

ENGINEERING JOURNAL

Article

Cognitive Wireless Sensor Networks: Intelligent Channel Assignment

Soamsiri Chantaraskul* and Chayaphon Tanwongvarl

The Sirindhorn International Thai-German Graduate School of Engineering (TGGS), King Mongkut's University of Technology North Bangkok, 1518 Pibulsongkram Road, Bangsue, Bangkok, Thailand

*E-mail: soamsiri.c.ce@tggs-bangkok.org

Abstract. One of the major problems in Wireless Sensor Networks (WSNs) is the traffic congestion caused by increasing number of devices sharing the limited spectrum of the ISM (Industrial, Scientific, and Medical radio) band. As a result, a new concept of Cognitive Wireless Sensor Networks (CWSNs) has been proposed in order to achieve reliable and efficient communication via spectrum awareness and smart adaption. Based on such concept, this paper proposes the intelligent channel assignment technique for channel management in CWSNs. The proposed method is based on the learning and prediction technique so called Policy Gradient together with our proposed virtual channel environment classification. Simulation model is used for the system performance evaluation. The simulation results show that our proposed intelligent channel assignment provides substantially higher performance in terms of system throughput and average packet end-to-end delay than the traditional IEEE 802.15.4 based system. It also outperforms the systems integrated with Episodic Reinforcement and GPOMDP learning technique.

Keywords: Wireless sensor networks, cognitive radio, channel assignment, policy gradient.

ENGINEERING JOURNAL Volume 21 Issue 1

Received 16 September 2016

Accepted 16 December 2016

Published 31 January 2017

Online at <http://www.engj.org/>

DOI:10.4186/ej.2017.21.1.279

1. Introduction

Nowadays, Wireless Sensor Networks (WSNs) receive high interest as important approach facilitating the ongoing works in many clusters such as industry, military, and medical. Essential examples are the healthcare system networks, which are needed to ensure that important alarming messages will be sent to central control or head office of the hospital. Based on WSN supporting standard like the IEEE 802.15.4, this message transfer will be done on a limited spectrum of the ISM band. One of the crucial challenges is that the ISM band is a free unlicensed band. Therefore, many communication devices are operating on the 2.4 GHz ISM band such as WiFi routers, Bluetooth devices, and wireless sensors. As a result, general usage environments of such WSN devices are likely to have many transceivers coexisting in the same area. This leads to unavoidable and higher interference to the devices. The coexistence issue within the 2.4 GHz ISM band has been studied, and the result shows that IEEE 802.15.4 devices, which are part of WSN, seem to have a severe impact from other systems, especially from the IEEE 802.11 devices operating on an overlapping frequency band in the same area [1, 2]. This is because IEEE 802.15.4 standard is designed for Low-Rate Wireless Personal Area Networks (LR-WPANs) with low transmission power devices.

Large number of work has been done and methods have been proposed to solve the coexistence problems. They can be categorized according to [1] into methods based on time management, space management, and channel management. Most of the management methods need the co-operation between sufferer and interferer. It can be implemented by time interval control, which means the unused time intervals are used by different network devices. The time management can also be implemented by traffic control. A larger packet interval can reduce the probability of packet conflict in time dimension. The space management is rather unrealistic when trying to change the space between the devices by moving them frequently. On the other hand, there is a close relationship between power fading and distance. Therefore, space management is often implemented by transmit power control and receiving power sensitivity control (CCA, Clear Channel Assignment, threshold control). However, these studies still lack of detailed design and implementation.

The channel management focuses on the usage of a dedicated interference-free channel. Channel management can be broken down into two types, central control and distributed control. The central control mechanism is aimed at broadcasting control channel message to the entire network embedded with information on suitable channel to be used. [3] presents active interference detection schemes, which ensure that the channel changing mechanism is done based on node clustering. The drawback of central control comes from the fact that it relies on availability of dedicated control channel, hence significantly limits the networks' scalability and robustness. Distributed control, on the other hand, does not depend on the node's coordination. However, all devices must operate with the identical algorithms. In [4], distributed channel selection algorithm for sensor nodes is presented in order to automatically adjust their channel in response to the 802.11 interference.

In the first two categories, the drawback comes from the complexity and performance limitation when implementing in the network with a large number of sensor nodes. Therefore, channel management is the one with many proposed approaches. In our previous work [5-7], several mechanisms have been proposed under the channel management category. [5] presents the dual-bridge technique, which has been proposed to enhance system performance by extending traditional single bridge node using dual-bridge nodes working together for data transfer across the interfered area. Multichannel CTP (Collection Tree Protocol) is proposed in [6-7] allowing multichannel usage on top of existing routing ability to deliver better system performance. These example mechanisms and other proposed mechanisms of the same scale offer certain system adaptations based on current traffic situation. Nevertheless, such mechanisms do not concern about the changing behavior of the environment, hence there is no long-term system adaptation.

Cognitive Wireless Sensor Networks (CWSN) implements cognitive radio techniques to enhance system performance by using opportunistic spectrum access. The concept enables real-time interaction with the environment in order to determine the appropriate system adaptation dynamically. In this work, the idea of cognitive capability in a cognitive radio is taken into account. Intelligent channel assignment is proposed for CWSN. The proposed mechanism extends a traditional cognitive concept to include learning and prediction process via an artificial intelligent approach.

The rest of this paper is organized as follows. In section two, learning techniques in cognitive radio are provided. Section three presents our proposed intelligent channel assignment. In section four, the simulation model utilized in this work for the performance evaluation and test scenario are given. Section five shows the simulation results and discussion. Finally, the paper is concluded in section six.

2. Learning in Cognitive Radio

From the overview of cognitive radio research, the overall system performance is influenced by the individual process. In the traditional cycle, the results obtained by one process are used by another process to reach the intended purposes. However, the extensive system performance can be enhanced, if such learnt result can be offered in a long term to provide more efficient algorithm used in every single process. This work is aimed at developing an engine that uses the learnt knowledge over time in order to support individual function for better overall system performance. The proposed method is to develop the learning engine and eventually create the prediction process. The prediction of radio environment is one of the most challenging aspects to implement in the cognitive radio system. Since the spectrum sensing cannot be done at all time, the prediction process should enhance the system performance in a long run. The new cognitive cycle proposed in this work is shown in Fig. 1.

