

A Question of Quality – VoIP, WebRTC or VoLTE?

Rebecca Copeland
Institute Mines Telecom,
Telecom Sud-Paris, France

Michael Copeland
Core Viewpoint Limited,
United Kingdom

Abstract— WebRTC introduces real-time media to any website with a few lines of code. This threatens the carriers' stronghold – Voice, but could also be an opportunity to change their DNA and offer QoS web-calling alongside managed Voice. This paper proposes QoS enhanced web calling architecture for collaborative network providers, who negotiate session policy across different domains, (unlike the silo OTTs). This QoSweb service is positioned between OTT VoIP and fully-managed services (VoLTE, Enterprise, Private-Mobile). It is an affordable mid-way quality service that would mostly appeal to SMEs, who value confidentiality and predictability. To select the most suitable (and affordable) service 'mode', the paper proposed a modelling tool that optimizes the service delivery according to the user context. The context is computed from QoS, Urgency, Security, and Affordability (QUSA) attributes of the request, which influence the decision in different ratios.

Keywords—WebRTC, QoS, Context Calling, VoLTE, TURN,

I. INTRODUCTION

A. The service environment metamorphosis

Voice over LTE (VoLTE) is now rolling out into mobile networks, but the service environment has already moved elsewhere. Equivalent services, such as RCS (Rich Communication Suite) have not grabbed a market share so far. Alternative communications methods (text, image, and social media) are eroding Voice popularity, due to the immersive power of social media. At the next stage, social media gains instant response, thus threatening the operators' stronghold – Voice (meaning here: conversational real-time voice and video). This is happening just when Mobile operators are implementing managed Voice over IP networks, so now both Telecom Voice and web Voice are carried over data networks - separated only by different 'service mode': managed or unmanaged. As LTE and multi-service access gateways are already access agnostic, and devices can run any application, selecting the service mode becomes 'a question of quality'.

With real-time web calling, Voice can be added to any web page with just a few lines of code, and the browser's native facility does the rest. Web technologies, such as HTML5, WebSocket, and WebRTC (Web Real Time Communications) are different from the old Internet VoIP, because they allow for any web page to act as the communication applications and remove the complexity of signaling. This ushers in 'context calling' from any browser app, where OSN (Online Social Network) become personal address-books, web shopping becomes 'yellow pages' of sellers, and web help page becomes a call center. In this multifaceted Voice paradigm, anything is callable, with instant connectivity. Since the ubiquity of the Internet is now sacrosanct, and its performance is soon to be enhanced by broadband 5G, the scene is set for an explosion of

context-based web calling services on a myriad of user devices, including cars, smart watches, and wearable devices. Web technologies are transforming the communication landscape yet again: first with social media, and now with real-time context-aware /social calling.

A tsunami of 'talking' web pages can be expected soon, where Voice is implemented as enhancements to existing online businesses, not as dedicated communication apps, but as additional website facilities. Businesses will provide context calling to engage with their customers and suppliers and get immediate results, sale or no-sale. To enable safe transaction completion, they will demand a reliable and confidential, yet affordable service, where Internet VoIP is not sufficient. Even in the consumer market, where fee is all important, users are concerned about protecting private data. Hence, a mid-way web calling service, pitched between VoLTE and Internet VoIP, will soon become a reality by sheer market forces. This market gap will be snapped up by web players, who are seeking ways to monetize 'free' Voice. Therefore, Telecom operators need to grab this opportunity very soon, accepting the new mode of business, and finding their own added value in this new formula.

Although it is simple to start real-time media, developing reliable business calling services requires much more effort and expertise. Many websites belong to entities that do not have the inclination, knowledge or resources to navigate through web Voice software. Such websites will also be keen to optimize the service delivery, upgrading calls to managed Mobile or Enterprise internal system, or demoting them to Internet VoIP, depending on the circumstances. Such websites may well be willing to pay for their connectivity and privacy, thus consumers still get a free service, but the websites will improve their transaction rate or their customers' experience. Operators hosting QoS enhanced web calling and switchable service mode would be able to offer a unique service that OTT web players cannot.

B. Managing QoS/Policy over the Internet

The commencing battle over Voice 'hegemony' is really a battle of QoE (Quality of Experience), where QoE encompasses policy, privacy and chargeability. VoLTE is delivered over managed networks with QoS guarantees, while Internet VoIP is only provided on the basis of 'best effort' and compromised privacy. However, this demarcation is eroding as web apps demand higher reliability and predictability, and even machine-to-machine services demand 'near-real-time' connectivity. As long as bandwidth-hungry applications continue to outstrip capacity, the quality and reliability of Internet VoIP will remain unpredictable. To this end, QoS-enhanced web calling service has been suggested in [18],

where the quality of the service delivery is determined by tunneled media via Internet media gateways that can enforce policy. These gateways control the media routing by TURN/STUN/ICE servers. The gateways plot a route to other networks with similar facilities, aiming to achieve end-to-end QoE improvements.

This QoS-enhanced webRTC (*QoSweb*) calling service is a hybrid, positioned between managed and unmanaged networks. The service cannot provide guarantees, as in VoLTE, it is not as reliable as Enterprise LAN, and it is not as secure as PMR, but it is better than OTT VoIP, and it can provide privacy for a fee. Therefore, QoSweb is now another service mode alternative, alongside OTT VoIP and managed networks.

