Anomaly Detection in UASN Localization Based on Time Series Analysis and Fuzzy Logic

Anjana P. Das[1] · Sabu M. Thampi[2] · Jaime Lloret[3]

Abstract

Underwater acoustic sensor network (UASN) offers a promising solution for exploring underwater resources remotely. For getting a better understanding of sensed data, accurate localization is essential. As the UASN acoustic channel is open and the environment is hostile, the risk of malicious activities is very high, particularly in time-critical military applications. Since the location estimation with false data ends up in wrong positioning, it is necessary to identify and ignore such data to ensure data integrity. Therefore, in this paper, we propose a novel anomaly detection system for UASN localization. To minimize computational power and storage, we designed separate anomaly detection schemes for sensor nodes and anchor nodes. We propose an auto-regressive prediction-based scheme for detecting anomalies at sensor nodes. For anchor nodes, a fuzzy inference system is designed to identify the presence of anomalous behavior. The detection schemes are implemented at every node for enabling identification of multiple and duplicate anomalies at its origin. We simulated the network, modeled anomalies and analyzed the performance of detection schemes at anchor nodes and sensor nodes. The results indicate that anomaly detection systems offer an acceptable accuracy with high true positive rate and F-Score.

Keywords Underwater sensor networks · Localization · Time series analysis · Anomaly detection · Fuzzy logic · Auto-regression

1 Introduction

The growing need of ocean exploration demands the development of smart underwater acoustic sensor network(UASN). The estimation of geographical location of sensor nodes in UASN is termed as localization. Localization of an UASN node is very challenging because of unpredictable node mobility, absence of GPS, propagation delay and low bandwidth of the acoustic channel. In the case of real-time surveillance and data collection, information regarding their location is vital for interpreting the data received from UASN nodes accurately.

UASN localization is well studied in the last decade [13, 21]. Localization techniques proposed for mobile UASNs can be grouped as range-free and range-based. As range-free techniques provide ambiguous location estimates, range-based techniques attract more attention. Most of the range-based schemes are anchor node based [8, 16]. SenLin et al. [44] proposed a localization scheme that does not require an even distribution of anchors. Beniwal et al. [4] introduced a time synchronization free localization scheme. They discussed an energy-efficient localization in [26]. Some localization techniques are suitable only for near surface networks [5, 14, 29]. Recently, researchers are focusing on localization schemes with mobility modeling and prediction [24, 27, 47]. Das and Thampi [9] proposed a fault resilient localization that exploit the spatial correlation property of underwater objects. However, the aforementioned techniques do not offer a robust, attack-resistant, and a secure communication path in localization.

UASN is vulnerable to different kinds of malicious activities because of the isolated and hostile nature of underwater environment. Also, an adversary can easily interrupt or deny packets through the unsecured and open acoustic channel [6, 7, 19]. This is very critical when the UASN system is deployed for defense applications. In practical cases, UASN is highly constraint to power, connectivity, and

✉ Anjana P. Das
anjanapdas@gmail.com

[1] University of Kerala, Kerala, India

[2] Indian Institute of Information Technology and Management-Kerala, Kerala, India

[3] Universidad Politecnica de Valencia, València, Spain

computation, which in turn become more vulnerable to faults and malicious activities. Moreover, in IoT applications, as the entanglement with the heterogeneous physical realm is very tight, security and privacy are the key requirements [25, 32–34]. Security threats cause immense degradation of data availability, thus affecting data freshness. As localization at a sensor node highly depend on freshness and integrity of data, false data injection ends up in wrong positioning. In large networks, forwarding and processing the erroneous wrong location information result in wastage of energy and time. Thus, to ensure accurate and timely localization, integrity and authenticity are very critical. The malicious activities or intrusion to a network normally show an unusual behavior. Hence, for accurate data interpretation and ignoring the false data, it is required to detect anomaly behavior in localization at their origin.

In this paper, we propose the first novel anomaly detection system for UASN localization. To identify and ignore malicious data at their origin, anomaly detection schemes are implemented in each sensor node and anchor node. We propose an auto-regressive prediction-based scheme for detecting anomalies at a sensor node (S-Node). A fuzzy inference system is designed to identify the presence of anomalous behavior at anchor nodes (A-Node). Once an anomaly is detected, the received data will be ignored and an alert packet is broadcast. We simulated the network and analyzed the performance of detection schemes. Results showed that anomaly detection schemes at A-Nodes and S-Nodes have good accuracy with less false alarms.

Our major contributions are

- An anomaly detection scheme is proposed for UASN localization
- The network is simulated and analyzed the performance of detection schemes

The rest of this paper is structured as follows: Section 2 summarizes related works, Section 3 explains network architecture and steps in localization, the proposed anomaly detection system is discussed in Sections 4 and 5 describes the experimental results and discussions, and Section 6 concludes the paper.

2 Related work

Even though energy efficiency, storage, data dissemination, time synchronization, routing and localization of UASN are well explored [40], its security aspects are scarcely addressed. Wormhole attacks in UASNs are studied in [18, 42, 43]. Dini and Duca [12] proposed a secure communication suit for UASN that offers confidentiality and integrity in routing. Authors also addressed [11] the security issues in network discovery. Atenieseet al. [3] proposed a security framework for secure routing in UASNs. Han et al. [15] proposed a trust model for UASN.

Secure localization schemes and frameworks proposed for WSN are well enough to address the issues with terrestrial networks [2, 17, 31, 37, 39, 45], but they are not enough to cope up with the challenges in UASN. As the unique characteristics and challenges of UASN are different from WSN [7], UASN requires trusted, reliable, and energy-efficient systems.

Anomaly detection mechanisms for terrestrial wireless sensor networks (WSN) can be classified as parametric-based and non-parametric based [35]. Parametric-based methods either follow statistical approaches or machine learning based classifiers that require a trained model. They are application specific and suitable only when the underlying data distribution is known. Moreover, machine learning based techniques cannot be used in UASN scenario because the behavior of the network is unpredictable and is not predefined in an isolated underwater environment. Non-parametric approaches do not require any prior knowledge about the distribution of data. It can be applied to networks having frequent changes in data distribution. Non-parametric approaches can be again classified as rule-based, density-based, and clustering-based. In density-based methods, population density distribution of the data is approximated and anomalies are tracked as those data points in low density areas. In clustering approaches, anomaly detection is performed on clustered data. Most of the rule-based approaches take information from data packets such as interval time between two packets, number of retransmissions, message payload size, and number of packet collisions [28, 38]. In UASN domain, a rule-based system with multiple parameters requires complex computation.

The approaches to deploy anomaly detection schemes in WSN can be categorized as centralized and distributed [36]. Distributed deployment is preferred in sensor networks. Centralized deployment creates high communication overhead, since all sensor nodes are supposed to send data to a gateway node or base station where anomaly detection is performed. Centralized anomaly detection consumes a high amount of node energy hence reducing the network lifetime.

Liu et al. [20] proposed anomaly detection, localization, and diagnosis scheme for WSN. Mamun et al. [23] proposed a voronoi diagram based network architecture to detect anomaly in WSN. Zheng et al. [46] introduced a trust-assisted anomaly detection and localization. The historic observations of anomaly detection are used to compute the trust of a link in the assumption that anomaly at one link is independent of others. They adopt an incremental probe selection based strategy for anomaly detection and localization that ends up in high communication overhead.

