

COUNTING THE NUMBER OF NODES IN OCEAN WIRELESS COMMUNICATION NETWORK

Abu Sadat Md. SAYEM¹ Md. Shamim ANOWER²

^{1,2}Department of Electrical & Electronic Engineering, Rajshahi University of Engineering & Technology, Bangladesh.
¹sayem_ruet_eee@yahoo.com, ²md.shamimanower@yahoo.com

Abstract: It is often essential to estimate the number of operating nodes in a wireless communication network (WCN), in which the nodes are deployed in different forms to cover small or large areas for a wide range of personal, scientific and commercial applications. The number of nodes may vary due to ad-hoc nature, power failure of nodes, or environmental disaster. In any communication network, it is important to estimate the number of nodes at any point in time for proper network operation and maintenance. Counting the number is very important for data collection, network maintenance and node localization. Also network performance depends on the area node ratio i.e. the number of operating nodes per unit area. Many estimation techniques are used to count the number of nodes in wireless communication networks, but in underwater environment they are not efficient. In this paper a cross-correlation based statistical signal processing approach for node estimation in underwater wireless communication network is proposed. In this method nodes are considered as acoustic signal sources and their number is calculated through the cross-correlation of the acoustic signals received at two sensors placed in the network. The mean of the cross-correlation function is used as the estimation parameter in this process. Theoretical and simulation results are provided which justify the effectiveness of the proposed approach.

Key words: Wireless communication network (WCN), Cross-correlation function (CCF), Estimation parameter, Mean of cross-correlation function, Node estimation, Underwater acoustic sensor networks (UASN).

1. Introduction

In underwater wireless sensor networks nodes are deployed for a variety of applications. These applications range from research to security purposes such as climatic data collection, pollution monitoring, environmental monitoring, seismic and acoustic monitoring to surveillance and national security, military and health care, discovering natural resources as well as locating man-made artefacts or extracting information for scientific analysis. In any type of underwater sensor network, proper network operation depends on the number of active nodes. So, estimating the number of active nodes is an important

issue in any sensor network. Moreover, maintenance and localization activities need exact estimation of nodes. So, the number of operating nodes is a very important factor in any network.

However, the number of operating nodes can vary with time due to various artificial as well as natural reasons (for example, some nodes might fail, some could be damaged, or batteries might fail). So, it is a matter of great interest for a communication network to know how many operating nodes or transmitters are available in the region at any point in time to ensure proper network operation as well as network maintenance (such as replacement of faulty nodes). At present there have been many node estimation techniques. For example, protocols [1-8] have been used to estimate the number of tag IDs in radio frequency identification (RFID) systems, which is a similar problem to the estimation of the number of nodes in wireless communication networks. Similarly, a Good-Turing estimator of node estimation for terrestrial sensor networks has been proposed in Budianu et al. [9-11], where each transmitting node transmits its ID in every slot according to a certain probability and the packet collection can be modeled as an i.i.d. (independent, identically distributed) sampling with uniform distribution by terrestrial sensor network with mobile access (SENMA) protocol (an ALOHA-like protocol). In this method they estimate the number of operating sensors by deriving an expression for it as a function of missing mass.

All of the abovementioned techniques are only effective for communication friendly networks such as RFID as well as terrestrial systems; these techniques do not take into account the capture effect, so they are difficult to apply in UASN. Later Howlader et al. [12, 13] proposed a node estimation technique taking the capture effect into account. The procedure is similar to probabilistic framed slotted

ALOHA [1]. But, it also suffers from long propagation delays, high path loss in underwater acoustic network.

So, it can be said that node estimation in underwater sensor network is not so easy as other type of networks. The existing procedures for the estimation of the number of nodes are not effective in underwater network due to underwater propagation characteristics [14] such as propagation delay, high absorption, and dispersion. These conventional protocol-based techniques are expensive, inefficient and time consuming for underwater network. Moreover, limited battery power, limited bandwidth, long and variable propagation delays, multi-path and fading problems and high bit error rates make the conventional techniques inefficient for underwater network.

In this paper, a simple novel estimation technique is proposed which is based on the cross-correlation of the acoustic signals received at two sensors in the network. The cross-correlation of signals received from random signal sources at each end of the channel is an important issue in current research and this technique has been investigated by the researchers [15-18]. However, none of these investigations has focused on the estimation of the number of signal sources.

