

# CONTEXT-AWARE MULTIMEDIA SERVICE COMPOSITION USING QUALITY ASSESSMENT

Alberto J. Gonzalez, Jesus Alcober  
Technical University of Catalonia/i2CAT Foundation  
Department of Telematic Engineering  
Esteve Terradas 7, 08860, Castelldefels, Spain  
{alberto.jose.gonzalez, jesus.alcober}@upc.edu

Ramon Martin de Pozuelo, Francesc Pinyol  
La Salle-Ramon Llull University  
GTM - Media Technologies Research Group  
Quatre Camins 2, 08022, Barcelona, Spain  
{ramonmdp, fpinyol}@salle.url.edu

Kayhan Zrar Ghafoor  
Universiti Teknologi Malaysia  
Faculty of Computer Sc. and Inf. Systems  
81310 Skudai, Johor D. T, Malaysia  
zgkayhan2@live.utm.my

## ABSTRACT

With the proliferation of multimedia capable devices, media services have to deal with heterogeneous environments where very different types of terminals wish to receive content anywhere and anytime. This situation motivates the appearance of multimedia services that adapt contents to the specific context of users. However, current Internet architecture is based on a rigid layered model which makes difficult to introduce new functionalities efficiently. To solve this, Service Oriented Architectures (SOA) appear with the goal of proposing new architectures based on services that can be invoked when and where necessary. This work introduces how SOA paradigm can be applied to context-aware multimedia communications. In addition, a scoring function for selecting different service implementations is presented and particularized for a case of selecting transcoding functions taking into account different quality assessment metrics.

**Index Terms**— Quality assessment, context-awareness, service composition, multimedia

## 1. INTRODUCTION

Multimedia content in the network is growing daily. There are emerging services which offer contents in multiple ways and different formats. Popularity of video services such as Youtube or Hulu, have made video traffic in the Internet the most present one. Moreover, the heterogeneity of devices connected to the network is also rising (e.g. handheld, PC, TV, etc.) and it usually requires the creation or the adaptation of services and resources specifically for each target platform. This situation leads to generic and static systems that do not provide the content adapted to the final device or too complex systems that require a big effort in development and maintenance tasks.

---

This work was supported by the TARIFA project of the i2CAT Foundation, the CENIT Programme (Spanish Ministry of Industry) under the i3media project and by the Spanish Government (MICINN) under research grant TIN2010-20136-C03. The authors would like to thank all participants from the i2CAT Foundation and Telefonica R&D for their help and support. Specially thanks to the GTM group of the Ramon Llull University and the BAMPLA group of the Technical University of Catalonia.

In this scenario, the necessity of context-aware systems arises. Their goal is to offer services adapted to the context of users (e.g. device capabilities, network context, user preferences). These services allow the maximization of the Quality of Service (QoS) and Quality of Experience (QoE) of the user, while allowing a more efficient usage of resources. However, their deployment requires precise and efficient monitoring and data management systems.

One promising approach to efficiently provide context-aware services is using Service Oriented Architectures (SOA), which divides Services in simpler ones and couples only those that are required or preferred for a specific context. Solutions based on this type of architectures provide some clear benefits: loose coupling, implementation neutrality, flexible configurability, granularity, task distribution, energy efficiency, efficient use of resources, etc.

The division of services could be done based on different aspects such as location, capabilities, functionalities, etc. In this paper we follow a role-based decomposition approach [1], and we divide complex services into indivisible or atomic functionalities. Examples of these functionalities are: encoding, acknowledgment or retransmission. Following this approach, in [2] authors call these functionalities Atomic Services (ASs), and specify that each of these services could be offered by different specific implementations, called Atomic Mechanisms (AMs). This separation between service definition and implementation should facilitate a loose coupling among services and a more flexible creation and reuse of complex services through the network. In order to overcome the limitations of current Internet TCP/IP layered stack [3], these principles are used by several Future Internet proposals aiming to define novel architectures in a clean slate manner.

Additionally, no matter how the decomposition is done, there are many approaches or visions of doing the composition process. However, there are some points to be taken into account and some common problems that appear in most of current techniques:

- Considering services as self-contained, self-describing, modular applications that can be published, located, and invoked across the network, service composition process can be defined as the combination of those ser-

vices required to create new processes and services.

- Fulfillment of preconditions when a service that can provide the desired effects exists [4].
- Generation of several effects. It is possible that a service request is associated to multiple effects that can be satisfied by different services.
- Knowledge and context data acquisition and management.