The inherent limitations of UASN should be considered in the design of the anomaly detection system so that energy consumption in the sensor nodes can be minimized and the lifetime of the network be maximized. The major constraints are acceptable accuracy, minimum computational power, and less false alarms. Therefore, to minimize energy utilization, instead of centralized or distributed deployment approach, we designed separate anomaly detection schemes for S-Nodes and A-Nodes. The detection schemes are implemented in every node, which enable identifying multiple and duplicate anomalies at its origin. Detection is progressed by processing and analyzing the received location data and does not require any additional communication packets. Apart from traditional complex rule based techniques, we designed a fuzzy rule based anomaly detection system at the A-Node where only the distance measurement derived from the location data is used for the rule generation. To scale with the unexpected nature of the hostile underwater environment, instead of trained machine learning model, we designed a statistical auto-regressive model to identify the anomalies at the S-Node.

3 UASN localization

The architecture of a typical UASN is shown in Fig. 1. The network is assumed to be having *n* S-Nodes geographically clustered. Each cluster consists of an A-Node and a group of S-Nodes in its coverage. S-Node is the sensor device deployed under water. A-Node is a surface-level node which is equipped with GPS receivers to support the positioning of S-Nodes. A-Nodes are monitored from remote terrestrial control stations through radio link. S-Nodes in a particular cluster are localized with the aid of the A-Node in that cluster. Figure 2 illustrates the steps in UASN localization procedure introduced in [8]. An A-Node sends its location data to all S-Nodes within its coverage periodically. Once an unusual event occurs, S-Node estimates its location based on the information retrieved from an A-Node. The estimated location information is embedded with the data packet and sent back to the A-Node. Therefore, localization procedure consists of two packet transfers, which are as follows:

1. The broadcast message from an A-Node to a desired S-Node
2. The location information along with the sensor data from S-Node to A-Node

3.1 Attack model

The malicious activities in the communication path from the S-Node to A-Node and from A-Node to S-Node lead to location estimation with false data and result in the wrong positioning that in turn affects the accuracy and integrity of data interpretation at the remote control station. Figure 3 shows the impact of external attack in localization procedure. We assume that the presence of malicious behavior in these communication paths exhibits an abnormal behavior in the received packet.

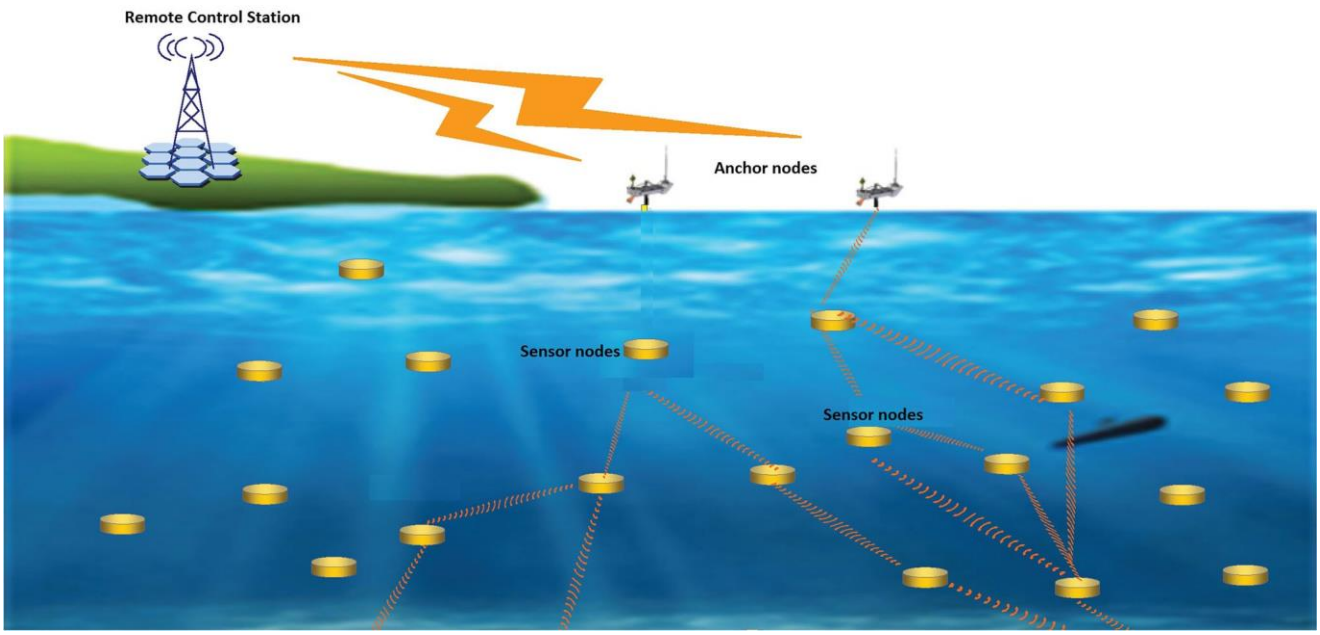

Fig. 1 UASN architecture

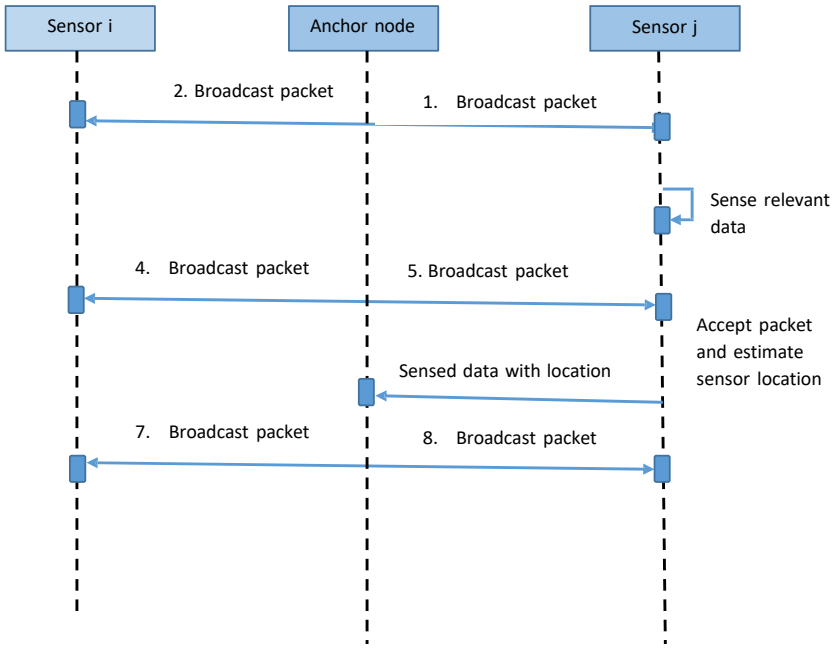

Fig. 2 Sequence diagram illustrating localization steps

4 Proposed anomaly detection system

As the role, functionality, mode of operation, and the format of sending and receiving packets are different for A-Node and S-Node, to ensure security with least computation and storage, separate anomaly detection systems are designed for A-Nodes and S-Nodes. The anomaly detection system is meant to examine the abnormal behavior of each incoming packet. Figure 4 illustrates the steps involved in anomaly detection enabled localization. The anomaly detection system at the A-Node and S-Node identifies the inconsistency in the behavior of the received packet. Once an A-Node receives a packet from the S-Node, the anomaly detection system checks its unusual behavior. If an anomaly is detected, the packet will be rejected and an alert message indicating the anomaly detection is broadcast. Otherwise, the A-Node accepts the packet. Similarly, anomaly detection system implemented in the S-Node examines the unusual nature of the incoming packet and accepts or rejects it according to the suggestion from the detection system.