C. Optimal Service Mode

It is debatable whether the notion of fee-for-quality and fee-for-privacy contradict the principles of ‘net neutrality’, which considers ‘two-speed Internet’ as a danger to equal opportunity and innovation. Since October 2015, an EU regulation [22][23] decrees support of uninhibited flow of Internet traffic, similarly to the current USA FCC’s rule, but it allows two significant exceptions: ‘class-based discrimination’ (by the type of service, e.g. emergency services); and ‘specialized services’ to Business for a fee. In fact, web calling services that allow full session interworking (unlike Over-The-Top ‘silos’) would serve the regulators’ goals even better than a blanket prohibition of QoS web. Inter-domain web calling will break the monopoly of large web players, allowing direct competition without service lock-in. Telecom players could also try their hand, alongside new web entrants, with flexible, global-reach, non-territorial, low-cost web services, which up till now were reserved only for unlicensed Internet developers.

The greatest appeal of web calling is to small business. This market segment desires more than only ‘best effort’ connectivity, but also reliability, predictability, security and privacy, at a reasonable cost. Web calling cannot match VoLTE or internal WLAN services in quality and predictability, since enhanced routing still does not match contractual guarantees and sophisticated network management, but it achieves better user experience than best-effort VoIP, and can offer privacy and policy-based delivery, with new business models and new types of bundling.

D. Benefits from Service Mode Selection

The service mode (Mobile, Enterprise, QoS-web calling, OTT) is no longer mandated by the access technology, the application or the device, because the same applications operate on any network and any device. As more service modes are made available, choosing between them becomes tricky, especially, where costs are balanced against reliability, flexibility against privacy, confidentiality against convenience. Organizations find it difficult to determine the optimal service mode, and even more difficult enforce it. Businesses need to protect their confidentiality as well as networking resources, balancing employees’ personal calls, which could be diverted to the Internet, against business calls which should be re-directed to a secure ‘line’ [9]. Service mode optimization could reduce SME costs significantly, by avoiding high roaming charges when low QoS would suffice, yet ensuring that the

appropriate levels of confidentiality and reliability are applied to important business calls.

For Mobile operators, switch service mode service is a trade-off between lost revenue and greater customer base, especially the new market of hosting communications for website context-calling. Operators who become web Calling Service Providers (wCSP) would gain subscribers’ loyalty, enticing them away from OTT, since they can get reasonably priced services on one hand, and value-for-money when necessary, on the other hand. The proposed selectable service mode is a chance for operators to *do more than web players* can – to switch between VoLTE and enhanced web calling. While this means that some VoLTE revenues are lost, operators will gain by entering the web arena, with global reach and global markets. The service mode selection will entice users away from OTT by offering greater affordability, while delivering improved QoS just when it is needed. The wCSPs could also offer user discovery, identity management, user privacy options, and more, thus beginning to serve the web SME market.

E. About this paper

The proposal in this paper is to create a mechanism for web Calling Service Providers (wCSPs) that enables them to select a service mode for each service request, regardless of the device, access network or application, but considering context and circumstances. A call may be best served by the suggested QoS-enhanced webRTC service (QoSweb), but could also be re-directed onto managed networks, or to no-guarantee VoIP services, if required. To determine the optimal service mode, the wCSPs needs to evaluate each service request, by profiling the call’s requirements of QoS (reliability, experience), Security (privacy, trust), Urgency (priority, criticality), or Affordability (no-fee, no-privacy, best-effort), i.e. QUSA. These QUSA profiles determine the best-fit service mode for each request.

The structure of the rest of the paper is: Section II: the state-of-the-art; Section III: Switchable service mode and QoSweb service architecture; Section IV, Defining service modes and call profiles; Section V: The procedure of selecting service mode from profiles; section VI: Summary.

II. STATE OF THE ART

The investigated areas include service request policy, always-best-connected, network re-selection, and modelling techniques. Several papers discuss optimizing network selection [8][9], but only as facilities for Network Operators to switch between transport carriers and off-loading Internet traffic. In [20], not only ‘best connected’, but also ‘best served’ are considered for the service mode, i.e. not only network status requirements, but it gives no solutions for gauging the context. A user-centric viewpoint is also taken in [20], and the enterprise choice is considered in [7]. In the author’s [11], Enterprise service request are profiled to determine business or leisure status, and the levels of risk (intruder, hacker, spy) are assessed, and in the author’s [12], the variability of credibility between sources of contextual evidence is incorporated in the attribute scores that utilize such sources.

Current standards of WebRTC do not allow for management of QoS. WebRTC signaling gateways are already appearing in various markets [17], providing protocols interworking to 3GPP Voice systems, but without the ability to determine policy for the web. To enable web apps to have selectable QoS levels, Internet traffic must be at least partly managed [6], as traffic should be routed via particular gateways that apply policy rules. Such gateways already exist to provide IPv4-IPv6 interworking, NAT (Network Address Translation) traversal etc. These gateways support TURN (Traversal Using Relays around NAT) as in RFC 5766, STUN (Session Traversal Utilities for NAT) as in RFC 5389, and ICE (Interactive Connectivity Establishment) as in RFC 5245. With ICE for offer/answer, negotiation, TURN for relaying traffic to enforce policy, and STUN to manage sessions and keep consistency, managing QoS communication is possible. The required QoS profiles can be negotiated by routers, utilizing well established protocols such as RSVP and Diffserve. The gateways can discover a path that matches the requirements and enforces policies. In this paper, the idea of QoS-enhanced web calling is examined further. Further study is still needed to standardize the initial cross-domain Internet based negotiation, and map the desired QoS/security parameters to the Internet gateways' enforcement capabilities, without encumbering the process with elaborate network management tools, as deployed for fully managed networks.