These problems denote a close relationship between optimal compositions and context-awareness. Thus, how to obtain and analyze the context is an important issue, in order to provide a good ground for the service composition process. In this framework, we review different methods that could be applied to evaluate the quality of multimedia services and how to apply multimedia quality assessment to enhance a multimedia service composition process.

Context-awareness features are especially relevant, as the ubiquity of mobile devices and the proliferation of wireless networks is enabling permanent access to the Internet at all times and all places. The next step to an Internet of Services is an Internet of context-aware Services.

The rest of this paper is organized as follows. In Section 2, authors provide background and related work, including a review of context-awareness, service composition approaches and multimedia quality assessment methods. Then, Section 3 describes the proposal for service composition and justifies the adopted principles. Section 4 details the multimedia quality assessment process used for scoring multimedia services. In Section 5 we present a proof of concept of the scoring function usage for a media transcoding service use case. Section 6 summarizes the results obtained from Section 5. Finally, Section 7 presents the conclusions.

## 2. RELATED WORK

According to the definition provided by Dey in [5], context is any information that can be used to characterize the situation of an entity, where an entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.

Context-awareness refers to the capability of an application or service to be aware of its physical environment or situation and responding intelligently (pro-actively or reactively) based on such context. So, it is important to compose services and dynamically adapt them according to the context information and changes in order to provide personalized and customized services to users. This should allow improving the QoS and QoE of users while optimizing the usage of network and computational resources.

Revising literature related to context-awareness [6][7], we define a consensus classification of the context according (but not limited) to the following:

- User context: user characteristics, user location, user preferences, and environmental constraints of the user (e.g. working place, home, etc.).
- Device context: type and capability of the device.
- Service context: service availability, minimum required QoS level for providing the service, and additional parameters that define specific attributes for a service.
- System resource context: CPU, memory, processor, disk, I/O devices, and storage.
- Network context: bandwidth, traffic, topology, and other parameters related to network performance.

On the subject of service composition, several approaches can be found in literature tackling service composition [8][9][10][11]. For example, in [10] authors propose a classification system in the form of taxonomy, for semantic web service composition approaches, that could be generalized to the global concept of service composition, and applied to compose services at network level.

Unlike other solutions, our approach uses the discovery process to find those services to be composed all along the end-to-end path attending to the requirements specified by the requester and, consequently, assuring a certain level of QoS.

Regarding to the quality assessment process, an overview on different metrics and techniques is presented.

Multimedia quality metrics can be classified into three groups taking into account in which way a reference signal is needed to measure the quality: (1) Full Reference (FR), which compares two entire signals, a reference signal (usually the original one) and a compared signal (usually the coded one); (2) Reduced Reference (RR), which only compares some signal characteristics (blocking, blurring, ringing, masking, etc.) previously detected; and (3) Non Reference (NR) metrics, which does not use any reference signal to determine the quality of a signal. Each of them is used depending on the availability of an undistorted signal (reference signal). The most used metrics are FR (e.g. PSNR), due to its low complexity, but as a drawback, both signals are needed: the original signal and the coded signal.

Additionally, methods for quality assessment can also be divided into two categories: a) objective and b) subjective. Objective methods aim to mathematically estimate the impairment introduced to media resources during compression or transmission whilst the Subjective ones aid in the statistical analysis of sample ratings generated by humans. In this work we used different objective quality metrics for estimating the quality of a received media resource. For video we used Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity (SSIM) [12]), while PSNR, audio SSIM [13] and Perceived Audio Quality (PEAQ) [14] were used for audio quality assessment.

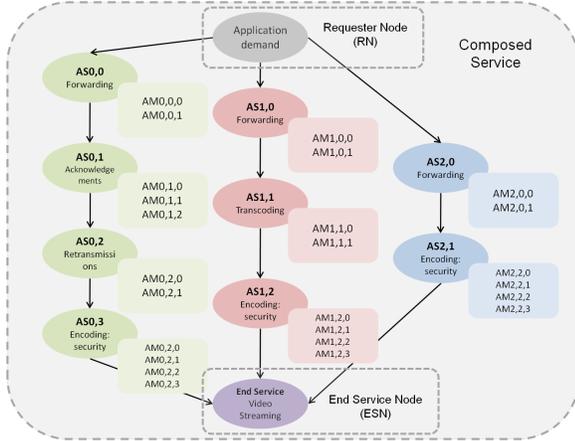


Fig. 1. Service Composition

### 3. COMPOSITION PROCESS

Taking into account a service framework as presented in [2] we define Composed Services (CSs) as a workflow of functionalities or Atomic Services (ASs) that could be implemented by different Atomic Mechanisms (AMs). It can be seen in detail in Figure 1. The composition process of these ASs consists on the process of selecting, allocating and combining those services to be executed along the path from Requester Node (RN) to the End Service Node (ESN). In this context, we propose a composition process orchestrated by the RN, to empower the consumer control of this process. RN will always decide which services to choose according to the discovered ones. Ideally, selection and allocation decision is done taking into account the cost of using it, with regard to RN priorities and requirements.

