applied sciences

MDPI

Article

Enhancing Unmanned Marine Vehicle Security: A Periodic Watermark-Based Detection of Replay Attacks

Guangrui Bian [1] and Xiaoyang Gao [2,*]

[1] School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China; bianguangrui@std.uestc.edu.cn
[2] School of Maritime Economics and Management, Dalian Maritime University, Dalian 116026, China
* Correspondence: gaoxiaoyang@uestc.edu.cn

Abstract: This paper explores a periodic watermark-based replay attack detection method for Unmanned Marine Vehicles modeled in the framework of the Takagi–Sugeno fuzzy system. The precise detection of replay attacks is crucial for ensuring the security of Unmanned Marine Vehicles; however, traditional timestamp-based or encoded measurement-dependent detection approaches often sacrifice system performance to achieve higher detection rates. To reduce the potential performance degradation, a periodic watermark-based detection scheme is developed, in which a compensation signal together with a periodic Gaussian watermark signal is integrated into the actuator. By compensation calculations conducted with all compensatory signals in each period, the position corresponding to a minimum value of the detection function can be derived. Then, the time that the attacks occurred can be ensured with the aid of the comparison between this position with the watermark signal in the same period. An application on a UMV is shown to demonstrate the effectiveness of the presented scheme in detecting replay attacks while minimizing control costs.

Keywords: unmanned marine vehicles; replay attack detection; periodic watermark signal; Takagi–Sugeno fuzzy system

check for updates

Citation: Bian, G.; Gao, X. Enhancing Unmanned Marine Vehicle Security: A Periodic Watermark-Based Detection of Replay Attacks. *Appl. Sci.* **2024**, *14*, 8298. https://doi.org/10.3390/app14188298

Academic Editor: Alessandro Lo Schiavo

Received: 21 August 2024
Revised: 7 September 2024
Accepted: 13 September 2024
Published: 14 September 2024

1. Introduction

Unmanned Marine Vehicles (UMVs) represent a category of highly automated, intelligent, cost-effective, and efficient maritime vessels that operate without human intervention, which have been found to have widespread applications in the area of the observation of marine hydrological data, detection of marine environments, and exploration of marine resources etc. [1]. Their developments have significantly advanced marine technologies [2–4], and also showcased enormous potential in scientific research, military applications [5,6], and civilian uses. In contrast to traditional vessels, UMVs can operate under harsh weather conditions and undertake monotonous and hazardous operations in complicated marine environments [7]. Such versatility has fostered the development and research within the domains of navigation control, fleet coordination, trajectory tracking control etc. [8,9].

The rapid development of UMVs has been inseparable from the significant growth of modern network communications [10]. UMVs are transitioning from reliance on traditional communication channels to integrated networks to enhance system efficiency. However, this transition also introduces significant cybersecurity challenges, such as network attacks [11,12]. Network attacks are not only increasingly common but also growing in complexity, leading to substantial risks to the UMVs systems [13,14]. Hence, the security challenges faced by UMVs are crucial, and the detection of attacks has significant research significance. In an open network environment, cybersecurity problems targeting UMV systems primarily manifest as denial of service (DoS) attacks and replay attacks [15]. Extensive investigations by Zhang et al. [16] and Fei et al. [17] have concentrated on countering DoS attacks on UMV systems. Unfortunately, unlike DoS attacks, replay attackers do not require prior knowledge of the control systems. These attackers typically retransmit system measurement data from periods of stability, rendering the attacks inherently covert and difficult to detect through standard methods [18]. The covert nature of replay attacks makes them challenging to detect with traditional methods, which contributes to their frequent occurrence in contemporary UMV systems.

