Introduction

There has been a lot of progress in the development of new wireless broadband technologies recently. 4G technologies such as mobile WiMAX and LTE are almost at the end of their commercial development phase, and are currently in their deployment phase. Clearwire, under the brand name Clear, has already deployed their mobile WiMAX service in major metropolitan areas such as Atlanta, Austin, Charlotte, Chicago, Las Vegas, Philadelphia and Seattle. Verizon has announced plans to deploy LTE in metropolitan cities of Boston and Seattle by the end of year 2010. Both WiMAX and LTE are competing standards that are going to make a strong push for a big market share for providing data services to mobile users. Although Gartner predicts that LTE will win out against WiMAX in the race to 4G adoption, even they acknowledge that neither of these technologies is going to grab a complete market share. So the 4G market space will be a combination of WiMAX and LTE wireless technologies.

In addition to the 4G network deployments, there has already been a large scale deployment of 3G networks (EV-DO CDMA by Sprint and Verizon, HSPA/UMTS by AT&T). Even when the 4G technologies become mature, the 3G technologies are not going to completely disappear from the market. So the eventual wireless broadband market space is going to consist of a combination of WiMAX, LTE and already existing 3G technologies. These technologies will co-exist with widely deployed variants of Wi-Fi (WLAN) technologies. This will lead to a wireless market space that is truly heterogeneous in nature.

With the advent of multi-radio capable devices (iPhone, netbooks etc.) and cognitive radios, a mobile device is now capable of using resources of all these heterogeneous technologies. To efficiently manage mobile device's use of resources in a heterogeneous network environment, an efficient radio resource management (RRM) scheme is imminent. This paper summarizes radio resource management frameworks, relevant standards, and optimization algorithms that exist in current literature.

RRM Frameworks

Common Radio Resource Management

The common radio resource management (CRRM) architecture has been proposed by 3GPP working group to facilitate interworking between UMTS and GPRS/EDGE networks. The CRRM architecture is a network-centric optimization approach based on a hierarchical centralized implementation. Figure 1 depicts the CRRM architecture solution. Each AWS/RAT in a given cell is managed by a local RRM entity. The RRM entity for UMTS is present either inside a radio network controller (RNC) or is connected to a RNC, and the RRM entity for GPRS/EDGE is present either inside a base station controller (BSC), or is connected to a BSC. The RNC is responsible for collecting and reconfiguring UMTS-related parameters and the BSC is responsible for collecting and reconfiguring GPRS/EDGE-related parameters for their operating cell. Both entities pass technology-specific measured parameters to the CRRM entity through its local RRM entity. The CRRM entities may also communicate with each other to gather parameters measured by RNC and BSC entities in adjacent cells. The parameters are processed by the decision engine in the CRRM entity. The CRRM entity uses the results of the decision engine to assist the RNC and BSC in configuration policies which help achieve some predefined network objective function.

In a CRRM scheme, some optimization criteria can be processed individually by the local RRM. The individually processed optimization criteria are not as efficient as optimization criteria jointly processed

by the CRRM. The number of joint optimization criteria processed by the decision engine in the CRRM is strongly determined by the degree of interaction between the RRM and the CRRM. The degree of interaction in turn is strongly determined by the level of coupling between the different access technology entities (BSC, RNC and WLAN). The tighter the coupling, the higher the degree of interaction is possible, and the higher the number of joint optimization criteria can be processed by the decision engine in the CRRM. The coupling strategies (open, loose, tight and very tight) are driven by the business agreements between the carriers. Some examples of business agreements are given in [ref].

