call retrials due to poor voice quality and no customer complaints are likely to be generated due to this cause. This value of R is the ideal design value for emergency services voice connections.

Unfortunately, many customers today routinely use mobile voice services (via a PLMN) for national calls, including emergency service calls. As indicated in Section 3 above, national cellular networks struggle to achieve an R rating of 70. This is largely due to the low rate encoding of the speech, combined with the need for at least one voice trans-coding stage from low rate to PSTN based G.711 encoded speech. Thus the normal performance target for national cellular voice calls is R greater than or equal to 65 for 95 % of all calls. Most national cellular to fixed network calls will achieve very close to this value (within the bounds of error associated with this type of analysis), especially if care is taken with the emergency contact centre termination facilities.

At R = 65, Table 2 shows that up to 3% of calls could be terminated early due to perceived poor voice quality. This is marginal for the emergency services call situation. However, it is still not excessive and within the margin of error for this type of subjective analysis. As mentioned previously, emergency services callers will tend to hold on to a call once answered and persist with any quality issues, rather than terminating the call early. The number of call retries is also expected to be only around 1% or less, which is probably acceptable. Similarly, the number of customer complaints concerning voice quality should be minimal. Hence a minimum R value of 65 is probably acceptable as a limiting value for emergency services voice communication and will encompass most of the forms of voice communication technology that end users are likely to use and expect to function satisfactorily in an emergency services calling environment. This includes:

- Most national cellular mobile voice connections, within satisfactory coverage areas,
- Single hop domestic satellite voice connections and
- A variety of voice over IP connections provided the IP environment used for these services is properly managed in terms of delay variation and packet loss.

Can one reduce the minimum value of R further and still expect reasonable emergency services voice call customer behavior? Based on Table 2 the answer appears to be NO! At a value of R = 60 the expectation of early terminated calls is getting to be very significant at 6%. The call retries are also expected to rapidly rise and customer complaints are starting to be become significant. No-one wants any customer complaints about emergency services call voice quality, as any such complaints are likely to also relate to poor emergency services response, with the consequential implications that this may have in terms of personal safety.

These figures, based on robust research over many years from around the world, certainly drive one to the conclusion that the network design value for R should be a minimum value of 65 for more than 95% of all national calls to ensure voice calls to emergency services contact centres are of satisfactory quality. This definition still leaves some leeway for a small percentage of calls to have poorer quality voice due to abnormal conditions, but any voice network intended to support emergency calling should not be designed to deliver a value of R less than 65 under normal conditions. Any voice services offered by service providers which are designed to deliver an R value of less than 65 for more than 95% of customer calls should NOT be designated as being suitable for emergency services use. They should use the same disclaimer as used by Skype today for its Skype Out services which specifically states that Skype itself is not a replacement for a traditional phone and cannot be used for emergency calling.

6. Extreme Values for R

Some service providers have proposed that the limiting value for R should be 50. They claim that in order to be inclusive to all service providers, it is best to set the limiting value for voice quality with this value of R, even for the support of emergency communications. Referring to Table 2, if voice calls were to be made over network connections with R=50, this would mean that around 18% of such voice calls would be terminated early due to poor voice quality and around 10% of such calls would result in retries. Most importantly, around 4% of these voice calls would result in some form of customer complaint to the service provider supplying the service. This potential outcome should not be acceptable to any service provider, and especially not for one who is supposedly supporting emergency services voice calls.

Returning to my example of Skype as a service provider, other than under exceptional circumstances (see below) I personally have certainly never experienced this level of poor quality voice performance when using this service, either on-net or off-net over a variety of broadband Internet connections, all around the world. Hence it would be difficult to understand why any service provider would attempt to offer a worse quality voice service than Skype, as Skype is available to all Internet users at very low prices. How would such a service provider compete in the market? As indicated previously, when used over a typical broadband connection, Skype offers a service which would be comparable with an R factor of 55 to 60, when averaged over say 100 voice calls. Certainly some calls must be terminated early due to voice quality and retries definitely do occur. However the proportions of these events correspond to those to be expected for a network designed for R in the range 55 to 60 rather than 50 or below.