The rising occurrence of replay attacks has drawn the interest of scholars, especially in the aspect of attack detection. Existing methods primarily include timestamp techniques and the addition of physical watermarks to control signals. Timestamp-based replay attack detection was first proposed in [19], where Kerberos leverages a reputable third-party authentication service to authenticate user identities. Subsequently, Greene et al. [20] proposed an enhanced remote keyless entry system employing timestamping and exclusive OR (EOR) encoding to robustly defend against replay attacks, significantly improving security over systems. Farha et al. [21] introduced a timestamp-based method to enhance ZigBee network security against replay attacks, ensuring effectiveness without impacting network efficiency. Furthermore, a Promela-based model checking approach for verifying timestamp-utilizing security protocols against replay attacks, effectively identifying vulnerabilities in the protocol, was proposed in [22]. Nonetheless, it had been suggested by Yang et al. [23] and Zhu et al. [24] that transmitted signals are vulnerable to interception by replay attackers, who can alter the timestamps, leading to noticeable decline in the efficiency of the detection strategy mentioned in [19–22]. To address this issue, Du et al. [25] added the continuous addition of Gaussian random noise to the control signals, causing noticeable variations in data residuals before and after an attack, therefore facilitating the detection of such attacks. Then, an approach for designing optimal physical watermark signals to detect replay attacks in linear time-invariant systems with unknown parameters was studied in [26]. These strategies adjusted the system's optimal control signals, trading off control costs for improved detection of attacks. Therefore, an optimized watermarking method that balances improved detection of replay attacks against control costs in networked control systems was discussed in [27]. Ma et al. [28] presented a watermarking method for efficiently detecting transient covert attacks in industrial cyber-physical systems, reducing system performance cost while detecting detection accuracy. A recursive watermark algorithm significantly safeguarding against replay attacks with remarkable efficiency and saving the control cost by embedding rapid, detectable watermarks in hard real-time channels was introduced in [29]. In [30], a watermark-based encoding and decoding method were also proposed to detect replay attacks on UMVs with less control cost. It is worth noting that while incorporating watermark signals into control systems of UMVs aids in replay attack detection, the approach also compromises control performance [25–30]. The advantages and limitations of different existing methods can be seen in Table 1. Therefore, it is interesting to study the schedule of watermarks to use control costs more effectively.

Table 1. Advantages and Limitations of Different Existing Methods.

Detection Method	Timestamp	Watermark	Vulnerable	Precise	Cost-Effective
Ref. [19]	✓		✓		
Ref. [20]	✓		✓		
Ref. [25]		✓		✓	
Ref. [26]		✓		✓	
Ref. [27]		✓		✓	
Ref. [27]		✓		✓	
Ref. [29]		✓		✓	
Proposed Method		✓		✓	✓

Inspired by the above-mentioned discussions, in this article, a periodic watermark signal method against replay attacks for UMV systems is investigated, which can precisely detect replay attacks and minimize the impact of the watermark signal on the control system's performance. Significant contributions of this paper are highlighted as follows.

(1) The proposed method utilizes the relative positions of watermark signals to detect replay attack, diverging from traditional watermarking methods in [30,31]. Compared with existing, this method significantly lowers control costs by reducing watermark signal variance.

(2) A novel periodic watermarking-compensation mechanism is constructed based on a Gaussian signal, contrasting with binary watermark signals employed in [32]. Comparatively, this mechanism can effectively enhance the detection rate of replay attacks.

Throughout this article, the superscripts 'T' and '-' denote the transposition and the inverse of a matrix.

2. Problem Formulation and Preliminaries

The considered system Structure is presented in Figure 1, which is comprised of UMVs, a Kalman filter, and an LQG controller for setpoint optimization. The following analysis will explore the Takagi–Sugeno (T-S) fuzzy modeling for UMVs, the replay attacks, and the applications of Kalman filters and LQG controllers.

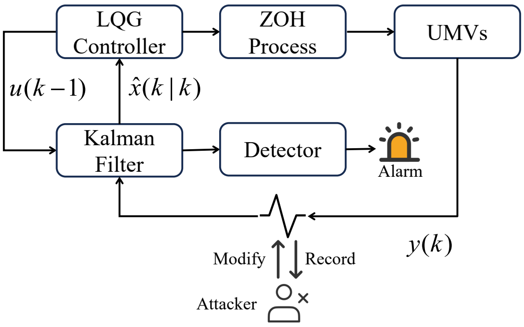

Figure 1. The UMVs diagram.

