Intelligent Approaches for Routing Protocols In Cognitive Ad-Hoc Networks

Suyang Ju

Submitted to the graduate degree program in
Electrical Engineering & Computer Science and the
Graduate Faculty of the University of Kansas
School of Engineering in partial fulfillment of the
requirements for the degree of Doctor of Philosophy

Dissertation Committee:

____
Dr. Joseph B. Evans, Advisor
Professor, EECS

____
Dr. Victor Frost
Professor, EECS

____
Dr. Erik Perrins
Professor, EECS

____
Dr. Xuewen Chen
Professor, EECS

____
Dr. Bozenna Pasik-Duncan
Professor, Mathematics

____
Data Defended

The Dissertation Committee for Suyang Ju certifies that this is the approved version of the following dissertation:

Intelligent Approaches for Routing Protocols In Cognitive Ad-Hoc Networks

Dissertation Committee:

Dr. Joseph B. Evans, Advisor
Professor, EECS

Dr. Victor Frost
Professor, EECS

Dr. Erik Perrins
Professor, EECS

Dr. Xuewen Chen
Professor, EECS

Dr. Bozenna Pasik-Duncan
Professor, Mathematics

Data Approved

Abstract

This dissertation describes the CogNet architecture and five cognitive routing protocols designed to function within this architecture. In this document, I first provide detailed modeling and analysis of CogNet architecture and then provide the detailed approach, mathematical analysis, and simulation results for each of the developed cognitive routing protocols.

The fundamental idea for these cognitive routing protocols is that a proper and adaptive network topology should be constructed from network nodes based on predictions using cognitive functions and past experience. The nodes in the cognitive radio network employ machine learning techniques to use past experience and make wise decisions by predicting future network conditions. The cognitive protocol architecture is a cross-layer optimized construct where the lower layer knowledge of the wireless medium is shared with the network layer.

This dissertation investigates several intelligent approaches for cognitive routing protocols, such as the multi-channel optimized approach, the scalability optimized cognitive approach, the multi-path optimized approach, and the mobility optimized approach. Analytical and simulation results demonstrate that network performance can be increased significantly by applying cognitive routing protocols.

Table of Contents

Table of Figures

Acknowledgements

First and foremost, I would like to express my sincere, cordial and affectionate thanks to my advisor, Professor Joseph B. Evans, for his unbounded support, enthusiasm, patience, encouragement and confidence that made my Ph.D studies at KU more fruitful and pleasant. When I came to KU in 2006, I have no research experience at all. Professor Joseph B. Evans guides me on my research in cognitive radio networks and gives me lots of very helpful advice and suggestions.

I am thankful to Professor Victor S. Frost and Professor Sam Shanmugan for their supportive supervision during my Ph.D study. I enrolled several their classes in the area of wireless communication and networking and gained advanced knowledge and practical experience, which is very helpful for my future career.

I specially thanks to Professor Xuewen Chen, Professor Victor Frost, Professor Bozenna Pasik-Duncan and Professor Erik Perrins for graciously serving on my dissertation committee.

Chapter 1

Introduction

1.1 Motivation

Cognitive radios (CRs), based on software-defined radios (SDRs), are considered to be the next-generation in wireless technology. They are able to monitor the communications environment, learn the channel characteristics based on the history of the environment, and perform the dynamic spectrum allocation based on the predicted future communications environment. Recent research on cognitive radio technology shows that CRs, that is, SDRs with embedded cognitive engines, are highly desirable, especially for dynamic spectrum access (DSA) in next-generation networks.

Many routing protocols have been proposed for mobile ad-hoc networks (MANETs). Most of them use an instantaneously estimated metric instead of an intelligently predicted metric for route selection. Therefore, they are not aware of history and unable to learn the trends of network changes based on past experience. For example, AODV, DSDV, and DSR adopt the number of hop counts as the metric for route selection. The nodes instantaneously estimate the hop counts for the paths between the source node and the destination node by flooding route request (RREQ) packets, so it is difficult from this learn the trends in the changes in the network and intelligently adjust the network topology. The nodes should intelligently perform routing functions to improve the network performance by avoiding unnecessary path failures or maximizing the overall network capacity.

In wired networks, the channel conditions of the links between the neighboring nodes are almost same because of the stability of the communications environment, so the value of hop counts is more suitable as the metric for route selection. For example, OSPF and BGP-4 are the dominant routing protocols for wired networks and they both adopt the number of hop counts as the metric for route selection. However, in wireless networks, the channel conditions experienced by the links vary significantly for different frequencies or different time periods because of the large-scale fading and small scale fading on a particular frequency at a given time and place. Many proposed routing protocols for wireless networks do not specifically consider the channel conditions or traffic load on a link. As a result, the nodes assume that all links have the same conditions when constructing the network topology, which is not reasonable for actual wireless scenarios. Therefore, it is necessary for the nodes to distinguish the links with different conditions to construct a proper network topology.

Many wireless routing protocols focus on scenarios with only one available frequency and only consider the network resource allocation in the spatial dimension but not in the spectrum dimension. However, in recent years, the cost of a wireless (e.g., 802.11) interface has been decreasing, which makes it feasible for the nodes to be equipped with multiple interfaces. Some multi-channel routing protocols use a static spectrum resource allocation approach by employing a license based spectrum allocation policy. Consequently, the spectrum resource is poorly allocated with spectrum holes. Therefore, a better approach would be for routing protocols to efficiently perform network resource allocation in both space and spectrum dimensions to exploit the multi-channel capability of the nodes.

A promising solution is to embed cognitive functions in software-defined radio technology to enable the nodes to learn and interact with the communications environment. For example, they could dynamically change the transmission or reception parameters by learning from past experiences and measurements of the current communications environment. However, most research focuses on solutions that modify the PHY and MAC layers and few efforts propose cognitive routing protocols that employ cognitive techniques to modify the behavior of the network layer. Cognitive routing protocols, a novel category of routing protocols, enable the nodes to learn from their past experiences and construct a proper and adaptive network topology by employing learning machines. They are designed for cognitive radio networks (CRNs).

1.2 Problem Statement

Cognitive routing protocols, a novel category of routing protocols, enable the nodes that form the network to learn from past experiences and construct a proper and adaptive network topology by employing learning machines. They are designed for and provide a critical capability for cognitive radio networks (CRNs). Some of the new challenges in cognitive routing are listed below:

1. What network architecture should be used to enable the cognitive radio to best perform the routing functions?
2. What parameters of the communications environment should the cognitive radio monitor and learn and what parameters should the cognitive radio predict?
3. How should joint network resource allocation in the space and spectrum dimensions be performed in order to construct a proper network topology?
4. What metric should be adopted by the nodes to intelligently and efficiently perform routing functions?
5. How should the tradeoff between knowledge accuracy and routing overhead be optimally adjusted?
6. How should the cognitive routing protocols be optimized based on the knowledge learned?

Many traditional routing protocols for wireless networks are single-channel routing protocols, therefore considering network resource allocation only in the space dimension. However, there are usually multiple available frequencies in CRNs. Consequently, cognitive routing protocols are multi-channel routing protocols, performing not only next-hop node assignment along a path in the spatial dimensions but also frequency assignment for the links in the spectrum dimension. In wireless networks, the channel conditions vary significantly for different links, different frequencies, or different time periods because of the small-scale fading and large-scale fading of the channel. As a result, it is difficult for the nodes to efficiently perform network resource allocation in both space and spectrum dimensions at a specific time period. Therefore, cognitive routing protocols must solve the challenging problem of how to jointly perform the network resource allocation in the spatial dimensions and the spectrum dimension to construct a proper network topology.

Many traditional routing metrics focus on how to maximize the network performance and a few of them specifically consider how to maximize the network stability. For example, the metric of hop counts might not be suitable for dynamic networks because the nodes along the selected path between the source node and the destination node might need to be physically distant to minimize the number of hop counts, making the communications links prone to breaking due to signal strength issues. There are two main drawbacks for traditional routing metrics. One drawback is that the nodes could not distinguish the links with different channel conditions. As mentioned before, in wireless networks, the channel conditions vary significantly for different links, different frequencies or different time periods because of small-scale and large-scale fading on a particular frequency. As a result, it is difficult for the nodes to improve the network performance and stability without carefully considering the impact of channel conditions on routing. The second drawback is that the nodes are not aware of history and thus are unable to learn from the trends in network changes based on past experience because traditional routing metrics are instantaneously estimated metrics instead of intelligently predicted metrics. As a result, it is difficult for the nodes to intelligently adjust the network topology to improve the network performance by avoiding unnecessary path failures or maximizing the overall network capacity. Therefore, cognitive routing protocols should solve the problem of what metric should be adopted by the nodes to intelligently and efficiently perform routing functions.

Many routing protocols proposed for multi-channel wireless networks utilize algorithms performed at a central server for spectrum allocation. Consequently, it is assumed that the global information on node position and spectrum distribution is readily available in a centralized database which enables the nodes to efficiently perform routing functions. However, it is difficult for the nodes to perform the routing functions by gathering the global information in a distributed manner if there is no centralized database. Consequently, it is necessary for the cognitive routing protocols to ensure that the distributed information at each node is up-to-date and consistent amongst the nodes to construct an adaptive and stable network topology. To make the network topology adaptive, routing updates should be triggered frequently to refresh the knowledge / memory of the nodes to accommodate the physical topology changes. On the other hand, to make the network topology stable, routing updates should be triggered only when necessary to avoid frequent network topology changes. Therefore, cognitive routing protocols working in a distributed manner should solve the problem of how to optimally adjust the tradeoff between knowledge accuracy and routing overhead.

1.3 Results Summary and Contributions

1.3.1 CogNet Architecture

The CogNet architecture was developed based on a cross-layer optimized network framework and was specifically designed for cognitive radio networks. It enables cognitive radios to share the network information between the lower three layers through a common database while efficiently processing the shared information using the cognitive engine which is attached to the common database. The cognitive engine in the CogNet architecture is primarily used for routing functions at the network layer in this work, but also serves as an example of the utility of the overall architecture. It contains four estimators for different purposes and a five-step sequential procedure is implemented to process the shared network information. The available parameters for routing functions, such as routing metrics, can be intelligently adjusted according to the cross-layer optimized feedback from the cognitive engine.

1.3.2 Multi-Channel Optimized Cognitive Routing Protocol

The spectrum aware routing protocol (SARP) is an on-demand cognitive routing protocol for cognitive ad-hoc networks. Focusing on the spectrum dimension, SARP is able to intelligently select the best frequency among all available frequencies for the links between the neighboring nodes based on the metric of the delay of RREQ packets by the multi-frequency selection function (MFSF) that performs offline learning to maximize channel diversity. In the space dimension, SARP is able to intelligently select the best path among all available paths between the source node and the destination node according to the metric of the throughput increment of a path by the multi-path selection function (MPSF) that performs offline learning using a neural network machine learning method. Simulation results show that the routing performance of SARP is better than MCRP and AODV in terms of network size, network dynamics and network spectrum utilization.

1.3.3 Scalability Optimized Cognitive Routing Protocol

The scalable cognitive routing protocol (SCRP) is an on-demand cognitive routing protocol that employs an intelligent flooding protocol, a novel approach for scalable flooding. A neural network machine learning method is adopted to make nodes aware of history. The intelligent flooding protocol saves routing overhead because nodes selectively flood RREQ packets along predicted strong links and over predicted good frequencies. The intelligent flooding protocol can be divided into two parts, the scalable space flooding protocol and the scalable spectrum flooding protocol. Analysis and simulation results show that SCRP scales well by network size, network dynamics and network spectrum.

1.3.4 Multi-Path Optimized Cognitive Routing Protocol

The cognitive multipath multi-channel routing protocol (CMMRP) is designed for a multi-channel environment where nodes can simultaneously use multiple interfaces to transmit packets over different frequencies. It employs cognitive functions to make nodes intelligently select multiple node-disjoint, edge-disjoint, and frequency-disjoint paths. Neural network machine learning is adopted to make nodes aware of history. CMMRP employs a modified path discovery protocol which can be divided into two parts, a space discovery method and a spectrum discovery method. Simulation results show that CMMRP significantly improves network reliability and performance.

1.3.5 Mobility-Aware Routing Protocol for Mobile Cognitive Networks

Traditional routing protocols trigger routing updates after the nodes detect a route failure. Even if the link conditions are getting worse, which means the link is likely to break in the future, the nodes will still transmit the packets along the current path. The mobility aware routing protocol (MARP) uses cognitive techniques to predict when the link is likely to break so that it can inform the previous hop to trigger the routing updates before the link breaks. In this way, the nodes are aware of the physical topology changes. Based on simulation results, MARP can increase overall performance of the network significantly.

1.4 Document Outline

The rest of this dissertation is laid out as follows. Chapter 2 offers detailed insight into the CogNet architecture, including modeling and analysis, which serves as the basis of all of the subsequently developed routing protocols. Chapter 3 provides a detailed description of the approach, mathematical analysis, and simulation results for the multi-channel optimized routing protocol. Chapter 4 describes the approach, mathematical analysis, and simulation results for the scalability optimized routing protocol. Chapter 5 provides details of the protocol operation as well as simulation results for the multipath optimized routing protocol. Chapter 6 describes the mobility aware routing protocol and simulation results for this approach. Chapter 7 provides conclusions and suggestions for future work.

Chapter 2

CogNet Architecture

2.1 Introduction

Cognitive radios (CR) developed based on software defined radios (SDR) are considered to be the next-generation radios. They are able to monitor the communication environment, learn the channel characteristics based on the environment history and perform the dynamic spectrum allocation based on the predicted future communication environment. Recent research [1-3] on cognitive radio technology shows that CRs with embedded cognitive engines are highly desirable, especially for dynamic spectrum access (DSA) [14] in the next-generation networks. However, most research [15,16] on CRs only focuses on the lower two layers of the OSI seven-layer protocol model, resulting in that only hardware part of CRs is considered. In other words, it is assumed that the OSI seven-layer protocol model is employed by the cognitive radio networks (CRN) with the higher layers left intact. However, with the traditional protocol models, CRs could not fully utilize the monitored information to efficiently perform the cognitive functions in each layer because it is lack of a mechanism to enable the cross-layer optimized feedback. Therefore, we argue that cross-layer optimized network architecture is necessary for CRNs to improve the network flexibility and performance.

The problem we address in this section is the design concepts and analysis of CogNet architecture which is how to share the network information learned from one layer with the other layers while enabling the cognitive engine to efficiently process the shared information. A multitude of cross-layer optimized network architectures [4-6] have been proposed by enabling the sharing of network information between layers. However, they do not meet the requirements or needs of CRNs where not only the sharing of network information between layers but also the processing and cognition of shared information should be considered. Therefore, cross-layer optimized network architectures without the cognitive engine cannot perform the cognitive functions to make CRN intelligent or cognitive. Our main contribution is CogNet architecture which is developed based on the cross-layer optimized network architecture. The proposed CogNet architecture is specially designed for CRNs by enabling CRs to share the network information between the lower three layers through a common database while efficiently processing the shared information using the cognitive engine which is attached onto the common database.

Our research experience shows that well-designed cognitive network architecture is definitely necessary for cognitive routing protocols and higher layer protocols. Cognitive engine in the proposed CogNet architecture is primarily used for routing function in network layer and is served as an example of use of CogNet architecture. However, the future work of cognitive engine should include but not limit to network layer, offering full cognition to all layers.

2.2 Related Work

In this section, we discuss related work of CogNet architecture, such as cross-layer optimized network architectures.

In [7-9], the authors showed that information exchange between layers can improve the network performance for wired networks. Cross-layer optimized feedback enables the information sharing and cooperation between layers, thus best utilizing the available network information. Although violating the traditional TCP/IP protocol model might incur additional overhead, cross-layer optimized network architecture is highly desirable because of the improved network flexibility and network performance. The available parameters of the functions in each layer can be properly adjusted to a desired value to maximize the network performance according to the feedback from the other layers. In [10], the authors illustrated the function and identified the adjustable parameters in each layer for cross-layer optimized network architecture and they are briefly listed as follows.

- Physical layer:

  Function: Transmit raw bit with minimum bit error rate.

  Parameters: Transmit power and coding/modulation.

- MAC layer:

  Function: Improve link reliability.

  Parameters: Error control code and frame length.

- Network layer:

  Function: Routing and addressing.

Parameters: Available routes and interfaces.

In [16], the authors categorized the cross-layer optimized network architecture according to the ways how traditional layered architecture is violated and they are listed as follows.

- Creation of new interfaces.
- Merging of adjacent layers.
- Design coupling without new interfaces.
- Vertical calibration across layers.

The authors also categorized the ways how cross-layer optimized network architecture is implemented and they are list as follows.

- Direct communication between layers [12].
- A shared database across the layers [6].
- Completely new abstractions [13].

From the related work, we find that cross-layer optimized network architecture is highly desirable by enabling the information sharing between layers. The proposed CogNet architecture is developed based on it.

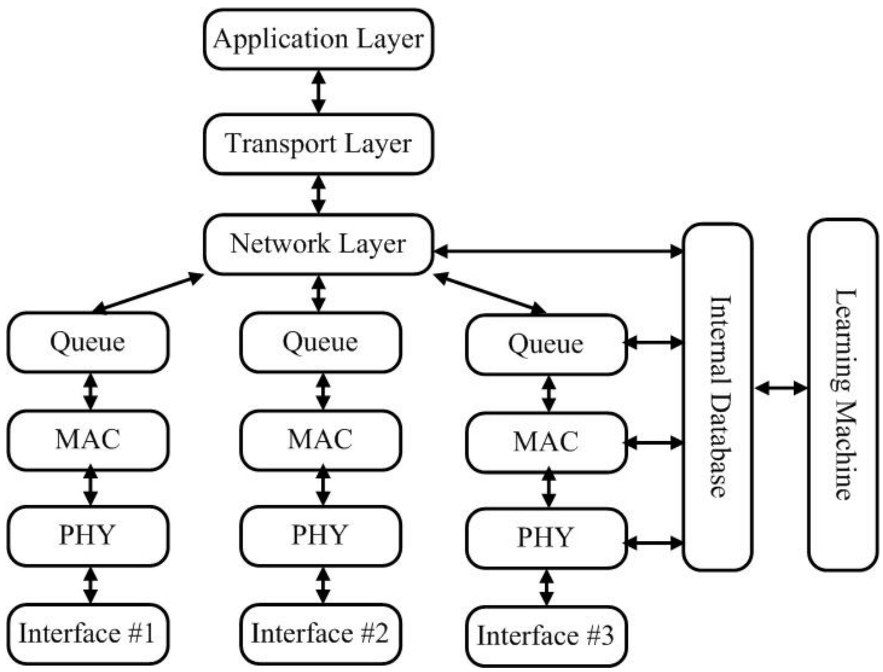


Figure 2.1 CogNet Architecture

2.3 Models

In this section, we discuss the details of our proposed CogNet architcture.

Figure 2.1 shows the proposed CogNet architecture and the details of it are explained as follows. Compared to the traditional TCP/IP protocol model, CogNet architecture is modified without any new layer inserted in between. Consequently, frame structures of the transmitted packets and designed functions in each layer are not completely changed. However, an internal database is attached onto the lower three layers, enabling CRs to perform cross-layer optimized functions with the shared network information. On the other hand, compared to the cross-layer optimized network architecture, it is modified by attaching a learning machine which is served as a cognitive engine onto the internal database. The learning machine performs the cognitive functions by processing the shared network information stored in the internal database.

One of the advantages of CogNet architecture is that it eliminates the restrictions of a dedicated centralized database in CRNs by equipping every CR with a separate internal database, which enables the learning machine to perform the cross-layer optimized cognitive functions in a distributed manner. Note that the queuing information is also collected by the internal database to estimate the load level.

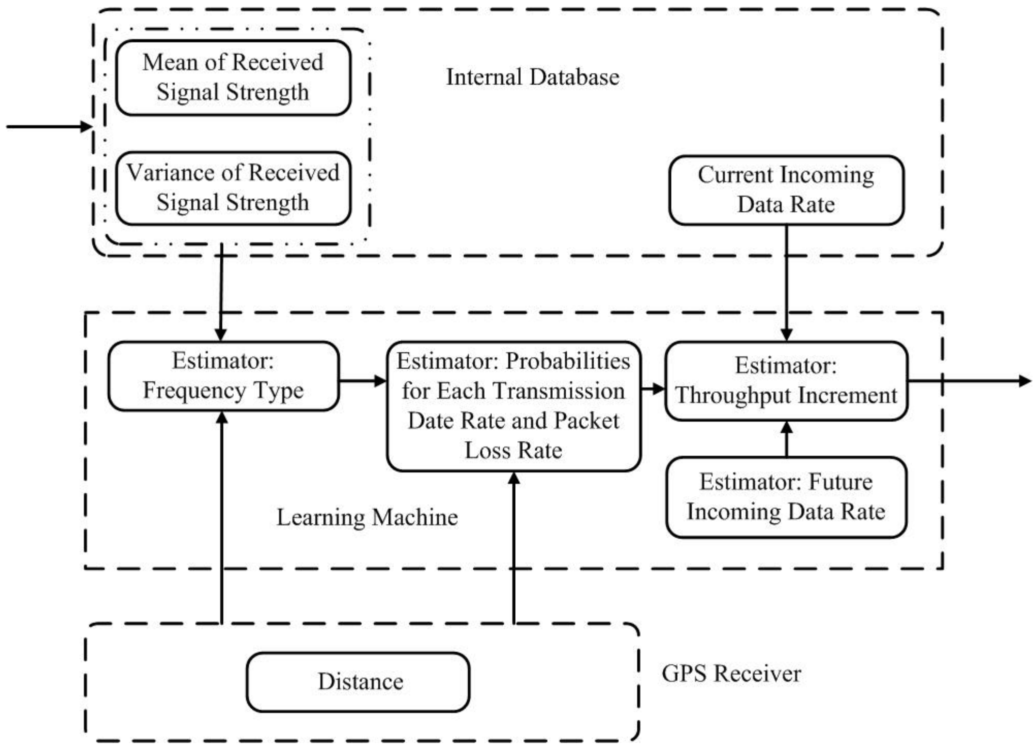


Figure 2.2 The Details of the Internal Database and Learning Machine

Figure 2.2 shows the details of the internal database and learning machine in CogNet architecture with the arrows indicating the directions of data flows. The internal database stores the network information such as the mean and variance of the received signal strength collected from PHY layer and current incoming data rate collected from MAC layer. The stored network information in each CR should be up-to-date and consistent with each other. The learning machine has four estimators, the estimator for the frequency type, the estimator for the probabilities for each data transmission rate and packet loss rate, the estimator for the future incoming data rate and the estimator for the throughput increment. The five-step sequential procedures of the estimators in the learning machine are explained as follows.

- Step #1: CRs estimate the type of transmission frequency based on the mean and variance of the received signal strength along with the corresponding transmission distance collected from GPS receiver.
- Step #2: CRs estimate the probabilities for each data transmission rate and packet loss rate based on the type of transmission frequency predicted from the former estimator and the transmission distance collected from GPS receiver.
- Step #3: CRs estimate the future incoming data rate based on the current incoming data rate, the predefined packet rate of the data application and the packet loss rate for the links along the path.
- Step #4: CRs estimate the throughput increment of a link between the neighboring CRs based on the current incoming data rates, the probabilities

for each data transmission rate and packet loss rate of a link and the future incoming data rate predicted from the former estimators.

- Step #5: CRs estimate the throughput increment of the path between the source CR and the destination CR based on the throughput increment of the links predicted from the former estimators.

Note that the throughput increment of the path between the source CR and the destination CR is used as the routing metric which is one of the adjustable parameters in network layer. The cognitive engine in the proposed CogNet architecture intelligently adjusts the routing metric in network layer according to the cross-layer optimized feedback from the lower layers.

2.4 Analysis and Implementations

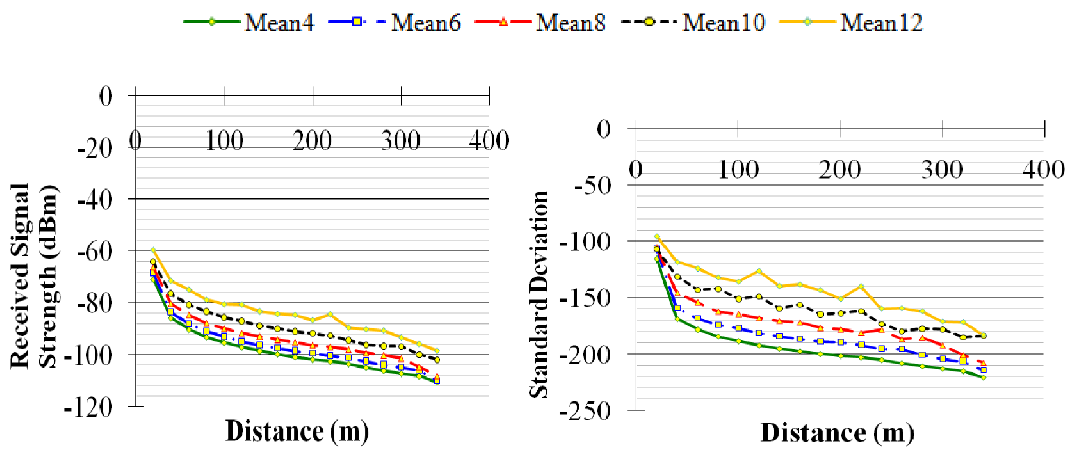
In this section, we provide the detailed analysis and implementations for each estimator.

2.4.1 Estimator #1: Frequency Type

The available frequencies in the communication environment can be categorized according to the channel characteristics. We assume that there are five types of frequencies with different shadowing means which are 4, 6, 8, 10 and 12 and same Ricean K factor as 16. In wireless networks, the received signal strength by CRs is heavily affected by the communication environment, therefore varying significantly for different frequencies or different time periods because of the large-scale fading and small-scale fading of a frequency. As a result, it is difficult for CRs to predict the type of transmission frequency based on the instantaneous received signal strength.

Figure 2.3 shows the mean and standard deviation of the received signal strength as a function of distance for frequencies with different channel characteristics in the simulated scenario where an active application stream established between two CRs transmitting 40,000 packets. Intuitively, we expect that the mean of received signal strength from CRs should be smaller when the shadowing mean of a frequency is larger. However, based on the simulation results, we find that the frequency with larger shadowing mean has larger mean and standard deviation of the received signal strength, which is against what we expect. The reason is that simulation results show characteristics of detectable signals only instead of all signals including undetectable signals. However, simulation results definitely prove that the type of transmission frequency is predictable if CRs monitor and learn the channel characteristics for a long enough time period.

The estimator for the frequency type is implemented by the neural network machine learning method which performs the off-line learning method and the details of its implementation are explained as follows. The estimator contains 17 neural networks and each neural network is responsible for processing the channel characteristics for every 20-meter transmission distance, which guarantees the data accuracy while covering the transmission distance from 0m to 340m, almost covering the entire transmission range in the simulated scenarios.



(a) Mean. (b) Variance.

Figure 2.3 The mean and standard deviation of the received signal strength as a function of distance for frequencies with different channel characteristics in the simulated scenario

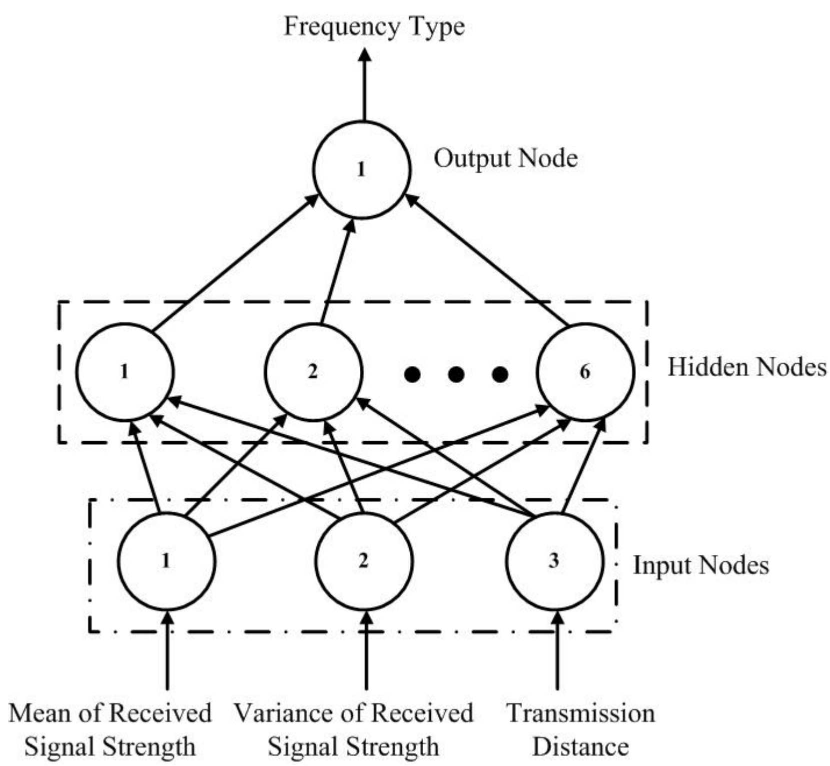


Figure 2.4 A neural network with multi-layer perceptrons

Figure 2.4 shows one of the neural networks in the estimator for the frequency type which has multilayer perceptrons. Each neural network in the estimator has three input nodes, six hidden nodes and one output node. The inputs for the neural network are the mean and the standard deviation of the received signal strength along with the corresponding transmission distance. The output for the neural network is the frequency type. The final decision of the type of the transmission frequency is based on the outputs from the 17 neural networks by selecting the frequency type with the greatest number of votes.

We denote w_{hj} as the weights in the first layer, w_{h0} as the bias weight and x_j as the inputs for each input node. The outcome of each hidden node is the weighted sum of each input node and given by

$$Z_h = sigmoid(w_h^T x) = \frac{1}{1+\exp[-(\sum_{j=1}^{3} w_{hj}x_j + w_{h0})]} \quad (2.1)$$

, h=1,2,...,6

We denote V_{ih} as the weights in the second layer and V_{i0} as the bias weight. The outcome of the output node is the weighted sum of each hidden node and given by:

$$y_i = v_i^T z = \sum_{h=1}^{6} v_{ih} z_h + v_{i0} \quad (2.2)$$

We employ the backpropagation algorithm to train the neural network. We denote r as the expected result and η as the learning factor. We calculate the gradient for the weights in the first layer:

$$
\begin{aligned}
\Delta w_{hj} &= -\eta \frac{\partial E}{\partial w_{hj}} \\
&= -\eta \sum \frac{\partial E}{\partial y} \cdot \frac{\partial y}{\partial z_h} \cdot \frac{\partial z_h}{\partial w_{hj}} \\
&= -\eta \sum -(r-y) \cdot v_h \cdot z_h (1-z_h) x_j \qquad (2.3)
\end{aligned}
$$

Based on (2.3), we find that the neural network machine learning method with sigmoid functions is a non-linear machine learning approach. This feature is necessary because of the non-linear relationship between the mean and standard deviation of the received signal strength with the corresponding distance. The mathematical analysis proves that the type of transmission frequency is predictable by neural networks using sigmoid functions.

We define the confidence rate as the number of received packets from the neural networks with correct prediction over the total number of received packets of CRs. We define the successful rate as the number of times when the estimator makes the correct prediction over the total number of times simulated, 50. Figure 2.5 shows the successful rate and confidence rate of the learning machine as a function of the number of received packets in the simulated scenario and they increase as the number of received packets increases. By off-line training the learning machine with several thousand of packets, CRs are able to make the prediction with enough confidence and successful rate. Therefore, simulation results prove that the type of transmission frequency are predictable based on past experience by using neural network machine learning method. However, it is difficult to guarantee the estimator to make the correct prediction because of the significant variance of the received signal strength in wireless networks.

2.4.2 Estimator #2: Probabilities for each data transmission rate and packet loss rate

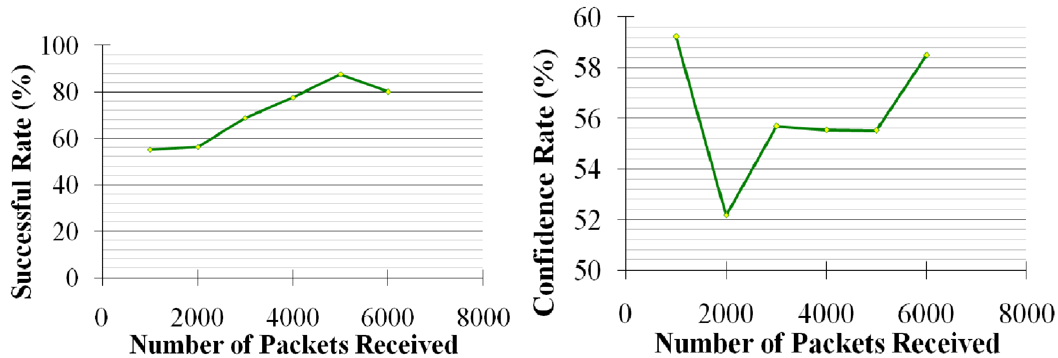
We assume that auto rate fallback is enabled at all CRs. As a result, the data transmission rate varies because of the large-scale fading and small-scale fading of a frequency, which makes it difficult for CRs to predict the data transmission rate for each packet. Figure 2.7 shows the probabilities for each data transmission rate and packet loss rate as a function of transmission distance in the simulated scenario where an active application stream established between two CRs transmitting 40,000 packets. Based on the simulation results, we find that the probabilities for packet loss rate and low data transmission rates increase as the transmission distance increases or as the shadowing mean of a frequency increases, which is what we expect. The simulation results prove that the variation of data transmission rate falls into a predictable pattern, which makes it possible for CRs to perform the off-line learning on the probabilities for each data transmission rate and packet loss rate.