In the proposed approach transmitted signals from a number of different random signal sources (nodes) within range are received by two sensors separated by a certain distance in the region; the received signals are summed at each of the two sensor locations, and these two signals are then cross-correlated. The estimation of the number of signal sources (assumed in our case the number of nodes in an underwater network) can be obtained based on the mean of the CCF. The process of cross-correlation might be affected by the sampling rate and the distance between sensors and the effects are investigated in this paper.

2. Formation of Cross-Correlation Function

Consider two receiving nodes surrounded by N transmitting nodes in a 3D space, as shown in Fig. 1. Assuming that the transmitting nodes are the sources of white Gaussian signals and are uniformly distributed over the volume of a large sphere inside a cube, the center of the sphere lays half way between the sensors, because only a sphere provides equal amounts of signals from every direction. The propagation velocity is constant, which in our case, the sound velocity Sp , in the medium.

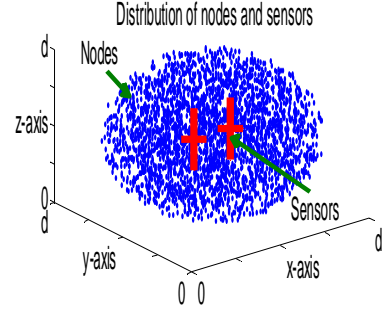


Fig. 1: Distribution of underwater nodes with N transmitting nodes.

Now, getting probe request, a node emits a very long Gaussian signal, which is recorded by the sensors with corresponding time delays. The signals in the sensors are cross-correlated, which takes the form of a delta function as it is a cross-correlation of two white Gaussian signals where one signal essentially is a delayed copy of the other. The position of this delta in the CCF will be the distance equal to the delay difference of the signals from the center of the CCF where the position is called a bin in this paper. This holds for all nodes and the formation of CCF for N number of nodes can be expressed as follows [19]:

If the transmitted signals from the nodes are denoted as $S_1(t), S_2(t), \dots, S_N(t)$ respectively, the corresponding delays to reach at sensor 1 are denoted as $\tau_{11}, \tau_{21}, \dots, \tau_{N1}$, and the corresponding attenuations are as $\alpha_{11}, \alpha_{21}, \dots, \alpha_{N1}$, the composite signal at sensor 1 can be expressed as

$$S_{r_1}(t) = \sum_{k=1}^N \alpha_{k1} S_k(t - \tau_{k1}) \quad (1)$$

Similarly, if the transmitted signals from the nodes are denoted as $S_1(t), S_2(t), \dots, S_N(t)$ respectively, the corresponding delays to reach at sensor 2 are denoted as $\tau_{12}, \tau_{22}, \dots, \tau_{N2}$, and the corresponding attenuations are as $\alpha_{12}, \alpha_{22}, \dots, \alpha_{N2}$, the composite signal at sensor 2 can be expressed as

$$S_{r_2}(t) = \sum_{k=1}^N \alpha_{k2} S_k(t - \tau_{k2}) \quad (2)$$

Assuming $\tau = d_{DBS}/Sp$ is the time shift in cross-correlation, then the CCF is

$$\begin{aligned}
C(\tau) &= \int_{-\infty}^{+\infty} S_{r_{i1}}(t) S_{r_{i2}}(t - \tau) d\tau \\
&= \int_{-\infty}^{+\infty} \sum_{k=1}^N a_{k1} S_k(t - \tau_{k1}) \sum_{k=1}^N a_{k2} S_k(t - \tau_{k2} - \tau) d\tau
\end{aligned} \tag{3}$$

which takes the form of a series of delta functions as it is a cross-correlation of two composite signals which are the summation of several white Gaussian signals.

One such obtained CCF with N ($=10000$) nodes is shown in Fig. 2.

