MCDM Method and Semantics Integration for Unified Communication Management and Control

Sajjad Ali Musthaq
Computer Engineering Dept.
Telecom Bretagne Brest France
Email: sajjad.mushtaq@telecom-bretagne.eu

Christophe Lohr
Computer Engineering Dept.
Telecom Bretagne Brest France
Email: christophe.lohr@telecom-bretagne.eu

Annie Gravey
Computer Engineering Dept.
Telecom Bretagne Brest France
Email: annie.gravey@telecom-bretagne.eu

Abstract—A dynamic decision making framework for multihomed infrastructure with convergence, unification and heterogeneity orientation is presented. Diversity and dimensionality of the multifaceted service, control, transport and access planes along with multiple objectives over the platform has led us to extend a MCDM method (Grey Relational Analysis). Extended method is then integrated with ontology for capturing semantics and dynamic variations. MCDM helps in achieving the multiple objectives on the basis of multi-disciplinary information domains. The ontology, on the other hand, overcomes the semantic matching and reasoning issues using conventional object-oriented techniques with the help of inference engine. A simple use-case involving the dynamic decision-based routing at private-public network border is presented to validate the proposed solution. The system gives higher throughput and lower call dropping probability with an add-on susceptible delay offering good QoS.

I. INTRODUCTION

Companies nowadays are approaching towards an all Internet Protocol (IP) paradigm by subscribing different access technology links from several service providers for reliability, redundancy and availability while providing good Quality of Service (QoS). Diversity and dimensionality of the company’s service, control, transport and network planes along with multiple objectives over the multihomed platform brings forth the challenge of devising a framework for such converged hybrid infrastructure. Unified Communication over such a platform promises an IP solution. Access technology convergence with heterogeneous network infrastructure, service unification, applications convergence with centralized control and management procedures are the main driving forces for stepping forward towards this unification and an all IP goal. The convergence of voice, video and data services onto the company’s network demands that the underlying infrastructure deliver high performance with resilience. Moreover, the infrastructure must provide effective QoS to ensure that latency sensitive traffic (e.g. multimedia traffic) is getting enough resources it requires even in the face of congestion.

Network Management and control has always been a challenging task over such multihomed and converged environments. Traffic mix, heterogeneous wireline/wireless access technology, fixed-to-mobile and vice versa convergence difficulties, versatile data representation, diverse QoS requirements and variations/dynamics over the network are some of the main complexities involved for network administration and control. The underlying problem becomes more complicated when rule based network management is desired over a platform providing promising QoS and Quality of Experience (QoE). QoS-centered architecture shown in Fig. 1 integrates multimedia, triple-play, quadruple and data services all together over the platform. These services are being offered by different service providers requiring diverse resources with varying set of QoS parameters. It merges different services, binds versatile control interfaces and combines a number of signaling/control/access/transport protocols in addition to unification of access technologies. Companies providing diverse services over the platform are linked to the rest of the world via different network accesses (links) and users of the platform may even access the services in a nomadic manner. It is therefore necessary to enable the company to implement policy rules for optimizing the utilization of its different access links while targeting the private-public network border traffic management issues.

Application, control and network information are to be shared within those distinct domains and across these planes for network management. The attributes and parameters related to resources, profiles, services and dynamics taking place over those planes are disseminated and exchanged via different interfaces and protocols. The information sharing and exploitation process becomes more delicate when its control and management has to be performed under the dictation of certain set of pre-configured rules. There are some frameworks [1], [2], [3] for rule-based management and control but they are static and/or semi-dynamic and are limited to device level configuration and control with different granularity. Moreover, in heterogeneous unified communication systems, there are vendor-specific concepts and implementation dependencies with different means of information representation and processing. These systems take into account a few information among the set of available parameters over the platform, while computing the rules (service profile, reliability information, time of the day, business objectives of the company, latest state of the links, user profiles and Service Level Agreement (SLA)). Moreover, a system capable of taking into account Service Level Specifications (SLS), e.g. susceptible delay, jitter and packet loss may not accommodate the technology specific information. Systems considering user, service and
QoS profiles do not compensate for dynamic context of the request. The multi-facet information stated above for rule-based network management and control comes from different sources with different dimensions, hence formulating a multi-criteria problem. The first challenge is to utilize the available information over the platform maximally so that the network management reflects dynamic control and effective resource utilization with good QoS and QoE. Moreover, the objective is to capture the semantics among these distinct domains (inter-domain and intra-domain) with all the dynamics and variations in order to avoid the conflicts and overlapping regarding the policies, resources and services over the platform. Finally, the use of existing technologies (e.g. Network Address Translation (NAT)ing, Domain Name Service (DNS) Cycling, Hashing, Proxying etc.) without introducing overheads in the protocol stack for policy enforcement is another aim targeted over the unified converged platform.

