Chapter 1: DEFINITION OF PHOENIX PROJECT

1.1.- Project aims

Phoenix Project is born with the aim of generating a transmission system that let at application level (source code, cipher) dialog with physical level (channel code, modulation) over the Ipv6 protocol stack. The main aim is optimize the audio/video transmission, exploiting a joint management between application and physical layers. The implementation of this system must be transparent to Network layer, and therefore, it must work according to actually Ipv6 protocol, exploiting the already existing network infrastructure.

To manage this type of architecture, is needed to transfer the information related to transmitted content and internal state of the system between the different identities , as well as obtain a value of the Service quality at destination node. This information must not impact over the net capacity of this system. It must be possible, particularly, doing a communication between the different protocolary levels, in order to permit the physical and application levels communicating between them.

The objective is jointly managing all the parameters that can be changed at source code, radio channel, network and cryptographic levels, using a source and channel joint code (Joint Source Channel Coding and Decoding, JSCC/D), in order to optimize the audio/video quality perceived from the final user.

Summarizing, the main objectives of Phoenix are:

- To develop innovative schemes to enable JSCC/D. This includes the development of flexible channel coding and modulation schemes and the adaptation of existing source coding schemes with respect their ability for JSCC/D and the development of new ones specifically optimized for this purpose.

- To build a global network architecture based on joint source channel (de)coding for future 4G systems.

- This objective includes the development of JSCC/D controllers which will jointly control the coding blocks and the development of the **network transparency** approach which will allow JSCC/D to be applicable in any kind of network and especially in the 4G ones that should be fully IP-based, to validate and to demonstrate the global architecture and the different building blocks that implement a real time end-to-end platform.

1.2.- Joint Code

An innovative global network architecture will be specified and implemented by Phoenix introducing the concept of **JSCC/D controllers**. These blocks will be the key elements of the architecture deploying a strategy suitable to achieve adaptive/dynamic optimization (efficiency/robustness) of end-to-end wireless links for the transport of interactive high speed packet data(i.e. video and graphics). Developing such a system concept represents a significant innovation as multiple components must be jointly controlled.

The Phoenix Project will investigate the impact of JSCC/D optimization techniques, from simple to more sophisticated ones, on the performance of practical network systems applications. The approach taken by Phoenix is to investigate existing coding schemes (MPEG-2, MPEG-4,H261,etc) with respect to their suitability for JSCC/D , to propose, if necessary, amendments and also to investigate new source coding schemes.

The design of flexible channel (de)coding/(de)modulation schemes and network transparency tools are enabling conditions to achieve the goal and are therefore part of project scope . Being aware of the foreseen fully IP_based structure of 4G, Phoenix pays special attention to let the **source encoder** running at the application layer and the **channel (de)coder/(de)modulator** at the physical layer talk through IP protocol, so as allow JSCC/D optimization of the wireless link. Phoenix allows joint source-channel coding optimization not only for a given source and a given channel, but for a broad range of sources and channel conditions

This platform, formed by different functional blocks, will be built around the targeted applications which involve multimedia streaming over IP networks: tele-working, teleconferencing, virtual presence, interactive networked games, multimedia data base consultation etc… In its initial phase, 4G represents an open area and it is essential for the European industry to constitute a strong **IPR** (Intellectual Property Rights) position in this.

1.3.- End User Point of View

The current users of wireless systems are used to receive and use efficient and safe services with a reasonable cost. But the requirements are always growing up, in terms of quality and characteristics of service. And also is needed to distinguish the different offers according to the fare and the offered service.

Capacity of using the available band of this proposed model, let us impose the limits of the band for certain services, so it permits to offer different categories of service. As well, the type of terminal will be a variable to be considered choosing the parameters of transmission, in order to optimize the data flow according to the capacity of the end receiver.

The service must be utilizable in all points of the Network, and therefore hand-over mechanisms will be provided in a completely transparent way for the end user. These mechanisms must guarantee acceptable and compatible times with the offered characteristics. In case it is needed, the parameters can be dinamicaly changed when the user is moved, in order to adapt him to the characteristics of the new Network and the new wireless channel.

Some of the applications that are though to be applicable on this project will be:

- **Video on Demand:** This server let the end user select a video content from a central server and see it on his terminal at any time. The end user has the full control of the flow, being able to stopping it, to going on, to backing off… Depending on the finality (education, entertainment….), the requirements will be different. This application (Video on Demand) requires the higher requisites.

- **Video Telephony and video conference:** Audio and Video real-time communication and full-duplex between two or more users in different locations. The required quality for this communication is typically at least telephonic type for the audio part, however for the video part the required quality can change a lot. In fact, there are systems in which video part is composed by static images with a refresh time (more or less some seconds of refresh period), or systems with fullmotion characteristics. Nowadays these systems are implemented over UMTS Networks, traditional telephonic network (5 frame/sec),or with solutions that are supported on a PC connected to Internet but with low guaranties over QoS and over the continuity of the service.