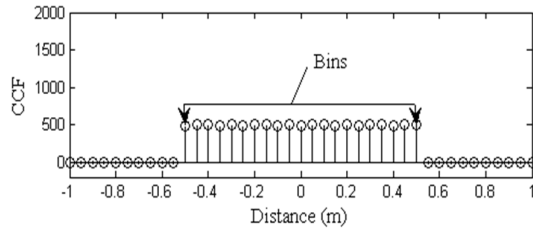


Fig. 2 : Bins in the cross-correlation process positions are in meters from the CCF center

3. The mean of the CCF

The mean of CCF is expressed by ensemble average of the signal cross-correlation in [15] as

$$\begin{aligned}
\langle C(t) \rangle &= Q_T T_r v \int_{-\infty}^{+\infty} d\vec{r}_s \frac{1}{|\vec{r}_b - \vec{r}_s| |\vec{r}_a - \vec{r}_s|} \\
&\times \delta\left(t + \frac{|\vec{r}_a - \vec{r}_s|}{S_p} - \frac{|\vec{r}_b - \vec{r}_s|}{S_p}\right).
\end{aligned} \tag{4}$$

where Q_T represents the acoustic power of the received signals from the nodes taken to be constant over time and space, and v the creation rate of the random nodes whose unit is unit time per unit volume, T_r total recording time, \vec{r}_s path length of node s from the origin, \vec{r}_a path length of first receiver from the origin, and \vec{r}_b the path length of second receiver from the origin.

4. Approach to binomial probability distribution

In the proposed estimation technique a simple analysis will show that the proposed cross-correlation

technique can be reframed to a probability problem using the well-known occupancy problem which follows the binomial probability distribution from which a parameter will be chosen to estimate the number of nodes of a network. Considering each delta function as a ball which occupies a bin according to the delay difference of corresponding recorded signals in the sensors, it is simple to model this cross-correlation problem as a probability problem based on the well-known occupancy problem, i.e., the problem of placing N balls in b bins. It is known from [20] that the occupancy problem follows the binomial probability distribution in which the parameters are the number of balls i.e. nodes, N , and the inverse of the number of bins, b .

Occupancy problems deal with the pairings of objects and have a wide range of applications in different fields containing probabilistic and statistical properties. The basic occupancy problem is about placing m balls into b bins [21]. If one threw some balls randomly towards several bins, the bins would be randomly filled by the balls, resulting in some bins being occupied by more than one ball, some by one while some may have none. In this work, the cross-correlation process for node estimation is reframed as this occupancy problem. It describes the reframing process as follows:

- In this process to obtain a CCF, N nodes create N number of delta functions which occupy the place in the correlation length where the length is divided by b number of bins as shown in Fig. 2.
- Some bins are empty i.e. not occupied by any delta function; some are occupied by only one and others are more than one.

Moreover, the formation of cross-correlation function to obtain node estimation satisfies the characteristics of binomial distribution as the number of trials i.e. the number of nodes is fixed, trials are independent in the sense that the nodes are sending independent Gaussian signal, there exist only two possible outcomes, success or failure, for every trial which indicates that delta for a particular node is occupying a bin or not, each trial has the same probability of success, p which is equivalent to $1/b$, where b is the number of bins. As the cross-correlation function follows the binomial distribution, its mean is easy to obtain from the theory of binomial distribution which is discussed in the following section.

5. Estimation of the number of nodes, N

It is discussed in the previous section that the cross-correlation function follows the binomial probability distribution where the parameters are the number of balls i.e. nodes, N , and the number of bins, b . Then the expected value, i.e. the mean, m of the CCF is defined as:

$$m = \frac{N}{b} \quad (5)$$

where b is the number of bins in the cross-correlation process and is obtained from the experimental setup with sampling rate, S_R , distance between sensors, d_{DBS} , and speed of propagation, S_p as [19]:

$$b = \frac{2 \times d_{DBS} \times S_R}{S_p} - 1 \quad (6)$$

Thus the estimation of N is obtained from equation (1) as:

$$N = b \times m \quad (7)$$

This is the relationship between the number of nodes, N , and the mean, m , of the CCF. Since the number of bins, b is known and m can be measured from the CCF, the number of nodes, N can be readily determined.