A framework for rule based network control and management has been presented. Routing rules computation (taking into account all the knowledge from distinct information domains and the dynamics taking place over the presented platform) at private-public multihomed network border is focused on while emphasizing the multimedia communication (voice, video). Multi Criteria Decision Making (MCDM) theory has been applied to handle the multidimensional problem with multiple objectives in addition to varying set of attributes and parameters. Ontologies [4] are integrated with MCDM paradigm to capture the inter and intra domain semantics and dynamics. The reasoner enables hypothesis generation using the facts, meanings and feedback captured from context, the information base and environmental conditional parameters. Reasoner can work with ontology space via the semantics fusion/normalization interface while being completely independent of their abstraction level using embedded axioms. The rest of the paper is organized as follows: In the following Section, platform’s architecture is described. Section III elaborates the system model, MCDM theory, ontology and their integration for the application of dynamic routing. In Section IV, the validity of the proposed solution is presented with simple scenario including an example. Section V includes the related work. Finally, Section VI, outlines the concluding remarks.

II. SYSTEM ARCHITECTURE

Multihoming oriented architecture shown in Fig. 1 provides a cost effective unified communication platform offering versatile access technology convergence while highlighting network traffic unification framework emphasizing private-public network border traffic management issues. It integrates devices and modules from different vendors over a single platform while offering diverse services for public

![Proposed Architecture](image-url)
and private (local) networks. The global objective is the accommodation of dynamic modifications/variations into the policy-computation (decision-making) criteria for control and management of a multihomed platform by using enhanced general methods/techniques and protocols. Service, control, network/transport and routing issues posing a multi-criteria problem are handled together without affecting the standard mechanisms and classical layered approach. The underlying platform stems from competitiveness cluster for handling dynamic routing (at higher layers e.g. session/application layer routing) and management issues at the network border in either direction (inbound and outbound traffic).

Policy Server (PS) is the main controller in the proposed architecture. It computes all the rules (decisions) by taking into account the static configurations and dynamics taking place over the platform, in addition to the policy enforcement supervision. The proposed dynamic decision engine partly constitutes the core of PS.

Session Border Controller (SBC) in the offered framework is primarily dedicated to multimedia communication. It provides a number of vendor specific functionalities depending on the requirements and its deployment. In addition to SBC’s standard functionalities, it is tweaked to act as a Policy Enforcement Point (PEP) in the proposed architecture.

Call Server (CS: an IP-based PBX/Softswitch) supports proxy, registrar, redirect and location services. CS here provides registration, user profile management and service control mechanism. It is modified to handle the user profile based Call Admission Control (CAC) functionality. Components of this platform (Fig. 1) are provided by partners: the platform’s service and application plane is realized by modules from Alcatel-Lucent whereas SBC and PS are developed and tweaked by two different teams at TELECOM Bretagne, Brest. For detailed functionality, information sharing and communication between different devices over the presented architecture, the reader is referred to [5], [6]. Data server offers conventional web, FTP and mailing services (irrespective of access protocols and technology) over the normative unified architecture. The protocol chosen to communicate the information/rules between PDP and PEP is Diameter with newly defined and developed Attribute Value Pairs (AVPs). SNMP trap/alarm-based information in our system is exploited to gauge the QoS parameters of access router interfaces. This paper addresses the policy (decision) based control and management in a highly dynamic, multidisciplinary and multivariate information domains (planes/layers) which are synchronized locally and asynchronous globally (Globally Asynchronous Locally Synchronous (GALS) system).

