Voice Monitoring

Ensuring a quality user experience on your voice applications

ekara
by ip-label

Summary

Introduction: Let's talk about quality

A few years ago, the CEO of a vast consulting empire held up his smartphone and told a journalist, ***"I run a group of 380,000 people with revenues of 30 billion dollars, all with this little device"***. Referring to the popular call application that he finds so essential, he called it his "day-to-day companion".

It just goes to show that voice applications are among the most necessary of the new and not-so-new technologies. Yet, despite the fact that internet telephony has been around for a while, **people may still experience poor speech quality on a video call** or sometimes be unable to elicit a proper response from an interactive voice response (IVR) system. These days, a good, clear connection should not be a remarkable occurrence. Nor should noise, lag, garbling, echo and other unpleasantness occur as frequently as they do.

Businesses must ensure that channels of voice communication are open to their workforce and customers. It's bad business if customers have to complain about a voice server before problems are noticed and fixed. It's a waste of time and productivity to try to conduct meetings with international teams on an echo-riddled conference call or when key people are unable to connect.

Voice applications are widely used in a variety of B2C and B2B situations. This white paper is intended for business audiences which implement:

- **any kind of professional telephony,**
- **on any type of infrastructure,**
- **internally, to other companies, or directly to customers,**
- **on mobile or desktop devices.**

We hope you will benefit from our white paper!

Your business relies on high-quality voice communications

CHAPTER 1

1. Ensuring the quality of voice applications and services

Since telephone and voice services are such intuitively necessary technologies, it's easy to see why they are some of **the most justifiable targets for investment in digital experience management (DEM).** But where does DEM come into play in the context of voice services and how can it help?

Simply put, the idea that most problems can be solved by increasing bandwidth misrepresents the reality of voice services.

Because problems cannot be solved so easily, professionals who deal with business telephony need effective ways of detecting anomalies and diagnosing them, with a view to continuous improvement of end-user experience.

Specialists of DEM for voice applications have the tools and experience to do this. They work **hand-in-hand with telephony providers** to help them live up to their service level agreements (SLAs). And they lend a **helping hand to businesses that rely on enterprise telephony** to make sure they get the best voice communications for their investment.

2. Mission-critical voice services for businesses

Among the types of company for which **high-quality voice services are mission-critical,** the following are just a few of the typical customers of DEM for voice applications:

- **enterprise-class businesses**
- **call/contact centers, telemarketing companies**
- **financial services** (banks, insurance companies, payment providers, etc.)
- **telecom and IT solution and service providers**
- **public services** (public health, transportation, tax administration, etc.)
- **e-learning/educational media and technology providers**
- **gaming operators/suppliers**

For call centers and telemarketing service providers, whose core business involves massive quantities of telephone work, **IVR and speech quality monitoring are essential.** Call queuing and call transfers ideally should work with **customer relationship management (CRM) software** to provide good customer service.

Companies in this sector, along with **CCaaS** operators (contact center as a service) and providers of business IT and telecom solutions, are among the key customers of DEM for call applications and all types of voice communications systems.

Because enterprise-class businesses and financial institutions, like many smaller companies, rely heavily on phone services, **they need high-performance telecommunications systems and software.** Their workforces use **fixed-line telephones** as well as **PCs (softphones)** and **smartphones** to communicate internally and externally. The clarity of communications and availability of connections is essential to business. Enterprise telephony providers therefore have to ensure **clear, fast, reliable phone calls, voicemail, telecommuting, video conferencing, and other voice services** on PCs as well as on mobile phones and tablets.

At the same time, providers of services to consumers, like banks, insurers, online retailers, educational institutions, and public services all need to make sure their voice technology is running properly, and that voice servers and agent hotlines are serving people well. **Callers must be able to hear and obtain the correct response at each step offered by interactive voice menus (IVR) on any device.**

In online learning or medical contexts, **e-learners or patients must be able to log on with their smartphone or computer** to use audiovisual sessions to interact with people and content.

For all of these sectors, this white paper takes a brief tour of the professional telephony landscape and reviews some of the trends and challenges of voice communication technology.

Next, it describes some of the most widely used ways of representing speech and audio quality from the end user's perspective, and goes behind the scenes into the technology of perceptual speech quality monitoring.

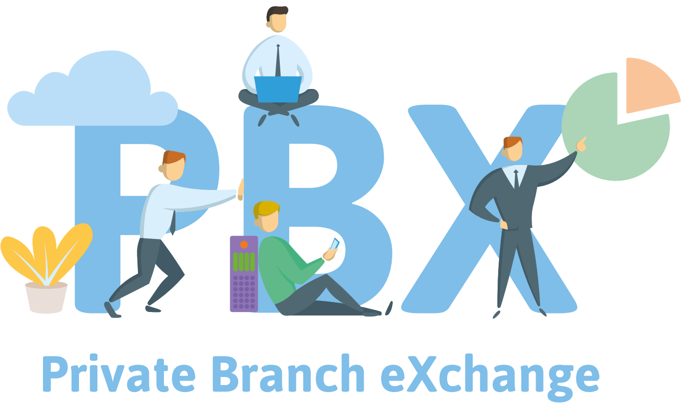
Lastly, it takes a look at typical uses of voice monitoring from real-life business cases.

Voice communications and today’s technologies

Enterprises have naturally been seeking ways to improve their systems at the lowest cost, and the specific constraints of each company are different. **The needs, legacies, and circumstances of voice communications** within a business must be weighed carefully.