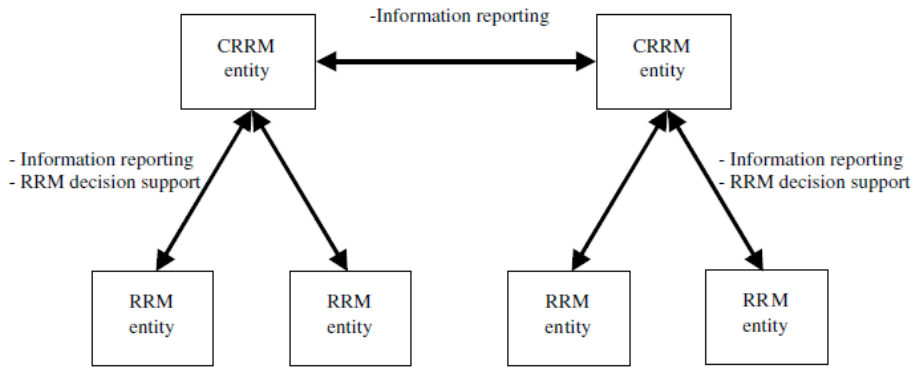


Figure 1. CRRM architecture functional model

Joint Radio Resource Management

The joint radio resource management (JRRM) architecture has been proposed by Siemens. The core concepts of this management system are multi-homing and service splitting. JRRM splits the service supported by the user terminal into fundamental part and enhanced part and the former is delivered by RAT with large coverage range, e.g. UMTS. JRRM is a network-centric centralized management scheme just like CRRM. JRRM allows joint managing of traffic streams between the involved networks and terminals through a very tight coupling. JRRM architecture with two underlying RATs is depicted in Figure 2. The central decision engine which consists of joint resource scheduler, joint session admission control, and joint load control (JOSCH, JOSAC and JOLDC) acts as network resource manager. The network resource manager uses the load information that it receives from the local RRMs and comes up with optimized configuration parameters. The optimized configuration parameters are communicated to the local RRM which uses these parameters to configure its local resource scheduler, session admission control and load control entities. Once the local parameters are configured, the user terminal is informed of the resources it can receive on each RAT by the respective local RRMs. The user terminal in turn configures itself to appropriate service splitting levels.

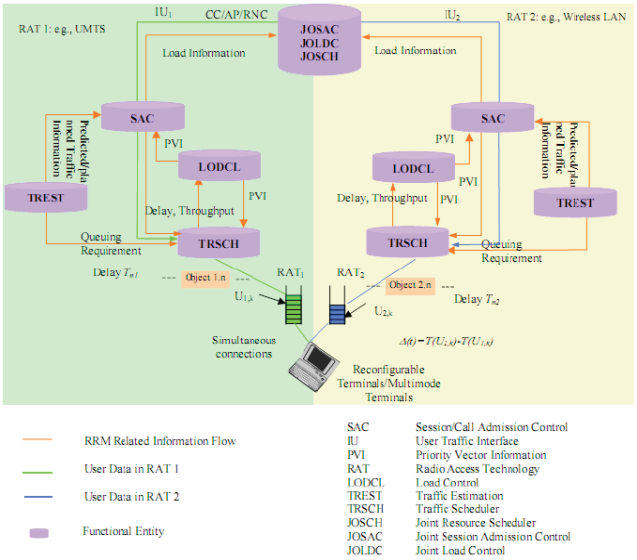


Figure 2. JRRM Architecture

Multi-access Radio Resource Management

The multi-access radio resource management (MRRM) scheme defines a network distributed mode that is currently supported between UMTS and GPRS/EDGE technologies according to 3GPP specifications. It also defines a network centralized hierarchical mode that is similar to CRRM scheme. The MRRM distributed mode is based on multi-agent implementation as depicted in Figure 3. The agents are responsible for collecting the status information of the underlying specific RAT and the information of neighboring agents. Based on the optimization criteria, optimization related input parameters are gathered at one specific agent in the system. The decision engine in the chosen agent deals with various issues such as coexistence, optimizations etc. In information gathering and result distributing process, two message passing interfaces are used, namely the Hr interface and the Ha interface. The Hr interface is used to carry information between the RAT and the corresponding agent. Through this interface, an agent collects radio network information such as parameters related to radio resource, parameters about performance etc. and sends control factors to the RAT. The Ha interface (Agent-Agent interface) is used to carry information concerning inter-network resource adjustment and coordination and status information of each other. The optimization results related to terminal configuration are communicated to the terminal through local RRMs.

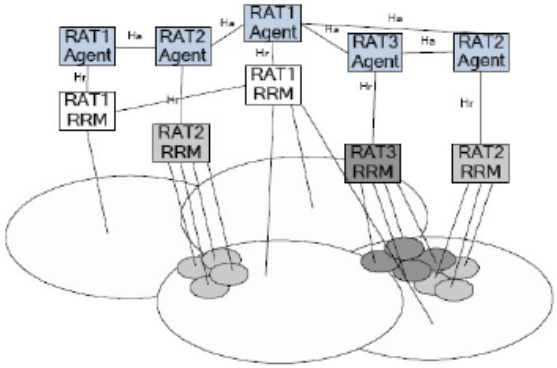
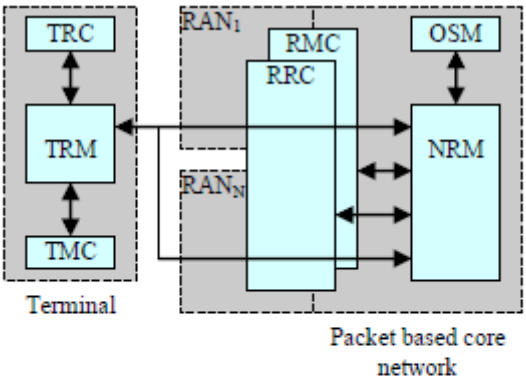


Figure 3. Agent-based Distributed MRRM architecture

Relevant Standards

IEEE P1900.4

The IEEE P1900.4 standard defines building blocks for enabling coordinated network-device distributed decision making which will aid in the optimization of radio resource usage, include spectrum access control, in heterogeneous wireless access networks []. The system architecture defined by the 1900.4 standard is shown in Figure 4.



