CHAPTER 2

RELATED WORK AND LITERATURE SURVEY

2.1 INTRODUCTION

The promotional qualities of MAS also amplify the expertise required to complete, expand and authenticate the MAS’s eminence goals and distinction with respect to the projected application domains. These requirements enlarge the research challenges as described in (Bond et al 1988, Katia 1998) in terms of the following aspects: design and development methodologies of MASs; communication and interactions models in the MAS environments; coherent decision making abilities; coordination, collaboration and cooperation issues; performance formulation, validation and justification claims and etc. As claimed in (Katia 1998), it is a major challenge to ensure about the MASs of coherent collective behaviors without unpredictable or harmful behaviors.

Collective performance is a global property of the MAS that could be measured in terms of efficiency, quality and consistency. Several methods for measuring the performance of MASs in these aspects have been proposed. Although many researchers have originated to investigate over the performance assessment schemes of the MASs, but the orientation towards metrics based approach is limited and of course no proper classification schemes are provided in this regard. This is due to two primary reasons: firstly, due the non-deterministic characteristics of MASs at the environment level and inherent complexity at the application level (Marc et al 2004);
secondly, due the complicated validation procedures of metrics based approaches.

In the former case, it is very well proved and realized that developing suitable metrics for quantifying the MAS performance attributes is a challenging issue, particularly for variable and uncertain characteristics of MASs. It is obvious in the real time environments as nobody can afford to create one system of best solution.

In the later case, for any kind of metrics based proposal, first the proposed metrics itself validated by using the standards of the same and the applicability or usability of the proposed metrics should be validated with respect to the anticipated domain. Although it is claimed that the metrics based evaluation models are the cheapest analytical way to settle the preferable environments, the process of exploring the correlation between the measurements done during tests and calculations based on the estimated metrics is so intricate (Dariusz et al 2008). It needs to conduct significant number of experiments to confirm about the measurements and interpretations. This two tire validation schemes complicates the intensity of metrics based proposals.

In spite of these decisive issues, there are considerable proposals concerned with metrics based MAS performance evaluation. At the course grained level, these works can be classified into two categories; framework based MAS performance evaluation models and application based MAS performance evaluation models. The following sections review the related works in the respective categories.
2.2 METRICS BASED MAS PERFORMANCE EVALUATION

2.2.1 MAS Framework Based Performance Evaluation

This category of evaluation models concentrates on MAS frameworks in isolation with the application environments. Most of the times, these evaluations may be served as blueprints and lead to make adequate compromisation(s) with respect to the applications. However, the same can be served as decision support tools to choose the better models for the planned applications. The framework based MAS performance evaluation schemes can be viewed in two different perspectives; methodology-specific performance evaluation and bespoke model-specific evaluation, which are discussed and reviewed as follows.

Methodology specific evaluation schemes: In the methodology specific evaluation points of view, as in case of with any new software engineering paradigm, the guidelines for MAS performance evaluation schemes are also provided along with the methodology proposals. Towards this orientation, a number of acceptable standard Agent Oriented Software Engineering (AOSE) methodologies have been proposed to define general and bespoke methodologies and offer guidelines for designing, developing and evaluating the Agent and Multi Agent Systems: GAIA – Generic Architecture for Information Availability) (Wooldridge et al 1999); MaSE - Multi-agent Systems Engineering (DeLoach 1999, Wood et al 2000); SODA - Societies in Open and Distributed Agent spaces (Omicini 2000); Tropos (Bresciani et al 2001); Prometheus (Lin et al 2002, Lin et al 2003); MESSAGE (Caire et al 2002).

Although these methodologies use different terminologies and provide different abstractions, all are recognized that the process of building MASs is radically different from the process of building other software systems.
(Franco et al 2003). In particular, they approved the thought that the MAS can be built in terms of an organized society of individuals in which each agent plays specific roles and interacts with other agents according to the protocols determined by the roles of the involved agents (Franco et al 2003). In this point of view, it is also explored that an organization is more than simply a collection of roles and that in order to effectively build a MAS in organizational terms, further organization-oriented abstractions are need to be devised in the context of a AOSE methodologies (Franco et al 2003).

