CHALLENGES OF APPLYING MARKOV DECISION PROCESS TO WIRELESS SENSOR NETWORKS: A REVIEW

P.K. Santhi[1], D. Selvi[1] and Y. Arputha Sahaya Rani[2]
[1]Post Graduate and Research Centre of Mathematics
[2]Central Library
Jayaraj Annapackiam College for Women (Autonomous), Periyakulam,
Theni District, Tamilnadu, India.

Abstract

Wireless sensor networks (WSNs) consist of autonomous and resource-limited devices. WSNs operate as stochastic systems because of randomness in the monitored environments. The behaviour of MSN can be modeled through Markov Process and transition probabilities are calculated. The challenges of using Markov Process to WSN discussed in this paper.

Keywords: Wireless sensor networks, Markov decision processes (MDPs).

1. Introduction

Nowadays Wireless Sensor Networks (WSNs) is very useful in smart cities. To make the smart cities, we must interact with the surrounding environment's dynamics and objects with the help of sensing systems. WSNs operate in stochastic (random) environments under uncertainty. Particularly, a sensor node, as a decision maker or agent, applies an action to its environment, and then transits from a state to another. The environment can encompass the node's own properties as well as many of the surrounding objects. In such an uncertain environment, the system dynamics can be modeled using a mathematical framework called Markov decision processes (MDPs) to optimize the network's desired objectives. MDPs entail that the system possesses a Markov property. In particular, the future system state is dependent only on the current state but not the past states. Recent developments in MDP solvers have enabled the solution for large scale systems, and have introduced new research potentials in WSNs. Since the last century there have been marked changes in the approach to scientific enquires. There has been greater realization that probability (or non deterministic) models are more realistic than deterministic model in many situations. Observations taken at different time points rather than those taken at a fixed period of time began to engage the attention of probabilities.

In section II, a discussion of MDP basic concept and its solution methods are presented. Then in section III, we discuss the benefits of MDP Model to WSN. The challenges of applying MDPs to WSNs are discussed in section IV. Finally the paper is concluded and summarized in section V.

2. Markov Decision Process

2.1. Stochastic Process

A Stochastic process is a probability model that describes the evolution of a system evolving randomly in time.

Definition

The set of possible values of a single random variables X_n of a stochastic process $\{X_n, n \geq 1\}$ is known as its state space.

Definition

A random process $\{X(t)\}$ is called a Markov process if for $(t_0<t_1<t_2<\ldots\ldots<t_n)$ we have $P\{X(t_n)=a_n/X(t_{n-1})=a_{n-1}, X(t_{n-2})=a_{n-2}\ldots X(t_1)=a_1\}=P\{X(t_n)=a_n/X(t_{n-1})=a_{n-1}\}$

i.e. if the future behavior of the process depends only on the present value but not on the past values, then the process is called **Markov Process**.

If the above condition is satisfied for all n, then the process $\{X(t)\}$ is called **Markov Chain**.

Types of stochastic Process

- Discrete time, discrete state space
- Discrete time, continuous state space
- Continuous time, discrete state space
- Continuous time, Continuous state space

Limiting Distributions

If a limiting distribution π exists, it satisfies

$$\pi_j = \sum_{i \in S} \pi_i p_{i,j} \quad , j \in S \text{ and } \sum_{j \in S} \pi_j = 1$$

There are many applications of Markov chains in WSNs, such as data aggregation and routing (1,7), duty cycle (8), sensing coverage (9), target tracking (10,11,12), MAC backoff operation (13), (14) and security (15,16,17). It is used for performanic analysis.

Markov Decision Process (MDP) is used for Stochastic Optimazation (i.e.) to obtain the best action to be taken given particular objectives and possibly a set of constraints.

It is used for decision making under uncertainity (18, 19). For WSNs, the MDP is used to Model the interaction between wireless sensor node (i.e. an agent) and their surrounding environment (i.e. a system) to optimize an energy control or a routing decision in WSNs.

Benefits of MDP models

- Using MDPs for dynamically optimizing the network operations to fit the physical conditions results in significantly improved resource utilization (1).
- The MDP model allows a balanced design of different objectives, for example, minimizing energy consumption and maximizing sensing coverage. Different works, e.g., (2,4), discuss the approaches of using MDPs in optimization problems with multiple objectives.
- MDP method can explore the temporal correlation of moving objects and predicting their future locations, e.g., (5,6).
- Therefore, the MDP model can be applied even for tiny and resource-limited nodes without any high computation requirements. Moreover, near-optimal solutions can be derived to approximate optimal decision policies which enable the design of WSN algorithms with less computation burdens.

- A Markov decision process (MDP) is an optimization model for decision making under uncertainty (18, 19). The MDP describes a stochastic decision process of an agent interacting with an environment or system. At each decision time, the system stays in a certain state s and the agent chooses an action a that is available at this state.