2.1. UMVs System Modeling

It follows from [33–35] that the equations describing dynamics and kinematics of the UMVs shown in Figure 2 are depicted as follows:

$$\begin{aligned} &\mathbb{M}\dot{\tau}(t)+\mathbb{N}\tau(t)+\mathbb{O}\varphi(t)=u(t)+\theta(t)\\ &\dot{\varphi}(t)=\mathbb{Z}(\omega(t))\tau(t) \end{aligned} \tag{1}$$

where $\tau(t) = \left[\tau_{surge}(t)\ \ \tau_{sway}(t)\ \ \tau_{yaw}(t)\right]^T$ denotes the vector of body-fixed linear and angular velocities, where $\tau_{\text{surge}}(t)$ stands for surge velocity, $\tau_{\text{sway}}(t)$ for sway velocity, and $\tau_{\text{yaw}}(t)$ for yaw velocity. $\varphi(t) = [x(t)\ \ y(t)\ \ \omega(t)]^T$ indicates the earth-fixed orientation vector, where $x(t)$ and $y(t)$ are positions, and $\omega(t)$ denotes the yaw angle. Control inputs are given by $u(t) = [u_1(t)\ \ u_2(t)\ \ u_3(t)]^T$. $\theta(t) = [\theta_1(t), \theta_2(t), \theta_3(t)]^T$ encapsulates wave-induced disturbances. $\mathbb{O} = \text{diag}\{O_{11}, O_{22}, O_{33}\}$ stands for mooring forces.

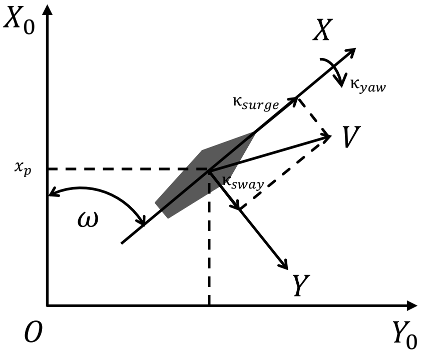

Figure 2. Reference Reference frames fixed to the Earth and the body.

Next, the UMVs system can be described as:

$$\dot{\tau}(t) = \mathbb{A}\varphi(t) + \mathbb{B}\tau(t) + \mathbb{D}u(t) + \mathbb{D}\theta(t) \tag{2}$$

where $\mathbb{A} = -\mathbb{M}^{-1}\mathbb{O}, \mathbb{B} = -\mathbb{M}^{-1}\mathbb{N}, \mathbb{D} = \mathbb{M}^{-1}$. In constructing the fuzzy dynamic model of the UMVs system, the yaw angle serves as the premise variable. In this case, Three pivotal yaw angle points $\omega(t)$ are selected at the origin and $\pm\pi/6$, respectively, and linearize the system at those three operating points, the fuzzy model of UMVs is given as follows:

R^1 : IF $\omega(t)$ is about 0 THEN

$$\begin{cases} \dot{x}(t) = A_1 x(t) + Bu(t), \\ y(t) = Cx(t). \end{cases} \tag{3}$$

R^2 : IF $\omega(t)$ is about $\pi/6$ THEN

$$\begin{cases} \dot{x}(t) = A_2 x(t) + Bu(t), \\ y(t) = Cx(t). \end{cases} \tag{4}$$

R^3 : IF $\omega(t)$ is about $-\pi/6$ THEN

$$\begin{cases} \dot{x}(t) = A_3 x(t) + Bu(t), \\ y(t) = Cx(t). \end{cases} \tag{5}$$

where

$$A_p = \begin{bmatrix} 0_{3\times 3} & Q_i \\ \mathbb{B} & \mathbb{A} \end{bmatrix}, \ B = \begin{bmatrix} 0_{3\times 3} \\ \mathbb{D} \end{bmatrix}, \ C = \begin{bmatrix} I_{3\times 3} & 0_{3\times 3} \end{bmatrix}, Q_1 = I_{3\times 3}, \ Q_2 = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ Q_3 = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Given that the position and angle can be measured with relative ease, the output matrix C is selected accordingly. Furthermore, the membership functions for the three rules are illustrated in Figure 3.