- **Voice over IP:** With this term all alternatives telephonic services to the traditional commuted network are grouped. It is an economic alternative, based on a Data Network but a good QoS is not ensured.

- **Videogame:** Entertaiment is also an important part of the market. Particularly, multiplayer services are considered, in which the communication is used to connecting people, so more player can add the players group.

- **Music on demand:** It is like Video on Demand but only with audio flow but with high quality. Wideband demanded is lower than the one demanded in VOD. Utilized algorithms are those ones related with audio compression.

These are some of the utility examples for this proposed architecture. But we have concentrated on video flow, because is the more critical type of flow due to the demanded wideband and to the fact that is real-time.

1.4.- Reference Architecture

1.4.1.- General Description

Next Figure presents the overall architecture considered by the project.

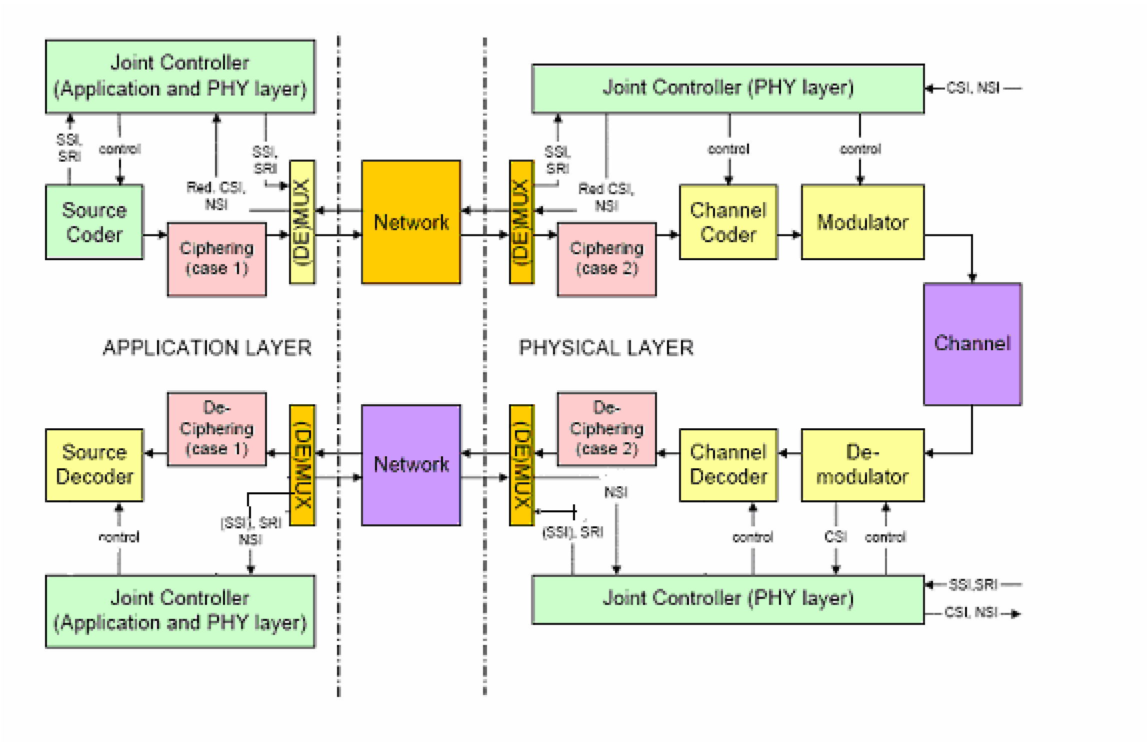

Figure 1.1 - End-to-end communication system over an IP based network

Rather than keeping the traditional approach of having each network layer working transparently from the other, it is proposed in the PHOENIX project to make the end-points aware of each other in order to perform a joint optimization of the use of available resources in the transmission chain. In practice, this means that the transmission chain components will exchange, **via the controllers**, information that they used to keep for themselves.

Phoenix project proposes to make the key elements of the architecture talk together through co-ordinating tools. The co-ordinating tools are named **"joint controllers",** which will implement the necessary controlling strategies and drive the whole communication chain by providing to each block the parameters to use (coding ratio, spectral efficiency, interleaving, video format,...). The function of the controllers is to use feedback from other layers (provided by corresponding controllers) and, also to modify the layer functionality according to the information. This provides mean for the whole system to adapt itself to changes in the separate **application, network and physical layers**. Control information can be exchanged between the system layers utilizing the **network transparency** method. The control information can be included into IP packets from which it is extracted by these special additional entities.

There are two joint controllers: at the application layer, that controls the source coding techniques and will be analyzed in this paper, and another one at physical layer that handles the channel coding and modulation techniques to be used on the wireless link.