TRC – Terminal Reconfiguration Controller
TRM – Terminal Reconfiguration Manager
TMC – Terminal Measurement Collector
RMC – RAN Measurement Collector
RRC – RAN Reconfiguration Controller
OSM – Operator Spectrum Manager
NRM – Network Reconfiguration Manager
RAN – radio access network

Figure 4. IEEE 1900.4 system architecture

Four management entities are defined on the network side: the Operator Spectrum Manager (OSM), the Random Access Network (RAN) Measurement Collector (RMC), the Network Reconfiguration Manager (NRM) and the RAN Reconfiguration Controller (RRC). OSM is the entity that enables operator to control dynamic spectrum assignment decisions to be made by NRM. RMC is the entity that collects RAN context information and provides it to the NRM. NRM is the entity that manages Autonomous Wireless Systems and terminals for network-terminal distributed optimization of radio resource usage and improvement in QoS. NRM makes RAN reconfiguration decisions and sends RAN reconfiguration requests to the RRC. RRC is the entity that controls reconfiguration of RANs based on requests from the NRM. To support scalable operation, RMC, NRM, and RRC may be implemented in a distributed manner.

Three management entities are defined on the terminal side: Terminal Measurement Collector (TMC), Terminal Reconfiguration Manager (TRM), and the Terminal Reconfiguration Controller (TRC). TMC is the entity that collects terminal context information and provides it to the TRM. Terminal context information is defined in the standard and may include the following: User preferences, Required QoS levels, Terminal capabilities, Terminal measurements, and Terminal geo-location information. TRM is the entity that manages the terminal for network-terminal distributed optimization of radio resource usage and improvement of QoS. TRC is the entity that controls the reconfiguration of terminal based on requests from TRM.

Correspondingly, six interfaces are specified to exchange messages between various entities in the 1900.4 system architecture as can be seen in Figure 4.

IEEE 802.21

The IEEE 802.21 standard facilitates the handover between different wireless networks in heterogeneous environments regardless of the type of medium. The goal of the IEEE 802.21 standard is to improve mobile nodes' usage experience by providing uninterrupted handover in heterogeneous networks. To provide uninterrupted handover, the handover procedures can use the information gathered from both the mobile terminal and the network infrastructure.

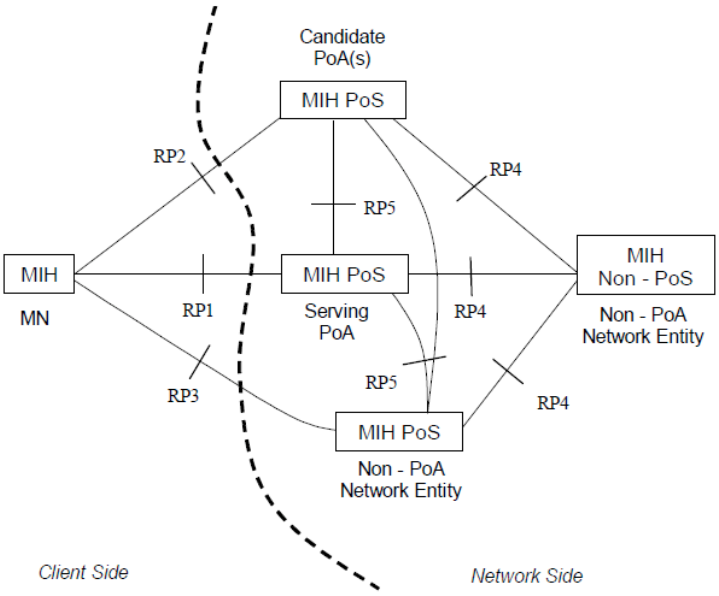


Figure 5. Media Independent Handover (MIH) Communication Model

The communication model of IEEE 802.21 is presented in Figure 5. As can be seen from the figure, communication reference points (RP1 to RP5) have been defined for communication between MIH-enabled Mobile Node (MN), PoA (Point of Attachment), PoS (Point of Service), non-PoA PoS, and non-PoA non-PoS entities. Communication can take place between client–network entities, and between network–network entities. The required optimization parameters are transported to the required entities using the MIHF (Media Independent Handover function) implemented at Layer 2.5 of each MIH enabled entity (Figure 6).