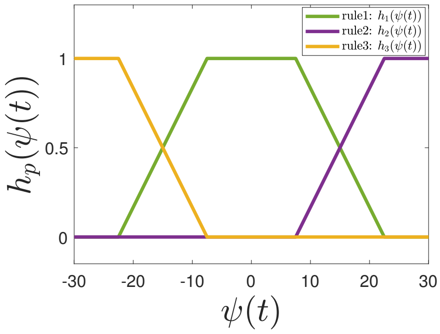

Figure 3. Membership functions.

The global T-S fuzzy model of the UMVs is derived using a singleton fuzzifier, product inference, and center average defuzzifier as described below:

$$\begin{cases} \dot{x}(t) = \sum_{p=1}^{3} \hbar_p(\varrho(t))(A_p x(t) + D_p u(t) + \theta(t)) \\ y(t) = Cx(t) + v(t) \end{cases} \tag{6}$$

where

$$\hbar_p(\varrho(t)) = \frac{v_p(\varrho(t))}{\sum_{p=1}^{3} v_p(\varrho(t))}, v_p(\varrho(t)) = M_{p1}(\varrho_1(t))M_{p1}(\varrho_2(t)), \hbar_p(\varrho(t)) \geq 0, \sum_{p=1}^{3} \hbar_p(\varrho(t)) = 1.$$

The discrete-time fuzzy system results from the periodic sensor sampling with a period of T_z and the application of the ZOH input mechanism:

$$\begin{cases} x(k+1) = \sum_{p=1}^{3} \hbar_p(\varrho(k))(A_{dp}x(k) + B_{dp}u(k) + \theta_{dp}(k)), \\ \quad y(k) = Cx(k) + v(k) \end{cases} \tag{7}$$

where $A_{dp} = e^{A_n T}$, $B_{dp} = \int_0^T e^{A_n T} B_n \, dt$, and C are defined as in Equation (6). The function $\hbar_p(\varrho(k))$ represents the value of $\hbar_p(\varrho(t))$ at the sampling instance $t = k$. The variables $\theta_{dp}(k)$ and $v(k)$ represent process noise and measurement noise, respectively. These noise components are mutually independent, with W and V serving as their respective covariance matrices.

In a UMVs system, optimal state estimation $\hat{x}(k|k)$ can be achieved by implementing a Kalman filter. Sensor data are forwarded to an estimator that applies the Kalman filter for estimating the state. The system is presumed to initiate operations at $-\infty$, allowing for the convergence of the Kalman filter into a constant-gain linear estimator. The update process is outlined in Algorithm 1.

Algorithm 1 Kalman Filter Utilized in T-S UMVs Model

```
Step 1: Input $A_{dp}, B_{dp}, C, W, V$
Step 2: $\hat{x}(k+1|k) = \sum_{p=1}^{3} \hbar_p(\varrho(k))(A_{dp}\hat{x}(k|k) + B_{dp}u(k))$
Step 3: $\hat{x}(k|k) = \hat{x}(k|k-1) + Kz(k)$
Step 4: $z(k) = y(k) - C\hat{x}(k|k-1)$
Step 5: $L = PC^T(CPC^T + R)^{-1}$
Step 6: $P = \sum_{p=1}^{3}(A_{dp}PA_{dp}^T + W) - A_{dp}PC^T(CPC^T + R)^{-1}CPA_{dp}^T$
Step 7: Output $\hat{x}(k|k)$
```

