Chapter 4

LITERATURE REVIEW

4.1 Introduction

The objective of this chapter is to review of literature available on semantic web and ants-based algorithms. Extensive research has already been done in the area of semantic web and agent technology. This section highlights the work of distinguished researchers and explores the challenges, which still need to be addressed.

Deneubourg et al. [55] thoroughly investigated the pheromone laying and following behavior of ants. In an experiment known as the “double bridge experiment”, the nest of the colony of ants was connected to a food source by two bridges of equal lengths. In such case, ants start to explore the surroundings of the nest and eventually reach the food source. Along their path between food source and nest, ants deposit pheromone. Initially, each ant randomly chooses one of the two bridges. However, due to random fluctuations, after some time one of the two bridges presents a higher concentration of pheromone than the other and therefore, attracts more ants. As the time the whole colony converges towards the use of same bridge. This colony-level behavior based on autocatalysis is used by ants to find the shortest path between a food source and their nest.

Nwana [49] provided a comprehensive discussion of ants characteristics and typology. The author identified a minimal list of three primary attributes: autonomy, learning, and cooperation. Autonomy refers to the principle that ants can operate on their own without the need for human guidance. For ant-based systems to be truly smart, ants
would have to learn as they react or interact with their external environment. In order to cooperate, agents need to possess a social ability i.e. the ability to interact with other agents and possibly humans via some communication language. Agents have been further distinguished as collaborative agents, interface agents, collaborative learning agents, and smart agents. In addition, agents can sometimes be labeled by their applications such as information agents, reactive agents, mobile agents, and hybrid agents.

Schoonderwoerd et.al [59] designed an ant-based system for telecommunication networks after the discovery of trail laying ability of ants. Author proposed to create mobile agents that are capable of traversing the network themselves and leaving behind information for the agents coming after them. Later ones can gather vital information about the paths and the network as a whole. In work, a simulated network models a typical distribution of calls between nodes; nodes carrying an excess of traffic can become congested causing calls to be lost. The network supports a group of ants that have simple-purposed tasks with no direct communication capability or explicit knowledge of the global goal. The ants move across the network between randomly chosen pairs of nodes as they move they deposit simulated pheromones as a function of two variables: (1) the distance from their source node and (2) the congestion encountered on their journey. They select their path at each intermediate node according to the distribution of simulated pheromones at that node. Calls between nodes are routed as a function of the pheromone distributions at each intermediate node. The ant-based control system was shown to result in fewer call failures than other methods such as fixed shortest-path routing and algorithmic mobile agent (instruction-specific), while exhibiting many attractive features of distributed control.
White et.al [77] described how multiple interacting swarms of adaptive mobile agents could be used to locate faults in networks. The authors proposed the use of mobile agents for fault finding in order to address the issues of client/server approaches to network management and control such as scalability and the difficulties associated with maintaining an accurate view of the network. The authors further defined three principal types of mobile agents: servlets, deglets, and netlets. Servlets are extensions or upgrades to servers that stay resident as integral parts of those servers. Mobile agents constituting servlets are sent from one component to another and are installed as code extensions at the destination component. Deglets are mobile agents that are delegated to perform a specific task and generally migrate within a limited region of the network for a short period of time. For example, undertake a provisioning activity on a network component. Netlets are mobile agents that provide predefined functionality on a permanent basis and circulate within the network continuously. These agents are small, mobile and communication is top-down, instead of peer-based.

Gianni et.al [31] presented AntNet, a novel ACO algorithm applied to the routing problem in connectionless communications networks. In AntNet artificial ants collectively solve the routing problem by a cooperative effort in which stigmergy plays a prominent role. Ants build local models of the network status and adaptive routing tables using indirect and non coordinated communication of information they collect while exploring the network.

