This item was submitted to [Loughborough's Research Repository](#) by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

A trust update mechanism based on reinforcement learning in underwater acoustic sensor networks

PLEASE CITE THE PUBLISHED VERSION

<https://doi.org/10.1109/tmc.2020.3020313>

PUBLISHER

Institute of Electrical and Electronics Engineers (IEEE)

VERSION

AM (Accepted Manuscript)

PUBLISHER STATEMENT

Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

LICENCE

All Rights Reserved

REPOSITORY RECORD

He, Yu, Guangjie Han, Jinfang Jiang, Hao Wang, and Miguel Martinez-Garcia. 2020. "A Trust Update Mechanism Based on Reinforcement Learning in Underwater Acoustic Sensor Networks". Loughborough University. <https://hdl.handle.net/2134/13026344.v1>.

A Trust Update Mechanism Based on Reinforcement Learning in Underwater Acoustic Sensor Networks

Yu He, Guangjie Han, *Senior Member, IEEE*, Jinfang Jiang, Hao Wang, and Miguel Martínez-García, *Member, IEEE*.

Abstract—Underwater acoustic sensor networks (UASNs) have been widely applied in marine scenarios, such as offshore exploration, auxiliary navigation and marine military. Due to the limitations in communication, computation, and storage of underwater sensor nodes, traditional security mechanisms are not applicable to UASNs. Recently, various trust models have been investigated as effective tools towards improving the security of UASNs. However, the existing trust models lack flexible trust update rules, particularly when facing the inevitable dynamic fluctuations in the underwater environment and a wide spectrum of potential attack modes. In this study, a novel trust update mechanism for UASNs based on reinforcement learning (TUMRL) is proposed. The scheme is developed in three phases. First, an environment model is designed to quantify the impact of underwater fluctuations in the sensor data, which assists in updating the trust scores. Then, the definition of key degree is given; in the process of trust update, nodes with higher key degree react more sensitively to malicious attacks, thereby better protecting important nodes in the network. Finally, a novel trust update mechanism based on reinforcement learning is presented, to withstand changing attack modes while achieving efficient trust update. The experimental results prove that our proposed scheme has satisfactory performance in improving trust update efficiency and network security.

Index Terms—Underwater acoustic sensor networks, reinforcement learning, trust update, environment model.

1 INTRODUCTION

UNDERWATER acoustic sensor networks (UASNs) are expected to play an increasingly important role in many fields, such as marine environmental monitoring, offshore exploration, auxiliary navigation, tsunami warning and marine military operations [1–4]. As shown in Fig. 1, a typical UASN generally comprises numerous underwater sensor nodes, which cooperate to accomplish environment awareness, information collection and packet transmission [5, 6]. Because of their open and unattended nature, sensor nodes can be easily compromised and attacked [7, 8]. Network security issues have gradually become the main obstacle restricting the development of UASNs.

Traditional security mechanisms such as key management and identity authentication are effective against external attacks. However, they are powerless to internal attackers who have successfully invaded the network [9–13]. An internal attacker can compromise normal nodes to obtain the keys and IDs, which allow them to neutralize most traditional security systems. Recently, various trust models have been investigated as effective tools for confronting internal attacks. Commonly, a trust model mainly includes

three parts: evidence accumulation, trust score computation and trust update. Considerable research has been done on the first two, whereas the latter still presents weakness in complex underwater environments.

Fig. 1. Structure of a typical underwater acoustic sensor network.

In this paper, we propose a novel *trust update mechanism based on reinforcement learning* (TUMRL) for UASNs. First, the impact of the underwater environment is quantified by a specific environment model, which regulates the trust score update process. Then, considering that the nodes with higher importance in the network may be preferentially subject to malicious attacks, the concept of *key degree* is presented – to protect key nodes by increasing their detection sensitivity to offensive attacks. Finally, the overall method is integrated into a decision-making trust score

- Yu He, Guangjie Han and Jinfang Jiang are with the Department of Information and Communication Systems, Hohai University, Changzhou, 213022, China. (e-mail: heyuhhu2018@outlook.com, hanguangjie@gmail.com, jiangjinfang1989@gmail.com).
- Hao Wang is with the Department of Computer Science and Technology, Hohai University, Nanjing, China. (e-mail: wanghaohu@outlook.com).
- Miguel Martínez-García is with the Dept. of Aeronautical and Automotive Engineering, Loughborough University, Leicestershire, LE11 3TU, U.K. (e-mail: m.martinez-garcia@lboro.ac.uk).

Manuscript received XXXX XX, 20XX; revised XXXX XX, 20XX.

update mechanism – by introducing reinforcement learning, and thereby achieving efficient trust update.

The main contributions of this paper are summarized as follows: (i) we propose an environment model to quantify the abstract environmental impact as a probabilistic quantity, which is used to mitigate trust misclassification caused by environmental factors, rather than offensive attacks; (ii) we propose the concept of *key degree* – to increase the trust sensitivity of key nodes in the network, thereby minimizing the losses caused by malicious attacks; (iii) reinforcement learning is integrated into the environment model and the key degree method to conform the process of trust update, while improving the efficiency and adaptivity of the trust update mechanism.

The rest of this paper is organized as follows. In Sec. 2, we provide an overview of related work about trust models. Sec. 3 introduces the network model and some reasonable assumptions. Then, detailed descriptions about the proposed scheme and simulation results are given in Secs. 4 and 5 respectively. Conclusions are drawn in Sec. 6.

2 RELATED WORK

Although research on trust models in UASNs is still in its infancy, trust models have been extensively studied – and proved to have important value in network security – in the terrestrial wireless sensor networks (TWSNs), cloud environments, and other related subjects. According to their differences in network structure, trust models can be divided into two main categories: 1) trust models in distributed networks, and 2) trust models in clustered networks.

Trust models in distributed networks: In a distributed network, each node calculates the trust of neighbor nodes (the nodes in the communication range) and maintains the trust update. The results of trust calculation are shared with other nodes through active broadcast or passive transmission, thereby achieving the global security. Ganeriwal et al. [14] first paid attention to the issue of trust security in sensor networks and proposed a reputation-based trust framework. In this framework, each node updates the reputation of neighbor nodes based on their past behaviors, and predicts the future behaviors of neighbors based on the latest reputation. In order to improve the accuracy of trust calculation, Yao et al. [15] presented a trust model based on direct interaction and recommendation trust. Both past behavior of the neighbor nodes and recommended trust evidence from other nodes are considered when updating the trust scores. To optimize the coupling between direct trust and indirect trust, Feng et al. [16] proposed a trust model based on modified evidence theory. In their system, each node first obtains the relevant direct trust and indirect trust, and then aggregates the information by way of fuzzy sets, which yield the final trust score.