2.2. Linear Quadratic Gaussian Controller

The LQG controller is utilized with the goal of minimizing the LQG cost is assumed as:

$$J = min \lim_{\ell \to \infty} \frac{1}{\ell} \mathbb{E}[\sum_{i=0}^{\ell-1} (x(k))^T Q x(k) + (u(k))^T R u(k)] \tag{8}$$

where $Q^T = Q$ and $R^T = R$ denote the state and control variables, respectively. Consequently, the derived optimal feedback control law is:

$$u(k) = -\sum_{p=1}^{3} \hbar_p(\varrho(k)) F \hat{x}(k|k) \tag{9}$$

The goal function given the optimal estimator and LQG controller is recast as:

$$J^{\star} = tr(SW) + tr[\sum_{p=1}^{3} (A_{dp}{}^T S A_{dp} + Q - S)(P - LCP)] \tag{10}$$

Controller enables the system to achieve a stable operational state. Consequently, parameters including L, P, F, and J can be computed offline.

2.3. Replay Attack Model

Mo et al. [31] initially defined a model for a replay attack. This model assumes that an attacker can capture a substantial amount of sensor data. During the attack, the real sensor measurement $y(t_1 + i)$ is obstructed and replaced by $y(t_1 + i) = y(t_0 + i)$, which is forwarded to the estimator as demonstrated in Figure 4, with $d = t_1 - t_0$ identified as the replay delay. The model is expressed as:

$$y(t_1 + i) = y(t_0 + i), 0 \leq i \leq T. \tag{11}$$

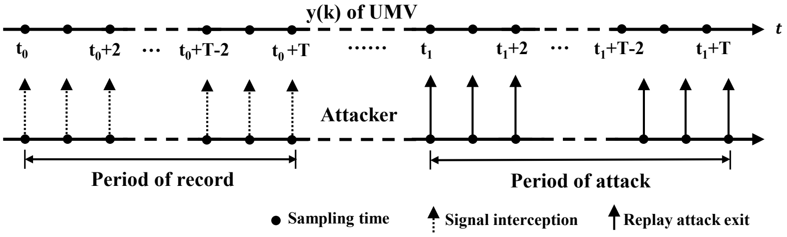

Figure 4. The timing diagram of the replay attacks.

Remark 1. *In UMV systems, two prevalent types of replay attacks are commonly discussed: Stuxnet-like Zhu et al. [24] and time-delay attacks, which often target network-transmitted signals such as $y(k)$ and $r(k)$. Stuxnet-like attacks, which involve manipulated control inputs, present significant detection challenges due to their reliance on historical feedback signals. Time-delay attacks disrupt system performance by modifying the poles of the closed-loop system. This study specifically focuses on replay attacks that affect sensor signals $y(k)$. Future research will investigate the simultaneous tampering of sensor data and control actions.*

Remark 2. *Replay attack threats to UMVs may come from state-sponsored groups, hacktivists, criminals, and insiders motivated by espionage, data theft, disruption, financial gain, and geopolitical leverage. The likelihood and success rate of replay attacks is contingent upon the security of the UMVs system's transmission network—specifically, whether it employs advanced encryption technologies to protect its communications and data transmissions and the complexity of the attacker's replay attack methods. When UMVs are fully deployed underwater, the inherent challenges of data transmission may reduce the success rate of such network attacks. Conversely, during docking operations or surface activities, UMV communications are more susceptible to the threat of replay attacks.*

Assumption 1. *In this paper's discussion of the replay attack scenario, the attack phase and the subsequent attack segment are temporally distinct with no overlap.*

3. A Periodic Watermark-Based Detection of Replay Attacks

3.1. Periodic Watermarking Detection Mechanism

As shown in Figure 5, the attacker intercepts the actual sensor measurement $y(k)$ and forwards $y^r(k)$ to the estimator or controller regularly. Furthermore, it has been shown in Yang et al. [23] that under certain conditions, these attacks can evade detection.

