

Large Deviation, Basic Information Theory for Wireless Sensor Networks

Kwabena Doku-Amponsah

Department of Statistics, School of Physical and Mathematical Sciences, CBAS, UG, Legon, Ghana

Article history

Received: 19-02-2017

Revised: 21-06-2017

Accepted: 13-10-2017

E-mail: kdoku-amponsah@ug.edu.gh

Abstract: In this research paper, we establish Shannon-McMillan-Breiman Theorem for Wireless Sensor Networks modelled as Coloured Geometric Random Networks. For, large n we show that a Wireless Sensor Network consisting of n sensors in $[0; 1]^d$ linked by an expected number of edges of order $n \log n$ can be transmitted by approximately $[n(\log n)^2 \pi^{d/2}/(d/2)!] H$ bits, where H is an entropy defined explicitly from the parameters of the Coloured Geometric Random Network. In the process, we derive a joint Large Deviation Principle (LDP) for the empirical sensor measure and the empirical link measure of coloured random geometric network models.

Keywords: Shannon-McMillan-Breiman Theorem, Joint Large Deviation Principle, Coloured Geometric Random Graph, Empirical Sensor Measure, Empirical Link Measure, Wireless Sensor Networks, Sensor Law, Near Entropy, Relative Entropy Sensor Graph, Mathematics Subject Classification: 94A15, 94A24, 60F10, 05C80

Introduction

Background

A Wireless Sensor Network (WSN) is a network formed by a large number of sensor vertices where each vertex is assigned a sensor capable of monitoring and controlling the physical environment. WSNs are regarded as a revolutionary information gathering method to build the information or communication system which will greatly enhance the reliability and efficiency of infrastructure systems. Compared with the wired solution, WSNs are easier to set up and it has greater exibility of tools. With recent innovative ideas on sensors development, WSNs will become the key technology for the internet of things. See, example Yinbiao *et al.* (2014; Chong and Kumar, 2003).

Recent advances in computer science, information theory and micro-electromechanical technology have resulted in a significant focus on a new WSN research. The new WSN research which begun two decade ago has been drawing attention of many communication experts internationally. This sensor network research has focused on networking technology and networked information processing appropriate for highly dynamic ad hoc environments and resource-constrained sensor nodes. See, Yinbiao *et al.* (2014). Many researches who have

tried to find a good network model for the WSN have suggested models that have their background from classical areas of theoretical computer science and applied mathematics. See, example (Stanley-Marbell *et al.*, 2008) and the references therein.

Recent studies, Kenniche (2010) and the references therein, have shown that a random schemes is the only way to setup the outsized number of sensors in remote areas and the random geometric graph or geometric random network is the most appropriate model for this network.

In this article we derive Shannon-MacMillan-Breiman Theorem (SMBT) or Asymptotic Equipartition Property, Dembo and Kontoyiannis (2002) for WSN modelled as Coloured Geometric Random Network (CGRN) models, using some of the large deviation techniques developed for studying information theory, Doku-Amponsah (2011), for networked data structures.

To be specific we derive a Large Deviation Principle (LDP) for the empirical sensor distribution and the empirical link distribution of the CGRN using (Doku-Amponsah, 2011), Theorem 3.3] and the methods developed therein. From this LDP we prove the Weak Law of Large Numbers (WLLN) for the empirical sensor distribution and the empirical link distribution. From the WLLN we derive the SMBT for CGRN as a model for the WSN.

The Coloured Geometric Random Graph Model

We shall look at a more complex model of random geometric networks, the CGRN in which the linking radius depends on the sensor of the vertices. The empirical sensor distribution and the empirical link distribution are our main object to be studied here.

For a probability distribution ν on X and a function $r_n: X \times X \rightarrow (0, 1]$ we can define the coloured random geometric network X with n nodes as follows: Sample sites Y_1, \dots, Y_n at according to the uniform probability distribution on $[0, 1]^d$. For each site Y_j , we assign sensor $X(Y_j)$ independently according to the sensor law μ . For given sensors, we link any two nodes Y_i, Y_j ($i \neq j$) by a link independently of everything else, if:

$$\|Y_i - Y_j\| \leq r_n [X(Y_i), X(Y_j)]$$

In this research paper we shall look at $r_n(a, b)$, for $a, b \in X$ as a link radius and all the time consider:

$$X = \left((X(Y_i), X(Y_j)) : i, j = 1, 2, 3, \dots, n \right), E$$

under the joint law of network and sensor. We look at X as CGRN with nodes Y_1, \dots, Y_n sample from the nodes space $[0, 1]^d$. For the purposes of this study we only consider the sparse, intermediate and dense cases i.e., the link radius r_n satisfies the assumption $nr_n^d(a, b) / \log n \rightarrow \lambda_{[d]}(a, b)$ for all $a, b \in X$, where $\lambda: X^2 \rightarrow [0, \infty)$ is a symmetric function, which is not identically equal to zero. The CGRN has been mentioned by Cannings and Penman (2003) as a likely extension to the coloured random graphs proposed and studied by Penman (1998).

The distance r_n plays a role similar to that of p_n in the coloured random graph model proposed in Penman (1998) and studied by Doku-Amponsah (2006). Based on one's choice of r_n , qualitatively, different types of behaviour can be seen. Note that, intuitively, the average degree scales with nr^d : To be specific, it can be shown that in the classical random geometric graph the ratio of the average degree divided by nr^d tends to a constant in probability as long as $n^2 r_n^d \rightarrow \infty$. MCDiarmid and Müller (2005). As a result of the interpretation of nr_n^d a distribution of the average degree, we refer to the case where $nr^d / \log n \rightarrow \lambda_{[d]} = 0$ as sparse case, the case $nr^d / \log n \rightarrow \lambda_{[d]}$ as the intermediate case(s) and $nr^d / \log n \rightarrow \lambda_{[d]} = \infty$ as the dense case.

With any coloured graph X we associate a probability distribution, the empirical sensor distribution $L_X^1 \in B(X)$, by:

$$L_X^1(a) := \frac{1}{n} \sum_{j=1}^n \delta_{X(Y_j)}(a), \text{ for } a \in X$$

and a symmetric finite distribution, the empirical link distribution $L_X^2 \in \tilde{B}_*(X^2)$, by:

$$L_X^2(a, b) := \frac{1}{n \log n} \sum_{(i, j) \in E} \left[\delta_{(X(Y_i), X(Y_j))} + \delta_{(X(Y_j), X(Y_i))} \right](a, b), \text{ for } (a, b) \in X^2$$

The total mass $\|L_X^2\|$ of the empirical link distribution is $2|E|/n \log n$.

For any finite or countable set X we denote by $B(X)$ the space of probability distributions and by $\tilde{B}(X)$ the space of finite distributions on X , both endowed with the weak topology.

The remaining part of the paper is organized as follows: Section 2 contain statement of our main results, Theorem 2.1 and Theorem 2.2. In Section 3 we proof the main results, 2.2 and 2.1.

Statement of Main Results

Throughout the remaining part of this research paper we assume $d \geq 2$ is finite.

Asymptotic Equipartition Property the Sparse and Intermediate for WSN

The underlying question is, how many bits are needed to store or transmit the information contained in a Wireless Sensor Network consisting of n sensors connected by number of links? This question can be answered by the SMBT for Wireless Sensor Networks, see Theorem 2.1. To state the SMBT we denote by P the distribution of the CGRN. We define the distribution $\lambda_{[d]} \sigma \otimes \sigma \in \tilde{B}(X \times X)$ by:

$$\lambda_{[d]} \sigma \otimes \sigma(a, b) = \lambda_{[d]}(a, b) \sigma(a) \sigma(b), \text{ for } a, b \in X$$

and write:

$$\int_{X^2} \lambda_{[d]} \sigma \otimes \sigma(da, db) = \sum_{a, b \in X} \sigma(a) \lambda_{[d]}(a, b) \sigma(b)$$

Theorem 2.1

Let X be a CGRN with sensor law ν and connection radius $r_n: X^2 \rightarrow [0, 1]$ satisfying $nr_n^d(a, b) / \log n \rightarrow \lambda_{[d]}(a, b)$, for some symmetric function $\lambda_{[d]}: X^2 \rightarrow [0, \infty)$ not identical to zero. Then, for every $\varepsilon > 0$:

$$\lim_{n \rightarrow \infty} P \left\{ \left| \begin{aligned} &-\frac{1}{n(\log n)^2} \log P(X) \\ &-\frac{\pi^{d/2}}{2(d/2)!} \int_{X^2} \lambda_{[d]} v \otimes v(da, db) \end{aligned} \right| \geq \varepsilon \right\} = 0$$