A different approach for distributed TWSNs was studied by Ren et al. [17]. In their method, subjective logic based on consensus techniques is adopted, to mitigate trust fluctuations caused by environmental factors. In addition, each node performs distributed trust storage by maintaining a geographic hash table, thereby reducing the cost of trust update. Further, Jiang et al. [18] proposed an efficient distributed trust model (EDTM) for TWSNs. In the EDTM,

the final trust result is integrated by communication trust, energy trust, and data trust to improve the accuracy. Each node maintains a sliding time window, which comprises several time-slots. Trust value is calculated in each time-slot and updated by sliding the time window iteratively. Lastly, the impact of poor link quality on trust evaluation was examined by Wu et al. [19]. In their work, a beta and link quality indicator (LQI)-based trust model (BLTM) was proposed. In the BLTM, the LQI between two nodes decides whether to update the trust score within a current period, and the weight of trust evidence based on the beta probability density is considered in the process of trust update.

Trust models in clustered networks: In a clustered network, the trust scores are maintained and updated by agent nodes (e.g. cluster heads) instead of the individual nodes. The agent nodes set trust scores for the remaining nodes within a given vicinity, and share their scores with other agent nodes – to achieve collaborative network security.

In order to minimize the overhead of trust models in terms of extra messages and time delays, Boukerch et al. [20] put forth an agent-based trust and reputation management scheme (ATRM) for TWSNs. In the ARTM, each node stores its trust information, and a mobile agent is responsible for trust calculation and trust update for each node in a local range. However, the reliability of the mobile agent leaves a great hidden peril for the network. For solving the high resource consumption of trust models in TWSNs, Shaikh et al. [21] proposed a lightweight group-based trust management scheme (GTMS). The trust score of a group of nodes is evaluated conjointly, instead of individual trust evaluation, to reduce the cost of trust records at each node. Nevertheless, the interaction between groups generates additional communication costs. Zhang et al. [22] presented a trust model for cloud environment to reduce management overhead and detect malicious nodes. In this scheme, nodes are partitioned into domains, to decrease storage and computational costs. Furthermore, a filter procedure is adopted to remove malicious trust evaluations and malicious nodes from a domain.

Another example is the trust management scheme for open systems introduced by Fan et al. [23], in which a feedback confirmation method – based on pairwise similarity – is developed to deal with dishonest ratings. In the same publication, a trust propagation strategy based on Susceptible-Infected-Recovered model is employed, to control the process of trust propagation. For the sake of secure data storage in cloud environment, Ghafoorian et al. [24] presented a role-based access control model based on trust and reputation, which provides accurate direct and indirect trust evaluation at the same time. The security goal that should be considered in an efficient trust-based system is presented, which has certain enlightening significance for the design of similar programs.

To achieve accurate and energy efficient trust evaluation in UASNs, an attack-resistant trust model based on multi-dimensional trust metrics (ARTMM) was proposed in [25]. In the ARTMM, the trust metrics including link trust, data trust, and node trust are analyzed to obtain an integrated trust. Moreover, the effect of communication channel and node mobility is utilized, as a factor to improve the accuracy

Fig. 2. Workflow schematic of the proposed TUMRL.

of trust evaluation. However, the impact of malicious attacks on the acquisition of trust metrics is neglected in this work.

Other researchers have also considered the uncertainties associated to the acoustic channel, dynamic network structure and weak link connectivity. For example, a trust model based on cloud theory (TMC) for UASNs was proposed in [26], which works as follows. First, in the process of generating trust evidences, malicious attacks are analyzed layer by layer. Then, the trust value is calculated based on the cloud model. The TMC effectively solves the ambiguity and uncertainty of trust. However, the lack of consideration for the changeable attack modes leaves hidden security risks. Therefore, we will consider an attack model with changing offensive modes in this paper. Additionally, we adopt a flexible trust update mechanism based on reinforcement learning, to incorporate adaptive properties to trust score evaluation.

The clustered network structure is of importance towards improving network scalability. However, assigning trust to third parties (cluster heads in unattended UASNs) generates new types of security challenges.

3 NETWORK MODEL AND ASSUMPTIONS

Because of the shortcomings present in clustered networks (Sec. 2), this study is based on a distributed UASN architecture (Fig. 1), where the underwater nodes are deployed at different depth levels by means of anchoring-buoys, and each node is allotted a unique ID. All underwater nodes are homogeneous – i.e., they all have the same but limited capabilities of energy, communication, computation, and storage. The locations of the underwater nodes are determined by established positioning algorithms [27, 28], and are periodically exchanged with neighboring nodes to update a neighbor table. The underwater nodes collect information, and send data packets to the surface sinks, by collaborating with neighbor nodes. Further, the data packets are transmitted to the base station on land via satellite relay. Since radio signals are strongly attenuated in saltwater, and often scattered by suspended particles, acoustic waves are the primary communication medium used in underwater environments.

As a result of the open and unattended nature of UASNs, they are highly vulnerable to offensive attacks. Once the attackers successfully invade a network, they can compromise and take control of normal nodes. Here it is assumed the worse case scenario where the attackers will first monitor and trace the network traffic, and select as targets the nodes situated in higher traffic hotspots. Moreover, the compromised nodes can launch attacks independently, and can also switch from one attack mode to another (DoS, selective forwarding, packet tampering, Sybil, wormhole attacks, etc.). As shown in Fig. 1, a normal node *B* is sending packets to a compromised node *A*. Assuming that *A* is performing a selective forwarding attack, the anomalous behavior of *A* can be detected by *B* through changes in packet delivery rate. If the attack mode of *A* is changed to packet tampering in a following cycle, the historical behavior evidence collected by *B* cannot truly reflect the attack type induced by *A*, which ultimately leads to inaccurate trust evaluation. In order to detect malicious nodes against internal attackers, each underwater node is embedded with a trust management program. The historical interaction information between nodes is used for trust calculation, and a sliding time window is maintained by each node during trust update (Fig. 5). The trust value is represented by a real number ranging from 0 to 1. The simulated network is initialized without compromised nodes, and with a trust score of 0.5.

4 TRUST EVALUATION ARCHITECTURE (TURML)

The overall workflow of the trust evaluation architecture – the TUMRL – is displayed in Fig. 2. In summary, the trust update cycle of a node starts by calculating the current environmental impact on trust score evaluation – according to the environment model – as an impact factor. If the impact factor is higher than a specified threshold, the trust update for the current time-slot is aborted. Otherwise, the interaction information stored in the sliding time window is used to compute the trust evidences, and the trust update based on reinforcement learning is executed. Finally, the key degree of nodes affects the probability of state transition, that is, the trust of the nodes with higher key degree is more

sensitive to offensive attacks, in order to minimize the attack damage.