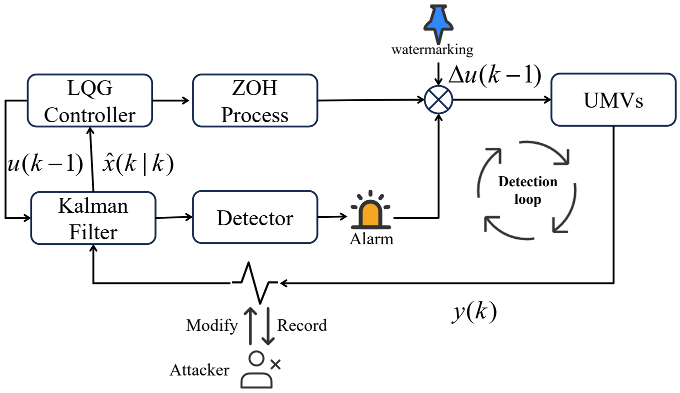

Figure 5. UMVs equipped with periodic watermarking algorithm.

Lemma 1. *In the scenario where a UMVs system faces a replay attack and equipping with a χ^2 detector, ensuring $\varepsilon = (A_{dp} + B_{dp}L)(I_n - LC)$ is Schur stable implies that the false alarm rate α and the positive detection rate β_k at time k will equalize over time, confirmed by $\lim_{k\to\infty}\beta_k = \alpha$.*

According to Lemma 1, the detector's alarm rate matches the false alarm rate, therefore complicating the confirmation of attacks upon alarms. To combat replay attacks, Mo et al. [31] introduced a technique of continuously adding Gaussian watermark signals $\Delta o(k)$ into the control inputs. It should be noted that, unless otherwise specified, the continuous watermark referenced hereafter is assumed to be a Gaussian signal. The actual control input can thus be described as:

$$u^r(k) = u(k) + \Delta o(k) \tag{12}$$

where $u^r(k)$ represents the actual control input of UMVs system added Gaussian watermark signals.

For UMV systems that are less responsive to watermarking, achieving an adequate detection rate might necessitate significantly increased watermark intensities Bian et al. [30]. To reduce the system control cost, a periodic watermarking mechanism is proposed in Figure 6.

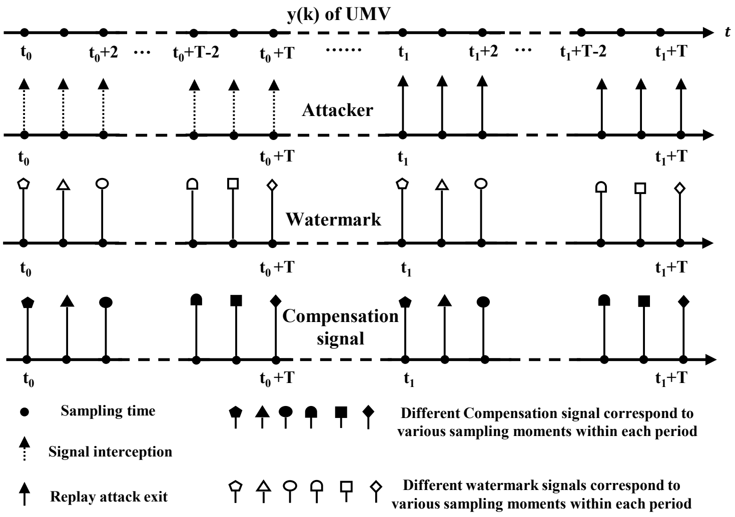
y(k) of UMV
t
t_0
t_0+2
···
t_0+T-2
t_0+T
······
t_1
t_1+2
···
t_1+T-2
t_1+T
Attacker
Watermark
Compensation signal
Sampling time
Signal interception
Replay attack exit
Different Compensation signal correspond to various sampling moments within each period
Different watermark signals correspond to various sampling moments within each period

Figure 6. A realization of a replay attack and periodic watermarking method.