In other words, in order to code a WSN in the given sparse or intermediate regime one needs with high probability, about $\lceil n(\log n)^2 \pi^{d/2} / (d/2)! \rceil H$ bits, where H is the entropy defined by:

$$H := \frac{1}{2 \log^2} \int_{X^2} \lambda_{[d]} v \otimes v(da, db)$$

For $d \in N$ we write:

$$\rho(d) := \begin{cases} \frac{\pi^{d/2}}{\Gamma\left(\frac{d+2}{2}\right)} & \text{if } d \geq 2 \\ 1 & \text{if } d = 1 \end{cases}$$

where, Γ is the gamma function. For the pair of distributions (σ, η) we define the near entropy (Kullback action) $\mathfrak{H}_{\lambda_{[d]}}$ by:

$$\begin{aligned} \mathfrak{H}_{\lambda_{[d]}}(\eta \| \sigma) &:= H(\eta \| \rho(d) \lambda_{[d]} \sigma \otimes \sigma) \\ &+ \rho(d) \| \lambda_{[d]} \sigma \otimes \sigma - \eta \| \end{aligned}$$

where, $H(\eta \| \tilde{\eta})$ means the relative entropy of the finite distribution η with respect to $\tilde{\eta}$.

Large-Deviation Principles in the Sparse and Intermediate CGRN

The following LDP is a key ingredient in the proof of our SMBT, see Theorem 2.1.

Theorem 2.2

Let X be a CGRN with sensor law $v: X \rightarrow (0, 1]$ and link radius $r_n: X \times X \rightarrow [0, 1]$ satisfying $nr_n^d(a, b) / \log n \rightarrow \lambda_{[d]}(a, b)$, with $\lambda_{[d]}: X \times X \rightarrow [0, \infty)$ symmetric. Then, for $n \rightarrow \infty$, the pair (L_X^1, L_X^2) satisfies an LDP in $B(X) \times \tilde{B}_*(X \times X)$ with speed:

- $n \log n$ and good rate function:

$$I_1^{[d]}(\sigma, \eta) = \frac{1}{2} \mathfrak{H}_{\lambda_{[d]}}(\eta \| \sigma) \tag{2.1}$$

whilst $\mathfrak{H}_{\lambda_{[d]}}(\eta \| \sigma) \geq 0$ and equality holds if and only if $\eta = \rho(d) \lambda_{[d]} \sigma \otimes \sigma$.

- n and good rate function”

$$I_1^{[d]}(\sigma, \eta) = \begin{cases} H(\sigma \| v) & \text{if } \eta = \rho(d) \lambda_{[d]} \sigma \otimes \sigma \\ \infty & \text{otherwise} \end{cases} \tag{2.2}$$

Derivation of Theorems 2.2 and 2.1

For any two points U_1 and U_2 uniformly and independently chosen from the space $[0; 1]^d$ write:

$$F(t) := P\{\|U_1 - U_2\| \leq t\}$$

where, $F(r_n(a, b)) = \rho(d) r_n^d(a, b)$, $a, b \in X^2$ i.e., the volume of a d -dimensional (hyper) sphere with radius $r(a, b)$ satisfying $r_n^d(a, b) / \log n \rightarrow \lambda_{[d]}(a, b)$. Let $p_n(a, b) = F(r_n(a, b)) = \rho(d) r_n^d(a, b)$ and:

$$C(a, b) = \rho(d) \lambda_{[d]}(a, b)$$

Then we have:

$$\begin{aligned} dP(X) &= \prod_{u \in V} v(X(Y_u)) \prod_{(u, v) \in E} F(r_n(X(Y_u), X(Y_v))) \\ &\prod_{(u, v) \notin E} (1 - F(r_n(X(Y_u), X(Y_v)))) \end{aligned} \tag{3.1}$$

$$\begin{aligned} &= \prod_{u \in V} v(X(Y_u)) \prod_{(u, v) \in E} p_n(X(Y_u), X(Y_v)) \\ &\prod_{(u, v) \notin E} (1 - p_n(X(Y_u), X(Y_v))) = d\tilde{P}(X) \end{aligned} \tag{3.2}$$

where, $\tilde{P}(X)$ is the law of coloured random graph X with the geometric plane $[0; 1]^d$ ignored.