The inconsistency of the received packet is examined by using the location data retrieved from the packet.

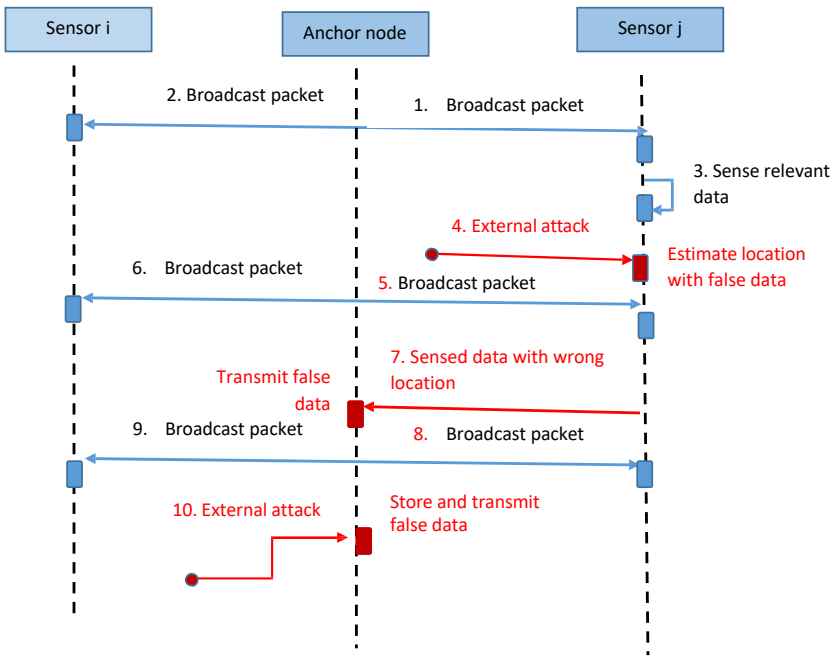

Fig. 3 Sequence diagram illustrating external attack in localization process

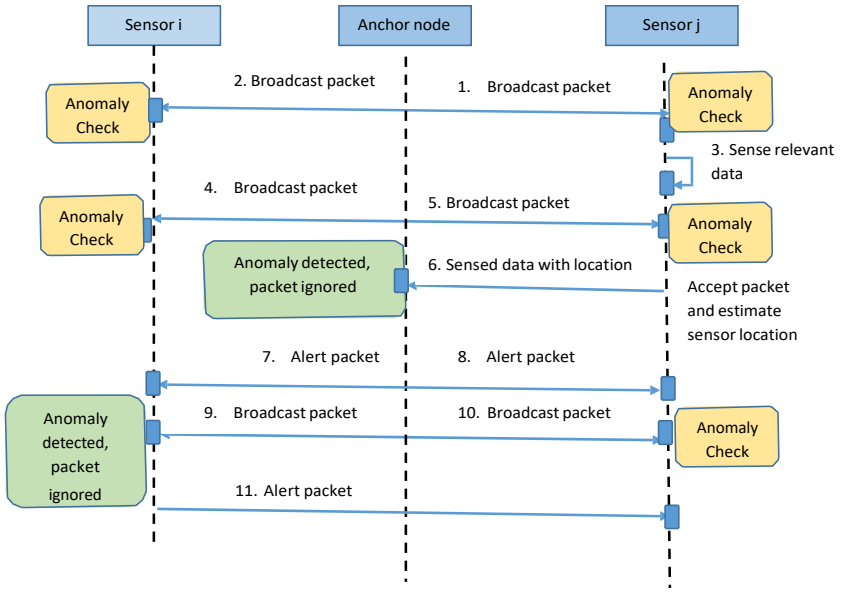

Fig. 4 Sequence diagram illustrating anomaly detection in localization

The variations in other packet-level parameters such as time-interval, the number of retransmissions, and packet collisions do not show a consistent behavior because of the hostile nature of the underwater environment. Also, retrieving and handling multiple packet-level parameters require high computation. Figure 5 illustrates the high level architectural view of detection systems designed for A-Node and S-Node. Since an A-Node sends its location data periodically to all S-Nodes, each S-Node is able to keep track of A-Node's geographical coordinates. Anomalies are detected by exploiting the statistical time series prediction principles on these stored coordinate data. For detecting anomaly at the S-Node, an anomaly index η is derived using fuzzy intelligence.

4.1 Anomaly detection system at an S-Node

The abnormal nature of the packet from an A-Node is determined by observing the previous location history of the A-Node. A-Node has a specific mobility pattern that can be viewed by analyzing its location data. The variation in location co-ordinates between any two consecutive packet transfers shows an almost similar behavior. It is assumed that the location data in a malicious packet is different from the mobility pattern of the A-Node and does not resemble the nature of the previous location data. Figure 6 depicts the detailed view of the detection system. Since the underwater objects follow a circular orbital motion, there will be least variation in the y-co-ordinate compared to x-co-ordinate