The estimator for the probabilities for each data transmission rate and packet loss rate is implemented by the table look-up method which performs the off-line learning method and the details of its implementations are explained as follows. Each CR constructs a look-up table with the off-line learned information regarding the probabilities for each data transmission rate and packet loss rate along with the corresponding transmission distance for each type of frequency. According to the sequential procedures of the estimators in the learning machine mentioned before, CRs make the prediction for the probabilities for each data transmission rate and packet loss rate based on the transmission distance collected by GPS receiver after predicting the type of the transmission frequency.

2.4.3 Estimator #3: Future Incoming Data Rate

For ease of implementation, it is assume that the packet rates of all data applications are exponentially distributed with a predefined mean. According to the sequential procedures of the estimators in the learning machine mentioned before, CRs predict the future incoming data rate based on the current incoming data rate, the predefined packet rate and the packet loss rate for the links along the path. The future incoming data rate is given by

$$Rate_{Future} = Rate_{Current} + Rate_{\mathrm{Pr}edefined} \cdot \prod_{Path} (1 - Packet\_Loss_{Link}) \qquad (2.4)$$



(a) Successful Rate. (b) Confidence Rate.

Figure 2.5 The successful rate and confidence rate of the learning machine as a function of the number of received packets in the simulated scenario

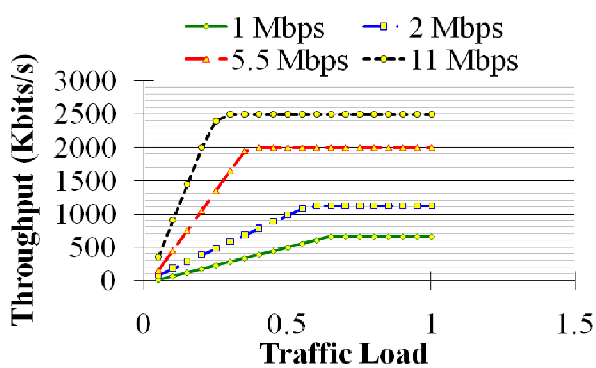
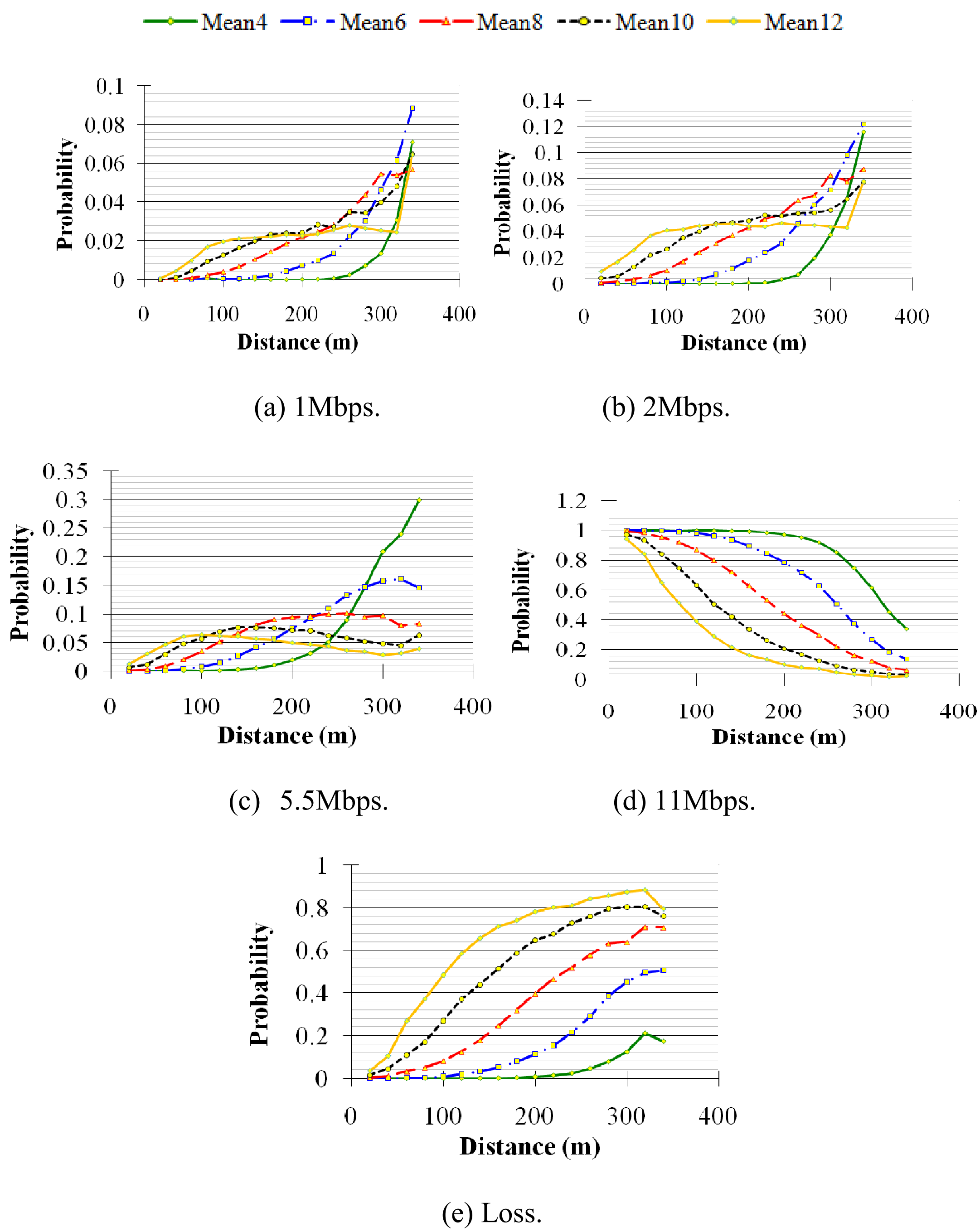


Figure 2.6 The throughput as a function of traffic load for each data transmission rate in the simulated scenario



(a) 1Mbps. (b) 2Mbps.

(c) 5.5Mbps. (d) 11Mbps.

(e) Loss.

Figure 2.7 The probabilities for each data transmission rate and packet loss rate as a function of transmission distance in the simulated scenario.

2.4.4 Estimator #4: Throughput Increment

Figure 2.6 shows the throughput as a function of traffic load for each data transmission rate in the simulated scenario where an active application stream established between two CRs transmitting 40,000 packets. Based on the simulation results, we find that different data transmission rates have different slopes of throughput increment and different saturation points for traffic load, which is what we expect. The simulation results prove that CRs are able to perform the off-line learning on the throughput increment of a link based on the current traffic load, future traffic load and the probability for each data transmission rate.

The estimator for the throughput increment of a link is implemented by the table look-up method which performs the off-line learning method and the details of its implementations are explained as follows. Each CR constructs a look-up table with the off-line learned information regarding the throughput along with the corresponding traffic load for each data transmission rate. CRs are able to predict the throughput increment of a link for each data transmission rate based on the current traffic load and future traffic load predicted from the former estimator. According to the sequential procedures of the estimators in the learning machine mentioned before, CRs predict the throughput increment of a link based on the throughput increment and probability for each data transmission rate of a link. The throughput increment of a link is given by

$$Throughput\_Inc_{Link} = \sum_{Rates} Throughput\_Inc_{Rate} \cdot P_{Rate} \qquad (2.5)$$

where $Throughput\_Inc_{Rate}$ and P_{Rate} are the throughput increment and the probability for each data transmission rate respectively.

CRs estimate the throughput increment of the path between the source CR and the destination CR based on the throughput increment of the links along the path. Suppose s is a source CR in the network and D is a set of destination CRs that are reachable from s. Let T(s, D) = (V, E) denotes the routing tree from s to CRs in D with CR set V and link set E. Every CR k∈V has a parent f(k) ∈V and a set of children c(k) = {j∈V : f(j=k)}, except that the source CR (root of the tree) has no parent and the destination CRs (leaves of the tree) have no children. We use e_k to denote link (f(k), k) and use P(i,j) to denote the sequence of links that connect j to i on the routing tree. . We define a set of link state variables Z_e for all inks e∈E on the routing tree T(s, D), which are used as the routing metrics for path selection or frequency selection after processing. The outcome variable I_k is the cumulative state variable experienced by the probe from s to CR k. We have

$$I_k = Min(I_{f(k)}, Z_{ek}) = \underset{e \in P(s,k)}{Min} Z_e \tag{2.6}$$

Based on (2.6), we find that the throughput increment of a path is determined by the minimum throughput increment of the links along the path. In communication networks, the bottleneck link along a path actually determines the network performance, which is the motivation of the proposed metric calculation algorithm. Therefore, the cognitive engine in the proposed CogNet architecture intelligently adjusts the routing metric, the throughput increment of a path, in network layer according to the cross-layer optimized feedback from the lower layers.

2.5 Conclusions

We have investigated the problem of how to share the network information learned from one layer with the other layers while enabling the cognitive engine to efficiently process the shared information. We have proposed CogNet architecture developed based on cross-layer optimized network architecture and provide the detailed modeling and analysis on it.

CogNet architecture enables CRs to share the network information between the lower three layers through a common database while efficiently processing the shared information using the cognitive engine which is attached onto the common database. Cognitive engine in the proposed CogNet architecture is primarily used for routing function in network layer and is served as an example of use of CogNet architecture. It contains four estimators for different purposes and a five-step sequential procedure is implemented to process the shared network information. The available parameters for routing functions, such as routing metrics, can be intelligently adjusted according to the cross-layer optimized feedback from cognitive engine.

The future work of cognitive engine in CogNet architecture should include but not limit to network layer, offering full cognition to all layers.

Chapter 3

Multi-Channel Optimized Cognitive Routing Protocol

3.1 Introduction

Currently, many routing protocols are proposed for mobile ad-hoc networks (MANET). Most of them adopt instantaneously estimated metric instead of intelligently predicted metric for route selection. Therefore, they are not aware of history and unable to learn the trend of network changes based on past experience. For example, AODV, DSDV and DSR adopt the value of hop counts as the metric for route selection. The nodes instantaneously estimate the value of hop counts for the paths between the source node and the destination node by flooding the RREQ packets, so it is difficult from them to learn the trend of network changes and intelligently adjust the network topology. The nodes should intelligently perform routing functions to improve the network performance by avoiding unnecessary path failures or maximizing the overall network capacity.

In wired networks, the channel conditions of the links between the neighboring nodes are almost same because of the stability of the communication environment, so the value of hop counts is suitable to be the metric for route selection. For example, OSPF and BGP-4 are the dominant routing protocols for wired network and they both adopt the value of hop counts as the metric for route selection. However, in wireless networks, channel conditions of the links vary significantly for different frequencies or different time periods because of the large-scale fading and small scale fading of a frequency. Many proposed routing protocols for wireless networks do not specially consider the channel conditions or traffic load of a frequency. As a result, the nodes assume that all links have the same conditions when constructing the network topology, which is not reasonable for actual wireless scenarios. Therefore, it is necessary for the nodes to distinguish the links with different conditions to construct a proper network topology.

Many wireless routing protocols focus on the scenarios with only one available frequency and only consider the network resource allocation in space dimension but not in spectrum dimension. However, in recent years, the cost of an 802.11 interface has been decreasing, which makes it feasible for the nodes to equip multiple interfaces. Some multi-channel routing protocols propose the static spectrum resource allocation approach by employing the license based spectrum allocation policy. Consequently, the spectrum resource is poorly allocated with spectrum holes. Therefore, the routing protocols should efficiently perform the network resource allocation in both space dimension and spectrum dimension to enable the multi-channel capability of the nodes.

In this section, we propose a novel spectrum-aware routing protocol (SARP) focusing on the scenarios of infrastructure-less multi-hop multi-channel CRNs.

Each node has the cognitive radio capability to individually detect spectrum opportunity (SOP), a set of frequency bands currently unoccupied and available for use. For ease of implementation and evaluation, we assume that all frequencies in the simulated scenarios are unoccupied and available for use because it does not affect the performance and effectiveness of the proposed routing protocol. The network information such as SOP from the lower layers of the cognitive radio should be shared with network layer to enable the nodes to perform cognitive routing functions, therefore requiring the cross-layer optimized architecture. The problem we address in this section is how to enable the nodes in CRNs to intelligently and efficiently perform routing functions and jointly perform network resource allocation in space dimension and spectrum dimension after SOP is detected.

The main contributions of this section include:

- A joint on-demand routing algorithm including two parts, multi-frequency selection function (MFSF) working in spectrum dimension and multi-path selection function (MPSF) working in space dimension.
- A cognitive cross-layer optimized routing architecture which enables the internal interaction between the network layer and the lower layers of a cognitive radio and the external interaction between the cognitive radios and the communication environment.

3.2 Related Work

In this section, we discuss the related work of SARP and some existing approaches to jointly allocate network resources in space dimension and spectrum dimension.

3.2.1 Multi-Hop Routing Protocol

In mobile ad-hoc networks (MANETs), the nodes are randomly distributed in the scenarios, resulting in multiple available paths between the source node and the destination node. A multitude of multi-hop routing protocols have been proposed for MANETs. The main idea of them is that the best path between the source node and the destination node should be selected according to the routing metric or algorithm.

Many traditional routing protocols adopt the value of hop counts as the metric for route selection and employ the shortest path algorithm. However, in [17], the authors point out that the metric of hop counts is not suitable for dynamic networks because it makes the network topology unreliable. In [18], the authors propose a routing protocol which adopts the link quality as the metric for route selection to improve network reliability. The metric of link quality is suitable for dynamic networks because it saves the routing overhead by reducing the complexity of route maintenance. In [19,20], the authors propose the routing protocols for MANETs which fall into the category of field-based routing protocols. The main idea of them is that every node in the network is assigned some degree of gradient according to the neighboring nodes and data packets are forwarded towards the nodes with the steepest gradient. Field-based routing protocols are suitable for dynamic networks because little overhead is incurred by the broken paths with local link repair. In [21], the authors propose a dynamic addressing routing protocol which makes the node addresses hierarchically allocated according to the physical topology. Consequently, routing loops are resolved easily and routing table storage is saved dramatically because of the hierarchical topology. However, the node addresses should be dynamically updated when the physical topology changes, which incurs considerable overhead. In [22-24], the authors propose the location-based routing protocols for MANETs which make the nodes equip GPS devices to obtain location information. The nodes perform directional flooding and data forwarding according to node positions. However, the mapping between the node addresses and the node positions usually incurs considerable overhead.

3.2.2 Multi-Channel Routing Protocol

In cognitive radio networks (CRNs), there are usually multiple frequencies available in the communication environment, therefore enabling the nodes to communicate over different frequencies to avoid the interference from the neighboring nodes. A multitude of multi-channel routing protocols have been proposed for MANETs. The main idea of them is that the best frequency should be selected for the links between the neighboring nodes according to the routing metric or algorithm.

Many papers provide the solutions which modify MAC or PHY layer instead of network layer. Consequently, the routing protocol is not aware of the multi-channel capability of the nodes, resulting in an improper network topology. Spectrum allocation by the multi-channel routing protocols is an emerging topic. In [25-28], the authors proposed the algorithms for spectrum allocation using the graph-coloring technique. However, the proposed algorithms are NP-hard problems and are not suitable for the networks working in a distributed manner because of the computation complexity. In [29], the authors proposed a frequency assignment and routing algorithm for multi-channel wireless networks by considering the traffic load and channel conditions of a frequency. However, the proposed algorithm is not suitable for CRNs because of the required central server. In [30], the authors proposed a multi-channel multi-interface routing protocol which is able to utilize all available frequencies even when the number of interfaces is smaller than the number of available frequencies. They introduced the concepts of fixed interface and switchable interface. The proposed routing protocol considers the channel diversity and channel switching delay but not the traffic load and channel conditions of a frequency.

3.2.3 Cognitive Routing Protocol

In recent years, several cognitive routing protocols have been proposed (e.g. [36-39]). The main idea of them is that a proper and adaptive network topology should be constructed by the nodes using the cognitive functions which make the prediction based on past experience. The nodes in CRNs employ machine learning techniques to learn past experience and make wise decisions by predicting future network conditions. The cognitive protocol architecture should be a cross-layer optimized architecture where the lower layer knowledge of wireless medium is shared with network layer. Currently, path selection and frequency selection are common topics for cognitive routing protocols.

In [31], the authors propose a capacity-based routing protocol for CRNs which adopts a novel metric for route selection by measuring and predicting the traffic pattern. In [32], the authors propose a routing protocol which adopts a novel metric predicted based on spectrum usage history to reflect the overall state. In [33], the authors propose a routing metric which reflects the conditions of spectrum resources such as channel availability and potential traffic delay. In [34], the authors proposed a spectrum aware mesh routing protocol (SAMER) which balances between long-term route stability and short-term opportunistic performance. It employs a novel metric for route selection to opportunistically route traffic over the paths with higher spectrum availability and quality. In [35], the authors proposed a high throughput spectrum aware routing protocol (SPEAR) which is a robust and efficient distributed channel assignment and routing protocol for dynamic spectrum networks based on two principles: integrated spectrum and route discovery for robust multi-hop path formation and distributed path reservations to minimize inter- and intra-flow interference.

3.3 Approach

In this section, we discuss the details of our proposed routing protocol.

3.3.1 Overview of SARP

The spectrum aware routing protocol (SARP) has two cognitive functions, intelligent multi-path selection function (MPSF) working in space dimension and intelligent multi-frequency selection function (MFSF) working in spectrum dimension. The nodes are aware of history and able to learn the trend of network changes based on past experience by the neural network machine learning method. They perform the routing functions in a distributed manner without a centralized database and the distributed information at each node should be up-to-date and consistent with each other. They intelligently and efficiently perform the routing functions and trigger the routing updates reactively when necessary. Therefore, SARP falls into two categories of routing protocols, on-demand routing and cognitive routing.

SARP is designed for the multi-channel scenarios where the nodes can simultaneously use multiple interfaces to transmit packets over different frequencies. Consequently, it should perform not only next-hop node assignment along a path in space dimension but also frequency assignment for the links between the neighboring nodes in spectrum dimension. SARP employs two novel intelligently predicted metrics for route selection and frequency selection respectively. In spectrum dimension, it is able to intelligently select the best frequency among all available frequencies for the links between the neighboring nodes according to the metric of the delay of RREQ packets by MFSF which performs the off-line learning to maximize channel diversity. In space dimension, it is able to intelligently select the best path among all available paths between the source node and the destination node according to the metric of the throughput increment of a path by MPSF which performs the off-line learning using the neural network machine learning method. SARP enables the nodes to select the best path between the source node and the destination node after selecting the best frequency for the links between the neighboring nodes, therefore performing MFSF and MPSF sequentially.

3.3.2 Intelligent Multi-Frequency Selection Function

As the name implies, the intelligent multi-frequency selection function (MFSF) is employed by SARP to enable the nodes to intelligently select the best frequency among all available frequencies for the links between the neighboring nodes according to a novel metric, the delay of route request (RREQ) packets.

We make some assumptions which are listed below. Each node in CRNs keeps track the list of available frequencies and equips three half-duplex interfaces which listen on different frequencies. Each interface is frequency-agile to quickly perform channel switching and has a separate queue to buffer the data packets and route control packets. The operating parameters for data transmission such as transmission power, modulation schemes and channel switching are properly processed by the MAC and PHY layer of the cognitive radio.

Many traditional multi-channel routing protocols enable the nodes to flood route control packets such as RREQ packets over a common control channel which might be the bottleneck of network performance for the scenarios with a large number of nodes. However, SARP solves this problem by MFSF which is performed at all nodes in a distributed manner, therefore eliminating the restriction of the common control channel and centralized database in CRNs. Note that the nodes should monitor as many frequencies as possible. The details of MFSF are explained as follows. Initially, the source node floods the RREQ packet to find the destination node by duplicating it and flooding a copy of it over each available frequency. After receiving a RREQ packet, the intermediate node records the frequency over which it is received if it is an unseen RREQ packet according the flooding ID and sequence number. Otherwise, the RREQ packet is discarded without any processing. The intermediate node relays the RREQ packet by repeating the above procedures, duplicating it and flooding a copy of it over each available frequency. The recorded frequency at each intermediate node is the selected frequency by MFSF for the link to the neighboring node. It is possible for a node to select different frequencies for the links to different neighboring nodes or relay a data packet by one or two interfaces over different frequencies. Actually, it is common that a relay node selects two frequencies for data transmission and data reception respectively to avoid internal interference.

$$Delay_{Packet} = Delay_{Queue} + Delay_{Transmission} + Delay_{Propagation} + Delay_{Switch} \quad (3.1)$$

$$Delay_{Queue} = \sum_{Queue} Delay_{Packets} \quad (3.2)$$

$$Delay_{Transmission} = Size_{Packet} / Capacity_{Channel} \quad (3.3)$$

$$Delay_{Prop} = Distance / Speed_{Light} \quad (3.4)$$

MFSF adopts the delay of RREQ packets as the metric for frequency selection and the detail analysis is illustrated as follows. (3.1)-(3.4) show the calculation of the packet delay. The packet delay is the sum of the queuing delay, packet transmission delay, packet propagation delay and channel switching delay between interfaces. The packet queuing delay of a packet is the sum of the transmission delay of all packets waiting in the queue before it. The packet transmission delay is given by the packet size over the channel capacity. The packet propagation delay is given by the transmission distance over the speed of light. We find that the calculation of packet delay is recursive because of the relationship between the packet delay and the packet queuing delay. Therefore, MFSF makes the nodes perform routing functions based on the average link quality instead of the instantaneous link quality by considering the transmission delay of all packets in the queue, which makes the constructed network topology stable. In wireless networks, the packet propagation delay is usually negligible because of the short transmission distance. Consequently, the packet delay is dominated by the packet queuing delay and packet transmission delay which reflect the traffic load and channel conditions of a frequency. The above analysis proves that the packet delay reflects the traffic load and channel conditions of a frequency. With MFSF, the delay of RREQ packets is used by the nodes to estimate the packet delay of a frequency. Therefore, the nodes are able to intelligently select the best frequency among all available frequencies for the links between the neighboring nodes according to the metric of the delay of RREQ packets to improve the network performance

Some other advantages of MFSF are listed as follows.

1) It is common that the number of interfaces equipped by a node is less than the number of available frequencies for the scenarios in CRNs, resulting in that an interface has to be switched alternately onto multiple frequencies for different data flows. The channel switching delay has little effect on the network performance if the traffic load of an interface is low because it has enough time to perform channel switching. However, it might have serious effect on the network performance if the traffic load of an interface is high because it has to perform channel switching frequently and delays the packet transmission. With MFSF, the nodes are able to intelligently select the best frequency for the links between the neighboring nodes according to the metric of the delay of RREQ packets which includes the channel switching delay. Therefore, MFSF carefully considers the impact of channel switching delay on routing by balancing the traffic load among the available interfaces of a node.
2) Channel diversity within an area heavily impacts the network performance in CRNs because there are usually multiple available frequencies. Consequently, the nodes should avoid the interference from the neighboring nodes by balancing the traffic load among the available frequencies to maximize the channel diversity. MFSF is a promising approach to efficiently solve this problem in a distributed manner.

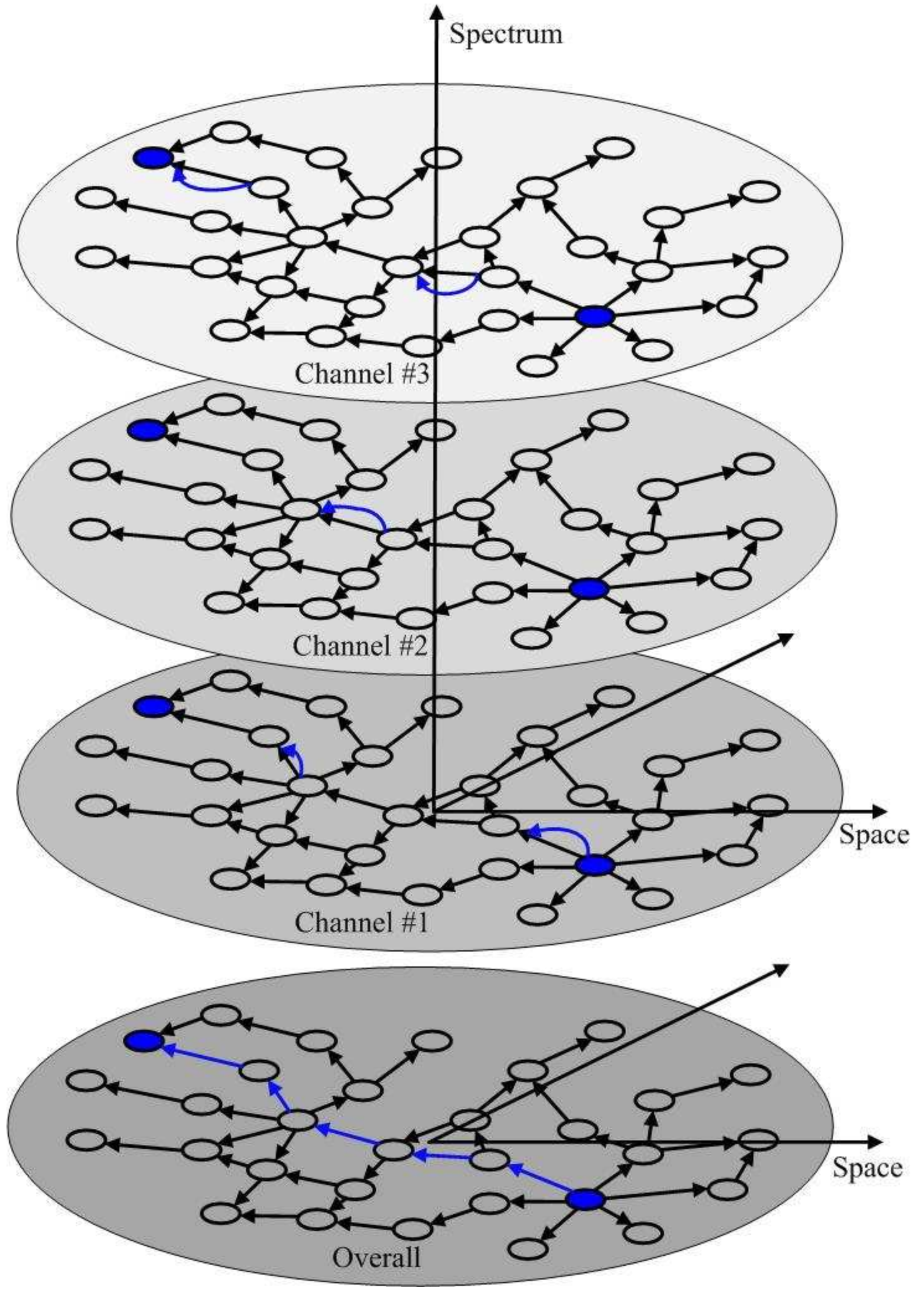


Figure 3.1 An example of the constructed network topology by SARP

Figure 3.1 shows an example of the constructed network topology by MFSF. Blue nodes are indicated as the source node and the destination node. Blue lines are indicated as the selected frequencies for the links between the neighboring nodes by MFSF for data transmission and black lines are indicated as the links for the flooding of RREQ packets. In the spectrum view, the nodes are able to properly allocate the spectrum to improve channel diversity.

3.3.3 Intelligent Multi-Path Selection Function

As the name implies, the intelligent multi-path selection function (MPSF) is employed by SARP to enable the nodes to intelligently select the best path among all available paths between the source node and the destination node according to a novel metric, the throughput increment of a path.

We define the throughput increment of a path as the predicted throughput after a new application joins minus the current throughput as (3.5) shows, therefore determining the future overall network performance. As expected, the path with the largest throughput increment should be selected for data transmission to improve the network performance. MPSF performs the off-line learning by neural network machine learning method, enabling the nodes to be aware of history and learn the trend of network changes.

$$Throughput_{increment} = Throughput_{Future} - Throughput_{Current} \tag{3.5}$$

We make some assumptions which are listed below. There are five types of frequencies which have different shadowing means which are 4, 6, 8, 10 and 12 and the same Ricean K factor as 16. The large-scale fading and small-scale fading of the frequencies are affected by the communication environment and node mobility. IEEE 802.11 is adopted as MAC protocol which has four possible data transmission rates which are 1Mbps, 2Mbps, 5.5Mbps and 11Mbps. Auto rate fallback is enabled at all nodes and the packet rates of all data applications are exponentially distributed with the same mean of 2ms.

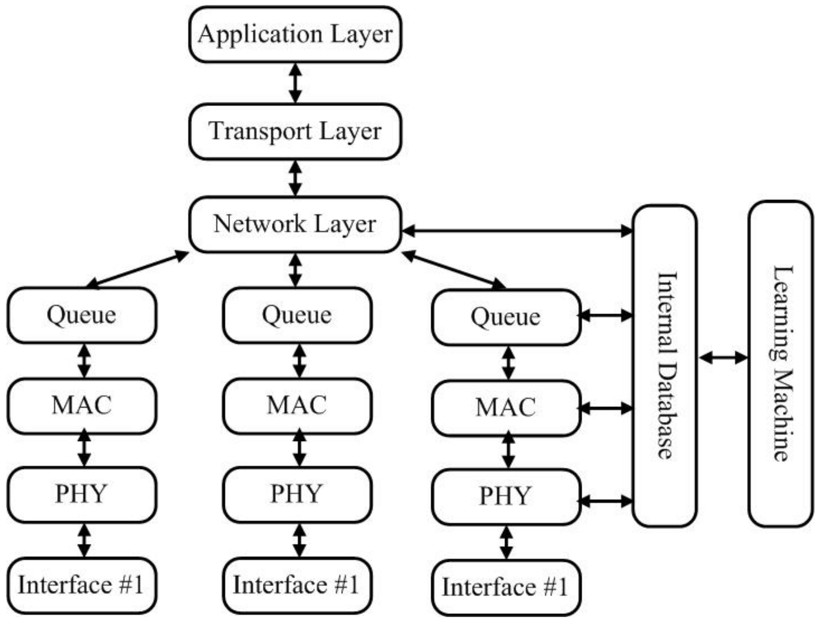


Figure 3.2 The protocol architecture for SARP

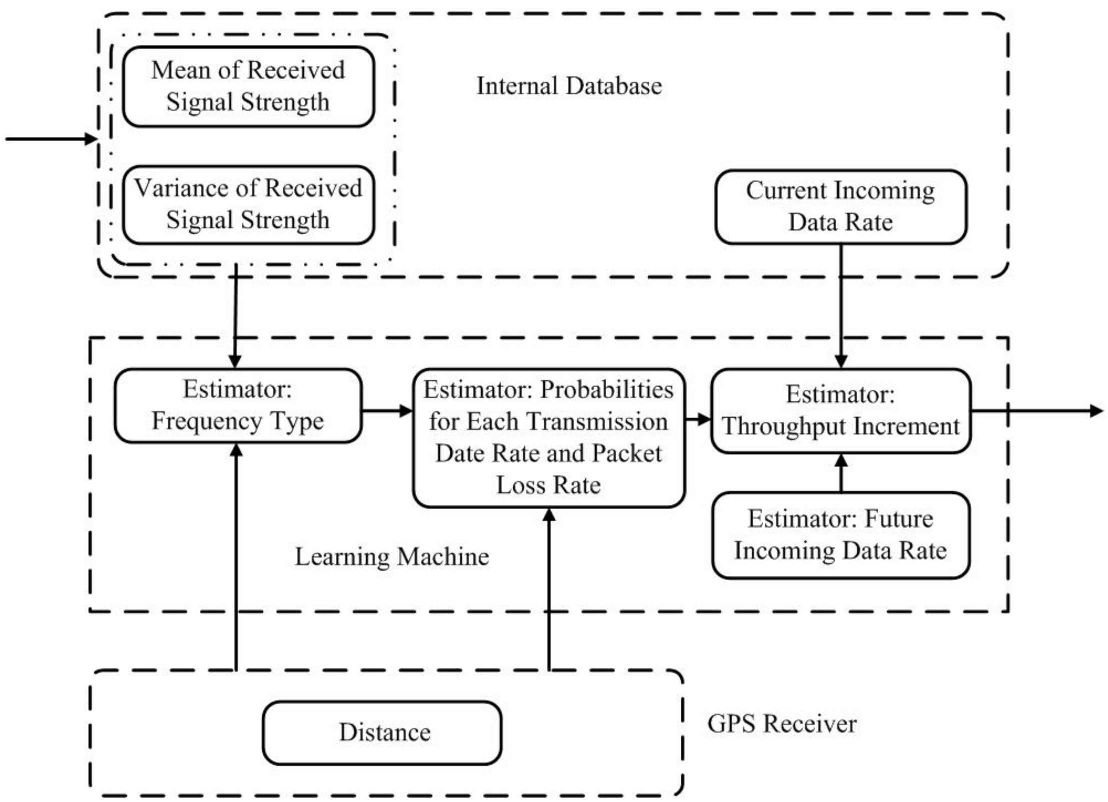


Figure 3.3 The details of the internal database and the learning machine

The nodes select the best path between the source node and the destination node after selecting the best frequency for the links, therefore performing MFSF and MPSF sequentially. The details of MPSF are explained as follows. Initially, the source node floods a RREQ packet with the recorded throughput increment of a path as null over all paths to find the destination node. After receiving the RREQ packet, the intermediate node checks if the recorded throughput increment in it is higher than the one recorded in the routing table entry for the source node if available. If the checking condition is true or the RREQ packet is unseen according the flooding ID and sequence number, it calculates the throughput increment of the extended path including the receiving link by selecting the minimum of the predicted throughput increment of the receiving link and the one recorded in RREQ packet. The RREQ packet will be discarded without any processing if the checking condition is false and the RREQ packet is seen before. The intermediate node relays the RREQ packet by recording the predicted throughput increment of the extended path in it. The following nodes to the destination node should repeat the above procedures to calculate the throughput increment of the paths between the source node and the destination node. Finally, the destination node selects the path with the highest throughput increment and sends an RREP packet back to notify the nodes along the path.