One thing that does not occur with Skype is complaints about poor voice quality, or certainly not in line with the expectations listed in Table 2. The value for money offered by this service (and many others like it) means that customers accept the limitations of the service and typically do not complain. In addition, Skype is a responsible service provider, in that they clearly alert the customer to the fact that the service is NOT suitable for the making of calls to emergency services contact centres. They accept that the voice service call quality can vary considerably and may not always be suitable when one of their customers wants to make an emergency voice call. This should also be the stance taken by any other service provider who offers a similar quality voice service to its customers. Again this would reinforce the need to have a voice quality limit based on an R value of greater than or equal to 65 for any form of voice communication with emergency services.

I have experienced a situation where Skype does exhibit voice quality which I estimate would correspond to an R value of less than 50. This occurred when I used Skype in conjunction with a bandwidth constrained shared user access technology. The access technology was CDMA2000 EVDO packet data over wireless, using a central Wellington cell site. During the period from 0930 to 1030 on most week days over a period of several months, I encountered congestion which consistently made Skype unusable. This same situation will occur for any user who is using a constrained bandwidth access service, with multiple users accessing a single access connection on a shared basis. My solution to this problem was simply to not use the cellular access for Skype calls during this heavy congestion period. I had other forms of voice communication available for me to use during these periods. This of course is quite different to a customer who wants to make a voice call to emergency services during this period of heavy congestion. If his is the only voice service available to the consumer then personal safety is highly likely to be compromised under these conditions.

7. Verifying Emergency Services Voice Quality

The use of R values for assessing voice performance does not provide a simple means to verify that service providers are achieving the required value at all times. Hence it will be necessary for service providers to demonstrate compliance with the requirement on a regular basis. The service provider would need to demonstrate this capability via the following steps:

- The transmission design for end-to-end national voice calls using the service providers network and any other transit networks used for accessing a certified emergency contact centre would need to be evaluated using the E-model,
- The resulting R value for the network design would need to achieve the minimum value for R designated for emergency services voice calls (assuming the end user is using a certified telephony device),
- If the result of this analysis showed that the R value could be achieved, then the service provided would be issued with a provisional compliance certificate,
- If the result of this analysis showed that the R value could not be achieved, then the service would not be issued with a compliance certificate, and that service provider would be requested to notify its customers that the voice service they are using is not suitable for the provision of voice access to an emergency services contact centre,
- The emergency service contact centre(s) will establish the means to measure voice call quality on a sample basis and will make such measures on incoming voice calls independent of the originating call service provider (standard techniques are available for this purpose)
- Based on the calling party's number, if calls from a specific service provider consistently provide call quality measures corresponding to R less than the agreed value, on a statistically valid basis, then the service provider would be notified that their status will be revoked unless the call quality is improved over the next 6 month measurement period,
- On the other hand, if the sample call quality related to any service provider exceeds the minimum R value over any 6 month period, then the provisional status will be upgraded to full certification for a period of 2 years,

- All service providers will need to verify their design voice performance every 2 years in order to retain full certification.

The above describes a very robust accreditation process, but may be considered to be too onerous by some parties. An alternative approach would be to use a self accreditation process, with the potential for periodic audit on an exceptions basis. Service providers would declare themselves to be compliant with the ECS Code following their own internal due diligence. Audit could be triggered by any party on any self accredited service provider, if that party considers that the voice quality being delivered in association with the provision of 111 calls is not compliant with the specification contained in the Code. To ensure that audit is only initiated under genuine conditions, the party requesting the audit would be liable for the costs associated with the undertaking of the audit by an independent, competent third party, if the service provider is proven to be operating within the requirements of the Code. Otherwise, if the service provider is found to be at fault it would need to pay for both the audit and remedy the problems identified with the service (assuming it wished to remain accredited).