1.4.2.- Source Coder

The PHOENIX project main goal is to effectively exploit the available bandwidth. The bandwidth may vary during the transmission and because of this, the

highest possible quality of the video is not always achieved. For achieving the project goal, the video coding tools have to be **adaptive** to the changing transmission conditions and resilient to transmission errors. Hence, coding efficiency, adaptability and robustness to transmission errors are important features for a video coder. The newest video coding schemes improved the overall efficiency remarkably compared to previous video coding technology. The two standardization bodies ISO/IEC MPEG and ITU-T have joint their effort to develop a new standard (MPEG-4 AVC and H.264 respectively), that outperforms the successor technology nearly by a factor of two in terms of compression efficiency.**H.264/AVC** standard includes many advances in terms of both, coding efficiency enhancement and flexibility for different networks and applications, but unfortunately, the standard does not include at this moment scalable coding tools, which would offer a technique to perform adaptability. **Scalability** refers to the capability of recovering physically meaningful video information by decoding only parts of a compressed bitstream. On the other hand, existing adaptive technology (e.g. the SNR scalable MPEG-4FGS) only has poor compression efficiency. In addition existing schemes are not designed for the optimization of joint source channel coding. The MPEG forum has recently announced a call for proposal on scalable video coding technology.
It must be taken into account that the Source Coder has to be suitable for JSCC/D. This requirement means that source coder has to exchange information with the joint controller.

1.4.3.- Source Decoder

The source decoder may utilize the control signal based on Decoder Reliability Information (**DRI**) for executing soft decoding or use DRI for error concealment . Source decoder may additionally pass the Source A posteriori Information (**SAI**) to the physical layer. This makes it possible to perform source controlled channel decoding or iterative channel-source coding.

Examples of typical video applications are given in Table 1.1 for classical video formats that will be considered in the context of PHOENIX.

Name	Number of pixels (width x height)	Typical application
Sub-QCIF	128 x 96	mobile applications
QCIF	176 x 144	streaming over Internet, mobile applications
CIF	352 x 288	films on CD, streaming over Internet
4-CIF	704 x 576	DVD

Table 1.1.- Examples of typical applications for classical video formats.

See also ANEX document [1.1]

1.4.4.- Ciphering

The major issue in securing transmission is to allow applications to build a secure end-to-end communication channel to prevent eavesdropping of the exchanged data. In future integrated networks, this service will rely on heterogeneous infrastructures, from network topology to fibre or wireless physical media. Furthermore new applications demand other security services, such us authentication for group communication or encryption for multimedia data protection. These new services express different security requirements (different security policies) with respect to different types of data. Key exchange is not dealt with in the same way as short lived text message delivery.

In other words, versatile communication applications require the security services to be implemented in a versatile manner.

The Network layer cryptographic transformations aim to protecting the data flow when carried on the network. Application-Layer encryption is not enough so we use Nertwork-Layer encryption to completely hide the stream. The price of this added protection is additional computational overhead, and depending on the encryption scheme used, worse tolerance of packet loss. The standard for network-layer encryption IPSec is only appropriate for unicast media sessions, so we have others encryption systems like IDEA, RC4, DES, 3DES, Blowfish, and AES. The principal design decision to make here is whether to use symmetric or asymmetric encryption. In practice, the two are often combined: an asymmetric algorithm is used to facilitate key exchange for the symmetric algorithm that encrypts the rest of the session

1.4.5.- IP Networking

The IP networking and network transparency requirements of Phoenix system includes the mandatory support for **IPv6 network protocol**. The IPv6 network protocol stack is based on widely known TCP/IP protocol suite definitions and includes the Internet networking protocol definitions, transport protocols, controlling and utility protocols and upper layer user service protocols. The Phoenix project will concentrate on the transport, control and service protocols which are required for robust video transmission. In order to define service access points and protocol triggers for cross-layer communications and network transparency, some control protocols and mechanism e.g. ICMP/IGMP and header compression mechanisms, can be studied in more detail.

Figure 1.2 represents the basic IP networking in case of real-time media streaming. The media is at first encapsulated by some method, for example MPEG or H.261. Then the Real-Time Protocol (RTP) is used for media transport to some transport layer protocol such as UDP or UDP Lite (TCP is not suitable for real-time purposes). RTP takes care of timing, loss detection, content labelling, talk spurts and encryption issues. The companion RTCP (Real-Time Control Protocol) transports some additional control data packets for signaling and QoS purposes. IP (Internet Protocol) at the network layer takes care of routing of the packets etc.

