Applied Soft Computing 38 (2016) 321–328

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Dynamic fuzzy logic and reinforcement learning for adaptive energy efficient routing in mobile ad-hoc networks

Saloua Chettibi*, Salim Chikhi

MISC Laboratory, SCAL Team, University of Constantine 2, Algeria

ARTICLE INFO

Article history:
Received 28 August 2013
Received in revised form 1 September 2015
Accepted 3 September 2015
Available online 3 October 2015

Keywords:
MANETs
Routing protocol
Fuzzy logic system
Dynamic membership function
Reinforcement learning

ABSTRACT

In this paper, a dynamic fuzzy energy state based AODV (DFES-AODV) routing protocol for Mobile Ad-hoc NETworks (MANETs) is presented. In DFES-AODV route discovery phase, each node uses a Mamdani fuzzy logic system (FLS) to decide its Route REQuests (RREQs) forwarding probability. The FLS inputs are residual battery level and energy drain rate of mobile node. Unlike previous related-works, membership function of residual energy input is made dynamic. Also, a zero-order Takagi Sugeno FLS with the same inputs is used as a means of generalization for state-space in SARSA-AODV a reinforcement learning based energy-aware routing protocol. The simulation study confirms that using a dynamic fuzzy system ensures more energy efficiency in comparison to its static counterpart. Moreover, DFES-AODV exhibits similar performance to SARSA-AODV and its fuzzy extension FSARSA-AODV. Therefore, the use of dynamic fuzzy logic for adaptive routing in MANETs is recommended.

1. Introduction

A Mobile Ad-hoc Network (MANET) is a set of wireless devices having the ability to communicate without referring to any central communication infrastructure. Nodes in MANET are constantly moving. This leads to frequent topological changes and links breakage. Also, the wireless communication medium is errors-prone with limited bandwidth capacity. It follows that finding and maintaining routing paths in MANET are very challenging tasks.

In MANET, communicating devices are battery powered. Unfortunately, in many practical usage scenarios, batteries cannot be replaced or recharged. Moreover, battery exhaustion of nodes may not only result in low connectivity but can also lead to network partitioning. With the aim of extending MANET lifetime, energy-aware routing protocols have been proposed. Three main approaches are identified in the related literature [1]. First, maximum lifetime routing that seeks to balance energy expenditure among mobile nodes. Second, power-save approach that searches to minimize energy loss during the inactivity periods. Third, power-control approach where nodes adjust their transmission power so that a good compromise is found between goals of: maximizing network connectivity and minimizing energy dissipation.

MANETs are highly dynamic and uncertain environments. On one hand, changes in MANETs' topology, data-traffic load, bandwidth and energy resources are frequent. On the other hand, due to MANETs' dynamic and to their distributed organization, available routing information is uncertain and incomplete. To ensure good network performance, a routing protocol for MANETs must change its routing policy online to account for changes in network conditions [2] and to deal with routing information imprecision [3]. In other words, a routing protocol for MANETs should be adaptive. Different adaptivity contexts may be considered such as traffic, energy and mobility. This paper presents adaptive energy-efficient routing solutions.

Majority of proposed adaptive routing protocols use techniques derived from computational intelligence (CI) discipline. However, due to limited energy and computation resources in MANETs, some CI techniques are more suitable. This includes swarm intelligence [4,5], reinforcement learning [6–9] and fuzzy logic [10]. The main focus of this paper is on the use of fuzzy logic and reinforcement learning. Particularly, the following questions are answered. Firstly, which CI paradigm is more appropriate for adaptive energy efficient routing in MANETs: fuzzy logic or reinforcement learning? Secondly, does the combination of both paradigms ensure significant improvement? In fact, research in adaptation and hybridization in computational intelligence is receiving growing interest. Examples of recent related works include extensions of: krill herd algorithm [11–16], particle swarm optimization [17,18] and cuckoo search algorithm [19,20].

To the best of our knowledge, all fuzzy-logic based routing protocols for MANETs use static membership functions. In this paper, a dynamic membership function is defined to enhance the adaptivity

* Corresponding author. Tel.: +213 699 65 33 05; fax: +213 34 50 13 92.
E-mail address: sa.chettibi@gmail.com (S. Chettibi).

http://dx.doi.org/10.1016/j.asoc.2015.09.003

of the proposed fuzzy logic system. Moreover, a fuzzy extension of a previously proposed RL based routing protocol for MANETs [21] is presented. Finally, simulation results demonstrate that the proposed dynamic fuzzy logic system ensures similar performance to those achieved by reinforcement learning system. Thus, the use of dynamic fuzzy logic for the task of adaptive energy aware routing in MANETs is more appropriate.

The rest of this paper is organized as follows. Section 2 discusses the related work covering adaptive energy aware routing in MANETs using fuzzy logic and reinforcement learning. Section 3 provides an overview of the basics of fuzzy logic and reinforcement learning. Moreover, the AODV routing protocol is presented. Section 4 describes the proposed adaptive fuzzy logic system for Route REQuests (RREQs) probability forwarding tuning problem. In Section 5, SARSA-AODV protocol [21] is presented with its fuzzy extension. Section 6 reports simulation results. Finally, Section 7 concludes the paper.

2. Related work

This section describes major related works to adaptive energy aware routing in MANETs using fuzzy logic and reinforcement learning. Clearly, both techniques are adequate to achieve adaptivity feature. However, no prior work has compared the performance of each technique. The present work comes to fill this gap.