Methods

- **Value iteration (VI):** This is the most efficiently and widely used method to solve an infinite time horizon discounted MDP This method has many advantages, e.g., quick convergence, ease of implementation, and is especially a very useful tool when the state space of MDPs is very large. Similar to the forward induction method of a finite· time horizon MDP, this approach was also developed based on dynamic programming (1).
- **Policy iteration (PI):** The main idea of this method is to generate an improving sequence of policies. It starts with an arbitrary policy and updates the policy until it converges. This approach consists of two main steps, namely policy evaluation and policy improvement. We first solve the linear equations to find the expected discounted reward under the policy π and then choose the improving decision policy for each state. (1)
- **Linear programming (LP):** Unlike the previous methods, the linear programming method aims to find a static policy through solving a linear program (20). The linear programming method is useful for MDPs with constraints since the constraints can be included as linear equations in the linear program (21).
- **Approximation method:** Approximate dynamic programming was developed for large MDPs. The method approximates the value functions (whether policy functions or value functions) by assuming that these functions can be characterized by a reasonable number of parameters. Thus, we can seek the optimal parameter values to obtain the best approximation, e.g., as given in (22, 23 and 24).

- **Online learning**: The aforementioned methods are performed in an offline fashion i.e., when the transition probability function is provided). However, they cannot be used if the information of such functions is unknown. Learning algorithms were proposed to address this problem (25).

3. Challenges of Applying MDPs to WSNs

The MDP Model is a powerful analytical tool to address stochastic optimization problems. The MDP framework has proved its applicability in many real world applications such as finance, agriculture, sports, etc (25, 30). However, there are still some limitations that need further research study.

3.1. Time Synchronization

Most existing studies assume perfect time synchronization among nodes. This assumption enables the network nodes to construct a unified MDP cycle (sense current state, make decision and take actions, sense new state, etc. Therefore, the clock of the node must be adjusted to a central timing device (31, 32) for time synchronization algorithms in WSNs. Besides, the clock may not be perfectly synchronized because of various delays. The mechanisms to address these issues must be developed.

3.2. The Course of Dimensionality

This is an inherent problem of MDPs when the state space and/or the action space become large (1). Consequently, we cannot solve MDPs directly by applying standard solution methods. Instead, approximate solutions are usually used. The work is present some examples of using approximate solutions to reduce the complexity of MDP-based methods in WSNs.

3.3. Stationarity and Time-Varying Models

It is assumed that the MDP's transition probabilities and reward function are time invariable. Nevertheless, in some systems, this assumption may be infeasible. There are two general methods to deal with non-stationary transition probabilities in Markov decision problems. In the first solution, an online learning algorithm, e.g., (33, 34), is used to update the state transition probabilities and the reward function based on the environment changes.

4. Summary

This paper has provided the extensive literature review related to an introduction Markov decision Process model and benefits of MDP solution methods. Then, challenges of applying MDPs to WSNs have been discussed. Finally, WSNs find new applications and serves as a key platform in many smart technologies and internet of things.

References

1. Markov Decision Processes with Applications in Wireless Sensor Networks: A Survey Mohammad Abu Alsheikh_Y, Dinh Thai Hoang_, Dusit Niyato_, Hwee-Pink Tany and Shaowei Liny_School of Computer Engineering, Nanyang Technological University, Singapore 639798 ySense and Sense-abilities Programme, Institute for Infocomm Research, Singapore 138632.
2. K. Chatterjee, R. Majumdar, and T. A. Henzinger, "Markov decision processes with multiple objectives," in Proceedings of the 23rd Annual Symposium on Theoretical Aspects of Computer Science. Springer, 2006, pp. 325–336.
3. K. Etessami, M. Kwiatkowska, M. Y. Vardi, and M. Yannakakis, "Multi-objective model checking of Markov decision processes," in Tools and Algorithms for the Construction and Analysis of Systems. Springer, 2007, pp. 50–65.
4. K. Chatterjee, "Markov decision processes with multiple long-run average objectives," in Proceedings of the 27th International Conference on Foundations of Software Technology and Theoretical Computer Science. Springer, 2007, pp. 473–484.
5. G.K. Atia, V.V. Veeravalli and J.A. Fuemmeler, "Sensor scheduling for energy-efficient target tracking in sensor networks," IEEE Transactions on Signal Processing, vol. 59, no. 10, pp. 4923–4937, 2011.
6. W.-L. Yeow, C.-K. Tham and W.-C. Wong, "Energy efficient multiple target tracking in wireless sensor networks," IEEE Transactions on Vehicular Technology, vol. 56, no. 2, pp. 918–928, 2007.

