Bringing Always Best Connectivity Vision a Step Closer: Challenges and Perspectives

Malamati Louta*[1], *Member IEEE*, Paolo Bellavista[2], *Senior Member IEEE*
[1]*Department of Informatics and Telecommunications Engineering, University of Western Macedonia, Greece*
[2]*DEIS, University of Bologna, Italy*

***Abstract*— 'Always Best Connectivity', which constitutes a key challenge in the context of 4G systems, solicits service access and provisioning through the most appropriate access network any time any place. In this relatively well-investigated area, the problem of dynamically selecting the most suitable network for a specific service, referred to as Access Network Selection (ANS), has recently received considerable attention. However, the several ANS proposals in the literature, which have explored relevant ANS criteria, methodologies, and techniques, point out that some related technical issues are still open challenges to be resolved. The aim of this paper is to identify and discuss on critical aspects and research challenges involved in the design of ANS decision schemes. At the same time, current research efforts are revisited and potential enabling technologies/solutions are highlighted, in particular the ones associated with cognition and advanced learning capabilities.**

***Index Terms*—Always Best Connectivity (ABC), vertical handover, access network selection, learning, cognitive networks.**

I. Introduction

Future communication systems will be increasingly complex, involving thousands of heterogeneous nodes with diverse features and various networking technologies with different characteristics and capabilities, interworking with the aim to provide users with ubiquitous access to advanced high-quality services in a cost efficient way at any time and any place, in line with the Always Best Connectivity (ABC) principle. The ABC concept provides users with the ability to connect each time with the most appropriate network in order to access the requested services according to user preferences, requirements and constraints, service/application and terminal profiles, network capabilities and related context. At the same time, users should remain agnostic of the heterogeneity of the underlying infrastructure as well as of its potential modifications, with service continuity, robustness/availability, and consistency maintained transparently.

The realization of the ABC vision falls within the realm of handover management procedures, which should be flexible and efficient, while involving complex multi-criteria considerations and trade-offs. Handover management involves 1a) deciding on the appropriate time to initiate a handover (by minimizing communication overhead and avoiding unnecessary handovers), 1b) selecting the most suitable access network for a specific service (the respective problem is referred to as Access Network Selection – ANS) and 1c) maintaining seamless service continuity, possibly in a robust way. Handover management is generally decomposed in three phases: 2a) information gathering, which involves detecting all available networks and collecting all relevant information for identifying the need/opportunity for handover, 2b) handover decision, which comprises the decision making process for selecting the most appropriate access network, and 2c) handover execution, which involves transition to the new network point of attachment, as depicted in Fig. 1.

The heart of the overall handover procedure (and of the ABC vision) is the second phase. Traditional handover decisions based on Received Signal Strength (RSS), adopted mostly in homogeneous environments, have demonstrated to be insufficient in highly heterogeneous and open networks [1]. Additional criteria should be considered, such as user requirements and preferences, terminal/service/application characteristics and capabilities, network conditions, economic costs, and security-related aspects. Taking into account the multiplicity and dynamic nature of the aforementioned aspects, as well as potentially unexpected situations, handover could imply an extremely complex decision process.

Handover procedures have received considerable attention in 4G-related research [2-10]. Recent standardization efforts provide a framework for seamless vertical handover support (e.g., IEEE 802.21, IEEE 1900.4, 3rd Generation Partnership Project -3GPP- and 3GPP2 IP Multimedia Subsystem -IMS), without however standardizing the algorithms and strategies to apply for decision making. In this context, a number of ANS schemes have been proposed, addressing the problem from different perspectives, having different objectives, utilizing different decision criteria, and applying different methodologies and techniques. In the light of the aforementioned aspects, the proposed ANS schemes lack unity [8], while a number of issues still need to be resolved. Even though some publications have surveyed vertical handover decision strategies (e.g., [9-10]), they mostly limit their attention to decision criteria and applied methodologies. To

Manuscript received August 19, 2011

M. Louta (corresponding author) is with the University of Western Macedonia, Karamanli & Ligeris str., Kozani, 501 00 Greece (phone: +30 24610 56566; fax: +30 24610 56501; e-mail: louta@ uowm.gr). P. Bellavista is with the University of Bologna, Italy (paolo.bellavista@unibo.it).

the best of our knowledge, there is no prior effort in the literature providing a comprehensive overview of aspects and issues to be considered when designing ANS solutions. Our aim is to cover this gap by identifying and discussing on critical issues and research challenges involved in the design of ANS schemes, while revisiting current research efforts and highlighting enabling technologies and solutions. Our ultimate goal is to contribute towards the definition of a commonly accepted ANS framework, by both providing a better understanding of the proposals published so far and pointing out relevant directions for future work.