In what follows, a detailed description of the building blocks of the TUMRL is given.

4.1 Environment Model

UASNs are usually deployed in complex, dynamic and harsh underwater environments. Thus the environmental conditions may highly affect the characteristics of node communication, data transmission and even network security. Therefore, it is necessary to consider the impact of underwater deployment environment on node trust update. In this section, we propose an environment model to quantify the impact of the underwater environment. The model quantifies the impact into two categories: the mobility of water flow and the instability of the acoustic communication. Based on the impact factors, a comprehensive environmental impact value is obtained, which regulates the trust update process.

4.1.1 The mobility of water flow

Under the influence of ocean currents, tides and other factors, underwater sensor nodes exhibit dynamic displacements in their relative positions. Hence a mobility model of underwater sensor nodes [29] is here adopted. As shown in Fig. 3, an underwater sensor node is anchored to a certain position O on the seabed by a tethered wire; O is the origin of the coordinate system. The position of the underwater sensor node and its subpoint on the undersurface are expressed as S and S' , respectively. The angle between the vector OS' and the X axis (φ), and the angle between the vector OS and the Z axis (θ) are also indicated. It is assumed that the speed of the underwater sensor node i at time-slot t is expressed as $v_i(t)$, which obeys the normal distribution $N(\mu_1, \sigma_1^2)$ and the effective range is truncated to $(0, 2\mu_1)$. The moving direction of node i at time-slot t is denoted as $(d\theta_i(t), d\varphi_i(t))$, where $d\theta_i(t)$ and $d\varphi_i(t)$ are assumed to obey uniform distributions; $U(0, \pi)$ and $U(0, 2\pi)$ respectively. The length of the tethered wire of node i is expressed as R_i . According to the current location $(x_i(t), y_i(t), z_i(t))$, the location of node i in next time slot is expressed as follows:

$$c_i(t+1) = \begin{cases} x_i(t) + v_i(t) \sin d\theta_i(t) \cos d\varphi_i(t) \\ y_i(t) + v_i(t) \sin d\theta_i(t) \sin d\varphi_i(t) \\ z_i(t) + v_i(t) \cos d\theta_i(t) \end{cases} \quad (1)$$

Due to the constraint imposed by the tethered wire ($|c_i(t) - O_i| \leq R_i$), the limit locus of node i in polar coordinates is $(R_i, \theta_i(t) + d\theta_i(t), \varphi_i(t) + d\varphi_i(t))$. The distance between node i and j is given by:

$$d_{ij}(t+1) = |c_i(t+1) - c_j(t+1)|. \quad (2)$$

The faster the distance between the two nodes changes with time, the stronger the impact of the mobility of water flow. Therefore, the impact of the mobility of water flow is defined as follows:

$$I_{mw} = \frac{|d_{ij}(t) - d_{ij}(t+1)|}{r}, \quad (3)$$

where r represents the communication radius of the underwater sensor nodes. When the numerator is greater than r ,

there is no communication between the two nodes, hence trust score update is not possible.

Fig. 3. Mobility model of underwater sensor node.

4.1.2 The instability of the acoustic communication

Instabilities in underwater acoustic communication are mainly caused by channel fading and environmental noise [30]. Underwater acoustic channels display both, distance and frequency selective fading. This is caused by spreading loss and by the signal-heat transfer increase – associated to larger transmission distances. Signal-heat transfer also increases with the transmission frequency. Based on the signal frequency f (in kHz) and the distance d_{ij} between transmitter i and receiver j , the attenuation occurring in an underwater acoustic channel can be estimated by the relation:

$$A(d_{ij}, f) = A_0 d_{ij}^k a(f)^{d_{ij}}, \quad (4)$$

where A_0 is the unit-normalizing constant, and k is the spreading factor, which is commonly set to 2 for spherical spreading, 1 for cylindrical spreading, and 1.5 for practical spreading. In addition, $a(f)$ (in dB/km) is the absorption coefficient, which can be estimated by a well-known empirical formula:

$$10 \log a(f) = 0.11 \frac{f^2}{1 + f^2} + 44 \frac{f^2}{4100 + f^2} + 2.75 \times 10^{-4} f^2 + 0.003. \quad (5)$$

Another factor that influences the quality of underwater acoustic channels is the plurality of noise sources present in underwater environments, which typically include: turbulence noise $N_t(f)$, ship-induced noise $N_s(f)$, wave motion noise $N_w(f)$ – caused by surface winds – and thermal noise $N_{th}(f)$. These noise categories can be described by Gaussian statistics, and estimated by the empirical relations:

$$\begin{aligned} 10 \log N_t(f) &= 17 - 30 \log f \\ 10 \log N_s(f) &= 40 + 20(s - 0.5) + 26 \log f - 60 \log(f + 0.03) \\ 10 \log N_w(f) &= 50 + 7.5w^{\frac{1}{2}} + 20 \log f - 40 \log(f + 0.4) \\ 10 \log N_{th}(f) &= -15 + 20 \log f, \end{aligned} \quad (6)$$

where $s \in [0, 1]$ is a factor representing the surface vessel activity frequency, and w is the wind speed (in m/s). The effective noise level at a frequency f is the sum of the contributions of the above factors:

$$N(f) = N_t(f) + N_s(f) + N_w(f) + N_{th}(f). \quad (7)$$

For moderate signaling bandwidth B and transmitted power P , the average signal to noise ratio (SNR) at the receiver is calculated as:

$$\text{SNR}(d_{ij}, f) = \frac{P/A(d_{ij}, f)}{N(f)B}, \quad (8)$$

with which the Rayleigh fading sub-channel – i.e., the probability of symbol error during the transmission – can be approximated:

$$P_e = \frac{1}{4\text{SNR}}. \quad (9)$$

The instability of underwater acoustic communication eventually causes errors in the packet transmission. Therefore, the impact of unstable acoustic communication is defined as follows:

$$I_{ac} = 1 - (1 - P_e)^N, \quad (10)$$

where N is the number of symbols in a packet.

Finally, according to the impact of the mobility of water flow and the instability of the acoustic communication, the integrated environmental impact I_e , that is, the output of the environment model, is defined:

$$I_e = k_1 I_{mw} + k_2 I_{ac}, \quad (11)$$

where $k_1, k_2 \in (0, 1)$ are the weights of the two factors, and $k_1 + k_2 = 1$.