Hence by the exponential equivalence, [(Dembo and Zeitouni, 1998), Theorem 4.2.13] and [(Doku-Amponsah, 2011), Theorem 3.3] we have Theorem 2.1 with rate functions $I_1^{[d]}$ and $I_2^{[d]}$.

Derivation of Theorems 2.1

Lemma 3.1

Let X be a CGRN with sensor law $v: X \rightarrow (0, 1]$ and link radius $r_n: X \times X \rightarrow [0, 1]$ such that $nr_n^d(a, b) / \log n \rightarrow \lambda_{[d]}(a, b)$, for $\lambda_{[d]}: X \times X \rightarrow [0, \infty)$ nonzero. Then, for any $\varepsilon > 0$ we have:

$$\lim_{n \rightarrow \infty} P \left\{ \sup_{a \in X} |L_X^1(a) - v(a)| \geq \varepsilon \right\} = 0$$

and:

$$\lim_{n \rightarrow \infty} P \left\{ \sup_{a \in X} |L_X^2(a, b) - v(a)\rho(d)\lambda_{[d]}(a, b)v(b)| \geq \varepsilon \right\} = 0$$

From Theorem 2.2(ii), we prove this lemma. To begin, we define a closed set:

$$F_1 = \left\{ (\sigma, \eta) \in B(X) \times \tilde{B}_\varepsilon(X \times X) : \sup_{a, b \in X} |\eta(a, b) - v(a)\rho(d)\lambda_{[d]}(a, b)v(b)| \geq \varepsilon \right\}$$

and:

$$F_2 = \left\{ (\sigma, \eta) \in B(X) \times \tilde{B}_\varepsilon(X \times X) : \sup_{a, b \in X} |\sigma(a) - v(a)| \geq \varepsilon \right\}$$

We observe that, by Theorem 2.2(ii):

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log P \{ (\mathcal{L}_X^1, \mathcal{L}_X^2) \in F \} \leq - \inf_{(\sigma, \eta) \in \mathcal{F}} I_2^{[d]}(\sigma, \eta)$$

where, $F = F_1 \cup F_2$.

This will be shown by contradiction that the right of (3.3) is negative. We assume there is a sequence (σ_n, η_n) in F such that $I_2^{[d]}(\sigma_n, \eta_n) \downarrow 0$. Then, because the function $I_2^{[d]}$ is a good rate function, we have all its level sets compact. Notice that, by lower semi-continuity of the mapping $(\sigma, \eta) \mapsto I_2^{[d]}(\sigma, \eta)$, there is a limit point $(\sigma, \eta) \in F$ with $I_2^{[d]}(\sigma, \eta) = 0$. Using Theorem 2.2(i), we have $H(\sigma||v) = 0$ and $\mathfrak{H}_C(\eta||\sigma) = 0$. This implies $\sigma(a) = v(a)$ and $\eta(a, b) = \rho(d)\lambda_{[d]}(a, b)v(a)v(b)$, for $a, b \in X$ which contradicts $(\sigma, \eta) \in F$. Hence as required.

Recall that $[n]$ is a fixed set of n vertices, say $[n] = \{1, \dots, n\}$, G_n is the set of all (simple) graphs with vertex set $[n]$ and edge set $E \subset \mathcal{E} := \{(u, v) \in [n] \times [n] : u < v\}$. Now we calculate the law $P_n: G_n(X) \rightarrow [0, 1]$ of X :

$$\begin{aligned} P(x) &= \prod_{u \in V} v(x(y_u)) \prod_{(u, v) \in E} F(r_n(x(y_u), x(y_v))) \\ &\quad \prod_{(u, v) \in E} (1 - F(r_n(x(y_u), x(y_v)))) \\ &= \prod_{u \in V} v(x(y_u)) \prod_{(u, v) \in E} \frac{F(r_n(x(y_u), x(y_v)))}{1 - F(r_n(x(y_u), x(y_v)))} \\ &\quad \prod_{(u, v) \in E} (1 - F(r_n(x(y_u), x(y_v)))) \end{aligned}$$