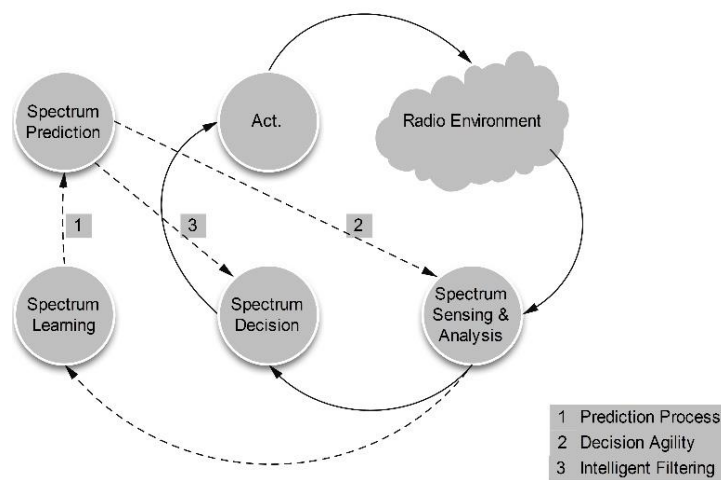


Fig. 1. Proposed cognitive cycle with spectrum learning and spectrum prediction.

The cycle includes the process of spectrum learning and spectrum prediction in addition to the traditional cognitive cycle. The spectrum prediction is used to process the learnt knowledge and predict what could happen according to the information from spectrum sensing & analysis and spectrum decision-making process. Appropriate mathematical models that support the learning and prediction process might include learning process such as Q-Learning, Markov Model, etc. The prediction process does not interfere with the normal operation. It can be operated periodically and the outcome can be applied and used to modify the equations and/or algorithms used by each process.

As mentioned, the first step in the cognitive cycle is the spectrum sensing & analysis. The accuracy and speed in recognition of radio environment are some of the biggest issues. With the reactive system addressed earlier, the event needs to take place before any changes can be made. As a result, cognitive radio terminal may end up with the delayed detection process, hence the system will continue using ineffective radio configuration for a certain time. This could result in degradation of the overall system performance. On the other hand, the system would benefit more from a proactive mindset that strives to make changes before they are needed. One important approach to solve the processing delay problem is to use radio environment prediction mechanism. By doing so, the cognitive radio can predict the availability of wireless media for transmission capacity estimation. In [8], a prediction of radio environment for cognitive radio is proposed. With the radio environment observation being done in various locations for the unlicensed band of 2.4 GHz, the result indicates that the radio channel occupation time presents long-range dependency with the time scale longer than 0.1 second. Hence, it is possible to predict the radio channel occupation time close to reality, in the condition when the time scale of time series is longer than 0.1 second.

Information from the prediction engine could also be used by the spectrum decision in order to provide an agile decision making. This might cover the new prediction index used to determine the wireless resource. The theory for simulating and understanding human cognition is introduced in [9] based on Adaptive Character of Thought (ACT) and it may be used in embedded algorithm of the prediction engine.

From the proposed mechanism as mentioned, the benefits brought by using the prediction process are introduced, although there are rather limited works proposed in the area. In addition, the previously proposed methods are still not efficient for most real-time applications, since it is difficult to implement such schemes on a standalone cognitive radio with limited processing and memory storage. Consequently, the prediction engine that improves overall system performance simultaneously while providing applicable implementation is important.

Generally, there are three approaches for machine learning including the unsupervised learning, the supervised learning, and the reinforcement learning, see Fig. 2 [10] for the illustration. Unsupervised learning is based on given properties and attempts to discover an appropriate model, which is the cluster data type. Supervised learning is a generated function that represents a granted data set based on the training data. This approach is suitable for neural networks or to support vector machines. Finally, reinforcement learning is a collection of algorithms designed to solve sequential decision problems. They are often used with a delayed reward. Therefore, this approach is also known as semi-supervised learning, in which the approaches can be divided into direct learning and indirect learning. In the direct learning approach, the mechanism attempts to find the transfer function directly in the policy space, such as Policy Gradient methods or Evolutionary Algorithms. The indirect approaches mostly use value function-based methods. This mechanism is intended to learn a value function in order to obtain an excellent policy.

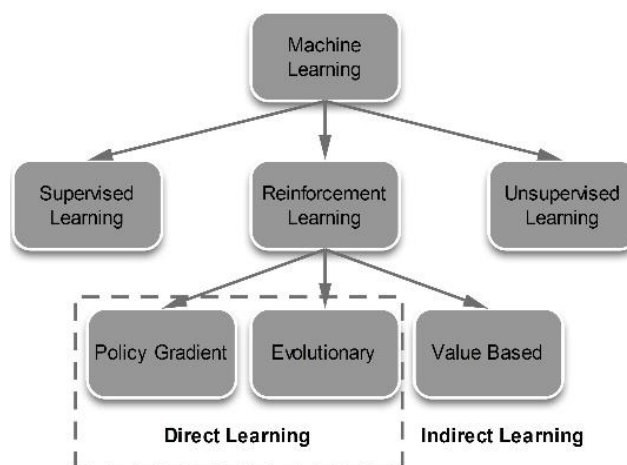


Fig. 2. The machine learning approaches.

As stated, channel management has several parameters to consider and the reinforcement learning mechanism is rather suitable to learn the problems of the channel assignment. However, majority of the proposed methods in the reinforcement learning community are not related more than three or four degrees of freedom and/or cannot handle parameterize policies [11]. On the other hand, the Policy Gradient mechanism is applicable for high-dimensional systems such as channel assignment.