4.2 Definition of Key Degree

UASNs are open systems, in which the communication between underwater nodes can be easily monitored by potential attackers through acoustic receivers. Therefore, an attacker can find hotspots in the network by monitoring, analyzing and tracing network traffic, and preferentially compromise nodes in the hotspots to optimize their offensive manoeuvres. In order to tackle the possibility of priority attacks, the concept of *key degree* is proposed to quantify the importance of a node to the network. Based on the difference in key degree, a differentiated trust update strategy is given in the next section, so as to achieve priority protection for the nodes with high importance.

Fig. 4 shows an example of node n_i which has only neighbor nodes n_1 , n_2 and n_3 . The nodes n_2 and n_3 are also neighbors to each other, so there is a direct acoustic link between them. Node n_1 cannot communicate directly with node n_2 , but it can communicate indirectly by routing through n_i , thereby there is an indirect link between them. Similarly, there is also an indirect link between node n_1 and node n_3 . It can be seen that once node i is removed from the network, the indirect links within the communication range will disappear. Therefore, the ratio of the number of indirect links to the total number of links is an indicator of the importance – or key degree – of node n_i .

Now, let's assume that the number of neighbors of node n_i is k , and the coordinates of the neighbors are

Fig. 4. Conceptual schematic of direct and indirect link connectivity with respect to a node n_i .

denoted as $C_i = \{c_1, c_2, \dots, c_k\}$. Within the communication range of node i , the total number of links – direct and indirect – can be computed by $\sum_{i=1}^{k-1} = \frac{k(k-1)}{2}$, and the set of direct links is expressed as $L = \{l_{mn} | |c_m - c_n| \leq r, \forall c_m, c_n \in C_i, m \neq n\}$, where l_{mn} represents the direct link between nodes m and n . Finally, the key degree of the node n_i is defined by

$$K_i = \frac{\frac{k(k-1)}{2} - |L|}{\frac{k(k-1)}{2}} = 1 - 2 \frac{|L|}{k(k-1)}, \quad (12)$$

where $|L|$ represents the cardinality of the set L .

4.3 Trust Update Based on Reinforcement Learning

The trust management system must not only confront challenges associated to the complexity of the underwater environment, but also various types of offensive attacks. Therefore, it is necessary to design an efficient and adaptive trust update mechanism. Reinforcement learning is a machine learning framework that has become ubiquitous in situations where online weight adaptation is required [31, 32]. In this case, the reinforcement learning model acquires information – to update the trust model parameters – by receiving environmental feedbacks iteratively. This characteristic makes it suitable for practical underwater environment, where the exact information is difficult to obtain. Therefore, reinforcement learning is introduced into the process of trust update. The process generally comprises three phases. First, an update regulation mechanism is presented to mitigate trust misclassification caused by environmental factors, rather than offensive attacks. Then, update benefit is accumulated based on the states of the evaluated nodes and the obtained trust evidences. Last, a state transition mechanism is elaborately designed to achieve optimal trust update.

4.3.1 Update regulation

As shown in Fig. 5, each node updates trust scores for neighbor nodes periodically, and each trust update cycle is divided into m update time-slots of equal length. That

Fig. 5. Trust update cycle with m update time-slots.

is, the node performs at most m update rounds in each trust update cycle. Each update time-slot records interaction information with other nodes during this period of time. Outside the trust update cycle, the node uses the results of the last trust update to evaluate other nodes. During each updating time-slot, the environmental impact I_e is first obtained by the TUMRL – according to the environment model mentioned in Sec. 4.1. If $I_e > \theta_e$ is satisfied, where $\theta_e \in (0, 1)$ represents the threshold of environmental stability, the updating time-slot is skipped, and the system proceeds to the next update time-slot. This is because when the environmental stability is poor – i.e., $I_e > \theta_e$ – the system could be affected by the environment and assign low trust scores to normal nodes.

If $I_e \leq \theta_e$ is satisfied, the trust evidences – including communication trust T_c , energy trust T_e and data trust T_d – are first calculated according to the relations (from our previous work in [33]):

$$T_c = \frac{2s + 1}{2(s + f + 1)} \quad (13a)$$

$$T_d = 2 \int_{v_j}^{\infty} f(v) dv \quad (13b)$$

$$T_e = \begin{cases} 0 & \text{if } E_{res} < \theta \\ 1 - |r_e - r_N| & \text{otherwise,} \end{cases} \quad (13c)$$

where $T_c, T_d, T_e \in [0, 1]$. s and f are the number of successful and failed communications in the time-slot, respectively. v_j represents the value in the packet received from the j -th neighbor, and $f(v)$ represents the degree of data differentiation. E_{res} is the surplus energy of the evaluated node, θ is the energy threshold, r_e is the current energy consumption rate and r_N is the normal energy consumption rate.

4.3.2 Benefit accumulation

After obtaining the trust evidences, the TUMRL performs trust score update through the reinforcement learning model. Specifically, it dynamically adjusts the weight proportion of the trust evidences to adaptively counter the unpredictable offensive attack modes mentioned in Sec. 3. The main process of trust update, which also occurs during a given time-slot is now described. First, a *reward* is obtained based on the current *state* and the latest performed *action*, and these values are fed to the Bellman Equation, in order to

TABLE 1
Parallel Table of States and Actions.

States	Expressions	Actions	Expressions
S_1	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	A_1	$\{T_c < 0.5, T_e \geq 0.5, T_d \geq 0.5\}$
S_2	$(\frac{1}{2}, \frac{1}{2}, 0)$	A_2	$\{T_c \geq 0.5, T_e < 0.5, T_d \geq 0.5\}$
S_3	$(\frac{1}{2}, 0, \frac{1}{2})$	A_3	$\{T_c \geq 0.5, T_e \geq 0.5, T_d < 0.5\}$
S_4	$(0, \frac{1}{2}, \frac{1}{2})$	A_4	$\{T_c < 0.5, T_e < 0.5, T_d \geq 0.5\}$
S_5	$(1, 0, 0)$	A_5	$\{T_c < 0.5, T_e \geq 0.5, T_d < 0.5\}$
S_6	$(0, 1, 0)$	A_6	$\{T_c \geq 0.5, T_e < 0.5, T_d < 0.5\}$
S_7	$(0, 0, 1)$	A_7	$\{T_c < 0.5, T_e < 0.5, T_d < 0.5\}$
		A_8	$\{T_c \geq 0.5, T_e \geq 0.5, T_d \geq 0.5\}$

update the accumulated benefit. Finally, the *state* is updated according to the latest accumulated benefit.