These above claims are also expressed by Quynh-Nhu et al (2003, 2004), through comprehensive multi-dimensional feature analysis frameworks, which have been provided as generic performance assessment models for MAS methodologies. To improve the relevance of the proposals, the evaluation criteria adopted are designed such that they are representative, case-generic and centered on the capabilities and usefulness of the individual methodologies. The significance of the frameworks has been improved by adding several evaluation criteria that are not accounted in other evaluation frameworks.

Jan et al (2005) suggested a framework to examine the gap between the modeling methodologies and platforms. The evaluations have been carried out to verify the appropriateness of methodologies with respect to the platforms in terms of relevant criteria. In this view, Arnon et al (2004) offered a comprehensive model for evaluating and comparing the performance of the AOSE methodologies. This model focused on four major and generic aspects of a methodology: concepts and properties; notations and modeling techniques; process; and pragmatics. GAIA methodology has been considered as the reference model of assessment to claim the merits of the proposed model. Onn et al (2001) presented an evaluation scheme for evaluating the MAS modeling techniques with respect to the aspects of distribution,
concurrency, testing and communication richness. Online auction has been chosen as the application domain for those evaluations. Luca et al (2002) proposed a framework to carry out an analysis over the agent-oriented analysis and design modeling methods. The attributes are classified systematically into three categories; internal attributes, interaction attributes and process attributes. It took qualitative evaluation criteria employing quantitative methods into consideration, particularly Goal-Question-Metric (GQM) method. These quantitative outcomes are the specialty of this approach and these also offered some new directions of research in the same field.

In summary, all these methodologies offer realistic recommendations for basic Agent and MAS concepts such as autonomy, goal oriented responsibilities, pro-activeness, re-activeness, collaboration, cooperation, negotiation and coordination along with qualitative performance evaluation schemes for the above concepts. On the other hand, several important concerns are not addressed in these methodologies, particularly for MASs, such as whether these methodologies can be applied for analysis and design of non-agent based systems, testing, debugging, quality assurance, estimating guidelines and quantification performance evaluation schemes (Wagner 2000, Khanh et al 2003, Dhavachelvan et al 2005, 2006).

**Model-specific evaluation schemes:** In the model-specific evaluation points of view, there are considerable researches have been carried out and appropriate recommendations are made to enhance the acceptability of the outcomes of those researches. Evangelin et al (2006) suggested a methodology that brought the idea of SPE (Software Performance Engineering) to evaluate the performance of MASs. Unified Modeling Language (UML) is used to model the performance of MAS and a general algorithm has been proposed to assess the performance of MASs, early in life cycles, in the software engineering
perspective. Though it was proved that the proposed approach performed better than the Object Modeling Technique (OMT) (Evangelin et al 2006), there were no guidelines suggested for performance quantification, particularly for MASs. But, this work described the necessity of evaluating the MASs at various stages. An architecture based evaluation scheme has been proposed by Gaku et al (1999) and it was developed particularly for evaluating the interactions between the search agents. Explicit characterizations are offered to define the related parameters; however quantification schemes were not presented for the same.

In (Paul et al 2003, Paul et al 2005, Paul et al 2006), a series of MAS architectural evaluation works have been carried out. In these proposals, MAS architectural styles are characterized according to specific set of properties, such as, types of controls used and the types of coordination schemes used. It is claimed that different architectural styles support different quality attributes to different extents and it is important to evaluate them in accordance to the quality attributes relevant to that application. The MAS architectures are evaluated in several dimensions, both domain dependent and independent performance-related attributes. Six important attributes related to dynamic and distributed resource allocation were identified and corresponding measurement schemes were instantiated for each of these attributes. These instantiations are studied in simulation experiments and measurements of the metrics are recorded. The measurements are then analyzed using the Analytic Hierarchy Process (AHP), which is a basic approach to select the most suitable alternative from number of alternatives evaluated with respect to corresponding criteria. The described Z-score method enables to compare the candidate architectures in terms of number of different criteria in the quantitative fashion.
A different approach for automatic design evaluation of MASs based on HLA (High-Level Architecture) federations and genetic algorithms has been suggested in (Sajal et al 2001). MAS design space dimensions that include number of each kind of agents, kinds of equipment carried out by each kind of agent and spatial formation of agents are specifically addressed. Hyuckchul et al (2002) proposed a design based evaluation model focused on cooperation and coordination aspects of MAS, in which MAS performance is modeled using POMDP (Partially Observable Markov Decision Process) in the context of DCSPs (Distributed Constraint Satisfaction Problems). The evaluation scope is limited to cooperation and coordination, but there was no clear distinction between these two facets.