The audit process should be undertaken by an independent and competent third party based primarily on a detailed review of the design of the voice service used for making 111 calls to emergency contact centers, using the best available international standards which would typically involve the E-Model. If the design is found to be appropriate or inappropriate relative to the relevant standards and the Code then this should conclude the audit in most cases. The result would be appropriate allocation of costs and any remedial action where required. However, it is recognized that by its very nature the requirements specified in the Code are statistical in nature and so there could be differences of view when using a deterministic audit approach, even when parameter variation is included in the analysis. If agreement cannot be reached based on the deterministic audit approach, then a second step could be undertaken based on actual measurement of voice quality associated with 111 calls. This would be a last resort approach due to the complexity and hence cost of making such measurements.

The above outline approach should ensure compliance with the Code on a self accreditation basis but retain the opportunity to ensure that the self accreditation is not abused by the service provider community. This should ensure the delivery of a compliance regime which can be implemented with low compliance cost.

8. Conclusions

There is a vast array of research and analysis on voice quality available over telecommunications networks through various standards organizations. A particularly good reference on this topic is that provided by ETSI in its Technical Report 250. This report provides the relationship between the calculated transmission rating factor R and the many measured subjective parameters that relate to customer perception of voice quality. It also goes one step further and provides the relationship between R values and expected customer reaction to poor voice quality. These latter comparisons are much more straightforward to understand and provide a very poignant reference for service providers.

Using these comparisons, it is demonstrated that the target values used in the PSTN of the past really are the limiting values that should be used for all future voice communications, especially those involved in some form of emergency situation. All networks designed to support voice calls that are designated as being suitable for use in making calls to emergency services facilities should achieve an R factor of greater than or equal to 65 for more than 95% of all national calls.

This requirement should not stop service providers from offering voice services which achieve an R factor with value less than 65. It simply means that such service providers must clearly signal to their customers that such a service should not be used to make calls to emergency communications facilities. This is nothing new – the precedent is already established with services such as Skype. When customers use the Skype Out service, they are warned not to use the service for emergency communication purposes. Given this warning, many millions of customers around the world still regularly use Skype both on-net and off-net to make voice calls, as the perceived value for money offered by the service is still very attractive. Many other providers of voice service across the Internet must also deal with this issue in a similar manner.

9. Recommendations

It is strongly recommended that voice services defined as being suitable for emergency communications be only provided on networks that are designed to achieve an R factor of greater than or equal to 65 for more than 95% of all national calls.

Networks that offer voice services that do not meet this criterion should be required to communicate this performance limitation clearly to their customers.

Appendix A: E-Model Advantage Factor A as defined in ITU Recommendation G.113

Appendix II of ITU Recommendation G.113 provides the most recent background material regarding the advantage factor, A. The advantage factor does not really deal with codec or signal processing distortion but rather with the relative ponderation of functionality and transmission quality in user expectations of services according to the type of user and the time. The advantage factor A represents an "advantage of access", introduced into transmission planning for the first time via the E-model ([ITU-T G.107] and [ETSI ETR 250]). This factor enables the planner to take into account the fact that customers may accept some decrease in quality for access advantage, e.g., mobility or connection into hard-to-reach regions. This value can be used directly in conjunction with all other impairment values and as an input parameter to the E-model.

Provisional A values are listed in Table 1 of [ITU-T G.107].These values are provisional since they have not been confirmed by subjective investigations to date. Therefore, the advantage factor A should be used with care and with respect to the specific situation of the user. The use of the advantage factor in transmission planning of networks and the selected values are subject to the planner's decision; however, the values in Table 1 of [ITU-T G.107] should be considered as the maximum upper limit for A.

The overall transmission quality as perceived by the user is influenced by the ease or difficulty to establish a connection. In certain cases, wireless systems have an advantage in that they allow spatial flexibility in the provision of service and, as a result, the user may discount the subjective impairments resulting from the speech transmission effects associated with wireless systems. Examples are mobile telephony and multi-hop satellite connections to hard-to-reach regions. However, the expectation factor may be asymmetric. For example, for a call from a mobile subscriber to a PSTN subscriber, the PSTN subscriber may expect PSTN quality while the mobile subscriber may expect mobile quality.

NOTE – In other documents the term "expectation factor" has frequently been used to express the same issue that "advantage factor A" stands for.