Berners-Lee et.al [68] introduced their vision of the semantic web, as an extension of the current web in which information has “well-defined meaning, hence better enabling computers and people to work in cooperation”. Semantic web is a rapidly evolving extension of the current web, where the semantics of information and services is
well-defined, making it possible for people and machines to precisely understand web content. So far, the fundamental semantic web technologies (content representation, ontologies) have been established and researchers are currently focusing their efforts on logic and proofs. Semantic web is intended for knowledge sharing among agents as well as humans. Ontologies play a vital role in achieving the goal of semantic web. Which express knowledge in a certain representation as well as in machine interpretable form were introduced and have grown considerably in number. To have a sustained growth of the semantic web and to have better interoperability between intelligent systems and applications it is highly desirable and is very critical to reuse existing ontologies.

Wu and Aberer [79] used swarm intelligence to create a model for the dynamic interactions between web servers and users which provide relevant feedback by browsing web pages. This model is being used for ranking web documents. A swarm intelligent module was added to the web server architecture. The trail laying feature used is composed of an accumulation feature that increases the pheromone amount when a user visits a page, and of an as reading feature in which pheromone is diffused to the pages that link to a certain page. The spreading feature does not comply with the ant colony optimization meta-heuristic.

Dan Schrage and Paul [18] propose an approach to scheduling that relies on a biologically-inspired optimization algorithm known as Ant Colony Optimization (ACO). This approach is easily combined with agent-based modeling to produce an accurate representation of the search environment. Moreover, it is one of the best known algorithms for solving a number of prioritized path-planning problems. Finally, ACO distinguishes itself from other search algorithms with its ability to adapt to dynamically changing circumstances in the search environment. Thus, not only is it well suited for the
initial planning problem, but it also offers the option of dynamic re-planning of a solution in response to changing battle space conditions.

Schneider [53] specified an abstract syntax and a formal semantics for the Web Ontology Language (OWL). In addition, author not only proposed a model-theoretic semantics in the form of a vocabulary extension to the RDF Semantics but also a mapping mechanism from the abstract syntax to RDF triples was also presented. However, the main focus had only been web ontology language and the work could not be made more generic.

Alcides Calsavara and Glauco Schmidt [5] proposed and defined a novel kind of service for the semantic search engine. A semantic search engine stores semantic information about web resources and is able to solve complex queries, considering as well the context where the web resource is targeted, and how a semantic search engine may be employed in order to permit clients obtain information about commercial products and services, as well as about sellers and service providers which can be hierarchically organized. Semantic search engines may seriously contribute to the development of electronic business applications since it is based on strong theory and widely accepted standards.

Kumar and Reddy [38] proposed an ant colony optimization algorithm for a multipurpose reservoir system. To tailor the ACO algorithm for their problem, a finite time series of inflows, classification of the reservoir volume into several class intervals, and defining the reservoir releases for each time period with respect to a predefined optimality criterion, were established. The ACO algorithm was compared to a real coded Genetic Algorithm (GA). It was shown that the ACO algorithm performed better than the GA.
Alonso et al. [8] describe a hypothetical architecture for the integration of unstructured, semi-structured and structured information from across an enterprise, using a combination of text analytics, XML and semantic web (RDF/OWL) technologies. The key components of the architecture are: an XML content and metadata repository; a semantic crawler and query engine; and a visualization layer. Oracle's support for managing XML content (Oracle) for full-text search and text analytics and for RDF storage, query and inference are proposed for providing the necessary capabilities. A prototype system named GIO is described operating on structured (RDF) content obtained from the DB digital library collection. The emphasis is clearly on making something work in the short term and there are some peculiarities in the way in which the RDF data is handled in the prototype. For example, the RDF/XML documents are not parsed as RDF but are stored and queried as pure XML with an associated OWL ontology being used in an ad hoc way to provide a document folder hierarchy. The capability to store and query native RDF graphs and to merge graph data without joins (which require additional information to be encoded in queries) has not been used. The work falls short of providing a complete vision for seamless enterprise information integration; for example it does not address decentralization, nor does it discuss mappings from XML or relational data into graph-based, ontology-aware data.

Ali et.al [7] used ant colony optimization to accelerate convergence of the Differential Evolution (DE) technique. Their methodology termed as Ant Colony Differential Evolution (ACDE) initializes the ant population using base learning techniques, utilizes a random localization methodology, and simulates the movement of ants to refine the best solution found in each iteration.
Sanchez [27] elaborated the necessity for integrating intelligent agents and semantic web and analyzed the potential benefits of this combination. They proposed a SEMMAS (semantic web services and multi-agent system) framework, which is an ontology based framework for seamlessly integrating intelligent agents & semantic web services.