We divide the composition process into the following stages: Filtering, Composing ASs and Scoring AMs. *a) Filtering:* This phase consists of filtering according to the requirements specified by the RN. This process is done at each node in the path from RN to ESN in order to propose to the RN the best services for the required communication. Moreover, when RN receives all the possible responses to the service discovery request it should validate if answers fulfill the specified requirements. Concretely, for a multimedia communication, this filtering phase can be done at the server side, taking into account the specific capabilities of the user which desires to visualize a streamed content. For example, considering the supported audio/video profiles that the server is able to generate and the client features, the best profile must be selected based on the end user context information such as network (e.g. bandwidth) and terminal capabilities (e.g. display resolution). *b) Composing ASs:* RN composes the services per each node (intermediate and end nodes). As seen in [10], several approaches can be adopted for service composition. Defining which are the most suitable for each case is out of the scope of this paper. However, it would be interest-

ing to propose benchmarks and comparisons of composition algorithms and techniques, in order to determine which are the best under specific conditions. *c) Scoring AMs:* In this phase the concrete AM that implements each AS is selected according to specific scoring functions, which take into account the QoS parameters and effects that they can provide and the priorities of RN. Several compositions can be produced to perform the same operation, so that the best suited to request priorities will be chosen. For instance, a reliable service can be provided by means of acknowledgments, error detection and retransmissions, or by applying forward error correction functions. Depending on the combinations, QoS may vary. Finally, those AMs that scored best will be selected and incorporated into a final composition. Section 4 describes how AMs can be scored taking into account different parameters. Concretely, we particularize this problem to the selection of the best codec to be used in a communication taking into account quality assessment metrics. However, it is just a first proof of concept of the framework explained until now, which is still being developed.

### 4. MULTIMEDIA QUALITY ASSESSMENT

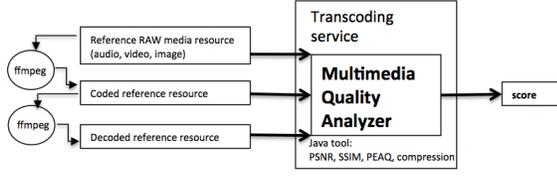
As said in Section 3, AMs need to be selected according to several parameters such as performance, quality of service or, in case of multimedia applications and services, parameters such as the perceptual quality. Quality assessment can be used for measuring the quality of multimedia communications. The goal is to select the best AM for each communication, and best means that it can provide the highest possible perceptual quality. This section proposes to use quality metrics for deciding which is the best AM to use when an AS of the type transcoding is used. This AS mainly consists on adapting a content taking into account the context of the requester. We introduce a scoring function that uses the measured objective quality and the compression ratio provided by different codecs. However, other parameters can be added, such as performance ones (e.g. CPU, energy consumption).

#### 4.1. Multimedia Quality Analyzer

We have developed a multimedia quality analyzer module (Figure 2) that calculates a score for each codec supported by the multimedia transcoding service. In this case, each codec corresponds to an AM, implementations of the AS named "transcoding". We use a FR system that can use the metrics defined in Table 1 for determining the obtained quality.

Media	Metrics		
Image	PSNR	SSIM	-
Video	PSNR	SSIM	-
Audio	PSNR	SSIM	PEAQ

Table 1. Full Reference metrics



**Fig. 2.** Quality Analyzer module

PSNR is an objective quality metric used to calculate the ratio between the maximum possible power of a signal (in this case an audio or video stream) and the power of the corrupting noise. It is commonly used to calculate the effect of losses in a video or audio signal. SSIM is a metric for calculating the similarity between two images, which lies in the assumption that human visual perception is highly adapted for extracting structural information from a scene. Its application to audio measurement is still being studied. Finally, PEAQ is a standardized algorithm for objectively measuring the perceived audio quality.

The inputs of the quality analyzer module are: (1) a media resource (image, video or audio) in raw format, (2) the same media resource coded with a supported codec, (3) the same media resource decoded to raw format.