2.1. General Assumption

Proposed in [10], Eq. (1) characterizes the probability of success to achieve a new state under the assumption of a specific action a_k in a state x_k and current time step k as reinforcement learning assumes the probability distribution. Initial assumption of the next state is referred to the probability distribution of the next state under the currently assumed action and state.

$$X_{k+1} \sim p(X_{k+1} | X_k, a_k) \quad (1)$$

Moreover, based on the Policy Gradient methods, all actions for the agent can be generated by a probabilistic policy:

$$a_k \sim \pi_\theta(a_k, X_k) \quad (2)$$

The index θ denotes the parameters of the policy. Based on the consideration of each policy π_θ , the value function and state-action function are given by the Bellman equations:

$$V^n(x) = r(x) + \alpha \sum_{x'} p(x' | x, \pi(x)) V^n(x') \quad (3)$$

$$Q^n(x, a) = \alpha \sum_{x'} p(x' | x, a) V(x') \quad (4)$$

where a denotes the discount factor for problems with unlimited horizon. H is for horizon (learning limit), which can be infinite, and at certain time learning system will receive reward, which is $r(x_n, a_n)$. The goal of Policy Gradient methods is to optimize the expectation of the cumulated reward function:

$$J(\theta) = \frac{1}{\alpha \Sigma} E \left\{ \sum_{k=0}^H \alpha_k r_k \right\} \quad (5)$$

The term α_Σ denotes a normalization factor to ensure that $\sum_{k=0}^H \frac{\alpha_k}{\alpha_\Sigma} = 1$. Note that we assume a value of $\alpha_k=1$ for all further equations and experiments. An additional representation of Eq. (5) computes this expectation as the integral over the state distribution and the integral over-all actions on the policy by

$$J(\theta) = \int_x d^\pi(x) \int_A \pi_\theta(a, x) r(a, x) dx da \quad (6)$$

The next sections use this equation to derive the Policy Gradient for the parameter θ .

2.2. Gradient Estimation

Refer to [10], the gradient estimation equation is

$$g \approx E \left\{ \left(\sum_{k=0}^H \nabla_\theta \log \pi_\theta(a_k, x_k) \right) \left(\sum_{l=0}^H r(a_l, x_l) \right) \right\} \quad (7)$$

Next, we can rewrite Episodic Reinforcement of the gradient estimation as

$$\left\{ \left(\sum_{k=0}^H \nabla_\theta \log \pi_\theta(a_k, x_k) \right) \left(\sum_{l=0}^H r_l - b \right) \right\} \quad (8)$$

* b = Estimated optimal baseline

Finally, GPOMDP gradient estimation is

$$\left\{ \sum_{l=0}^H \left(\sum_{k=0}^l \nabla_\theta \log \pi_\theta(a_k, x_k) \right) (r_l - b) \right\} \quad (9)$$

Figure 3 illustrates the map of True Policy Gradient. The background plot represents the rewards map, and the foreground plot represents the map of Policy Gradient estimation. The blue plot at the background shows the calculated reward map. The initializing value is represented by red dots and the appropriate policy for suitable estimation is represented with green crosses. When comparing with other estimations i.e. Episodic Reinforcement and GPOMDP, the expected reward map of the True Policy Gradient provides the best fittings than other methods since the suitable values are in the same level when comparing with other estimations [12]. Note that the maps for Episodic Reinforcement and GPOMDP can be seen in our previous paper [12], which also shows the simulation results for the comparison of prediction accuracy. Simulation results illustrated that the True Policy Gradient offers the highest estimation accuracy while Episodic Reinforcement and GPOMDP provides lower and the lowest value, respectively.

3. Proposed Intelligent Channel Assignment

Based on the method of cooperative spectrum sensing, types of Cognitive Radio Networks (CRNs) can be categorized into centralized and decentralized networks. In centralized cooperative spectrum sensing, network coordinator collects local observations from the sensor nodes then decides on the spectrum occupancy by using decision fusion rules and informs the sensor nodes on which channel to access [13, 14]. In decentralized cooperative spectrum sensing, sensor nodes within a CRN exchange their local spectrum sensing results among themselves without requiring a backbone or network coordinator [15]. In this paper, the centralized approach is employed. The reason is that the learning process requires more power consumption, hence it is unsuitable to be done by the end devices, which are battery powered. Moreover,

the system will gain benefit from learning efficiency. This is because the central control will have full knowledge of the system environment by obtaining information from all sensor nodes.

The intelligent channel assignment proposed in this work is based on the learning and prediction technique so called True Policy Gradient estimation combined with our virtual channel environment classification. The general architecture of the proposed intelligent channel assignment consists of three major components, which are WSN nodes, virtual channel environment classification engine, and learning and decision making engine. Illustrated in Fig. 4 is the proposed channel assignment learning flow for WSNs.

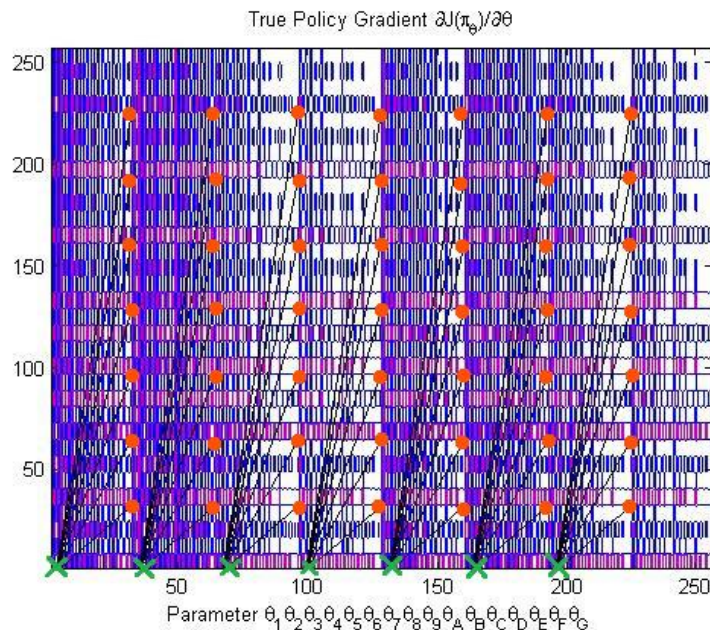


Fig. 3. Map of the True Policy Gradient estimation.