Jose et al (2007) recommended a different approach in the series of design based evaluation models concentrated on communication and coordination aspects of MAS. It belongs to the assessment category of model checking, in which the abstract MAS architecture is modeled as a discrete-event system using Petri nets and structural analysis. Deadlock avoidance in the multi-agent system is considered as a key property in this work, and it is evaluated using liveliveness and boundedness properties of the Petri net model.

MAS structure based performance evaluation model using software metrics was proposed by Dariusz et al (2008). The focus is limited to the structure of MAS and the corresponding connection (communication) oriented metrics were proposed. Another design based evaluation model as a recent evolution in the series of MAS performance evaluation schemes has been suggested by Qin et al (2009). It belongs to the assessment category of theorem proving, which used the Vicsek’s Model to formulate and evaluate the components of MASs. This approach offered an understanding mechanism of MAS in terms of phase transition in a group of self-driven particles and the consensus protocols.
Glenn et al (2005) presented a model to evaluate the performance of MAS in terms of dynamic runtime properties at on-demand agent generation scenarios. This work proved that the required number of agents generated can be reduced as long as the plan structure is not linear and fans out. An MAS evaluation model from the software engineering perspective has been proposed by Caire et al (2002). The aim of the model is to cope with important issues connected to the transition from the design phase to the implementation phase. No quantification mechanism is being proposed to assess the performance of MAS.

Regis et al (2001) described a framework, which is suitable environmental space for evaluating the coordination and adaptive qualities of MASs w.r.t the environmental space. Michael et al (2001) suggested a set of guidelines for evaluating the coordination aspects of MASs. The coordination ability of MAS has been discussed in terms of delegations point of view; social delegation and task delegation. Though there is no guidelines were given for quantifying the delegation aspects, but it provided a distinct and useful idea for deriving new measurement schemes for quantifying the coordination (delegation) abilities of MASs.

In summary, all the reviewed methodologies succeeded in offering guidelines for MAS performance quantification engineering to some extent, but they suffered with Babylonic mismatch between the proposed evaluation schemes and the intended objectives. In addition to that, only very few classified the performance attributes and provided appropriate measurement schemes, but the scope of coverage is so limited in these respects. From these perspectives, the proposed research work suggests an enhanced model for evaluating the MASs at various stages by using appropriate quantification mechanisms. This proposal is aimed to offer a set of application-independent
2.2.2 MAS Application Based Performance Evaluation

This type of MAS performance evaluation methods focuses on MAS based applications’ perspective. These evaluations may be served as the preferable mechanisms, but suffered with domain dependent issues. A method has been proposed by Gaku et al (1999) for evaluating the interaction between the search agents. Here the interaction protocols are considered for evaluating the interactions between the agents within the framework. In spite of qualitative descriptions, quantification schemes were not suggested to evaluate the intended attributes.

In the simulation point of view, a performance assessment model for MAS has been proposed by Robert et al (2000). This proposal is concerned with evaluating the performance of the member agents of MAS, particularly the multi-agent control approach only and for this, an experimental set up has been suggested for the same of the application specific MAS. Arne et al (2008) offered mere level qualitative guidelines for evaluating the MAS’s performance, particularly for MAS communicational issues with respect to the temporal aspects. A simulation middleware has been developed for assessing the performance of the MASs communicational issues with reference to the FIPA standard. Though there are no proper guidelines for quantifying the specified attributes, these guidelines can be served as a directional tool for developing appropriate quantification schemes for measuring the anticipated performance attributes.

A set of unique schemes of evaluation have been presented to compare and validate the performance of the multi agent platforms (Raquel et al 2007,
Anders 2004). The former one provided qualitative comparisons intended to evaluate the mobile agent platforms and the later one intended to evaluate the ubiquitous agent platforms in different scenarios and these models can help the developers to choose the appropriate platform suites based on the requirements. In (Bora 2001), a framework has been proposed to evaluate the performance of the MAS, particularly transaction management in a distributed environment. A test bed has been constructed such that to offer a homogeneous environment for testing different DDBM protocols on heterogeneous DBMS environments. The test bed offered flexibility by allowing the changes in the parameters of current protocols, as well as exchanging the protocols wholly at run-time.