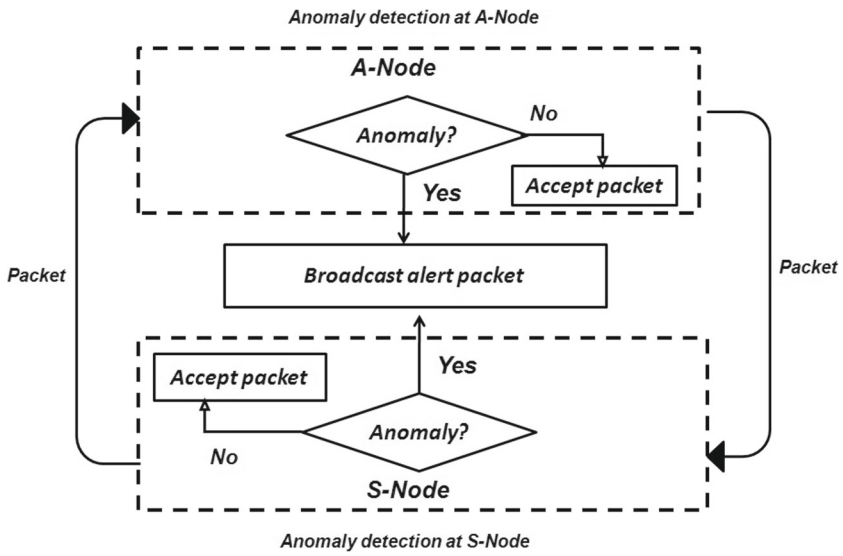

Fig. 5 Anomaly detection system

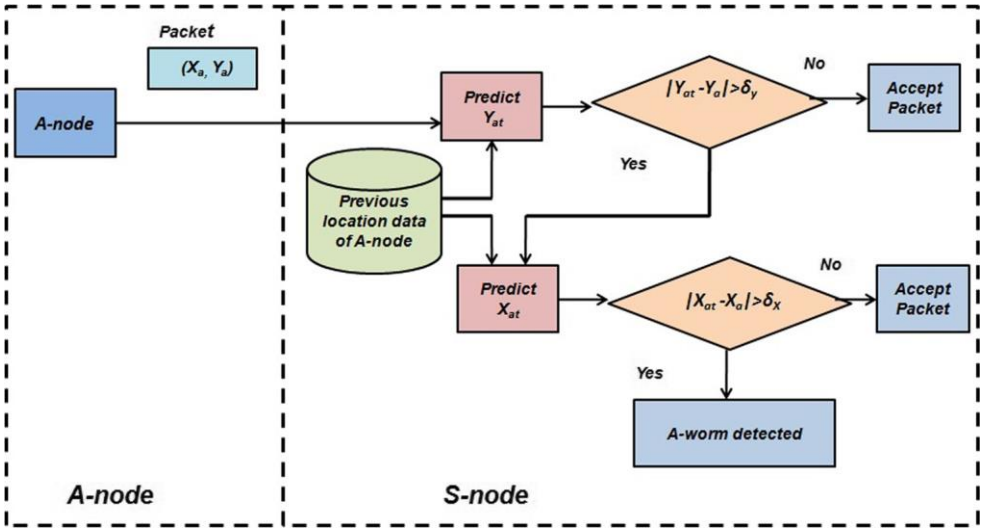

Fig. 6 Anomaly detection system at an S-Node

with respect to time [41]. Hence, the y-co-ordinate of A-Node in the incoming packet is examined first. For that, the y-coordinate at time t is predicted and analyzed. If the standard deviation, σ_Y of the predicted value and the received value is greater than the y-co-ordinate threshold δ_Y, then the x-coordinate will be predicted and analyzed in a similar manner. Otherwise, the packet will be accepted. If the standard deviation, σ_X of x co-ordinate prediction, is also greater than the x-coordinate threshold, δ_X, the received packet will be rejected and an alert packet confirming anomaly detection is broadcast. Otherwise, the packet will be accepted.

4.1.1 Prediction using AR

As an A-Node sends the packet at regular intervals, each S-Node retrieves co-ordinate information of A-Node from that received packet and stores it as a time series data. Let the series be $(X_{a0}, Y_{a0}), (X_{at}, Y_{at}), (X_{a2t}, Y_{a2t})..., (X_{aT}, Y_{aT})$ where, t is the time lag between two packet receptions and T is the current packet reception time. Table 1 displays the format of A-Node's co-ordinate data stored in an S-Node. In UASN domain, the mobility pattern of an object is with respect to the wave motion. Hence, the location (X_{ap}, Y_{ap}) exhibits a temporal correlation with its lagged data $(X_{a(p-t)}, Y_{a(p-t)})$. Therefore, it is possible to predict the future values of this time series data by observing the behavior of its previous values. However, in the harsh UASN environment, the probability of occurrence of the unexpected factors affecting the mobility behavior of a node is high. The impact of unexpected external forces affecting the wave motion in an UASN environment can be modeled as an unpredictable stochastic factor. Hence, we exploit an auto-regressive (AR) model to predict the future location. The AR model is applied for a time varying random process. AR provides the predicted future value of a parameter depending on its own previous values and an imperfectly predictable stochastic factor. Since the mobility model of UASN nodes have dependency with external domain characteristics, it can be better mapped into an AR model than other auto-regressive moving average models. Let p be the order of an AR model; then the prediction of x-coordinate of an A-Node is

$$X_{aT} = \beta + \beta_1 X_{aT-1} + \beta_2 X_{aT-2} + ... + \beta_p X_{aT-p} + \varepsilon \quad (1)$$

where, β is the constant, $\beta_1, \beta_2, ..., \beta_p$ are the parameters of the AR model, and ε is the error term. Similarly, Y_{aT} can also be predicted.

Table 1 Format of A-Node location history stored in an S-Node

Time	Location
0	(x_{a0}, y_{a0})
t	(x_{at}, y_{at})
2t	(x_{a2t}, y_{a2t})
⋮	⋮

4.2 Anomaly detection system at A-Node

A fuzzy inference system is designed to identify the presence of anomalous data at the A-Node. An S-Node

Table 2 Format of location data stored in an A-Node

Time	Node id	Location	Immediate neighbors
0	S_0	(x_{s0}, y_{s0})	A set of S_i s where, $S_i \in \{S_0 S_1, ..., S_n\}$
	⋮	⋮	⋮
	S_n	(x_{sn}, y_{sn})	A set of S_i s where, $S_i \in \{S_0 S_1, ..., S_n\}$

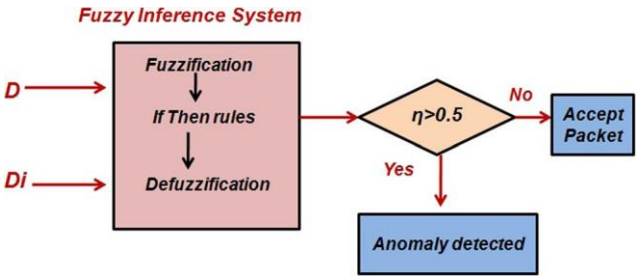

Fig. 7 Anomaly detection system at an A-Node

sends a packet to the A-Node only when it senses an unusual or critical event in the environment. Hence an A-Node does not have the time series location data of all S-Nodes. Also, it is hard to store the location history of S-Nodes in a resource constraint UASN. Instead of that, A-Nodes keep track of the information of the immediate neighbors of all S-Nodes. In UASN domain, S-Nodes exhibits a spatial correlation property [24]. The coordinates of an S-Node are positively correlated with the coordinates of its immediate neighbors. Therefore, immediate neighbors exhibit similar variation in their mobility pattern. This property is exploited to design the fuzzy rules based anomaly detection technique. Initially, at the deployment stage, an A-Node stores the location information and immediate neighbors of each S-Node. Immediate neighbors of an S-Node are other S-Nodes that have the least euclidean distance from it. Table 2 presents the format of data stored in an A-Node. On each incoming packet from S-Node, A-Node checks its trustiness using the fuzzy inference system. Figure 7 illustrates the detection mechanism. The incoming packet has the coordinates of the sender S-Node. Once it is identified as a packet from a trusted S-Node, the location of that S-Node is updated. The Mamdani fuzzy model is used [22] to generate the anomaly index. Two inputs are given to the fuzzy system, D and D_i. Let the S-Node, S_j send a packet to the A-Node and the location is (x_{sj}, y_{sj}). D is the distance between the present and the previous location of S_j. Let the stored location of S_j be (x_{sj}, y_{sj}). D_i is the distance between the present location of S_j and the stored location of its alive immediate neighbor. The alive immediate neighbor is an S-Node in S_j 's immediate neighbor set whose location is recently updated in the stored database. Let S_k be the alive immediate neighbor whose stored location is (x_{sk}, y_{sk}). The input parameters to the fuzzy inference system are calculated as follows:

$$D = \sqrt{(x'_{sj} - x_{sj})^2 + (y'_{sj} - y_{sj})^2} \quad (2)$$

$$Di = \sqrt{(x'_{sj} - x_{sk})^2 + (y'_{sj} - y_{sk})^2} \quad (3)$$

The output of the fuzzy inference system is the anomaly index η, which determines whether to reject or accept the packet. The dependence of η on D and D_i can be expressed by the following four fuzzy rules:

1. If D is *Very near* or *Near* and D_i is *Very near* or *Near*, then the output is *Normal*
2. If D is *Very near* or *Near* and D_i is *Far* or *Very Far*, then the output is *May be Normal*
3. If D is *Far* or *Very Far* and D_i is *Very near* or *Near*, then the output is *May be Anomaly*
4. If D is *Far* or *Very Far* and D_i is *Far* or *Very Far*, then the output is *Anomaly*

The membership functions *Very near*, *Near*, *Far*, and *Very Far* for the variables D and D_i are defined as $\mu_{VN}(x)$, $\mu_N(x)$, $\mu_F(x)$, and $\mu_{VF}(x)$based on the threshold values a, b, c, d, e, and f. It is assumed that the mean time-interval between two packet transfers from an S-Node is 100 s. The distance traveled by all S-Nodes in 100 s is observed, and its minimum value, mean value, and maximum value are taken for the parameters a, b, and c respectively. The parameters d, e, and f are assigned in accordance with a, b, and c. Hence a trapezoidal membership function is used for *Very near*, *Near*, *Far*, and *Very Far*.

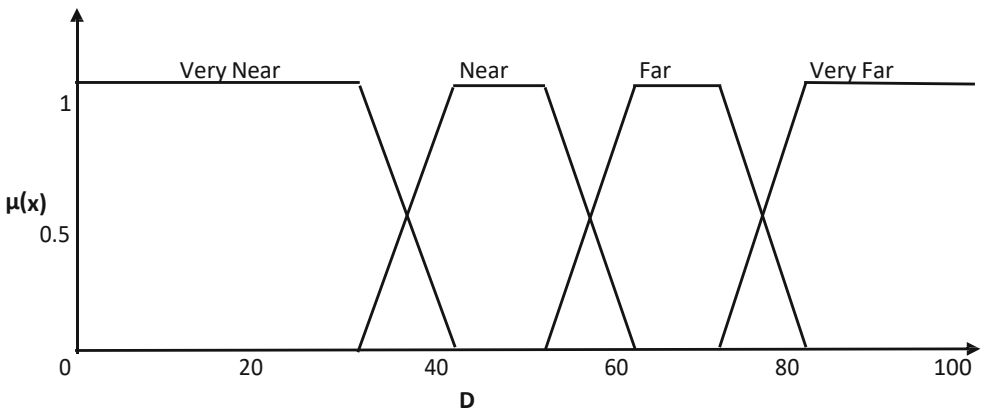

Fig. 8 Membership function of D

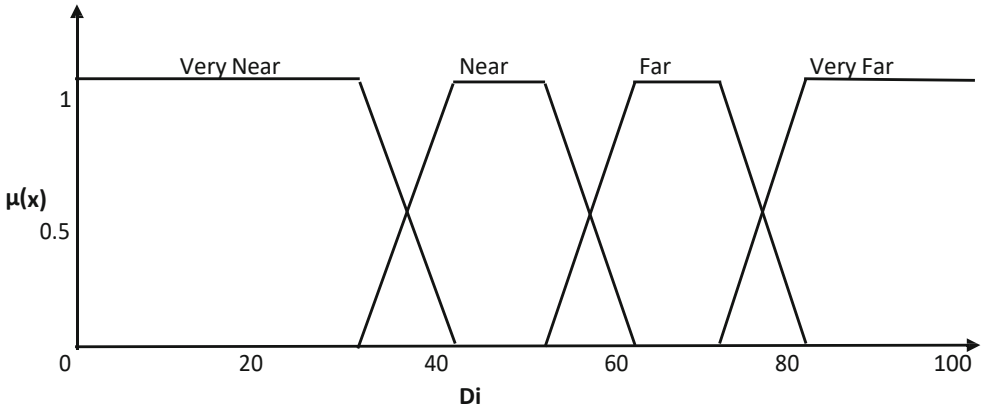

Fig. 9 Membership function of *Di*

The membership function for *Very near*, *Near*, *Far*, and *Very Far* are defined as follows:

$$\mu_{VN}(x) = \begin{cases} 1, & D < a \\ \frac{a-D}{b-a}, & a \le D \le b \\ 0, & Otherwise \end{cases} \quad (4)$$

$$\mu_{N}(x) = \begin{cases} \frac{D-b}{b-a}, & a \le D \le b \\ 1, & b \le D \le c \\ \frac{c-D}{d-c}, & c \le D \le d \\ 0, & Otherwise \end{cases} \quad (5)$$

$$\mu_{F}(x) = \begin{cases} \frac{D-d}{d-c}, & c \le D \le d \\ 1, & d \le D \le e \\ \frac{e-D}{f-e}, & e \le D \le f \\ 0, & Otherwise \end{cases} \quad (6)$$

$$\mu_{VF}(x) = \begin{cases} \frac{D-f}{f-e}, & e \le D \le f \\ 1, & D > f \\ 0, & Otherwise \end{cases} \quad (7)$$

Figures 8 and 9 shows the membership functions of D and Di, respectively. After applying fuzzy inference rules, the obtained result is subjected to defuzzification process. Centroid method is used for defuzzification. The output of defuzzification process is a crisp value η, which indicates the probability of occurrence of abnormal signal. The membership function of η is shown in Fig. 10. To optimize false negatives, the threshold of η is set as 0.5. In each incoming packet, η is computed and compared with the threshold. If $\eta > 0.5$, an alert message indicating anomaly detection is broadcast.

5 Experimental results and discussion

The performance of anomaly detection systems at the A-Node and S-Node is analyzed by conducting separate sets of simulation experiments.

5.1 Anomaly detection at the S-Node

Anomalous nature of the signal from an A-Node is modeled as the data packet with random location data, that is location data having random values of X and Y coordinates. Packets

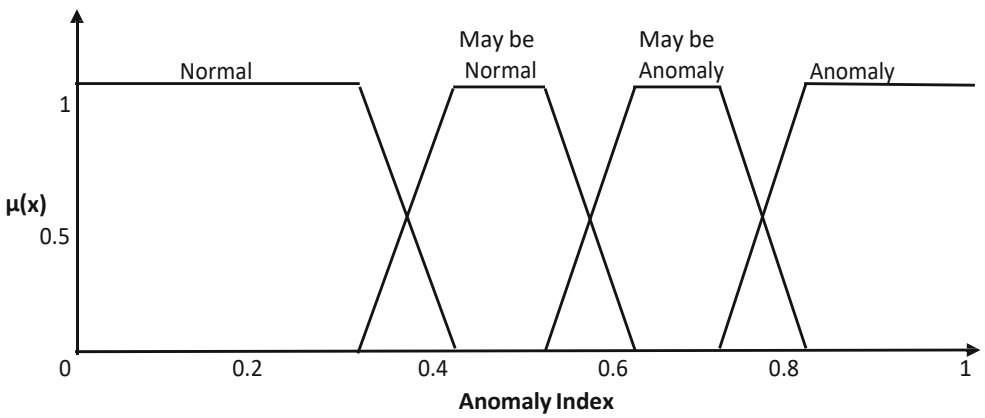

Fig. 10 Membership function of anomaly behavior index

Table 3 Simulation parameters used

Simulator : AquaSim	
Simulation parameter	Value
Communication channel	UnderWater channel
Attenuation model	Thorp's attenuation
Communication range	100 m
Packet size	50 bytes
Mobility model	RPGM
Network Layer	Vector based forwarding
MAC	Broadcast mac
Bit rate	10 kbps
Frequency	25 kHz
Initial Energy	10000 J
Simulation time	700 s