MIHF defines three different services: Media Independent Event Service (MIES), Media Independent Command Service (MICS) and Media Independent Information Service (MIIS). MIES provides events triggered by changes in the link characteristic and status. MICS provides the MIH user necessary commands to manage and control the link behavior to accomplish handover functions. MIIS provides information about the neighboring networks and their capabilities.

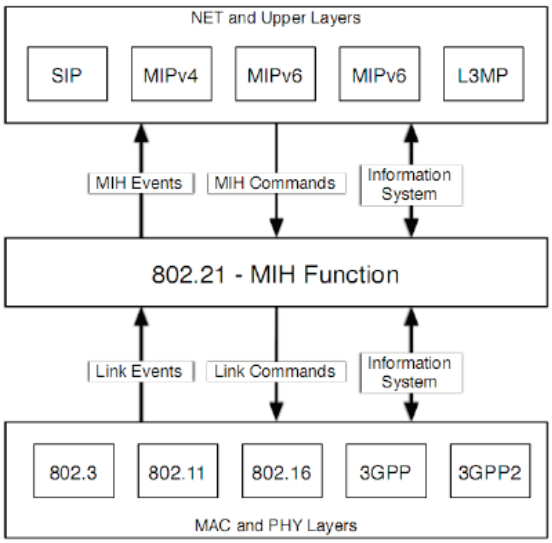


Figure 6. MIH Architecture

The list of event, command and information service parameters that can be passed between various network and terminal entities are included in Table 5 (Page 41), Table 7 (Page 45) and Table 10 (Page 50, 51) of the IEEE 802.21-2008 standard [ref]. The link status varies with time and MN mobility. Information provided by MICS is dynamic information composed of link parameters such as signal strength and link speed and should be periodically updated; whereas, information provided by MIIS is less dynamic or static in nature and is composed of parameters such as network operators and higher layer service information. MICS and MIIS information could be used in combination by the MN/network to facilitate the handover. A number of commands are defined in this standard to allow the MIH users to configure, control, and retrieve information from the lower layers including MAC, Radio Resource Management, and PHY.

It is interesting to note that both standards (IEEE P1900.4 and 802.21) call for obtaining optimization input information from several RANs that might have a same measurement parameter but which might still be uncorrelated. These standards do not talk about specifying any mechanisms to normalize these parameters which would produce values that could be readily used by the optimization engine. It is left up to the optimization engine to normalize the measured parameters.

Optimization Algorithms

The optimization engine takes the input parameters and processes them according to some specific criteria. The optimization engine can be as simple as one that uses policy based algorithm, or it can be as complex as a full-fledged cognitive engine implementing various optimization methodologies. A brief description of various optimization algorithms found in literature is provided next.

Policy Based Algorithms

Policy based algorithms can be categorized into two different approaches (i) Service based policy (ii) Radio network based policy. In the service based policy, different application services are mapped to different access technologies. For example, a voice service is always allocated to cellular access technology (CDMA) whereas a data service is always allocated to WLAN access technology (IEEE 802.11).

In a radio network based policy, given that the amount of radio resources necessary for an indoor user in UMTS is considerably higher than for an outdoor user, the policy would define that the indoor users are allocated to WLAN while the outdoor users are allocated to UMTS. Note that there would be several mechanisms suitable for estimating whether a user requesting service is indoor or outdoor (e.g. located-based services, comparison between estimated path loss and required path loss, etc.).

Algorithm for joint optimization of Load-Balancing (APs) and Battery Lifetime (MNs)

Vertical handoff decision algorithms in heterogeneous wireless networks to jointly optimize the overall traffic load (load balancing) among all attachments points (APs and BSs) and the overall battery lifetime of mobile nodes is studied in [ref]. The vertical handoff decision algorithm is implemented at the Vertical Handoff Decision controller (VHDC) located in the access networks. The decision inputs for the VHDCs are obtained via Media Independent Handover Function defined in IEEE 802.21 standard. For VANETs, a route selection algorithm is also devised to select the most appropriate AP or BS based on the joint optimization function. A flow-chart for high level procedure used by VHDC based on joint optimization requirements (α = Battery Lifetime, β = Load Balancing) is given in Figure 7.