Figure 3.2 shows the protocol architecture for SARP and the details of it are explained as follows. Compared to the TCP/IP protocol architecture, it is slightly modified without any new layer inserted in between. However, an internal database and a learning machine are attached onto the lower three layers. The modified protocol architecture for SARP enables the nodes to perform cross-layer optimized routing functions because the internal database can be easily accessed by the lower three layers of the cognitive radio to share the network information and the learning machine performs the cognitive functions by processing the network information stored in the internal database. MPSF eliminates the restriction of the centralized database in CRNs because every node equips the separate internal database and the learning machine to perform cross-layer optimized cognitive functions in a distributed manner. Note that the information on the queues should be collected for MFSF.

Figure 3.3 shows the details of the internal database and learning machine with the arrows indicating the direction of data flows. The internal database stores the network information such as the mean and variance of the received signal strength collected from PHY layer and current incoming data rate collected from MAC layer. The stored network information should be up-to-date and consistent with each other. The learning machine has four estimators, the estimator for the frequency type, the estimator for the probabilities for each data transmission rate and packet loss rate, the estimator for the future incoming data rate and the estimator for the throughput increment. The steps of the estimation are listed as follows.

The nodes estimate the type of transmission frequency based on the mean and variance of the received signal strength along with the corresponding distance collected from GPS receiver.

The nodes estimate the probabilities for each data transmission rate and packet loss rate based on the type of transmission frequency predicted from the former estimator and the transmission distance collected from GPS receiver.

The nodes estimate the future incoming data rate based on the current incoming data rate, the predefined packet rate of the data application and the packet loss rate for the links along the path.

The nodes estimate the throughput increment of a link based on the current incoming data rates, the probabilities for each data transmission rate and packet loss rate of a link and the future incoming data rate predicted from the former estimators.

The nodes estimate the throughput increment of the path between the source node and the destination node based on the throughput increment of the links by MPSF.

3.4 Mathematical Analysis

In this section, we provide mathematical analysis of SARP.

Assertion I: By employing MFSF, the nodes are able to predict based on past experience and select the frequency with the best channel conditions for the links between the neighboring nodes by using the novel metric, the delay of RREQ packets.

Analysis: As mentioned before, the nodes adopt the delay of RREQ packets as the metric for frequency selection by MFSF which performs the off-line machine learning to maximize channel diversity. We make some assumptions on the queue model which are listed below. Each interface has one dedicated queue and the nodes employ the FIFO queue algorithm by inserting the packets to be transmitted at the end of the queue of the corresponding interface. Each data flow follows the M/M/1 queue model with an infinite waiting room. In other words, the packet arrival rate of each data flow is Poisson distributed with mean λ_i per second and the packet service time of each data flow is exponentially distributed with the identical mean $1/\mu_i$ per second. We denote λ as the mean of the packet arrival rate of the shared queue, $1/\mu$ as the mean of the packet service time of the shared queue and τ as the mean of the channel switching delay of an interface.

Situation I: The queue of the interface is utilized by only one data flow.

According to the queue theory, if the queue of the interface follows the M/M/1 queue model, the packet delay is exponentially distributed with mean

$$\bar{D} = \frac{1}{\mu_i - \lambda_i} \tag{3.6}$$

Based on (3.6), we find that the packet service time and packet arrival rate of the data flow determine the average packet delay. In wireless networks, they are affected by the channel capacity and traffic load. The mathematical analysis proves that the channel conditions of the frequency are predictable in terms of the channel capacity and traffic load by estimating the average packet delay of the queue.

With MFSF, the delay of RREQ packets is used to predict the channel conditions by estimating the average packet delay of the queue. A node floods a RREQ packet over a frequency by inserting it at the end of the queue of the corresponding interface. Therefore, for the first situation, the nodes are able to predict based on past experience and select the frequency with the best channel conditions for the links between the neighboring nodes.

Situation II: The queue of the interface is shared by multiple data flows.

In CRNs, it is common that the number of interfaces equipped by a node is less than the number of available frequencies for the scenarios, resulting in that an interface has to be switched alternately onto multiple frequencies for different data flows. As a result, the packet service time and packet arrival rate of the shared queue of the interface are affected by multiple data flows. Based on the probability theory, the packet arrival rate of the shared queue is still Poisson distributed with mean

$$\lambda = \sum_{i} \lambda_i \tag{3.7}$$

On the other hand, the packet service time of the shared queue is still exponential distributed because the mean of the packet service time of all data flows are identical. The mean of it is given by

$$\mu = \mu_i \tag{3.8}$$

According to the queue theory, the shared queue follows the M/M/1 queue model and the packet delay including channel switching delay is exponentially distributed with mean

$$\bar{D} = \frac{1}{\mu - \lambda} + \tau \tag{3.9}$$

Based on (3.9), we find that the packet service time and packet arrival rate of the shared queue dominate the average packet delay. In wireless networks, they are affected by the channel capacity and traffic load. The mathematical analysis proves that the channel conditions are predictable in terms of the channel capacity and traffic load by estimating the average packet delay of the shared queue.

As mentioned before, a node floods a RREQ packet over a frequency by inserting it at the end of the queue of the corresponding interface. With MFSF, The delay of RREQ packets is used to estimate the average packet delay of the shared queue which well reflects the overall status of the channel conditions by considering the channel switching delay of an interface. Therefore, for the second situation, the nodes are able to predict based on past experience and select the frequency with the best channel conditions for the links between the neighboring nodes.

Assertion II: By employing MPSF, the nodes are able to predict the throughput increment of a link based on their past experience by using off-line machine learning method after predicting the type of transmission frequency.

Analysis: The detail analysis on the bit error probability for multilevel baseband transmission schemes is illustrated as follows. The sequence of message bits $\{b_k\}$ are modulated onto the M-level amplitude of a periodic pulse waveform. The transmitted signals $X_T(t)$ and as the received signals $X_r(t)$ are given by

$$X_T(t) = \sum_{j=-\infty}^{\infty} A_j P_T(t - jT_s) \tag{3.10}$$

$$X_r(t) = \sum_{j=-\infty}^{\infty} A_j P_r(t - jT_s) + n_0(t) \tag{3.11}$$

where $A_j = \pm A, \pm 3A, \pm 5A..., \pm(M-1)A$ and $n_0(t)$ is the additive white Gaussian noise. Note that $P_T(t)$ and $P_r(t)$ are the response of the transmit filter $H_T(f)$ and receive filter $H_T(f)H_C(f)H_R(f)$ to the input pulse $P_M(t)$ respectively which is the unit energy pulse of arbitrary shape confined to the time interval $[0, T_s]$.

According to the network communication theory, the bit error probability P_{eb} for the M-level PAM is given by:

$$P_{eb} \le Q\left(\sqrt{\frac{2E_b}{N_0}\frac{3\log_2(M)}{M^2-1}}\right) \tag{3.12}$$

where E_b is the energy per bit and N_0 is the noise power spectral density. Based on (3.12), we find that the bit error probability is determined by the received signal strength, the noise power and the modulation schemes. As mentioned before, auto rate fallback is enabled at all nodes. In actual scenarios, the nodes should keep the bit error probability below a threshold by determining a proper modulation scheme and transmit power for the varied channel conditions.

In wireless networks, the variance of the received signal strength from the transmit node is affected by the small-scale and large-scale fading of a frequency caused by the communication environment of the neighboring area. It is reasonable to assume that the characteristics of the communication environment vary slowly in terms of the transmission duration. As a result, the variation of the received signal strength and the noise power falls into a predictable pattern. We denote B as the channel bandwidth, S as the received signal power and N as the noise power of an additive white Gaussian noise. According to the Shannon-Hartley theorem, the channel capacity C is given by:

$$C = B\log(1+\frac{S}{N}) \tag{3.13}$$

Based on (3.13), we find that the channel capacity also falls into a predictable pattern under the assumptions we made and if the modulation scheme and the transmit power are properly determined. Therefore, the mathematical analysis proves that the nodes are able to predict the throughput increment of a link based on their past experience by using off-line learning method after predicting the type of transmission frequency.

Assertion III: By employing MPSF, the nodes are able to predict based on past experience and select the best path among all available paths between the source node and the destination node by using the novel metric, the throughput increment of a path.

Analysis: Suppose s is a source node in the network and D is a set of destination nodes that are reachable from s. Let T(s, D) = (V, E) denotes the routing tree from s to the nodes in D with node set V and link set E. Every node k∈V has a parent f(k) ∈ V and a set of children c(k) = {j∈V : f(j=k)}, except that the source node (root of the tree) has no parent and the destination nodes (leaves of the tree) have no children. We use e_k to denote link (f(k), k) and use P(i,j) to denote the sequence of links that connect j to i on the routing tree. The probing model is used to analyze the logical routing tree constructed by SARP and is explained as follows. The source node floods a probe to search for the destination node when there is no path between them. The intermediate node relays the received probe through all of its children. We define a set of link state variables Z_e for all inks e∈E on the routing tree T(s, D), which are used as the routing metrics for path selection or frequency selection after processing. Depending on the ways to process the set of state variables of the links along a path, traditional routing metrics are classified into additive and multiplicative.

The routing metrics such as packet delay and hop counts are typical additive metrics. The outcome variable T_k is the cumulative state variable experienced by the probe from s to node k. We have

$$T_k = T_{f(k)} + Z_{ek} = \sum_{e \in P(s,k)} Z_e \qquad (3.14)$$

The routing metrics such as packet loss rate and link stability are typical multiplicative metrics. The outcome variable L_k is the cumulative state variable experienced by the probe from s to node k. We have

$$L_k = L_{f(k)} \cdot Z_{ek} = \prod_{e \in P(s,k)} Z_e \qquad (3.15)$$

However, MPSF employs the throughput increment of a path as the metric for route selection. The novel metric is not additive or multiplicative, therefore requiring a new way to process the set of state variables. The outcome variable I_k is the cumulative state variable experienced by the probe from s to node k. We have

$$I_k = Min(I_{f(k)}, Z_{ek}) = \underset{e \in P(s,k)}{Min} Z_e \qquad (3.16)$$

Based on (3.16), we find that the throughput increment of a path is determined by the minimum throughput increment of the links along the path. In communication networks, the bottleneck of a path determines the network performance. Therefore, the nodes are able to predict based on past experience and select the best path among all available paths between the source node and the destination node by using the novel metric, the throughput increment of a path.

3.5 Simulations

In this section, we provide the simulation results of SARP using Qualnet 4.0. To show the benefits of SARP, it is compared with ad-hoc on demand routing protocol (AODV) and a simple multi-channel routing protocol (MCRP).

We implement two versions of SARP, non-optimal version and optimal version. The optimal version of SARP guarantees the nodes in CRNs to know the type of each frequency without performing the off-line learning and some predefined paths between the source node and the destination node are manually selected. However, the non-optimal version of SARP cannot guarantee the nodes to correctly predict the type of each frequency and the nodes have to dynamically determine the paths between the source node and the destination node.

The details of MCRP are explained as follows. MCRP is an on-demand multi-channel routing protocol. It enables the nodes to flood RREQ packets over the common control channel instead of over all available frequencies like SARP. However, the common frequency is not always available at all nodes in actual scenarios, therefore resulting in bad network performance or unstable network topology. In spectrum dimension, MCRP lets the nodes simply negotiate with each other to randomly select a frequency without considering the channel conditions. In space dimension, it lets the nodes select the best path according to the value of hop counts instead of the throughput increment of a path.

We present following metrics with 95% confidence intervals of measured values to compare the network performance of SARP with MCRP and AODV.

- Overhead: The number of RREQ packets received over all of the available frequencies, which dominates the number of route control packets.
- Throughput: Average rate of successful message delivery measured in Kbits per second, which reflects the network reliability and performance.

3.5.1 Impact of Network Size

In the first experiment, we show how network performance is affected when the number of nodes increases. We created a scenario which has 12 applications distributed in a 600m by 1500m region. There are six available frequencies which have different shadowing means such as 4, 6, 8, 10 or 12 and same Ricean K factor as 16. Automatic rate fallback is enabled at all nodes. UDP and IEEE 802.11 are employed as TCP and MAC protocol respectively. Application packet rate is exponentially distributed as a mean of 2ms. We vary the number of nodes and compare 3-interface SARP with 3-interface MCRP and 1-interface AODV.

Figure 3.4 and figure 3.5 show the comparison of throughput and overhead respectively as a function of number of nodes. As expected, overhead increases and throughput decreases as the number of nodes increases.

Compared to 3-interface MCRP and 1-interface AODV, 3-interface SARP increases overhead considerably. The reason is explained as follows. SARP performs spectrum allocation by MFSF which lets the nodes flood a copy of RREQ packet over each available frequency to estimate the channel conditions, which incurs considerable overhead. However, MCRP enables the nodes to flood RREQ packets over the common control channel and performs spectrum allocation by letting the nodes simply negotiate with each other to randomly select a frequency for data transmission. AODV is a single-channel routing protocol without performing the spectrum allocation. As a result, it is reasonable for 3-interface SARP to have more overhead than 3-interface MCRP and 1-interface AODV.

Compared to 3-interface MCRP, 3-interface SARP increases throughput considerably. The reason is explained as follows. In spectrum dimension, SARP is able to intelligently select the best frequency among all available frequencies for the links between the neighboring nodes according to the metric of the delay of RREQ packets by MFSF which performs the off-line learning to maximize channel diversity. In space dimension, SARP is able to intelligently select the best path among all available paths between the source node and the destination node according to the metric of the throughput increment of a path by MPSF which performs the off-line learning using the neural network machine learning method. On the other hand, MCRP enables the nodes to flood RREQ packets over the common control channel which is the bottleneck for scenarios with a large number of nodes. In spectrum dimension, MCRP lets the nodes simply negotiate with each other to randomly select a frequency for data transmission without considering the channel conditions. In space dimension, MCRP lets the nodes select the best path among all available paths according to the value of hop counts without distinguishing the links with different conditions. Compared to 1-interface AODV, 3-interface SARP increases throughput significantly because it is a multi-channel routing protocol instead of a single-channel routing protocol. The simulation results show that it has about 10 times the throughput as 1-interface AODV.

The optimal version of SARP performs better than the non-optimal version of SARP. The reason is explained as follows. The optimal version of SARP guarantees the node to know the type of each frequency and some predefined paths between the source node and the destination node are manually selected. However, the non-optimal version of SARP enables the nodes to make the prediction on the frequency type with a successful rate of about 80%. Consequently, the nodes might select some non-optimal paths between the source node and the destination node because of the inaccurate knowledge on the communication environment.

The above results show that SARP has better network performance than MCRP and AODV for scenarios with a large number of nodes.

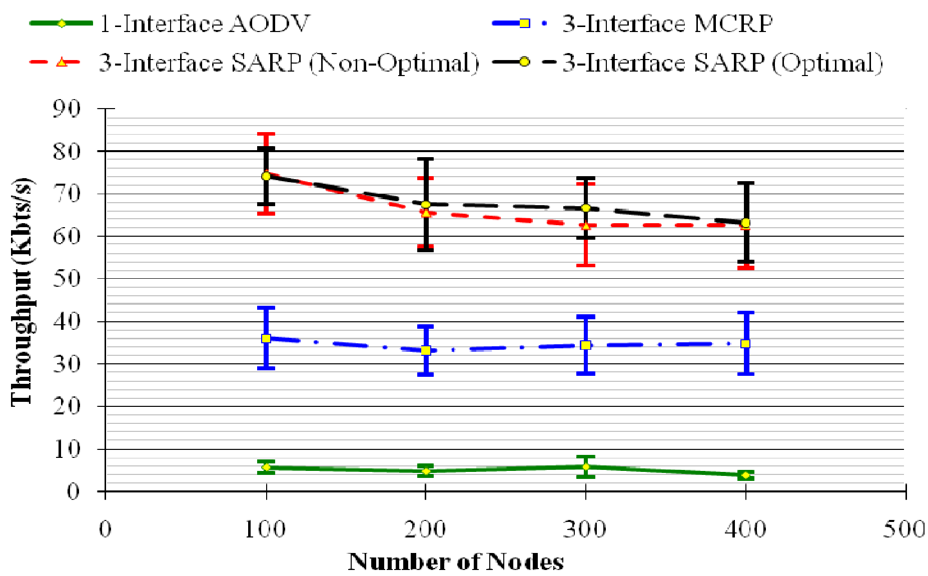


Figure 3.4 Comparison of throughput as a function of number of nodes.

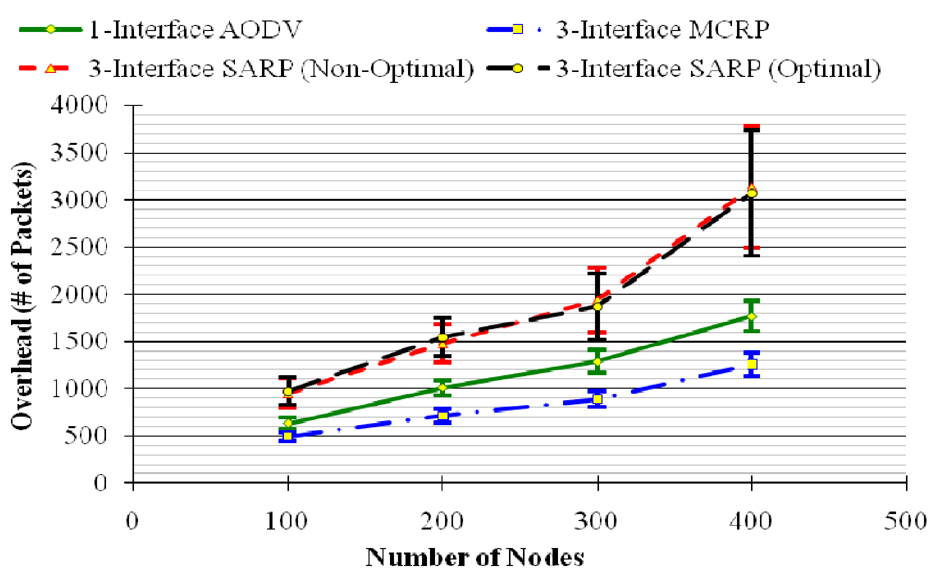


Figure 3.5 Comparison of overhead as a function of number of nodes.

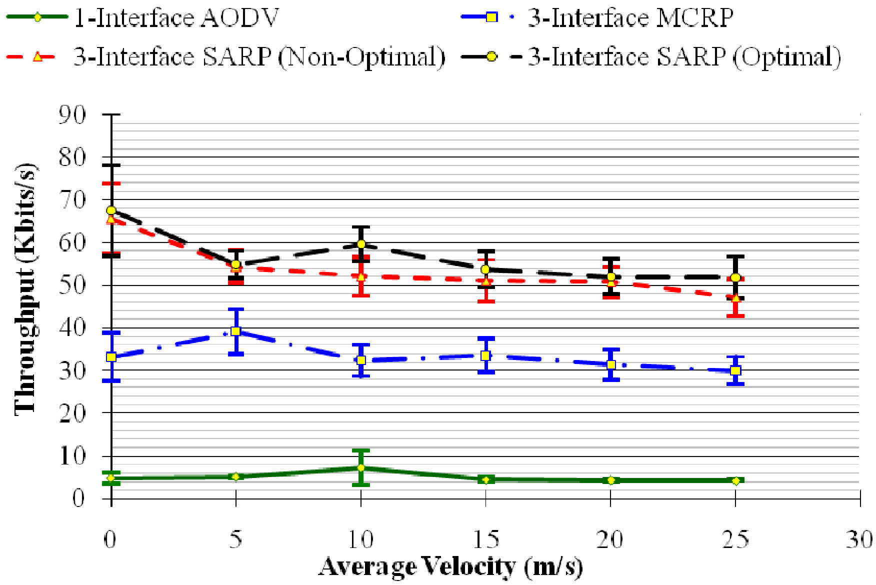


Figure 3.6 Comparison of throughput as a function of average velocity.

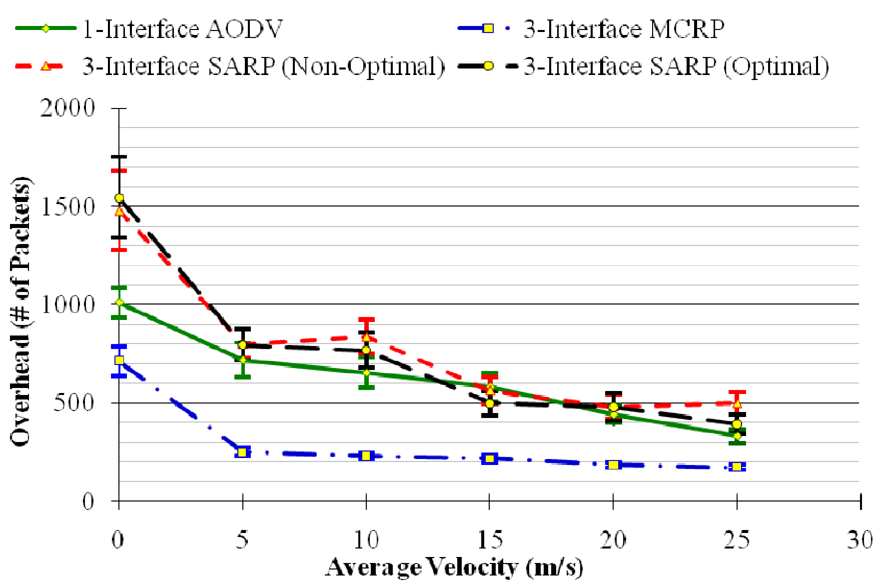


Figure 3.7 Comparison of overhead as a function of average velocity.

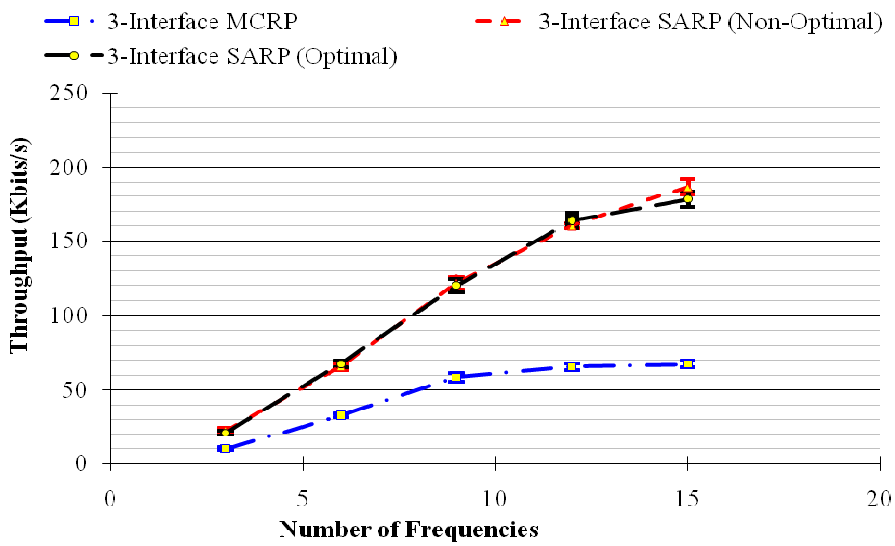
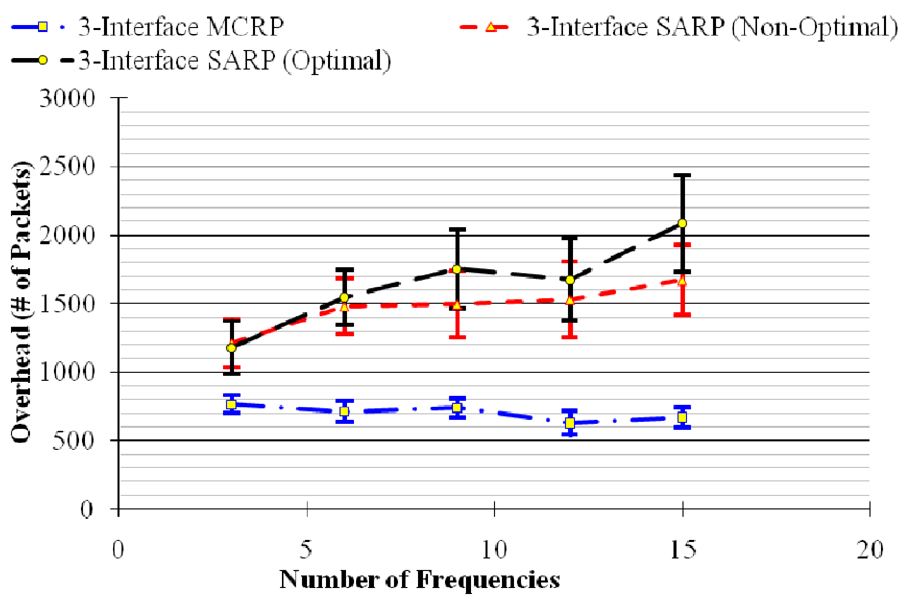


Figure 3.8 Comparison of throughput as a function of number of frequencies.



.

Figure 3.9 Comparison of overhead as a function of number of frequencies.

3.5.2 Impact of Network Dynamics

In the second experiment, we show how network performance is affected when node velocity increases. We created a scenario similar as the first experiment which has 12 applications and 200 nodes distributed in a 2500m by 2500m region. We vary node velocity and compare 3-interface SARP with 3-interface MCRP and 1-interface AODV.

Figure 3.6 and figure 3.7 show the comparison of throughput and overhead respectively as a function of average velocity. As expected, the throughput decreases as the average velocity increases because the link breakage happens frequently.
Compared to 3-interface MCRP and 1-interface AODV, 3-interface SARP increases overhead considerably. On the other hand, compared to 3-interface MCRP and 1-interface AODV, 3-interface SARP increases throughput considerably. The optimal version of SARP performs better than the non-optimal version of SARP. The reason is similar as the first experiment. These results show that SARP has better network performance than MCRP and AODV for scenarios with high node velocity.

3.5.3 Impact of Network Spectrums

In the third experiment, we show how network performance is affected when the number of frequencies increases. We created a scenario similar as previous experiment which has 12 applications and 200 nodes distributed in a 600m by 1500m region. We vary the number of frequencies and compare 3-interface SARP with 3-interface MCRP.

Figure 3.8 and figure 3.9 show the comparison of throughput and throughput respectively as a function of number of frequencies. As expected, the overhead and throughput increase as the number of frequencies increases.

Compared to 3-interface MCRP, 3-interface SARP increases overhead considerably. The reason is explained as follows. 3-interface SARP performs spectrum allocation by MFSF which lets the nodes flood a copy of RREQ packet over each frequency to estimate the channel conditions, which incurs considerable overhead. Therefore, the overhead increases as the number of frequencies increases. However, 3-interface MCRP enables the nodes to flood RREQ packets over the common control channel. Therefore, the overhead is almost constant as the number of frequencies increases.

Compared to 3-interface MCRP, 3-interface SARP increases throughput considerably. The reason is explained as follows. In spectrum dimension, SARP is able to intelligently select the best frequency among all available frequencies for the links between the neighboring nodes according to the metric of the delay of RREQ packets by MFSF which performs the off-line learning to maximize channel diversity. In space dimension, SARP is able to intelligently select the best path among all available paths between the source node and the destination node according to the metric of the throughput increment of a path by MPSF which performs the off-line learning using the neural network machine learning method. The performance of optimal version and non-optimal version of SARP are almost same. These results show that SARP has better network performance than MCRP for scenarios with a large number of frequencies.

3.6 Conclusions

We have investigated the problem of cognitive routing in multi-channel mobile ad-hoc networks and proposed a novel spectrum aware routing protocol (SARP) which is an on-demand cognitive routing protocol.

In spectrum dimension, SARP is able to intelligently select the best frequency among all available frequencies for the links between the neighboring nodes according to the metric of the delay of RREQ packets by MFSF which performs the off-line learning to maximize channel diversity. In space dimension, SARP is able to intelligently select the best path among all available paths between the source node and the destination node according to the metric of the throughput increment of a path by MPSF which performs the off-line learning using the neural network machine learning method. Simulation results show that the routing performance of SARP is better than MCRP and AODV in terms of network size, network dynamics and network spectrums.

Chapter 4

Scalability Optimized Cognitive Routing Protocol

4.1 Introduction

Scalability is an important property of a network, which is defined as the ability of a network to handle growing amounts of work in a graceful manner. For wired networks, it is usually measured in terms of network size. However, for wireless networks, this property should be measured in terms of network size, network dynamics and network spectrums. Scalability of a network is affected by many factors of network designs. In this section, we mainly consider the scalability of a routing protocol which is defined as how routing protocol affects the scalability of the network.

In wired networks, the major consideration on scalability of a routing protocol is focused on the number of routing table entries it can handle. The size of a wired network is usually large, even as large as the internet covering the whole world. Therefore, a scalable routing protocol for wired network should maintain a reasonable routing table size and perform efficient routing table look-up algorithm to minimize the table look-up time. Otherwise, the routing table processing time might seriously impact total end-to-end packet delay. Currently, BGP-4 combined with OSPF-2 is the routing protocol backing the core routing decisions on the internet. It employs several useful mechanisms to scale the network, which are briefly explained as follows.

- Network hierarchy: BGP-4 is a typical hierarchical routing protocol, which effectively reduces the routing table size. BGP-4 is a path vector routing protocol, which views the network as a set of anonymous systems (AS). AS is a collection of connected IP routing prefixes under the control of one or more network operator. Within an AS, the routers employ interior gate way protocol (IGP). Between AS, the routers employ border gateway protocol (BGP). The choice of IGP for an AS is independent with others, which enables the configuration more flexible. BGP routers only need to consider the IGP routes within the AS and the BGP routes between AS. Therefore, this mechanism significantly reduces the number of routing table entries the router needs to handle.
- Route reflector: Full meshed network is required by IBGP working within an AS to make sure the network information is consistent between IBGP routers, so that BGP is always synchronized. However, the practical issue is that the number of connections required might be extremely large in a large AS. The excessive routing overhead might significantly degrade the network performance. Route reflector is served as a central server for a cluster of directly connected routers and ensures that the updated network information from a peer router is consistent with the other routers.

Therefore, this mechanism effectively saves the routing overhead because only the router reflector needs to flood the copy of network information.

- BGP Confederations: In practice, the network size of a single AS might be too large for an internal router to handle. BGP confederation provides an approach to further divide an AS into multiple internal sub-AS. These internal small AS are still viewed as a single AS from the external world. In other words, BGP confederation makes an AS a hierarchical network and the whole set of AS multiple hierarchies, which improves the scalability of a network by minimizing the routing table entries in a router.

In wireless networks, the scalability of a routing protocol should be focused on the network size, network dynamics and network spectrums. It is obvious that the considerations are different compared to wired networks where network size is the major issue because wireless communication environment is much more complicated than that of wired networks. Some of the new features are listed as below:

- Open communication environment: In wireless network, the nodes have to broadcast signals to transmit packets which usually lead to interference to its neighboring nodes. If too many neighboring nodes are transmitting packets simultaneously, a node might be unable to receive a packet correctly because of the high packet error rate. A node might transmit data packets for application or control packets to maintain the network. We consider data packets as throughput packets and control packets as overhead packets. The ratio of overhead packets and data packets should be minimized to improve the network performance. The routing packets from network layer are counted as overhead packets, which are used by the nodes to maintain the network topology. Therefore, a scalable routing protocol should minimize the routing overhead.