Table 4 Performance metrics

Metric	Equation
Accuracy	$\frac{t_P + t_N}{t_P + t_N + f_P + f_N}$
Precision	$\frac{t_P}{t_P + f_P}$
True positive rate (TPR)	$\frac{t_P}{t_P + f_N}$
False positive rate (FPR)	$\frac{f_P}{f_P + t_N}$
F-Score	$2 * \frac{Precision * TPR}{Precision + TPR}$
Receiver operating characteristics (ROC)	Curve generated by plotting TPR against FPR. The curve above the diagonal line and approaching 1 represents good performance.
Area under the curve (AUC)	It is the area under the ROC curve. The value approaching 1 represents good performance.

with false data is simulated and the performance of the detection mechanism is evaluated.

Simulation settings The network of 20 nodes is simulated in a $200 \times 200 \times 200\ m^3$ space. The simulator used is AquaSim[1], which is a tool used for UASN simulation and research [10]. Table 3 lists out the various simulation parameters. As UASN nodes exhibit group mobility behavior, reference point group mobility model [30] is applied to map the mobility behavior. The UASN network is simulated in such a way that the A-Node sends packet to all S-Nodes every 50 s. Anomalies are modeled as follows:

- A packet with location data as (100,100) is sent to all S-Nodes after 500 s
- A packet with location data as (200,200) is sent to all S-Nodes after 600 s

We considered the location data up to 250 s for learning process of the predictor algorithm in the assumption that the probability of occurrence of anomalies at the initial stage of network deployment is less. After 250 s, abnormal behavior of each packet reception is examined. To find the order of an AR model, the auto-correlation and partial auto-correlation functions of the A-Node data are observed, and AR model with order 1 is selected. The least square method is used to implement the AR model.

Let t_P is the true positive count, t_N is the true negative count, f_P is false positive count, and f_N is false negative count. The performance metrics used for the analysis of detection schemes are listed in Table 4. We considered detection at S-Node 2 as a test case and analyzed the performance at S-Node 2. Table 5 displays the results of detection. To optimize the false negatives, the thresholds δ_X and δ_Y are set as 10. Observing the results of detection in all packet reception cases, an accuracy of 90%, precision of 66.6%, and true positive rate of 100% are obtained when δ_y and δ_x are set as 10.

5.1.1 The impact of δ_y

δ_y has a significant impact on the processing overhead associated with anomaly detection at an S-Node. The detection algorithm progresses only when the estimated variation is greater than δ_y. Even though lower δ_y ensures an accurate trusted detection system, it increases the processing power. A low δ_y value increases false positives without affecting true positives. We analyzed the impact of δ_y on accuracy, precision, and sensitivity of detection. Table 6 depicts the variation in performance of the detection system with different values of δ_y. Figure 11 represents the ROC (receiver operating characteristics) curve obtained for different values of δ_y. The computational complexity of an AR(1) model with n members is O(n). The energy utilization in each detection phase depends on the complexity of running the AR(1)model. So, energy utilization linearly depends on data storage. Since we designed the system without specifying the storage limit, the processing overhead increases with time, which means processing overhead is linearly dependent on network lifetime.

5.1.2 The impact of δ_x

As the horizontal movement of nodes with respect to the ocean wave is less compared to the vertical variation, the displacement of x with time is less. Thus, even if the small

[1]http://obinet.engr.uconn.edu/wiki/index.php/Aqua-Sim

Table 5 Results of anomaly detection at S-Node 2

Time (S)	Received A-Node coordinates	Predicted Y	σ_y	Predicted X	σ_x	Result 1-Anomaly 0-Normal
300	(173.04, 137.25)	126.8736948	21.38053	126.8736948	32.644507	1
350	(134.47, 193.66)	135.6120027	16.12406	135.6120027	0.8075179	0
400	(125.45, 168.76)	129.8952341	10.35466	129.8952341	3.1432552	0
450	(131.17, 173.46)	129.0838737	7.031253	-	-	0
500	(127.36, 134.2)	129.5552317	21.06258	129.5552317	1.5522632	0
550	(100, 100)	126.9827525	24.4018	126.9827525	26.9827525	1
550	(122.43, 71.98)	126.9827525	37.06783	126.9827525	3.2192822	0
600	(118.19, 18.4)	127.8613525	24.98678	27.8613525	6.8386789	0
650	(200, 200)	127.5414826	180.7961	127.5414826	72.4585174	1
650	(103.45, 19.55)	127.5414826	0.244753	-	-	0

value of δ_x requires more processing power, to ensure a trusted detection system, the value of δ_x should be as small as possible. We analyzed the impact of δ_x. Table 6 depicts the variation in performance of the detection system with different values of δ_x. Figure 12 represents the ROC curve obtained for different values of δ_x. Results show 90% as accuracy, 0.94 as AUC value, and 0.8 as F-Score for (δ_x, δ_y) combinations like (10, 10), (10, 20), and (20, 10).

5.2 Anomaly detection at A-Node

To evaluate the performance of an anomaly detection scheme at an A-Node, we conducted simulation experiments by modeling anomaly behavior as transmitting packets from S-Nodes with random coordinate values.

5.2.1 Simulation settings

A network of 10 nodes is simulated in a $500 \times 500 \times 500$ m^3 space using AquaSim [1]. Table 3 lists the simulation parameters used. Generally, an S-Node sends data to A-Node only when an unexpected event occurs. Hence, the UASN network is simulated in such a way that S-Nodes send packets to A-Node randomly. Anomalies are modeled as random packet transfers from random locations.

The values of parameters D and D_i are estimated and given to the fuzzy intelligence system to generate the anomaly index, η. To optimize false negatives, the value of η is taken as 0.5. Table 7 presents the results of the detection system. The detection system showed 100% TPR with 72% accuracy and 41.66% precision when η is 0.5.