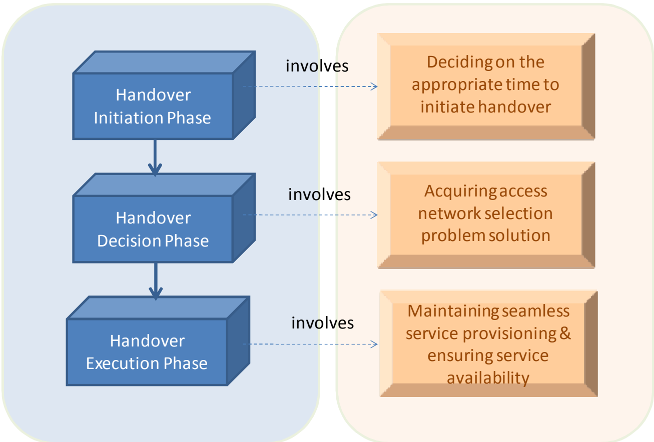

Fig. 1: Handover phases.

II. ANS-Related Standardization Efforts

Recent 4G standardization efforts provide mechanisms for seamless vertical handover support. In this section, we aim to highlight the standardization advancements related to ANS.

The IEEE 802.21 standard for local and metropolitan area networks specifies media access independent mechanisms that facilitate handovers between heterogeneous networks. Specifically, IEEE 802.21 enables cooperative handover decision making by providing common information representation across different networks (e.g., via XML) and standardizing Media Independent Handover Function (MIHF), a mechanism for information exchange through MIH messages between mobile terminals and network attachment points. Related to the network selection process, a mobile node or a network entity makes a decision to connect to a specific access network based on information obtained by exploiting MIHF services and policies configured in the mobile node and/or the network. Handover control, policies, and other algorithms involved in handover decision making are assumed to be handled by a network selector entity and do not fall in the scope of this standard. Suitable amendments are required to existing standards of different media-specific technologies (e.g., IEEE 802.3, IEEE 802.11, IEEE 802.16, and 3GPP/3GPP2 systems) to satisfy the requirements identified by the IEEE 802.21 standard.

In 3GPP TS 23.402 and TS 24.312 specifications, Release 11, a new network element called Access Network Discovery and Selection Function (ANDSF) is defined in System Architecture Evolution (SAE) to support access network discovery and selection. ANDSF is optionally deployed in the network infrastructure and contains data management and control functionality to assist mobile nodes in the network selection process through the provisioning of operators' mobility-related policies. Upon a mobile node's request, ANDSF may provide a list of access networks available in the mobile node vicinity, including info about access network identifier and technology type. Mobility policies are operator-defined rules and preferences that affect mobility decisions taken by the mobile node. They may indicate whether a specific technology type or a specific access network identifier is preferable to another, under which conditions mobility is restricted from one network type to another and under which conditions the policies to be enforced are valid.

The IEEE 1900.4 standard defines the architectural building blocks enabling network-device distributed decision making for optimized radio resource usage in heterogeneous wireless access networks. Specifically, the standard aims to improve the overall composite capacity and quality of service of wireless systems in a multiple radio environment by defining suitable architecture and protocols to facilitate radio resource optimization, including dynamic spectrum access control. Proper network and terminal reconfigurations are employed based on information exchanged between network and mobile terminals.

IMS is designed to provide robust multimedia services to roaming users over diverse access networking technologies. The IMS architecture was collaboratively formed by 3GPP and the Internet Engineering Task Force (IETF), is access independent, while session control lies on the Session Initiation Protocol (SIP). 3GPP Voice Call Continuity targets handovers between the IMS packet switched domain and the circuit switched domain of GSM/UMTS.

III. ANS Design

In this section, a systematic classification of critical issues and challenges for ANS design is presented.

A. ANS Objectives and Control

The first issue to be considered when designing an ANS mechanism is the determination of the problem objectives in conjunction with the entity that will undertake the responsibility and control of the whole procedure. From the network operators' perspective, highly competitive and open environments should encompass mechanisms that will assist them in accounting for their interests, i.e., offering, at a given period of time, adequate quality services in a cost-efficient manner. Such network-centric mechanisms are often associated with efficient management of network resources when fulfilling users' requests. In addition, user-centric schemes, which address the ANS problem from the users' side, aim at assisting and enabling users by focusing on the satisfaction of user requirements, preferences, and constraints.

Network-centric approaches may be implemented in three complementary ways: a) adopting proper optimization techniques, b) employing network cooperation, and c) exploiting network reconfigurations to adapt to changing conditions through cognitive networking [11]. The whole procedure is typically controlled by a network operator-related entity, residing in the network operator's domain, which holds the intelligence to fulfill the task.

Through network cooperation it is possible to successfully handle excessive traffic demand by directing users to alternate networks/technologies that can offer the requested services according to current user/application requirements and preferences. In this perspective, ANS techniques should take into account (among other issues) potential agreements between network operators. However, network cooperation has to face some possible limitations [11]. Massive deployments of network elements are required to lessen inter-network operators' dependencies, which in turn increase the necessitated (and potentially risky) investments and the respective capital expenditure. Additionally, it should be noted that massive deployments of heterogeneous network elements tend to result in their low utilization, thus, leading to inefficient resource usage.