- Unstable communication environment: In wireless network, the major issue is the unstable communication environment which is caused by many factors such as node mobility, multipath fading and interference from the neighboring nodes. The varying receiving signals lead to the varying bit error rate and packet error rate. It is very common to retransmit a packet when encountering a corrupted packet. For data packets, the nodes are not able to limit the data packet transmissions to make the communication environment relatively clean. However, the nodes by employing a scalable routing protocol are able to limit the control packet transmissions.
- Unstable topology: One big advantage of wireless networks is that it enables network dynamics. The nodes are able to transmit with each other while moving. However, one big issue of network dynamics is the unstable network topology. The nodes should maintain the network topology by repairing the broken links when the network topology changes, which usually incurs routing overhead. Therefore, it is common that the network performance degrades and the routing overhead increases as the nodes are moving faster. However, a scalable routing protocol should be able to efficiently adapt the routing topology to accommodate the physical

topology, while minimizing the routing overhead.

- Multi-channel Environment: In cognitive radio networks (CRN), the nodes are able to select or transmit simultaneously over multiple frequencies. This big advantage of CRN improves the network performance and makes networks topology very flexible. However, the big issue is how to select among the set of frequencies. One solution is to measure the conditions of all frequencies by transmitting control packets over each frequency and select the best frequency for data transmission. This approach incurs significant routing overhead and degrades the network performance. Therefore, this approach is not scalable as the number of available of frequencies increases. In this section, we provide a novel approach to solve this problem.

Based on the new features of the wireless communication environment illustrated above, we argue that scalability of a routing protocol for wireless networks should be focused on the routing overhead it can handle as the network size, network mobility and network spectrum increases.

In practice, the size of wireless networks is usually small compared to wired networks. Therefore, the size of routing table and routing table look-up time should not be the major issue.

In this section, we provide a novel solution to solve the scalability problem of a routing protocol by cognition. Cognitive routing protocols, a novel category of routing protocols, enable the nodes to learn the past experience and construct a proper and adaptive network topology by employing the learning machines. We propose a novel scalable cognitive routing protocol (SCRP) focusing on the scenarios of infrastructure-less multi-hop multi-channel CRNs. Each node has the cognitive radio capability to individually detect spectrum opportunity (SOP), a set of frequency bands currently unoccupied and available for use. For ease of implementation and evaluation, we assume that all frequencies in the simulated scenarios are unoccupied and available for use because it does not affect the performance and effectiveness of the proposed routing protocol. The network information such as SOP from the lower layers of the cognitive radio should be shared with network layer to enable the nodes to perform cognitive routing functions, therefore requiring the cross-layer optimized architecture.

The problem we address in this section is how to save routing overhead to make MANET scalable in terms of network size, network dynamics and network spectrum while maintaining the network performance. Many routing protocols have been proposed to solve this problem. However, most of them trade performance for scalability.

The main contributions of this section include:

- Space flooding protocol: It is used to make the flooding scalable by controlling the flooding region.
- Spectrum flooding protocol: It is used to make flooding scalable by limiting the number of frequencies used for control packets

4.2 Related Work

In this section, we discuss related work of SCRP and approaches to make MANETs scalable.

4.2.1 On-demand Routing Protocols

A multitude of on-demand routing protocols have been proposed (e.g. [40-45]). This approach takes the advantage that the frequency of the session initiation query might be less than the frequency of topology changes. In other words, the nodes without active sessions will not trigger routing updates when topology changes, which greatly reduces routing overhead.

Unlike table-driven routing protocols, routing updates are triggered reactively when links break instead of being triggered periodically. In [46], simulation results show that the routing performance of AODV is better than DSDV [47]. However, if the topology changes too fast, the advantage of on-demand routing becomes drawback. Therefore, on-demand routing is suitable for static or slowly changed network topologies. It is a simple and efficient approach to make MANET scalable.

4.2.2 Hierarchical Routing Protocols

Many hierarchical routing protocols have been proposed (e.g. [48-50]). This approach is inherited from wired networks where the network size is a big issue. Multi-level hierarchies are adopted in network topologies to save routing table storage but side effects include non-optimized topologies. However, this approach is suitable for some scenarios when cluster head along with its slaves maintain a relative stable topology compared to the overall network topology. In this case, only the connections between cluster heads need to be updated. Therefore, hierarchical routing is commonly used in wired networks, but may not be suitable for dynamic network topologies because significant routing overhead might be generated to repair broken links for a cluster.

4.2.3 Location-based Routing Protocols

A multitude of location-based routing protocols have been proposed (e.g. [51-53]). This approach takes the advantage of the preprocessed network topology by GPS devices. The idea is that nodes use GPS devices to obtain destination location information before the connection is initiated. Directional flooding and data forwarding are the mechanisms used in location-based routing protocols, which improves network performance by effectively preventing the data from flooding in all directions. However, the mapping between node address and location information usually incurs additional routing overhead or processing delay compared to ordinary topology-based routing protocols. In spite of that, location-based routing is an efficient approach to make MANETs scalable in terms of network size and network dynamics.

4.2.4 Field-Based Routing Protocols

Several field-based routing protocols have been proposed for MANET (e.g.

[54,55]). This approach takes the advantage that little routing overhead is generated when network topology is changed if the broken links are locally repaired. The idea is that nodes are assigned some degree of gradient according to their neighboring nodes. When some link breaks, the nodes are able to reset the degree of gradient based on its new neighboring nodes. Data packets are forwarded towards the nodes with the steepest gradient. Field-based routing is especially suitable for highly dynamic network topologies. Therefore, it is a scalable routing protocol in terms of network dynamics. It is another approach to make MANETs scalable.

4.2.5 Dynamic Addressing Routing Protocols

A dynamic addressing routing protocol was proposed in [56]. This approach takes the advantage of hierarchical routing by employing hierarchical addressing. The idea is that node addresses are hierarchically allocated according to physical topology. Therefore, the nodes address implies the physical topology relationship. Consequently, routing loops are resolved easily because of the structural address. Another advantage of this approach is that substantial routing table storage is saved because of the hierarchical addressing. However, one disadvantage is that node addresses have to be dynamically updated when physical topology is changed, which might incur routing overhead. Dynamic addressing routing is an approach to make MANETs scalable by network size.

4.2.6 Cognitive Routing Protocols

In recent years, several cognitive routing protocols have been proposed (e.g. [57-60). The main idea of them is that a proper and adaptive network topology should be constructed by the nodes using the cognitive functions which make the prediction based on past experience. The nodes in CRNs employ machine learning techniques to learn past experience and make wise decisions by predicting future network conditions. The cognitive protocol architecture should be a cross-layer optimized architecture where the lower layer knowledge of wireless medium is shared with network layer. In [62], the authors propose a capacity-based routing protocol for CRNs which adopts a novel metric for route selection by measuring and predicting the traffic pattern. In [63], the authors propose a routing protocol which adopts a novel metric predicted based on spectrum usage history to reflect the overall state. In [64], the authors propose a routing metric which reflects the conditions of spectrum resources such as channel availability and potential traffic delay. In [65], the authors proposed a spectrum aware mesh routing protocol (SAMER) which balances between long-term route stability and short-term opportunistic performance. It employs a novel metric for route selection to opportunistically route traffic over the paths with higher spectrum availability and quality. In [66], the authors proposed a high throughput spectrum aware routing protocol (SPEAR) which is a robust and efficient distributed channel assignment and routing protocol for dynamic spectrum networks based on two principles: integrated spectrum and route discovery for robust multi-hop path formation and distributed path reservations to minimize inter- and intra-flow interference.

Currently, path selection and frequency selection are common topics for cognitive routing protocols. We argue that cognitive routing is another promising approach to make MANETs scalable.

4.3 Approach

In this section, we discuss the details of our proposed routing protocol.

4.3.1 Overview of SCRP

As illustrated in the previous section, on-demand routing is able to effectively save the routing overhead when the frequency of the session initiation query is less that the frequency of topology changes. In practice, this mechanism is very useful. However, if the density of node distribution in an area is high, a flooding of control packets from the source node to the destination node might heavily impact the network performance. In wireless networks, the source node has to unconsciously flooding control packets along all links or even through all available frequencies to find the shortest path to reach the destination node. For some scenarios where the network topology is unstable, the flooding of control packets might seriously affect the overall network performance because the control packets will consume too much network bandwidth. Therefore, the flooding of control packets is a big issue in wireless networks.

Similar to on-demanding routing protocols, SCRP enables the nodes to trigger the routing updates reactively when necessary. However, unlike the ordinary on-demand routing protocols, the proposed SCRP enables the nodes to equip the cognitive engine to intelligently flood the control packets. The cognitive engine is implemented by neural network machine learning method which is able to provide a node with pre-processed network information before it joins a network and continuously update the network information after it joins a network. Cognitive radios with cognitive engine are aware of history and able to learn the trend of network changes based on past experience. They perform the routing functions in a distributed manner without a centralized database and the distributed information at each node should be up-to-date and consistent with each other. In this section, the cognitive engine for the nodes performs the scalable flooding protocol.

SCRP is designed for the multi-channel scenarios where the nodes can simultaneously use multiple interfaces to transmit packets over different frequencies. Consequently, it should perform not only next-hop node assignment along a path in space dimension but also frequency assignment for the links between the neighboring nodes in spectrum dimension. The proposed scalable flooding protocol can be viewed as two cognitive functions, scalable space flooding protocol working in space dimension and scalable spectrum flooding protocol working in spectrum dimension. In space dimension, the cognitive radios should minimize the flooding region. In spectrum dimension, the cognitive radios should minimize the transmission frequencies to avoid the interference. However, the challenge of SCRP is how to save the routing overhead while maintaining the network performance.

4.3.2 Scalable Space Flooding Protocol

As the name implies, the scalable space flooding protocol is used to make the flooding scalable by controlling the flooding region. It is an evolved version of our own work. In the previous effort, we developed a new metric, the throughput increment. It is defined as the predicted throughput after a new application joins minus the current throughput. It is this metric that determines future overall throughput. The predicted throughput increment is estimated based on a predicted channel type and channel capacity using a neural network machine learning method. Simulation results show that throughput increment is an excellent metric for path selection.

In this work, the metric, throughput increment, is utilized for the scalable space flooding protocol. The idea is that the links in the network are categorized into L levels according to the value of link metric. Inspired from QoS-oriented routing protocol, the control packets for SCRP have a new field indicating the desired metric level L_i by the source node. In this section, we assume that L_0 is the lowest metric level and L_{L-1} is the highest metric level. If the lowest metric level is indicated by the control packets, the intermediate nodes should relay them unconditionally. However, if some higher metric level is indicated by the control packets, the intermediate nodes should estimate the channel conditions of the corresponding receiving link before they relay the packets. If the estimated channel condition of the receiving link does not meet the desired metric level of the source node, the intermediate nodes will suppress the control packets. Otherwise, they relay the control packets.

The detailed description of the scalable space flooding protocol is illustrated as follows. When a source node needs to transmit some data packets to a destination node, it first checks its forwarding table. If there is an existing forwarding table entry for the destination node, the source node simply transmits the data packets to the next hop node indicated by the table entry. However, if the destination node is not reachable from the source node, the procedures of the scalable space flooding protocol starts. The steps of the protocol procedures are listed as follows.

The source node sets the desired metric level with the highest metric level L_{L-1} in the control packets and floods the control packets to the neighboring nodes. However, if the source node does not receive the acknowledgement packet from the destination node within the timeout period, it constructs another control packet and set the desired metric level as L_i which is one level lower than the previous metric level. If the source node sets the desired metric level with the lowest metric level L_0 in the control packets, it should be able to receive the acknowledgement packet from the destination node if they are reachable.

When a node receives a control packet, it first checks if the destination address is same as its address. If they are not same, it means that the receiving node is the intermediate node. Intermediate node always estimates the channel conditions of the receiving link before it determines whether to relay the control packets. There are three cases which the intermediate nodes need to consider. First, it has not seen any control packet from the source node before. In this case, if the estimated channel condition of the receiving link is better than the desired metric level of the source node, the intermediate will relay the control packet. Otherwise, the control packet will be suppressed. Second, the intermediate node has already seen some control packet from the source node before. In this case, it still estimates the channel condition of the receiving link as before. However, it only relays the control packets when the newly estimated channel condition is better than memorized estimated channel condition and is also better than the desired metric level of the source node. Otherwise, the control packet will be suppressed.

If the destination address is same as its address, it means that the receiving node is the destination node. There are two cases which the destination node needs to consider. First, it has not transmitted an acknowledgement packet to the source node before. In this case, it should construct an acknowledgement packet for the source node without estimating the channel conditions of the receiving link because this receiving link might be the only link available to the source node. Second, it remembers that it has transmitted an acknowledgement packet to the source node before and a new route request packet is received from the same source node. In this case, the destination node needs the check the channel conditions of the receiving link because the new link might be a better alternative link to the source node. If the estimated channel condition of the receiving link is better than the channel condition indicated by the existing routing table entry. The destination node will replace the old entry with the new entry and transmit another acknowledgement packet back to the source node. So the source node is informed that a better alternative path exists.

Figure 4.1 shows an example of the scalable space flooding protocol. In this scenario, the cognitive engine categorizes the links into four levels, $L_0 \sim L_3$. The arrows indicate the flows of the packet transmission. Yellow, red, green and black nodes are intermediate node. The two blue nodes are the source node and the destination node. Black nodes will relay the control packets with the desired metric level from $L_0 \sim L_3$. However, yellow nodes only relay the control packets when L_0 is the desired metric level. In general, the accumulated path condition will be worse as the path between the source node and the intermediate node becomes longer. Figure 1 clearly shows the advantage of the scalable space flooding protocol which is able to intelligently limit the flooding region. For the first trial, the source node sets the desired metric level with the highest metric level L_3. When black nodes receive the control packets, it relays the control packets since the estimated channel condition of the receiving link is better than the desired metric level indicated. However, the control packets are suppressed by the green nodes and the flooding region of the control packets are limited. Luckily, one of the black nodes relays the control packet to the destination node. It constructs an acknowledgement packet and transmits it back to the source node within the timeout period. The connection between the destination node and the source node is established without disturbing the yellow nodes, red nodes and some of the green nodes, which saves the routing overhead.

Figure 4.2 shows another example of the scalable space flooding protocol. In this scenario, the source node is not lucky and it has to transmit the control packets multiple times until L_0 is set in the desired metric level. When L_0 is reached, all intermediate nodes have to relay the control packets. However, black nodes have to relay the control packets four times, green nodes have to relay the control packet three times and red nodes have to relay the control packets twice when the source node is decreasing its desired metric level from three to zero. In other words, in this specific scenario, scalable space flooding protocol performs even worse because more routing overhead is incurred compared to the traditional flooding algorithm where all intermediate nodes relay the control packets unconditionally. However, in the mathematical analysis section, we will show that this is the worst case of space flooding protocol and is not the general case.

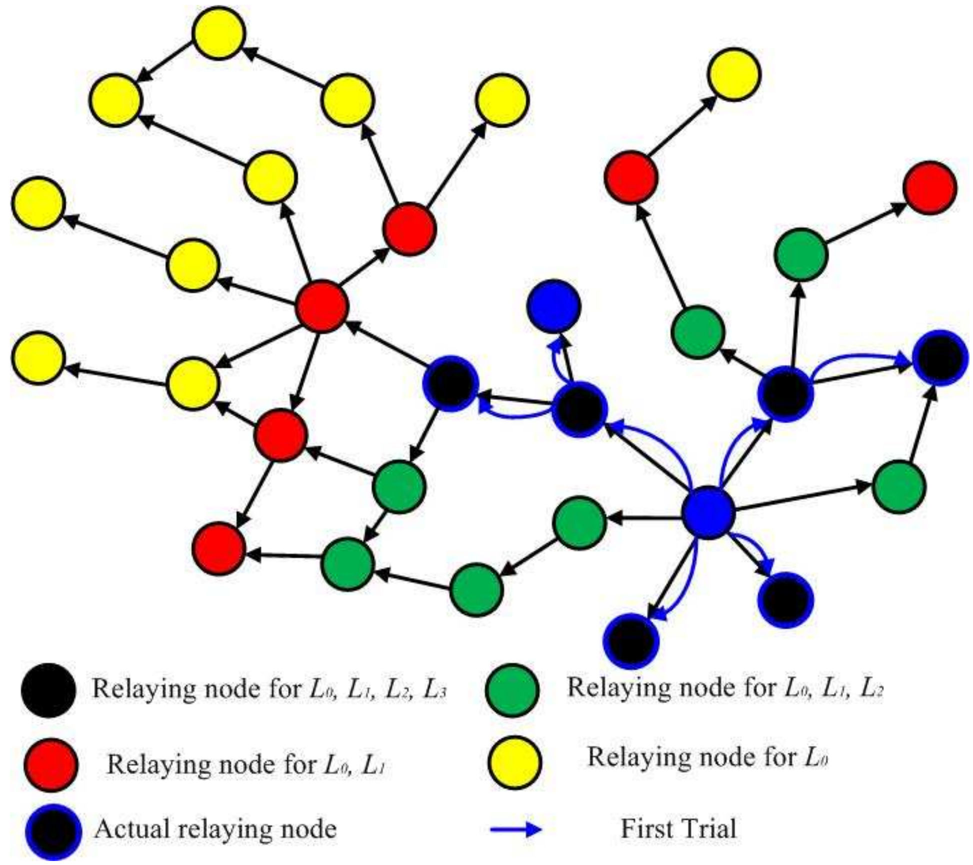


Figure 4.1 Example of scalable space flooding protocol when there is a path with high level of metric between source node and destination node.

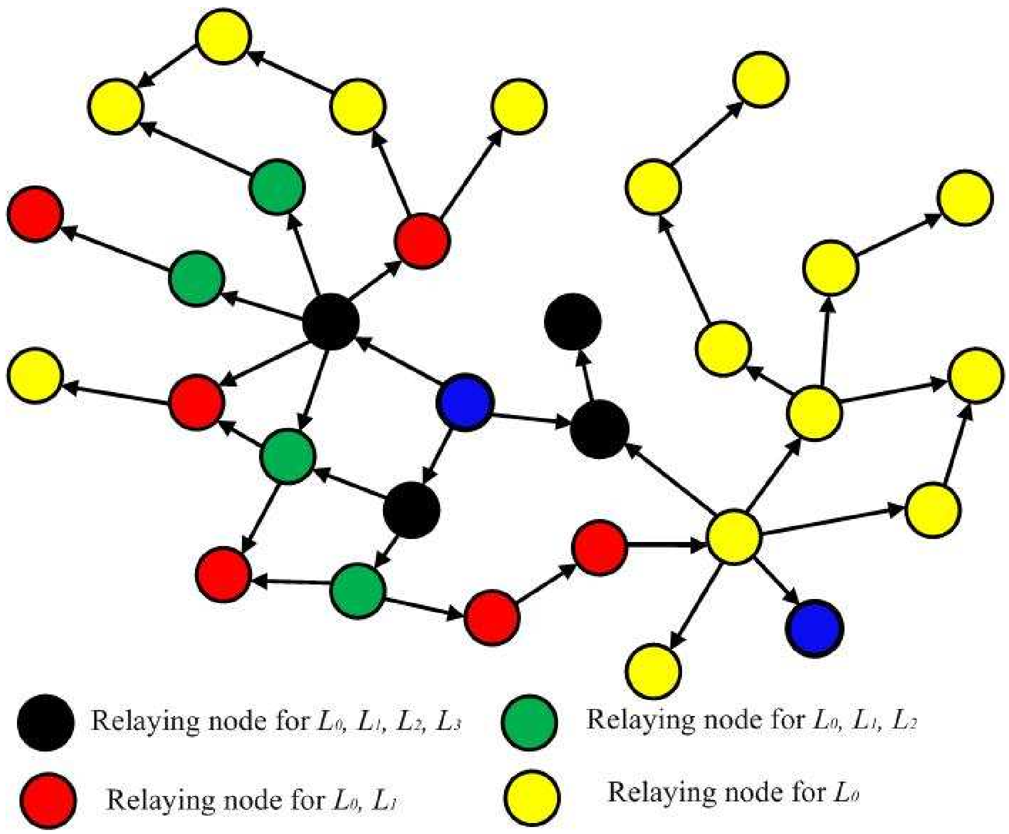


Figure 4.2 Example of scalable space flooding protocol when there are only paths with the lowest level of metric between source node and destination node.

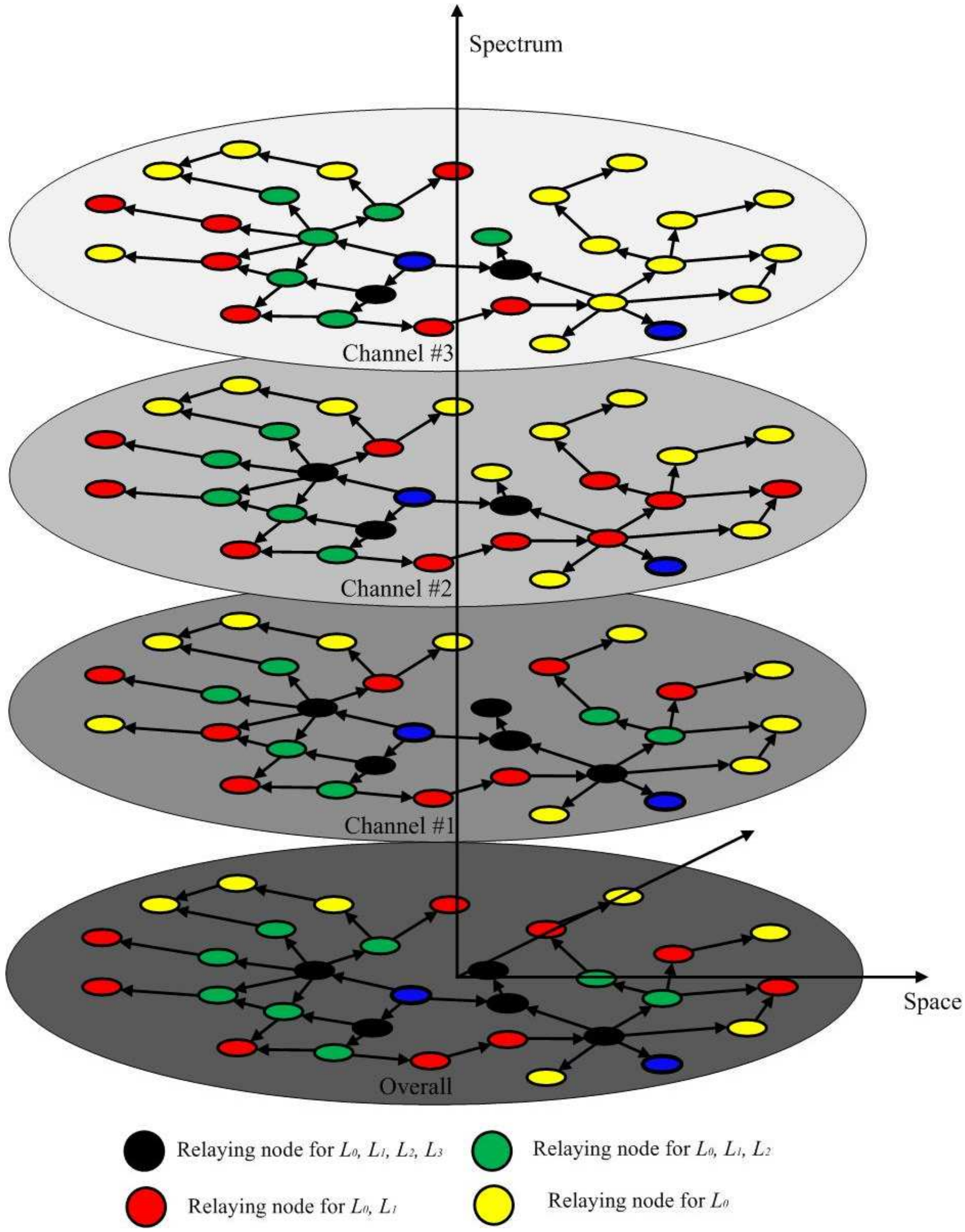


Figure 4.3 Example of the overall effect of scalable spectrum flooding protocol combined with scalable space flooding protocol.

4.3.3 Scalable Spectrum Flooding Protocol

As the name suggests, the scalable spectrum flooding protocol is used to make flooding scalable by limiting the number of frequencies used for control packets. It is a modification of some of our previous work. In that work, we developed a metric, the delay of RREQ packets, for frequency selection. In wireless communication, packet delay is mainly determined by queuing delay which is affected by traffic load and channel capacity. Delay of RREQ packets is used to estimate queuing delay. The frequency over which delay of RREQ packets is small is predicted to have good channel conditions. Simulation results show that it is an excellent metric for frequency selection.

In this work, this metric, delay of RREQ packets, is utilized for the scalable spectrum flooding protocol. The idea is that the intermediate nodes determine which frequencies should be selected to relay the control packets by its internal frequency ranking system. The internal ranking system for each node is continuously updated as the communication environment changes. In this section, we assume that there are F frequencies available and each node should select f frequencies which is a small subset of F to transmit the control packets. The internal ranking system of a node grants the frequency a score from 0 to F according to the channel conditions.

The ranking system works as follows. Each node is constantly monitoring F frequencies to gather their channel conditions. When the source node constructs a control packet, it sets the sequence number and the destination node ID. There are three cases which the intermediate nodes need to consider. First, the intermediate node received a control packet which it has not seen before. In this case, it means that a new connection request is received. It also means that the receiving frequency F_i and the previous hop node N_i might be the best frequency and the best link if it is relaying the data packets because the control packet is firstly received from this link and this frequency. Therefore, the internal ranking system will grant the receiving frequency the highest score, F. Second, the intermediate node received a control packet which it has seen before but through a different frequency F_j. In this case, it means that some alternative frequency exists for the link. The ranking system of the node will grant this frequency a score according to how many times it has seen the control packets before. If it has seen the control packet k times before, it will grant the receiving frequency a score F-k. Third, the intermediate node received a control packet which it has seen before and through the same frequency F_j. In this case, the ranking system will do nothing but the scalable space flooding protocol needs to handle the control packet. In a word, duplicated control packets from different frequencies still affect the internal ranking system of a node, which is an effective way to make sure the ranking system up-to-date.

As mentioned before, the intermediate node needs to select f frequencies to flood the control packets. The selected subset of frequencies is further divided in to two parts, intelligently selected frequencies F_c and randomly selected frequencies F_r. Intelligently selected frequencies are chosen based on its internal ranking system by picking the top F_c frequencies. Randomly selected frequencies are chosen from the remaining frequencies excluding the intelligently selected frequencies. It is obvious that the choice of the radio R between F_c and F_r is a tradeoff. If R is high, the probability that the neighboring nodes select the same set of frequencies for control packets will high because the channel conditions of the neighboring area are similar. In this case, it might result in run time network congestion if some neighboring nodes select the same frequency for data transmission. On the other hand, if R is low, the internal ranking system will be ineffective. Therefore, the balance of the radio R is important, so that the nodes can effectively take the advantage of the internal ranking system while avoiding the run time network congestion.

Figure 4.3 shows an example of the overall effect of the scalable spectrum flooding protocol combined with the scalable space flooding protocol. In this scenario, each node should select three frequencies to relay the control packets. The cognitive engine categorizes the links into four levels, $L_0 \sim L_3$. The arrows indicate the flows of the packet transmission. Yellow, red, green and black nodes are intermediate node. The two blue nodes are the source node and the destination node. Black nodes will relay the control packets with the desired metric level from $L_0 \sim L_3$. However, yellow nodes only relay the control packets when L_0 is the desired metric level. The view for each layer of frequency is illustrated in figure 1 and figure 2. However, the bottom layer of figure 4.3 shows the overall effect of the whole network after combining the channel conditions of three frequencies. In this scenario, it is obvious that the source node is able to find the destination node after the first trial. Therefore, the scalable spectrum flooding protocol is able to make the network scalable by limiting the number of frequencies used for the control packets, while enhancing the performance of the scalable space flooding protocol.

4.4 Mathematical Analysis

In this section, we discuss the mathematical analysis of scalable cognitive routing protocol (SCRP) and compare it with the traditional non-scalable routing protocols.

We model the communication network as a connected graph $G = (V, E, F),$ where V represents the set of vertices in G, E represents the set of edges connecting V and F represents the set of frequencies available in G.

The assumptions we made are listed below.

- The routing overhead is defined as the number of route request packets (RREQ) received by the nodes in a network. The number of RREQ packets usually dominates the total number of control packets in a network because the nodes flood RREQ packets in all directions, whereas route replay (RREP) packet is transmitted only along the selected path between the source node and the destination.
- The routing metric is uniformly divided into L levels. The probability of channel conditions of a link has a uniform distribution covering the entire L levels, which is independent of the location of the link, the frequency of the link and the estimated time.

4.4.1 Simple Case 1

In case 1, we assume that G is constructed by a linked list of V and F=1.In other words, we are considering a single-channel communication environment and the constructed network topology is simple. In this case, $E = V - 1$. According to the definition of routing overhead we made, in this case, routing overhead is the number of edges used to relay RREQ packets. Without loss of generality, we assume that the source node is in the middle of the linked list and the destination node might be on the left side or the right side of the source node.

4.4.1.1 Traditional non-scalable routing protocols

For traditional non-scalable routing protocols, the best case, worst case and average case of routing overhead are same as O(E) because at least half of the nodes in G have to unconditionally relay the control packets even though the destination node is on the other side. In other words, the nodes are not aware of the communication environment. Therefore, the routing overhead is at least E/2.

4.4.1.2 Best case for SCRP: O(1)

The best case of routing overhead for SCRP is O(1), which is explained as follows. If the destination node is just one hop away from the source node, the source node can find the destination node after the first trial. If the neighboring node of the source node on the other side of the destination node suppressed the control packets by applying the scalable space flooding protocol, only one routing overhead is incurred at this node and no additional routing overhead is incurred at its following nodes. Therefore, total number of routing overhead for the best case is just one.

4.4.1.3 Worst case for SCRP: O(E)

The worst case of routing overhead for SCRP is O(E), which happens when the source node is not lucky and has to transmit the control packets L times until L_0 is set in the desired metric level. As mentioned before, if the source node sets the desired metric level with the lowest metric level L_0 in the control packets, it should be able to receive the acknowledgement packet from the destination node if they are reachable. If the nodes on the other side of the destination node did not suppress the L control packets and relay them all the way to the end of the path, the number of overhead incurred on this side is L* E/2. The worst case happens when the communication environment is very bad on the destination node's side but it is very nice on the other side. In a word, total number of routing overhead for the worst case is at least L* E/2. It is obvious that the constant factor is L which is worse than traditional non-scalable routing protocols.

4.4.1.4 Average case for SCRP: O(E)

The average case of routing overhead for SCRP is O(E), which is explained as follows. As mentioned before, the source node will gradually decrease the desired metric level from the highest level L_{L-1} to the lowest level L_0. In other words, the source node will flood the control packets for at most L time. Let's consider the routing overhead on the other side of the destination node for each trial of flooding because the analysis of the routing overhead on the destination node's side is similar.