Therefore, we have in the case of Theorem 2.1:

$$\begin{aligned} -\frac{1}{n(\log(n))^2} \log P(x) &= -\int_X \frac{\log v(a)}{(\log n)^2} \mathcal{L}_X^1(da) \\ &\quad -\frac{1}{2} \int_{X^2} \frac{\log(F(r_n(a, b))/(1 - F(r_n(a, b))))}{\log n} \mathcal{L}_X^2(da, db) \\ &\quad -\frac{1}{2} \int_{X^2} \frac{\log(1 - (F(r_n(a, b))))}{(\log)^2/n} \mathcal{L}_X^1 \otimes \mathcal{L}_X^1(da, db) \\ &\quad -\frac{1}{2} \int_{X^2} \frac{\log(1 - (F(r_n(1, b))))}{(\log n)^2} \mathcal{L}_\Delta^1(da, da) \end{aligned}$$

Now in the first case the integrands $\frac{-\log v(a)}{(\log n)^2}$,

$$\frac{-\log(1 - (F(r_n(a, b))))}{(\log n)/n} \quad \text{and} \quad \frac{-\log(1 - (F(r_n(1, b))))}{(\log n)^2} \quad \text{all}$$

approach zero, while $\frac{-\log(F(r_n(a, b)))/(1 - F(r_n(a, b)))}{\log n}$

$\rightarrow 1$, for all $a, b \in X$. Hence Theorem 2.1 follows from Theorem 3.1.

Conclusion

In Information theory, the Shannon-McMillan-Breiman Theorem (SMBT) or Asymptotic Equipartition Property is the analog of the strong law of large numbers. See, Cover and Thomas (1991). We have found in this research paper an SMBT, see Theorem 2.1, for the Wireless Sensor Networks. It is a direct consequence of the weak law of large numbers, see Lemma 3.1, for the empirical sensor distribution and empirical link distribution of a Wireless Sensor Networked data structure. The SMBT will allow us to partition output sequence of a Wireless Sensor Networked data structure into two sets, the typical set, where the sample entropy is close to the true entropy and the non-typical set, which contains the other sequence. Further coding theorems could be developed for the Wireless Sensor Network from this SMBT.

Conflict of Interest

The author declares that he has no conflict of interest.

Acknowledgement

The author thanks the reviewers of the paper for their suggestions which have helped improve this paper.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and there are no ethical issues involved.

Reference

- Cannings, C. and D.B. Penman, 2003. Models of Random Graphs and their Applications Hand-Book of Statistics 21. In: Stochastic Processes: Modeling and Simulation, Shanbhag, D.N. and C.R. Rao (Eds.), Elsevier, pp: 51-91.
- Chong, C.Y. and S.P. Kumar, 2003. Sensor networks: Evolution, opportunities and challenges. Proc. IEEE, 91: 1247-1256. DOI: 10.1109/JPROC.2003.814918
- Cover, T.M. and J.A. Thomas, 1991. Elements of Information Theory. 99th Edn., Wiley, New York, ISBN-10: 0471062596, pp: 542.
- Dembo, A. and L. Kontoyiannis, 2002. Source coding, large deviations and approximate pattern matching. IEEE Trans. Inform. Theory, 48: 1590-1615. DOI: 10.1109/TIT.2002.1003841
- Dembo, A. and O. Zeitouni, 1998. Large Deviations Techniques and Applications. 1st Edn., Springer, New York, ISBN-10: 0387984062, pp: 396.
- Doku-Amponsah, K., 2006. Large deviations and basic information theory for hierarchical and networked data structures. PhD Thesis, Bath.
- Doku-Amponsah, K., 2011. Asymptotic equipartition properties for simple hierarchical and networked structures. ESAIM: PS, 16: 114-138. DOI: 10.1051/ps/2010016
- Kenniche, H., 2010. Random geometric graphs as model of wireless sensor networks. Proceedings of the 2nd International Conference on Computer and Automation Engineering, Feb. 26-28, IEEE Xplore Press, Singapore, pp: 103-107. DOI: 10.1109/ICCAE.2010.5451758
- MCDiarmid, C.J.H. and T. Müller, 2005. Colouring random geometric graphs DMTCS Proceedings of the European Conference on Combinatorics, Graph Theory and Applications, (GTA' 05), pp: 1-4.
- Penman, D.B., 1998. Random graphs with correlation structure. PhD Thesis, Sheffield.
- Stanley-Marbell, P., T. Basten, J. Rousselot, R.S. Olive and H. Karl *et al.*, 2008. Decotignie system models in wireless sensor network. ES Reports, Technische Universiteit Eindhoven.
- Yinbiao, S., K. Lee, P. Lanctot, F. Jianbin and H. Hao *et al.*, 2014. Internet of things: Wireless sensor networks. IEC.