1. Conventional networks and digital telephony

A number of phone systems come into play in modern corporate telephony. A corporate telephone system may be built on a **conventional PBX or Centrex, traditional key systems,** or an **IP PBX or hybrid IP/TDM PBX**, **along with enterprise voice applications, computer-telephony integration (CTI) technology,** and so forth, to deliver voice communications and services.



Such a variety of business phone systems results from the coexistence of several types of transport networks or communications channels. More businesses than you might think still rely on the analog public switched telephone network (PSTN) or "plain old telephone system" ("POTS"), even as it is already being phased out in many countries.

The old analog landline telephone network was originally designed specifically to carry voice communications. Over the years, PSTN was enhanced with digital capabilities to carry data traffic as well. **The integrated services digital network (ISDN),** for example, made it possible to **transmit voice alongside video, data, and digital services.** Many companies use a combination of PSTN + digital subscriber line (DSL) or have built their services around ISDN.

Now, of course, there are also all-IP networks, used for the internet, and these IP networks can carry voice as well. In this case, voice is transformed into data packets and transmitted as data. This gave rise to **telephony over internet protocol (ToIP)** which, for the purposes of this white paper, is grouped together with **voice over internet protocol (VoIP).**

1-1 Legacy systems and digital transformation

The experience of DEM specialists shows that analog fixed-line business telephony is alive and well in diverse legacy systems, alongside digital telephony. At the same time, a recent study confirms that for corporate professionals in the 5 European countries surveyed (Spain, Germany, Italy, France, and the United Kingdom), the desk phone is still by far the preferred tool for professional use, ahead of smartphones and PCs* ...and **traditional telephony is still a major channel of voice communications.** This is why **it is important to plan for connectivity between VoIP solutions and PSTN**, at least until the analog telephone network has become a thing of the past.

To date, **less than one-third of businesses in Italy, France and Spain have made the move to "full VoIP"** (more than one-third in the UK, and edging up to 40% in Germany) so far. An average of around 15% of businesses in all five countries combined are planning to transition to VoIP in the year ahead*.

Those businesses need to **know how their planned digital telephony compares to their existing telephone systems**. They will also need to make the transition smooth for their users, as well as ensure better service on VoIP telephony in order to justify investments in new systems. Such enterprises are among the main customers of voice application performance management services.

2. VoIP

As stated above, one major trend in business communications over the past decades has been to **migrate to digital communications, where phone calls are delivered using the internet protocol (ToIP or VoIP)** – a move which paralleled developments in the consumer telephony landscape.

* *Findings from a report from 2019 commissioned by Snom from Norstat, an independent research institute*

At first the main attraction of VoIP with respect to conventional telephone networks was **cost-savings** (free long-distance calls on softphones, for example). In terms of the physical infrastructure, VoIP installations can replace the expensive copper wires used by conventional corporate telephone networks. VoIP can be transmitted over any type of internet access (cable, ADSL, fiber, wireless, and mobile).

By migrating their phone systems to VoIP on all-IP networks, enterprises can **maintain their infrastructure with software**, instead of having to buy and replace hardware to obtain new features.

Over the years, the reasons for migrating from traditional phone systems to VoIP telephony have evolved. According to a recent study, cost is still the most important factor, but **better speech quality and increased mobility have become very strong motivators***. Another main selling point nowadays is the variety of services VoIP can provide, especially in computer-telephony integration (CTI) contexts. Enterprises can also opt for hosted VoIP solutions, which may offer telephony as part of **united communications (UC) systems**.

2-1 SIP (and RTP)

One VoIP technology that gets a lot of attention is **SIP (session initiation protocol)**. During a phone call, **SIP deals with signaling**, so that one device can "tell" another device what it wants to do. **RTP (realtime transport protocol) takes care of transporting the audio data.**

Business VoIP systems can use SIP trunking to connect to SIP provider server networks on the internet, for example, or direct SIP connections to a PSTN gateway or IP-PBX. **This way businesses can handle PSTN phone calls on their VoIP systems to ensure that communications don't get "lost" between the two technologies.**

Voice applications can be developed to integrate via SIP with almost any telephone system regardless of hardware. Furthermore, **SIP enhances network security** across devices and facilitates the delivery of hosted VoIP for enterprises.

There's a lot to say about SIP and RTP; suffice it simply to note that SIP is a VoIP technology that has changed the way CTI (computer-telephony integration) is delivered.

* *Findings from a report from 2019 commissioned by Snom from Norstat, an independent research institute*

3. CTI and customer-facing services

While CTI is not such a new technology, it used to be very expensive and complicated to implement back in the days when it was primarily based on proprietary systems. With SIP-enabled customer service and widespread IP environments, **CTI options became more affordable and scalable.** CTI means voice communications interact with other technologies to enhance efficiency and productivity.

Customer-facing services like **IVR, automatic callback, and click-to-call** are obviously a main concern for call centers, as well as for any business which offers customer care on the phone. **Wherever customer experience is at stake, CTI can be what differentiates a business from its competitors.** To take the example of contact centers, integrating CRM into the call flow for automated instantaneous display of customer information is a key requirement for providing smooth, fast service to customers on the phone.

4. UC and productivity in the enterprise

Finally, taking business communications one step further, providers of **unified communications (UC) solutions see telephony itself as just one part of a single platform** which unifies contact capabilities beyond the telephone to all devices (PCs, smartphones, etc.) and a whole range of software.