According to the probability theory, when the highest desired metric level L_{L-1} is set by the source node, the average routing overhead R_{L-1} is calculated as:

$$R_{L-1} = \frac{L-1}{L}*1+\frac{1}{L}*(\frac{L-1}{L}*2+\frac{1}{L}*(\frac{L-1}{L}*3+...+\frac{1}{L}*(\frac{L-1}{L}*(\frac{E}{2}-1)+\frac{1}{L}*\frac{E}{2}))...)) \quad (4.1)$$

$$= \frac{L-1}{L}*1+\frac{L-1}{L^2}*2+\frac{L-1}{L^3}*3+...+\frac{L-1}{L^{\frac{E}{2}-1}}*(\frac{E}{2}-1)+\frac{1}{L^{\frac{E}{2}}}*\frac{E}{2} \quad (4.2)$$

It is obvious that each term of R_{L-1} is computed by the probability that previous nodes relay the control packet and the current node suppressed the control packet times the corresponding number of routing overhead incurred. For example, the meaning of the first term $\frac{L-1}{L}*1$:

= Probability (Control packets stop at 1st node)

* Number (routing overhead)

The meaning of the second term $\frac{L-1}{L^2}*2$:

= Probability (Control packets doesn't stop at 1st node)

* Probability (Control packets stop at 2nd node)

* Number (routing overhead)

After understanding the meaning of R_{L-1}, we perform further computation based on formula (4.1):

$$R_{L-1} = \frac{L-1}{L}*1+\frac{1}{L}*(\frac{L-1}{L}*2+\frac{1}{L}*(\frac{L-1}{L}*3+...+\frac{1}{L}*(\frac{L-1}{L}*(\frac{E}{2}-1)+\frac{1}{L}*\frac{E}{2}))...))$$
$$= \frac{L-1}{L}*1+\frac{1}{L}*(\frac{L-1}{L}*2+\frac{1}{L}*(\frac{L-1}{L}*3+...+\frac{1}{L}*(\frac{L-1}{L}*(\frac{E}{2}-1)+\frac{1}{L}*(\frac{E}{2}-1)+\frac{1}{L}))...))$$
$$= \frac{L-1}{L}*1+\frac{1}{L}*(\frac{L-1}{L}*2+\frac{1}{L}*(\frac{L-1}{L}*3+...+\frac{1}{L}*((\frac{E}{2}-1)+\frac{1}{L}))...))$$
$$= \frac{L-1}{L}*1+\frac{1}{L}*(\frac{L-1}{L}*2+\frac{1}{L}*(\frac{L-1}{L}*3+...+\frac{1}{L}*(\frac{L-1}{L}*(\frac{E}{2}-2)+\frac{1}{L}*(\frac{E}{2}-1)+\frac{1}{L^2}))...))$$
$$= \frac{L-1}{L}*1+\frac{1}{L}*(\frac{L-1}{L}*2+\frac{1}{L}*(\frac{L-1}{L}*3+...+\frac{1}{L}*(\frac{L-1}{L}*(\frac{E}{2}-2)+\frac{1}{L}*(\frac{E}{2}-2)+\frac{1}{L}+\frac{1}{L^2}))...))$$
$$= \frac{L-1}{L}*1+\frac{1}{L}*(\frac{L-1}{L}*2+\frac{1}{L}*(\frac{L-1}{L}*3+...+\frac{1}{L}*((\frac{E}{2}-2)+\frac{1}{L}+\frac{1}{L^2}))...))$$

$$=\frac{L-1}{L}*1+\frac{1}{L}*(\frac{L-1}{L}*2+\frac{1}{L}*(\frac{L-1}{L}*3+...+\frac{1}{L}*((\frac{E}{2}-2)+\frac{1}{L}+\frac{1}{L^2}))...))$$

...

$$=\frac{L-1}{L}*1+\frac{1}{L}*(\frac{L-1}{L}*2+\frac{1}{L}*2+\frac{1}{L}+\frac{1}{L^2}+...+\frac{1}{L^{\frac{E}{2}-3}})$$

$$=\frac{L-1}{L}*1+\frac{1}{L}*(2+\frac{1}{L}+\frac{1}{L^2}+...+\frac{1}{L^{\frac{E}{2}-3}})$$

$$=\frac{L-1}{L}+\frac{1}{L}+\frac{1}{L}+\frac{1}{L^2}+...+\frac{1}{L^{\frac{E}{2}-2}}$$

$$=1+\frac{1}{L}+\frac{1}{L^2}+...+\frac{1}{L^{\frac{E}{2}-2}}$$

$$\approx\frac{1}{1-\frac{1}{L}}=\frac{L}{L-1}\approx 1$$

We found that if the number of metric level L is large compared to 1, the average routing overhead for the first trial R_{L-1} is roughly equal to 1. Another thing we found is that the average TTL of control packet is $\frac{L}{L-1}\approx 1$, if L is large compared to 1 because the routing overhead for a linked list graph is equal to TTL of control packet.

If the source node does not receive the acknowledgement packet from the destination node within the timeout period, it constructs another control packet and set the desired metric level as L_i which is one level lower than the previous metric level. Similar to the proof and calculation of R_{L-1}, it is easy to find that the average routing overhead $R_{L-i}\approx i$ and its average TTL of control packet is i.

At this point, we found the average routing overhead for each trial and its corresponding average TTL of control packet. We assume that the number of metric levels L is larger than the maximum TTL of control packets. For ease of analysis we let N as $\dfrac{E}{2}$, the average routing overhead R for SCRP is:

$$R=\frac{1}{N}*1+\frac{N-1}{N}*(\frac{2}{N}(1+2)+\frac{N-2}{N}(\frac{3}{N}(1+2+3)+\ldots$$
$$+\frac{3}{N}(\frac{N-2}{N}(1+2+\ldots+(N-2))+\frac{2}{N}(\frac{N-1}{N}(1+2+\ldots+(N-1))+\frac{1}{N}(1+2+\ldots+N)))\ldots)) \tag{4.3}$$

$$=\frac{\frac{(N-1)!}{(N-1)!}*1}{N^1}\sum_{i=1}^{1}i+\frac{\frac{(N-1)!}{(N-2)!}*2}{N^2}\sum_{i=1}^{2}i+\frac{\frac{(N-1)!}{(N-3)!}*3}{N^3}\sum_{i=1}^{3}i+\ldots$$
$$+\frac{\frac{(N-1)!}{2!}*(N-2)}{N^{N-2}}\sum_{i=1}^{N-2}i+\frac{\frac{(N-1)!}{1!}*(N-1)}{N^{N-1}}\sum_{i=1}^{N-1}i+\frac{(N-1)!}{N^{N-1}}\sum_{i=1}^{N}i$$

$$=\sum_{i=1}^{i=N-1}\frac{\frac{(N-1)!}{(N-i)!}i\sum_{j}^{j=i}j}{N^{i-1}}+\frac{(N-1)!}{N^{N-1}}\sum_{j}^{j=N}j \tag{4.4}$$

It is obvious that each term of R is computed by the probability that the trail is successful times its corresponding routing overhead. The probability that the trail is successful is relative to the TTL of control overhead. Formula (4.4) shows the general form of R and we perform further computation based on formula (4.3):

$$R=\frac{1}{N}*1+\frac{N-1}{N}*(\frac{2}{N}(1+2)+\frac{N-2}{N}(\frac{3}{N}(1+2+3)+\ldots$$
$$+\frac{3}{N}(\frac{N-2}{N}(1+2+\ldots+(N-2))+\frac{2}{N}(\frac{N-1}{N}(1+2+\ldots+(N-1))+\frac{1}{N}(1+2+\ldots+N)))\ldots))$$
$$=\frac{1}{N}*1+\frac{N-1}{N}*(\frac{2}{N}(1+2)+\frac{N-2}{N}(\frac{3}{N}(1+2+3)+\ldots$$

$$=\frac{1}{N}*1+\frac{N-1}{N}*(\frac{2}{N}(1+2)+\frac{N-2}{N}(\frac{3}{N}(1+2+3)+\ldots$$
$$+\frac{3}{N}(\frac{N-2}{N}(1+2+\ldots+(N-2))+\frac{2}{N}(\frac{N-1}{N}(1+2+\ldots+(N-1))$$
$$+\frac{1}{N}(1+2+\ldots+N-1)+\frac{1}{N}*N)))\ldots))$$

$$=\frac{1}{N}*1+\frac{N-1}{N}*(\frac{2}{N}(1+2)+\frac{N-2}{N}(\frac{3}{N}(1+2+3)+\ldots$$

$$+\frac{3}{N}(\frac{N-2}{N}(1+2+...+(N-2))+\frac{2}{N}(1+2+...+(N-1)+1)))...))$$

$$=\frac{1}{N}*1+\frac{N-1}{N}*(\frac{2}{N}(1+2)+\frac{N-2}{N}(\frac{3}{N}(1+2+3)+...$$

$$+\frac{3}{N}(\frac{N-2}{N}(1+2+...+(N-2))+\frac{2}{N}(1+2+...+(N-2)+N)))...))$$

$$=\frac{1}{N}*1+\frac{N-1}{N}*(\frac{2}{N}(1+2)+\frac{N-2}{N}(\frac{3}{N}(1+2+3)+...+\frac{3}{N}(1+2+...+(N-2)+2)...))$$

$$=\frac{1}{N}*1+\frac{N-1}{N}*(\frac{2}{N}(1+2)+\frac{N-2}{N}(\frac{3}{N}(1+2+3)+...+\frac{3}{N}(1+2+...+(N-3))+N)...))$$

...

$$=\frac{1}{N}*1+\frac{N-1}{N}*(\frac{2}{N}(1+2)+\frac{N-2}{N}(\frac{3}{N}(1+2+3)+...+\frac{N-3}{N}(1+2+3+N)))$$

$$=\frac{1}{N}*1+\frac{N-1}{N}*(\frac{2}{N}(1+2)+\frac{N-2}{N}(1+2+3+N-3))$$

$$=\frac{1}{N}*1+\frac{N-1}{N}*(1+2+N-2))$$

$$=\frac{1}{N}*1+\frac{N-1}{N}*(1+N))$$

$$=1+N-1=N=\frac{E}{2} \tag{4.5}$$

We found that for simple case 1, the average routing overhead for SCRP with L trials is $\frac{E}{2}$ which is same as the traditional non-scalable routing protocols.

4.4.2 General Case 2

We assume that G is constructed by a general graph where each node has at least one child and F=1.In other words, we are considering a single-channel communication environment and $V-1 \le E < V^2$. C_i represents the number of child for the node V_i and the average number of child for each V_i is C. We assume that the maximum TTL of control packet is P.

4.4.2.1 Traditional non-scalable routing protocols

For traditional non-scalable routing protocols, the best case, worst case and average case of routing overhead are same as O($\frac{C^{P+1}}{C-1}$). The nodes in G have to unconditionally relay the control packets. If the maximum TTL of control packet is P, a control packet will be flooding for P times. Since the average number of child for each V_i is C, the total number of routing overhead R is:

$$R = C + C^2 + ... + C^P = \sum_{i=1}^{P} C^i \quad (4.6)$$

$$= \frac{C(C^P - 1)}{C-1} \approx \frac{C^{P+1}}{C-1} \quad (4.7)$$

4.4.2.2 Best case for SCRP: O(C)

The best case of routing overhead for SCRP is O(C), which is explained as follows. If the destination node is just one hop away from the source node, the source node can find the destination node after the first trial. If all neighboring nodes of the source node suppressed the control packets by applying the scalable space flooding protocol, the total number of routing overhead is only C and no additional routing overhead is incurred at their following nodes. Therefore, total number of routing overhead for the best case is just one.

4.4.2.3 Worst case for SCRP: O($\frac{C^{P+1}}{C-1}$)

The worst case of routing overhead for SCRP is O(C^P), which happens when the source node is not lucky and it has to transmit the control packets L times until L_0 is set in the desired metric level. As mentioned before, if the source node sets the desired metric level with the lowest metric level L_0 in the control packets, it should be able to receive the acknowledgement packet from the destination node if they are reachable. If the nodes that cannot reach the destination node did not suppress the L control packets and relay them all the way to the end of the path, the number of overhead incurred on this side is L* $\frac{C^{P+1}}{C-1}$. The worst case happens when the communication environment is very bad on the path to the destination node but it is very nice on the path that cannot reach the destination node.

4.4.2.4 Average case for SCRP: <<O(C^P)

Similar to the simple case 1, it is easy to find that the average TTL of control packet for ith trial is i. Unlike traditional non-scalable routing protocols, the distribution of the location of destination nodes heavily impacts the average routing overhead R. Therefore, we will compare the average routing overhead R for two kinds of distributions, uniform distribution according to the network area and uniform distribution according to path length between the source node and the destination node.

If the location of destination node follows the uniform distribution according to the network area, the average routing overhead R is:

$$R=\frac{C}{C^{P}}*C+(1-\frac{C}{C^{P}})*(\frac{C^{2}}{C^{P}}(C+C^{2})+(1-\frac{C^{2}}{C^{P}})(\frac{C^{3}}{C^{P}}(C+C^{2}+C^{3})+\ldots$$

$$+(1-\frac{C^{P-3}}{C^{P}})(\frac{C^{P-2}}{C^{P}}(C+C^{2}+\ldots+C^{P-2})+(1-\frac{C^{P-2}}{C^{P}})(\frac{C^{P-1}}{C^{P}}(C+C^{2}+\ldots+C^{P-1})$$

$$+(1-\frac{C^{P-1}}{C^{P}})(C+C^{2}+\ldots+C^{P})))\ldots))$$

$$=\frac{C}{C^{P}}*C+(1-\frac{C}{C^{P}})*(\frac{C^{2}}{C^{P}}(C+C^{2})+(1-\frac{C^{2}}{C^{P}})(\frac{C^{3}}{C^{P}}(C+C^{2}+C^{3})+\ldots$$

$$+(1-\frac{C^{P-3}}{C^{P}})(\frac{C^{P-2}}{C^{P}}(C+C^{2}+\ldots+C^{P-2})+(1-\frac{C^{P-2}}{C^{P}})(\frac{C^{P-1}}{C^{P}}(C+C^{2}+\ldots+C^{P-1})$$

$$+(1-\frac{C^{P-1}}{C^{P}})(C+C^{2}+\ldots+C^{P-1})+C^{P}-C^{P-1})))\ldots))$$

$$=\frac{C}{C^{P}}*C+(1-\frac{C}{C^{P}})*(\frac{C^{2}}{C^{P}}(C+C^{2})+(1-\frac{C^{2}}{C^{P}})(\frac{C^{3}}{C^{P}}(C+C^{2}+C^{3})+\ldots$$

$$+(1-\frac{C^{P-3}}{C^{P}})(\frac{C^{P-2}}{C^{P}}(C+C^{2}+\ldots+C^{P-2})$$

$$+(1-\frac{C^{P-2}}{C^{P}})(C+C^{2}+\ldots+C^{P-1}+C^{P}-C^{P-1}))\ldots)))$$

$$=\frac{C}{C^{P}}*C+(1-\frac{C}{C^{P}})*(\frac{C^{2}}{C^{P}}(C+C^{2})+(1-\frac{C^{2}}{C^{P}})(\frac{C^{3}}{C^{P}}(C+C^{2}+C^{3})+\ldots$$

$$+(1-\frac{C^{P-3}}{C^{P}})(\frac{C^{P-2}}{C^{P}}(C+C^{2}+\ldots+C^{P-2})$$

$$+(1-\frac{C^{P-2}}{C^{P}})(C+C^{2}+\ldots+C^{P-2}+C^{P}))\ldots)))$$

$$=\frac{C}{C^{P}}*C+(1-\frac{C}{C^{P}})*(\frac{C^{2}}{C^{P}}(C+C^{2})+(1-\frac{C^{2}}{C^{P}})(\frac{C^{3}}{C^{P}}(C+C^{2}+C^{3})+\ldots$$

$$+(1-\frac{C^{P-3}}{C^{P}})(C+C^{2}+\ldots+C^{P-3}+C^{P})$$

$$\ldots$$

$$
\begin{aligned}
&= \frac{C}{C^P} * C + (1-\frac{C}{C^P}) * (\frac{C^2}{C^P}(C+C^2) + (1-\frac{C^2}{C^P})(\frac{C^3}{C^P}(C+C^2+C^3) + (1-\frac{C^3}{C^P})(C+C^2+C^3+C^P))) \\
&= \frac{C}{C^P} * C + (1-\frac{C}{C^P}) * (\frac{C^2}{C^P}(C+C^2) + (1-\frac{C^2}{C^P})(\frac{C^3}{C^P}(C+C^2+C^3) + (1-\frac{C^3}{C^P})(C+C^2+C^3) + C^P - C^3)) \\
&= \frac{C}{C^P} * C + (1-\frac{C}{C^P}) * (\frac{C^2}{C^P}(C+C^2) + (1-\frac{C^2}{C^P})(C+C^2+C^P)) \\
&= \frac{C}{C^P} * C + (1-\frac{C}{C^P}) * (C+C^P) = C^P
\end{aligned}
\tag{4.8}
$$

We found that if the location of destination node follows the uniform distribution according to the network area, the average routing overhead R for SCRP with L trials is smaller than the routing overhead $\frac{C^{P+1}}{C-1}$ for the traditional non-scalable routing protocols.

If the location of destination node follows the uniform distribution according to the path length between the source node and the destination node, the average routing overhead R is:

$$
\begin{aligned}
R &= \frac{1}{P} * C + (\frac{P-1}{P}) * (\frac{2}{P}(C+C^2) + (\frac{P-2}{P})(\frac{3}{P}(C+C^2+C^3) + \ldots \\
&\quad + \frac{3}{P}(\frac{P-2}{P}(C+C^2+\ldots+C^{P-2}) + \frac{2}{P}(\frac{P-1}{P}(C+C^2+\ldots+C^{P-1}) + \frac{1}{P}(C+C^2+\ldots+C^P)))\ldots))
\end{aligned}
\tag{4.9}
$$

$$
\begin{aligned}
&= \frac{1}{P} * C + (\frac{P-1}{P}) * (\frac{2}{P}(C+C^2) + (\frac{P-2}{P})(\frac{3}{P}(C+C^2+C^3) + \ldots \\
&\quad + \frac{3}{P}(\frac{P-2}{P}(C+C^2+\ldots+C^{P-2}) + \frac{2}{P}(\frac{P-1}{P}(C+C^2+\ldots+C^{P-1}) \\
&= \frac{1}{P} * C + (\frac{P-1}{P}) * (\frac{2}{P}(C+C^2) + (\frac{P-2}{P})(\frac{3}{P}(C+C^2+C^3) + \ldots \\
&\quad + \frac{3}{P}(\frac{P-2}{P}(C+C^2+\ldots+C^{P-2}) + \frac{2}{P}(C+C^2+\ldots+C^{P-1} + \frac{\frac{1!}{0!}C^P}{P}))\ldots))) \\
&= \frac{1}{P} * C + (\frac{P-1}{P}) * (\frac{2}{P}(C+C^2) + (\frac{P-2}{P})(\frac{3}{P}(C+C^2+C^3) + \ldots + \frac{3}{P}(\frac{P-2}{P}(C+C^2+\ldots+C^{P-2}) \\
&\quad + \frac{2}{P}(C+C^2+\ldots+C^{P-2}) + \frac{\frac{2!}{1!}C^{P-1}}{P} + \frac{\frac{2!}{0!}C^P}{P^2})\ldots)))
\end{aligned}
$$

$$= \frac{1}{P} * C + (\frac{P-1}{P}) * (\frac{2}{P}(C + C^2) + (\frac{P-2}{P})(\frac{3}{P}(C + C^2 + C^3) + ...$$

$$+ \frac{3}{P}(C + C^2 + ... + C^{P-2} + \frac{\frac{2!}{1!}C^{P-1}}{P} + \frac{\frac{2!}{0!}C^P}{P^2})...)))$$

...

$$= \frac{1}{P} * C + (\frac{P-1}{P}) * (\frac{2}{P}(C + C^2) + (\frac{P-2}{P})(\frac{3}{P}(C + C^2 + C^3)$$

$$+ \frac{P-3}{P}(C + C^2 + C^3) + \frac{\frac{(P-3)!}{(P-4)!} * C^4}{P} + \frac{\frac{(P-3)!}{(P-5)!} * C^5}{P^2} + ... + \frac{\frac{(P-3)!}{1!}C^{P-1}}{P^{P-4}} + \frac{\frac{(P-3)!}{0!}C^P}{P^{P-3}}))) \backslash$$

$$= \frac{1}{P} * C + (\frac{P-1}{P}) * (\frac{2}{P}(C + C^2) + (\frac{P-2}{P})(\frac{3}{P}(C + C^2 + C^3)$$

$$+ \frac{P-3}{P}(C + C^2 + C^3) + \frac{\frac{(P-3)!}{(P-4)!} * C^4}{P} + \frac{\frac{(P-3)!}{(P-5)!} * C^5}{P^2} + ... + \frac{\frac{(P-3)!}{1!}C^{P-1}}{P^{P-4}} + \frac{\frac{(P-3)!}{0!}C^P}{P^{P-3}}))$$

$$= \frac{1}{P} * C + (\frac{P-1}{P}) * (\frac{2}{P}(C + C^2) + (\frac{P-2}{P})(C + C^2 + C^3 + \frac{\frac{(P-3)!}{(P-4)!} * C^4}{P} + ... + \frac{\frac{(P-3)!}{0!}C^P}{P^{P-3}}))$$

$$= \frac{1}{P} * C + (\frac{P-1}{P}) * (C + C^2 + \frac{\frac{(P-2)!}{(P-3)!}C^3}{P} + \frac{\frac{(P-2)!}{(P-4)!} * C^4}{P^2} + ... + \frac{\frac{(P-2)!}{0!}C^P}{P^{P-2}}) \backslash$$

$$= C + \frac{\frac{(P-1)!}{(P-2)!}C^2}{P} + \frac{\frac{(P-1)!}{(P-3)!}C^3}{P^2} + \frac{\frac{(P-1)!}{(P-4)!} * C^4}{P^3} + ... + \frac{\frac{(P-1)!}{0!}C^P}{P^{P-1}}$$

$$= \sum_{i=1}^{P} \frac{\frac{(P-1)!}{(P-i)!}C^i}{P^{i-1}} < \frac{\frac{(P-1)!}{0!}C^P}{P^{P-1}} \approx \frac{\sqrt{2\pi(P-1)}(\frac{P-1}{e})^{P-1}C^P}{P^{P-1}} \approx \frac{C^P}{e^{P-1}} << C^P \tag{4.10}$$

We found that if the location of destination node follows the uniform distribution according to the path length between the source node and the destination node, the average routing overhead R for SCRP with L trials is much smaller than $\frac{C^{P+1}}{C-1}$ for the traditional non-scalable routing protocols.

4.5 Simulations

In this section, we provide simulation results for SCRP using Qualnet 4.0.

To show the benefits of SCRP, it is compared with the spectrum-aware routing protocol (SARP), which is our own previous work, and the well-known AODV.

SARP is described as follows. It is an on-demand cognitive routing protocol. Neural network machine learning is used to make nodes aware of history. It employs two cognitive functions, an intelligent multi-interface selection function (MISF) and an intelligent multi-path selection function (MPSF). The throughput increment is adopted by MPSF for path selection and the delay of RREQ packets is adopted by MISF for frequency selection. Simulation results show that SARP improves network performance. However, the main drawback of SARP is lack of scalability. RREQ packets are unconsciously flooded along all links and over all frequencies, which incurs significant overhead.

We present the following metrics, with 95% confidence intervals, to compare the performance and scalability of SCRP with SARP and AODV.

- Overhead: Number of RREQ packets received by nodes, which dominates the number of route control packets.
- Throughput: Average rate of successful message delivery measured in Kbits per second, which reflects network performance.

4.5.1 Scalability with Network Size

In the first experiment, we show how performance is affected when the number of nodes increases. We created a scenario which has 12 applications distributed in a 600m by 1500m region. We vary the number of nodes. There are three available frequencies. Automatic rate fallback is enabled. UDP is employed as the transport layer protocol. IEEE 802.11 is employed as the MAC protocol. We compare 3-interface SCRP with 3-interface SARP and 1-interface AODV.

Figure 4.4 and figure 4.5 show the comparison of overhead and throughput respectively as a function of number of nodes.

As expected, routing overhead is increased as the number of nodes increases and throughput is decreased as the number of nodes increases. In this experiment, the space flooding protocol is working and spectrum flooding protocol is not working because the number of available frequencies is three which is equal to the number of interfaces. In other words, this experiment only shows how performance is affected by the space flooding protocol.

Compared to 3-interface SARP, 3-interface SCRP saves routing overhead dramatically to be almost same as 1-interface AODV. The reason is as follows. SCRP intelligently selects a subset of nodes as relay nodes. However, SARP and AODV use all nodes as relay nodes. They do not control the flooding region. Because of this, 3-interface SARP incurs much more routing overhead than 3-interface SCRP. On the other hand, 1-interface AODV is not a multi-channel routing protocol. Although 1-interface AODV seems to have almost same routing overhead as 3-interface SCRP, it is able to utilize only one frequency rather than three frequencies. Further, the routing overhead of 3-interface SARP and 1-interface AODV is increases faster as the number of nodes increases. The routing overhead of 3-interface SCRP increases slower as the number of nodes increases because of the space flooding protocol.

These results show that the performance of SCRP scales better in terms of network size than SARP and AODV.

Compared to 3-interface SARP, 3-interface SCRP improves network performance considerably. The reason is as follows. The space flooding protocol intelligently selects nodes on potentially best paths as relay nodes and the other nodes as suppressed nodes. It does not sacrifice much performance and avoids interference between nodes because many nodes become suppressed nodes. Network performance benefits from the space flooding protocol especially for scenarios with a large number of nodes. On the other hand, compared to 1-interface AODV, 3-interface SCRP increases network performance significantly because 3-interface SCRP is able to utilize three frequencies and 1-interface AODV is able to utilize only one frequency. In other words, 3-interface SCRP should have much more throughput than 1-interface AODV because of the advantages of multi-channel routing protocols. The simulation results show that it has about five to eight times the throughput as 1-interface AODV. These results show that network performance of SCRP is better than SARP and AODV for scenarios with a large number of nodes.

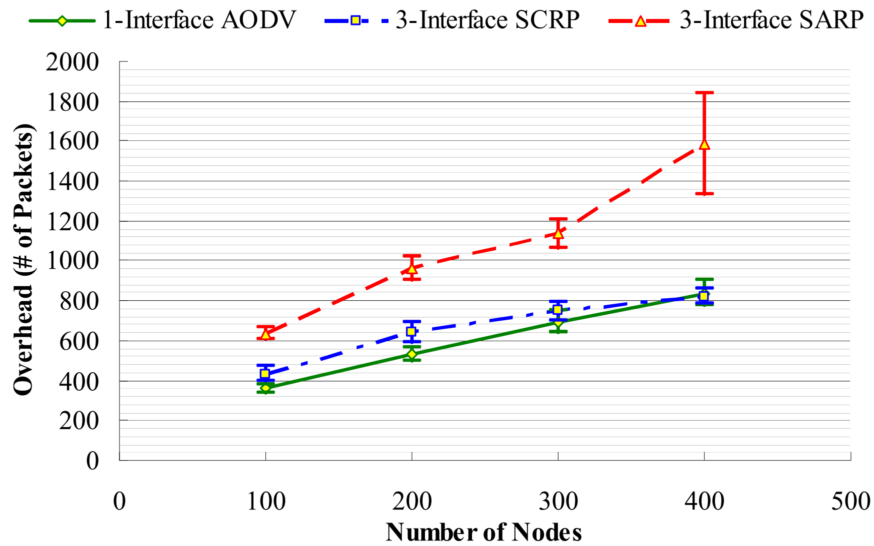


Figure 4.4 Comparison of overhead as a function of number of nodes

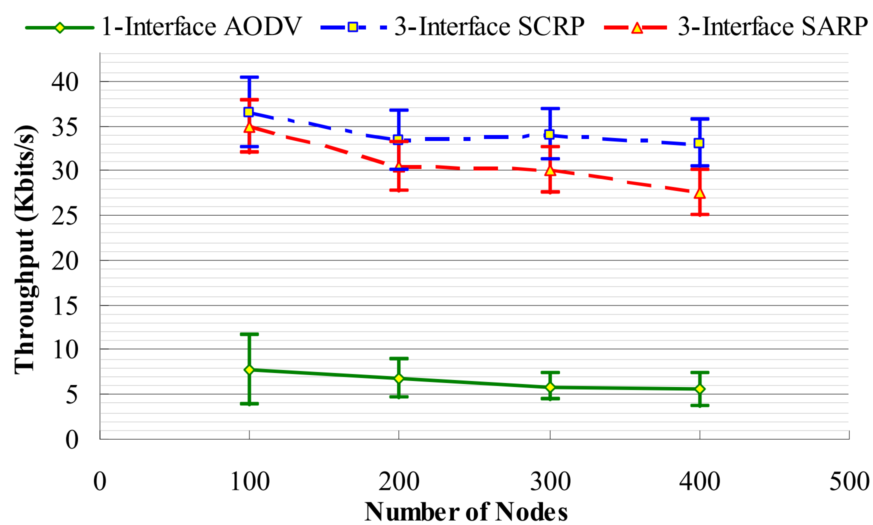


Figure 4.5 Comparison of throughput as a function of number of nodes

4.5.2 Scalability with Network Dynamics

In the second experiment, we show how performance is affected when node velocity increases. We created a scenario which has 12 applications and 200 nodes distributed in a 2500m by 2500m region. We vary node velocity. There are three available frequencies. Automatic rate fallback is enabled. UDP is the transport layer protocol and IEEE 802.11 is employed as the MAC protocol. We again compare 3-interface SCRP with 3-interface SARP and 1-interface AODV.

Figure 4.6 and figure 4.7 show a comparison of overhead and throughput respectively as a function of average velocity.

As expected, throughput is decreased as average velocity increases. Again in this experiment, the space flooding protocol is exercised and the spectrum flooding protocol is not because the number of available frequencies is three that is equal to the number of interfaces, so this only shows how performance is affected by the space flooding protocol.

Compared to 3-interface SARP and 1-interface AODV, 3-interface SCRP saves significant routing overhead. The reason is similar to the first experiment. In MANETs, nodes trigger routing updates frequently because of the dynamic network topology. The space flooding protocol intelligently selects nodes on potentially best paths as relay nodes and the other nodes as suppressed nodes, which saves routing overhead. On the other hand, network performance benefits from the space flooding protocol for highly dynamic network topologies. Compared to 3-interface SARP, 3-interface SCRP increases throughput considerably. Compared to 1-interface AODV, it increases throughput significantly. These results show that routing performance of SCRP scales better in terms of network dynamics than SARP and AODV.

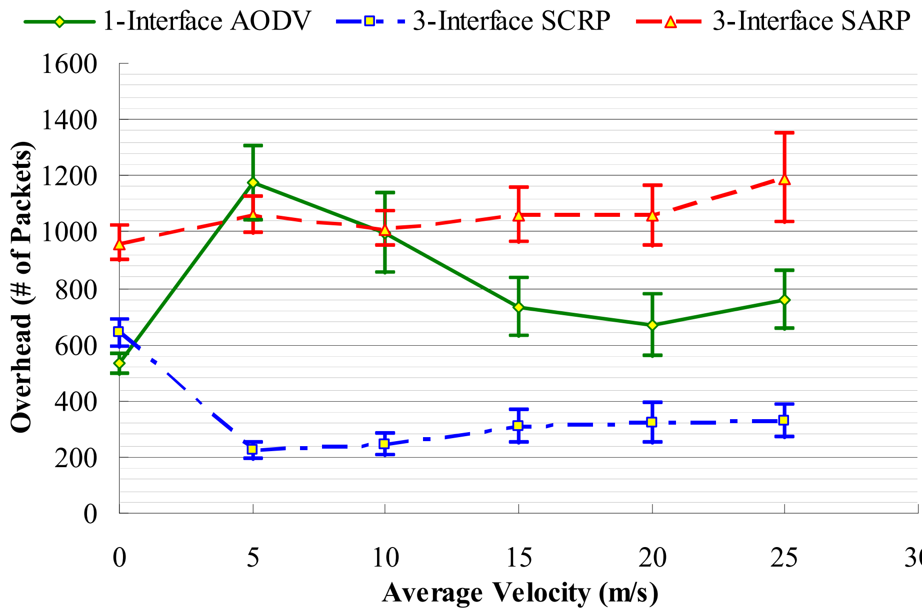


Figure 4.6 Comparison of overhead as a function of average velocity

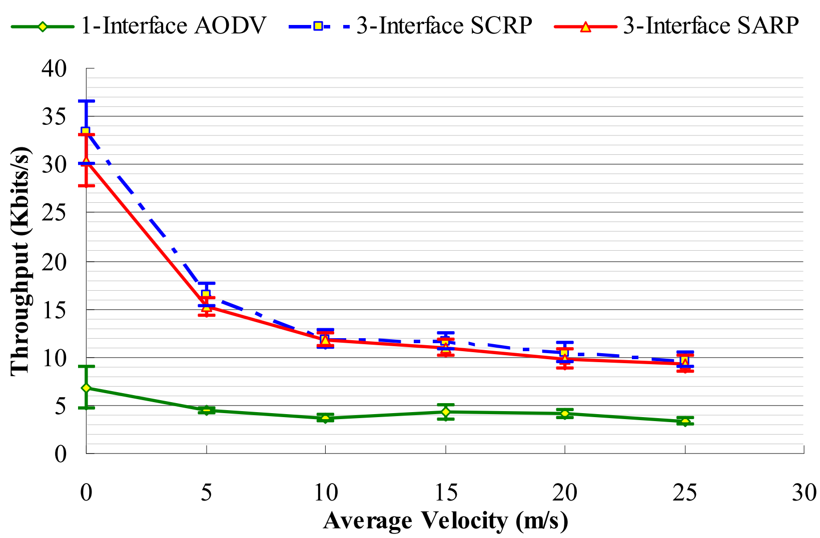


Figure 4.7 Comparison of throughput as a function of average velocity

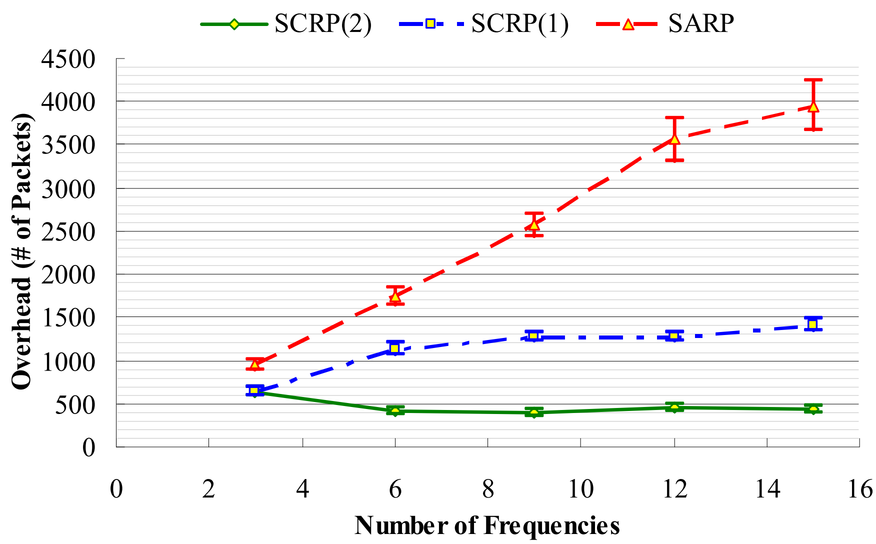


Figure 4.8 Comparison of overhead as a function of number of frequencies

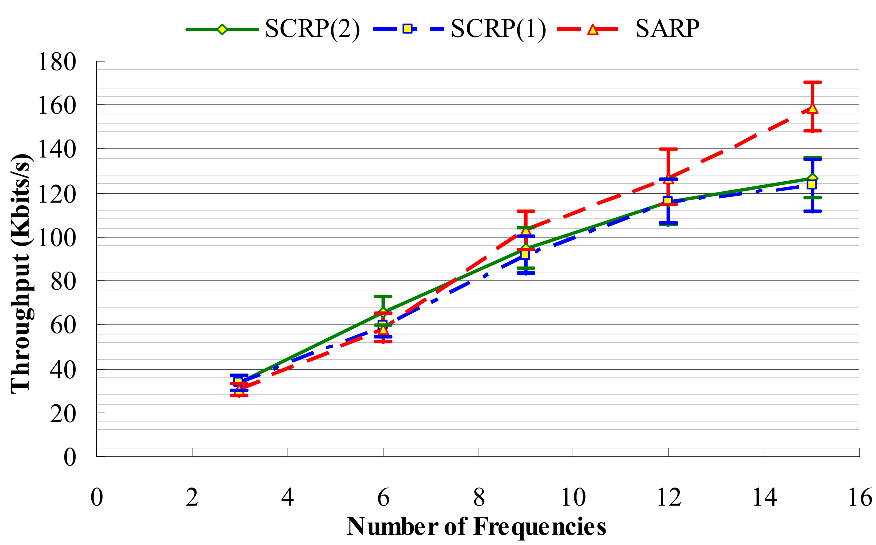


Figure 4.9 Comparison of throughput as a function of number of frequencies

4.5.3 Scalability with Network

4.5.4 Spectrums

In the third experiment, we show how performance is affected when the number of frequencies increases. We created a scenario which has 12 applications and 200 nodes distributed in a 600m by 1500m region. We vary the number of frequencies. The frequencies have different shadowing means, such as 4, 6, 8, 10 or 12, and the same Ricean K factor, 16. Automatic rate fallback is enabled. UDP is employed as the transport layer protocol And IEEE 802.11 as the MAC protocol. To show the benefits of the spectrum flooding protocol, we compare SCRP employing a two-dimensional intelligent flooding protocol, denoted as "SCRP(2)", with SCRP employing only the space flooding protocol, shown as "SCRP(1)", and SARP .
Figure 4.8 and figure 4.9 show a comparison of overhead and throughput respectively as a function of number of frequencies. As expected, throughput is increased as the number of frequencies increases.