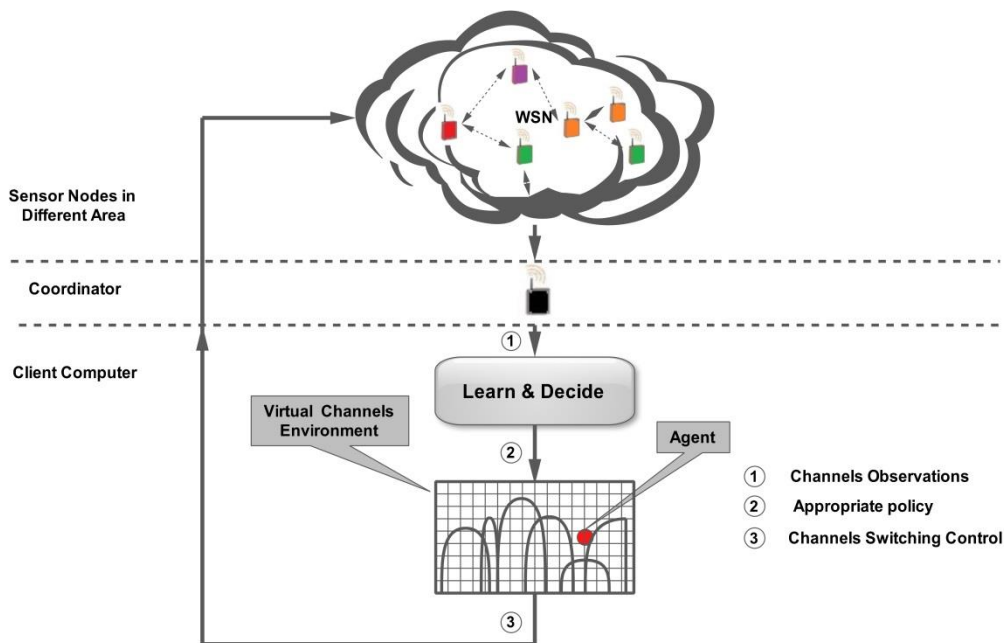


Fig. 4. A channel assignment learning flow for wireless sensor networks.

3.1. Learning and Decision Making

Table 1 illustrates the internal process for learning and decision making engine with the Policy Gradient being utilized.

Table 1. Client computer using policy gradient.

INPUT: N =NUMBER OF STATES(NUMBER OF CHANNEL) , M =NUMBER OF ACTIONS(LIFT-DOWN,RIGHT-DOWN), α = DISCOUNTED REWARD	
1.	While <i>System running</i> do
2.	Get the observation O_t from sensor nodes
3.	Create automatically transition probability table of the sixteen state problem based on actions
4.	Automatically create reward table based on channel level selection
5.	Calculate expected reward map from $\theta_1, \theta_2, \theta_3, \dots, \theta_{16}$
6.	Obtaining Policy Gradient estimation
7.	Learning process for policy optimization
	end
1.	Set initial channel
2.	Get action form optimized policy
3.	Get appropriate channel
4.	Send control message to a coordinator for changing the new frequency

In general, there are two types of a decision-making process including deterministic approach and stochastic approach. Deterministic approach offers the predictable results, while the stochastic approach is suitable for non-static environment, which does not provide accurate prediction of how the result will come out. However, stochastic approach could explain the probability of such result using unlimited processing in order to relate distributed parameters together. Based on the Reinforcement Learning, most of the problems are a discrete type, which is the type of stochastic approach such as the spectrum usage of the ISM frequency band. Table 1 explains the process for computing solution, in this case, suitable channel for WSN to be utilized. When the channel environment is observed, step 3 and 4 will determine system environment limitation and possible criteria, which will be used in step 5 for probability estimation. In Step 5, the engine calculates all potential solutions based on Markov process, which analyses parameters' behaviour for future prediction by adapting the parameters $\theta_1, \theta_2, \theta_3, \dots, \theta_{16}$ (policies). These parameters are non-stationary. The final step is plotting the discounted distribution during the Markov process (Expected reward for this work) based on the specified problem and policy.

Note that State (x) is a current channel usage status (16 possible states). Action (a) is the decision, which will change according to channel usage that has less interference. Based on channel observation mapping, direction channel changing can either be to-the-left or to-the-right-hand side depending on a result of the next state. Reward (r) is the result of changing channel, and the value depends on whether the new channel has less or more interference than the current state.

Step 6 and 7 are intended to find resulting parameters based on the formulas shown in the previous section, Eq. (7), Eq. (8), Eq. (9). Subsequently, the learning process will obtain appropriate policy, which will best suit the environment. The system will set the initial channel as the current channel that the sensor network is currently using then calculate the next suitable channel based on the obtained policy.

3.2. Virtual Channel Environment Classification

The operational example of our proposed virtual channel environment mapping and classification is shown in Fig. 5. The process begins when sensor nodes send channel observation information to the network coordinator. Coordinator then maps information received from all sensor nodes into a single grid called channel observation map, which concludes the current state of the channels. In the next step, learning and decision making engine will deliver action according to the input information from the channel observation map. Then, an action together with channel observation map will be input to the classification process. Through the classification process, a positive or negative response will be offered. The response will be handled as a reward for the action chosen. Then, it will be sent back to the learning and decision making

engine for further optimization. Here, observation is the collection of all channels' status in the area operated by wireless sensor networks. State is the currently chosen channel by sensor network and policy is the suitable condition for current channel environment, which impacts the section of action to resolve interference on currently used channel.