The complexity of sophisticated UC suites means additional demands on internal IT staff as well as a whole new set of required skills. This is why many enterprises turn to **UCaaS**, additionally benefitting from capex-to-opex possibility and scalable capacity, as is the case with many solutions provided “as a service”. SOA-enabled **UC makes telephony just one more software service, which can interact** with data applications, knowledge management tools, and more in a person-oriented architecture.

5. Where voice services fall short of expectation

In the shift to digital communications, many factors which affect audio quality – how well the speaker can be heard by the listener – have not been resolved, and often are even perceived as worse with digital technologies, from the user's point of view.
Why? **Because in the same way as analog networks were originally designed to carry voice traffic (not data), the original purpose of IP networks was to transmit data (not voice).**

Nevertheless, IP-based networks can handle voice traffic when VoIP technology transforms voice into digital signals. However, instead of allocating a dedicated circuit for each call, as on a PSTN, **IP packet-switching networks divide voice calls into packets** that are sent individually. Packets may be lost (**packet loss**) or arrive unevenly spaced (**jitter**). These and other factors, mentioned below, may garble voice communication or cause echoes or dropped calls and effectively undermine the audio quality of telephone calls.

One journalist sums up the situation in laypeople's terms:

> *Voice call quality has gone down over the past two decades. Mobile phones have added convenience and a million other things, but they have done away with a wired network dedicated solely to voice communication, as well as the large microphones and speakers of old landlines, which featured decades of refinements to improve call quality. Cell phones, on the other hand, rely on tiny, awful microphones, and tiny, awful speakers for calls, and tiny, awful allocations of bandwidth. Mobile networks can be compromised by everything from streaming video, to the presence of a tree or wall, to the weather. Most importantly, they've dropped the emphasis on voice quality. Nobody seems to care any more whether they can hear a pin drop.*
> *Sure, much communication has moved over to text, email, and social platforms, but everyone still needs to talk on the phone sometimes.*

Dan Nosowitz, "It It Might Be Time to Update the Old 'Alfa-Bravo-Charlie' Spelling Alphabet" https://www.atlasobscura.com/articles/best-spelling-alphabet 12 July 2019

So it is not unusual for callers to experience some kind of **interference, distortion, lag, or reception problem**, whatever the underlying cause may be. Because of sporadic speech quality, people have learned to cope, using various techniques to make themselves understood, such as spelling words out with the civil aviation spelling alphabet.

Mishaps with voice/video technology are so common that mainstream media pokes fun at digital call quality. Screenwriters use sudden disconnections of voice or video calls as a plot device for comic relief, to signify a troubled relationship, or to enhance suspense in TV series and movies. But in a crucial real-life conference call or webinar, unreliable voice application performance is detrimental to productivity and erodes customer trust – there's nothing funny about it.

In addition, speech recognition and speech-to-text functionalities are still far from perfect, even if great strides have been made. One study found that **at least 1/3 of people who use voice interaction wind up texting something inappropriate.** Funny voice texting failures are posted all over the web, second only to disconcerting spell check blunders.

A similar situation in a professional setting could be more damaging.
When a customer tries to speak to the voice server and is not understood, it is not merely awkward. **If the voicebot cannot process the customer's speech or synthesize a spoken response, the experience is infuriating, stress-inducing and time-consuming.**

IVR: Please say the name of the department you are trying to reach.
Client: sportswear
IVR: <pause> Please say the name of the department you are trying to reach.
Client: sportswear
IVR: <pause>
Client: SPORTSWEAR. SPORTSWEAR!
IVR: <pause> We're sorry. <pause> The department you requested does not exist. Please hold for a customer service agent, or hang up and dial again.

The title of one blog article, "Exasperation, or the new normal?", says it all. Call quality exasperation has become a 'thing' in the digital age, and people are loath to give up landline phone service.

> *"I had heard rumours that the telephone companies might try and phase out fixed lines....God forbid that should ever happen."*
>
> *http://www.franklinfarm.fr/blog/exasperation-or-the-new-normal _ 31 July 2017.*

Yet the fact is that landlines are scheduled to disappear. People therefore need to be assured that digital telephony works. They want to know that improving the quality of telephone services is a priority for businesses and telephone systems providers and operators.

Understanding and improving user experience: MONITORING & TESTING

CHAPTER 3

In response to the frustrating voice situations everyone is all too familiar with, research institutions, telecommunications regulatory organisms, and digital experience management specialists have teamed up to devise ways of understanding the quality of voice traffic with a view to improving it.

While desire for technological progress is itself a major impetus, businesses and their communications providers are motivated by the practical considerations of **productivity and customer satisfaction**. It is with these business-oriented goals in mind that tools and services have been developed to monitor and measure the performance of voice applications.

To best understand how users experience voice communications, DEM specialists implement **various tools, methods and services to investigate call quality, identify anomalies, and propose solutions.** Among the most useful for measuring the quality of voice applications are:

◊ **Synthetic monitoring** (also called active monitoring) **simulates voice traffic with automated transactions to measure availability** of a conferencing service or a contact center operator, for example, and **detect degraded speech quality from the end user's standpoint in real time. This is the most complete way of testing voice applications,** and currently is the most widespread.

◊ **Load testing measures how well the infrastructure handles voice traffic,** and is particularly helpful for **testing before/after upgrades or a migration to a new system.**

Other approaches may be proposed by voice DEM experts depending on your testing requirements. In some cases, **real-user monitoring (RUM) can contribute to the overall picture of network traffic performance down to the last mile**. RUM also can be a useful technology for testing connectivity in mobile situations, for instance, and for investigating the quality of call apps on mobile networks.