Cognitive wireless access networks are expected to overcome the aforementioned drawbacks by adapting to changing conditions so as to handle user requests in the most appropriate manner. Cognitive networks employ proper reconfigurations to the current Radio Access Technology (RAT) or activate an alternate one by triggering appropriate software. They promise to be characterized by self-* properties (self-configuration, self-optimization, self-healing, self-protection, self-management), following the autonomic computing principle. Specifically, they include mechanisms for selecting their configuration; they form their future behavior in accordance with system goals, profiles, policies, and current operational context, while they exploit knowledge from previous interactions to adapt to external stimuli and optimize performance.

User-centric approaches generally fall within two distinct categories: a) Mobile Controlled Network Assisted (MCNA), according to which a user-related entity residing in the mobile terminal domain undertakes the ANS task, by exploiting network-related information and b) Network Controlled Mobile Assisted (NCMA), where a network–related entity considers information and measurements gathered from the terminal to decide on the current "best" access network.

Network-centric as well as NCMA user-centric approaches face limitations in the context of the deregulated and highly competitive telecommunications market. Specifically, users may be unwilling to reveal to network operators their preferences, requirements, and constraints, in fear of the operators (unfairly) capturing the whole surplus of the market. Additionally, a dilemma is posed to the network operators that should be trusted to find the best possible network among the available ones (even those belonging or administered by competitors), overcoming conflicting goals and business policies. Furthermore, security issues should be carefully considered, as context transfer between different operators may be involved. MCNA user-centric approaches are more flexible, relieving the network from significant complexity, while they are considered to be an imperative property of 4G ABC environments.

Our view is that co-existence and potential interworking (e.g., by means of a negotiation phase) of both types of user/network–centric mechanisms would facilitate the realization of the ABC vision. The two associated types of ANS problems have different and potentially contradicting, yet strictly related, objectives. They are addressed by entities that may act in a selfish manner, while the decisions reached have interdependencies to take into account. Thus, an ANS problem instance may be initiated and controlled by a network operator to handle potential QoS degradations and/or efficiently utilize network resources, by reconfiguring specific network parameters/elements or re-directing users to another RAT (either of his own or belonging/administered by an affiliated network operator), aiming primarily at revenue maximization. In addition, users should be enabled to initiate an ANS process, especially in case they feel the current selection does not fulfill their requirements/preferences, seeking to maximize their satisfaction in terms of perceived QoS and imposed cost. A first attempt to jointly model network selection and resource allocation problems can be found in [4].

B. ANS Initiation

Careful consideration should be given to the conditions that should hold for initiating an ANS problem instance, as its solution involves computing/communications cost increase, while inefficient solutions could lead to unnecessary handovers, with possible detrimental impact on QoS, system capacity, and signaling load. ANS should be activated in case of: a) a new service request, b) considering an active service session in case of b1) identification of a new wireless access network that under certain conditions could constitute a better alternative to the current one, b2) QoS degradation below a certain threshold, c) imperative and robustness-related conditions, such as failure of the current RAT interface, network failure or failure of handover execution, and finally d) after user intervention, as illustrated in Fig. 2.

A still open question is whether ANS problem instances should be handled independently for each service request. In general, it is possible to select two different networks to access two different services requested by a user at the same time. However, this entails several technical challenges (e.g., security-based, interference-based, power-based) that still need to be solved [2]. Similar issues apply to the envisioned and complex future scenarios of enabling multi-path multi-homing, where even the same single application can synergically exploit more than one access network.

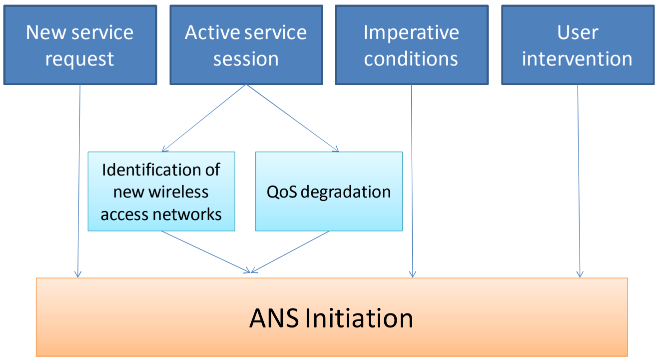

Fig. 2 ANS mechanism activation

C. ANS Decision Criteria

A critical factor for designing ANS mechanisms is the decision parameters to be evaluated. In general, ANS solutions should satisfy user preferences, requirements, and constraints (both QoS- and budget-related), take into account service/application characteristics, consider terminal/network capabilities, operational context, and economic cost. Some of the criteria are considered static in the sense that their changes do not often incur (e.g., user profiles, terminal characteristics), while others are more dynamic (e.g., network conditions) [1].

ANS schemes in the literature have identified various network characteristics as potential criteria, while subsets of them have been used in their decision making strategies. They may be grouped as: a) link quality, evaluated considering indicators such as RSS, Carrier to Interference Ratio (CIR), Signal to Interference Ratio (SIR), Signal to Noise and Interference Ratio (SNIR), b) network availability, considering coverage, bandwidth availability, and call blocking probability, c) QoS-related aspects, considering throughput, delay, latency, jitter, Bit Error Rate (BER), packet loss ratio, average number of retransmissions per packet, and d) network reliability, considering call dropping and handover failure probabilities.