(1) and (2) are used to evaluate the compression ratio and to obtain the file with losses due to the effect of coding. 3) is used to compare the resulting resource with the original one in raw format (input of the multimedia analyzer) and to measure the differences and impairments. In our system, this process is performed offline, when the system starts or a new codec is added to the system. Then, the system performs all the analyses and stores the results (scores) into a table which is looked up when necessary by the decision-making algorithm.

#### 4.2. Score parameter definition

The use of lossy codecs allows the compression of the resource size. But, intrinsically, it also reduces the user quality perception. So, it must be found a trade-off of the compression ratio and the perceptual quality.

A way to decide which codec is better than another is to consider the perceptual quality and the compression ratio of a coded media resource. Thus, it can be said that a codec is better than other if this presents a better perceptual quality and compression ratio relationship. This can be expressed according to:

$$score = A * perceptual\_quality + (1 - A) * compression\_ratio, \\ where 0 \leq A \leq 1$$

A is a weight, which determines the relevance of each parameter considered in the scoring function. Hence, the relevance of each parameter can be changed. The weight that

specifies an equitable relationship between perceptual quality and compression ratio is obtained for  $A = 0.5$ .

The score parameter is defined in the  $\mathbb{R}$  set and can take values from -1 to 1:

$$score \in \mathbb{R}, -1 \leq score \leq 1$$

where -1 and 1 indicates respectively the worst and the best perceptual quality and compression ratio relation.

The compression ratio parameter is defined in the  $\mathbb{R}$  set and it can also take values between -1 and 1:

$$compression\_ratio \in \mathbb{R}, \\ -1 \leq compression\_ratio \leq 1$$

where -1 indicate that there is no compression between the original and the coded resource, but there has been an increment in the total number of bits, and a positive value (less than 1) indicates a reduction of the total number of bits.

The compression ratio parameter mathematical expression is defined in (1).

$$compression\_ratio = \\ \frac{(original\_resource\_num\_of\_bits - coded\_resource\_num\_of\_bits)}{(original\_resource\_number\_of\_bits)} \quad (1)$$

The perceptual quality parameter can take values between 0 and 1 and it is also defined in the  $\mathbb{R}$  set:

$$quality \in \mathbb{R}, \\ 0 \leq quality \leq 1$$

where 0 indicates, in perceptual quality terms defined by ITU-R in [15], very annoying perceptual quality, and 1 indicates no difference between the original and coded resource.

Some of the considered quality metrics do not take values between the defined ranges. So, they must be normalized. The quality metrics to be normalized are:

$$0 \leq PSNR \leq \infty, \\ -4 \leq PEAQ \leq 0$$

It is not necessary to normalize the SSIM quality metric as its output range fits into the perceptual quality parameter range. More details are shown in [16].

#### 4.3. Combining audio and video

The scoring of audiovisual contents should take into account the relationship between audio and video, not only considering their individual scores in an independent manner. The goal is to avoid bad combinations of audio and video profiles, for instance when obtaining combined profiles with very good audio and very poor video (or viceversa).

$$score(audioQ, videoQ) = (score_A + score_V) - \\ \frac{|(A * score_V) - (B * score_A)|}{\sqrt{A^2 + B^2}} \quad (2)$$

In (2) we define a combined score as the sum of individual qualities and the subtraction of each point (videoQ and audioQ) to the line defined by the optimal quality ( $Ax - By = 0$ ). A and B coefficients are defined by the R parameter (3). This parameter can be introduced by default (system administrator) or by the user as a preference.

$$A = 0.5 * B = R, R \in [0, 0.5] \\ A = (1 - R) * B = 0.5, R \in [0.5, 1] \\ R \in [0, 1] \quad (3)$$

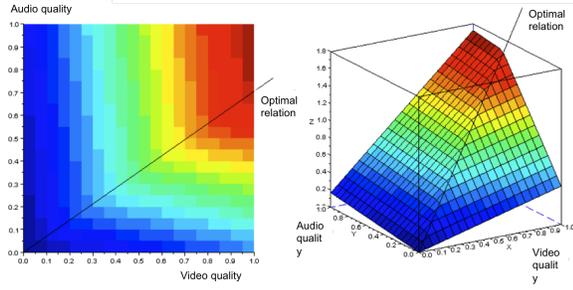


Fig. 3. Score combination

Figure 3 shows how the quality function looks like for an example value of  $R=0.66$ , which corresponds to a 4:3 relation between audio and video. Thus, we give a little bit more priority to video than to audio. However, this relation can be tweaked according to user preferences.