III. PROBLEM FORMULATION AND SYSTEM MODEL

The proposed system promises to mitigate the cumulative multi-facet effects due to disparate diversity, myriad dimensionality with multiple objectives within GALS. The information sources representing different domains (e.g. service, control, transport planes etc.) might be highly structured and synchronized locally (inter-domain/plane) but may have higher probability of asynchronous and un-structured information representation globally (intra-domains/planes). QoS profile of the links, user authentication/authorization profiles, service variants/profiles, business objectives of the company, fluent dynamics over the multihomed platform and traffic management issues at private-public network border constitutes a multi-disciplinary problem. The information coming from different sources with different dimensions reflects the complexity of the underlying problem when a single rule/decision has to be computed on the basis of multi-dimensional and multi-disciplinary information. Moreover, the situation becomes more mingled when multiple objectives have to be targeted in addition to the varying inter/intra-domain semantics that have to be captured for effective and efficient rule-based control and management.

Conventional solutions to handle such scenarios in multihomed converged environment are either user-centric or motivated for efficient resource utilization over the platform and/or they are centered towards application optimization for desired QoS. However, to cope with all these multi-criteria goals and objectives, MCDM is used. The choice of the technique and its impact on the decision-making is not within the scope of this work and the reader is referred to [7] for an overview of this particular domain. Ontologies are used for representing the individual domain knowledge with semantics. These ontologies formally and explicitly specify the conceptualization via terms and relationships. The domains, presenting the knowledge base are complex and might have large number of sub-domains with frequent variations and dynamics. This complexity is handled by using the fusion/normalization technique (via inference engine). The ontologies with multiple inheritance are managed with automated reasoner using a set of axioms. Moreover, the ontologies are representation of relationships between classes and/or instances within a particular domain and as such, they cannot provide the knowledge based reasoning and feedback mechanism. The reasoner accommodates the feedback mechanism. Interoperability, scalability and extensibility while targeting multiple objectives and goals over the platform is addressed by integrating MCDM with ontology using conventional object oriented techniques.
handling the multi-criteria and multidimensional information coming from different sources with multiple objectives. The underlying system precisely illustrates different components and their interactions. However, the aim here is to highlight the use of ontology in coordination with MCDM over the unified platform for capturing the semantic variations alongside multiple criteria and versatile objectives.

A. Multi Criteria Decision Theory and Ontology Integration for Decision Engine

We focus on multimedia services over the multihomed unified platform due its QoS demands and control. Moreover, it is not possible to present an ontology on a piece of paper totally representing a particular domain. So a simple use-case is presented to understand the ontology and MCDM integration due to space limitation. It is worthwhile to mention here that Session Initiation Protocol (SIP)-based voice communication routing over the platform for rule (decision) computation and its enforcement is emphasized. SIP is a request response signaling protocol. Its infrastructure is highly open and flexible facilitating the services [8]. It allows the creation, modification and termination of service sessions independently of the underlying data-link layer technologies and transport protocols.

Two user Bob (local) and Alice (remote) are supposed to communicate (outgoing SIP communication) with each other by using the available resources on the platform as shown in Fig. 1. Fig. 3 illustrates the hierarchy of the desired goal, the criteria, sub-criteria and the available alternatives. Bob initiates the communication and sends an initial INVITE to the CS. Resource priority tag indicating the user profile is added and the request is forwarded to the SBC. Communication type is evaluated at SBC and this information (resource priority tag, communication type) is sent to the PS. Decision engine within the PS maps the corresponding profiles from the central profile base and then generates a trigger and takes the snapshot of the platform’s information for further processing. Particular context overrides the specific instance of the knowledge base (reliant on the semantics fusion/normalization unit which in turn is dependent on inference engine) that embody the overloaded ontology space (containing concepts, classes relations and instances) for the ongoing context under the influence of embedded axioms and pre-defined set of rules. MCDM in coordination with ontology capture the knowledge relating to the service/control/transport planes, diverse network provider’s capabilities, environmental constraints, business goals and dynamic variations over the underlying network infrastructure.

A glimpse of service, QoS and SLA ontologies is shown in Fig. 4. Relevant goal, criteria and sub-criteria are adapted adequately and the utility function of the MCDM method is formed. Weights of the corresponding attributes (representing the criteria and sub-criteria) are computed keeping in view the context and the business rules in consent with the reasoner.