User profiles may designate, for each service, the features of primary interest and an associated set of corresponding quality levels. A quality level can be seen as the specification of the (perhaps range of) quality parameter values relevant to the service feature. Users could also specify a maximum price (tariff) that can be afforded for service access (willingness to pay), a minimum value concerning the anticipated user satisfaction from accessing a service through an access network, a preferred network technology/operator, and a minimum reputation value a network operator should hold. Service/application characteristics as encoded in the respective profile may involve distinct features (e.g., audio, video) composing the service/application as well as corresponding QoS parameters. Terminal capabilities may involve service features in conjunction with corresponding quality levels that could be supported by the terminal currently in use.

Distinct service/application profiles associating features at different QoS levels may be published by network operators along with the corresponding cost for service access. Users may be allowed to select one of the published profiles or, instead of being limited to a fixed set of inflexible choices, they may initiate a negotiation process with the network operator, in line with the personalization aspect of the Beyond 3G vision. Users may be assisted in selecting the best profile and/or in negotiating by a specialized entity, residing on their terminals and acting on their behalf.

The cost of access network resources constitutes a major decisive factor in ANS. Network operators may exploit different billing schemes based on duration and/or volume of data. Most proposed schemes use cost as a static criterion. We believe that cost should be a parameter dynamically formed on the operator’s side, ideally in the context of each user request, taking into account specific factors, such as current resource utilization. In general, cost determination requires mapping requested service features and corresponding QoS levels to resources necessitated from the network.

Finally, contextual information may comprise current network load conditions, terminal velocity, terminal location, and remaining battery lifetime to support power utilization efficiency in the overall selection process.

D. ANS Methodology & Algorithms

The decision methodology followed to determine the most appropriate access network is another significant factor for ANS. However, in the related literature, no commonly accepted classification of the proposed algorithms exists. For example, in [10], the authors classify algorithms in four groups based on the main criterion used: RSS-, bandwidth-, cost function-, and combination-based. Combination-based solutions include algorithms that utilize a rich set of inputs and apply machine learning techniques in the decision process. In [9], algorithms are classified into either basic or advanced, by considering the set of parameters used for ANS, the latter category including context-based, fuzzy logic, and neural networks to interpret imprecise information and also Multi-Criteria Decision Making (MCDM) algorithms. In [1], strategies are classified into five main categories: function-based, user-centric, multiple attribute decision, based on fuzzy logic and neural networks, and context-aware strategies. User-centric strategies focus on user satisfaction, aiming to maximize a utility function, while function-based strategies aim to minimize the utilization of network resources, however failing to evaluate user satisfaction. Context-aware strategies combine additional criteria, while information concerning the operational environment is gathered, managed and evaluated in order to reach proper decisions on handover initiation and ANS.

We claim that ANS is inherently a MCDM problem, thus, it could be solved adopting multi-objective (MODM) and/or multi-attribute (MADM) related methodologies and algorithms (e.g., multi attribute utility theory methods, outranking approaches such as ELimination Et Choix Traduisant la REalité −ELECTRE- and Preference Ranking Organization METHod for Enrichment Evaluations -PROMETHEE-, Analytic Hierarchy Process -AHP-, Grey Relational Analysis -GRA-, Technique for Order Preference by Similarity to Ideal Solution -TOPSIS-, Simple Additive Weighting -SAW-, and weighted product model -WPM). MCDM algorithms can be used in combination with fuzzy logic when input attribute values are not clearly defined, thus enabling advanced decision methodology based on imprecise or incomplete data.

As final notes, all the aforementioned classifications do not incorporate a category for policy-based strategies, even though various systems have been proposed in related research literature (e.g., [3]) and have been adopted by recent ANS-related standard specifications (e.g., IEEE MIHF and 3GPP ANDSF). Policy-based systems are claimed to be sufficient for handling complexities in 4G systems, permitting to avoid sophisticated decision models and cost functions.

E. ANS Evaluation

To the best of our knowledge, there is lack of a commonly accepted evaluation methodology, thus constituting comparative performance evaluation a cumbersome task. This may be attributed to the fact that ANS evaluation in different research works follows specific and distinct objectives, leading to the adoption of different input parameters, decision criteria, methodologies, and performance evaluation metrics.

The performance metrics most commonly used for evaluation purposes comprise the number of handovers performed, handover success/failure rate, delay associated with the three handover phases, and packet loss. A collection of the performance metrics considered in different ANS schemes can be found in [9].

In most cases, performance evaluation is conducted via simulation; in fact, it is very difficult (if not impossible) to develop quantitative analytical models of the proposed mechanisms for several complexity-related motivations, e.g., in many cases a large number of decision criteria are considered. The authors build different simulation environments utilizing network simulators such as ns-2, OPNET, MATLAB or in some cases self-designed testbeds.