We now describe the states, actions, and rewards of the TUMRL in more detail. The weight proportion of the various examples of trust evidences are given in Table 1. The states, which are defined as $S_i, i = 1, 2, \dots, 7$, are the different combinations of the latest trust evidences. The actions are denoted by $A_j, j = 1, 2, \dots, 8$. For example, *state* = S_1 means that the current trust score of the evaluated node is $T = \frac{1}{3}T_c + \frac{1}{3}T_e + \frac{1}{3}T_d$, and *action* = A_1 means that the latest trust evidences satisfy $T_c < 0.5, T_e \geq 0.5$ and $T_d \geq 0.5$. The corresponding rewards table is also given (Table 2), in which $S_i A_j = v$ signifies that when the current *state* is S_i and the *action* is A_j , the *reward* is v .

TABLE 2
Reward Table.

R	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
S_1	2	2	2	1	1	1	0	10
S_2	1	1	0	2	0	0	1	10
S_3	1	0	1	0	2	0	1	10
S_4	0	1	0	0	0	2	1	10
S_5	2	0	0	1	1	0	1	10
S_6	0	2	0	1	0	1	1	10
S_7	0	0	2	0	1	1	1	10

Actions performed over different states produce different results, hence they have different rewards. For example, when the *state* is S_1 , *actions* A_1, A_2 , and A_3 have in common that there is only one trust evidence below the threshold 0.5 (Table 1). This situation is likely to be caused by a compromised node, so the *reward* is set to 2. In A_4, A_5 , and A_6 , there are two types of trust evidences below the threshold, which is less likely to occur because of compromised nodes – as these can only perform one type of attack at a time. The three types of trust evidences in A_7 are all below the threshold, so the *reward* is set to 0. The three types of trust evidence in A_8 are all above the threshold, which is the desired state corresponding to a healthy UASN, so the *reward* is set to 10.

After the *state*, *action* and *reward* in current time-slot are

obtained by the TUMRL, the Bellman Equation (Eq. 14) is employed to update the accumulated benefit:

$$Q(S, A) = (1 - \alpha) Q(S, A) + \alpha [R + \gamma \max_{A'} Q(S', A')], \quad (14)$$

where $\alpha \in [0, 1]$ is the learning rate – which weighs the past experience against the current learning – and $\gamma \in [0, 1]$ is the discount factor – indicating the importance of the future benefit. Additionally, $Q(S, A)$ represents the accumulated benefit, and $\max_{A'} Q(S', A')$ denotes the maximum accumulated benefit under all *actions* when the next *state* is S' .

4.3.3 State transition

In the TUMRL, the choice of the next *state* S' is of great significance for dealing with offensive attacks. The state transition mechanism consists of two stages: (1) the initial state transition, that occurs in the first effective update time-slot (the time-slot that is never skipped due to the environment model) of each update cycle; and (2), the probabilistic state transition that occurs in the subsequent update time-slots.

TABLE 3
Initial State Transition.

Next state	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
Non-key node	$S_1 S_2$	$S_1 S_2$	$S_1 S_3$	$S_1 S_2$	$S_1 S_3$	$S_1 S_4$	$S_1 S_2$	S_1
	$S_3 S_5$	$S_4 S_6$	$S_4 S_7$	$S_5 S_6$	$S_6 S_7$	$S_5 S_7$	$S_3 S_4$	
Key node	S_5	S_6	S_7	S_2	S_3	S_4	S_1	S_1

The initial state transition is shown in Table 3. According to the definition of key degree given in Sec. 4.2, the key node is defined as the evaluated node whose key degree is higher than the average key degree. Non-key nodes are those whose key degree is lower than the average. In Table 3, in most cases, non-key nodes can select one of the four *states* as the next *state* with equal probability, and the initial transition *state* of the key node is unique. The purpose of this is to increase the reaction rate of key nodes to attacks, in order to minimize the losses caused by compromised nodes.

As an example, assuming that there is a key node i and a non-key node j queuing for state transition, and that the current *state* of both nodes is S_1 , their latest *action* is A_1 , then the next *state* of the key node is S_5 , and the next *state* of the non-key node can be chosen to be S_2 (Table 3). According to the Table 1, *action* = A_1 means that $T_c < 0.5$, $T_e \geq 0.5$, and $T_d \geq 0.5$. Since only the communication trust T_c is below the threshold, it can be suspected that the evaluated node is performing an attack that causes communication failure, e.g. selective forwarding attack. In addition, from the *states* S_1 , S_5 and S_2 , we can infer that the trust scores in the current time-slot are $T_i^k = T_j^k = \frac{1}{3}T_c + \frac{1}{3}T_d + \frac{1}{3}T_e$, and for next time-slot are $T_i^{k+1} = T_c$ and $T_j^{k+1} = \frac{1}{2}T_c + \frac{1}{2}T_e$. According to $T_i^{k+1} - T_j^{k+1} = \frac{1}{2}(T_c - T_e) < 0$, it can be inferred that although the trust value of the key node and the non-key node is the same in the current time-slot, the trust value of the key node is lower than that of the non-key node in

the next time-slot, which shows that the key node presents lower time delays when reacting offensive attacks.

Although the initial state transition can quickly react to attacks, it is unable to cope with the changing attack mode mentioned in Sec. 3. In the subsequent time-slots of the update cycle, probabilistic state transition is used to address this problem. The main idea of the probabilistic state transition is to make a decision on the current state transition by using the accumulated benefit of the historical state transition, and the higher the accumulated benefit, the greater the probability that the corresponding *state* is selected. Suppose that the current *state* is S , and the latest *action* is A in the current time-slot (a time-slot after the first effective time-slot of the current update cycle), the probabilistic state transition is defined as follows:

$$P(S'|S, A) = \frac{Q(S', A)}{Q(\hat{S}, A)}, \quad (15)$$

where $P(S'|S, A)$ is the probability of transition from the *state* S to *state* S' when the *action* is A , $Q(\hat{S}, A)$ is the sum of the accumulated benefit of all *states* when the *action* is A , and $Q(S', A)$ is the accumulated benefit when the *action* is A and the *state* is S' .

A correct *state* selection can get a greater *reward*, thereby accelerating the benefit accumulation mechanism (Eq. 14). Therefore, the *state* with a high accumulated benefit is often the *state* that can effectively deal with a specific attack mode. Although a compromised node may change its attack mode in different cycles, the result of the probabilistic state transition makes the TUMRL to show preference towards *states* with greater accumulated benefit. Therefore, the probabilistic state transition can effectively deal with the changing attack mode of the compromised node.