B. Grey Relational Analysis (GRA)

GRA is a decision-making technique that is based on grey system theory. Originally developed by Deng [9], Grey theory is widely applied in fields such as systems analysis, data processing, modeling and prediction, as well as control and decision-making. It is an effective mathematical means to deal with systems characterized by conflicting and partial information. Grey relation refers to the uncertain relations among things, among elements of systems, or among elements and behaviors. Due to its ability to use reference attribute vector, it is being applied in the proposed decision-making system in outsourcing mode. Moreover, the platform’s latest conditions and the context of the request are taken into account while constructing the reference vector.

To avoid the stringent mathematics, 6 attributes are chosen for the application of MCDM method (in coordination with ontology) on 4 alternative links over the multihomed platform shown in Fig. 1. SIP-based voice call will be routed to one of the 4 alternative links by using the proposed decision engine. \( L_1, L_2, L_3 \) and \( L_4 \) are the four links and \( UR, D, J, PL, TB \) and \( AB \) are 6 attributes representing the voice/video Utilization Ratio, Delay, Jitter, Packet Loss, Total Bandwidth.
The values of these attributes are obtained from the SNMP traps and the Service Level Agreements (SLAs) of the corresponding links over the platform. As the parameters involved in the DM come from different sources, the units representing the values are different. We need to normalize these parameters in order to make them unit-less. The attributes having bigger values are divided by the largest value in the corresponding column vector, while the smaller range attribute (e.g., D is in milliseconds) is divided by the smallest value in the corresponding column vector. The normalized Decision Matrix is given by

\[ DM = \begin{bmatrix} UR_1 & D_1 & J_1 & PL_1 & TB_1 & AB_1 \\ UR_2 & D_2 & J_2 & PL_2 & TB_2 & AB_2 \\ UR_3 & D_3 & J_3 & PL_3 & TB_3 & AB_3 \\ UR_4 & D_4 & J_4 & PL_4 & TB_4 & AB_4 \end{bmatrix} \]

The normalized reference vector is given by

\[ \text{Ref} \left( L \right) = \begin{bmatrix} UR & D & J & PL & TB & AB \end{bmatrix} \]

Now the distance between the corresponding normalized reference vector entities and the normalized Decision Matrix entities is calculated as follows:

\[ \Delta_{UR_i} = \left| UR_i - UR \right|_i, i = 1, 2, 3, 4 \] (5)

The \( \Delta \) Decision Matrix is obtained by applying Eq. 5 to the corresponding entities in the normalized Decision Matrix and the normalized reference vector:

\[ \Delta_{DM} = \begin{bmatrix} \Delta_{UR_1} & \Delta_{D_1} & \Delta_{J_1} & \Delta_{PL_1} & \Delta_{TB_1} & \Delta_{AB_1} \\ \Delta_{UR_2} & \Delta_{D_2} & \Delta_{J_2} & \Delta_{PL_2} & \Delta_{TB_2} & \Delta_{AB_2} \\ \Delta_{UR_3} & \Delta_{D_3} & \Delta_{J_3} & \Delta_{PL_3} & \Delta_{TB_3} & \Delta_{AB_3} \\ \Delta_{UR_4} & \Delta_{D_4} & \Delta_{J_4} & \Delta_{PL_4} & \Delta_{TB_4} & \Delta_{AB_4} \end{bmatrix} \] (6)

Grey Relation Coefficients (GRCs) representing the measurement of similarity of an attribute to its reference are calculated (e.g. for voice/video Utilization Ratio of a link \( UR \)) as follows:

\[ GRC_{UR_i} = \frac{\Delta_{\min} + \alpha \Delta_{\max}}{\Delta_{UR_i} + \alpha \Delta_{\max}}, i = 1, 2, 3, 4 \] (7)

where \( \alpha \in [0, 1] \); \( \Delta_{\min} \) and \( \Delta_{\max} \) are calculated as follows:

\[ \Delta_{\max} = \max \left( \Delta_{UR_1} + \Delta_{D_1} + \Delta_{J_1} + \Delta_{PL_1} + \Delta_{TB_1} + \Delta_{AB_1} \right) \]

\[ \Delta_{\min} = \min \left( \Delta_{UR_1} + \Delta_{D_1} + \Delta_{J_1} + \Delta_{PL_1} + \Delta_{TB_1} + \Delta_{AB_1} \right) \] (8)