Table 6 Performance comparison of anomaly detection system at S-Node 2 with different δ_y and δ_x combinations

δ_y	δ_x	Accuracy (%)	Precision (%)	TPR (%)	F-Score	AUC
5	10	80	50	100	0.66	0.88
10	10	90	66.6	100	0.8	0.94
20	10	90	66.6	100	0.8	0.94
30	10	80	50	50	0.5	0.69
10	5	80	50	100	0.66	0.88
10	20	90	66.6	100	0.8	0.94
10	30	80	50	50	0.5	0.69

5.2.2 The impact of η

As the value of η plays an important role in the performance of the detection system, we analyzed the performance

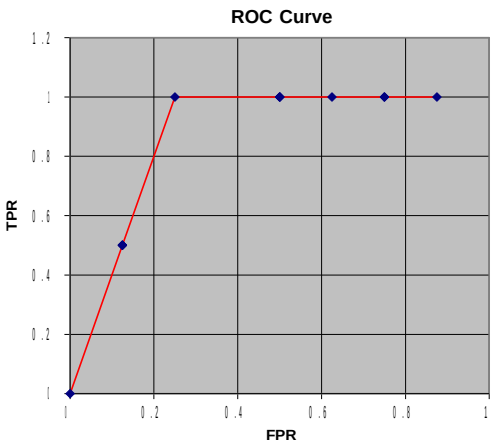

Fig. 11 ROC curve of anomaly detection at S-Node 2 with different δ_y values

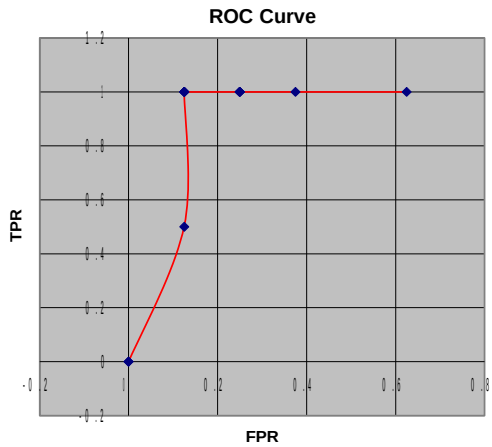

Fig. 12 ROC curve of anomaly detection at S-Node 2 with different δ_x values

Table 7 Results of anomaly detection at A-Node

Time (s)	Sender S-node id	Location *(x, y)*	η
50	S_1	(64.2689, 60.3963	0.1737
90	S_8	(65.6001, 165.793)	0.1745
100	S_0	(158.702, 160.919)	0.1737
110	S_2	(71.7167, 69.7725)	0.1737
150	S_1	(88.7677, 69.2657)	0.1737
150	Pretending S_1	(200, 200)	0.8573
170	S_9	(383.257, 224.097)	0.5
200	S_3	(19.3594, 79.8404)	0.8755
250	S_2	(36.6057, 81.3597)	0.1806
250	Pretending S_2	(100, 100)	0.8759
260	Pretending S_4	(250, 250)	0.5
300	S_7	(95.2372, 130.8)	0.65
310	S_2	(27.9219, 79.5283)	0.1737
330	S_3	(21.2074, 79.4102)	0.1737
360	Pretending S_5	(200, 200)	0.8722
400	S_4	(52.8232, 66.9643)	0.1828
410	S_3	(27.2177, 108.976)	0.174
420	S_9	(216.231, 238.912)	0.8615
440	S_0	(107.455, 103.68)	0.8598
440	S_5	(78.8085, 108.997)	0.65
450	S_6	(48.0862, 64.3274)	0.7479
460	S_7	(70.5413, 97.5305)	0.1737
470	Pretending S_3	(100, 150)	0.8759
480	S_2	(25.117, 114.204)	0.1847
490	S_1	(134.379, 74.6605)	0.45

Table 8 Performance comparison of anomaly detection system at A-Node with different η values

η	Accuracy (%)	Precision (%)	TPR (%)	F-Score	AUC
0.2	68	38.46	100	0.55	0.8
0.3	68	38.46	100	0.55	0.8
0.4	68	38.46	100	0.55	0.8
0.5	72	41.66	100	0.58	0.82
0.6	76	44.44	80	0.57	0.77
0.7	80	50	80	0.61	0.8
0.9	80	-	0	-	0.5

variation and behavior of the detection system with respect to η. We analyzed accuracy, precision, TPR, F-Score, and the ROC curve of the detection system for different η values. Table 8 shows the performance variation of detection algorithm with respect to η. Figure 13 shows the ROC curve. Even though higher η gives good accuracy, it has least TPR. The maximum AUC value is obtained when η=0.5 with F-Score 0.58.

The whole concept of anomaly detection is derived based on the spatial correlation property of underwater objects. In the anomaly detection system at A-Node, the anomaly index η depends on the input parameters, D and Di , which are derived directly. As S-Nodes do not show a periodic nature in the packet transfer, normalized D and Di with respect to packet interval time, can provide more robust results. Moreover, the complexity level analysis of detection schemes help to observe the variations in network performance.

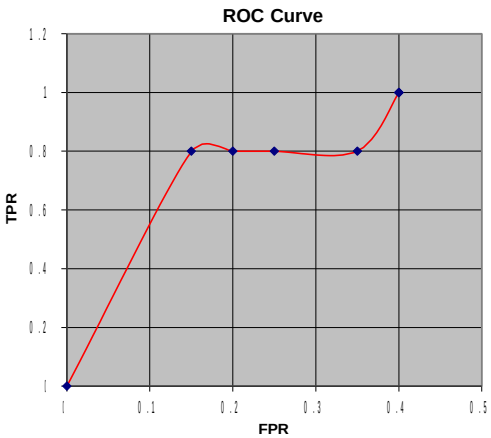

Fig. 13 ROC curve of anomaly detection at A-Node with different η values

6 Conclusion

We proposed the first novel anomaly detection scheme for UASN localization. The detection scheme is implemented in all nodes so that multiple, and duplicate anomalies can be detected at their origin. The statistical time series prediction principles are applied to identify an anomaly at S-Node. Anomaly index is derived using the fuzzy inference system to identify the presence of anomalies at A-Node. Detection schemes offered good accuracy with less false alarms. The inherit domain properties of UASN is considered in the design of algorithms so that it become scalable enough to build a trusted secure UASN platform in future.

Acknowledgements The authors thank the Higher Education Department, Government of Kerala, for the research fellowship.