1. Synthetic Monitoring

Synthetic monitoring, also called **active monitoring**, is currently **the state of the art for voice transactions.** Synthetic monitoring **automates the process of collecting user-experience data about the quality of voice services.** Instead of asking human callers for their feedback after the fact and then processing their subjective responses, specialists use robots to collect and feed back data in real time.

To do this, **robots are set up to perform and measure the actions of both the caller and the receiver ('bidirectional'), including SMS and synthesized speech** where applicable. The robots run custom scripted **'scenarios' or transactions** on real devices **to interact with the voice service as a human would in real life, at pre-defined frequencies around the clock.**

This kind of monitoring can measure any type of voice application usage context: **fixed-line telephones, PC softphones, conference calls, mobile phones, IVR, voicemail, text-to-speech, speech recognition, and so forth.**

Some typical use cases of synthetic voice monitoring:

- **For IVR servers,** testing may involve robots listening to DTMF sequences ("answering"), sending synthesized speech messages, listening to voice messages, hanging up, issuing DTMF sequences ("dialing"), sending text messages, and so forth.

- **For voice calls and voicemail,** communication can be tested on standard telephones, mobile phones, or softphones to check availability and performance in both directions..

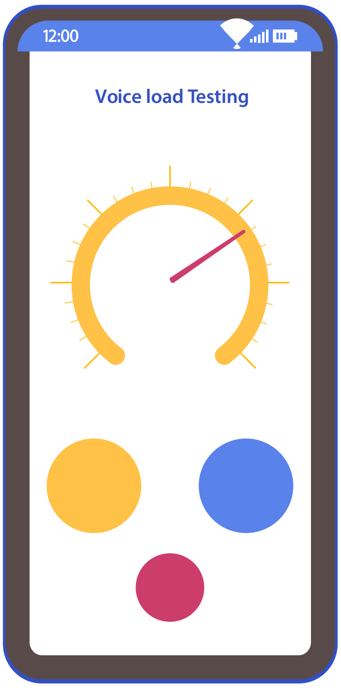
- **CTI installations** can be monitored with multiple-mode scenarios (for testing web, mobile, etc.), for instance to simulate call center operators and test their availability.

Synthetic monitoring robots can test short calls to obtain metrics (availability, dial tone, busy signal, garbled message, dropped calls, call set-up time, post dial delay), as well as **long calls** (for stability measurements and more) **on any kind of communications network** (analog/PSTN, as well as ISDN, GSM, GPRS, 4G).

This kind of monitoring and real-time measurement is ideally adapted to the requirements of technical accuracy on real-life networks and the derivation of subjective performance values.

2. Load Testing

Load is as much a challenge for voice communications as it is for any other web or business application. **For voice applications, load testing should be considered along with monitoring**, for example before and after software/hardware updates or platform migrations to help with change management.



The approach is similar to synthetic monitoring, insofar as **robots run scenarios to simulate user traffic**. With load testing, however, the focus is on **infrastructure capacity in targeted campaigns, rather than continuous surveillance of end-to-end performance.**

Load testing can be conducted on **LAN/VLAN infrastructures and with a WAN link** to identify any snags or contention (for example, in RTP traffic, at firewalls, etc.). It is a good practice prior to any production rollout of VoIP systems and terminals.

With these kinds of tools and services, DEM providers monitor and measure **performance from the point of view of the end user.** The resulting metrics enable DEM specialists to diagnose problems and point you in the right direction to fix them.

CHAPTER 4

Understanding and improving user experience: METRICS & INDICATORS

Metrics are the ‘outputs’ of monitoring. Measurements are made (by a robot, by the browser, on a network component, etc.) during automated use of a voice application or service. These measurements are processed or calculated in accordance with accepted standards to ensure their representativeness, take certain conditions into account, factor in user perception, and deal with other particularities.

Below is an overview of some of the different types of metrics encountered in the context of monitoring speech quality on VoIP services.

1. Voice data transmission metrics

The quality of voice communications depends on the seamless transmission of audio data back and forth from the sender (caller) to the receiver (callee). Performance problems can occur at any or several of the many steps in the call stream.

To take a simplified example of a SIP server/network on the internet or a corporate LAN or WAN (dedicated networks): the call must be set up, the signal converted into data, the packets transmitted in an RTP/UDP audio stream, and the confirmation returned. Each of these steps involves a series of negotiations or processing: to connect the call, to convert/encapsulate the voice flow into data/packets/frames, to sequence their transmission, and so forth.

There are tools to measure the different steps to keep track of how voice traffic is doing. Among the most commonly measured and calculated transmission details for voice applications are network parameters such as:

- **latency (delay) represents the time between the moment a voice packet is sent from one endpoint and the moment it arrives at destination**; it is perceptible to the user as a delay or echo.
- **jitter (delay variation or deviation) occurs when packets in a stream are unevenly spaced** such that the delay between packets varies.
- **packet loss means that some of the packets transmitted do not reach the destination**; this is perceptible to the user in the form of dropped calls, extraneous noise, or lags in the conversation.
- **error rate measures the number of errors in data units with respect to the total number of data units transmitted.**
- **SIP session details from HTTP traces**
- and others...