Context calling needs to glean information from concurrent activities, spatial-temporal factors and call history. In [1], the OpenSocial specifications will allow OSNs to interwork. This effort has now moved to W3C [4], with the charter to leverage social media: to increase autonomy over personal data; to enable ad-hoc communication circles; and to utilize social context for triggering actions, including context-calling.

The reTHINK¹ project is concerned with peer-to-peer webRTC session management. The framework devolves session processing to the endpoints, empowering users to maintain their own identity. This unilateral session control involves a single service provider, without cross-entity interaction, so the recipients' wishes are ignored. This paper makes the case for a 'phase 2' architecture that provides for a bi-lateral session setup, where QoS/policy requirements of *both* parties are considered, enabling this solution to work for business too.

III. THE PROPOSED ENVIRONMENT

A. The GENOME Choice

It is becoming increasingly difficult to determine the best service mode due to the convergence of access layers, the ubiquity of web apps, and the blurring of business and personal communications. Hence, a recommendation service that selects the best-fit service mode is proposed. This paper proposes a new GENOME: '*Good Enough Network Of Moving-Endpoints*', as in the author's blog [15]. The Genome service provides selectable '*good enough*' service, with a lower price tag than VoLTE, but still better than the '*best effort*' OTT services. Good-enough services have succeeded against best-

in-class before, where there was an attractive trade-off, e.g. Mobile services (ubiquity despite lower quality) won over Fixed Networks with 'five-nines' reliability. Similarly, introducing enhanced web by carriers may carve up a new market segment between VoIP and VoLTE, for users who look for an affordable service with appropriate privacy and confidentiality, and acceptable, 'just-right' quality.

The Genome approach bridges over the divide between Telecom and web. While VoLTE connects subscribers to other mobile operators' subscribers, VoIP web apps deliberately create 'silos' that are only accessible to their own registered users. By contrast, Genome applications will support an open and interoperable architecture that is not bound by federated agreements. While VoLTE allows for policy decision and policy enforcement, Internet Voice has no mechanisms for negotiating policy between the parties or enforcing them. By contrast, Genome applications enable negotiating policy, and the session set up respects both parties' preferences.

The notion of hosting web calling services is already practiced in the web world. Skype is offering its calling services to other web applications, for example to Facebook. Skype also encourages websites to offer clickable links that pass call information (at least the destination) from the website but initiate a connection on Skype. However, such hosting is based on proprietary interfaces, and does not allow ecosystem with fair competition. For Telecom operators to become web Calling Service Providers (wCSPs), they need flexible hosting platforms and simple APIs that companies can easily integrate in their website, as Skype does, but enable the interaction between different calling services, to maintain free choice.

B. Network Positioning of QoS-web Calling

Managed networks determine the media path by traffic management systems. They rely on traffic-exchange agreements between network providers, which are monitored, measured and accounted. By contrast, The Internet-style reciprocal traffic exchange has no underpinning contracts and guarantees, and no means of monitoring performance against such guarantees. Internet traffic is carried over numerous alternative and unpredictable paths, where the passage of traffic is reciprocal, with no guarantees. The remarkable resilience and QoS achieved by the Internet is due to this ability to self-determine alternative paths. Currently, Internet network providers do not accept policy requirements from each other (unlike Telecom operators), but they do manage QoS/policy within their own sectors.

The idea of tunneling through the Internet via TURN/STUN gateways detracts from the Internet resilience if fewer paths are available, i.e. paths without such gateways are excluded. Therefore, tunneled web routing requires universal participation of network providers in a loose collaboration agreement, which negotiates the path, but without explicit contracts, and without crippling inter-accounting and monitoring. To achieve better-than-best-effort, there must be large enough participating networks, where routing is no longer just 'best-effort', but it cannot be 'guaranteed' either.

Considering these limitations, it is suggested that the wCSP should orchestrate network delivery, including the appropriate

¹ The reTHINK project is European Union's Horizon 2020 research and innovation project.

access provider and network providers. The wCSP can specify the 'via gateway' parameter that incorporates partnering network providers for the media paths. However, to enable call completion, the traffic could still traverse non-compliant networks, if the enforcement of policy is not universal, and consequently, QoS bands guarantees cannot be given. QoS may be boosted by network providers, if they route tunneled traffic partly on their managed networks, to increase the quality for their local destinations.

The quality of VoLTE is superior to webRTC calling in travelling at high speed. Unlike VoLTE, WebRTC calling does not have cell management and cell handover, therefore it lacks the mechanism that provides smooth service continuation. This deficiency is somewhat alleviated at low speeds by WiFi handover standards (802.21) [2]. The management of IP address changes is improved by 'Dynamic DNS', through the performance is still too slow, thus web calling still requires a static location or only slow moving devices.

C. QoSweb Calling Session set up

To deliver enhanced services over the web, the service provider has to discover the destination user's location and to negotiate the policies for the connection parameters, according to the parties' requirements and preferences. **Figure 1** shows the process of setting up a session with service delivery negotiated between wCSP1 and wCSP2. Once agreed, the originating wCSP1 launches the call towards the destination user. If the call requires no QoS, path (1) is selected, over the open Internet. If certain levels of QoS and security are required, path (2) is selected, via collaborative networks that can enforce policy, assuming that only Internet service modes are considered. The delivery is entrusted to the first Network Provider in the path, who routes towards the destination network via the special policy enforcing gateways.