F. ANS & Negotiation

Considering that a new connection/handover request is usually refused in case the selected network is heavily loaded, since the network cannot accommodate new user requests (at least without degrading the quality of currently connected users) and user requests could be rejected by network operators to maximize long-term revenues or reserve resources for their "premium" users, a new ANS problem instance initiation may be needed. Therefore, to guarantee successful handovers as much as possible, a negotiation phase between users and network operators may be necessitated. Additionally, as already mentioned, personalization of service offerings requires some form of interaction between users and network operators. However, negotiation functionality is not commonly found in related literature. We refer to [4] and [5] as related studies comprising a limited and extended set of negotiation capabilities, respectively.

Careful consideration should be given to the design of the negotiation phase, taking into account user speed (that may result in an imperative handover) and specific time-related constraints. Multi-round negotiations serve better the personalization issue, while take-it-or-leave-it offers should be considered when time limitations with approaching deadlines apply. At this point it should be noted that negotiation phase could be overrun in case of an imperative handover due to severe QoS degradation in order to avoid call dropping. In any case, a scheme for estimating whether a user will be admitted to a new network on the basis of its current load and estimated load difference would be helpful.

G. ANS & Power Utilization Efficiency

In the literature, power utilization efficiency has been addressed mostly in the information gathering phase, while it may also form a decision criterion in the ANS decision process. Intelligent interface management and adaptive handover initiation solutions (e.g., [4, 6]) involve keeping one interface active at a time for communication, turning off high-power consumption rate interfaces (e.g., WiFi, WiMAX) in case the remaining battery lifetime is less than a predefined threshold or even adapting the interface activating interval so that power consumption is decreased; the objective is to prolong device lifetime. Additionally, user location plays an important role in the application of power-saving criteria, defining proper policies according to user preferences.

H. ANS & Reputation

The establishment of trust constitutes an issue of outmost importance for the success of 4G environments. Service provisioning calls for a high degree of cooperation among diverse actors, who, seeking for the maximization of their welfare, may misbehave, thus leading to the deterioration of the overall system performance.

In general, misbehavior (i.e., deviation from regular functionality, which may be unintentional due to faults or intentional if selfish parties wish to take advantage of given situations) can significantly degrade system performance. Traditional models aiming to avoid strategic misbehavior may be inadequate or even impossible to apply due to the complexity, heterogeneity, and high variability of the considered environments. Reputation mechanisms may be employed to provide a "softer" security layer, considered to be sufficient for many applications. They can establish trust by exploiting learning-from-experience concepts to obtain a reliability value of participants in the form of rating stemming from other entities' view/opinion. Reputation metadata may be disseminated to a large number of participants to adjust their strategies and behavior, rewarding good players and penalizing bad ones (i.e., providing incentives). In general, reputation mechanisms may be centralized/decentralized and can either be based on direct encounters or take into account information disseminated from other parties on the basis of their past experiences with the entity under evaluation.

We believe that the reputation of network operators, with respect to a specific access network, should form another criterion to be taken into account in ANS. A simplified form of a reputation criterion has been used in [6] to eliminate access networks from the candidate list. However, no details are provided with respect to how this list is formed or how it is updated in case the network operators exhibit good behavior in the future. In [7], to speed up the vertical handover decision process, ANS is performed based on the reputation of network operators, which reflects the QoS perceived by previous users for different pre-defined service classes, in terms of BER, delay, jitter, and bandwidth. The proposed reputation system is based on simplified rating functions, with no support at all for users' personalization.

I. ANS & Robustness

ANS mechanisms should be able to efficiently handle a large number of decision criteria, an increasing number of networking technologies, possibly imprecise data and/or partial knowledge, and an uncertain and highly dynamic environment. In addition, they should be capable of reacting and effectively following changing environment conditions, while they should be able to learn from knowledge acquired in the past. In the light of the aforementioned, a critical aspect to take into account in ANS design is robustness, in terms of both capability to promptly react to failures and ability to achieve stable ANS decisions under conditions of partial and possibly imprecise knowledge about the execution environment. To this purpose, policy-based solutions, in combination with machine learning techniques, seem to constitute good candidates for a viable answer to the robustness challenge, which is anyway still open in the ANS systems available at the state-of-the-art.

In Table I, we summarize the basic characteristics of some representative ANS schemes presented in the literature. We have decided to clearly depict "Context" and "Cognition/Learning" characteristics, as they are widely recognized as relevant properties of future ANS solutions.