As we are emphasizing on voice communication (outbound calls) and to meet the QoS requirements of voice, (Delay and Jitter are given more weight), we choose the weights corresponding to each attribute in the Decision Matrix. The available bandwidth is coupled with user profile loaded from the profile base (in case of gold profile, it is highly desirable to choose the link with good available bandwidth so \( AB \) and \( U \) will also be given suitable weight values). These assigned weights illustrate the relative importance of each attribute in Decision Matrix such that:

\[ W = W_{UR} + W_D + W_J + W_{PL} + W_{TB} + W_{AB} = 1 \] (10)

The weighted GRC coefficient representing an attribute column is given by:

\[ GRC_{wUR_i} = W_{UR} \times GRC_{UR_i}, i = 1, 2, 3, 4 \] (11)

The resulting weighted GRC matrix is given by equation 12. The GRC value for individual link is calculated as follows:

\[ Coef \left( GRC \right)_i = GRC_{wUR_i} + GRC_{wD}, \; GRC_{wJ}, \; GRC_{wPL_i}, \; GRC_{wTB}, \; GRC_{wAB_i}, i = 1, 2, 3, 4. \] (13)

The Candidate Link with the highest GRC coefficient value is the final decision, i.e., the best link for the request. The links \( L_1, L_2, L_3 \) and \( L_4 \) characterized by attributes \( UR, D, J, PL, TB \) and \( AB \) are represented by the values shown in Table I. \( D, J \) and \( PL \) are given higher weights due to voice call (outgoing) while keeping in view the required bandwidth judged from the codec negotiated during the call setup. For the application of GRA on the links represented by the corresponding row vectors in Table I, all the steps are gone through in the order stated above in this section. The links are ranked with GRC values as mentioned in Table II.

<table>
<thead>
<tr>
<th>UR</th>
<th>D</th>
<th>J</th>
<th>PL</th>
<th>TB</th>
<th>AB</th>
</tr>
</thead>
<tbody>
<tr>
<td>%age</td>
<td>Milliseconds (ms)</td>
<td>%age</td>
<td>Megabits per second (Mbps)</td>
<td>Mbps</td>
<td></td>
</tr>
<tr>
<td>L_1</td>
<td>49.65</td>
<td>145</td>
<td>21</td>
<td>22</td>
<td>100</td>
</tr>
<tr>
<td>L_2</td>
<td>31.84</td>
<td>130</td>
<td>24</td>
<td>26</td>
<td>100</td>
</tr>
<tr>
<td>L_3</td>
<td>20.81</td>
<td>110</td>
<td>18</td>
<td>13</td>
<td>100</td>
</tr>
<tr>
<td>L_4</td>
<td>25.00</td>
<td>155</td>
<td>32</td>
<td>28</td>
<td>100</td>
</tr>
</tbody>
</table>

**TABLE I**

Links With Corresponding Parametric Values

<table>
<thead>
<tr>
<th>GRC Value</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4624</td>
<td>2</td>
</tr>
<tr>
<td>0.4107</td>
<td>3</td>
</tr>
<tr>
<td>0.8101</td>
<td>1</td>
</tr>
<tr>
<td>0.3981</td>
<td>4</td>
</tr>
</tbody>
</table>

**TABLE II**

GRC Values and the Corresponding Grading of Alternative Links.

**IV. Solution Validation**

Dynamic SIP-based call routing decision-making at private-public network border under the control of Decision Engine is investigated. The validation and proof of the concept is accomplished by using the architecture shown in Fig. 1 with
4 links of 100Mbps each. The system offers the provision of on-the-fly (on line) and off-line decision computation depending on the chosen enforcement mode. SIPp [10] is used to generate extensive SIP requests (INVITE messages). It is a configurable traffic generator and is extensible via a simple XML configuration language. Call model with User Agent Client (UAC) at CS sends an INVITE to softswitch at CS that adds the resource priority tag and forwards the request to SBC. SBC analyzes the request to judge its communication type and pairs it with the resource priority tag sent by the CS. Resource priority type and communication type tuple is mapped to an appropriate user profile from central profile base at PS. It is important to mention here that a random number is generated to send the codec information along with the SIP message. The bandwidth requirement of the call is judged from the codec information and the request is forwarded to an appropriate link (decision can be computed by following the criteria in sections III-A and III-B). Network Address Translation (NAT) is enabled at SBC and the decision is enforced during NAT implementation. Details about the design and development of the parser for embedding the calculated decisions during NATing are avoided due to space limitations. The remote UAC responds with 100 TRYING, 180 RINGING and 200 OK. UAC then sends an ACK and the call is established. The UAC closes the communication after a variable timespan by sending a BYE which is acknowledged by the SIP server with 200 OK. Wireshark [11] is used to capture the traffic at different interfaces (links). OriginLab [12] is used for data analysis from the captured file.