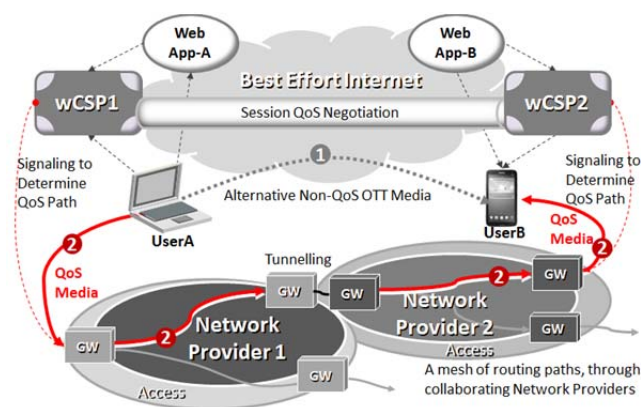


Figure 1. QoS Enhanced WebRTC Architecture

This architecture involves cross-domain wCSP to wCSP signaling negotiation, which is yet to be standardized. The session establishment procedure allows the recipient, *not only* the call originator, to have a say in the manner of the session delivery. For example an individual with no privacy preference may call the Police, where higher levels of confidentiality and reliability are demanded, so the session parameters must take

the called-party's preferences, not the caller's. In business-to-business or personal calls, the differences must be negotiated, to decide which ones prevail.

D. wCSP Collaboration

In current VoIP services, users choose the calling application, and the rendezvous is performed within the application. Only one VoIP application is involved, and it manages both communication 'legs'. The service mandates that *both* users are registered to the same calling service. This silo style model is favored by Internet VoIP due to its viral power, and is the reason why web players resist opening their APIs.

As a step further from the OTT approach, the reTHINK Project interworking is facilitated by dynamically downloading software onto the device of the non-subscribing parties, so that the caller's service provider can control the session, as shown in **Figure 2**. This does not impose the requirement to join, but forces a temporary subscription. The 'push compatibility' technique of downloading software to the other party assumes that called parties are willing to accept such downloads from those purporting to be web calling service providers. This will require a vigorous verification of callers' wCSP, who may be anywhere on the globe. This process can only be governed by one wCSP (the downloading side), and the called party's preferences and policies are not consulted at all. This *unilateral session setup* approach is quick-to-market, needing no standards before launching, but the QoE of non-subscribing recipients leaves much to be desired.

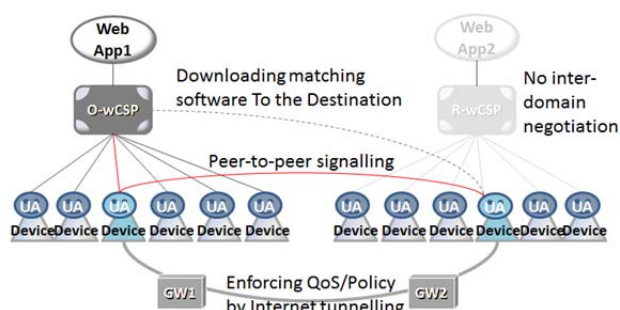


Figure 2. The unilateral model for session setup

QoS tunneling can still be implemented, but the session setup ignores the recipients' policy requirements, so it is not likely to be acceptable for business-to-business calling. With growing concerns of cyber attacks and cyber espionage, it is doubtful that the business sector will accept frequent software downloads that they have not tested, and allow external service providers to run their operations on the corporate servers or the employees' devices. Even trusted wCSPs increase device vulnerability, let alone unknown ones from anywhere in the world. Although the consumer market may accept the unilateral model (if the price is right), it is more than likely that this method will not be favored by the business world.

A solution which is more complex, but can apply to the enterprise market involves bilateral web calling setup, as in **Figure 3**. The bilateral architecture allows for full negotiation of QoS and policy between wCSPs. This requires new Internet standards, since webRTC currently lacks the mechanism to do

that. In the bilateral model, each leg of the call is managed within its own wCSP domain and its own user preferences, so that session parameters are mutually agreed.

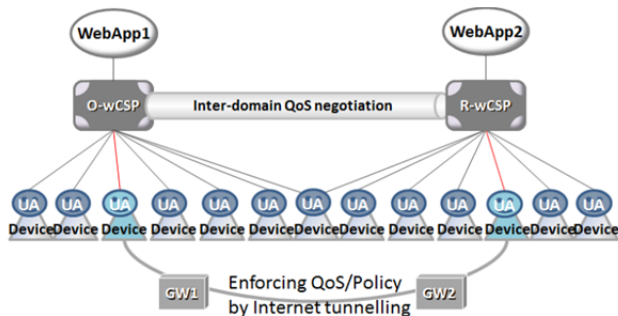


Figure 3. Bi-lateral model for session setup (consulting the Recipient)

This bilateral QoS/policy negotiation is still different from the Telecom traditional inter-entity dialogue, because only one service provider is in charge of orchestrating the delivery path. This relies on the Internet style routing, rather than network management tools that require full collaboration.