2.1. Adaptive energy aware routing in MANETs using fuzzy logic

Typically, an energy-aware routing protocol based on fuzzy reasoning uses FLS either for adjusting some routing parameters or for estimating the energy-cost of a routing path. For example, in [22], fuzzy path selection power-based AODV (FPSP-AODV) is proposed. To select a path for routing, number of hops, bandwidth and node remaining power are used to evaluate its cost. In [23], a fuzzy-based virtual backbone (FVB) routing scheme for large-scale mobile ad hoc networks is presented. FVB aims to maximize the network lifetime. For this purpose, the authors have developed a FLS with the following inputs: node residual energy, traffic and mobility. The FLS output indicates the node eligibility to be a cluster-head. Abirami et al. [24] suggested using a FLS to choose energy efficient paths after a route discovery procedure. The FLS inputs are battery cost and power consumption of discovered paths. Hiremath et al. [25] designed an adaptive energy efficient reactive routing protocol where mobile nodes use fuzzy residual-energy thresholds to decide RREQ forwarding. Chettibi and Chikhi [26] have extended OLSR proactive routing protocol with a FLS for energy aware routing. Remaining energy and expected residual lifetime are used by a node to adjust its willingness parameter that reflects its ability to act as a router.

2.2. Adaptive energy aware routing in MANETs using reinforcement learning

A mobile node implementing a RL-based routing protocol learns either how to adjust some routing parameters or how to make routing decisions (i.e. choosing next-hop or path for routing). The energy aware-routing problem in MANETs was tackled as a RL problem in [21,27,28]. In [27], each node learns how to choose its next-hop for routing. This is by using a stochastic gradient descent RL algorithm. A node decision depends on its neighbors' selfishness and remaining energy. Naruephiphat et al. [28] have proposed a RL-model that aims to balance objectives of maximizing nodes lifetime and minimizing energy consumption. Each source node runs the first-visit on policy Monte Carlo (ONMC) RL algorithm to learn how to choose among: the minimum-energy path, the max–min residual battery path, and the minimum-cost path. Chettibi and Chikhi [21] have formulated the problem of RREQs forwarding rate in an energy aware route discovery procedure as a reinforcement learning task. Each node decides the amount of RREQs to forward according to its expected residual lifetime. The learning goal is to balance energy consumption among nodes as a means for extending network lifetime.

3. Applied methods and routing protocol

In this section, preliminaries on fuzzy logic systems and reinforcement learning technique are described. Since all the proposed routing policies in this paper are implemented on the top of AODV protocol, Section 3.3 is dedicated to its presentation.

3.1. Fuzzy logic

Human beings are able to make decisions even in the presence of imprecise or incomplete knowledge. Fuzzy logic allows approximating human reasoning. In fuzzy set theory, initialized by Zadeh [29], an element can be a member of a set with a certain degree. A fuzzy logic system (FLS) uses fuzzy sets to make decisions. A FLS can be seen as an expert system that encompasses a set of linguistic fuzzy rules. Fuzzy rules follow this general pattern: *If premises(s) Then conclusion(s).* In a fuzzy rule, premises and conclusions correspond to fuzzy input and fuzzy output sets, respectively.

As shown in Fig. 1, a FLS contains four main modules. First, fuzzification module that transforms the crisp numbers inputs into fuzzy sets. This is by using membership functions. Second, knowledge base stores the IF-THEN rules. Third, inference engine is used to establish fuzzy conclusions. Finally, DEFuzzification module transforms the obtained fuzzy conclusion into a crisp value.

Mamdani [30] and Takagi Sugeno (TSK) [31] are the most utilized FLS models in the literature. They mainly differ in the output structure. In Mamdani systems, both inputs and outputs are fuzzy propositions using linguistic variables. In TSK models, system outputs are numerical values rather than linguistic variables. The outputs can be constants, polynomials, functions or differential equations.

3.2. Reinforcement learning

Reinforcement learning (RL) [32] is an efficient method for discovering policies in Markovian sequential decisions tasks. Whenever an RL agent takes an action, the environment responds by a reward or a punishment signal. The feedback signal gives an indication about the quality of undertaken actions. As depicted in Fig. 2, an RL agent interacts with its environment in discrete time steps. At a time step t, the environment state is s_t. The RL agent chooses an action a_t. Consequently, it receives a feedback r_t and a new state s_{t+1} is determined. This cycle is repeated until that the learning agent converges to an optimal policy maximizing the expected future reward.

When state transition and reward functions are known, dynamic programming [32] can be successfully applied to find an optimal policy. However, in practice, RL agents do not have a complete knowledge about their environments' models. In such circumstances, temporal difference (TD) and Monte-Carlo (MC) RL algorithms [32] are more suitable.

In this paper, SARSA, a well-known TD algorithm (see Fig. 3), is used. It is based on evaluation of action-value function denoted by $Q^{\pi}(s, a)$. This latter estimates the expected future reward to the agent when it performs a given action, a, in a given state, s, and follows the policy π thereafter. At every time step, SARSA updates the action-value function Q^{π} using the quintuple $(s_t, a_t, r_t, s_{t+1}, a_{t+1})$, which gives rise to the name of the algorithm. SARSA is an on-policy RL algorithm. It uses the learned policy not only to take decisions