6. Results and discussion

Both theoretical and simulation results of the estimation of the number of nodes using this novel signal processing approach using cross-correlation are provided in Fig. 3 to Fig. 5. Simulations have been performed in the Matlab programming environment. Fig. 3 to Fig. 5 show the theoretical and corresponding simulated results for the estimation of the number of nodes in a network in terms of the estimation parameter m of CCF, which show that the simulations match the theory properly and is the indication of effectiveness of the process. The solid lines indicate the theoretical results and the circles the corresponding simulated results. The variations of b in the three different Figs. are as a result of varying d_{DBS} (considering sampling rate and propagation speed constant). The distances between the sensors are: 0.0625m in Fig. 3, 0.1875m in Fig. 4 and 0.4m in Fig. 5. The other parameters are radius of the sphere is 2000m, $N=1, 10, 20, \dots, 100$, signal length is 10^6 samples, signal propagation speed is 1500m/s, and sampling rate $S_R = 180$ kSa/s.

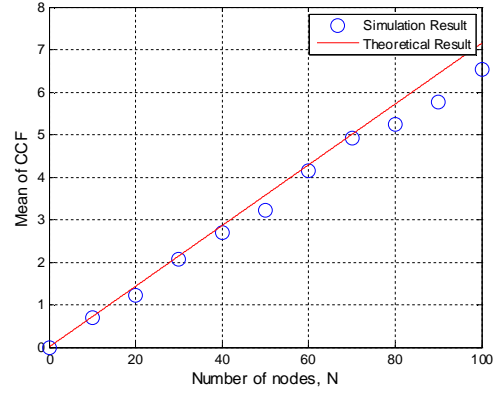


Fig. 3. Mean of CCF versus number of nodes, N for $b = 14$, $d_{DBS} = 0.0625$ m

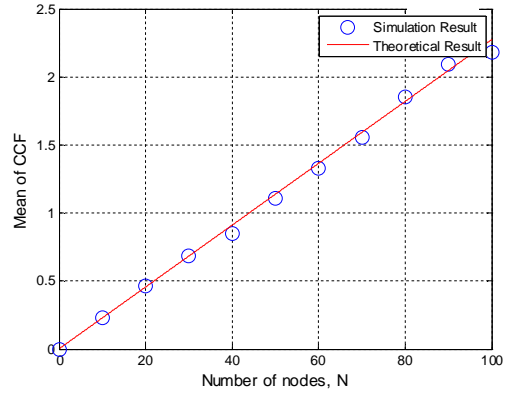


Fig. 4. Mean of CCF versus number of nodes, N for $b = 44$, $d_{DBS} = 0.1875$ m

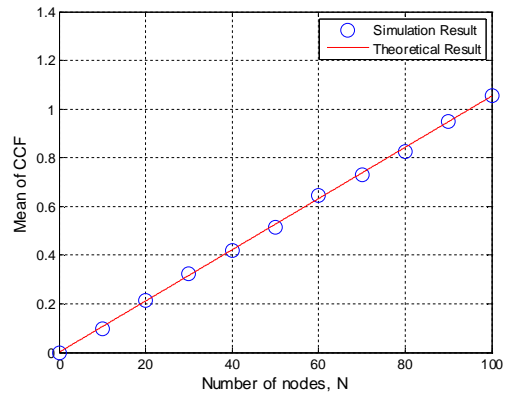


Fig. 5. Mean of CCF versus number of nodes, N for $b = 95$, $d_{DBS} = 0.4$ m

The above mentioned results show that the simulated and corresponding theoretical results are very close to each other, which indicates that the process is effectively applicable for estimation. At the same time, it is clear that the number of bins, b has an effect on the estimation parameter, which is depicted in the estimation expression equation (7). It can be seen that the value of the estimation parameter is lower in case of higher b and vice-versa and the simulated lines are more closer with the theoretical lines. It is also obvious from the results that a good approximation of the number of nodes, N , can be obtained for any dense network. The shown results are for 100 nodes and for more number of nodes estimation is possible in the same way.

Now, another approach will be taken, the sampling rate will be increased and the process will be repeated. A comparison will be observed for the proposed estimation process for the same number of bins as before.