TABLE I. MAIN CHARACTERISTICS OF ANS SCHEMES IN THE LITERATURE

Ref. No.	**Objectives/ Control**	**Decision Criteria**	**Methodology& Algorithms**	**Evaluation**
[1]	User-centric Mobile Controlled	User preferences Network characteristics	MCDM/ Fuzzy with Policies	No details are provided
[2]	User-centric Mobile Controlled	Network characteristics Terminal capabilities User preferences	MCDM	5 networking technologies (UMTS, WLAN – IEEE 802.11a, b, n) Performance evaluation metrics: preference value P of each network
[3]	Network–centric Mobile Controlled	Network characteristics Terminal capabilities	Policy	3 networking technologies (GPRS, WLAN - 802.11b, LAN) Simulation environment: LCE-CL testbed Performance evaluation metrics: number of handovers, handover delay
[4]	User–centric & network-centric Mobile Controlled	User preferences Network characteristics Application requirements Terminal capabilities	MCDM & Q-Learning	3 networking technologies (UMTS, GPRS, WLAN) Performance evaluation metrics: handover failure probability, network revenue rate
[5]	User-centric & network-centric Mobile Controlled	User preferences Network characteristics Terminal capabilities	MCDM	1 networking technology (IEEE 802.11) Performance evaluation metric: negotiation time
[6]	User–centric Mobile Controlled	User preferences Network characteristics Terminal capabilities	MCDM	3 networking technologies (UMTS/HSDPA, WiMAX, WiFi) Simulation environment: OPNET Performance evaluation metrics: application specific (interruption duration ratio for streaming applications)
[7]	User–centric Mobile Controlled	Network Reputation (calculated on the basis of network QoS characteristics with respect to application requirements)	Single Criterion Decision Making	2 networking technologies (UMTS, WLAN) Simulation environment: MATLAB Performance evaluation metrics: reputation of each network for different types of applications, handover delay

Ref. No.	**Negotiation**	**Power Utilization Efficiency**	**Reputation**	**Context**	**Learning**
[1]	No	No	Yes	Yes	No
[2]	No	No	Yes	Yes	No
[3]	No	No	Yes	Yes	No
[4]	Yes (Limited)	Yes	No	No	Yes (operator's side)
[5]	Yes (Fully)	No	Yes	Yes	No
[6]	No	Yes	Yes (simplified form – as a criterion in the elimination phase)	Yes	No
[7]	No	Yes (partially)	Yes (is the only decision criterion considered based on the quality of the candidate network)	Yes	No

IV. ANS & COGNITION - OUR VISION

Incorporation of learning/cognitive capabilities into networks is considered to be a major step towards efficient management of increasing complexity and heterogeneity. The term cognitive networks is used here to indicate networks able to sense their context of operation, analyze, reason and plan, make a decision, and act in accordance with the decision reached, while they learn from previous experience. Cognitive networks have the ability to think, learn, remember, and adapt to changing conditions in order to achieve end-to-end goals and objectives, thus being self-aware.

Cognitive networks may be centralized or distributed. The centralized approach has significant computing, communication, time, and storage advantages, but may suffer from the classical disadvantages of centralized architectures (e.g., performance bottlenecks and single points of failure). Distributed cognitive networks may be formed as a collection of cognitive entities, which incorporate intelligent functionality, have reasoning capabilities, are characterized by autonomy, social ability, learning from experience, and adaptivity, while they interact with other components and act in a reactive/proactive way to accomplish their goals.

Extending the intelligent control loop of an autonomic system following the "monitor, analyze, plan and execute" sequence, a generic architectural framework of a cognitive system may comprise the following modules: *Sensing Module*, *Reasoning Module*, *Learning Module*, *Decision Module*, *Act Module*, and *Policy Module*, as illustrated in Fig. 3. A cognitive system continuously senses its environment to identify potential conditions that could affect its operation status. The *Sensing Module* aggregates, correlates, and filters data, until a condition that should be further analyzed is identified. The observations captured by the *Sensing Module* will be processed and analyzed by the *Reasoning Module*, while they will be also fed to the *Learning Module* that is able to learn and remember useful observations, which can aid the decision making process in the future. The *Reasoning Module* determines potential actions to be taken based on observations, knowledge acquired through the *Learning Module*, and policies stored in the *Policy Module*. The *Decision Module* decides on the actions to be taken, exploiting also learning from experience. The *Act Module* executes the output of the *Decision Module*. The *Learning Module* can learn from several sources, e.g., from collected information, strategies, decisions, and feedback received; it can correlate and infer from this knowledge.

Focusing on handover management, the cognitive entities may be classified into two main categories that are in principle in conflict: the user-related cognitive entity, acting on behalf of the user, and the network-related one, acting on behalf of the network operator. The two entity categories have different and possibly contradicting objectives, are considered to be selfish and interested in maximizing their owner's profit, whilst they may cooperate to assist each other in the decision process or even to reach a joint ANS decision. Both entity types comprise the aforementioned six cognitive modules and may be implemented as an extension to MIH entities in the IEEE 802.21 standard or as an extension to ANDSF-related entities if working in compliance with 3GPP specifications. Specifically, the Sensing Module identifies available access networks, measures and aggregates QoS-related information, observes current context (e.g., user's velocity/location, battery status) and forwards these metadata to the Reasoning Module. The Reasoning Module decides whether a handover process should be initiated, taking into account communication overhead and potential bouncing effects, based on information received from the Sensing Module and knowledge acquired from the Learning Module. Additionally, it eliminates candidate access networks in case they do not satisfy specific policies. For example, a network may not be able to provide a requested service or satisfy the minimum QoS requirements imposed by the user (e.g., minimum data rate). Additionally, in case a user moves at a high speed, networks with small coverage range may be withdrawn from the candidate list. Furthermore, the Reasoning Module may constrain the candidate list, in case the reputation of a candidate network operator operating a specific access network is too low.