Information consensus based evaluation schemes are the recent approaches for evaluating the performance of high-dimensional MAS environments mainly focused on the projected applications (Xiao et al 2007, Qin et al 2009). Through these works, consensus problems for multi-variable agent systems are studied and the structures of consensus functions in a MAS environment have been characterized and a set of necessary and sufficient conditions for a MAS to solve a consensus problem have been presented. These approaches investigated and proved that the information consensus or agreement upon certain quantities of interest is one of the important problems encountered in decentralized control of multi-agent systems and offered a different idea in assessing the MAS performance compared to its predecessors.

A Record/Replay approach was suggested to evaluate the performance of MAS based on the application (Marc et al 2004). It is a hybrid model that evaluates the performance of MAS in two phases; traces the non-deterministic events at first and then to analyze the deterministic events. It supports the post-mortem analysis and suitable for evaluating predefined scenarios only. In
the series of run time environment evaluation, tool architecture has been proposed to verify the properties of MAS based on testing the runtime behavior of the agents in a specified application environment (Denis et al 2006).

An another approach of debugging mechanism for MAS has been proposed by David et al (2003), which used the protocol specifications at runtime for reporting any discrepancies in interactions compared to the predefined sets. These works described and categorized a range of bugs found in actual multi-agent systems and made to be compared with the intended systems’ behaviors. Johan et al (2005) presented an evaluation mechanism based on the significances of information theory – information fusion and information market, particularly for coordination issues in the MAS environments. Nathan et al (2005) proposed an approach that combined queuing analysis and application-level control to engineer a generic framework that is capable of self-optimizing its domain-specific utility to assure application-level survivability. The scope of these works limited with only coordination and communicational aspects of MAS.

In summary, all the reviewed approaches provide a reasonable support for basic performance evaluation over the MASs with distinct and efficient concerns. Irrespective of these supports and recommendations, the scope for quantification procedures is so limited in all the discussed categories. These restrictions are discussed and investigated in (Ruben et al 2004). It is also observed that for any kind of evaluation proposals, at least one suitable application domain should be chosen to prove the claims of the proposals. From these perspectives, the work reported in this thesis, recommended an improved model for evaluating the performance attributes of MASs at different junctures by applying suitable measurement theory and the proposed claims are proved with respect to an appropriate application.
2.3 SUMMARY

All the reviewed methodologies succeeded in offering guidelines for MAS performance quantification engineering to some extent, but they suffered with Babylonic mismatch between the proposed evaluation schemes and the anticipated objectives. Most of the works in the related category used only a particular set of agents either reactive or deliberative that composed the MAS environments. Very few used the hybrid agents that are encouraged to formulate generic assessment models rather than the bespoke recommendations. In case of performance quantification, the derivatives of measurement theories are not proper, which are to be reviewed and enhanced appropriately.

In addition to that, though there are considerable works and standards proposed in the domain of metrics based MAS performance evaluations, but the scope of coverage is so limited with respect to the specific set of operational issues. At the outset, as discussed earlier, metrics based evaluations are always noteworthy, because if they were confirmed by experiments, it would allow them to be used to valuate many other realizations without actual tests implementations (Dariusz et al 2008).

The primary advantage of the metrics based approaches is to combine any of the core performance evaluation schemes of MAS. This is also applicable for evaluating both deterministic and non-deterministic attributes of MAS by quantifying them with appropriate amendments. The main problem associated with metrics based evaluation is to verify the correlation between measurements done during tests and the performance estimations exclusively based on metrics.
From these perspectives, this work suggests an enhanced model for evaluating the MASs at various stages by using appropriate quantification mechanisms. This proposal is aimed to offer a set of application-independent fine grained quantification schemes, which will act as a generic model for evaluating the performance attributes of MASs. It is also observed that for any kind of evaluation proposals, at least one suitable application domain should be chosen to prove the claims of the proposals. Hence, in this point of view, a suitable application environment is identified to prove the proposed claims and thereby to transfer it into real time environments.