Compared to SCRP(1) and SARP, SCRP(2) saves routing overhead significantly. The reason is that SCRP(2) intelligently selects a subset of frequencies as flooding frequencies. Routing overhead is almost constant as the number of frequencies increases because of the spectrum flooding protocol. On the other hand, SCRP(1) and SARP select all frequencies as flooding frequencies. As a result, routing overhead is increased as the number of frequencies increases. These results show that the performance of SCRP(2) scales better in terms of network spectrum than SCRP(1) and SARP.

Compared to SCRP(1) and SARP, SCRP(2) has almost same throughput, because the spectrum flooding technique intelligently selects the top two frequencies in total score as flooding frequencies and randomly selects another frequency to avoid similar frequency selection with neighboring nodes. In other words, SCRP(2) lets nodes flood RREQ packets over a set of potentially best frequencies instead of all frequencies. The simulation results show that these routing protocols have almost the same throughput, which means the spectrum flooding protocol sacrifices little performance. These results show that the network performance of SCRP(2) can be almost same as SCRP(1) and SARP even though nodes flood RREQ packets over a limited set of frequencies.

4.6 Conclusions

We have investigated the problem of scalable routing in multi-channel mobile ad-hoc networks. We have proposed a novel scalable cognitive routing protocol (SCRP) which is an on-demand cognitive routing protocol.

SCRP employs the modified intelligent flooding protocol, a novel approach for scalable flooding. Neural network machine learning is adopted to make nodes aware of history. Lower layer knowledge is shared with the network layer to help SCRP work properly and efficiently. An intelligent flooding protocol saves significant routing overhead because nodes selectively flood RREQ packets along predicted strong links and over predicted good frequencies. The intelligent flooding protocol can be divided into two parts working over two dimensions, the space flooding and spectrum flooding protocols. Simulation results show that the routing performance of SCRP scales better than SARP and AODV in terms of network size, network dynamics and network spectrums.

Chapter 5

Multi-Path Optimized Cognitive Routing Protocol

5.1 Introduction

Multipath routing protocols have been proposed in many papers for both wired networks and wireless networks. In wired networks, the major consideration is how nodes utilize multiple paths such as back up redundant paths and load balancing among available paths. Currently, OSPF and BGP-4 are the dominant routing protocols for wired networks. They both include multipath capabilities. Based on desired metrics, nodes are able to balance load among paths or support redundant paths. On the other hand, mobile ad-hoc networks (MANETs) have two new characteristics, wireless communication and dynamic physical topologies. Nodes have to suffer interference from other nodes because of the broadcast nature of wireless communications. As a result, network performance degrades if neighboring nodes transmit packets over the same frequency. This problem becomes serious for multipath routing protocols when nodes simultaneously use multiple paths to transmit packets. Also, dynamic physical topology usually incurs considerable overhead to repair broken paths especially for multipath routing protocols. Nodes should select stable links to save routing overhead and maintain network reliability. On the other hand, it is hard to perform optimal load balancing in dynamic physical topologies because significant overhead has to be generated to inform nodes of updated conditions for each path. Considering the characteristics of MANETs, we argue that multipath routing protocols on MANETs should be focused on how nodes select multiple paths to improve network reliability and performance instead of how nodes optimally utilize multiple paths.

The problem we address in this section is how nodes select multiple node-disjoint, edge-disjoint, and frequency-disjoint paths. Many multipath routing protocols have been proposed to allow nodes to select multiple node-disjoint and edge-disjoint paths. However, few of them provide solutions for how nodes can select multiple frequency-disjoint paths.

Our main contribution is the cognitive multipath multi-channel routing protocol (CMMRP). It falls into a novel category of routing protocols, cognitive routing protocols. Each node predicts future conditions of links and frequencies based on past experience. Multiple disjoint paths are discovered one by one by triggering RREQ packets multiple times from the source node, trading routing overhead for network reliability and performance.

5.2 Related Work

In this section, we discuss work related to CMMRP.

5.2.1 Disjoint Multipath Routing Protocols

A multitude of disjoint multipath routing protocols have been proposed (e.g. [67-72]). The main idea is that a source node selects multiple non-overlapped paths to transmit packets to a destination node. In [67], the authors proposed a maximally disjoint multipath routing protocol, an on-demand routing protocol. Unlike traditional protocols where intermediate nodes discard duplicated RREQ packets, the proposed routing protocol lets intermediate nodes relay duplicated RREQ packets to make destination node know all possible paths. A destination node selects multiple disjoint paths according to recorded traversed paths of RREQ packets and sends RREP packets back to source node through the corresponding paths.

Disjoint multipath routing protocols improve network reliability because the possibility of multiple disjoint paths breaking simultaneously is much lower than possibility that one path breaks. However, the main drawback is the significant routing overhead incurred when intermediate nodes relay all duplicated RREQ packets.

5.2.2 Meshed Multipath Routing Protocols

Many meshed multipath routing protocols have been proposed (e.g. [73-77]). The main idea is that a source node selects multiple overlapped paths to transmit packets to a destination node. Unlike disjoint multipath routing protocols, they let intermediate nodes have multiple paths between source node and destination node. As a result, there will be a large number of overlapping paths between source node and destination node constructed via links selected by intermediate nodes. In [73], the authors argue that meshed multipath routing protocols are more reliable than disjoint multipath routing protocols. However, the difference is small because the selected paths are overlapping but not independent.

Meshed multipath routing protocols improve network reliability considerably because of the large number of overlapping paths between source node and destination node. However, the main drawback is the excessive overhead incurred when intermediate nodes transmit replicated data packets over multiple links to the destination node.

5.2.3 Cognitive Routing Protocols

In recent years, several cognitive routing protocols have been proposed (e.g. [88-91]). The principle idea is that nodes are able to make wise decisions by predicting future network conditions based on past experience. Machine learning techniques are adopted to make nodes aware of history. Lower layer knowledge of wireless medium is shared with network layer. Currently, path selection and spectrum selection are common topics in cognitive routing protocols. Little prior work has focused on multipath routing. We argue that cognitive routing is a promising approach to make nodes intelligently perform multipath routing.

5.3 Approach

In this section, we discuss the details of our proposed routing protocol.

5.3.1 Overview of CMMRP

The cognitive multipath multi-channel routing protocol (CMMRP) is an on-demand disjoint multipath routing protocol.

CMMRP lets nodes trigger routing updates reactively when links break. Unlike the other on-demand multipath routing protocols where multiple paths are discovered at once by triggering one RREQ packet from source node, CMMRP allows multiple paths to be discovered one by one by triggering RREQ packets multiple times from the source node. The benefits of modification are:

- Nodes efficiently control the number of paths.
- CMMRP can downgrade to a single-path routing protocol.
- Nodes efficiently perform paths maintenance in dynamic networks.
- Nodes confidently discover frequency-disjoint paths.

CMMRP is designed for multi-channel environments where nodes can simultaneously use multiple interfaces to transmit packets over different frequencies. Unlike the other disjoint multipath routing protocols where nodes select multiple node-disjoint and edge-disjoint paths over one frequency, CMMRP employs cognitive functions to allow nodes to intelligently select multiple node-disjoint, edge-disjoint, and frequency-disjoint paths over multiple frequencies. Frequency-disjoint paths indicate that neighboring nodes on different paths transmit packets over different frequencies, which avoids interference from the other nodes. Neural network machine learning is adopted to make nodes aware of history, which makes CMMRP a cognitive routing protocol. Lower layer knowledge is shared with the network layer to help CMMRP work properly and efficiently.

The path discovery protocol of CMMRP can be divided into two parts, space discovery and spectrum discovery protocol, working in the two dimensions, space and spectrum.

5.3.2 Space Discovery Protocol

As the name implies, space discovery protocol is used to discover multiple node-disjoint and edge-disjoint paths in space dimension. It is further developed based on our own work. In previous work, we developed a new metric, throughput increment, which is defined as predicted throughput after a new application joins minus current throughput. It is the throughput increment that determines future overall throughput. Predicted throughput increment is estimated based on predicted channel type and predicted channel capacity using neural network machine learning method. Simulation results show that throughput increment is an excellent metric for path selection.

In this work, space discovery protocol utilizes the developed metric to select paths after multiple disjoint paths are discovered. For ease of implementation, source node desires two disjoint paths to destination node.

Space discovery protocol works as follows. Source node floods RREQ packets reactively when number of paths to destination node is less than desired number. Unlike the other disjoint multipath routing protocols, intermediate nodes do not relay duplicated RREQ packets. Consequently, destination node knows a subset of possible paths to source node and selects the best path from them based on the developed metric. To discover two disjoint paths, source node has to flood RREQ packets again. Each intermediate node checks whether it already had a valid routing table entry to destination node. If so, it does not relay RREQ packets because it is already selected by destination node on one of the paths. Otherwise, it relays RREQ packets if they are not duplicated. Destination node selects the best path from updated known paths based on the developed metric. This approach guarantees source node to discover multiple node-disjoint and edge-disjoint paths.

Space discovery protocol makes multiple disjoint paths discovered one by one, which makes it possible for source nodes to independently determine the number of disjoint paths to destination nodes. This approach also enables CMMRP to downgrade to a single-path routing protocol, which makes it compatible with some other routing protocols. Compared to the other disjoint multipath routing protocols, CMMRP reduces routing overhead significantly by employing space discovery protocol. Traditional disjoint multipath routing protocols let intermediate nodes relay duplicated RREQ packets to make destination node know all possible paths. Routing overhead is mainly determined by the number of neighboring nodes because it determines the number of duplicated RREQ packets received by a node, which usually means a large amount of routing overhead especially for dense networks. Unlike them, CMMRP makes multiple paths discovered one by one by triggering RREQ packets multiple times from source node. Intermediate nodes discard duplicated RREQ packets. By applying space discovery protocol, routing overhead is mainly determined by the number of times that source node triggers RREQ packets. As a result, CMMRP does not generate significant overhead for dense networks.

5.3.3 Spectrum Discovery Protocol

As the name suggests, spectrum discovery protocol is used to discover multiple frequency-disjoint paths in spectrum dimension. It is further developed based on our own work. In previous work, we developed a metric, delay of RREQ packets, for frequency selection. In wireless communication, packet delay is mainly determined by queuing delay which is affected by traffic load and channel capacity. Delay of RREQ packets is used to estimate queuing delay. The frequency over which delay of RREQ packets is small is predicted to have good channel conditions. Simulation results show that it is an excellent metric for frequency selection.

In this work, spectrum discovery protocol utilizes the developed metric to discover multiple frequency-disjoint paths. For ease of implementation, each node equips two interfaces which can be used simultaneously for packet transmission over different frequencies.

Spectrum discovery protocol works as follows. Nodes monitor as many frequencies as possible. Each node floods RREQ packets over all available frequencies to make neighboring nodes know conditions of each frequency. By applying the developed metric, when receiving a RREQ packet with a new sequence number or flooding ID, node records receiving frequency in the corresponding routing table entry and discards the rest duplicated RREQ packets. The recorded frequency is the most possible one which is different with the frequencies selected by neighboring nodes because allocated frequencies with interference from transmitting nodes tend to have large delay of RREQ packets. Each node on path selected by space discovery protocol transmits RREP packet back to source node over the recorded frequency to notify previous-hop node of the frequency selected by spectrum discovery protocol.

CMMRP is designed for multi-channel environment. It uses a novel approach, cognition, to discover multiple frequency-disjoint paths. Nodes estimate and predict future conditions of frequencies based on past experience. Spectrum discovery protocol utilizes the developed metric. Traditional multipath routing protocols only improve network reliability. However, by employing spectrum discovery protocol, CMMRP improves both network reliability and performance because spectrum diversity is maximized with frequency-disjoint paths, which avoids interference from neighboring nodes. Also, nodes are able to confidently discover frequency-disjoint paths because multiple paths are discovered one-by-one. Spectrum can be gradually allocated to each path according to updated environment. Otherwise, it is difficult to guarantee that neighboring nodes on different paths select different frequencies simultaneously in distributed manner.

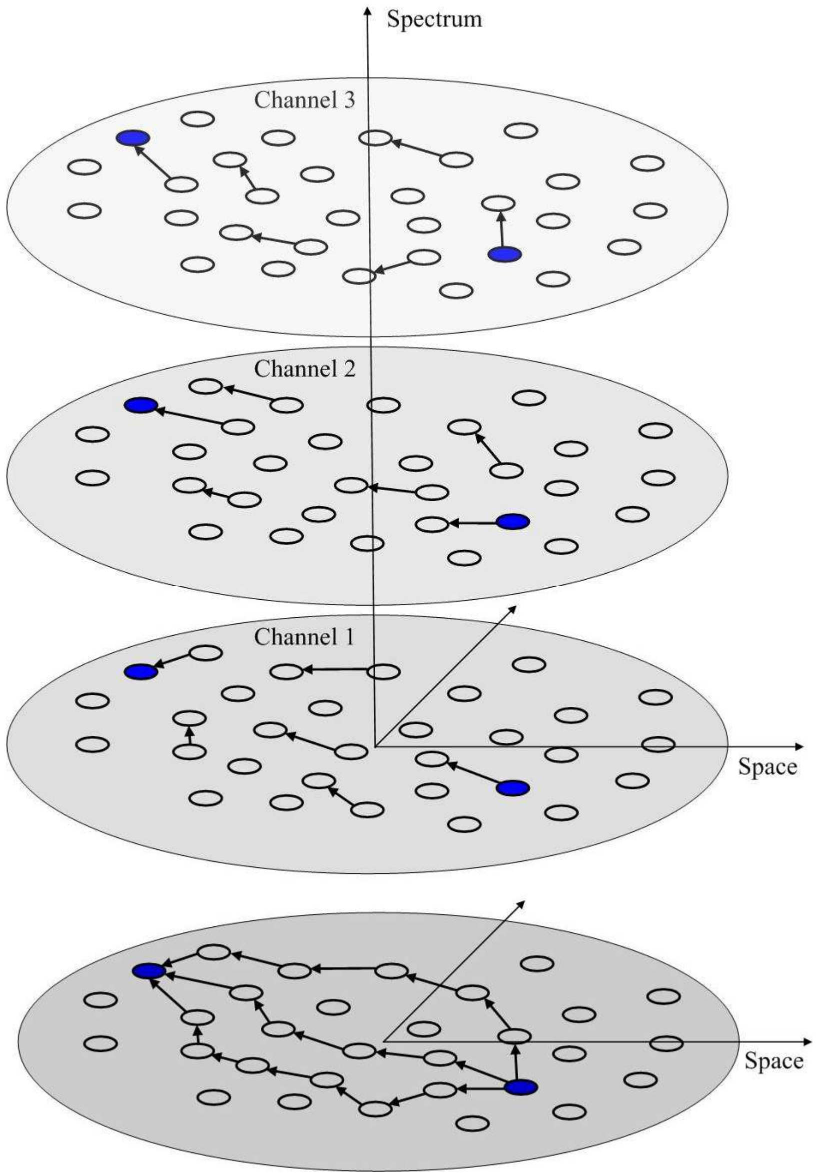


Figure 5.1 An example of constructed network topology by space discovery protocol and spectrum discovery protocol

Figure 5.1 shows an example of constructed network topology by space discovery protocol and spectrum discovery protocol. It clearly shows the benefits of them. Blue nodes are indicated as source node and destination node. In spectrum view, nodes select frequency-disjoint paths for packet transmission. In space view, source node selects node-disjoint and edge-disjoint paths for packet transmission.

5.3.4 Paths Maintenance

In MANET, links tend to break because of dynamic physical topology. As a result, an efficient paths maintenance protocol is demanded to maintain network reliability.

Most traditional multipath routing protocols have to replace all valid paths between source node and destination node as long as one path breaks because source node initiates one RREQ packet to discover multiple paths at once. Otherwise, source node has to wait until number of valid paths is below a threshold to trigger RREQ packets. By applying traditional approaches, multipath routing protocols have to either generate significant routing overhead or degrade network performance. The problem becomes serious for MANET where network topology changes quickly. Unlike them, CMMRP provides an efficient paths maintenance protocol which is suitable for MANET. It makes multiple paths discovered one by one by triggering RREQ packets multiple times. Source node is able to efficiently control the number of valid paths by repairing one broken path each time, which usually incurs a small amount of routing overhead because of space discovery protocol.

5.3.5 Paths Usage

In wired networks, major consideration on multipath routing protocols is focused on paths usage because of the static network topology. However, in MANET, we argue that major consideration should be focused on path discovery.

CMMRP does not let nodes back up any redundant path because it shows little benefit in MANET. To balance load among multiple disjoint paths, source node should clearly know the conditions of each path. In [81], the authors proposed a routing protocol using WCETT as metric for path selection. We proposed a new metric, throughput increment, for path selection. CMMRP is able to balance load among multiple paths according to our developed metric. However, it shows little benefit in MANET because source node cannot continuously update conditions of each path. Otherwise, significant overhead has to be generated. Therefore, for ease of implementation, CMMRP equally balance load among multiple paths.

5.4 Simulations

In this section, we provide simulation results of CMMRP using Qualnet 4.0. To show the benefits of CMMRP, it is compared with spectrum-aware routing protocol (SARP) which is our own previous work and a simple multipath routing protocol (MRP).

MRP is an on-demand multipath single-channel routing protocol. SARP is described as follows. It is an on-demand cognitive single-path multi-channel routing protocol. Neural network machine learning method is used to make nodes aware of history. It employs two cognitive functions, intelligent multi-interface selection function (MISF) and intelligent multipath selection function (MPSF). The metric, throughput increment, is adopted by MPSF for path selection and the metric, delay of RREQ packets, is adopted by MISF for frequency selection. Simulation results show that SARP improves network performance.

We present following metrics with 95% confidence intervals of measured values to compare network reliability and performance of CMMRP with SARP and MRP.

- Overhead: Average number of RREQ packets received per frequency which dominates the number of route control packets.
- Throughput: Average rate of successful message delivery measured in Kbits per second which reflects network reliability and performance.

5.4.1 Impact of Network Size

In the first experiment, we show how network reliability and performance are affected when the number of nodes increases. We created a scenario which has 6 applications distributed in a 600m by 1500m region. We vary the number of nodes. There are ten available frequencies. Frequencies have different shadowing means such as 4, 6, 8, 10 or 12 and same Ricean K factor as 16. Automatic rate fallback is enabled. UDP is employed as transport layer protocol. IEEE 802.11 is employed as MAC protocol. We compare CMMRP with SARP and MRP.

Figure 5.2 and figure 5.3 show the comparison of overhead and throughput respectively as a function of number of nodes. As expected, overhead is increased as number of nodes increases because node density increases.

Compared to SARP and MRP, CMMRP increases overhead considerably. The reason is explained as follows. CMMRP is a disjoint multipath routing protocol and SARP is a single-path routing protocol. By employing CMMRP, multiple paths are discovered one by one by trigger RREQ packets multiple times, which incurs more overhead than SARP. On the other hand, CMMRP is a multi-channel routing protocol and MRP is a single-channel routing protocol. By employing CMMRP, each node equips two interfaces which monitor ten available frequencies and floods RREQ packets over all frequencies. CMMRP generates more average overhead per frequency than MRP because possibility that route control packets are transmitted over a long distance increases when number of frequencies increases.

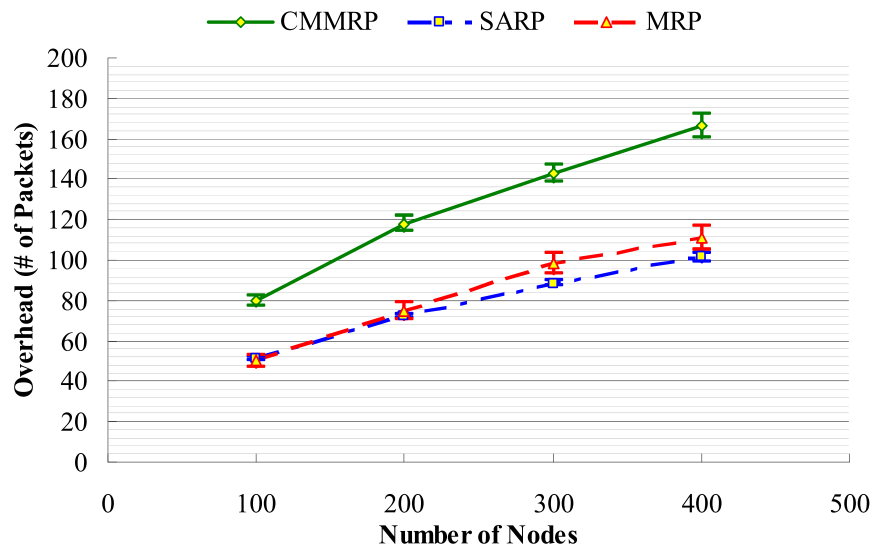


Figure 5.2 Comparison of overhead as a function of number of nodes

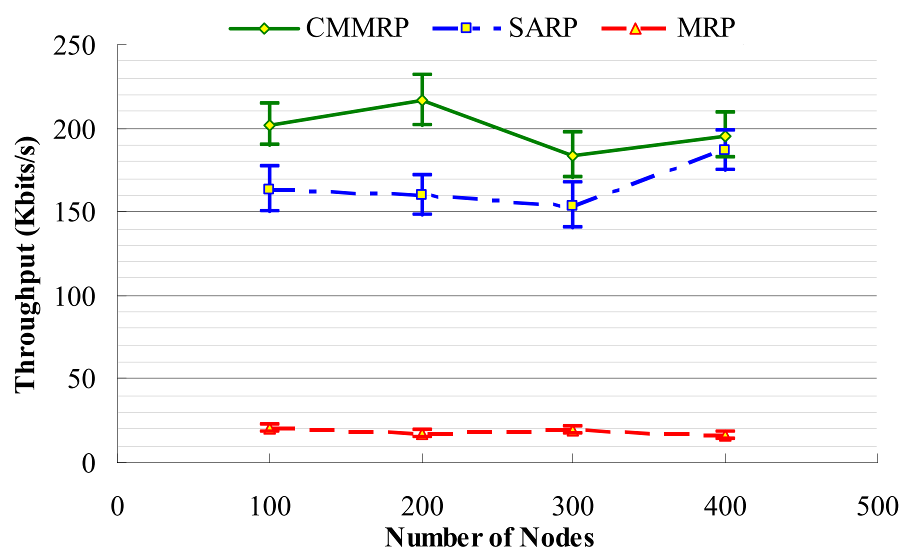


Figure 5.3 Comparison of throughput as a function of number of nodes

Compared to SARP, CMMRP increases throughput considerably. The reason is explained as follows. CMMRP is a disjoint multipath routing protocol and SARP is a single-path routing protocol. By employing CMMRP, source node discovers multiple node-disjoint, edge-disjoint and frequency disjoint paths, which improves network reliability and performance. Compared to MRP, CMMRP increases throughput significantly. The reason is explained as follows. Like most multipath routing protocols, MRP can utilize only one frequency. As a result, interference between nodes is serious. On the other hand, CMMRP is a multi-channel routing protocol. Nodes discover multiple frequency-disjoint paths over ten frequencies. Consequently, interference between nodes decreases dramatically because frequency diversity is maximized by spectrum discovery protocol. These results show that CMMRP has better network reliability and performance than SARP and MRP for scenarios with a large number of nodes.

5.4.2 Impact of Network Dynamics

In the second experiment, we show how network reliability and performance are affected when node velocity increases. We created a scenario similar as the first experiment. 200 nodes are distributed in a 2500m by 2500m region. We vary node velocity. We compare CMMRP with SARP and MRP.

Figure 5.4 and figure 5.5 show the comparison of overhead and throughput respectively as a function of average velocity. As expected, overhead is increased and throughput is decreased as average velocity increases because links breakage happens frequently.

Compared to SARP, CMMRP increases overhead considerably. The reason is explained as follows. In MANET, nodes trigger routing updates frequently to repair broken paths. CMMRP generates considerable overhead especially for high node velocity to perform paths maintenance because nodes have to maintain multiple paths. Routing overhead of SARP increases slower than CMMRP as average velocity increases because it maintains fewer path than CMMRP. Routing overhead of CMMRP increases slower than MRP as average velocity increases because it makes multiple disjoint paths discovered one by one by triggering RREQ packets multiple times from source node, which is an efficient paths maintenance protocol.

Throughput is decreased as average velocity increases because of the dynamic network topology. Compared to SARP, CMMRP increases throughput considerably. Compared to MRP, CMMRP increases throughput significantly. The reasons are similar as the first experiment. These results show that CMMRP has better network reliability and performance than SARP and MRP for scenarios with high node velocity.

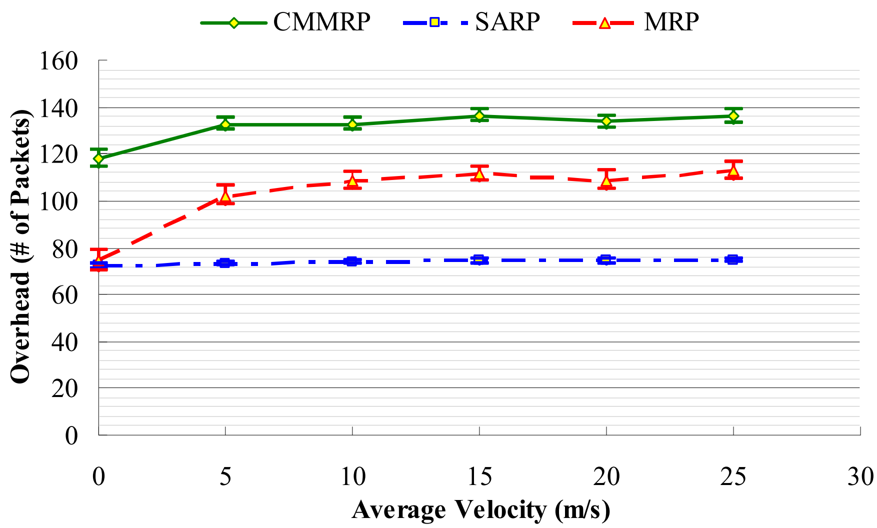


Figure 5.4 Comparison of overhead as a function of average velocity

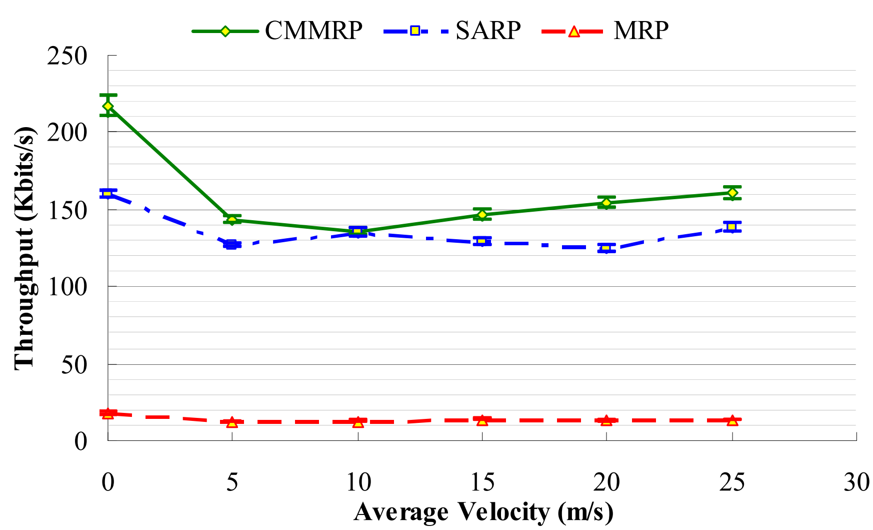


Figure 5.5 Comparison of throughput as a function of average velocity

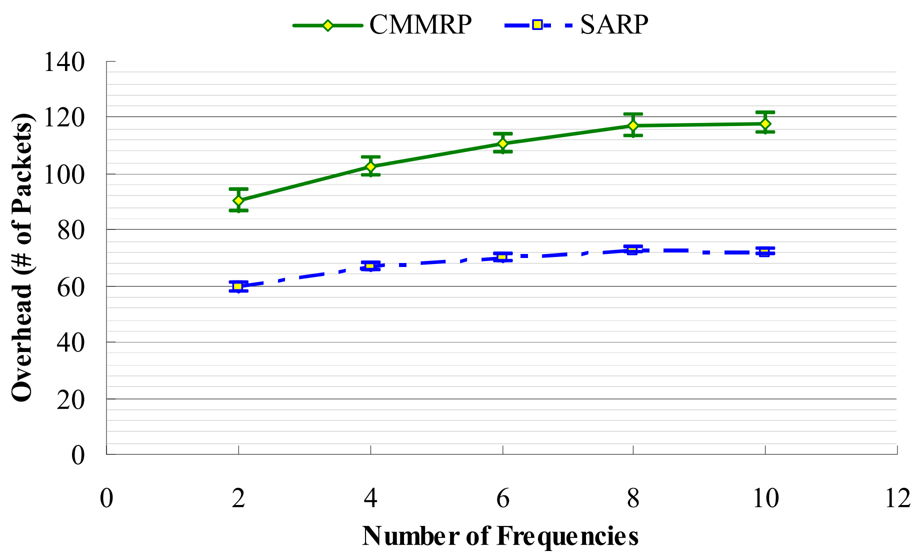


Figure 5.6 Comparison of overhead as a function of number of frequencies

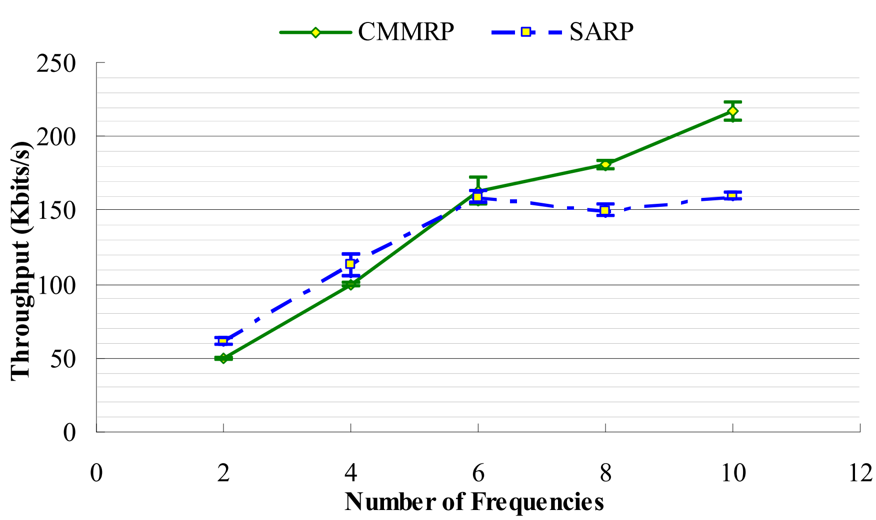


Figure 5.7 Comparison of throughput as a function of number of frequencies

5.4.3 Impact of Network Spectrums

In the third experiment, we show how network reliability and performance are affected when the number of frequencies increases. We created a scenario similar as previous experiments. 200 nodes are distributed in a 600m by 1500m region. We vary the number of frequencies. We compare CMMRP with SARP.