## 5. PROOF OF CONCEPT

We have implemented a quality assessment tool (Multimedia Quality Analyzer) in Java that is able to perform the calculation of the score parameter defined in Section 4 in order to verify its behavior and to test different common codecs against network loss effect. The quality assessment tool is able to use PSNR, SSIM and PEAQ metrics. The testbed scenario is composed of three basic elements: (1) streaming media server, (2) streaming client where the media resource is analyzed, and (3) a controlled network over which losses are introduced.

The media resource server acts as a video streaming server. It has been used the FFmpeg transcoding software (libavcodec 52.10.0)[17]. This transcoder allows transcoding multimedia resources with a wide range of supported codecs. FFmpeg also supports streaming over a network interface. In our case, we used UDP/RTP to stream the content over the network and to notice the loss effect. In order to remotely control this transcoder a web-service interface which publishes the transcoding service has been deployed.

The Analyzer Client is a Java application that realizes requests to the transcoding web-service and receives the streaming sent by the server. Once it receives the coded video resource, it decodes the video and analyzes its perceptual quality. The decoding process is done using the FFmpeg framework too. It is mandatory that the client gets the original resource in raw format to allow the analyzer module to perform the resource analysis. The Controlled Network consists on a PC running the DummyNet [18] network emulator, which permits to emulate networks with a specific bandwidth and Packet Loss Rate (PLR). The analyzed codecs, configuration and input resources are shown in Table 2, 3 and 4 respectively. It has been chosen these multimedia resources because they are those used in typical quality assessment studies.

The packet loss rates applied in the video and audio tests were: 1%, 3%, 5% and 10%. Image analysis considers that if there is a loss in the transmission, all the image is lost.

Image	JPEG, GIF, PNG
Video	MPEG-1 video, MPEG-2 video, MPEG-4 part 2, H.263, H.264, WMV1, WMV2
Audio	MP3, AAC, AC3, Vorbis

Table 2. Tested codecs (AMs)

Video	Audio
Bitrate: 1024kbps	Bitrate: 128kbps
Frame rate: 25fps	Sampling frequency: 44100Hz
GoP size: 12	Bits/sample: 16bits
Quantification scale variation:	Coding quality parameter: default codec

Table 3. Configuration parameters

## 6. RESULTS

Results shown in Figure 4 are focused on audio because all the mentioned metrics can be used in audio analysis (Table 1) and also due to space limitations. A deeper insight on results can be found in [16].

From the PEAQ scoring results it can be concluded that the best audio codec in terms of quality is AC3, followed by AAC. However, in terms of compression ratio the best one is Vorbis, although it is the worst in terms of quality. Thus, if the score parameter is considered, the best scored one is the AC3 codec. The scoring function allows to order different implementations of a specific function, this case a transcoding service implemented by different codecs, in order to determine which is the best one. Moreover, the use of SSIM for audio quality assessment is still being studied [13].

## 7. CONCLUSIONS

Notice that this research wants to take special relevance for Future Internet network architectures, which can be used in the development of distributed real-time systems, and will permit the allocation of network services according to each situation and not in a monolithic way. Thus, services must be allocated all along the route, executing just the desired service at each hop, section of hops or end-to-end. Hence, this research should help to pave the way to highly flexible networks, by efficiently applying service oriented approach to networking, resource optimization and service composition.

Concretely, we only present a proof of concept of the proposed framework, where we particularize a general expression for scoring services in the case of selecting a multimedia codec taking into consideration quality assessment metrics. However, it is important to notice that other parameters can be added to the scoring function. Additionally, each AS can propose a specific scoring function in order to select the best AM that is able to provide it. The scoring of an AMs can be done by each node in the network or, if the profile information for all the nodes is available in the network, it can be done by

Image	Lena
Video	Foreman
Audio	Vocal quartet, Instrument flute

Table 4. Tested resources

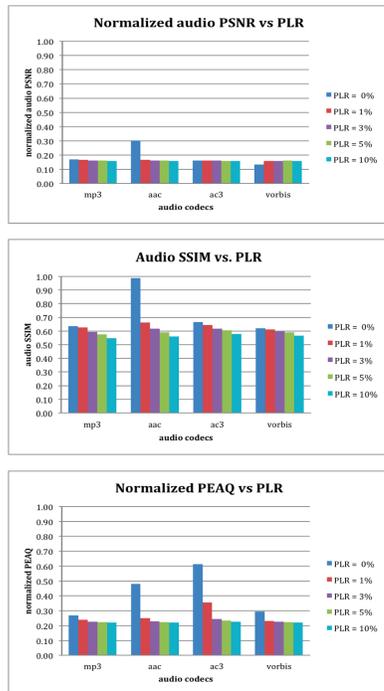


Fig. 4. Testbed results

specific external nodes that can carry out this task. Current services using these techniques will be able to adapt content taking into account the context of users intrinsically.