5 SIMULATION RESULTS AND ANALYSIS

In this study, the proposed TUMRL, and other established trust models, were simulated on MATLAB R2017a to evaluate and compare their performance. First, the performance of the different mechanisms in the TUMRL was evaluated. Specifically, both the impact of the key degree mechanism, on the reaction rate to offensive attacks, and the impact of the environment model on the detection accuracy rate of compromised nodes were verified. Then, the performance of the TUMRL was compared with other related work: the ARTMM [25], TMC [26] and BLTM [19] trust models. To the best of our knowledge, the ARTMM and the TMC represent the state-of-the-art in trust modelling for UASNs. The BLTM has a simple update mechanism, which partly has motivated this study. The performance was compared in terms of detection accuracy rate, false alarm rate and energy efficiency. The deployment area was set to $500 \times 500 \times 500 \text{ m}^3$, from which 500 sensor nodes were randomly deployed in the area, and the communication radius of sensor nodes was set to 100 m.

5.1 Performance of TUMRL

5.1.1 Evaluation of Key Degree

The efficacy of key degree assignment during trust updates was evaluated, by measuring the time delay required by the

Fig. 6. Efficacy of key degree assignment on the TUMRL under various attack modes.

Fig. 7. Performance of the TUMRL, with and without environment model, for an increasing proportion of compromised nodes.

TUMRL to react to attacks. Three typical attack modes were simulated to evaluate the performance of key nodes and non-key nodes. The attack modes were selective forwarding, DoS and packet tampering, which mainly affect communication success rate, energy consumption rate and packet error rate respectively. The nodes operated normally from 0 to 100s, when attacks were launched by the compromised nodes during subsequent 50s. The trust value decreased faster than that of the non-key nodes during the attack (Fig. 6). This is because the state transition mechanism makes key nodes more sensitive to anomalous behavior, and the attack consequences are mitigated by quickly reducing their trust value. In the period from 100 to 150s, the curve of key node in Fig. 6(a) is smoother than that in the other two figures. The reason is that the selective forwarding attack not only affects communication success rate, but also causes abnormal energy consumption. Thus, the key node in Fig. 6(a) has more choices for transitioning among states.

5.1.2 Evaluation of Environment model

In order to evaluate the performance of the environment model, a set of controlled experiments was designed. The detection accuracy rate was evaluated with an increasing proportion of compromised nodes, under two conditions: with and without environment model. The detection ac-

curacy rate was taken to be the *true positive rate* of compromised node classification. In Fig. 7, it is clear that the trust model with environment model outperforms the trust model without environment model. This is because the environment model can quantify the impact of the underwater fluctuations, to reduce the frequency of trust update when these are large, thereby alleviating the trust decline caused by environmental factors effectively. In addition, the environmental models under different positioning accuracy (100%, 85%, 70%) are simulated to observe how errors in the underwater positioning algorithms would affect the proposed security algorithm. As shown in Fig. 7, the detection accuracy rate decreases with the reduced positioning accuracy. This is because the positioning accuracy directly affects the mobility factor I_{mw} in the environmental model, to increase the decision error of the environmental model, thereby reducing the detection accuracy to compromised nodes.

5.2 Comparison with related work

5.2.1 Comparison of Detection Accuracy Rate

The TUMRL was compared with the TMC, ARTMM and BLTM in terms of detection accuracy rate. The detection accuracy rate was measured under the simulated condition that 30% randomly selected nodes are compromised. As shown in Fig. 8, the TUMRL performs worse than the other schemes before about 180s. The memory properties introduced by the reinforcement learning method in the TUMRL imply that the TUMRL states need to be populated, before achieving sufficient evidence accumulation. Therefore, the performance of the TUMRL is relatively poor initially, but it shortly reaches higher performance. Thus the TUMRL is more robust in a realistic setting, where importance is on the long term credibility of the sensor data.

In order to assess the impact on trust evaluation caused by the changing attack mode, the performance of the detection accuracy rate was simulated with changing frequency between attack modes. For example, a changing frequency of 2 means that the compromised node changes the attack mode twice during each time cycle. As shown in Fig. 9, when the changing frequency gradually increases from 0 to 2, the detection accuracy rate of the TUMRL presents a slower decline as compared to the other trust models. Once the attack mode of a compromised node changes, the difference in performance of this node can be detected by

Fig. 8. Detection accuracy vs. simulation time of the examined trust models (TUMRL, TMC, ARTMM and BLTM).

its neighbor nodes, which is reflected in the trust evidence in the TUMRL. Moreover, the weight of the trust evidences can be flexibly changed to confront the unpredictable attack mode. Therefore, the simulations suggest that the performance of the TUMRL is more robust in detecting compromised nodes, under the condition that the attack mode is changeable.

Fig. 9. Detection accuracy rate with changing frequency of attack mode.

5.2.2 Comparison of False Alarm Rate

For determining the impact of the dynamic environment on trust evaluation, the TUMRL was compared with the TMC and ARTMM by evaluating their false alarm rate – under the simulated condition of different moving speeds between the sensor nodes¹. As shown in Fig. 10, the simulations produce lower false alarm rate for the TUMRL, when the speed of sensor nodes was gradually increased from 0 to 50 m/s. Thus the use of an environment model seems to be an effective extension of trust models, when monitoring underwater dynamic environments. As the speed of the sensor nodes

1. Since the BLTM is designed for TWSN, it was not considered in this comparison.

Fig. 10. False alarm rate vs. moving speed of sensor nodes for the compared models.

increases, so does the response of the environment model. Once the output exceeds the specified threshold, the trust scores are not updated during the current time-slot. Hence the TUMRL distinguishes the impact originated from attacks from environment fluctuations. However, there are no effective mechanisms in the TMC and ARTMM to complete a similar distinction. Therefore, when the node movement speed increases, that is, the impact of the environment on the trust evaluation increases, the false alarm rate gradually increases in the TMC and ARTMM.

Fig. 11. False alarm rate vs. network density.

An additional simulation was implemented to exhibit the effect of the distribution density of UASN on the false alarm rate. As shown in Fig. 11, the TUMRL outperforms other schemes when the distribution density of sensor nodes is greater than 100. However, the performance of false alarm rate for the TUMRL is worse than that of the other schemes for smaller densities. The reason is that the sensor nodes lack sufficient information for trust score evaluation when the distribution density of sensor nodes is relatively low. Further, the TURML has higher requirements on the amount of data. Hence the TUMRL is more robust in compact

Fig. 12. Energy efficiency with different proportion of compromised nodes.

networks, where information between the nodes highly contributes to the overall health factor of the network.

5.2.3 Comparison of Energy Efficiency

With respect to energy efficiency, the four models (TUMRL, TMC, ARTMM and BLTM) were compared under different percentages of compromised nodes and network densities. Since the energy consumption of a node, when receiving or sending packets, is much greater than the energy consumption for storage and computation, only the communication energy consumption is considered here. The energy efficiency is defined as the ratio of the remaining energy of the network over the initial energy after detecting all the compromised nodes.