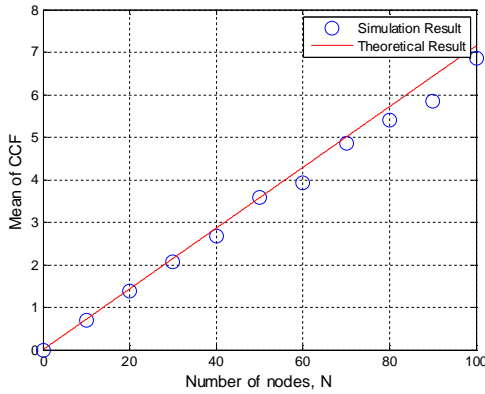


Fig. 6. Mean of CCF versus number of nodes, N for $b = 14$, $d_{DBS} = 0.0313$, $S_R = 360kSa/s$

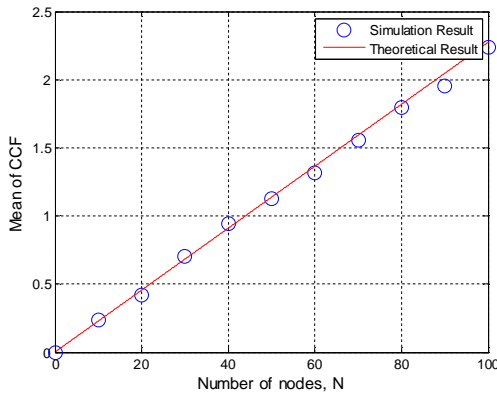


Fig. 7. Mean of CCF versus number of nodes, N for $b = 44$, $d_{DBS} = 0.0938m$, $S_R = 360kSa/s$

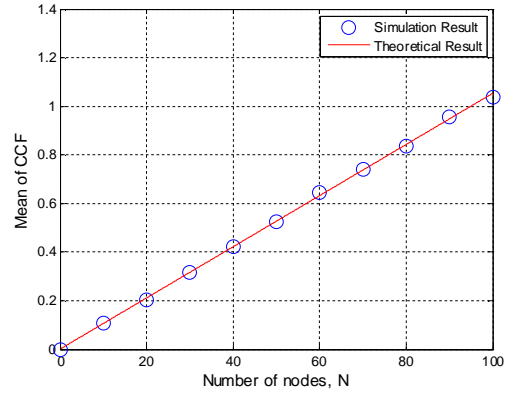


Fig. 8. Mean of CCF versus number of nodes, N for $b = 95$, $d_{DBS} = 0.2m$, $S_R = 360kSa/s$

From Fig. 6 to Fig. 8 it can be observed that improvement in result occurs with the increased in number of bins as before.

In conclusion, it can be said that better result occurs for the increased number of bins. So, the number of bin has a great effect on the result. Also, it is observed that the result depends on the variation of the number of bins but independent on how bin number is varied. Bin number can be varied by varying distance between sensors, d_{DBS} ; sampling rate, S_R and propagation speed, S_P as expressed in equation (6).

Now, the results of estimated number of nodes, N (estimated) with respect to exact number of nodes will be shown.

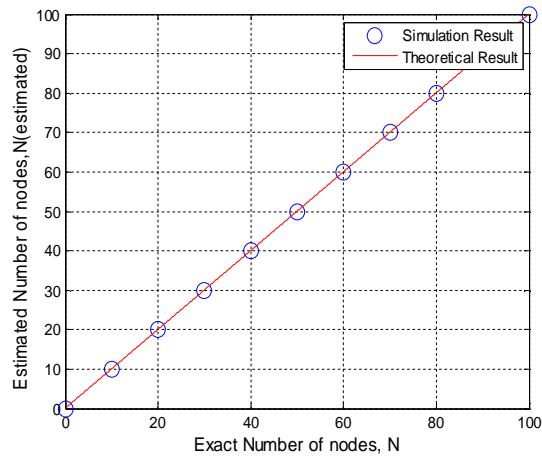


Fig. 9. Comparison of theoretical and simulated number of estimated nodes

Fig. 9 shows the comparison of theoretical and simulated number of estimated nodes (for bin number 119). In this Fig., the solid line indicates the theoretical result and the circles the corresponding simulated results. From Fig. 9, it can be seen that, the theoretical and simulated results are very close to each other, which signify the validity of the proposed approach.