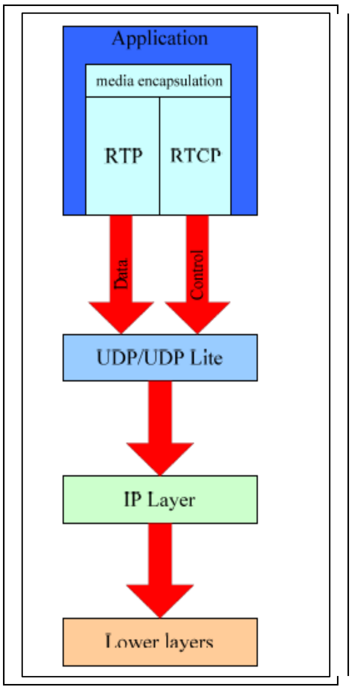

Figure 1.2 : IP networking & streaming media

1.4.6.- Network Transparency

Network Transparency is a basic concept that allows the support of a JSCC/D technique. The goal is twofold. From one side, to transparently cross a generic network node, as well as to permit the communication between different layers of specific concerned devices. The communication between the different layers of specific node refers to the cross-layer communication design. On the other side, the goal is not to require a generic network element to be aware of non-standard JSCC/D functionalities which inevitably would entail a score of software and hardware updates. These are fundamental requirements to be addressed in order to open the way towards a wide deployment of joint source and channel (de)coding techniques in the next generation IP wireless networks. The Network Transparency mechanism derives from two jointly working designs; cross-layer communications and network transparent communications.

Some mechanisms that could implement the concept of Network Transparency have been briefly described in paragraph 3.4, such as IPv6 data packets and extension headers, ICMPv6 messages, and direct socket-to-socket communication; another possible method relies on the introduction of adaptation layers at the receiver side to allow for the exchanges implicated by joint decoding. Assuming that only the source and destination terminals, as well as the Tx and Rx radio nodes, are JSCC/D aware and hence enhanced, several different mechanisms can be adopted to transmit control/signalling information between and inside them.

The Network Transparency concept is a basic and important aspect when building the network structure we want to implement. With this aim, all the strategies have been collected to transfer the information between all the entities that concern the joint code system, trying the network structure to be totally transparent to the joint controllers. Managing the communication between all the entities at the end node and particularly between the different layers of OSI is required to achieve these objectives.

The cross layer communication design is usually defined as a layer co-operation method, in which the protocols that belong to different layers share the control information. This kind of approach, which causes inevitably the breakage of traditional strict layered OSI protocol model, can overcome the network performance problems especially with wireless networks. The full-scale cross-layer communication model provides the interlayer signaling between the protocols in all the layers as illustrated in Figure 1.3. This kind of communication solution, where all of the protocols use the state information throughout the stack and adapt their performance according to the state, can of course enhance the performance of the whole system.

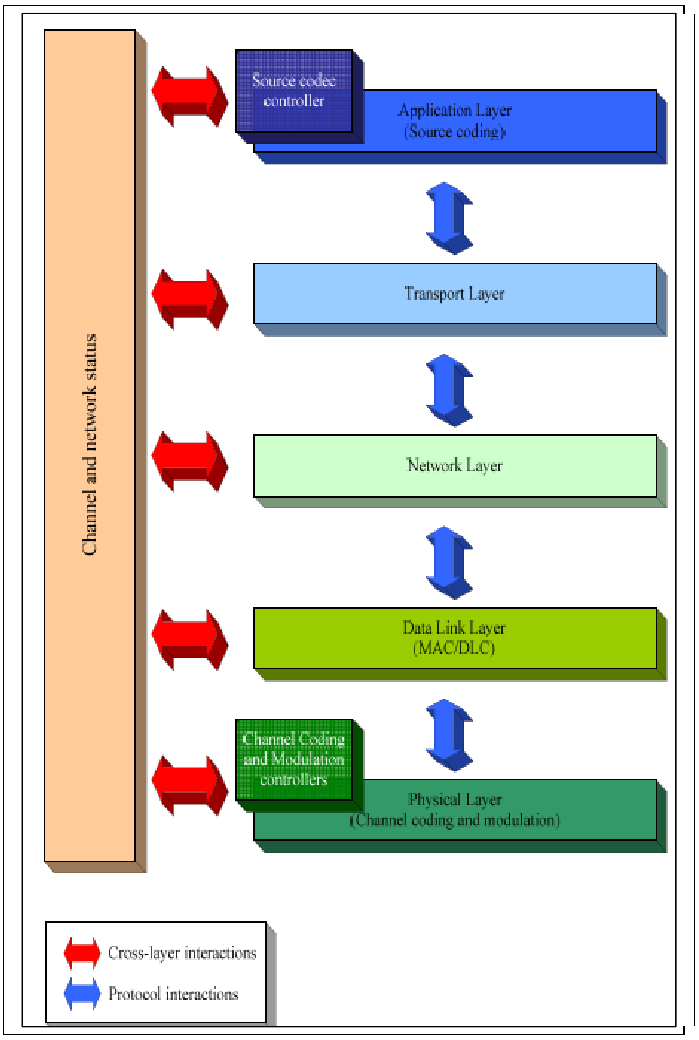

Figure 1.3 : Cross-layer protocol stack model and interactions

On the other hand, the full-scale cross-layer mechanism will make the system and protocol code highly complex and hard to maintain. If the cross-layer communication model is not well designed, it can also decrease the system performance if unintentional interactions are produced. In order to optimize the system, a proper casespecific solution between the full-scale cross-layer communication and layer triggering methods should be found.