Throughput of each link is plotted with and without decision engine (i.e. using built-in LB in SBC) as shown in Fig. 5. It is observed that there is a significant improvement in the throughput for each link with decision engine while performing SIP-based call routing.

The retransmission mechanism within SIPp is turned off when INVITE messages are sent in order to know that a call has been dropped. The aggregated call dropping probability (for the 4 links shown in Fig. 1 with the proposed decision engine has lowest/lower value in the outsourcing mode (on-line/off-line) than the ordinary SBC’s built-in LB as shown in Fig. 6. The delay introduced by the system with and without Decision Engine is also calculated over the same test bed. The graph shown in Fig. 7 indicates that addition of Decision Engine in the system introduces some overhead (delay). This calculation is performed in outsourcing enforcement mode (online and offline) due to more dynamics involved in that particular mode (for detailed functionality, information sharing and communication between different devices over the presented architecture in outsourcing and provisioning modes respectively, the reader is referred to [5], [6]). The delay increases almost linearly as the number of calls/requests increases and is small enough having very little impact on services. It is due to the fact that the decisions are being executed and enforced during call/connection setup time. It has the highest/higher value in the outsourcing mode (on-line/off-line) than the ordinary built-in SBC’s LB.
Fig. 7. Delay (Milliseconds) Introduced by the System Without and With Decision Engine (Outsourcing Mode (Online and Off-line)).

V. RELATED WORK

MCDM [13], [14], [15] and ontology [16], [17], [18] has been used for rule-based network management and control independently. To the best of our knowledge, there is no reported work that uses MCDM in coordination with ontology for dynamic network management. There are commercial and proprietary solutions available for multimedia traffic management at higher layers (above OSI transport layer). Publicly available information does not reveal the decision-making mechanism and the LB algorithms. The core design and lower-level functionality are hidden because of commercial implications. However some vendors provide Software Development Kit (SDK) for customization of the specific solution with limited interaction and access to the core [19], [20]. Some products offer partial dynamicity with limited controls, while others are enforcing static decisions/rules. F5 networks [21] uses NAT for Load Balancing the SIP traffic to multiple links with static configurations. The proposed solution in this work accommodates the dynamic behavior of the platform and the context with the provision of provisioning and outsourcing (online and off-line) mode decision-making by integrating MCDM theory and ontology.

VI. CONCLUSIONS

QoS profile of the links, user authentication/authorization profiles, business objectives of the company, reciprocal SLAs with providers, technology specific and technology independent information over converged platform, fluent dynamics over the multihomed platform and traffic management issues at private-public network border constitutes a multi-disciplinary problem with multi-face orientation. The normative scalability of the service, control and network/transport planes in converged network environment cannot be guaranteed without semantic reasoning and introducing the object oriented techniques while modifying/updating and/or adding new datasets. The information coming from different sources with different dimensions reflects the complexity of the underlying problem when a single decision has to be taken on the basis of multi-dimensional and multi-disciplinary information. Ontological engineering in conjunction with MCDM is devised using object oriented reasoning. MCDM theory is used to address the multi-criterion and multidimensionality problem with multiple objectives while the semantic variations and the reasoning are captured by applying conventional object-oriented concepts to ontologies. The proposed solution is tested for dynamic routing decision-making at the private-public network border (SIP-based multimedia traffic). Decisions are computed outsourcing mode (on-line and off-line) in order to validate the dynamism. Existing standards and mechanisms are used for decision enforcement without introducing overheads in the protocol stack. Throughput of the individual links improved significantly where the resources are being used efficiently and effectively at the cost of susceptible delay. Aggregated call dropping probability with the proposed Decision Engine has lower values than the SBC’s built-in load balancer for call routing.

REFERENCES