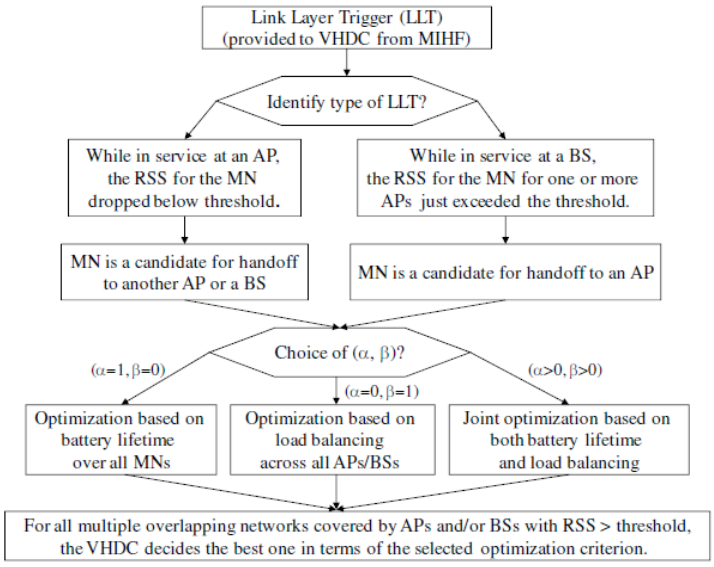


Figure 7. Flow chart for high-level procedure used by the VHDC

For battery optimization criteria, a Max L (maximize the battery lifetime of all MNs in the network) and for load balancing criteria, a Min F (minimize the variance of average load experienced by an AP/BS) equation is formulated as a mixed integer problem. The joint optimization function (Opt-G) is formulated as the weighted sum of the Max L and Min F equations. The mixed integer problem is solved using TOMLAB and CPLEX for an example topology with 2 BSs, 5 APs and 50-100 MNs. Simulation results for the optimization functions are compared with that of traditional SSF approach (MN connects to an AP with which it receives strongest signal strength) and is presented in terms of average remaining battery lifetime and coefficient of variance (standard deviation of loads observed at APs divided by mean load).

Fuzzy-Neural Algorithm

The fuzzy-neural algorithms are good at explaining how to reach decisions from imprecise information by using the fuzzifier and defuzzifier rules and the inference engine concept. This concept has been demonstrated to work better for cases where there are several input variables from different access technologies that are not very strongly related. A study of three technologies UMTS, GPRS/EDGE, and WLAN using the fuzzy-neural concept has been provided in [ref]. Figure 8 depicts the block diagram of the proposed Fuzzy-neural algorithm. A detailed explanation of the fuzzy-neural concept is provided in [ref]. In addition, [ref] does a comparison study on Fuzzy-Neural algorithm that uses signal strength, resource availability and mobile speed as input parameters and shows that it outperforms other JRRM algorithms that only try to use least loaded criteria (LJRRM), and lowest path loss criteria (PLJRRM).

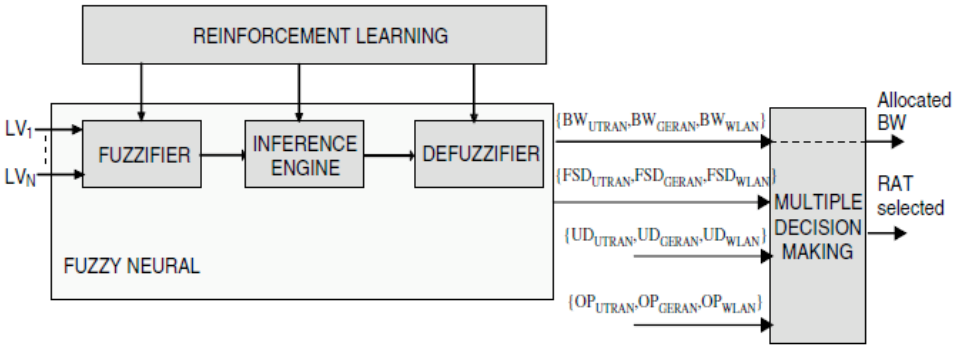


Figure 8. Block diagram of Fuzzy Neural algorithm

The list of algorithms listed is not exhaustive and several other approaches exist. In addition to these algorithms, real test-bed implementations are being studied by RiWCoS and Aragorn projects in Europe. The RiWCoS implementation is based on IEEE 802.21 implementation for message exchange. The Aragorn implementation is based on Cognitive Resource Manager implementation. More information about both those projects is provided in [ref]. A few vertical handover optimization test-beds have also been implemented using IEEE 802.21 architecture [ref][ref].