<u>MobileMan</u>

MobileMan (Mobile Metropolitan Ad Hoc Network [4]) is a project that aims to exploit careful cross-layer design in mobile **ad hoc** networks (manets). In its approach, the protocols belonging to different layers can cooperate by sharing network status information while still maintaining separation between the layers in protocol design. The vertical Network Status layer uniformly manages the cross-layer interaction. It aims to optimize the overall network performance by increasing local interaction between protocols, decreasing remote communications, and consequently saving network bandwidth.

3GPP cross-layer communication model

(3rd Generation Partnership)

The generic cross-layer communication model, which provides communication mechanism across the whole layer structure, was illustrated in Figure 1.3. The Figure 1.4 represents correspondingly the UMTS network and service architecture.

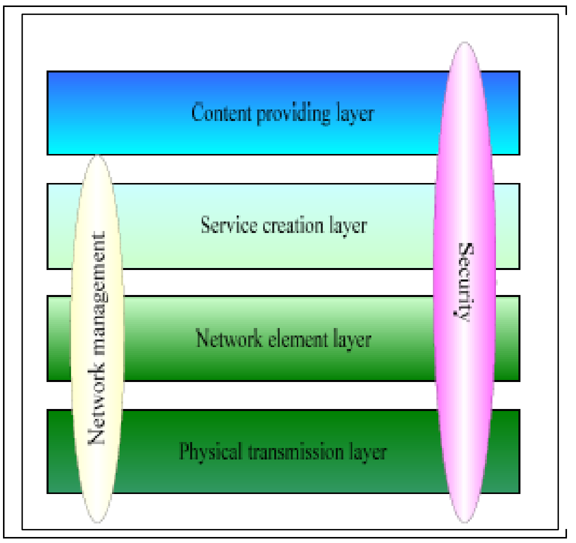

Figure 1.4. UMTS network architecture

The current specification of 3GPP access stratum defines controlling mechanism, which enables cross-layer signaling between physical layer, medium access control, radio link control and radio resource control entities.

The Figure 1.5 represents the signaling lines between the access stratum entities. The cross-layer design of 3GPP is carried out only at the lowest layer of system i.e. at the physical transmission layer, which corresponds to the physical and data link layers of the OSI model. In Phoenix system the interconnection between physical transmission layer, packet data protocols and service and content providing layers need to be created.