Depending on their intended purpose, these metrics may be named differently or be supplemented by additional indicators. Some such commonly encountered metrics are **availability, call setup, call throughput,** etc.

In addition, **metrics may be broken down into detailed measures.** For example, latency may be broken down into processing delay, queueing delay, transmission delay, and propagation delay. Availability, another important metric, may consist of separate measured times for the user journey as a whole in addition to timings for each step.

No single metric alone is enough to represent the performance of voice communications on any network. Nor does measuring these various parameters provide an adequate picture of the end user's experience of voice applications, whether on a fixed-line telephone, a computer (softphone) or a mobile device.

To obtain an accurate representation of how the user experiences a voice service or phone call, **attention must be paid to quality as the end user perceives it**. Read on...

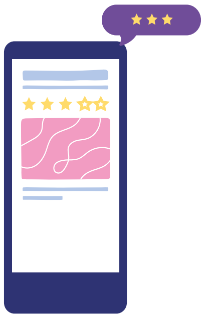
2. End-user perceptual quality ratings

The purpose of monitoring voice services is to measure not only how well voice traffic is transported over networks, but to assess the quality of the communication as users perceive it. Without good speech quality, a user's experience can only be bad.

In addition to the various metrics representing impairment factors, there are also **standards for measuring the end-user's perception of the quality of voice services**, i.e. perceptual quality. To represent the overall quality of experience, two of the values that are often referred to are **MOS and R-Factor.**

- **MOS (mean opinion score)** is **a rating of perceptual speech quality** on a scale of 1 (bad) to 5 (excellent). Experts sometimes use MOS specifically for insights into codec performance to check how compression affects call quality in terms of user experience. MOS measures are generally conducted under controlled conditions, like in the calibrated environments where synthetic monitoring is implemented.

- **R-Factor (rating factor)** is **a quality metric which yields the percentage of users satisfied with the quality of a received voice signal**. Its value is calculated based on metrics such as latency, jitter, and packet loss. Typical scores range from 50 (bad) to 90 (excellent).



How do we get end-user quality of experience indicators like MOS or R-Factor from automated testing?

Simply put, these user-side ratings are calculated according to quality models, on the basis of the measurements collected during monitoring. These quality models are our next topic.

2-1 Objective quality models and perceptual quality

Objective quality models are algorithms which approximate or estimate the level of quality as humans subjectively perceive and assess it. Most objective quality models feature in recommendations from the ITU-T (the International Telecommunication Union's Telecommunication Standardization Sector), which oversees telecom and computer protocol specifications.

Some of the most frequently used standard speech quality assessment models nowadays are **POLQA, PESQ** (Perceptual Evaluation of Speech Quality, for any type of network) and **E-Model** (for IP networks). There are others, including ITU-T standards such as **PSQM**; another example is **PAMS**, developed by British Telecom.

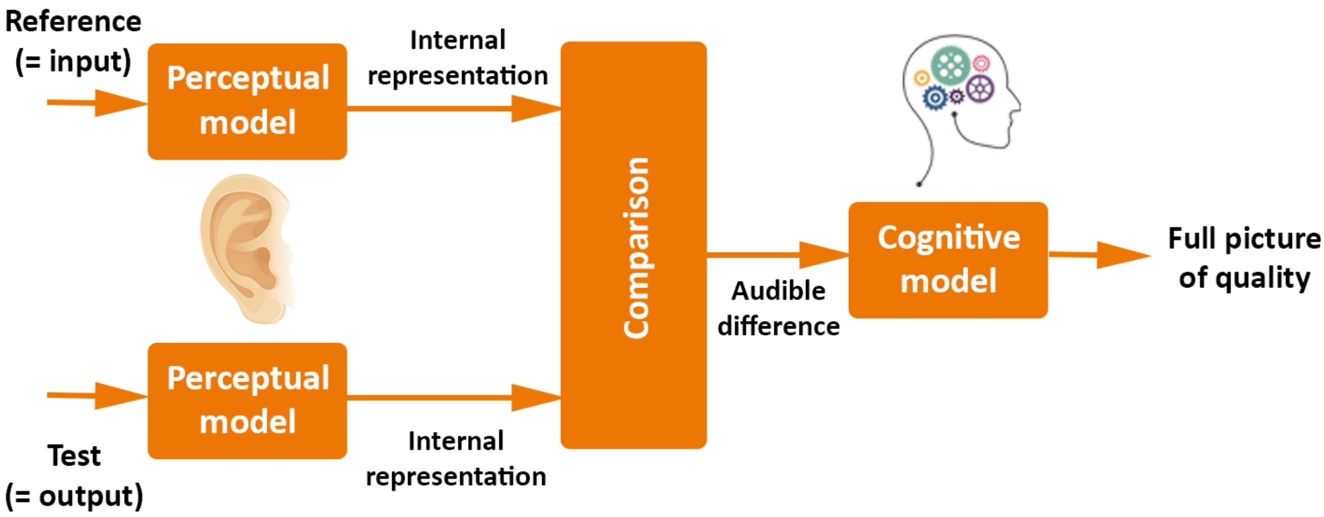
These algorithms are categorized according to the information that is made available to them. Here are two types of approaches:

- **NR (no reference)**: **A speech-quality test algorithm can model subjective tests using only the audio signal itself.** These NR (or "no reference") algorithms are useful mainly as estimates or for transport-stream analysis.