7. Analysis of error in estimation

As the proposed cross-correlation based approach is a statistical technique, the statistical error, the coefficient of variation (CV), is used as its error in estimation in order to fully assess the accuracy of the proposed estimation technique. To obtain a simulated CV of estimation, a simulation process is run 1000 times for a particular N and b . From these 1000 values of estimated N , the standard deviation and mean of estimation and, thus, the corresponding CV, is obtained. In this case firstly, the mean, m , of the CCF from 100 iterations, and then the estimated \hat{N} using the expression of N related to this m , are obtained. Secondly, to obtain the CV, the same process is continued 1000 times without any change in parameters and the values of all estimated \hat{N} are recorded. Finally, the CV for one iteration is obtained from the ratios of the standard deviation to the mean of those values as [22]:

$$CV = \frac{\sigma(\hat{N})}{\mu(\hat{N})}$$

Now, after u iteration, the standard deviation and, thus, the CV, are reduced to $\frac{1}{\sqrt{u}}$ so that the CV after the u^{th} iteration is

$$CV = \frac{1}{\sqrt{u}} \left(\frac{\sigma(\hat{N})}{\mu(\hat{N})} \right)$$

Now, CV of the proposed estimation technique will be calculated and the result will be compared with the previous cross-correlation based method proposed by Anower et al. [19] which is also a signal processing approach. In the comparison it is considered a very long fixed signal length, N_s , of 158093 samples, Sampling rate 390000 HZ, signal propagation speed 1500m/s, bin number 119 and $d_{\text{DBS}}=0.25\text{m}$.

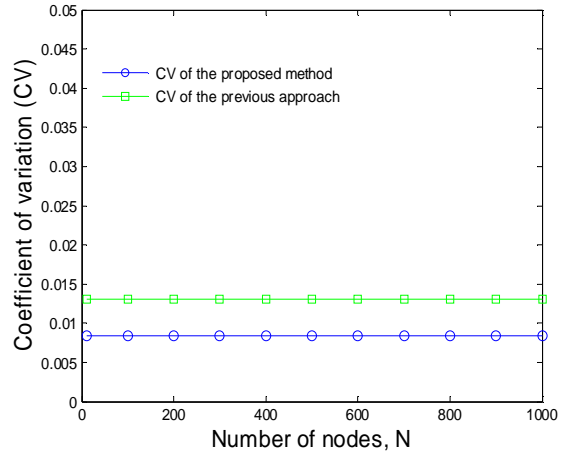


Fig. 10. Comparison of CV of the proposed method with previous technique.

From the above result it can be observed that the proposed estimation technique gives better accuracy than previous technique. The case behind this better accuracy can be analyzed by looking at the theory regarding error estimation. Equation (7) gives $N = b \times m$. So, $CV(N) = CV(b \times m)$

$$= CV(m) \text{ as } b \text{ is constant.}$$

and $CV(m)$ is inversely proportional to b [22]. Here, the expression $CV(N) = CV(m)$ is depicted directly but in previous method [19] this case was assumed, it is the main reason behind the better accuracy of the proposed estimation technique.

8. Conclusion

It can be concluded that this signal processing approach would be an effective alternate of the conventional protocols to estimate the number of nodes in any type and size of underwater networks. Previously used techniques are inefficient, complex and time consuming in harsh underwater environment. The proposed method is effectively applicable to any type of dense underwater network. Error in estimation of the number of nodes is investigated and it has been observed that the accuracy of the proposed approach shows satisfactory performance. Thus the proposed cross-correlation based technique might be the useful alternate of the existing techniques.