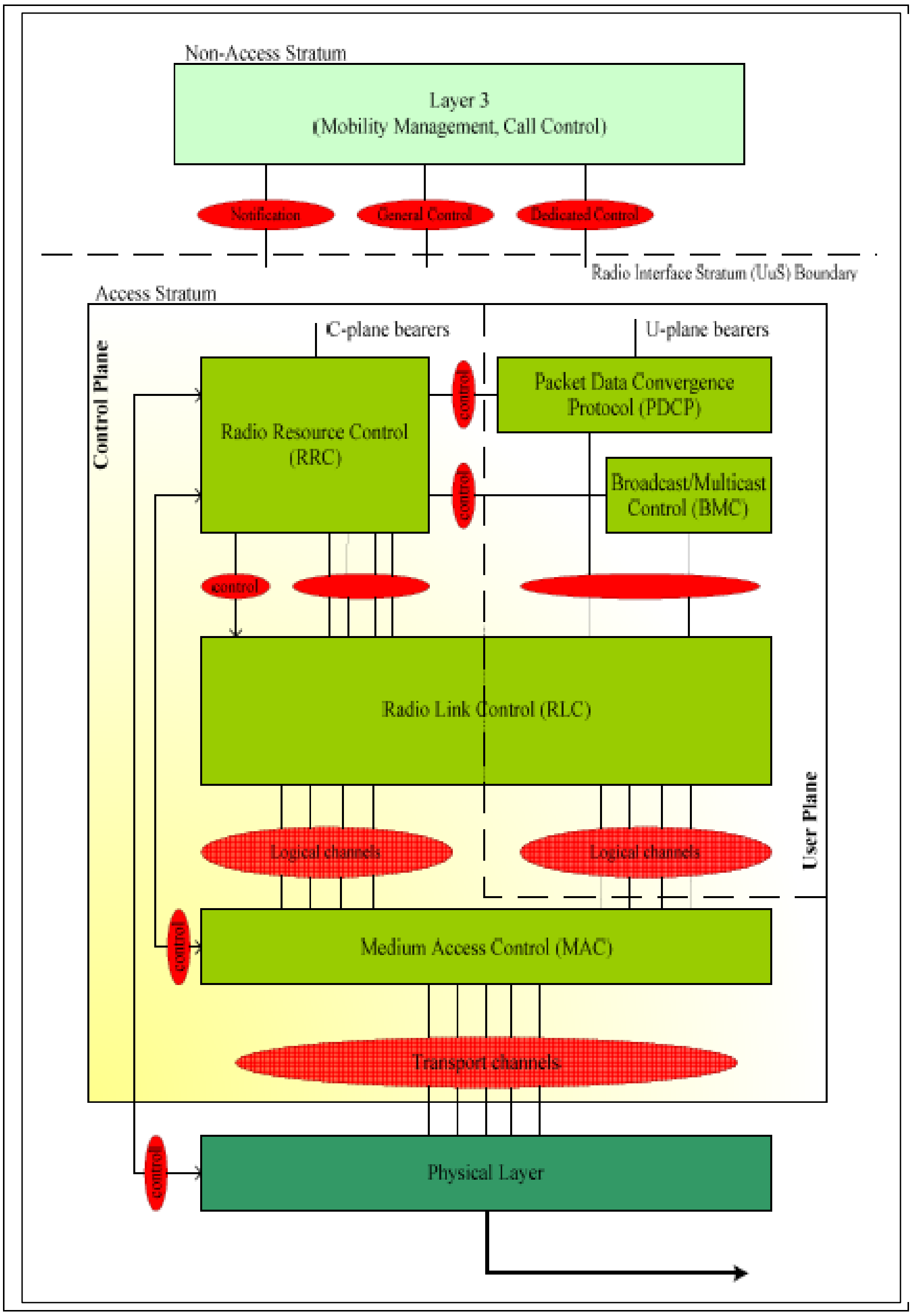

Figure 1.5 : 3GPP Radio interface protocols and cross-layer control lines

3GPP/UMTS bearer and QoS architecture and mechanisms, which are illustrated in Figure 1.6, can guarantee the quality of service for most of the transmission cases. However, these mechanisms may not fully meet the case for joint source and channel coding and new mechanism and signaling (network transparency) over the UTRA radio interface (Uu) and UMTS interface between SGSN and RNC (Iu) and through core network needs to be developed as well as cross-layer communication within UTRAN network architecture.

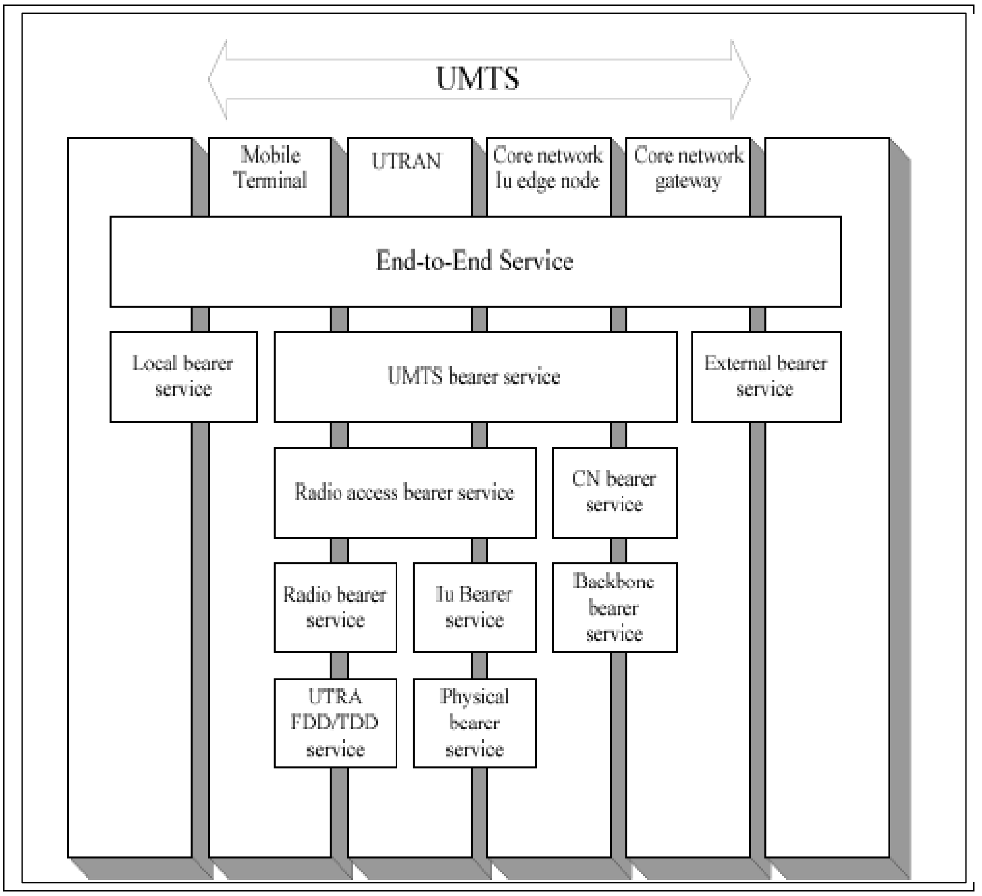

Figure 1.6 UMTS QoS architecture

1.4.7.- Physical layer controller Module

The Physical layer controller's task is to optimize the channel coding, interleaver and modulator parameters according to the information collected from the system. Only control of channel coding through convolutional channel codes was considered in the basic simulation chain. New options have been developed in the current version of the simulation chain in order to also control channel coding with different classes of codes and modulation parameters. In particular, specific modules for controlling channel coding through LDPC codes and to provide multi-carrier modulation adaptation (adaptive loading) are under development.