References

1. Aqua-sim. [online] http://obinet.engr.uconn.edu/wiki/index.php/Aqua-Sim. (Accessed 5 November 2015)
2. Anjum F, Pandey S, Agrawal P (2005) Secure localization in sensor networks using transmission range variation. In: 2005. IEEE international conference on Mobile adhoc and sensor systems conference. IEEE, pp 9
3. Ateniese G, Capossele A, Gjanci P, Petrioli C, Spaccini D (2015) Secfun: Security framework for underwater acoustic sensor networks. In: OCEANS 2015-Genova. IEEE, pp 1–9
4. Beniwal M, Singh RP, Sangwan A (2016) A localization scheme for underwater sensor networks without time synchronization. Wireless Personal Communications 88(3):537–552
5. Chen P, Ma H, Gao S, Huang Y (2015) Ssl: Signal similarity-based localization for ocean sensor networks. Sensors 15(11):29702–29720
6. Cong Y, Yang G, Wei Z, Zhou W (2010) Security in underwater sensor network. In: 2010 international conference on Communications and mobile computing (CMC). IEEE, vol 1, pp 162–168
7. Das AP, Thampi SM (2015) Secure communication in mobile underwater wireless sensor networks. In: Proceedings of the IEEE International Conference on Advances in Computing, Communications and Informatics, India, pp 2164–2173
8. Das AP, Thampi SM (2015) Single anchor node based localization in mobile underwater wireless sensor networks. In: Algorithms and architectures for parallel processing. Springer, pp 757–770
9. Das AP, Thampi SM (2016) Fault-resilient localization for underwater sensor networks. Ad Hoc Networks
10. Das AP, Thampi SM (2016) Simulation tools for underwater sensor networks: a survey. Netw Protocol Algorithm 8(4):41–55
11. Dini G, Duca AL (2011) Seflood: a secure network discovery protocol for underwater acoustic networks. In: 2011 IEEE symposium on Computers and communications (ISCC). IEEE, pp 636–638
12. Dini G, Lo Duca A (2012) A secure communication suite for underwater acoustic sensor networks. Sensors 12(11):15133–15158
13. Garcia M, Sendra S, Atenas M, Lloret J (2011) Underwater wireless ad-hoc networks: a survey. Mobile ad hoc networks: Current status and future trends 379–411
14. Gomez JV, Sandnes FE, Fernandez B (2012) Sunlight intensity based global positioning system for near-surface underwater sensors. Sensors 12(2):1930–1949
15. Han G, Jiang J, Shu L, Guizani M (2015) An attack-resistant trust model based on multidimensional trust metrics in underwater acoustic sensor network. IEEE Trans Mob Comput 14(12):2447–2459
16. Han G, Jiang J, Shu L, Xu Y, Wang F (2012) Localization algorithms of underwater wireless sensor networks: a survey. Sensors 12(2):2026–2061
17. Jiang J, Han G, Zhu C, Dong Y, Zhang N (2011) Secure localization in wireless sensor networks: a survey. JCM 6(6):460–470
18. Kong J, Ji Z, Wang W, Gerla M, Bagrodia R, Bhargava B (2004) On wormhole attacks in under-water sensor networks: a two-tier localization approach. UCLA Computer Science Department Technical Report 4005
19. Kong J, Ji Z, Wang W, Gerla M, Bagrodia R, Bhargava B (2005) Low-cost attacks against packet delivery, localization and time synchronization services in under-water sensor networks. In: Proceedings of the 4th ACM workshop on Wireless security. ACM, pp 87–96
20. Liu J, Djurdjanovic D, Marko KA, Ni J (2009) A divide and conquer approach to anomaly detection, localization and diagnosis. Mech Syst Signal Process 23(8):2488–2499
21. Lloret J (2013) Underwater sensor nodes and networks. Sensors 13(9):11782–11796
22. Mamdani EH, Assilian S (1975) An experiment in linguistic synthesis with a fuzzy logic controller. Int J Man-Mach Stud 7(1):1–13
23. Mamun Q, Islam MR, Kaosar M (2014) Anomaly detection in wireless sensor network. JNW 9(11):2914–2924
24. Meiqin L, Xiaodong G, Senlin Z (2015) Localization based on best spatial correlation distance mobility prediction for underwater wireless sensor networks. In: 34Th chinese control conference, pp 7827–7832, Hangzhou
25. Miorandi D, Sicari S, De Pellegrini F, Chlamtac I (2012) Internet of things: vision, applications and research challenges. Ad Hoc Netw 10(7):1497–1516
26. Mirza D, Schurgers C (2008) Energy-efficient ranging for post-facto self-localization in mobile underwater networks. IEEE J Sel Areas Commun 26(9):1697–1707
27. Ojha T, Misra S (2013) Mobil: a 3-dimensional localization scheme for mobile underwater sensor networks. In: Proceedings of the IEEE National Conference on Communications, pp1–5, New Delhi
28. Onat I, Miri A (2005) A real-time node-based traffic anomaly detection algorithm for wireless sensor networks. In: Systems communications, 2005. misc. IEEE, pp 422–427
29. Park D, Kwak K, Kim J, Chung WK (2015) Underwater sensor network using received signal strength of electromagnetic waves. In: Proceedings of the IEEE International Conference on Intelligent Robots and Systems,Hamburg, Germany
30. Pei G, Gerla M, Hong X, Chiang CC (1999) A wireless hierarchical routing protocol with group mobility. In: proceedings of the IEEE Wireless Communications and Networking Conference, pp 1538–1542, New Orleans
31. Poovendran R, Wang C, Roy S (2007) Secure localization and time synchronization for wireless sensor and ad hoc networks, vol 30. Springer Science & Business Media, Berlin
32. Qiu T, Chen N, Li K, Atiquzzaman M, Zhao W (2018) How can heterogeneous internet of things build our future: A survey. IEEE Communications Surveys & Tutorials

33. Qiu T, Qiao R, Wu D (2018) Eabs: an event-aware backpressure scheduling scheme for emergency internet of things. IEEE Trans Mob Comput 17(1):72–84
34. Qiu T, Zhao A, Xia F, Si W, Wu D, Qiu T, Zhao A, Xia F, Si W, Wu D (2017) Rose: Robustness strategy for scale-free wireless sensor networks. IEEE/ACM Trans Netw (TON) 25(5):2944–2959
35. Rajasegarar S, Leckie C, Palaniswami M (2008) Anomaly detection in wireless sensor networks. IEEE Wireless Communications 15(4):34–40
36. Rajasegarar S, Leckie C, Palaniswami M, Bezdek JC (2006) Distributed anomaly detection in wireless sensor networks. In: 2006. ICCS 2006. 10th IEEE Singapore international conference on Communication systems. IEEE, pp 1–5
37. Rasmussen KB, Capkun S, Cagalj M (2007) Secnav: secure broadcast localization and time synchronization in wireless networks. In: Proceedings of the 13th annual ACM international conference on Mobile computing and networking. ACM, pp 310–313
38. da Silva APR, Martins MH, Rocha BP, Loureiro AA, Ruiz LB, Wong HC (2005) Decentralized intrusion detection in wireless sensor networks. In: Proceedings of the 1st ACM international workshop on Quality of service & security in wireless and mobile networks. ACM, pp 16–23
39. Srinivasan A, Wu J (2007) A survey on secure localization in wireless sensor networks. Encyclopedia of Wireless and Mobile communications
40. Tariq M, Latiff MSA, Ayaz M, Coulibaly Y, Wahid A (2016) Pressure sensor based reliable (psbr) routing protocol for underwater acoustic sensor networks. Ad Hoc Sensor Wirel Netw 32(3-4):175–196
41. Thurman HV, Trujillo AP, Abel DC, McConnell R (1999) Essentials of oceanography. Prentice Hall, Englewood Cliffs
42. Wang W, Kong J, Bhargava B, Gerla M (2008) Visualisation of wormholes in underwater sensor networks: a distributed approach. Int J Secur Netw 3(1):10–23
43. Zhang R, Zhang Y (2010) Wormhole-resilient secure neighbor discovery in underwater acoustic networks. In: INFOCOM, 2010 Proceedings IEEE, pp 1–9. https://doi.org/10.1109/INFCOM.2010.5462093
44. Zhang S, Zhang Q, Liu M, Fan Z (2014) A top-down positioning scheme for underwater wireless sensor networks. Sci China Inf Sci 57(3):1–10
45. Zhang Y, Liu W, Fang Y, Wu D (2006) Secure localization and authentication in ultra-wideband sensor networks. IEEE J Sel Areas Commun 24(4):829–835
46. Zheng S, Baras JS (2011) Trust-assisted anomaly detection and localization in wireless sensor networks. In: 2011 8th annual IEEE communications society conference on Sensor, mesh and ad hoc communications and networks (SECON). IEEE, pp 386–394
47. Zhou Z, Peng Z, Cui JH, Shi Z, Bagtzoglou A (2011) Scalable localization with mobility prediction for underwater sensor networks. IEEE Trans Mob Comput 10(3):335–348