References

1. Kodialam, M. and Nandagopal, T.: *Fast and Reliable Estimation Schemes In RFID Systems*. In: Proc. Mobicom '06, September 23-26, 2006, Los Angeles, California, USA.
2. Bonuccelli, M.A., Lonetti, F. and Martelli, F.: *Tree slotted aloha: a new protocol for tag identification in RFID networks*. In: Proceedings of IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (2006), pp.603–608, Jun. 2006.
3. M. A.-I. Center, Draft protocol specification for a 900 MHz class 0 radio frequency identification tag, <http://www.epcglobalinc.org>, February 2003.
4. Law, C., Lee, K. and Siu, K.Y.: *Efficient memoryless protocol for tag identification (extended abstract)*. In: Proceedings of 4th International workshop on Discrete Algorithms and methods for mobile computing and communications (DIALM '00), New York, USA, 2000, pp. 75–84.
5. Ryu, J., Lee, H., Seok, Y., Kwon, T. and Choi, Y.: *A Hybrid Query Tree Protocol for Tag Collision Arbitration in RFID systems*. In: Proceedings of IEEE International Conference on Communications'07, pp. 5981–5986, Jun. 2007.
6. Myung, J., Lee, W. and Srivastava, J.: *Adaptive binary splitting for efficient RFID tag anti-collision*. In: IEEE Communications Letters, vol. 10 (3), pp.144–146, Mar. 2006.
7. Myung, J., Lee, W., Srivastava, J. and Shih, T.K.: *Tag-Splitting: adaptive collision arbitration protocols for RFID tag identification*. In: IEEE T. Parall. Distr., Vol. 18 (6), pp. 763-775, Jun. 2007.
8. Bonuccelli, M.A., Lonetti, F., Martelli, F.: *Perfect tag identification protocol in RFID networks*. In: http://arxiv.org/PS_cache/arxiv/pdf/0805/0805.1877v1.pdf, May 13, 2008.
9. Budianu, C., Ben-David, S. and Tong, L.: *Estimation of the Number of Operating Sensors in Large-Scale Sensor Networks with Mobile Access*. In: IEEE Transactions on Signal Processing, Vol. 54, No. 5, May 2006.
10. Budianu, C., and Tong, L.: *Estimation of the number of operating sensors in a sensor network*. In: Presented at 2003 Asilomar Conf. Signals, Systems, Computers. <http://acsp.ece.cornell.edu/pubC.html/>
11. Budianu, C. and Tong, L.: *Good–Turing estimation of the number of operating sensors: A large deviations analysis*. In: International Conference Acoustics, Speech, Signal Processing (ICASSP), Montreal, QC, Canada, May 2004.
12. Howlader, M.S.A., Frater, M.R. and Ryan, M. J. : *Estimating the number of neighbours and their distribution in an underwater communication network (UWCN)*. In: MILCIS2007, Canberra, 20–22 November 2007.
13. Howlader, M.S.A., Frater, M.R. and Ryan, M.J.: *Estimation in Underwater Sensor networks taking into account capture*. In: IEEE Oceans'07, Aberdeen, Scotland, 18-21 June 2007.
14. Liu, L., Zhou, S. and Jun-Hong C.: *Prospects and problems of wireless communication for underwater sensor networks*. In: Wireless Communication Mobile Computer 2008, Published online in Wiley InterScience. DOI= <http://doi.wiley.com/10.1002/wcm.654>
15. Roux, P., Sabra, K., Kuperman, W. and Roux, A.: *Ambient noise cross correlation in free space: theoretical approach*. In: J. Acoustic Soc. Am., 117, 79–84, 2005.
16. Snieder, R.: *Extracting the Green's function of attenuating heterogeneous acoustic media from uncorrelated waves*. In: J. Acoustic Soc. Am., Vol. 121, No. 5, 2637-2643, May 2007.
17. Godin, O. A.: *Recovering the Acoustic Green's function from ambient noise cross correlation in an inhomogeneous moving medium*. In: Physical Review Letters, The American Physical Society, 97, 054301(2006).
18. Sabra, K.G., Roux, P. and Kuperman, W. A.: *Emergence rate of the time-domain Green's function from the ambient noise cross-correlation function*. In: J. Acoustic Soc. Am., Vol. 118, No. 6, 3524–3531, December 2005.
19. Anower, M.S., Motin, M.A., Sayem, A.S. Md. and Chowdhury, S.A.H.: *A node estimation technique in underwater wireless sensor network*. In: Proceedings of the 2nd International Conference on Informatics, Electronics and Vision (ICIEV 2013), 17 - 18 May, Dhaka University, Dhaka, Bangladesh.

20. Vogt, H.: *Efficient object identification with passive RFID tags*. In: *Lecture Notes in Computer Science*, Springer-verlag, vol. 2414, 2002. Available from: <<http://www.vs.inf.ethz.ch/publ/papers/pvc02rfid.pdf>>
21. Feller, F.: *An Introduction to Probability Theory and its Applications*. John Wiley, 1968.
22. Anower, M. S.: *Estimation using cross-correlation in a communications network*. In: Ph.D. dissertation, SEIT, University of New South Wales at Australian Defense Force Academy, Canberra, 2011.