IV. DEFINING QUSA PROFILES

A. Service Modes Auto-Selection

The Genome service provides context-based choice of appropriate service mode which is just ‘good enough’ for the particular request of service. QoSweb calling provides mid-way option between VoIP and VoLTE, or between VoIP and Enterprise WLAN, with greater QoS and reliability than OTT services, but without the guarantees of VoLTE and the reliability of WLAN/LAN. The motivation to select a service mode is not just the Quality of Experience (QoE), but it is often based on price, or ‘Affordability’, which is a more subjective term. Decisions may be solely determined by Security, i.e. confidentiality and privacy of both data and call logs. Some public services are highly sensitive to the ‘Urgency’ rating, so that essential calls are certain to get the priority they need. Hence, each service mode has a number of characteristics that need to be matched to the service request type, user status and the overall circumstances, i.e. to the service context status. As mentioned above, this context is expressed by the relative QUSA levels of QoS, Urgency, Security, and Affordability.

The service mode selection considers the balance between all the QUSA profiles. For example, enterprises would prefer to optimize network usage by applying ‘force-on-net’ or ‘force-off-net’ [8], choosing the enterprise WLAN instead of Mobile to save costs but preserve confidentiality, or pushing personal heavy usage during work hours to the free Internet, when it requires no security. A consumer’s best option may well be Internet VoIP, if privacy is not necessary, but the same user may also be on duty for emergency services, where the call priority (Urgency) is paramount.

The service delivery modes are determined not just by QUSA profiles, but also by proximity to access nodes, membership in organizations’ calling systems and the ability to switch between service modes. Most users have access to Internet VoIP and consumer Mobile, and the QoS-enhanced

web calling is added to that. Employees also have access to enterprise calling services on internal WLAN/LAN or to PMR (Private Mobile Radio). Switching to alternative service modes depends on the wCSP arrangements, e.g. PMR and Enterprise. In addition, service availability also depends on the service ubiquity, i.e. being out-of-range.

In Figure 4, the modelled QUSA profiles (QoS, Urgency, Security and Affordability) are shown, with five modelled service modes, within managed and unmanaged networks.

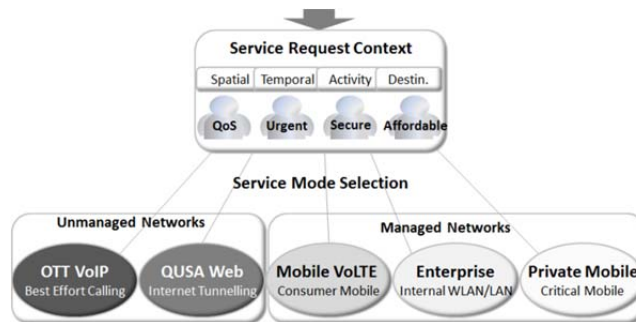


Figure 4. Service Modes selected from QUSA Profiles

B. Context Attributes from Heterogeneous Sources

To facilitate optimization of service mode, context information has to be collected. Traditional networks as well as OTT apps have only limited number of variables per call request, and setting up delivery parameters is governed mostly by network status. Now, much more information is on-tap about requests’ circumstances and user’s associated activities, which is generated by users’ intensified digital life. Information is gathered from context calling situations, social media, environment sensors and historical or concurrent behavior, without divulging private information. Large web players are already using such information. They exploit their huge repositories of user data, gained from searching engines, messaging and social media [15]. The new-style web calling service providers should use such knowledge strictly to enhance service delivery, and gain users’ trust by a responsible and transparent service, thus benefitting users with low cost services, while protecting their privacy.

Additionally, such context information is derived from sources of widely varying levels of credibility, and this must be factored in the computation. When computing QUSA profiles, some doubt exists in every category, and contradictory evidence may be encountered. Nevertheless, it is possible to derive the QUSA profiles by computing scores from contributing attributes, while considering the credibility of the data, and taking into account conflict as well as corroboration.

C. QUSA Profile Descriptions

The request context is profiled by the main requirements for the QUSA profiles, which vary request by request:

QoS (Quality, QoE and reliability) is an important differentiator between the best-effort Internet and QUSA-Web. While net neutrality dictates that QoS will not be ‘throttled’ [23], arguably, there is no barring to QoS being enhanced. Required QoS depend on business or leisure, and the requested

media, and the necessary response time varies in certain applications more than others.

Urgency (criticality and call priority) is a requirement for essential and time-critical connections, usually for business callers, but also for emergency services. Urgency is detected by destinations, but also by intense patterns of initiating service requests. Call priority is best achieved in dedicated networks (Enterprise, PMR).

Security (confidentiality and privacy) is judged by the need to protect against intruders and hackers, and protect business confidentiality as well as user privacy. This is an important differentiator for SMEs and ‘super-users’, but is not necessary in personal casual interactions.

Affordability represents the requirement of cost constraints, but more often, the preference for the trade-off, favoring no-fee service in returns for accepting no-privacy and only ‘best-effort’ QoS. Affordability requirements may be oblique, i.e. no need for QoS and security, but it is also evident from the type of destination (e.g. casual surfing) and user circumstances (on holiday). It is also determinable from user preference tables, indicating personal and business policies.

To profile these characteristics, they are quantified by computed QUSA levels, which are interpreted for each service mode. The profiles represent the requirements, while the service modes represent expected achievable levels. To determine the most suitable service mode, the request profiles are computed first. These profile levels are weighted by a particular mix of impact rates that differentiates service modes from each other.

V. SERVICE MODE SELECTION PROCESS

A. QUSA Profiling Procedures

The values of the QUSA profiles are derived from context attributes that describe the request status and the user situation. Such information is gathered from attributes drawn from the request details, observed spatial and temporal aspects and their inferences, and from what is known about the user’s status and activities. Data sources vary greatly in their trustiness and faithfulness, so the credibility of the sources should be factored in the assessment of the request. In addition, the Intensity of the observations is measured, e.g. distance for ‘remote location’ or level of security for a ‘confidential’ attribute.