- **FR (full reference) algorithms compare a test sample to a reference audio file which is provided to them.** FR algorithms offer the highest accuracy for dedicated testing in production; in other words, **FR models are used to represent end-user experience.**

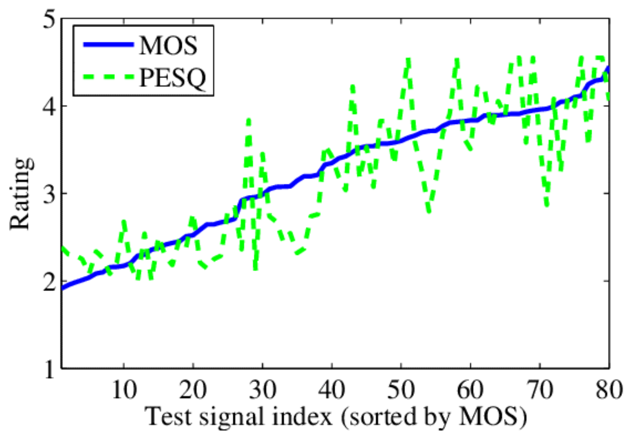
PESQ and POLQA algorithms are full-reference standards which primarily calculate MOS for listening/talking situations on any kind of telecommunications network. **They are used to derive a full picture of the quality of the user's experience.**

The diagram below shows the steps involved in a full-reference approximation of the quality perceived by the user.



Source: Opticom GmbH

Human perception of the acoustic quality of voice traffic over the network can be expressed theoretically as a MOS rating, on a scale of 1 to 5, as mentioned above. In real life, however, the maximum score varies according to the **type of technology or network** involved (GSM, VoIP, PSTN, ISDN) and especially the **compression method** (codec) used.



Source: researchgate.net

According to Voip-info.org, the ratings for VoIP calls most often fall between **3.5** (fair, where impairment is slightly annoying) to **4.2** (good, where impairment is merely perceptible). There is room for improvement!

In addition to MOS, objective quality models can also derive a number of secondary speech indicators, including **noise index, loss index and offset index**. These may be expressed as a percentage with respect to a reference file.

In a nutshell, these different steps and calculations are how automated monitoring produces meaningful speech quality metrics. These metrics paint a picture of the quality perceived by users of voice applications and telephone services.

Understanding and improving user experience: ANALYSIS & EXPERTISE

CHAPTER 5

Measuring service and collecting the right metrics are only two steps leading to problem detection and substantiation. The next step, toward problem resolution, is to **analyze metrics** to identify the type of problem and pinpoint its origin.

On the basis of objective measurements, an experienced application performance specialist can offer advice on how to improve the situation. This is whole point of monitoring voice applications and gathering metrics: **to improve the quality of voice communications and deliver the best possible user experience.**

Specialists refer to many types of data to carry out their analysis. These include **diagnostic information from call detail records (**timestamp, duration, loss, latency, etc.), **packet capture** ("pcap", an umbrella term for various APIs used for capturing network traffic), and **additional data from network audits.**

These types of information are weighed alongside perceptual quality indicators from synthetic monitoring to provide meaningful, in-depth analysis of the speech quality experienced by end users.

Because the relationship is not easy to ascertain between a metric, on the one hand, and a performance issue and its root cause, on the other, **calling on experts can be a huge timesaver.** An DEM specialist can use indicators and run tests to **accurately diagnose root causes and prescribe a treatment.** Sometimes the problem can be straightforward; an experienced analyst systematically investigates several areas.

Some usual suspects:

- Codecs: **the way voice traffic is encoded and decoded** is an important factor which determines the quality of VoIP communications.

- Configuration/traffic management: tuning **firewall or load balancer configuration** can help speed up VoIP traffic flow; likewise, troubleshooting **DSCP** or optimizing **bandwidth management** (packet or traffic shaping) may be required to smooth VoIP communications.

- Hardware: **components** may be defective or inefficient, **cables** may be poorly installed or deteriorated, router problem, etc.

- Software: **overdue updates** or other issues

- Buffer: **size, jitter correction, etc.**

Nevertheless, every problem cannot be solved simply by investigating the usual suspects. Determining the best configuration can be tricky, and any change in configuration requires further testing to ensure that it yields an improvement. Expert analysis can point you in the right direction and help you document the results of changes you make.

Furthermore, **a performance issue may be caused by several factors**, some of which may be much more costly than others to resolve. DEM specialists can help target priority areas for improvement, thereby **saving you not only time, but money as well.**

1. Voice application monitoring in the real world

Experience shows that most of the needs for voice application quality monitoring can be boiled down to just a few questions. Enterprises – and the business telephony providers they rely on – want to know:

- **Is service good?**

If not:

- **What is the problem (how can it be fixed)?**

and

- **Where is the problem (whose responsibility is it)?**

Questions may be phrased differently (Are calls from abroad getting through? Why are users complaining? Is there some problem with mobile phone access? Are calls being dropped? Why do so many transactions remain uncompleted?), but the concern always boils down to these three questions.

The following case studies offer a sampling of the kinds of needs that have been addressed with DEM voice application monitoring solutions and services. Voice application performance specialists **adapt to the specific technical and budgetary requirements of their customers**, and offer tips and advice based on experience.

"These days, customers don't buy products to ensure and improve speech quality – they buy services which implement tools that are appropriate for the need at hand," explains Didier Revelle, a voice application performance specialist at ip-label. *"One of the great things about working in this field is the creative part, where we learn about the customer's specific issues and figure out how to use our tools to help make their voice technology work better for them."*

In other words, voice application performance services are always tailored to the customer's needs. Tools or products alone ensure objective measurement and calculate indicators, but knowing which tools to use and how to use them is a matter of experience. Analysis by specialists is paramount for improving the quality of experience on voice applications.