Figure 5.6 and figure 5.7 show the comparison of overhead and throughput respectively as a function of number of frequencies. As expected, overhead and throughput are increased as number of frequencies increases.

Overhead is increased as number of frequencies increases because possibility that route control packets are transmitted over a long distance increases when number of frequencies increases. Compared to SARP, CMMRP increases overhead considerably. The reason is similar as previous experiments.

Throughput increases as number of frequencies increase because nodes are able to utilize more frequencies to increase frequency diversity. Compare to SARP, CMMRP has almost same throughput when the number of frequencies is smaller than 6 and it has more throughput when number of frequencies is bigger than 6. The reason is explained as follows. When the number of frequencies is small, nodes cannot effectively perform spectrum discovery protocol to discover multiple frequency-disjoint paths. On the other hand, when the number of frequencies is big, CMMRP makes nodes efficiently perform spectrum discovery protocol to maximize frequency diversity, which improves network performance. However, 6 frequencies are enough for SARP to allocate spectrum because it is a single-path routing protocol.

These results show that CMMRP can improve network performance when number of frequencies is big enough.

5.5 Conclusions

We have investigated the problem of multipath routing in multi-channel mobile ad-hoc networks. We have proposed a novel cognitive multipath multi-channel routing protocol.

CMMRP lets nodes trigger routing updates reactively when links break. It is designed for multi-channel environment where nodes can simultaneously use multiple interfaces to transmit packets over different frequencies. It employs cognitive functions to make nodes intelligently select multiple node-disjoint, edge-disjoint and frequency-disjoint paths. Neural network machine learning method is adopted to make nodes aware of history. Lower layer knowledge is shared with network layer to help CMMRP work properly and efficiently. Path discovery protocol of CMMRP can be divided into two parts, space discovery protocol and spectrum discovery protocol. They work in two dimensions, space and spectrum. Simulation results show that CMMRP significantly improves network reliability and performance.

Chapter 6

Mobility-Aware Routing Protocol for Mobile Cognitive Networks

6.1 Introduction

In mobile ad-hoc networks, the physical topology of the nodes tends to change because of the nodes mobility. Consequently, the link conditions between the nodes also tend to change. In order to gain better performance, the logical topology, that is, the routing table used by the routing protocols should adapt fast to the physical topology changes.

The traditional routing protocols [86-94] will trigger the routing updates only after the nodes detect the route failure. If the distance between the nodes is small, the available bandwidth might be large because of the strong received signal strength. On the other hand, if the link between the nodes is about to break, the available bandwidth might be small because of the weak received signal strength. Therefore, in traditional routing protocols, the condition of a link has to be rather bad before the routing table is updated which decreases the overall network performance. If the routing protocol is aware of the mobility, the nodes will be able to trigger the routing updates intelligently before the link breaks. The nodes might find a better next hop with larger available bandwidth. The network performance might be improved.

6.2 Related Work

Most of the work [96-99] on mobile wireless ad hoc networks focuses on the routing topologies and scalability issues. Few efforts have considered the mobility-aware routing protocol (MARP).

In [82], the authors propose a routing algorithm called adaptive distance vector (ADV) for mobile ad-hoc networks. In this algorithm, the frequency and the size of the routing updates can be changed dynamically based on the network load and mobility conditions. In ADV, the routing updates frequency increases as the mobility velocity increases. From the results provided in the paper, ADV is able to improve the throughput and decreases the end-to-end delay. However, we argue that the mobility velocity is not a clear sign to change the routing updates frequency because a cluster of nodes might move with the same mobility velocity and in the same direction. In this case, the nodes do not need to trigger the routing updates because the relative physical topology is not changed.

In [83], the authors propose a preemptive routing maintenance algorithm. The authors argue that the algorithm combines the best of the table-driven and on-demand routing protocol, because the hand-off is initiated early to minimize the delay and the routing update is triggered as needed to minimize the overhead. In [84], the authors propose a router handoff algorithm for mobile ad-hoc networks. The node tries to locally find an alternate next hop before it floods the route request. In this way, the overhead incurred can be minimized. In [85], the authors propose an algorithm called preemptive AODV (PrAODV). The algorithm combines two pre-emptive mechanisms, schedule a rediscovery in advance and warn the source before the path breaks. The algorithms proposed in [83-85] use the received signal strength as a sign to predict when the link is likely to break. However, we argue that the received signal strength is also not a clear sign to determine when to trigger the routing updates because of its large variance. The received signal strength reflects both the small scale fading and large scale fading. The routing protocols will have too many overheads if the routing table changes fast, which might decrease the network performance.

In mobile ad-hoc networks, if a link along the path breaks, the source node has to stop transmitting the packets, which decrease the throughput and increase the end-to-end delay because of the longer queue length. From the related works, we found that if the routing protocol is aware of the mobility and triggers the routing updates before the link breaks, the network can not only avoid the potential path failure but also might find a better next hop with larger bandwidth. In this way, the overall performance of the network can be improved.

6.3 Approach

Compared to the previous works, our MARP uses slope of the throughput rather than the received signal strength to predict when the link is likely to break. In mobile ad-hoc networks, the variance of the received signal strength tends to be large. However, the variance of the slope of the throughput is relative small compared to the received signal strength. It is relatively easy to keep track of the slope of the throughput. Therefore, our MARP should gain better performance.

Figure 6.1 shows the relationship between the throughput and the load for different type of links. It is obvious that the link with larger bandwidth will have larger slope of the throughput. However, after the link is saturated, the slope of the throughput becomes zero. Therefore, the slope of the throughput can reflect both the link conditions and the load of a link.

In our MARP, the nodes need to predict the slope of the throughput based on the history. The algorithm used to predict the slope of the throughput will be illustrated later. By keeping track of the slope of the throughput, the nodes are able to predict that the link is likely to break as long as the slope of the throughput decreases. The reason is that when the distance between the nodes becomes larger or the interference from the neighboring nodes becomes larger, the slope of the throughput decreases. In most cases, the link condition will be even worse in the future, which causes the link breaks finally. Basically, our MARP uses the change of the slope of the throughput as the sign of the physical topology changes. Therefore, as long as the slope of the throughput decreases, the node should trigger the routing updates to find a better next hop instead of waiting until the link breaks.

In our MARP, the nodes will perform the local optimization if they have determined to trigger the routing updates. It means that the intermediate node only needs to inform the previous hop rather than the source node. It is the previous hop that needs to flood the RREQ packets. Therefore, the local optimization might be transparent to the source node. In other words, the source node can keep transmitting the packets to the destination node without worrying about the link failure, because the routing updates are performed before the link fails. Therefore, our MARP performs seamless handoff.

Routing table should not only be as fresh as possible but also be as stable as possible. Otherwise, the nodes have to switch the next hop too frequently, which might incur too much overhead. Most recent MAC protocols [103-105] tend to use the received signal strength to determine which node should access the medium. The received signal strength can reflect both the small scale fading and the large scale fading. It meets the demands of MAC protocol, so that the nodes are able to perform opportunistic scheduling. However, routing protocol does not need to consider the small scale fading. It only needs to consider the large scale fading to make routing table relatively stable. Therefore, using the slope of the throughput to determine when the nodes should trigger the routing updates might be suitable for routing protocols, because it can only reflect the large scale fading effects and tends to have smaller variance compared to the received signal strength.

Our MARP also combines the best of the on-demand and table-driven routing protocol. It is able to gradually increase number of the routing updates as the physical topology of the network changes faster. When the physical topology changes slowly, the slope of the throughput also changes slowly. Few routing updates might be triggered. Therefore, the unnecessary overhead involved is minimized. On the other hand, when the physical topology changes fast, the slope of the throughput also changes fast. Many routing updates might be trigger. In this way, the routing table is able to adapt fast to the physical topology changes. Therefore, our MARP tries to minimize the unnecessary overhead and also adapt fast to the physical topology changes.

At this point, we know the advantages of choosing the slope of the throughput as the sign to predict when the link is likely to break. In order to predict the slope of the throughput, we use the neural network machine learning method [100]. Based on our previous work [101], we are able to predict the future throughput given the potential source load based on the current packet loss rate and the current end-to-end delay. Equation (6.1) illustrates how to calculate the slope of the throughput. We denote S as the slope of the throughput, FT as the predicted future throughput, CT as the current throughput, FL as the future load and CL as the current load. In this way, the nodes are able to predict the slope of the throughput.

$$S = (FT - CT)/(FL - CL) \tag{6.1}$$

After triggering the routing updates, the nodes might receive several RREP packets, even from the upstream nodes, because the upstream nodes are transparent to the local optimization. However, we argue that there will be not routing loop problem involved. The reason is that our MARP uses the intelligent multi-path selection algorithm [101] to determine whether the existing route should be preempted by the new route. The path from the upstream nodes to the destination nodes cannot be better than the path from the current node to the destination node. Therefore, the existing route will not be preempted by the upstream nodes.

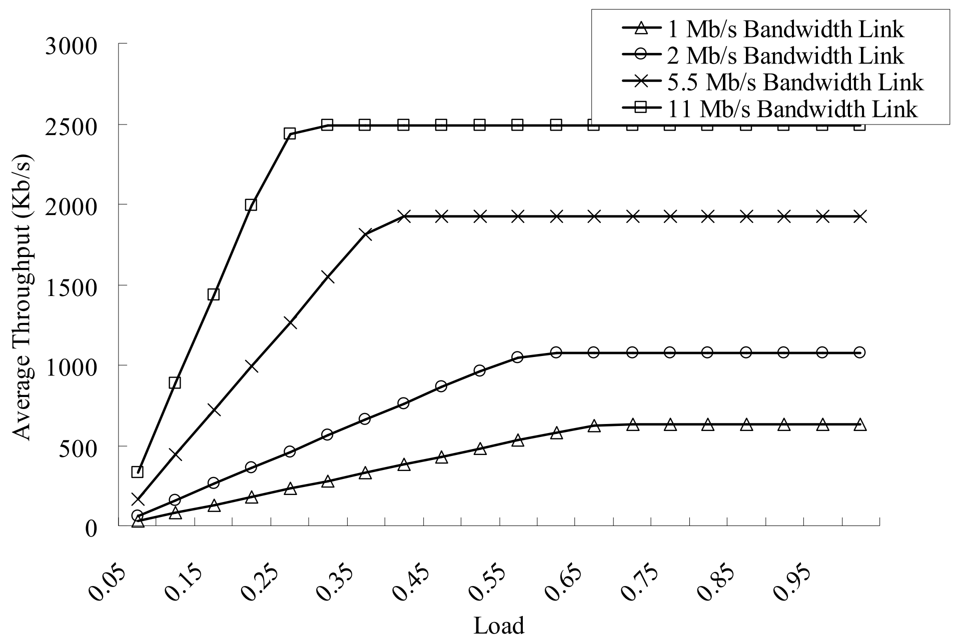


Figure 6.1: Average network-wide throughput as a function of load

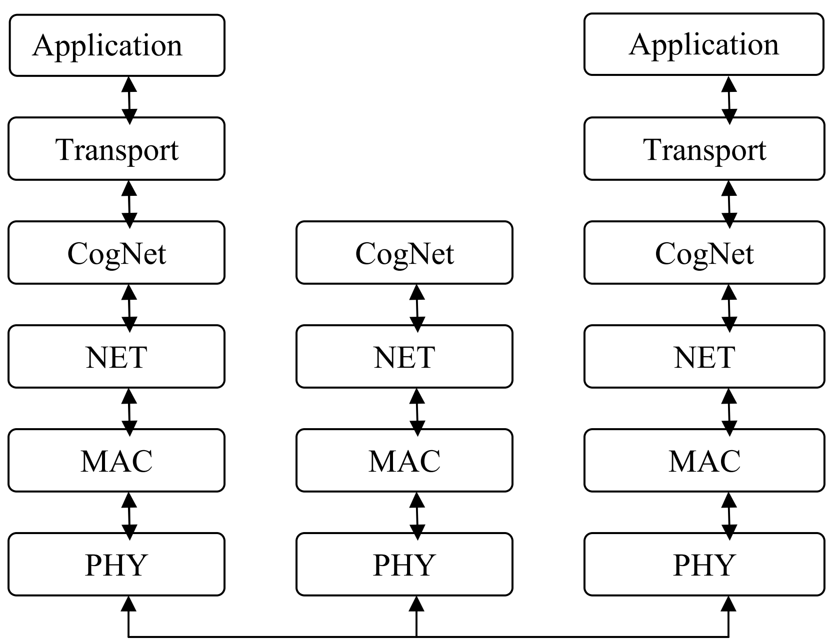


Figure 6.2: The CogNet layer architecture

6.4 Implementations

In last section, we mentioned that we use our previous work [101] to predict the slope of the throughput. In that work, we added a new layer between the network layer and the transport layer. This new layer is called the CogNet layer. In this approach, the new layer is responsible for maintaining the end-to-end delay, the packet loss rate and the number of packets sent of the links for the destination nodes, predicting the type of links for the neighbors, predicting the load level of links for the neighbors and predicting the slope of the throughput. Figure 6.2 shows the protocol architecture we used in our previous work. We implement the routing framework based on AODV.

The receiver keeps track of the slope of the throughput for the link between the receiver and its previous hop. The receiver needs to learn the history for two seconds before it determines the changes of the slope of the throughput. In this way, the variance of the slope of the throughput can be reduced. Whenever the slope of the throughput decreases, the receiver sends a warning indication to its previous hop. This warning signal is very similar to the RREQ packet. The only difference is that this warning signal contains the destination node address along the path rather than the sender of the RREQ. The radius of the flooding region of the warning signal is only 1-hop. Therefore, it has only little impact on the aggregate network performance.

When the previous hop receives the waning signal, it floods the RREQ packets to the destination node. The radius of the flooding region is same as the hop count of the existing path to the destination node. In our MARP, the nodes will not retry this RREQ packet which performs the local optimization, because there will be much more overhead incurred. However, if the node receives another warning signal from the next hop, the node needs to flood the RREQ packets again.

6.5 Simulations

In this section, we provide simulation results with different mobility velocities, obtained using Qualnet 4.0 [102].

Our scenario has a 1000m by 1000m region. There are 56 nodes in this region which is randomly distributed. There are six application sources. We assume the applications generating traffic for the simulation have an exponential distribution inter-arrival rate with a mean of 1.5 ms. User datagram protocol (UDP) is used as the transport layer protocol. Because the nodes experience different wireless conditions, the bandwidth for each node can be different. The bandwidth between nodes might be 1 Mbps, 2 Mbps, 5.5 Mbps or 11 Mbps. The node queue size is 20,000 bytes. Auto fallback rate is enabled. Mobility velocities of 5 m/s, 10 m/s, 15 m/s, 20 m/s, 25 m/s and 30 m/s were evaluated. Simulations were run at least 45 times for each velocity. AODV was used as the basis for comparison of the performance of our MARP.

Figure 6.3 shows the distribution of the average throughput for each time as a function of mobility velocity. Figure 6.4 shows the average throughput for each speed as a function of the mobility velocity. From these two figures, it is obvious that our approach is able to increase the throughput significantly.

Figure 6.5 shows the percentage increased of the throughput as the function of the mobility velocity. The throughput is increased by about 30% when the mobility velocity is high. However, when the mobility velocity is low, the performance is roughly similar to AODV. The reason is that when the mobility velocity is low, the physical topology changes slowly. The local optimization is seldom performed. In this situation, the MARP has the similar behavior to AODV. On the other hand, when the mobility velocity is high, the local optimization is triggered very frequently. Compared to AODV, the routing table in our MARP can adapt faster to the physical topology changes. Therefore, it is able to attain much higher throughput. Figure 6.6 shows the standard deviation of the average throughput as the function of the mobility velocity. Compared to the mean of the average throughput, the standard deviation of the throughput is small.

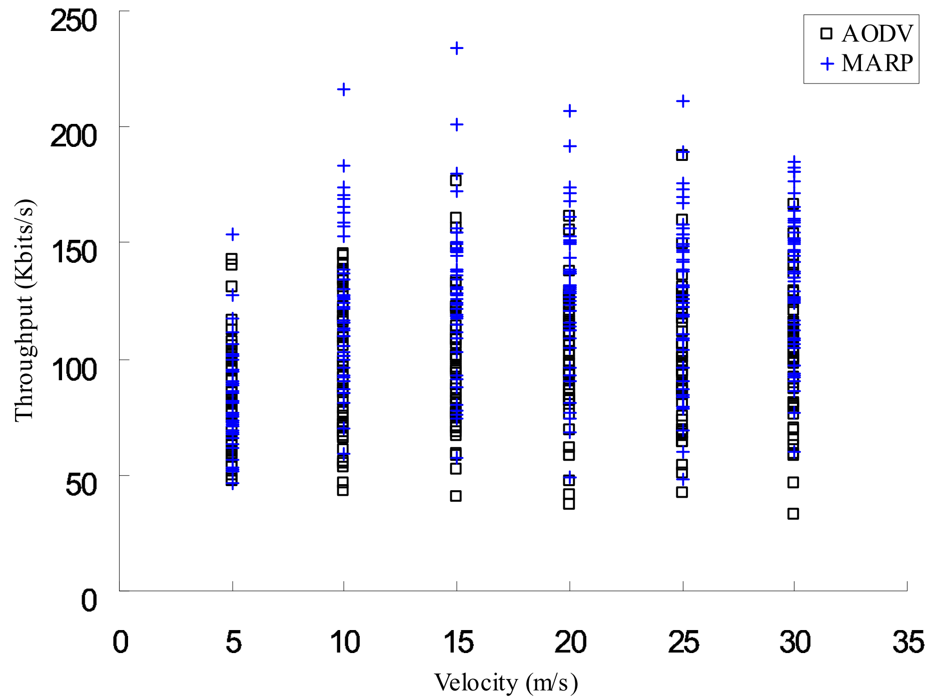


Figure 6.3: The distribution of the average throughput as a function of mobility velocity.

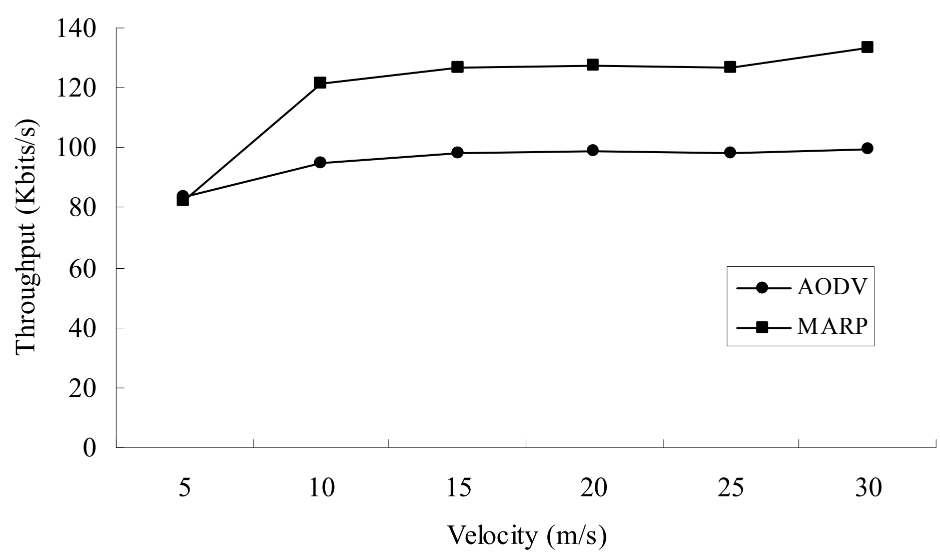


Figure 6.4: The average throughput as a function of mobility velocity.

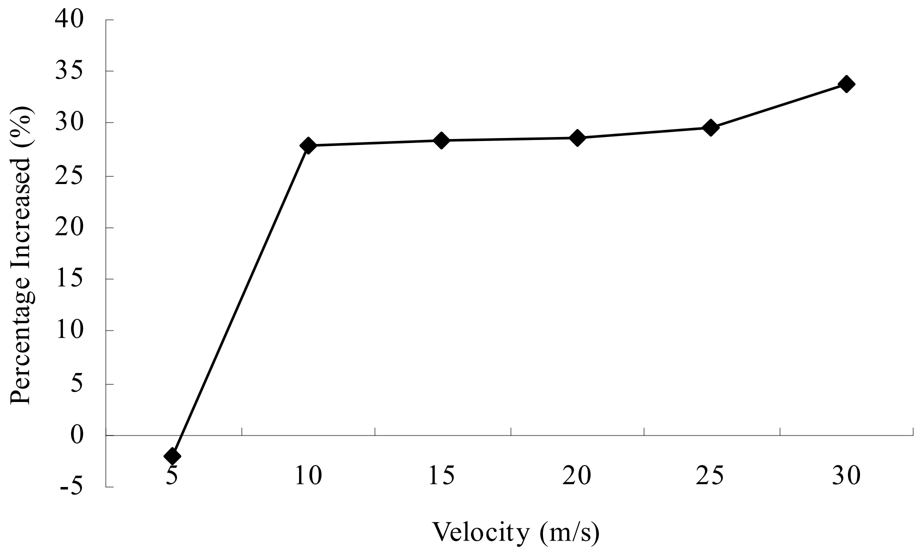


Figure 6.5: The percentage increased of the throughput as a function of mobility velocity

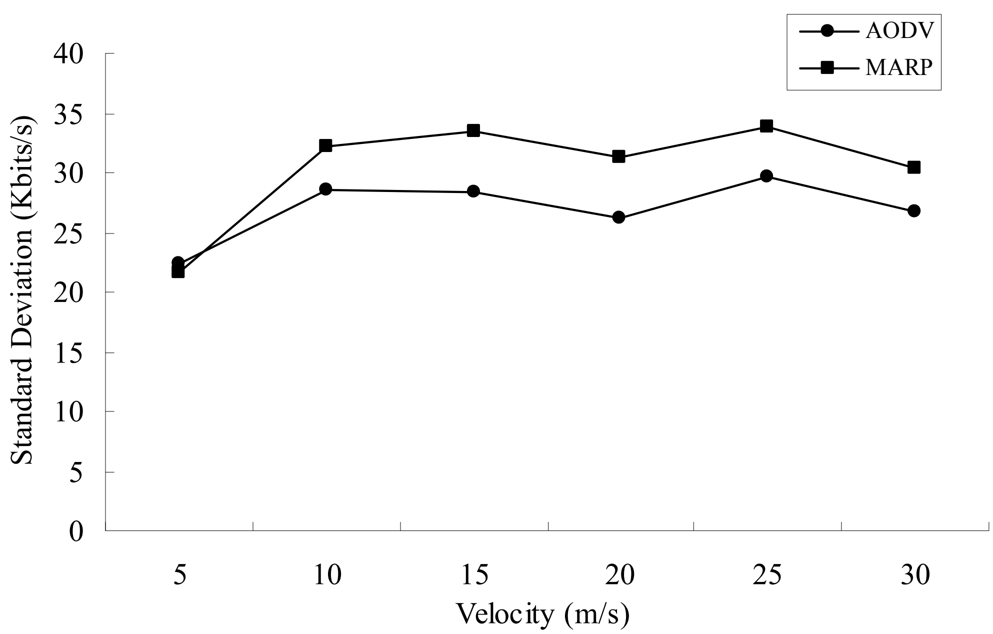


Figure 6.6: The standard deviation of the throughput as a function of mobility velocity

Figure 6.7 shows the distribution of the average end-to-end delay for each time as a function of mobility velocity. Figure 6.8 shows the average end-to-end delay for each speed as a function of the mobility velocity. These results suggest that our approach is able to decrease the end-to-end delay significantly.

Figure 6.9 shows the percentage decrease of the end-to-end delay as a function of the mobility velocity. The end-to-end delay can be decreased by about 33%. When the mobility velocity is high, the local optimization is triggered very frequently. Compared to AODV, our MARP will lead to fewer link failures, so the average queue length will be smaller, which results in the decreased end-to-end delay. On the other hand, when the mobility velocity is low, the MARP is also able to decrease the end-to-end delay significantly. The reason is that the nodes use the slope of the throughput as the measure to determine when they should trigger routing updates. The slope of the throughput combines both the link conditions and the loading on the link. Therefore, when the mobility velocity is low, the MARP is able to perform load balancing so that the end-to-end delay is decreased. Figure 6.10 shows the standard deviation of the average end-to-end delay as the function of the mobility velocity. Compared to the mean of the average end-to-end delay, the standard deviation of the end-to-end delay is small.

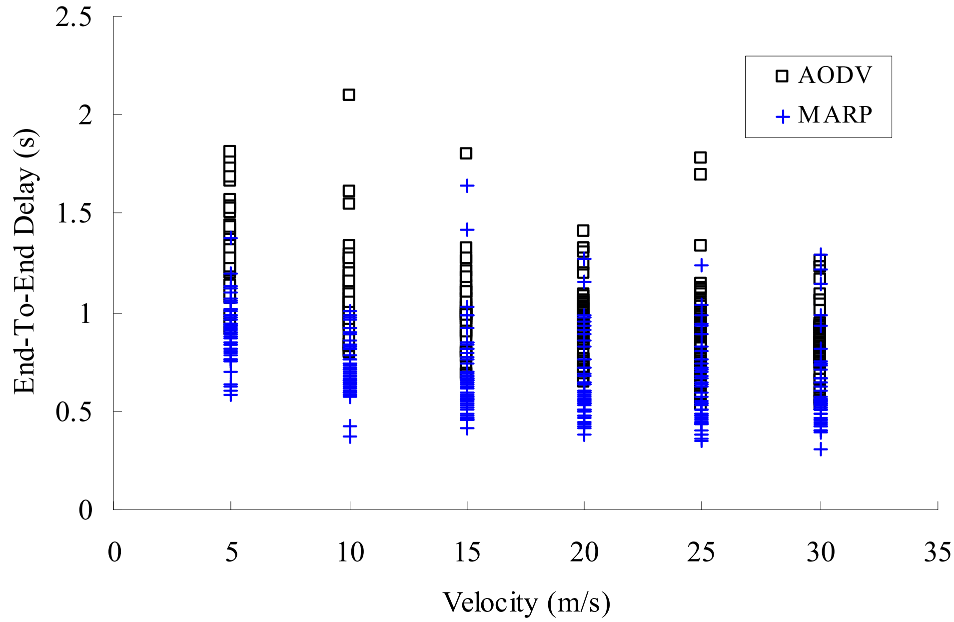


Figure 6.7: The distribution of the average end-to-end delay as a function of mobility velocity

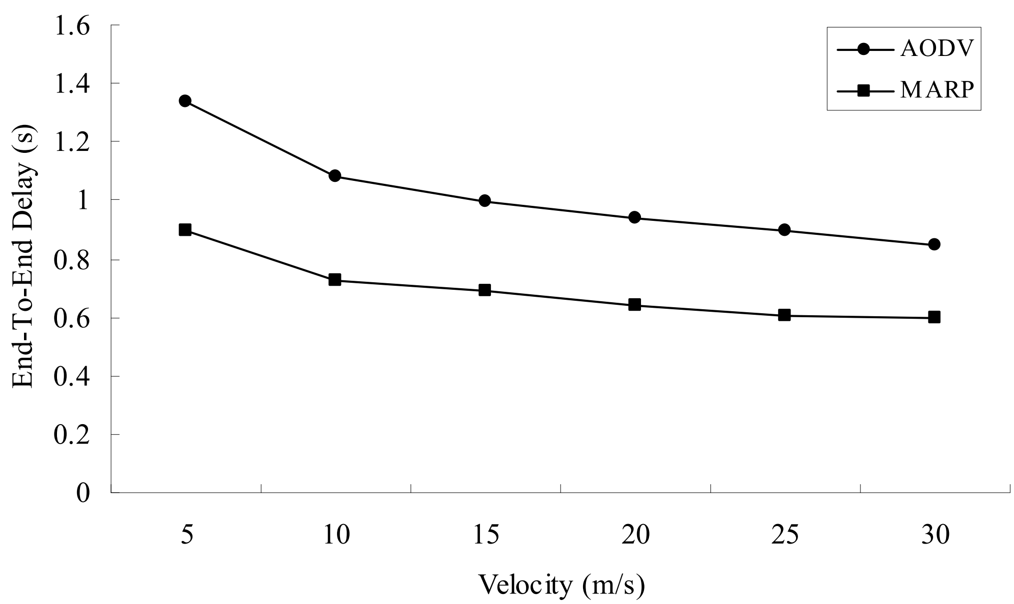


Figure 6.8: The average end-to-end delay as a function of mobility velocity

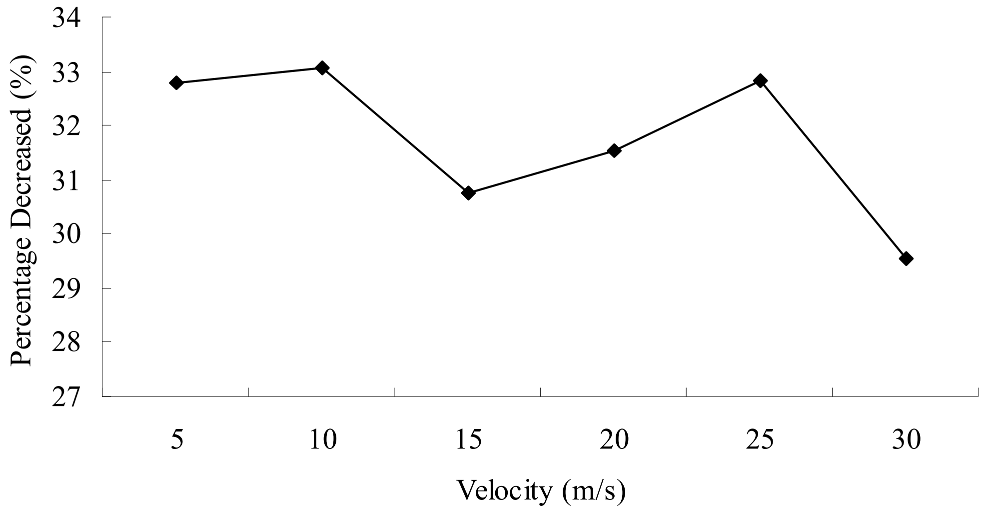


Figure 6.9: The percentage decreased of the end-to-end delay as a function of mobility velocity

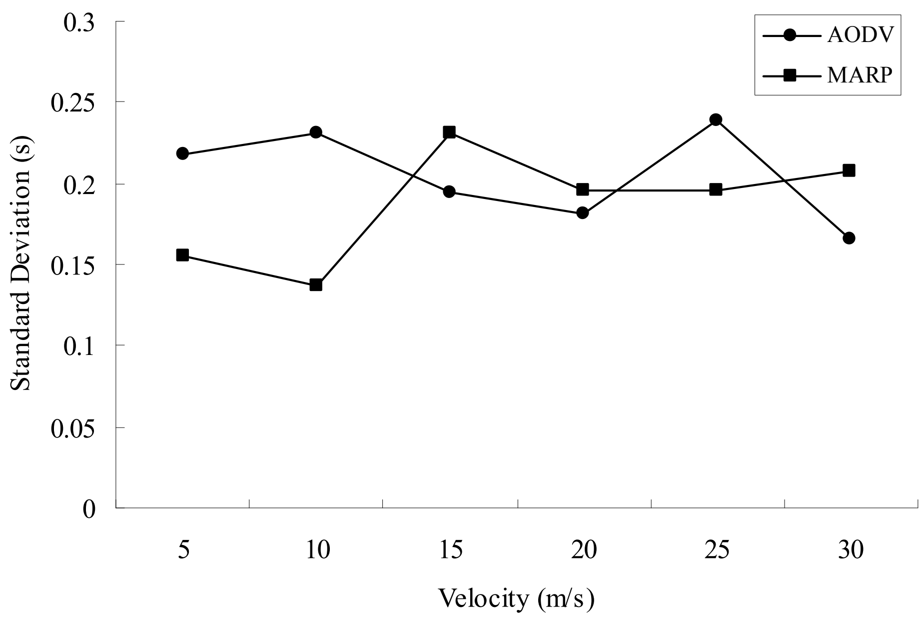


Figure 6.10: The standard deviation of the end-to-end delay as a function of mobility velocity

In the following discussion, the overhead is defined as the number of RREQ packets initialized and forwarded, because it is the RREQ packets that dominate the route control packets.