Viorica et al. [73] presented two ant-inspired methods applied in two phases of the automatic web service composition process, named as the discovery and the selection of the optimal composition phases. To improve the efficiency and accuracy of services discovery process the authors organized the set of available services into service clusters. The clustering method was inspired by the way ants cluster read corpses to build cemeteries and sort larvae. The clustering method is based on a set of metrics for evaluating the similarity and dissimilarity level between two services by considering their semantic descriptions. In the case of composition, this approach combines a service composition graph model with the ant colony optimization metaheuristic to find the optimal composition solution.

Aiden Dipple [4] defines stigmergy as a biological term used when discussing insect or swarm behaviour and describes a model supporting environmental communication separately from artefacts or agents. This phenomenon is demonstrated in the behavior of ants and their food gathering process when following pheromone trails. What is interesting with this mechanism is that highly organized societies are achieved without an apparent management structure. Stigmergic behavior is implicit in the web where the volume of users provides a self-organizing and self-contextualization of content in sites which facilitate collaboration. However, the majority of content is generated by a minority of the web participants. The authors create a model of web
stigmergy, identifying virtual pheromones and their importance in the collaborative process.

Gross et.al [42] considered a variant of the double bridge experiment in which one bridge is significantly longer than the other. In this case, the stochastic fluctuations in the initial choice of a bridge are much reduced and a second mechanism plays an important role. The ants choosing by chance the short bridge are the first to reach the nest. The short bridge receives the pheromone earlier than the longer one and this fact increases the probability that further ants select the shorter path rather than the longer one.

Francesco et.al [26] presented a state of art about ants-based system with particular attention to industrial plants problems. The state-of-art shows that the main fields of application for ACO algorithms are the routing problems but in last few years, different applications have been developed to solve industrial plants problems and in particular, production scheduling problems (job shop, flow shop, open shop). Their work focused on the integration of modeling & simulation and ant theory for supporting real time production scheduling in job shop systems.

Tamma et.al in [72] have explored various issues related to knowledge sharing among agents in distributed & heterogeneous environments. Their work invalidates the assumption that agent posses complete knowledge of the environment at design time due to the dynamism and scale of the environment (i.e. web).

Stuckenschmidt et.al in [32] described an approach for exploiting partially shared ontologies in multi-agent communications by translating private concepts into shared ones while ensuring some formal properties. Their approach enables agents on the WWW to exchange semantic information while relying on internally provided mappings between
ontologies. However, this approach still needs sophisticated communication protocols that agents can use to find out what are the ontologies shared among them and also the options for re-writing queries.

Hung and Yang [28] presented an intelligent search engine with semantic technologies. This research has combined description logic inference system and digital library ontology to complete intelligent search engine.

Gardarin et.al [80] discussed SEWISE is an ontology-based web information system to support retrieval. According to domain ontology, SEWISE can map text information from various web sources into one uniform XML structure and make hidden semantic in text accessible to program. The textual information of interest is automatically extracted by web wrappers from various web sources and then text mining techniques such as categorization and summarization are used to process retrieved textural information.

Kravari et.al [35] describe an intelligent agent as software programs intended to perform tasks more efficiently and with less human intervention. They are considered the most prominent means towards realizing the semantic web vision [Hendler, 2001]. The gradual integration of multi-agent systems (MAS) with semantic web technologies will affect the use of the web in the future; its next generation will consist of groups of intercommunicating agents traversing it and performing complex actions on behalf of users. Thus, intelligent agents are considered to be greatly favored by the interoperability that semantic web technologies aim to achieve. Intelligent agent will often interact with other agents. However, it is unrealistic to expect that all inter-communicating agents will share a common rule or logic representing formalism.
Nicholas et.al [47] investigated the integration of the nascent web services infrastructure with the richer semantics of the semantic web in particular through the use of more expressive languages for service description. In their implementation of service descriptions, the existing web services specifications are more concerned with the signature of services. Such signatures comprise the types of the parameters of the service (typically expressed in terms of XML Schema data types) rather than with any form of ontological classification of the services.