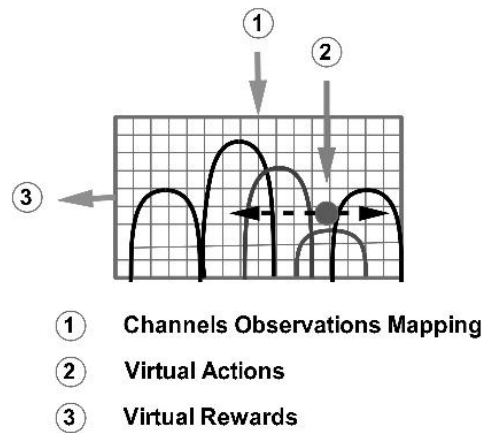


Fig. 5. Virtual channels environment classification.

In the classification process, basic classification approach is used based on [16]. If the sensor node is close to IEEE 802.11 node, the packet loss will increase as illustrated in Fig. 6. Therefore, the algorithm can be written as

$$y = \left\{ \begin{array}{lll} r_{v1} & \text{if} & a_p > c_p \quad : \text{positive - overlap} \\ r_{v2} & \text{if} & a_p < c_p \quad : \text{negative - overlap} \\ r_{v3} & \text{if} & a_p = c_p \quad : \text{overlap} \\ r_{v4} & \text{if} & \text{only } a_p \quad : \text{none - overlap} \end{array} \right\} \quad (10)$$

The output is $y \in \{r_{v1}, r_{v2}, \dots, r_{vc}\}$, where r_v is the virtual reward, and c is the number of classes. Input $x \in \{a_{p1} : c_{p1}, a_{p2} : c_{p2}, \dots, a_{pn} : c_{pn}\}$ is data, where a_p is the agent power position and c_p is channel power position.

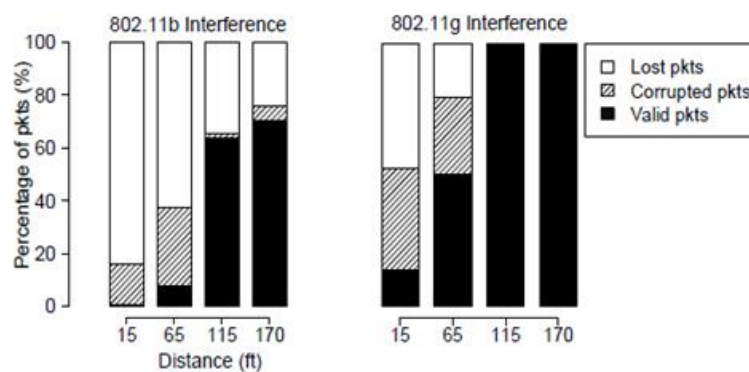


Fig. 6. Percentage of packets reception ratio between IEEE 802.11 and IEEE 802.15.4 [16].

4. Simulation Model

To observe the performance of our proposed learning approaches for CWSN, the simulation model is developed. The simulation model for CWSNs employed here is based on the open-source Castalia simulator [17], which has the basis of the OMNeT++ platform [18]. The main features of Castalia are the

channel model based on empirically measured data, and the advanced radio model based on really low-power radios. The WSN test scenario used here is given in Fig. 7, and the simulation parameters used for the simulation are given in Table 2.

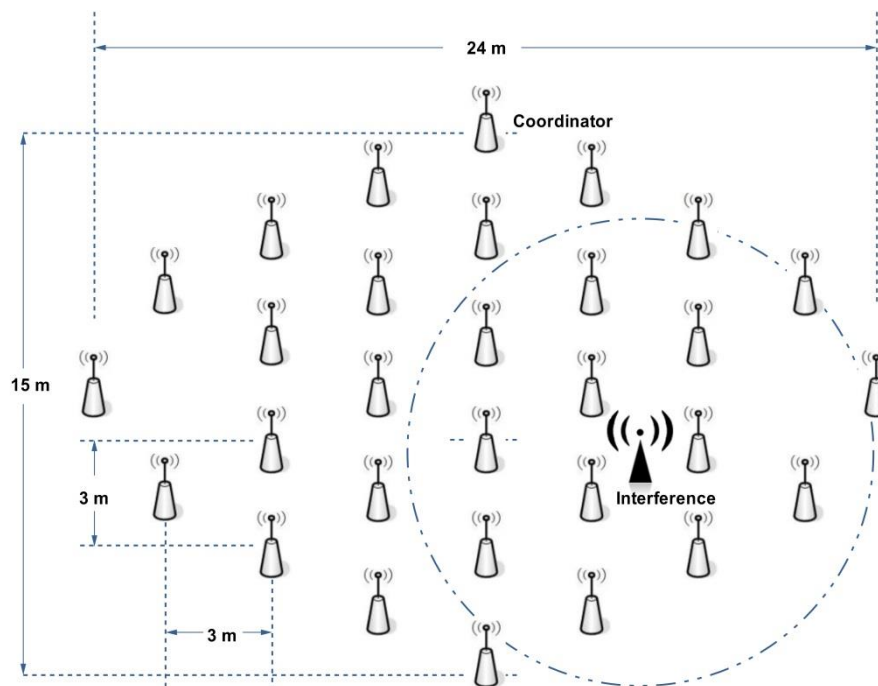


Fig. 7. Simulation test scenario for the IEEE 802.15.4 based WSNs with interference.

Table 2. Simulation parameters.

Parameters	Device	
	IEEE 802.15.4	Interference
Application : Packet Rate	1-20 (packet/sec)	200 (packet/sec)
MAC : Data Rate	250 kbps	1024 kbps
Radio : Output Power	-25 dBm	-10 dBm
Carrier Frequency	2450 MHz	2450–2460 MHz

5. Simulation Results

Based on the simulation model and test scenario illustrated in section four, tests have been done, and the simulation results can be illustrated in the following subsections.

5.1. Simulation for Systems with and without Interference

In this subsection, the simulation has been done for the traditional IEEE 802.15.4 based system and the CWSN embedded with our proposed method in two scenarios, including no interference case and interference case. Simulation results are presented in Figs. 8-10.