1-1 SLAs in VoIP telephony

Case study

A small company with a workforce of 50 provides professional VoIP telephony services and software to contact centers. The required level of availability and speech quality of the telephone services it provides is defined in service level agreements (SLA).

The VoIP company wished to monitor the availability and speech quality of the services it was providing to one of its customers, a call center. It asked an DEM specialist for help. In response to the need, a synthetic monitoring probe was set up to test a telephone number using a voice scenario to measure quality every 15 minutes. The purpose was to test calls using a voice message to analyze end-user quality, and provide availability and MOS indicators.

The monitoring solution was able to supply metrics which gave both parties (the VoIP provider and the contact center) objective indicators of the quality of service delivered with respect to the SLA. This was beneficial to both sides, and helped to locate the source of incidents. Following this outcome, the telephony provider decided to implement the DEM solution each time any of its customers required SLA monitoring.

1-2 IVR monitoring and diagnostics for insurance policyholders

Case study

A leading international insurance company specializing in car and health coverage on the phone and on the internet needed to make sure its IVR service was running properly. The insurer's phone service lets customers manage their policies and obtain roadside assistance at night.

The insurance company had noticed that its IVR was sometimes unavailable, and customers were complaining that they could not talk to an agent to ask for roadside help. The insurer called on its DEM partner to set up a voice monitoring solution to identify the problem. They decided together to have the DEM specialist deploy a robot at one location to conduct SaaS synthetic monitoring of the insurer's IVR services.

The robot ran 2 IVR DTMF testing scenarios. One scenario dialed up to test the availability of car policy management in 5 steps, running at a 5-minute frequency from 8 am to 10 pm. The other scenario placed calls to test the availability of agents for roadside assistance at night in 2 steps, at 30-minute intervals, from 11 pm to 8 am.

The diagnostic information supplied by the DEM solution enabled the insurer's teams to find and resolve the problems that had been occurring in their IVR services. This case was only one instance of the insurer's successful collaboration with its DEM specialist. After years of synthetic monitoring in various situations, the insurer decided to upgrade to include mobile app monitoring in its DEM package.

1-3 Testing of IT helpdesk availability and speech quality around the world

Case study

A leading global telecom operator/contact solutions provider ensures telephony services for an internationally renowned IT service company with a presence in 30 countries. The IT firm discovered that in some countries people were having trouble reaching their local support helpdesk on the phone. The only way the IT firm could check the problem internally was by calling its various offices abroad, but this could be done only manually and not necessarily during office hours or in the right language.

The telephony provider wanted to offer the IT firm a service for monitoring the availability of its support service in each country, including Australia, India, Argentina, Brazil, the USA, Canada, the UK, the Nordic countries, and across continental Europe, from Spain to Poland. The IT firm wanted real-time alerts whenever service disruptions occurred on phone calls.

The telecom provider consulted an DEM specialist to determine the most cost-efficient way to monitor performance at the IT firm's various locations. The voice DEM team found that instead of deploying a fleet of robots, the need could be addressed first with synthetic monitoring limited to a combination of direct-call and call forwarding.

This would be the least expensive way to test connections to the IT firm's different helpdesks on 30 phone numbers every 15 minutes. The call scenarios would run during office hours in the time zones specific to each of the 30 countries and test the speech quality of messages in the local language.

The SaaS solution set up by the DEM specialist would include scenario scripting, upkeep of the solution, provision of a real-time monitoring console, statistics and reports, and an alerting system, as requested by the telecom provider's customer.

The simplified approach was successful; the DEM specialist's monitoring system yielded value at the lowest cost. The telecom operator later wished to continue its collaboration with the DEM specialist, to offer its customer more than call connection metrics and incident alerts. The next step in the telecom operator's quest for excellence would be continuous testing of IVR operation for the IT firm.

1-4 Monitoring the availability and performance of a public health hotline

Case study

A provider of a public health service hotline wanted to make sure its telephone service was available and performing well at all times. Registered nurses take calls to advise and orient patients seeking medical help in order to make doctors' appointments more efficient.

To ensure that the quality of experience was consistently good and be alerted whenever service quality deteriorated, the provider consulted an DEM specialist about how to approach this need proactively. Together they determined that the solution would consist of two-way synthetic monitoring with voice robots to test the hotline at regular intervals around the clock, from the patient's side and from the nurse's side.

The voice DEM specialist configured scenarios in consultation with the hotline manager. The 'patient' robot would place a call, listen for a response from the 'nurse' robot, record it, measure the response times of the call, and calculate perceptual speech quality against a reference message. Then the test would run in the other direction: the 'patient' would simulate a voice interaction (play a recorded message), and the 'nurse' robot would listen to it to measure its quality in MOS terms.

The hotline provider's designated staff would receive alerts whenever issues occurred, and would also receive reports with MOS indicators as well as availability and response time metrics to document the quality experienced by hotline users on both sides (nurse and patient). This is one example of the most straightforward kind of proactive synthetic monitoring in a speech context.

1-4 Ensuring IVR functionality for mobile and landline callers

Case study

A few years back, a provider of IVR platforms and other contact center solutions for enterprises was seeking to prove and improve the quality of the services and infrastructure it provides to businesses. Its need was to monitor its IVR platform from a PSTN, an ISDN, and 3 mobile telephone networks, and to be informed of any internal problem on its LAN.