The attribute definitions depend on available information, which varies between wCSPs, i.e. mobile operators, enterprises or web players. Enterprises have better view of employees’ associated activities; web player have better surfing patterns; and Mobile operators monitor better travelling and roaming. Such data may be considered private, but it is only used internally to improve the service for the user, while the privacy policy ensures that user data is not commercially exploited.

Calculating profile scores starts by accumulating the credibility rates of the relevant sources that contribute to an observation. The attributes scores are computed in step (1) from their valid observations and their observed intensities, where several observations may be combined. The QUSA profiles are computed in step (2) from the attributes’ scores, where attributes are prioritized per profile. In step (3), the

service modes scores are produced from the QUSA profiles, which are weighted according to the service mode templates. As shown in **Figure 5**, each service mode interprets the QUSA profiles differently, e.g. Affordability is paramount for OTT VoIP but not for PMR, and Urgency is judged higher in Mobile than in QUSA Web service.

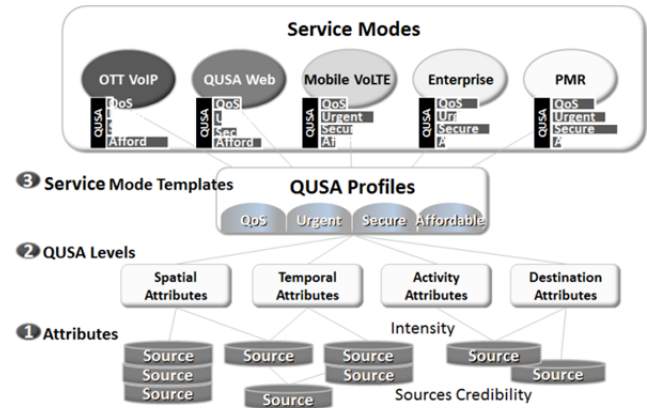


Figure 5. Procedure for computing Service Mode

B. Computing Context Attributes

Context attributes for the service request vary by service provider access to various records, log files, historical archives, location information, and so on. It is assumed that the service is fully supported by the enterprise or SME, in order to maximize the benefits, therefore the calling service will have access to employee information and corporate data, beyond the immediate service request. Common data sources are corporate directories and calendar (general holidays and personal holidays), GPS location, apps login, WLAN login and so on. The credibility level of each source is computed from their digital characteristics, as shown in **Table 1**.

Table 1: Computing Credibility Ratios of Sources

Sources		DPI	GPS	Corp. App	Dir P2P	Destin Type	Favor	Calendar	Login App	Recent log	Pattern
Credibility	Reliability	0.131	0.013	0.074	0.039	0.019	0.007	0.064	0.101	0.036	0.027
	Stability	0.000	0.000	0.060	0.034	0.016	0.006	0.000	0.023	0.001	0.012
	Timeliness	0.037	0.008	0.037	0.015	0.017	0.008	0.076	0.114	0.130	0.021
Accuracy	Fidelity	0.045	0.030	0.002	0.035	0.006	0.029	0.021	0.000	0.015	0.026
	Robust	0.035	0.022	0.076	0.010	0.004	0.020	0.010	0.024	0.033	0.052
	Error	0.013	0.000	0.002	0.022	0.013	0.001	0.001	0.000	0.002	0.002
Precision	Resolution	0.008	0.095	0.000	0.000	0.000	0.000	0.015	0.000	0.000	0.030
	Spatial/Time	0.000	0.069	0.037	0.000	0.000	0.000	0.021	0.026	0.001	0.032
	Range	0.005	0.062	0.000	0.000	0.000	0.000	0.009	0.000	0.001	0.020

The characteristics of the sources break down to confidence (reliability, stability, timeliness), Accuracy (fidelity, robustness, error management), and Precision (resolution, spatio/temporal proximity, and relevant range). Each characteristic is broken down further to a measurable index or a detailed obtainable estimate [12], e.g. equipment time-between-failure is a measure of reliability, data timeliness or ‘freshness’ is indexed between 1-10 according to the update

cycles etc. In this way, sources credibility rate is a starting point for the attribute score. Actual observation readings from the sources are valued by their levels of intensity, rated by their credibility. An attribute may contain several aggregated observations, each computed by their own sources credibility and their own observed intensity.

For the purpose of a generic model, four areas of attributes have been identified: Spatial factor, Temporal factor, Activity and Destination. As a collaborative service, it is assumed that the service provider can associate various activities with a specific user identity, including user-provided qualifying tables that are used to interpret the readings, e.g. tagged locations ('home', 'office'), business or personal destinations (applications or contact names/numbers), type of destinations (confidential, low-priority) and more. **Figure 6** shows an example of a computed specific service request (record 201).