Compared to the method in Bian et al. [30], the primary advantage of this periodic watermarking method lies in its reliance on comparing the relative magnitudes of the detection functions for compensation signals at various times within the same cycle rather than their absolute values. At time $k = t_0$, a periodic random watermark is introduced into the control variable, and a corresponding compensation signal is designed. As the system operates, this compensation signal develops periodicity that aligns with the watermark's cycle. Furthermore, during each operational cycle of the UMV, there is a direct correlation between the compensation signal and the watermark.

3.2. Construction of Periodic Watermark Signal

In contrast to the approach presented in Fang et al. [32], which injects watermarks of varying sizes at periodic intervals, this study introduces a method that continuously injects a periodic Gaussian random signal into the control signal. The watermark signal's period is denoted by T, where $T \in \mathbb{N}^r$. The series of periodic watermark signals is represented by $\Delta o_i = [\Delta o_0, \Delta o_1, \cdots, \Delta o_{T-1}]$, with each $\Delta o_i (i = 0, \ldots, T-1)$ being independently drawn from a Gaussian white noise signal characterized by a zero mean and a covariance matrix Σ_o. Assuming the system operates steadily at time t_0, a signal is continuously injected into the control variable. At any given time k, this injected watermark signal Δo is defined as follows:

$$\Delta o(k) = \Delta o(i), i = (k - t_0) \bmod T \tag{13}$$

where $(k - t_0) \mod T$ represents the remainder of k divided by T, and $0 \leq (k \mod T) \leq T - 1$. After injecting the periodic watermark signal, the UMVs state equation can be expressed as:

$$\begin{cases} x^r(k+1) = \sum_{p=1}^{3} \hbar_p(\varrho(k))(A_{dp}x^r(k) + B_{di}(u(k) + \Delta o(k)) + \theta_{di}(k)) \\ \quad y^r(k) = Dx^r(k) + v(k) \end{cases} \tag{14}$$

3.3. Construction of Periodic Compensation Signal

To construct the auxiliary system for obtaining the measurement compensation signal $y^f(k)$ after watermark insertion ($k > t_0$):

$$\begin{cases} x^f(k+1) = \sum\limits_{p=1}^{3} \hbar_p(\Delta o(k))\Big(A_{dp}x^f(k) + B_{di}\Delta o(k)\Big) \\ \quad y^f(k) = -Cx^f(k) \end{cases} \tag{15}$$

where $x^f(t_0) = 0$.

Theorem 1. *Assume A_{dp} is stable, there exits an $n_0 \in \mathbb{N}^r$ such that $A_{dp}^{(n_0-1)T} \approx 0$. By introducing watermark signals into the UMVs system over n_0 periods, the compensatory signal $y^f(k)$ achieves periodicity for $k \geq t_0 + n_0T$, mirroring the periodic nature of the watermark. A series of these compensatory signals $\mathbb{Y}^f = [\mathbb{Y}^f(0), \mathbb{Y}^f(1), \ldots, \mathbb{Y}^f(T-1)]$. Therefore, each compensatory signal within a period corresponds directly to its associated watermark signal, as demonstrated by the relationship $o(i-1) \leftrightarrow \mathbb{Y}^f(i)$, where $i = (k - t_0) \mod T$.*

Proof. To demonstrate that the compensation signal $y^f(k)$ becomes periodic for $k \geq t_0 + n_0T$, it suffices to prove that the compensation state $x^f(k)$ becomes periodic for $k \geq t_0 + n_0T$. When $k \in [t_0, t_0 + T - 1]$, the compensation signal state can be expressed as:

$$\begin{aligned} x^f(k) &= \textstyle\sum_{p=1}^{3} \hbar_p(\varrho(k))(A_{dp}x^f(k) + B_{dp}\Delta o(k)) \\ &= \textstyle\sum_{i=0}^{k-t_0} \sum_{p=1}^{3} \hbar_p(\varrho(k))\Big(A_{dp}x^f(k)^i B_{dp}\Delta o(k - t_0 - i)\Big) \end{aligned} \tag{16}$$