Moreover, the study presented here introduces a way of enabling context-aware communications in the context of Future Internet architectures based on services. Thanks to these architectures new functionalities can be added in an easy and flexible way, allowing the proliferation of new applications while adapting architectures to past, present and newcoming requirements. Regarding to streaming services, advanced video coding techniques such as MDC (Multiple Description Coding), SVC (Scalable Video Coding) or MVC (Multiview Video Coding) for next 3D formats could be placed in the network, and instantiated only if required, to enable transparent media aware networks and save network resources consumption.

## 8. REFERENCES

- [1] R. Braden, T. Faber, and M. Handley, "From protocol stack to protocol heap: role-based architecture," *ACM SIGCOMM Computer Communication Review*, vol. 33, no. 1, pp. 1722, 2003.
- [2] X. Sanchez-Loro, J. Casademont, J. Paradells, J.L. Ferrer, and A. Vidal, "Proposal of a clean slate network architecture for ubiquitous services provisioning," in *Future Information Networks, 2009. ICFIN 2009. First International Conference on*, 2009, pp. 54–60.
- [3] John Day, *Patterns in Network Architecture: A Return to Fundamentals*, 2008.
- [4] Dennis Schwerdel, Abbas Siddiqui, Bernd Reuther, and Paul Müller, "Composition of self descriptive protocols for future network architectures," in *Proceedings of the 2009 35th Euromicro Conference on Software Engineering and Advanced Applications*, Washington, DC, USA, 2009, SEAA '09, pp. 571–577, IEEE Computer Society.
- [5] A. K. Dey, "Understanding and using context," *Personal and ubiquitous computing*, vol. 5, no. 1, pp. 47, 2001.
- [6] R. Ocampo, L. Cheng, Z. Lai, and A. Galis, "ContextWare support for network and service composition and self-adaptation," *Mobility Aware Technologies and Applications*, p. 8495, 2005.
- [7] M. Autili, V. Cortellessa, A. D. Marco, and P. Inverardi, "A conceptual model for adaptable context-aware services," 2006, p. 1533.
- [8] L. Yang, J. Huai, T. Deng, H. Guo, and Z. Du, "QoS-aware service composition in service overlay networks," in *Web Services, 2007. ICWS 2007. IEEE International Conference on*, 2007, pp. 703–710.
- [9] S. Dustdar and W. Schreiner, "A survey on web services composition," *International Journal of Web and Grid Services*, vol. 1, no. 1, p. 130, 2005.
- [10] S. K. Garg and R. B. Mishra, "TRS: system for recommending semantic web service composition approaches," in *Information Technology. ITSIM 2008. International Symposium on*, 2008, vol. 2, pp. 1–5.
- [11] B. Benatallah, Q. Z. Sheng, A. H.H Ngu, and M. Dumas, "Declarative composition and peer-to-peer provisioning of dynamic web services," in *icde*, 2002, p. 0297.
- [12] Zhou Wang, Ligang Lu, and Alan C. Bovik, *Video Quality Assessment Based on Structural Distortion Measurement*, 2004.
- [13] S. Kandadai, J. Hardin, and C. D Creusere, "Audio quality assessment using the mean structural similarity measure," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, 2008, pp. 221–224.
- [14] T. Thiede, W. C. Treurniet, R. Bitto, C. Schmidmer, T. Sporer, J. G. Beerends, C. Colomes, M. Keyhl, G. Stoll, K. Brandenburg, et al., "PEAQ-The ITU standard for objective measurement of perceived audio quality," *Journal-Audio Engineering Society*, vol. 48, no. 1/2, pp. 329, 2000.
- [15] ITU-R, "Method for objective measurements of perceived audio quality," Recommendation BS.1387-1, International Telecommunication Union, 2001.
- [16] O. Sole Molina, "Multimedia quality assessment," M.S. thesis, Universitat Politècnica de Catalunya, 2009, directed by Alberto J. Gonzalez.
- [17] "FFMPEG transcoder," <http://ffmpeg.org>, Dec. 2010.
- [18] "Dummynet," <http://info.iet.unipi.it/luigi/dummynet/>, Dec. 2010.