In Fig. 8, system performance is monitored in terms of system throughput as packet generation rate is being varied for the traditional IEEE 802.15.4 based system in comparison with the system implemented with our proposed method. The dotted lines are the results observed from the no interference scenario. It can be seen that without being interfered, both systems offer very close throughput performance. This is because although our proposed technique needs the central control to collect status information from end

nodes and perform learning and prediction mechanism, the detected information is designed to be sent along with the regular beacon communication. Hence, there is no extra overhead on transferring of such control information. The solid lines represent the simulation results obtained under interference case. It can be seen that at high-traffic generation rate, system throughput of a traditional IEEE 802.15.4 based WSN is substantially lower than that offered by the system with the proposed methods. This is because the learning and prediction process allows the system to switch to operate on the channel with lower interference.

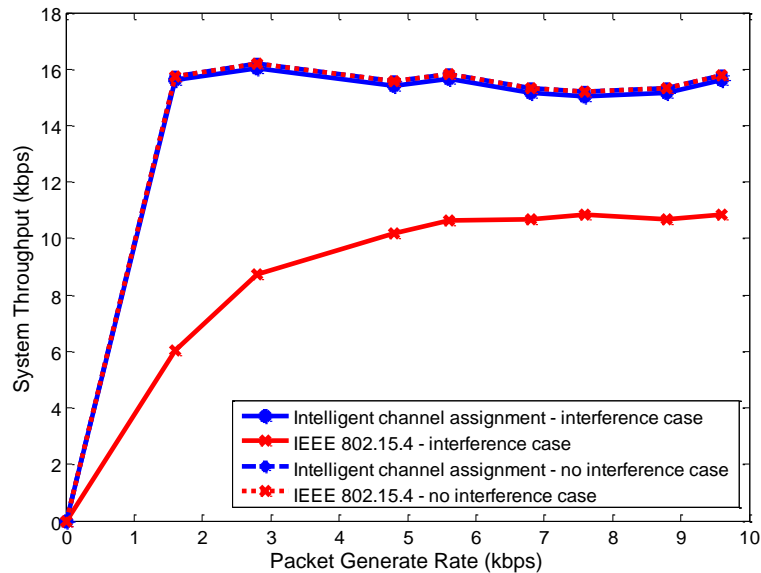


Fig. 8. System throughput performance comparison.

Figures 9 and 10 present the simulation results in terms of average packet end-to-end delay performance comparison and application level latency performance comparison, respectively. It can be seen in Fig. 9 that both traditional system and system with proposed techniques perform similarly under no interference case for the reason stated before. However, under interference scenario, system integrated with the proposed intelligent channel assignment outperforms conventional systems with rather low average delay for all levels of traffic load. This is due to the intelligent adaptability of the proposed scheme.

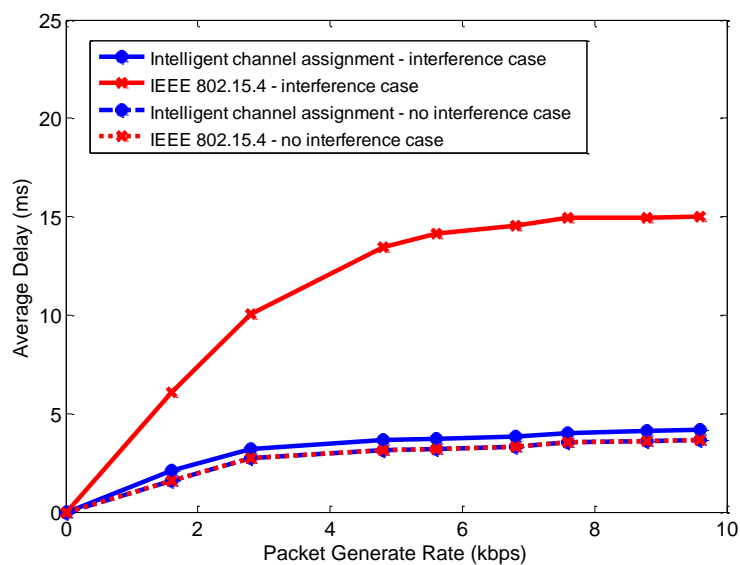


Fig.9. Average packet end-to-end delay performance comparison.

Figure 10 presents the simulation results for system performance in terms of application level latency captured as simulation progress. Bars with a dotted pattern represent results under no interference case. It can be seen that under no interference both system offer a similar range of latency at low level. Bars with solid colour represent results under interference scenario. In this case, the traditional system offers generally higher level of latency in comparison with no interference case and the values are violating throughout the monitoring time. This is caused by the transmission is being disrupted by the interference. The system with proposed intelligent channel assignment performs differently in this interference case. It offers slightly high latency at the beginning and declining latency as time progresses. This is because the proposed system needs to learn from the initial experience to be able to adapt itself in this case assigning more suitable channel as interference occurs in order to maintain overall system performance.

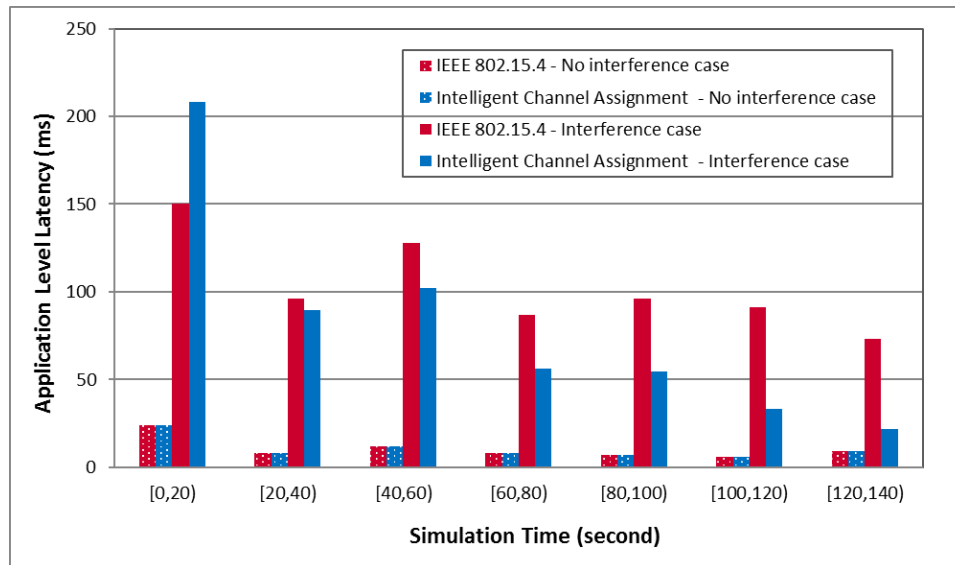


Fig. 10. Application level latency (Packet Generate Rate = 7.6 kbps).