Nigel and Wendy et.al [48] described the evolution of a web that consisted largely of documents for humans to read to the one that included data and information for computers to manipulate. The semantic web is a web of actionable information derived from data through a semantic theory for interpreting the symbols. The semantic theory provides an account of “meaning” in which the logical connection of terms establishes interoperability between systems.

Yuxiao et.al [82] suggested agents are programs acting on behalf of another person, an entity or a process. Intelligent agents are widely known and are useful for application automation and electronic commerce. Hendler [68] indicated that ontology-based intelligent agents could obviously enhance application integration and thus improve e-commerce.

Leger et.al [1] elaborated application areas of semantic web. Semantic web technology is more and more often applied to a large spectrum of applications where domain knowledge is conceptualized and formalized as support for diversified processing operated by machines. Moreover, through subtle joining of human reasoning (cognitive) and mechanical reasoning (logic-based), it is possible for humans and machines to share complementary tasks. Corporate portals and knowledge management, e-commerce, e-
work, healthcare, natural language understanding and automated translation, information search, data and services integration, social networks and collaborative filtering, knowledge mining, etc are areas of applicability of semantic web.

Motta et.al [24] examined the current semantic web applications and introduce seven dimensions for analyzing them. They compare some older and newer systems which they call the next generation of semantic web applications. As result, they concluded that semantic web applications will have to deal with increasing heterogeneity of semantic sources and new web technologies such as social tagging and web services. They focused on usage of ontologies which are discussed as one of the dimensions because it is a key technology for all semantic web applications.

Biswa et.al [62] introduced next generation of semantic web applications. Such applications achieve their tasks by automatically retrieving and exploiting knowledge from the semantic web as a whole. Unlike, early semantic web applications which gathered and engineered knowledge at design time these new applications explore the web to discover ontologies relevant to the task at hand. Because dynamic knowledge reuse replaces the traditional knowledge-acquisition task, they can potentially reduce the application development cost. Additionally, as these applications can use any semantic information available online, they’re not necessarily bound to a particular domain. Any application that wishes to explore large-scale semantics must perform the following tasks: find relevant source and select appropriate knowledge.

Diplomarbeit et.al [20] presented building blocks for evolving existing web services standards into semantic web services technology. Semantic web services are considered to be the next step in the evolution of web services. In addition to web services, semantic web services feature formalized and machine understandable
descriptions of their capabilities which allow them to be understood and processed by automatic algorithms. Such services are envisioned to enable sophisticated tasks such as automated discovery of services by matching specified service requests with a large pool of service providers, automated composition of services to instantiate high-level descriptions of complex tasks by a sequence of calls to simpler services, and service monitoring to evaluate the quality of work provided by services.

Richard et.al [71] have identified some of the major challenges such as: (i) the availability of content, (ii) ontology availability, development and evolution, (iii) scalability, (iv) visualization to reduce information overload, and (v) stability of semantic web languages.

Dorigo et.al [43] described an artificial ant colony capable of solving the traveling salesman problem (TSP). Ants of the artificial colony are able to generate successively shorter feasible tours by using information accumulated in the form of a pheromone trail deposited on the edges of the TSP graph. Simulations demonstrate that the artificial ant colony is capable of generating good solutions to both symmetric and asymmetric instances of the TSP. The existing methods like simulated annealing, neural networks, and evolutionary computation are example of the successful use of a natural metaphor to design an optimization algorithm.

Holland et.al [50] have proposed the Ant-Based Control (ABC) for network management. The approach is destined for load balancing in circuit switched telecommunication networks and is based on regular transmission of ants from all nodes in the network to collect network information. Based on the collected information load balancing is performed.
Maniezzo et.al [75] introduced a routing approach for packet switched telecommunication networks which can be regarded as an extension of the earlier used approaches. The authors distinguished between two different kinds of ants which are regularly transmitted where one of the ants collects information in the network and the other refresh routing information.