201	Sources	For Service Mode:											Weighting Table				Aggregation Table					
		QoSweb	Acc Data	OSN Data	GPS	Mobi le	WiFi	Corp App	Dir	Type	email Login	Calendar	Login App	Hist.	Pattern	QoS	Urgent	Secure	Afford	QoS	Urgent	Secure
Credibility		0.99	0.64	0.71	0.87	0.39	0.58	0.29	0.03	0.01	0.42	0.69	0.49	0.48					0.142	0.005	0.021	0.077
Spatial																			0.090	0.051	0.043	0.071
Temporal			0.54																0.084	0.010	0.040	0.017
Activity		0.98								0.93									0.009	0.009	0.014	0.020
Destination		0.80																	0.007	0.006	0.010	0.001
Totals													0.009	0.045	0.014	0.001						

Figure 6. Example of QUSA calculation for a service request

The factors are shown in the 1st column, and attributes in 2nd column. The data sources (2nd row) and their assessed credibility (3rd row) are combined with the intensity (values in shaded areas). Attributes may have one or more readings, but not all attributes are relevant, and of those, not all are observed. The QUSA scores are computed by weighting the observations according to their impact within each of the profiles (Weighting Table), then aggregating the profile total (Aggregation Table). The resulting QUSA levels are given in the 3rd row of the aggregation table (far right).

C. Selecting the Service Mode

Having produced the QUSA profile scores, the service mode can be established. This requires customizable impact ratios, as in **Table 2a**, that determine the effect of each profiled QUSA requirement on the service mode. For example, Affordability impact on the OTT service mode is high, but Urgency and Security are low. For selecting QoSweb, both QoS and Affordability are high, while Mobile will be chosen if Affordability is low (i.e. not important), but Urgency is high. The Enterprise service mode will be selected where a reliable and high quality service is paramount, while PMR favors Security above all.

Table 2b contains five cases, with five sets of input (in the shaded 3rd column) from the QUSA model, where profile scores per case are computed for QoS, Urgency, Security and Affordability. The results of case 201 are those shown in Figure 6 3rd row, (Aggregation Table). These values are

weighted by the Service Mode impact ratios (above), to account for the different influence they have on the resulting service mode score. Since not all service mode are available even within one Genome service, the unavailable service modes are neutralized when the availability indicator = 0 (1st row per case) so that only available services are considered, e.g. case 201 has no PMR and case 205 is out-of-range for the Enterprise WLAN. Finally, the weighted profile scores per service mode are totaled per case, revealing the highest service mode score (highlighted).

Table 2a: Service Mode Impact Rates							
Service Mode	Service Modes:		QoSweb	OTT	Enterprise	Mobile	PMR
	Profiles		Impact Rates				
Service Mode	Quality:		0.39	0.12	0.42	0.20	0.22
	Urgency:		0.08	0.00	0.18	0.50	0.33
	Security:		0.12	0.05	0.33	0.30	0.45
	Affording:		0.41	0.83	0.07	0.00	0.00
Table 2b: Computing Service Mode from QUSA Profiles							
Case201: Availability			1	1	1	1	0
QoSweb	Q	0.142	0.056	0.017	0.060	0.028	
	U	0.005	0.000	0.000	0.001	0.002	
	S	0.021	0.003	0.001	0.007	0.006	
	A	0.077	0.032	0.064	0.005	0.000	
	1 Totals		0.088	0.081	0.072	0.037	0.000
Case202: Availability			1	1	1	1	1
PMR	Q	0.212	0.083	0.025	0.089	0.042	0.047
	U	0.253	0.020	0.000	0.045	0.126	0.083
	S	0.275	0.033	0.014	0.091	0.082	0.124
	A	0.009	0.004	0.007	0.001	0.000	0.000
5 Totals		0.134	0.046	0.210	0.232	0.234	
Case 203: Availability			1	1	1	1	1
Enterprise	Q	0.181	0.071	0.022	0.076	0.036	0.040
	U	0.070	0.006	0.000	0.013	0.035	0.023
	S	0.152	0.018	0.008	0.050	0.046	0.068
	A	0.066	0.027	0.054	0.005	0.000	0.000
3 Totals		0.117	0.082	0.138	0.113	0.126	
Case 204: Availability			1	1	1	1	1
Mobile	Q	0.239	0.093	0.029	0.100	0.048	0.053
	U	0.324	0.026	0.000	0.058	0.162	0.107
	S	0.272	0.033	0.014	0.090	0.082	0.122
	A	0.034	0.014	0.028	0.002	0.000	0.000
4 Totals		0.157	0.069	0.231	0.267	0.257	
Case205: Availability			1	1	0	1	1
OTT	Q	0.110	0.043	0.013		0.022	0.024
	U	0.076	0.006	0.000		0.038	0.025
	S	0.118	0.014	0.006		0.035	0.053
	A	0.229	0.094	0.190		0.000	0.000
2 Totals		0.150	0.206	0.000	0.092	0.099	

D. Simulation

The determination of QUSA profile has been simulated by 200 elaborated cases, each with a different mix of observations and sources. Several algorithms have been used to gauge relative accuracy, but there is no available real-world data to underpin this work.

E. Feasibility Discussion

The ability and willingness to determine the optimal service mode by context and execute the mode re-selection varies depending on the type of wCSP: mobile operator, web player or even an enterprise. Operators can offer more versatile connectivity, including VoLTE and PMR in their range, but may not be in possession of full employees' information for accurate profiling. Web player can switch between best-effort

and QoSweb, but may have difficulty in exercising force-on-net and force-off-net.

The accuracy of the QUSA profiles is key to the success of the selectable service mode. Clearly, wider range of sources can corroborate observations and add confidence, but collecting observations requires some operational effort. To examine how different deployment may fair, each of the data sources was eliminated from the scoring process, except those which are deemed to be always available. **Table 3** shows the case for a mobile operator, with always available knowledge of the subscription account details, GPS, mobile and WiFi access, and the service-generated historical logs (in blue). If OSN information is not collected, no change occurred in the given sample (green). Likewise, user directories, diaries, destination types and email login did not impact the outcome, in isolation. However, logging to corporate applications and other apps seems to be significant. Note that this does not mean that certain sources are redundant, and that the table does not show the cumulative effects when several sources are not available at the same time.