As shown in Fig. 12, when the proportion of malicious nodes is less than 8%, the energy efficiency of the TUMRL is slightly inferior to that of the other three schemes. As stated, the reinforcement learning adopted in the TUMRL is more effective in dense attacks; few attacked nodes only produce a small increase in energy expenditure. However, as the proportion of compromised nodes increases, the energy consumption caused by attacked nodes far exceeds the computational costs of the reinforcement learning model. The energy efficiency of the TUMRL is better than that of the other schemes when the proportion of compromised nodes is higher than 12%.

The energy efficiency was also analyzed under different network densities (Fig. 13). Although the performance of the TUMRL is similar to the other schemes – and even slightly lower when the distribution density of sensor nodes is less than 90 nodes/km³. For higher densities, the TUMRL outperforms the other methods. As before the reason is that the reinforcement learning algorithm in the TUMRL requires sufficient data.

6 CONCLUSION

This study explores the problem of effective trust update, while facing unstable underwater environment fluctuations and offensive manoeuvres – consisting of switching attack

Fig. 13. Energy efficiency with different distribution density.

modes. A novel trust update mechanism based on reinforcement learning is proposed. In the scheme, the impact of underwater environment is first analyzed and quantified by an environment model. The quantified result is used to regulate and improve the trust score update mechanism. Further, the concept of *key degree* is introduced. The key degree of the sensor nodes determines their relative priority during trust score update. The trust score update is completed by way of a reinforcement learning model, which integrates the role of the environment model and the key degree – to adapt to switching attack modes and anomalous behaviour from compromised nodes. Moreover, the experimental analysis indicates that the TUMRL offers increased performance with respect to previously established work, in terms of complex underwater environment management and defective node detection. This is particularly true when measuring the long term performance of the TUMRL, in networks with relatively high density, networks subject to widespread global attacks and varying attack conditions.

ACKNOWLEDGMENTS

The work is supported by the National Key Research and Development Program, No.2018YFC0407900 and the National Natural Science Foundation of China under Grant No. 61971206, the Open fund of State Key Laboratory of Acoustics under Grant SKLA201901.

REFERENCES

- [1] J. Heidemann, M. Stojanovic, and M. Zorzi, "Underwater sensor networks: applications, advances and challenges," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 370, no. 1958, pp. 158–175, Jan. 2012.
- [2] M. Erol-Kantarci, H. T. Mouftah, and S. Oktug, "A survey of architectures and localization techniques for underwater acoustic sensor networks," *IEEE Communications Surveys & Tutorials*, vol. 13, no. 3, pp. 487–502, Mar. 2011.
- [3] C. Lin, G. Han, M. Guizani, Y. Bi, and J. Du, "A scheme for delay-sensitive spatiotemporal routing in

- sdn-enabled underwater acoustic sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 9280–9292, Jul. 2019.
- [4] R. Diamant, P. Casari, and S. Tomasin, "Cooperative authentication in underwater acoustic sensor networks," *IEEE Transactions on Wireless Communications*, vol. 18, no. 2, pp. 954–968, Dec. 2018.
- [5] E. M. Sozer, M. Stojanovic, and J. G. Proakis, "Underwater acoustic networks," *IEEE journal of oceanic engineering*, vol. 25, no. 1, pp. 72–83, Jan. 2000.
- [6] D. Pompili, T. Melodia, and I. F. Akyildiz, "A cdma-based medium access control for underwater acoustic sensor networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 4, pp. 1899–1909, May 2009.
- [7] K. Chen, M. Ma, E. Cheng, F. Yuan, and W. Su, "A survey on mac protocols for underwater wireless sensor networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1433–1447, Mar. 2014.
- [8] S. Jiang, "State-of-the-art medium access control (mac) protocols for underwater acoustic networks: A survey based on a mac reference model," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 1, pp. 96–131, Nov. 2017.
- [9] M. Xu and L. Liu, "Sensevault: A three-tier framework for securing mobile underwater sensor networks," *IEEE Transactions on Mobile Computing*, vol. 17, no. 11, pp. 2632–2645, Mar. 2018.
- [10] G. Han, J. Jiang, N. Sun, and L. Shu, "Secure communication for underwater acoustic sensor networks," *IEEE communications magazine*, vol. 53, no. 8, pp. 54–60, Aug. 2015.
- [11] X. Chen, K. Makki, K. Yen, and N. Pissinou, "Sensor network security: a survey," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 2, pp. 52–73, Jun. 2009.
- [12] D. He, N. Kumar, J. Chen, C.-C. Lee, N. Chilamkurti, and S.-S. Yeo, "Robust anonymous authentication protocol for health-care applications using wireless medical sensor networks," *Multimedia Systems*, vol. 21, no. 1, pp. 49–60, 2015.
- [13] Y. Xiao, M. Peng, J. Gibson, G. G. Xie, D.-Z. Du, and A. V. Vasilakos, "Tight performance bounds of multihop fair access for mac protocols in wireless sensor networks and underwater sensor networks," *IEEE Transactions on Mobile Computing*, vol. 11, no. 10, pp. 1538–1554, Sept. 2011.
- [14] S. Ganeriwal, L. K. Balzano, and M. B. Srivastava, "Reputation-based framework for high integrity sensor networks," *ACM Transactions on Sensor Networks (TOSN)*, vol. 4, no. 3, pp. 1–37, Jun. 2008.
- [15] Z. Yao, D. Kim, and Y. Doh, "Plus: Parameterized and localized trust management scheme for sensor networks security," in *2006 IEEE International Conference on Mobile Ad Hoc and Sensor Systems*. IEEE, Oct. 2006, pp. 437–446.
- [16] R. Feng, X. Xu, X. Zhou, and J. Wan, "A trust evaluation algorithm for wireless sensor networks based on node behaviors and ds evidence theory," *Sensors*, vol. 11, no. 2, pp. 1345–1360, Jan. 2011.
- [17] Y. Ren, V. I. Zadorozhny, V. A. Oleshchuk, and F. Y. Li, "A novel approach to trust management in unattended wireless sensor networks," *IEEE Transactions on Mobile Computing*, vol. 13, no. 7, pp. 1409–1423, Feb. 2013.
- [18] J. Jiang, G. Han, F. Wang, L. Shu, and M. Guizani, "An efficient distributed trust model for wireless sensor networks," *IEEE transactions on parallel and distributed systems*, vol. 26, no. 5, pp. 1228–1237, Apr. 2014.
- [19] X. Wu, J. Huang, J. Ling, and L. Shu, "Bltm: beta and lqi based trust model for wireless sensor networks," *IEEE Access*, vol. 7, pp. 43 679–43 690, Mar. 2019.
- [20] A. Boukerch, L. Xu, and K. El-Khatib, "Trust-based security for wireless ad hoc and sensor networks," *Computer Communications*, vol. 30, no. 11-12, pp. 2413–2427, Sept. 2007.
- [21] R. A. Shaikh, H. Jameel, B. J. d'Auriol, H. Lee, S. Lee, and Y.-J. Song, "Group-based trust management scheme for clustered wireless sensor networks," *IEEE transactions on parallel and distributed systems*, vol. 20, no. 11, pp. 1698–1712, Dec. 2008.
- [22] P. Zhang, Y. Kong, and M. Zhou, "A domain partition-based trust model for unreliable clouds," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 9, pp. 2167–2178, Mar. 2018.
- [23] X. Fan, L. Liu, M. Li, and Z. Su, "Grouptrust: dependable trust management," *IEEE Transactions on Parallel and Distributed Systems*, vol. 28, no. 4, pp. 1076–1090, Sept. 2016.
- [24] M. Ghafoorian, D. Abbasinezhad-Mood, and H. Shakeri, "A thorough trust and reputation based rbac model for secure data storage in the cloud," *IEEE Transactions on Parallel and Distributed Systems*, vol. 30, no. 4, pp. 778–788, Sept. 2018.
- [25] G. Han, J. Jiang, L. Shu, and M. Guizani, "An attack-resistant trust model based on multidimensional trust metrics in underwater acoustic sensor network," *IEEE Transactions on Mobile Computing*, vol. 14, no. 12, pp. 2447–2459, Feb. 2015.
- [26] J. Jiang, G. Han, L. Shu, S. Chan, and K. Wang, "A trust model based on cloud theory in underwater acoustic sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 1, pp. 342–350, Dec. 2015.
- [27] W. Zhang, G. Han, X. Wang, M. Guizani, K. Fan, and L. Shu, "A node location algorithm based on node movement prediction in underwater acoustic sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 3166–3178, Jan. 2020.
- [28] T. Zhang, G. Han, L. Yan, B. Liu, and L. Shu, "Optimal design of beacon array for long baseline positioning system used in manned deep-sea submersibles," *IEEE Access*, vol. 7, pp. 140 411–140 420, Sept. 2019.
- [29] H. Chang, J. Feng, and C. Duan, "Reinforcement learning-based data forwarding in underwater wireless sensor networks with passive mobility," *Sensors*, vol. 19, no. 2, p. 256, Jan. 2019.
- [30] G. Han, S. Li, C. Zhu, J. Jiang, and W. Zhang, "Probabilistic neighborhood-based data collection algorithms for 3d underwater acoustic sensor networks," *Sensors*, vol. 17, no. 2, p. 316, Feb. 2017.
- [31] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.
- [32] L. P. Kaelbling, M. L. Littman, and A. W. Moore,