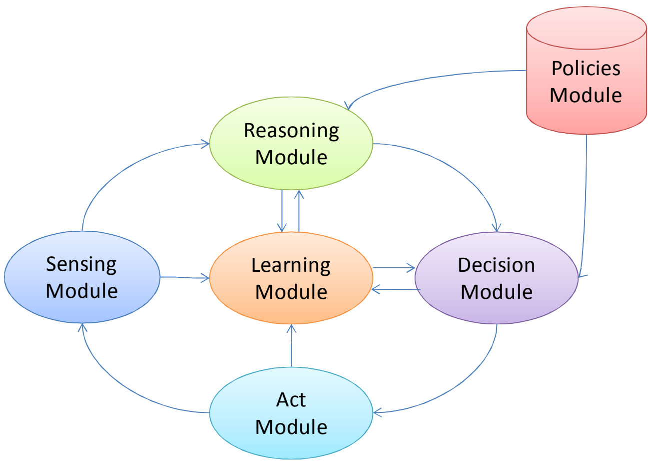

Fig. 3 Architectural framework of a cognitive system.

The output of the Reasoning Module (filtered candidate list) is forwarded to the Decision Module in order to decide on the most appropriate access network. The Decision Module takes into account knowledge inferred from the Learning Module and additional policies specified in the Policy Module. Such a policy could be "in case user's location is home or office, give lower significance to the remaining terminal's battery", or "in case user is risk adverse, give higher relevance to reputation of network operator, while in case user is willing to accept the associated risk, give lower significance to reputation". Different methodologies and algorithms (e.g., policy-, MCDM-, fuzzy logic-based, or potential combination of those) could be applied.

The Learning Module should incorporate mechanisms for correlating observations and inferring knowledge from the contextual environment of operation to aid the decision process. Without being exhaustive, such mechanisms may assist in a) deciding on handover initiation, b) building the reputation of each network operator for each of the access technologies it provides on the basis of users' past experiences and feedback returned, c) estimating the probability of user's admittance to a network in case handover is decided, prior to its execution, d) predicting traffic generated due to users' requests, e) forming the strategy that should be followed in case of negotiation, f) estimating the price that should be posed for providing network resources, g) acquiring the most appropriate reconfiguration or selecting the most appropriate network operator considering network cooperation paradigm, and h) modifying policies in accordance with user and/or network operators' preferences.

Finally, the Act Module performs the handover. User's experience will be fed to the Learning Module in order to update the respective knowledge. In Fig. 4, we graphically illustrate the cognitive handover management process along with the cognitive modules that undertake the responsibility for fulfilling each task.