7. Y.M. Ko and N. Gautam, "Epidemic-based information dissemination in wireless mobile sensor networks," IEEE/ACM Transactions on Networking, vol. 18, no. 6, pp. 1738–1751, 2010.
8. K.-W. Lee, V. Pappas, and A. Tantawi, "Enabling accurate node control in randomized duty cycling networks," in Proceedings of the 28th International Conference on Distributed Computing Systems. IEEE, 2008, pp. 123–132.
9. C.Y. Ma, D.K. Yau, N.K. Yip, N.S. Rao and J. Chen, "Stochastic steepest descent optimization of multiple-objective mobile sensor coverage," IEEE Transactions on Vehicular Technology, vol. 61, no. 4, pp. 1810–1822, 2012.
10. L. Liu and J.-F. Chamberland, "Cross-layer optimization and information assurance in decentralized detection over wireless sensor networks," in Proceedings of the 40th Asilomar Conference on Signals, Systems and Computers. IEEE, 2006, pp. 271–275.
11. J. Zhang, G. Zhou, S.H. Son, J.A. Stankovic and K. Whitehouse, "Performance analysis of group based detection for sparse sensor networks," in Proceedings of the 28th International Conference on Distributed Computing Systems. IEEE, 2008, pp. 111–122.
12. L. Zheng, Y. Yao, M. Deng, and S.-T. Yau, "Decentralized detection in ad hoc sensor networks with low data rate inter sensor communication," IEEE Transactions on Information Theory, vol. 58, no. 5, pp. 3215– 3224, 2012.
13. P. Park, P. Di Marco, P. Soldati, C. Fischione and K.H. Johansson, "A generalized Markov chain model for effective analysis of slotted ieee 802.15.4," in Proceedings of the 6th International Conference on Mobile Adhoc and Sensor Systems. IEEE, 2009, pp. 130–139.
14. Sikandar, S. Kumar, and G.U.K. Maurya, "Optimizing delay for MAC in randomly distributed wireless sensor networks," in Intelligent Computing, Networking, and Informatics. Springer, 2014, pp. 609– 618.
15. C, Paschalidis and Y. Chen, "Anomaly detection in sensor networks based on large deviations of Markov chain models," in Proceedings of the 47th IEEE Conference on Decision and Control. IEEE, 2008, pp. 2338–2343.

16. R. Di Pietro, L. V. Mancini, C. Soriente, A. Spognardi, and G. Tsudik, "Data security in unattended wireless sensor networks," IEEE Transactions on Computers, vol. 58, no. 11, pp. 1500–1511, 2009.
17. D. Chen, Z. Zhang, F.-H. Tseng, H.-C. Chao, and L.-D. Chou, "A novel method defends against the path-based DOS for wireless sensor network," International Journal of Distributed Sensor Networks, vol. 2014, 2014.
18. R. Bellman, Dynamic programming. Princeton University Press, 1957.
19. M. Puterman, Markov decision processes: Discrete stochastic dynamic programming. Hoboken, NJ: Wiley, 1994.
20. F.D. Epenoux, "A probabilistic production and inventory problem," Management Science, vol. 10, pp. 98-108, October 1963.
21. E. Altman, Constrained Markov decision processes. Chapman and Hall/CRC Press, 1999.
22. W.B. Powell, Approximate dynamic programming: Solving the curses of dimensionality. Wiley-Interscience, 2011.
23. R.S. Sutton and A.G. Barto, Introduction to reinforcement learning. MIT Press Cambridge, 1998.
24. D.P.D. Farias and B.V. Roy, "The linear programming approach to approximate dynamic programming," Operations Research, vol. 51, pp. 850-865, December 2003.
25. O. Sigaud and O. Buffet, Markov decision processes in artificial intelligent: MDPs, beyond MDPs and applications. Wiley-IEEE Press, 2010.
26. S. Chobsri, W. Sumalai, and W. Usaha, "A parametric POMDP framework for efficient data acquisition in error prone wireless sensor networks," in Proceedings of the 4th International Symposium on Wireless Pervasive Computing. IEEE, 2009, pp. 1–5.
27. C. Pandana and K. R. Liu, "Near-optimal reinforcement learning framework for energy-aware sensor communications," IEEE Journal on Selected Areas in Communications, vol. 23, no. 4, pp. 788–797, 2005.

28. "An analytical approach towards cooperative relay scheduling under partial state information," in Proceedings of the 31st IEEE International Conference on Computer Communications. IEEE, 2012, pp. 2666–2670.
29. Aprem, C. R. Murthy, and N. B. Mehta, "Transmit power control policies for energy harvesting sensors with retransmissions," IEEE Journal of Selected Topics in Signal Processing, vol. 7, no. 5, pp. 895–906, 2013.
30. D.J. White, "A survey of applications of Markov decision processes," The Journal of the Operational Research Society, vol. 44, pp. 1073-1096, November 1993.
31. Markov decision processes with their applications. Springer, 2008.
32. M.L. Sichitiu and C. Veerarittiphan, "Simple, accurate time synchronization for wireless sensor networks," in Proceedings of the IEEE Wireless Communications and Networking, Vol. 2. IEEE, 2003, pp. 1266–1273.
33. E. Elson and D. Estrin, "Time synchronization in wireless sensor networks," Ph.D. dissertation, University of California, Los Angeles, 2003.
34. E.A. Feinberg and A. Shwartz, "Handbook of Markov decision processes: Methods and applications." Kluwer Academic Publishers, 2002.
35. Q. Hu and W. Yue, Markov decision processes with their applications.
36. Y. Yu and S. Mannor, "Online learning in Markov decision processes with arbitrarily changing rewards and transitions," in Proceedings of the International Conference on Game Theory for Networks, May 2009, pp. 314-322.
37. G. Neu, "Online learning in non-stationary Markov decision processes," Ph.D. dissertation, Budapest University of Technology and Economics, Hungary, 2013.