Account	OSN	GPS	Mobile	WiFi	Corp. App	Dir	Destin Type	email Login	Calendar	App Login	Recent	Pattern	Diary	Work Hrs	Event
Blue	Green	Blue	Blue	Blue	Orange X	Green	Green	Green	Blue	Orange X	Blue	Blue	Green	Blue	Green

VI. SUMMARY

In this paper, an enhanced web calling service architecture (QoSweb) is proposed, with a service that selects the best service mode, the Genome. QoSweb is positioned as mid-way between best-effort OTT VoIP and managed networks, i.e. VoLTE, Enterprise WLAN/LAN and private mobile (PMR). As services are now access agnostic and operate on multitude of devices, the choice of the service mode is no longer predicated on device, but can be switched by instructing the device client to relaunch the request on an alternative service mode. This opens up a rare opportunity for Telecom operators to become web Calling Service Providers, and offer optimal service mode selection, which is just 'good-enough' for the purpose. Operators could host web calling on behalf of numerous websites of entities that have neither the skills nor the inclination to get involved in Voice communications in returns for losing some VoLTE revenues in the SME market, who may otherwise defect to OTT solutions anyway.

The service mode is optimized according to users' circumstances and requirements, based on analysis of four categories (QoS, Urgency, Security, and Affordability), from attributes that describe the request and the user context in terms of spatial-temporal status, destinations and logged user activities. The model computes scores per service request, taking into account the variable credibility of the resources. These QUSA characteristics are then used to determine the optimal service mode for the request.

ACKNOWLEDGMENT

The author wishes to acknowledge the reTHINK projects which inspired this study. This work has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 645342, project reTHINK. Special thanks are also due to Stephane Tuffin from Orange Labs, France, for his insightful comments and for introducing me to the idea of QoSweb calling service.

REFERENCES

- [1] Social Open Federation <https://www.w3.org/blog/2014/12/opensocial-foundation-moves-standards-work-to-w3c-social-web-activity/>
- [2] [23] H Marques, J Ribeiro, P Marques, J Rodriguez "Simulation of 802.21 Handovers Using ns-2" Journal of Computer Systems, Networks, and Communication.
- [3] Belton V, Stewart T "Multiple Criteria Decision Analysis – An Integrated Approach" Springer Science & Business Media 2001
- [4] W3C Social Web <http://www.w3.org/2013/socialweb/social-wg-charter>
- [5] <http://telecoms.com/417622/google-mvno-to-launch-this-week-report/>
- [6] S Bécot, E Bertin, J-M Crom, V Frey, S Tuffin "Communication Services in the Web Era" ICIN 2013
- [7] Canan'ea, I., Mariz, D., Kelner, J., Sadok, D "An On-line Access Selection Algorithm for ABC Networks Supporting Elastic Services" IEEE WCNC 2008.
- [8] Chen Y, Deng C, Yang Y "Access Discovery in Always Best Connected Networks" Knowledge Acquisition and Modeling Workshop, 2008.
- [9] Copeland R, Crespi N "Selecting Access Network for BYOD enterprises with Business Context (eBCR) and enterprise-centric ANDSF" Mobilware, Springer 2012.
- [10] Copeland R, Crespi N "Classifying and Aggregating Context Attributes for Business Service Requests- No One-Size-Fits-All" SCC, IEEE 2014.
- [11] Copeland R, Crespi N "Identifying Risk Profiles and Mitigating Actions for Business Communication Services" CNSM, IFIP/ IEEE/ ACM 2014.
- [12] Copeland R, Crespi N "Rating Credibility of Sources for Profiling Risk and Business Context of Service Requests" SSC IEEE 2014.
- [13] Coyle G "The analytic hierarchy process (AHP)" Pearson 2004.
- [14] Kellokoski, J., Hamalainen, T "User-Centric Approach to Always-Best-Connected Networks" IEEE ICUMT 2011.
- [15] <http://www.connect-world.com/index.php/component/acymailing/archive/view/listid-1-connect-world-eletter/mailid-388-the-new-genome-good-enough-network-of-moving-endpoints>
- [16] <http://precursorblog.com/?q=content/google-buys-jibe-force-android-web%E2%80%99s-default-means-communications>
- [17] <http://www.sonus.net/en/resources/press-releases/sonus-introduces-virtualized-webrtc-solution>
- [18] Janczukowicz E, Tuffin S, Braud A, Bouabdallah A, Fromentoux G, Bonnin J M "Approaches for Offering QoS and Specialized Traffic Treatment for WebRTC" Advances in Communication Networking, pg 59-69 Springer 2014
- [19] Ji, Z, Ganchev, I, O'Droma, M "An iWBC Consumer Application for 'Always Best Connected and Best Served': Design and Implementation" IEEE Transactions on Consumer Electronics 2011.
- [20] Kellokoski, J., Hamalainen, T "User-Centric Approach to Always-Best-Connected Networks" IEEE ICUMT 2011
- [21] 3GPP TS 23.203 V9.9.0 (2011-06) Policy and charging control architecture (Release 9)
- [22] <https://ec.europa.eu/digital-agenda/en/eu-actions>
- [23] BEREC "Guidelines for quality of service in the scope of net neutrality" 2012.