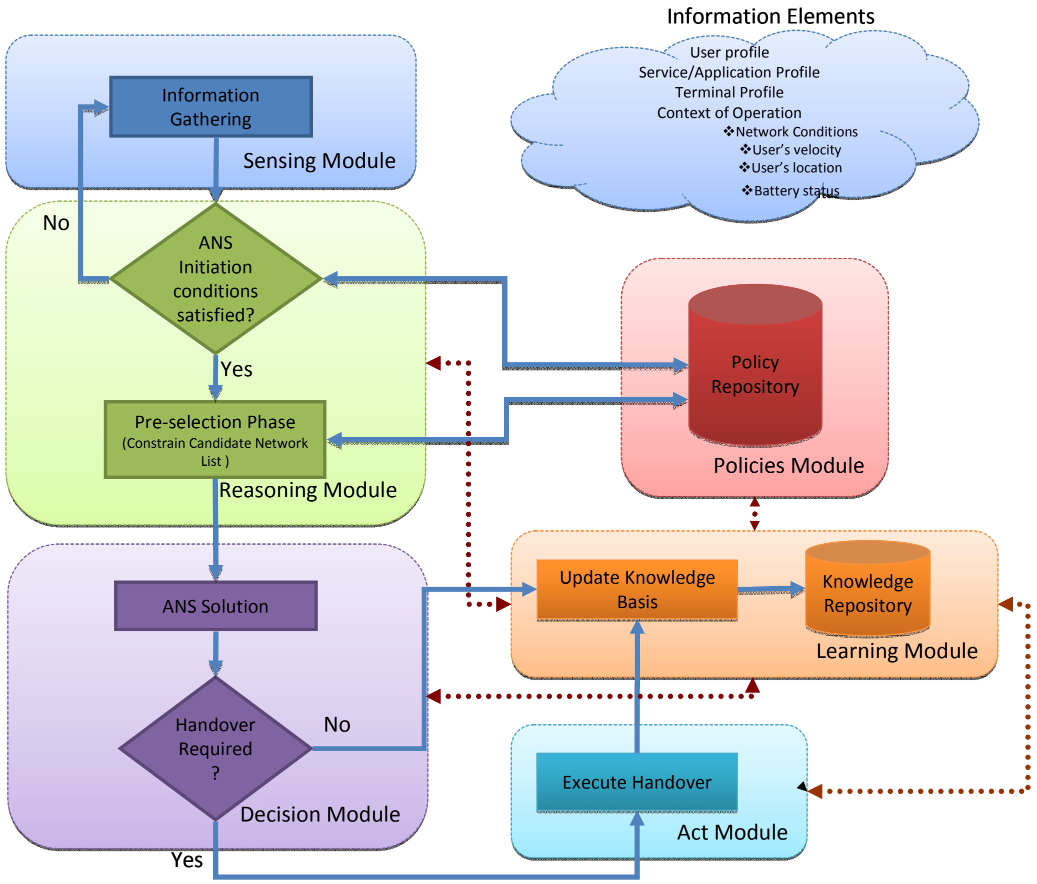

Fig. 4 Cognitive Handover Management Process.

V. Conclusion

Next generation wireless networks are migrating to 4G systems, involving heterogeneous access networks, such as mobile communications systems (2G, 2.5G, 3G, 3.5G), Long Term Evolution (LTE), Wireless Local/Metropolitan Access networks (WLAN/WMAN), Wireless Personal Area Networks (WPAN), ad-hoc networks, sensor networks, short range communications as well as Digital Video/Audio Broadcasting (DVB/DAB), while soliciting their cooperation. Enabling users to select and access services through the most appropriate access network remains a challenging endeavor. Motivated by the fact that the proposed ANS schemes lack unity, in this paper we have formed a comprehensive list of critical aspects that should be considered when designing an ANS solution. Concurrently, we have revisited current research efforts and identified suitable enabling technologies.

In addition, we claim the suitability of the introduction of cognition and advanced learning capabilities, acting as a catalyst for improving the quality of ANS decisions. Furthermore, careful design of a negotiation process could serve for successful handovers and personalization, while the reputation of network operators should be taken into account in the overall ANS decision process.

References

[1] M. Kassar, B. Kervella, G. Pujolle, "An Overview of Vertical Handover Decision Strategies in Heterogeneous Wireless Networks", Computer Communications, vol. 31, no. 10, 2008, pp. 2607-2620.

[2] F. Bari, V. C. M. Leung, "Automated Network Selection in a Heterogeneous Wireless Network Environment", IEEE Network, vol. 21, no. 1, 2007, pp. 34-40.

[3] P. Vidales, J. Baliosian, J. Serrat, G. Mapp, F. Stajano, A. Hopper, "Autonomic System for Mobility Support in 4G Networks", IEEE Journal on Selected Areas in Communications, vol. 23, no. 12, 2005, pp. 2288-2304.

[4] Q. Song, A. Jamalipour, "A Quality of Service Negotiation-based Vertical Handoff Decision Scheme in Heterogeneous Wireless Systems", European Journal of Operational Research, vol. 191, no. 3, 2008, pp. 1059-1074.

[5] E. Bircher, T. Braun, "An Agent-Based Architecture for Service Discovery and Negotiations in Wireless Networks", in Proc. of the 2nd Int. Conf. Wired/Wireless Internet Communications, 2004, LNCS 2957, pp. 295–306.

[6] Q.-T. Nguyen-Vuong, N. Agoulmine, Y. Ghamri-Doudane, "A User-centric and Context-aware Solution to Interface Management and Access Network Selection in Heterogeneous Wireless Environments", Computer Networks, vol. 52, no. 18, 2008, pp. 3358-3372.

[7] M. Zekri, B. Jouaber, D. Zeghlache, "On the Use of Network QoS Reputation for Vertical Handover Decision Making", in Proc. of the IEEE Globecom Workshop on Advances in Communications and Networks, 2010, pp. 2006-2011.

[8] M. Louta, P. Zournatzis, S. Kraounakis, P. Sarigiannidis, I. Demetropoulos, "Towards Realization of the ABC Vision: a Comparative Survey of Access Network Selection", in Proc. of the IEEE Int. Symp. Computers and Communications, 2011, pp. 472-477.

[9] J. Marquez-Barja, C.T. Calafate, J-C. Cano, P. Manzoni, "An Overview of Vertical Handover Techniques: Algorithms, Protocols and Tools", Computer Communications, vol. 34, no. 8, 2011, pp. 985-997.

[10] X. Yan, Y. A. Sekercioglu, S. Narayanan, "A Survey of Vertical Handover Decision Algorithms in Fourth Generation Heterogeneous Wireless Networks", Computer Networks, Vol. 54, 2010, pp. 1848-1863.

[11] P. Demestichas, V. Stavroulaki, D. Boscovic, A. Lee, J. Strassner, "m@Angel: Autonomic Management Platform for Seamless Cognitive Connectivity to the Mobile Internet", IEEE Communications Magazine, vol. 44, no. 6, 2006, pp. 118–127.

.