5.2. Simulation for Systems with Proposed Technique and Other Learning Techniques

This subsection provides simulation results for the system performance comparison of three WSN systems under interference scenario, including system with our proposed intelligent channel assignment (labelled as Policy Gradient in Figs. 11-13), system integrated with Episodic Reinforcement as a learning technique, and system integrated with GPOMDP as a learning technique.

In Fig. 11 and Fig. 12, system performance is monitored in terms of system throughput and average packet end-to-end delay as packet generation rate is being varied for the system implemented with our proposed method (labelled as Policy Gradient), and system with other learning techniques, including Episodic Reinforcement and GPOMDP. The result from a traditional system is also given (labelled as IEEE 802.15.4). From our previous work [12], which focuses on the selection of suitable learning approach, results from the study show that True Policy Gradient outperforms the other two techniques with 94% prediction accuracy. However, the proposed system includes virtual channels environment classification step, which further prevents wrong decision to be implemented. As a result, the improvement in terms of system throughput as shown in Fig. 11 is not significantly higher than using Episodic Reinforcement and GPOMDP.

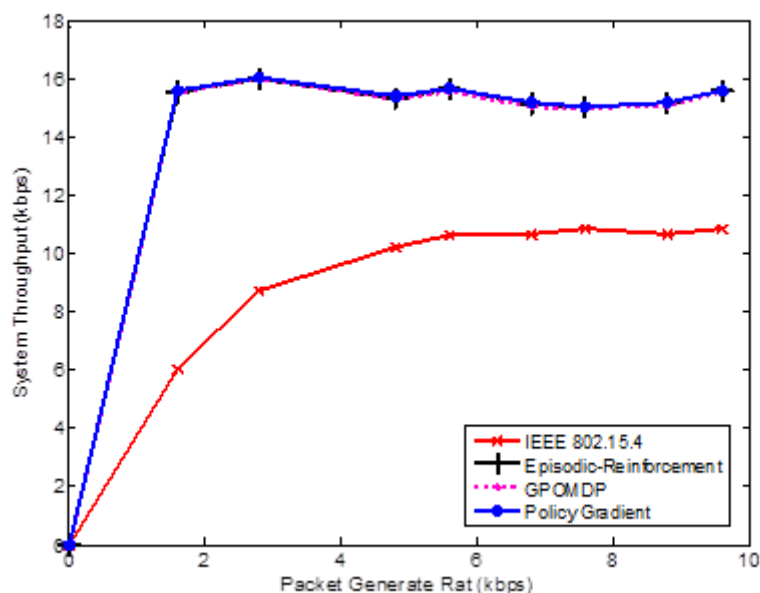


Fig. 11. System throughput performance comparison.

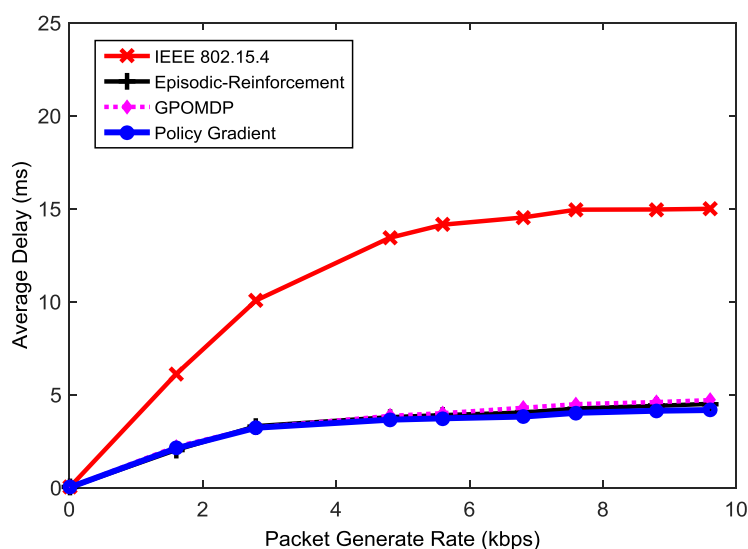


Fig. 12. Average packet end-to-end delay performance comparison.

It is obvious from the simulation results that under interference environment, the three systems implemented with learning techniques outperform the traditional IEEE 802.15.4 based system both in terms of throughput and average packet delay. With more accurate learning and prediction process done by the proposed method, the proposed intelligent channel assignment based system offers the best throughput and delay performance, especially at high traffic load.

Figure 13 presents the simulation results for system performance comparison in terms of application level latency captured as simulation progress. As mentioned in the previous subsection, traditional system offers a similar level of latency throughout monitoring time since the system does not adapt itself according to environmental change. On the other hand, the other three systems are implemented with learning algorithms, hence the latency level declines at times. In this aspect, the level of latency among the three learning techniques are very similar.