1.4.8.- Channel Co-Decoder

The basic chain (See attached document) only considered convolution coding, with Viterbi decoding or MAP decoding. In the second version of the simulation chain, beside these convolutional codes, LDPC codes will be taken into consideration, with the option of using Belief Propagation or Min-Sum decoding algorithms. Hard or Soft output can be considered. The option of using Turbo codes has also been taken into account.

1.4.9.- Modulator

Advanced high-rate services such as high-resolution interactive video and audio applications are unlikely to be supported by the 3G systems owing to their limited bitrate capabilities. These bitrate-related challenges remain to be solved by future 4G-type mobile broadband systems. The most recent version of the IMT-2000 3G standard is in fact constituted by a family of five independent standards. The current solutions on the market aim for attaining international roaming and flexibility by engineering tri-band solutions, but **reconfigurable solutions** would be more desirable.

For the sake of high bandwidth efficiency, jointly optimised iterative source-coding, channel-coding and modulation will be used, where technically feasible. Trellis Coded Modulation (TCM),Turbo TCM (TTCM), Bit-Interleaved Coded Modulation (BICM) and iteratively decoded BICM (BICM-ID) schemes will be studied. A significant coding gain may be achieved without bandwidth expansion, when exchanging information between the Variable-Length Coded (VLC) video stream for example and the coded modulation decoders with the advent of iterative decoding. With the aid of using independent interleavers for the In-phase and Quadrature phase components of the complex-valued constellation, further diversity gain may be achieved. The performance of the proposed schemes will be evaluated when communicating over both AWGN and Rayleigh fading channels.

The above-mentioned coded modulation schemes may also be rendered adaptive to channel-quality variations, transmitting a higher number of bits during instances of high channel quality and a low number of bits during low channel quality spells. The associated effective throughput in the various coded modulation modes may be 1, 2, 3 or 5 bits per symbol, associated with coding rates of 1/2, 2/3, 3/4 and 5/6, respectively. The accurate estimation of the channel quality is of high importance, which may be based on the Mean Squared Error (MSE) at the channel equaliser's or multi-user detector's output. More explicitly, the residual signal deviation from the error-free transmitted phasors at the DFE's output reflects the instantaneous channel quality of the time varying wideband fading channel. Hence, given a certain instantaneous channel quality, the most appropriate modulation mode can be chosen according to this residual signal deviation from the error-free transmitted phasors at the DFE's output. A practical modem mode switching regime adapted to the specific requirements of wireless video telephony may be employed, where a suitable modulation mode is chosen at the receiver on a Burst-by-Burst (BbB) adaptive basis and it is then communicated to the transmitter by superimposing the requested modulation mode identifier code onto the terminal's reverse direction transmission burst. Hence, the actual channel condition is outdated by one TDMA/TDD frame duration. Specifically, four different adaptive modulation modes, namely **4QAM, 8PSK, 16QAM and 64QAM** may be invoked and protected by the above-mentioned coded modulation schemes.

In high speed wireless data applications, the orthogonal frequency division multiplexing (OFDM) modulation scheme has been considered, due to its relatively simple receiver structure compared to single carrier transmission in frequency selective fading channels. Multicarrier modulation is particularly suitable for transmission in presence of inter-symbol-interference (ISI). OFDM allows flexibility to channel conditions, through adaptive loading techniques. OFDM modulation is adopted by IEEE for the extension of the 802.11 wireless LAN standard to the 5GHz band (IEEE802.11a) providing data rates up to 54Mb/s and for IEEE802.11g. ETSI adopted the OFDM scheme for the High Performance LAN physical layer standard (HIPERLAN2) as well.

Recently, attractive transmit diversity techniques known as Space-Time Block Coding (STBC) and Space-time Trellis Coding (STTC) have been advocated for providing additional diversity gains as well as for overcoming the limited capacity offered by the hostile fading wireless channel. STBCs do not provide coding gain, unless concatenated with an outer channel code. The main benefit of STBCs is the provision of full diversity with the aid of a simple decoding scheme. In contrast to STBCs, STTC relies on the joint design of channel coding, modulation transmit diversity as well as optional receiver diversity schemes. The design rules that are proposed resulted in maximum diversity gain for systems employing two transmit antennas.