Figure 6.11 shows the distribution of the average overhead for each time as a function of mobility velocity. Figure 6.12 shows the average overhead for each speed as a function of the mobility velocity. From these two figures, it can be seen that our MARP will incur more overhead compared to AODV. Also, the average overhead increases as the mobility velocity increases because when the mobility velocity increases, more local optimization might be triggered. Figure 6.13 shows the percentage increased of the overhead as the function of the mobility velocity. The average overhead is increased by 10%-50%. The percentage increase tends to be larger when the mobility velocity is high. There is a tradeoff between the overhead cost and the network performance. Figure 6.14 shows the standard deviation of the average overhead as a function of the mobility velocity. The standard deviation of the average overheads increases as the mobility velocity increases. The reason is that the MARP triggers the routing updates based on prediction. For different mobility behaviors, the overhead incurred might be quite different, especially when the mobility velocity is high.

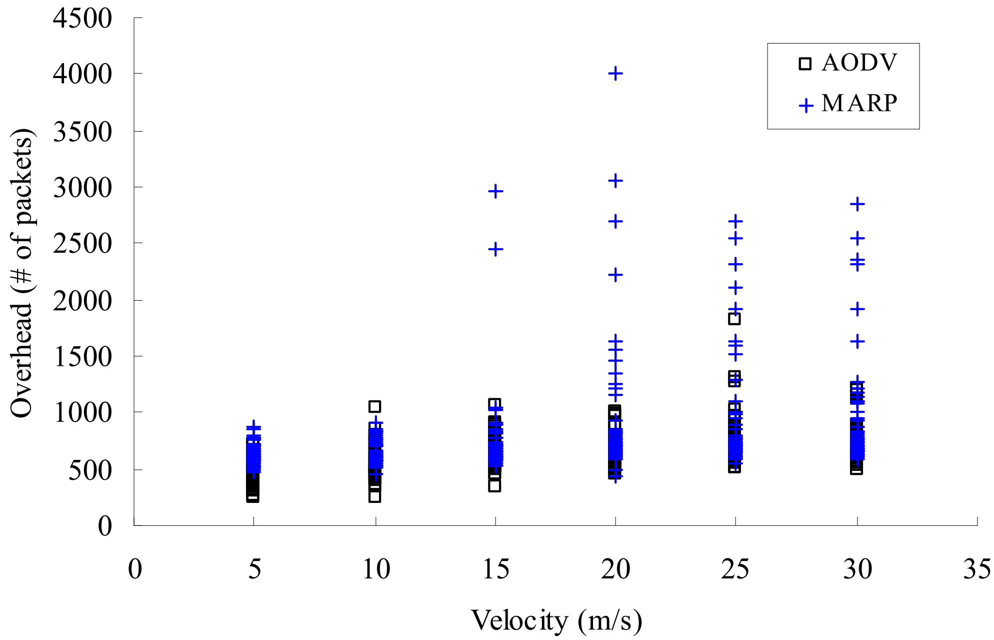


Figure 6.11: The distribution of the average overhead as a function of mobility velocity

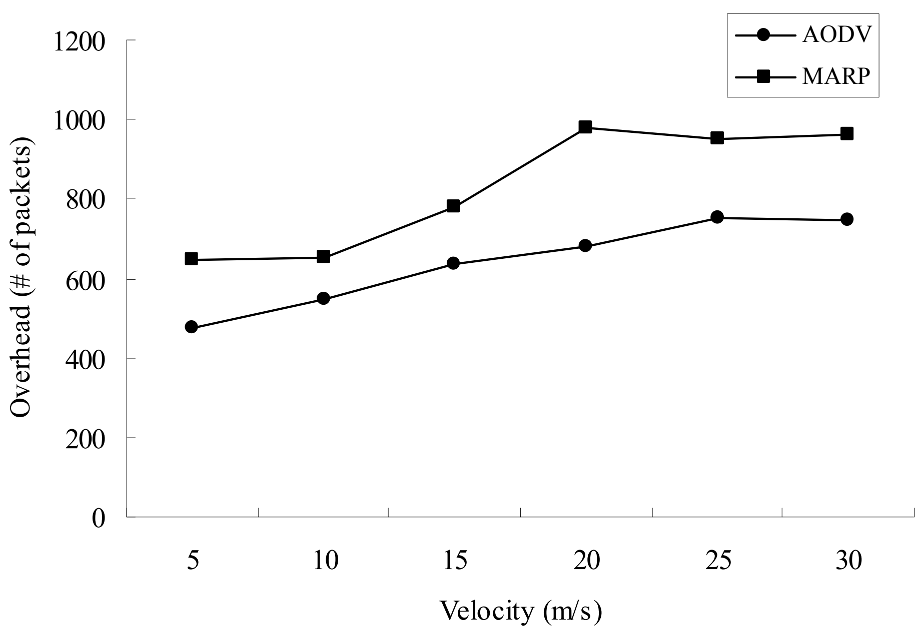


Figure 6.12: The average overhead as a function of mobility velocity

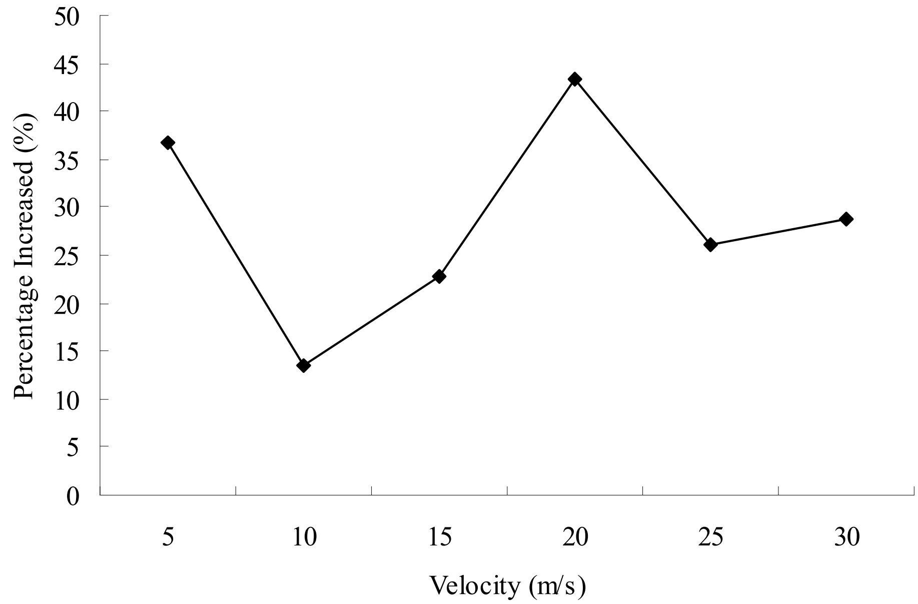


Figure 6.13: The percentage increased of the overhead as a function of mobility velocity

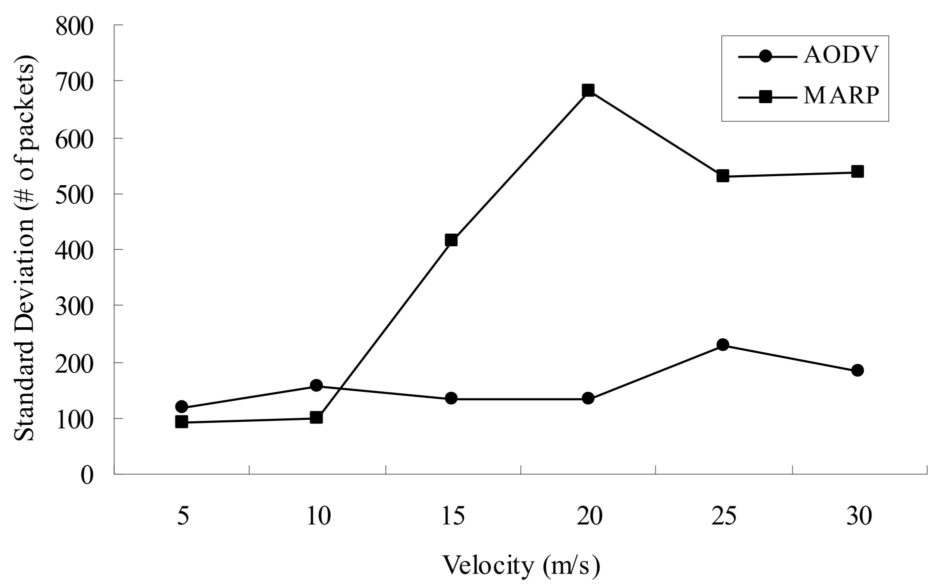


Figure 6.14: The standard deviation of the average overhead as a function the mobility velocity

In this section, we provide the results on the simulation. From these results, we found the MARP is able to increase the throughput by about 30% and decrease the end-to-end delay by about 30%. And the cost of the improvement is that the overhead is increased by 10%-50%. However, the key factor for the MARP to increase the overall performance is that the nodes are able to find a better next hop after they trigger the routing updates. Actually, if the nodes are not able to find a better next hop, the overall network performance might be worse than the traditional routing protocols. The reason is that some of the bandwidth might be wasted by the routing updates. Therefore, for mobile ad-hoc network in which the nodes are distributed sparsely, the MARP cannot work very well. However, if the nodes are distributed densely, the MARP will increase the performance significantly.

Also, in the simulations, we found that the links might still break with our MARP. One of the reasons is that our machine learning method is based on prediction. The accuracy is about 80%-90% as we illustrated in our previous work [20]. It means that our MARP might miss some of the potential link breaks. Another reason is that the nodes will not retry the RREQ for local optimization. Therefore, if the nodes did not receive the RREP, they will miss the potential local optimization. However, we found that if the nodes retry the local optimization, much more overhead might be incurred, because in many situations, there are actually no better next hops than the current next hop. That is why we let the nodes in our MARP not retry the local optimization.

Finally, compared to the performance of the traditional routing protocol, the performance of our MARP tends to have larger variance. The reason is that the performance of our MARP is determined by the mobility behavior of the nodes in the networks. It is the mobility behavior that determines when to trigger the routing updates. Therefore, in some situations, the routing updates might be triggered very frequently and many overhead might be incurred. In these situations, the performance of our MARP should be much better than the performance of the traditional routing protocols.

Chapter 7

Conclusion

7.1 Lessons Learned

In this dissertation, I proposed several intelligent approaches for routing protocol in cognitive ad-hoc networks. This dissertation begins with a detailed modeling and analysis of the new CogNet architecture and then offers a detailed description of the approach, mathematical analysis, and simulation results for the several routing protocols developed in the course of this work.

As discussed earlier, the fundamental concept for these cognitive routing protocols is that a proper and adaptive network topology should be constructed by the nodes using cognitive functions that make predictions based on past experience. The nodes in CRNs employ machine learning techniques to learn past experiences and make wise decisions by predicting future network conditions. The cognitive protocol architecture presented here is a cross-layer optimized architecture where the lower layer knowledge of the wireless medium is shared with the network layer.

Based on the simulation results, it was demonstrated that the network performance could be increased significantly by use of cognitive routing protocols. Multiple ways of using cognitive functions were explored, such as the multi-channel optimized approach, the scalability optimized cognitive approach, the multi-path optimized approach, and the mobility optimized approach. The benefits of the new protocols were substantial.

7.2 Future Work

This section discusses some opportunities for future work based on this dissertation.

In chapter 2, I provide the detailed modeling and analysis of the CogNet architecture. However, the detailed implementation is much more complicated than simulation. When the CogNet architecture is implemented in real devices, more implementation details with fewer assumptions must be considered.

In chapters 3, 4, 5, and 6, I provide a detailed description of the approach, mathematical analysis, and simulations results for the cognitive routing protocols. Of course, if the routing protocols are evaluated in experiments with actual wireless devices, the communication environments will be much more complex than were captured in these simulations and analyses. Demonstration and testing of the cognitive routing protocols in an experimental testbed will be very challenging and enlightening.

References

[1] M. M. Buddhikot, “Cognitive Radio, DSA and Self-X: Towards Next Transformation in Cellular Networks”, IEEE Symposium on New Frontiers in Dynamic Spectrum, April, 2010.

[2] H. Song and X. Lin, “A Novel DSA-Driven MAC Protocol for Cognitive Radio Networks”, Wireless Sensor Network, Vol.1, No.2, pp. 1945-3078, July 2009.

[3] C. Zou and C. Chigan, “On game theoretic DSA-driven MAC for Cognitive Radio Networks”, Computer Communications, Vol. 32, Issue. 18, Dec 2009.

[4] V. Kawadia and P. R. Kumar, “A Cautionary Perspective on Cross Layer Design”, IEEE Wireless Communication., vol.12, no. 1, pp. 3–11. Feb. 2005.

[5] S. Shakkottai, T. S. Rappaport, and P. C. Karlsson, “Cross-Layer Design for Wireless Networks”, IEEE Communication. Mag., vol. 41, no. 10, pp. 74–80, Oct. 2003.

[6] V. T. Raisinghani and S. Iyer, “Cross-Layer Design Optimizations in Wireless Protocol Stacks”, Comp. Communication., vol. 27, pp. 720–24, 2004.

[7] D.D. Clark, “The Structuring of Systems using Upcalls”, in: ACM Symposium on Operating Systems, pp. 171–180, 1985.

[8] G.H. Cooper, “The Argument for Soft Layer of Protocol”, Technical Report TR-300, Massachussets Institute of Technology, Cambridge, MA, 1983, May.

[9] D.D. Clark, D.L. Tennenhouse, “Architectural Considerations for New Generations of Protocols”, in: ACM SIGCOMM, Philadelphia, PA, pp. 200–208, 1990.

[10] V. T. Raisinghani and S. Iyer, “Cross-Layer Design Optimizations in Wireless Protocol Stacks”, Computer Communications, Vol. 27, Issue. 8 pp. 720-724, May 2004.

[11] V. Srivastava and M. Motani, “Cross-Layer Design: A Survey and the Road Ahead”, Vol. 43, Issue. 12, pp. 112-119, Dec, 2005.

[12] Q. Wang and M. A. Abu-Rgheff, “Cross-Layer Signaling for Next-Generation Wireless Systems”, IEEE Wireless Communication and Network Conference, New Orleans, LA, Mar. 2003.

[13] R. Braden, T. Faber, and M. Handley, “From Protocol Stack to Protocol Heap — Role-Based Architecture”, Hot Topics in Network, Princeton, NJ, Mar. 2002.

[14] Q. Zhao and B. M. Sadler, “A Survey of Dynamic Spectrum Access”, Signal Processing Magazine, Vol. 24, No. 3. pp. 77-89, 2007.

[15] H. Su and X. Zhang, “CREAM-MAC: An Efficient Cognitive Radio-Enabled Multi-Channel MAC Protocol for Wireless Networks”, World of Wireless, Mobile and Multimedia Neworks, 2008.

[16] G. Feng, W. Chen and Z.Cao, “A Joint PHY-MAC Spectrum Sensing Algorithm Exploiting Sequential Detection”, Signal Processing Letters, 2010.

[17] D. J. De Couto, D. Aguayo,, B. A. Chambers, and R. Morris, “Performance of multihop wireless networks: Shortest path is not enough”, ACM SIGCOMM,

Proceedings of the First Workshop on Hot Topics in Networks (HotNets-I), Princeton, New Jersey, October 2002,

[18] R. Dube, C. D. Rais, K. Y. Wang and S. K. Tripathi, “Signal stability-based adaptive routing for ad hoc mobile networks”, IEEE Personal Communication Magazine, pp. 36-45, Feb, 1997.

[19] V. Lenders, M. May and B. Platter, “Density-based vs. Proximity-based Anycast Routing for Mobile Networks”, IEEE INFOCOM, Barcelona, Spain, April 2006.

[20] V. Park and S. Corson, “Temporally-Ordered Routing Algorithm (TORA)”, IETF Internet Draft, July 2001.

[21] J. Eriksson, “DART: Dynamic Address Routing for Scalable Ad Hoc and Mesh Networks”, IEEE/ACM Transactions on Networking, Vol. 15, Issue 1, pp. 119-132, February 2007.

[22] B. Karp and H. Kung, “Greedy Perimeter Stateless routing”, Proceedings of Sixth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom 2000), 2000.

[23] P. Bose, P. Morin, I. Stojmenovic and J. Urrutia, “Routing with Guaranteed Delivery in Ad-Hoc WIreless Networks”, ACM Wireless Networks, November 2001.

[24] J. Gao, L. J. Guibas, J. Hershburger, L. Zhang, A. Zhu, “Geometric Spanner for Routing in Mobile Networks”, Proceedings of the 2nd ACM Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc 2001), pp. 44-55, October 2001

[25] K. Leung and B.-J. Kim, "Frequency assignment for IEEE 802.11 wireless networks", in IEEE Vehicular Technology Conference, Oct. 2003.

[26] P. Mahonen, J. Riihijarvi, and M. Petrova, "Automatic channel allocation for small wireless local area networks using graph coloring algorithm approach", in IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Sept. 2004.

[27] A. Mishra, S. Banerjee, and W. Arbaugh, "Weighted coloring based channel assignment for WLANs", in Mobile Computing and Communications Review, July 2005.

[28] A. Mishra, V. Brik, S. Banerjee, A. Srinivasan, and W. Arbaugh, "A client-driven approach for channel management in wireless LANs", IEEE INFOCOM, 2006.

[29] Ashish Raniwala, Kartik Gopalan, and Tzi-cker Chiueh, "Centralized channel assignment and routing algorithms for multi-channel wireless mesh networks", Mobile Computing and Communications Review, vol. 8, no. 2, pp. 50–65, April 2004.

[30] Pradeep Kyasanur and Nitin H. Vaidya "Routing in Multi-Channel Multi-Interface Ad Hoc Wireless Networks", Technical Report, December 2004.

[31] H. Khalife, S.S. Ahuja, N. Malouch, M. Krunz, "Probabilistic path selection in opportunistic cognitive radio networks", IEEE GLOBECOM, November 2008.

[32] B. Zhang, Y. Takizawa, A. Hasagawa, A. Yamauchi, and S. Obana, “Tree-based routing protocol for cognitive wireless access networks”, IEEE WCNC 2007, pp. 4207-4211, March 2007.

[33] G. Zhu, M.D. Felice, I.F. Akyildiz, “STOD-RP: a spectrum-tree based on demand routing protocol for multi-hop cognitive radio networks”, IEEE GLOBECOM, November 2008.

[34] I. Pefkianakis, S.H.Y. Wong, and S. Lu, “SAMER: Spectrum Aware Mesh Routing in Cognitive Radio Networks”, IEEE DySPAN, Oct. 2008.

[35] A. Sampath, L. Yang, L. Cao, H. Zheng, and B. Y. Zhao, “High Throughput Spectrum-aware Routing for Cognitive Radio based Ad-hoc Networks”, CrwownCom 2008, 2008.

[36] K. R. Chowdhury and M. D. Felice, “Search: A Routing Protocol for Mobile Cognitive Radio Ad-Hoc Networks”, Computer Communications, vol. 32, Issue 18, pp. 1983-1997, December 2009.

[37] H. Khalife, S. Ahuja, N. Malouch and M. Krunz, “Joint Routing and Spectrum Selection for Multihop Cognitive Radio Networks”, Tech. Rep., 2007.

[38] M. Sharma, A. Sahoo and K. Nayak, “Channel Selection Under Interference Temperature Model in Multi-Hop Cognitive Mesh Networks”, IEEE DySPAN, 2007.

[39] G. Cheng, W. Liu, Y, Li and W. Cheng, “Joint On-Demand Routing and Spectrum Assignment in Cognitive Radio Networks”, IEEE ICC, 2007.

[40]D. B. Johnson and D. A. Maltz, “Dynamic Source Routing in Ad Hoc Wireless Networks”, Mobile Computing, Kluwer Academic Publishers, vol. 353, pp. 153-181, 1996.

[41]C. E. Perkins and E. M. Royer, “Ad Hoc On-Demand Distance Vector Routing”, Proceedings of IEEE Workshop on Mobile Computing Systems and Applications, 1999, pp. 90-100, February 1999.

[42]C. K. Toh, “Associativity-Based Routing for Ad Hoc Mobile Networks”, Wireless Personal Communication, vol. 4, no. 2, pp. 1-36, March 1997.

[43]R. Dube, C. D. Rais, K. Y. Wang and S. K. Tripathi, "Signal Stability-Based Adaptive Routing for Ad Hoc Mobile Networks", IEEE Personal Communications Magazine, pp. 36-45, February 1997.

[44]W. Su and M. Gerla, “IPv6 Flow Handoff in Ad Hoc Wireless Networks Using Mobility Prediction”, Proceedings of IEEE GLOBECOM 1999, pp. 271-275, December 1999.

[45]R. S. Sisodia, B. S. Manoj and C. Siva Ram Murthy, “A Preferred Link-Based Routing Protocol for Ad Hoc Wireless Networks”, Journal on Selected Areas in Communications, vol. 17, no.8, pp. 14-21, March 2002.

[46]N. S. M. Usop, A. Abdullah and A. F. A. Abidin, “Performance Evaluation of AODV, DSDV & DSR Routing Protocol in Grid Environment”, IJCSNS International Journal of Computer Science and Network Security, vol.9, no.7, pp. 261-268, July 2009.

[47] E. E. Perkins and P. Bhagwat, "Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers", Proceedings of ACM SIGGOMM 1994, pp. 234-244, August 1994.

[48] C. C. Chiang, H. K. Wu, W. Liu and M. Gerla, "Routing in Clustered Multi-hop Mobile Wireless Networks with Fading Channel", Proceedings of IEEE SICON 1997, pp. 197-211, April 1997.

[49] Elizabeth M. Belding-Royer "Hierarchical routing in ad hoc mobile networks," Wireless Communications and Mobile Computing, vol. 2, pp. 515-532, Sept 2002.

[50] A. Iwata, C. C. Chiang, G. Pei, M. Gerla and T. W. Chen, "Scalable Routing Strategies for Ad Hoc Wireless Networks", IEEE Journal on Selected Areas in Communications, vol. 17, no.8, pp. 1369-1279, August 1999.

[51] B. Karp and H. Kung, "Greedy Perimeter Stateless routing", Proceedings of Sixth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom 2000), 2000.

[52] P. Bose, P. Morin, I. Stojmenovic and J. Urrutia, "Routing with Guaranteed Delivery in Ad-Hoc WIreless Networks", ACM Wireless Networks, November 2001.

[53] J. Gao, L. J. Guibas, J. Hershburger, L. Zhang, A. Zhu, "Geometric Spanner for Routing in Mobile Networks", Proceedings of the 2nd ACM Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc 2001), pp. 44-55, October 2001

[54] V. Lenders, M. May and B. Platter, "Density-based vs. Proximity-based Anycast Routing for Mobile Networks", in IEEE INFOCOM, Barcelona, Spain, April 2006.

[55] V. Park and S. Corson, "Temporally-Ordered Routing Algorithm (TORA)", IETF Internet Draft, July 2001.

[56] J. Eriksson, "DART: Dynamic Address Routing for Scalable Ad Hoc and Mesh Networks", IEEE/ACM Transactions on Networking, Vol. 15, Issue 1, pp. 119-132, February 2007.

[57] K. R. Chowdhury and M. D. Felice, "Search: A Routing Protocol for Mobile Cognitive Radio Ad-Hoc Networks", Computer Communications, vol. 32, Issue 18, pp. 1983-1997, December 2009.

[58] H. Khalife, S. Ahuja, N. Malouch and M. Krunz, "Joint Routing and Spectrum Selection for Multihop Cognitive Radio Networks". Tech. Rep., 2007.

[59] M. Sharma, A. Sahoo and K. Nayak, "Channel Selection Under Interference Temperature Model in Multi-Hop Cognitive Mesh Networks", in Proceedings of IEEE DySPAN Conference, 2007.

[60] G. Cheng, W. Liu, Y, Li and W. Cheng, "Joint On-Demand Routing and Spectrum Assignment in Cognitive Radio Networks", in Proceedings of IEEE ICC Conference, 2007.

[61] http://www.scalable-networks.com

[62] H. Khalife, S.S. Ahuja, N. Malouch, M. Krunz, "Probabilistic path selection in opportunistic cognitive radio networks", IEEE GLOBECOM, November 2008.

[63] B. Zhang, Y. Takizawa, A. Hasagawa, A. Yamauchi, and S. Obana, “Tree-based routing protocol for cognitive wireless access networks”, IEEE WCNC 2007, pp. 4207-4211, March 2007.

[64] G. Zhu, M.D. Felice, I.F. Akyildiz, “STOD-RP: a spectrum-tree based on demand routing protocol for multi-hop cognitive radio networks”, IEEE GLOBECOM, November 2008.

[65] I. Pefkianakis, S.H.Y. Wong, and S. Lu, “SAMER: Spectrum Aware Mesh Routing in Cognitive Radio Networks”, IEEE DySPAN, Oct. 2008.

[66] A. Sampath, L. Yang, L. Cao, H. Zheng, and B. Y. Zhao, “High Throughput Spectrum-aware Routing for Cognitive Radio based Ad-hoc Networks”, CrwownCom 2008, 2008.

[67] Sung-ju Lee and Mario Gerla, “Split Multipath Routing with Maximally Disjoint Paths in Ad hoc Networks” IEEE ICC , vol. 10, pp. 3201-3205, August 2001.

[68] X. Li and L. Cuthbert, “Node-Disjoint Multipath Routing and Distributed Cross-Layer QoS Guarantees in Mobile Ad Hoc Networks”, Proceedings of 7th ACIS international conference on software engineering, pp. 243-248, 2006.

[69] X. Li and L. Cuthbert, “On-Demand Node-Disjoint Multipath Routing in Wireless Ad Hoc Network”, 29th annual IEEE international conference on local computer networks, pp. 419-220, 2004.

[70] A. M. Abbas and B. N. Jain, “Path Diminution in Node-Disjoint Multipath Routing for Mobile Ad Hoc Networks in Unavoidable with Single Route

Discovery", International journal of ad hoc and ubiquitous computing, vol. 5, No. 1, pp 7-21, 2010.

[71] S. Ramasubramainian, H krishnamoothy and M. Krunz, "Disjoint Multipath Routing Using Colored Trees", Computer Networks, Vol. 51, pp. 2163-2180, June 2007.

[72] X. Li and L. Cuthbert, "Stable Node-Disjoint Multipath Routing with Low Overhead in Mobile Ad Hoc Networks", 12th IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunications Systems, pp. 184-191, 2004.

[73] Marc Mosko and J.J. Garcia-Luna-Aceves "Multipath Routing in Wireless Mesh Networks" in First IEEE Workshop on Wireless Mesh Networks (WiMesh 2005).

[74] S. De. C. Qiao and H. Wu, "Meshed Multipath Routing: An Efficient Strategy in Sensor Networks", Wireless communications and networking, vol. 3, pp, 1912-1917, New Orleans, LA, 2003.

[75] D. S. Nandiraju, N. S. nandiraju and D. P. Agrawal, "Adaptive State-Based Multi-Radio Multi-Channel Multipath Routing in Wireless Mesh Networks", Pervasive and mobile computing, vol. 5, pp. 93-109, 2009.

[76] L. Bonori, M. D. Felice, A. Molinaro and S. Pizzi, "Joint Channel Assignment and Multipath Routing for Multi-Radio Mesh Networks", Proceedisng of 29th IEEE internation conference on distributed computing systems workshops, pp. 476-481, 2009

[77] K. R. Chowdhury and M. D. Felice, “Search: A Routing Protocol for Mobile Cognitive Radio Ad-Hoc Networks”, Computer Communications, vol. 32, Issue 18, pp. 1983-1997, December 2009.

[78] H. Khalife, S. Ahuja, N. Malouch and M. Krunz, “Joint Routing and Spectrum Selection for Multihop Cognitive Radio Networks”. Tech. Rep., 2007.

[79] M. Sharma, A. Sahoo and K. Nayak, “Channel Selection Under Interference Temperature Model in Multi-Hop Cognitive Mesh Networks”, in Proceedings of IEEE DySPAN Conference, 2007.

[80] G. Cheng, W. Liu, Y, Li and W. Cheng, “Joint On-Demand Routing and Spectrum Assignment in Cognitive Radio Networks”, in Proceedings of IEEE ICC Conference, 2007.

[81] R. Draves, J. Padhye and R. Zill, “Routing in Multi-Radio, Multi-Hop Wireless Mesh Networks”, International conference on mobile computing and networking, pp. 114-128, Philadelphia, PA, USA, 2004.

[82] R. Boppana and S. Konduru. “An Adaptive Distance Vector Routing Algorithm for Mobile, Ad Hoc Networks,” In IEEE Infocom 2001, Mar 2001.

[83] T. Goff, N. B. Abu-Ghazaleh, D. S. Phatak, R. Kahvecioglu, “Preemptive routing in Ad Hoc Networks,” In ACM SIGMOBILE, Roma, 2001.

[84] P. Srinath et al., “Router Handoff: A Preemptive Route Repair Strategy for AODV”, Personal Wireless Communications, IEEE International Conference, 2002.

[85] A. Boukerche and L. Zhang "A performance evaluation of a pre-emptive on-demand distance vector routing protocol for mobile ad hoc networks", wireless communications and mobile computing, Vol. 4, 2004, pp. 99–108.

[86] Elizabeth M. Belding-Royer "Hierarchical routing in ad hoc mobile networks," Wireless Communications and Mobile Computing, vol. 2, pp. 515-532, Sept2002.

[87] Mario Joa-Ng and I-Tai Lu "A peer-to-peer zone based two level link state routing for mobile ad hoc networks," IEEE Journal on Selected Areas in Communications, vol. 17, No.8, Aug 1999.

[88] Charles E. Perkins Elizabeth M. Royer "Ad-hoc On-Demand Distance Vector Routing," in MILCOM'97 panel on Ad Hoc Networks, Nov, 1997.

[89] Charles E. Perkins and Pravin Bhagwat "Highly dynamic destination sequenced distance vector routing for mobile computers," In Proc. Of the ACM SIGCOMM, Oct 1994

[90] Ben Y.Zhao, Yitao Duan and ling Huang "Brocade: Landmark routing on overlay networks," In Proceedings of 1st International Workshop on Peer-to-Peer Systems (IPTPS) , March 2002.

[91] Liliana M. Arboleda C. and Nidal Nasser "Cluster-based routing protocol for mobile sensor networks," In QShine'06, The Third International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks, Aug 2006.

[92] Navid Nikaein, Houda Labiod and Christian Bonnet "DDR-distributed dynamic routing algorithm for mobile ad hoc networks," In MobiHOC 2000, pp 19-27.

[93] Navid Nikaein, Christian and Neda Nikaein “HARP- hybrid ad hoc routing protocol,” International Symposium on Telecommunications, IST 2001, 2001.

[94] Atsushi Iwata, Ching-Chuan Chiang and Guangyu Pei “Scalable routing strategies for ad hoc wireless networks,” IEEE Journal on Selected Areas in Communications, vol. 17, No.8 Aug 1999.

[95] Dipankar Raychaudhuri, Joseph B. Evans Srinivasan Seshan “CogNet- An architectural foundation for experimental cognitive radio networks within the future internet” MobiArch ’06, December, 2006, San Francisco CA, USA.

[96] Richard Draves, Jitendra Padhye and Brian Zill “Routing in multi-radio, multi-hop wireless mesh networks,” In Proceedings of MOBICOM 2004, Philadelphia.

[97] Venugopalan Ramasubramanian, Zygmunt J. Haas and Emin Gun Sirer “SHARP: A hybrid adaptive routing protocol for mobile ad hoc networks,” In Proceedings of the 4th ACM International Symposium on Mobile Ad Hoc Network and Computing pp 303-314, 2003.

[98] Mingliang Jiang, Jinyang Li and Y. C. Tay “Cluster based routing protocol (CBRP),” Internet Draft (work in progress), 8/98.

[99] Brad Karp and H. T. Kung “Greedy perimeter stateless routing for wireless networks,” In Proc. ACM/IEEE MobiCom, August 2000.

[100] Hsinchun Chen “Machine learning for information retrieval: Neural networks, symbolic learning, and genetic algorithms ,” In Journal of the American Society for Information Science, pp. 194-216 1995.

[101] Suyang Ju and Joseph B. Evans “Intelligent multi-path selection based on parameters prediction”, IEEE CogNet 2008 Workshop, Beijing, China, May 2008.

[102] http://www.scalable-networks.com

[103] Raymond Yim, Neelesh B. Mehta Andreas F. Molisch, “Efficient multiple access using received signal strength and local channel information”, IEEE WCNC 2008 proceedings

[104] Razvan Pitic and Antonio Capone, “An opportunistic scheduling shceme with minimum data-rate guarantees for OFDMA”, IEEE WCNC 2008 proceedings.

[105] J Del Prado and S. Choi, ‘‘Link adaptation strategy for IEEE 802.11WLAN via received signal strength measurement,’’ ICC03, May 2003, pp. 1108-1113.