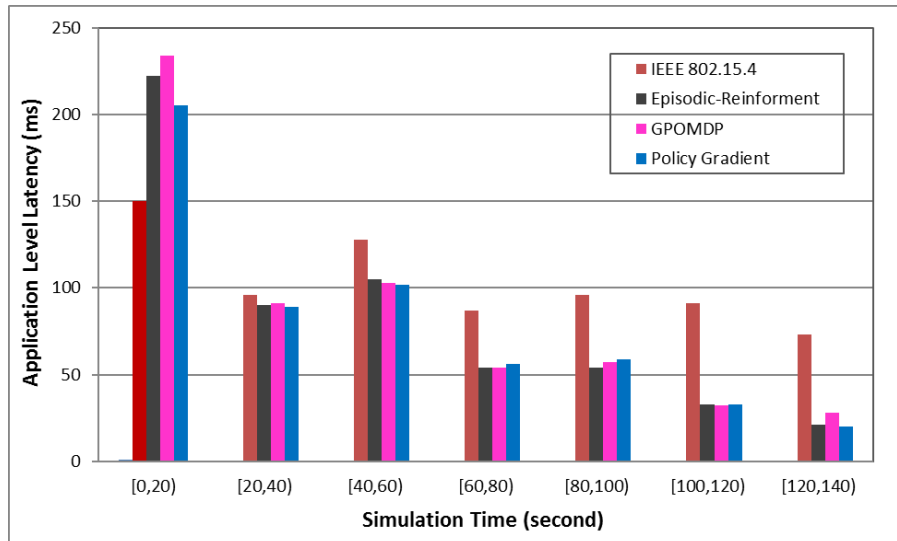


Fig. 13. Application level latency (Packet Generate Rate = 7.6 kbps).

6. Conclusions

This paper proposes an intelligent channel assignment for WSN based on the True Policy Gradient under our proposed virtual channels environment. The proposed mechanism adopts a cognitive radio concept of intelligent spectrum awareness and proposes suitable learning and prediction technique in order to obtain the optimal channel assignment in CWSNs. As a result, the system has higher tolerance to the interference. System performance is evaluated through the developed simulation model. The simulation results show that under interference environment our approach offers better system performance in terms of system throughput and average packet end-to-end delay in comparison with traditional IEEE 802.15.4 based system as well as the system integrated with other two learning techniques, i.e. Episodic Reinforcement and GPOMDP.

Acknowledgement

This research was funded by King Mongkut's University of Technology North Bangkok, Contract no. KMUTNB-GOV58-48 and supported by a scholarship of TOT Public Company Limited of Thailand.

References

- [1] D. Yang; Y. Xu, and M. Gidlund. "Coexistence of IEEE802.15.4 based networks: A survey," in *Proc. IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society*, Nov. 2010, pp. 2107–2113.
- [2] A. Sikora and V. F. Groza, "Coexistence of IEEE802.15.4 with other systems in the 2.4 GHz-ISM-Band," in *Proceedings of the IEEE Instrumentation and Measurement Technology Conference*, May 2005, pp. 1786-1791.
- [3] M. S. Kang, J. W. Chong, H. Hyun, S. M. Kim, B. H. Jung, and D. K. Sung, "Adaptive interference-aware multi-channel clustering algorithm in a ZigBee network in the presence of WLAN interference," presented at *Int. Symposium on Wireless Pervasive Computing 2007*, Feb. 2007.
- [4] S. Pollin, M. Ergen, M. Timmers, A. Dejonghe, L. Van der Perre, F. Catthoor, I. Moerman, and A. Bahlai, "Distributed cognitive coexistence of 802.15.4 with 802.11," in *Proc. IEEE Int'l Conf. Cognitive Radio Oriented Wireless Networks and Comm. (CROWNCOM '06)*, June 2006, pp. 1-5.
- [5] C. Tanwongvarl and S. Chantaraskul, "Dual-bridge: A mode-adaptive operation for multi-channel utilization in wireless sensor networks," in *Proc. IET International Conference on Information and Communications Technologies (IETICT 2013)*, 2013, pp. 481–486.
- [6] C. Buengbon, C. Tanwongvarl, and S. Chantaraskul, "Multi-channel collection tree protocol for wireless sensor networks," in *Proc. 2013 10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, 2013, pp. 1–5.

- [7] A. Phokaew, C. Tanwongvarl, and S. Chantaraskul, "Adaptive multi-channel CTP for wireless sensor networks," in *Proc. Electrical Engineering Congress (iEECON), 2014 International*, 2014, pp. 1–4.
- [8] K. Takeuchi, S. Kaneko, and S. Nomoto, "Radio environment prediction for cognitive radio," in *Proc. of the CrownCom, 2008*, Singapore, May 2008, pp. 1–6.
- [9] Carnegie Mellon University, (2009). Welcome to ACT-R. [Online]. Available: <http://act-r.psy.cmu.edu/>
- [10] A. Witsch, "Applying policy gradient reinforcement learning to optimise robot behaviours," Diss. Faculty of Electrical Engineering and Computer Science, University of Kassel, Germany, 2010.
- [11] J. Peters and S. Schaal, "Reinforcement learning of motor skills with policy gradients," *Neural Networks*, vol. 21, no. 4, pp. 682–697, May 2008.
- [12] C. Tanwongvarl and S. Chantaraskul, "Performance comparison of learning techniques for intelligent channel assignment in cognitive wireless sensor networks," in *Proc. The 7th International Conference on Ubiquitous and Future Networks (ICUFN 2015)*, Sapporo, Japan, July 2015, pp. 503-507.
- [13] V. Moghiss, M. R. Meybodi, and M. Esnaashari, "An intelligent protocol to channel assignment in wireless sensor networks: Learning automata approach," in *Proc. 2010 International Conference on Information Networking and Automation (ICINA)*, 2010, vol. 1, pp. V1–338–V1–343.
- [14] M. Bkassiny, S. K. Jayaweera, Y. Li, and K. A. Avery, "Optimal and low-complexity algorithms for dynamic spectrum access in centralized cognitive radio networks with fading channels," in *Proc. Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, pp. 1–5.
- [15] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 589–600, Apr. 2007.
- [16] C. Liang, "Interference characterization and mitigation in large-scale wireless sensor networks," Doctoral dissertations, Johns Hopkins University, Baltimore, Maryland, 2011.
- [17] Castalia [Online]. Available: <https://castalia.forge.nicta.com.au/index.php/en/> [Accessed Feb 2015].
- [18] N. LTE, O. released, I. released, I. released and O. released, 'OMNeT++ Discrete Event Simulator - Home', Omnetpp.org, 2014. [Online]. Available: <http://www.omnetpp.org>. [Accessed: Feb 25, 2015].