“Reinforcement learning: A survey,” *Journal of artificial intelligence research*, vol. 4, pp. 237–285, 1996.

- [33] G. Han, Y. He, J. Jiang, N. Wang, M. Guizani, and J. A. Ansere, “A synergetic trust model based on svm in underwater acoustic sensor networks,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 11, pp. 11 239–11 247, Nov. 2019.

Hao Wang is currently pursuing his Ph.D. degree from the Department of Computer Science and Technology at Hohai University. He received his B.S. degree in Electronic Information Science and Technology from Nanjing Agricultural University, China, in 2015. His current research interests include location privacy protection.

Yu He is currently pursuing his Ph.D. degree from the the Department of Information and Communication Systems at Hohai University. He received his B.S. degree in Internet of Things Engineering from Hohai University, Nanjing, China, in 2017. His current research interests include security of underwater acoustic sensor networks.

Miguel Martínez-García received a BSc degree in Mathematics from the Polytechnic University of Catalonia (UPC), Spain in 2013, a MSc in Advanced Mathematics and Mathematical Engineering (MAMME) from the same university in 2014, and a PhD in Engineering from the University of Lincoln, UK, in 2018. He also worked as a researcher since 2017, both at the University of Lincoln and at the Advanced Virtual Reality Research Centre (AVRRC) in Loughborough, UK. In 2020 he became a Lecturer in Human-

machine Systems at Loughborough University. His research interests include human-machine integration, machine learning, artificial intelligence, intelligent signal processing, and complex systems – with particular focus in the analysis of non-linear signals representing phenomena of interest between humans and machines.

Guangjie Han [S'03, M'05, SM'18] is currently a Professor with the Department of Information and Communication System, Hohai University, Changzhou, China and a Distinguished Professor of Dalian University of Technology, Dalian, China. He received the Ph.D. degree from Northeastern University, Shenyang, China, in 2004. In February 2008, he finished his work as a Postdoctoral Researcher with the Department of Computer Science, Chonnam National University, Gwangju, Korea. From October 2010 to October 2011, he was a Visiting Research Scholar with Osaka University, Suita, Japan. From January 2017 to February 2017, he was a Visiting Professor with City University of Hong Kong, China. His current research interests include Internet of Things, Industrial Internet, Machine Learning and Artificial Intelligence, Mobile Computing, Security and Privacy. Dr. Han has over 400 peer-reviewed journal and conference papers, in addition to 160 granted and pending patents. Currently, his H-index is 43 and i10-index is 141 in Google Citation (Google Scholar). Total citation of his papers by other people is more than 6900+ times.

Dr. Han is a Fellow of the UK Institution of Engineering and Technology (FIET). He has served on the Editorial Boards of up to 10 international journals, including the IEEE Network, IEEE Systems, IEEE/CCA JAS, IEEE ACCESS, Telecommunication Systems, etc. He has guest edited a number of special issues in IEEE Journals and Magazines, including the IEEE JSAC, IEEE Communications, IEEE Wireless Communications, IEEE Transactions on Industrial Informatics, Computer Networks, etc. Dr. Han has also served as chairs of organizing and technical committees of many international conferences. He had been awarded 2020 IEEE Systems Journal Annual Best Paper Award, 2017-2019 IEEE ACCESS Outstanding Associate Editor Award. He is a Senior Member of IEEE.

Jinfang Jiang is currently an associate professor with the Department of Information and Communication System, Hohai University, Nanjing, China. She received her Ph.D degree in Information and Communication Engineering from Hohai University, China, in 2015. Her current research interests include wireless sensor networks, mobile cloud computing, and information security technologies.