A primary concern was to identify any problems and determine their root causes. Is it a telecom problem outside of the IVR infrastructure (mobile or fixed-line operators)? Is there any problem connecting to the voice server within the IVR infrastructure (IP WAN, IP LAN)?

The contact solutions provider contacted an DEM specialist to find out what kind of technical solution could be set up. The specialist prescribed a solution which consisted of two synthetic monitoring robots, each with three application queues. One robot would run call scenarios to test access to the IVR from the three mobile networks. The other robot would use two queues to generate calls on the analog and digital landline networks, and a third queue for direct VoIP access to the local network.

The IVR-specific indicators that were collected included call set-up time, duration of each step (delay before receiving the expected message), availability of the IVR tree end to end and for each step, and error type in the event of failure (with a recording of what the robot heard, as a diagnostic aid).

The DEM specialist was able to propose a solution for the range of voice technologies (mobile, ISDN, PSTN, VoIP) involved in the IVR provider's request. The types of metrics and diagnostic information collected by the robots were aligned with the IVR provider's needs.

1-5 Videoconferencing quality on a dedicated national network

Case study

An organism which manages a national network dedicated to communications in the fields of technology, education, and research wished to check its videoconferencing system to make sure that remote participants are able to set up and log on to video conferences.

The need was to test this dedicated national network from end to end, to measure the availability and performance of every step that a real user would perform to set up and/or join a video conference using a softphone. The network managers contacted an DEM specialist to find out how they could test their videoconferencing system to measure the quality delivered to users on both sides (conducting the conference and attending the conference).

The prescribed solution consisted of 2 synthetic monitoring robots to test the steps to set up a videoconference, and to test the quality of interactions during a conference. One robot would run scenarios scripted to log on to the network, locate the videoconferencing setup page, reserve a time and choose invitation recipients, and initiate a conference. The other robot would receive the invitation, and log on to the conference.

In real life, different users connect from different devices and systems (PC, Mac, tablet, smartphone, telephone/ISDN, GDS, SIP, H.323, IP, etc.), so the robot would run scenarios configured to join the video conference using a variety of protocols. Once the videoconference was open, a number of measurements would be made of the quality of experience, including speech quality.

The robots conduct response time and availability measurements from the end user's perspective, at every step of the softphone's interaction with the national research network. The resulting metrics show exactly which step is slow or fails, and alerts are set to notify personnel of degraded service. End-user quality indicators are calculated to represent the quality of experience delivered to users of the videoconferencing platform.

1-6 Measuring audio + video quality with RUM and crowdsourcing

Case study

A leading telecom operator teamed up with research institutions, a European DEM provider, and public authorities to measure the audio and video quality of streaming media. The project aimed to test the perceptual quality of mobile networks and services. For this it used objective real-user monitoring (RUM) and subjective crowdsourcing.

The DEM provider's R&D team delivered the metrology network for the project, and the partners worked on an application for monitoring mobile streaming video. This app, which users downloaded onto their device, automatically connected to mobile networks and collected anonymized data about the connection as users rode mass transit and commercial rail lines. In addition to gathering data from the device with RUM, the app allowed interactive input from passengers (crowdsourcing).

The immediate aim was to combine automated monitoring of quality of service with quality of experience data from real users to obtain results that are perceptually accurate and predictive for mobile telecom providers.

But that was just the first step. Additional, wider research possibilities grew from the project, whose greater relevance to voice services, gaming or learning platforms, and other digital media is evident. End-user monitoring strives to improve audio and visual communications, of course, but it also serves a larger purpose: to drive and support future developments.

Conclusion: Voice application monitoring in the foreseeable future

Voice monitoring will most likely advance further toward multimedia monitoring (video, audio, text). Solutions currently exist for each of these needs; they are already being combined to provide an overall view of quality and performance in CTI contexts, and developments are in store for wider use in UC.

The quality of experience delivered to users of voice communication services is more than the sum of technical operations provided by systems and networks. Because user experience is perceptual by nature, technical metrics are not enough to represent the quality experienced by users.

Advances continue to be made in end-user quality of experience of voice and video applications. The R&D departments of DEM solution vendors play an active part, alongside telecom operators and research institutions, in developing ways to assess the performance of audiovisual communications. (The case study above is just one example.)

In the meantime, consumers may begin to expect operators, providers and manufacturers to implement and publish quality standards for call applications and such functionalities as IVR, voice texting and video calls, not to mention online medical/health services, voice orders, and many more.

Demand for this kind of monitoring will grow as voice technologies develop. At present, many of these services are still “newish”, and the consumer market is not yet extremely demanding.

That time will come soon.

About the ip-label group

Since 2001, ip-label has been helping enterprises in more than 25 countries to manage and optimize the performance of their critical applications (website, business apps, mobile apps, voice apps, etc.), using software solutions for synthetic monitoring (robots) and real-user monitoring (RUM). Its user-centric DEM (Digital Experience Management) products and services empower organizations to improve the availability and response times of their applications, motors of business success, with a view to growing their audience, revenues, and productivity.

Want to know more about application performance management?

Request a demo

ekara
by ip-label

ekara.ip-label.com
+33(0)1 77 49 53 00
marketing@ip-label.com

Paris I Madrid I Stockholm I Helsinki I Shanghai I Tunis