Typically, the detection of the STTC symbols is constituted by independent channel equalization and space-time trellis decoding operations. However, instead of implementing the channel equalization and STTC decoding operations independently, the performance of the STTC system can be further improved by performing both operations iteratively.

1.5.- Network Transparency

To achieve the Network Transparency, the following control/Signalling signals have been used:

1.5.1.- SSI

Source significance information (SSI) is generated by the **source encoder**. The SSI specifies the priority of a certain part of the bitstream and length of that part. The application controller can use the SSI to adjust UEP mode. In practice, SSI information shall vehicle the importance of each class of the bitstream in terms of final rendering. Typically, this information will permit distinguish parts of the bitstream that are mandatory for correct (eventually partial) reconstruction, from those that will only provide a better visual result. SSI is strictly related to the data stream so it must be synchronised with it and its frequency, the introduced overhead as well, is somehow proportional.

1.5.2- CSI

CSI is generated by the concerned Radio Receiver node and effectively exploited by all the transmitting nodes, both at the radio interface (for the channel coders and modulators) and at the source terminal (for the source coder). CSI is about the actual condition of a wireless channel crossed by the video stream; possibly more CSI flows are generated when more wireless hops are present along the concerned routing path to the destination terminal(s).

CSI goes in the reverse path with respect to the video data packets, hence it is not strictly synchronised with them (although, it must be up-to-date enough to be helpful). Furthermore, presumably the CSI frequency should be much lower than the packets rate to be considered almost negligible in terms of additional overhead. While the physical layer controller needs detailed channel state information, the application layer controller may use a reduced subset of information: CSI fed to the application layer controller are not necessarily the same as those to be fed to the physical layer controller.

1.5.3.- DRI

DRI is generated by each radio receiver and is collected by the destination terminals in order to better tune the source decoding process and hence improve the resulting QoS [5]. Further levels of complexity are introduced when more wireless interface are present along the data path.

DRI provides further elements related to the channel decoding process. It is based on the concept of "fuzzy" decision, where the final result about the value of a bit is not simply either 0 or 1, but also the level of certainty of it. This means that the introduced overhead can be even higher than the coded video stream rate, for sure not lower. Thus, its benefits and cost must be accurately evaluated.

Some sort of integration methods, that produces a single set of DRI, can be envisioned, but in principle each DRI flow can be treated separately. DRI must be strictly synchronised with the video stream and needs to be available at the receiver terminals at the same time.

1.5.4.-- SRI

Source a priori information (SRI) is further information produced by the source encoder that are exploited at the destination side and possibly also by the other entities concerned in the JSCC/D chain along the data path at the radio transmitter nodes, in order to optimise the QoS resulting from the decoding process of the video stream.

Example of SRI is for instance the encoding type or parameters that were used at the emitter side, elements which can be used by the decoder to decoder better, in particular in intelligent source decoding context.

The related overhead can be even of the same order of the coded stream, depending on the provided precision; however, a better compromise would lead to lower amount of signalling avoiding an excessive waste of both computational and communication resources.

SRI is synchronized with the associated video stream, more and more in relation of its accuracy.

1.5.5.- SAI

SAI (source a posteriori information) results from the analysis of the decoding process of the video stream, hence generated by the destination terminal(s) and it is exploited at the radio receiver(s) for iterative decoding techniques [6], or more specifically to set the working parameters of the channel decoder and demodulator module(s) or even at the source terminal, in order to improve the performance of the said entities and the resulting QoS as well.

SAI travels from the destination terminal(s) to the source one, i.e. not strictly synchronised with the video data packets; however, a relation between it and the concerned video fragment must be enforced, more and more depending on the provided accuracy.

The related overhead can be even of the same order of the coded stream, depending on the targeted level of detail; however, a better compromise would lead to lower amount of signalling avoiding an excessive waste of both computational and communication resources. It is to be noticed that in many systems the uplink direction is characterized by narrower band connections, particularly into the access part of the network, thus in most cases (*e.g.* video streaming from a server to one or more end-user terminals), there is a more stringent constraint than for the source a priori or decision reliability information.

In a multicast scenario, some sort of integration methods applied at the multicast tree branching nodes that produces a single set of SAI can be envisioned, but in principle each flow can be treated separately.

1.5.6.- NSI

Network state information (NSI) consists basically in the real time transport control protocol (RTCP) information. Interesting monitor information for RTCP information include: round-trip estimation time, packet loss ratio and inter arrival jitter. Round-trip estimation time is used for estimating interactiveness and control delays for end to end connection. The whole transmission and packet loss ratio during short period of time can be used to controlling transmission. Inter arrival jitter can be used for estimating difference in packet spacing at the receiver compared to the sender for a pair of packets. The application level controller can adjust QoS control, error resilience and UEP mode based on NSI. Also the controller can utilize NSI to control bit-rate, spatial resolution and temporal resolution.