White et.al [70] suggested another routing algorithm for circuit switched networks. The approach is based on three classes of ants. The first class collects information, the second class allocates network resources based on the collected information and the third class gets allocated resources free after usage.

Gambardella et.al [30] have proposed an ingenious optimization method. The ant system, which they applied to classical optimization problems, such as the traveling salesman problem, the quadratic assignment problem or the job-shop scheduling problem, with reasonable success: this method, as a general heuristic, can be compared to simulated annealing.

Kroon et.al [37] developed a hierarchical routing system based on ant-based control algorithm. This algorithm is uses the principles that ants living in ant colonies use.

Ridge et.al [23] describes a multi-agent system architecture that would implement an established and successful nature-inspired algorithm, Ant Colony System (ACS) in a parallel, asynchronous and decentralized environment. It suggests how obstacles may be overcome using a pheromone infrastructure and some modifications to the original algorithm. The authors have addressed the two major problems this algorithm. Firstly, the algorithm is highly synchronous. Each ant performs its own tour and local pheromone update in turn. The algorithm waits for all ants to finish this phase before a global
pheromone update is performed. Once the global pheromone update is complete, all ants begin the tour building phase again. Secondly, the global pheromone update requires a centralized comparison of all the ant tours to acquire the global knowledge of the system’s best tour.

Crespo and Molina [10] propose semantic overlay networks. In this approach, the resources at each peer are classified according to the concepts of a classification hierarchy. There is a semantic overlay network for every concept, which a peer joins if it possesses a significant amount of resources related to that concept. The queries are also classified, and each query is sent to those semantic overlay networks that correspond to the concept of the query. If they do not provide enough results, the query will additionally be sent to the semantic overlay networks corresponding to the super-concepts.

Schelthout et.al [39] tested whether the principle of synthetic pheromone can be employed for coordination in distributed agent-oriented environments. Their framework is based on the idea of object spaces known from concurrent computing.

Labella et.al [67] presents a foraging-inspired approach for division of labor between a group of robots performing object search and retrieval. The goal is to distributive determine which agent is suited best for carrying out a certain task using indirect communication only.

Aquín et.al [46] present an overview of the intelligent systems, a semantic web search engine providing various functionalities not only to find and locate ontologies and semantic data online, but also to explore the content of these semantic documents. Beyond the simple facade of a search engine for the semantic web, we show that the
availability of such a component brings new possibilities in terms of developing semantic applications that exploit the content of the semantic web.

Darieby et al. [22] proposed a system of intelligent ants to reduce management of traffic on the network (no bandwidth-exhaustive client/server message exchanges), enable more robust response to problems (timely, intelligent problem-solving behavior of the agents), reduce administrative overhead and cost due to function delegation (agents perform tasks autonomously), and allow low-level problems to be dealt with locally at the network component (reducing both the processing load on the manager computer and the traffic carrying instructions and data regarding the problems).

Ferguson et al. [25] used an economic model for ants in telecommunication networks. According to the authors, computer networks are being used by growing and increasingly heterogeneous set of network components (computers, channels, and users) that have diverse Quality of Service (QoS) requirements. To support the diversity in large distributed networks, the tasks of efficient service provisioning and optimal resource allocation become very complex but can be accomplished via decentralization.

In biology, stigmergy describes a mechanism of indirect communication where the actions of individuals affect the behavior of others (and their own). This communication mechanism describes the apparent cooperative behavior of insects’ nest building and food gathering activities. For example, the food gathering activities of ants are structured around the use of pheromone trails, where the ants are triggered to perform food gathering tasks. To find the most recent and relevant food source the ants follow particular paths based on the strength of any given trail. The interesting communication here is not only the explicit signal in the pheromone (to gather food) but the implicit signal through the level of decay: information within the trails themselves show which
trail will currently lead to a food source opposed to trails leading to a depleted food source. Combining bio-inspired designs and algorithms based on stigmergy with social network analysis might facilitate the creation of a more sophisticated web application.

4.2 Conclusions

This chapter provided the review of literature available on semantic web and ants-based algorithms. Extensive research has already been done in the area of semantic web and agent technology. This chapter highlighted the work of distinguished researchers and explores the challenges, which still need to be addressed.