When $k = t_0 + nT + i$, with $i \in [0, T-1]$ and $n \in \mathbb{N}^r$, the compensation signal state can be represented as:

$$\begin{aligned} x^f(k) =& x^f(t_0 + nT + i) \\ =& \textstyle\sum_{p=1}^{3} \hbar_p\varrho(k)(A_{dp}x^f(t_0 + nT + i - 1))^i B_{dp}\Delta o(t_0 + nT + i - 1)) \\ =& \textstyle\sum_{p=1}^{3} \hbar_p\varrho(k)(({A_{dp}}^2(A_{dp}x^f(t_0 + nT + i - 3))^i B_{dp}\Delta o(i-3))) + A_{dp}B_{dp}\Delta o(i-2)) + B_{dp}\Delta o(i-1) \\ =& \textstyle\sum_{p=1}^{3} \hbar_p\varrho(k)({A_{dp}}^{(n-1)T} + x^f(t_0 + T) + x^f(t_0 + (l-1)T + i)) \end{aligned} \tag{17}$$

Given the stability of the transition matrix A_{dp}, an integer $n_0 \in \mathbb{N}^r$ is identified for which $A_{dp}^{(n_0-1)T} \approx 0$. This value of n_0 is determined solely by the matrix A_{dp}. Consequently, it is observed that for $n \geq n_0$, the state $x^f(t_0 + nT + i) = x^f(t_0 + (n-1)T + i)$. We can achieve

$$x^f(k) = x^f(k - T) \tag{18}$$

It is indicated in (18) that the state values of the compensatory signal are deduced to become periodic after n_0 operational periods, reflecting the periodicity of the watermark signal. With the independence of watermark signals within each period, the state of the compensatory signal is confirmed to be distinct within any given period. Given that the compensatory signal $y^f(k)$ is defined as $y^f(k) = -Cx^f(k)$ and output matrix C is constant, the compensatory signal is periodic, with each cycle exhibiting distinct values.

In accordance with (15), it is observed that for $k = t_0 + n_0T + i$, where i spans from 0 to $T-1$, a precise correlation is established between the value of the compensatory signal $\mathbb{Y}^f(i)$ and the watermark signal value $\Delta o(k-1) = \Delta o(i)$ present in the measurement data. It has been confirmed that for any distinct i and j, $\text{Prob}(\Delta o(i) = \Delta o(j)) = 0$. Thus, a one-to-one correspondence between compensatory signal $\mathbb{Y}^f(i)$ and the watermark signal sequence $\Delta o(i-1)$ is established. □

3.4. Detection of Replay Attacks

At time k, the measurement $y^r(k)$ is received and augmented by adding the compensatory signal $\mathbb{Y}^f(m)$, resulting in a compensated measurement represented as $y^e(k,m) = y^r(k) + \mathbb{Y}^f(m)$. This process results in the modified measurement being defined as $y^e(k,m) = y^r(k) + \mathbb{Y}^f(m)$. The prediction state estimate from the Kalman filter at this point is designated as $\hat{x}^e(k|k-1)$. Subsequently, the residual is computed as $r^e = y^e(k,m) - Cx^e(k|k-1)$. This residual is then employed to construct the chi-squared detection function $g(k,m) = (r^e(k,m))^T\Sigma_r^{-1}r^e(k,m)$, where Σ_r is defined as $CPC^T + V$. The effectiveness of the proposed detection method is confirmed in this paper, and it is demonstrated that the subsequent theorem is applicable.

Theorem 2. *At time k, where $k = t_0 + nT + i$, $i \in [0, T-1]$, and $n \geq n_0$, the control watermark signal is $\Delta o(i-1)$. Introducing different compensatory signals $\mathbb{Y}^f(m)$, $m \in [0, T-1]$, results in the chi-square detection function $g(k)$ satisfying $g(k,i) < g(k,j)$ for $j \in [0, T-1]$ and $j \neq i$. This indicates that assume compensatory signal $\mathbb{Y}^f(m)$ is injected with $m = i$, the detection function value $g(k,m)$ is minimized.*

Proof. According to Theorem 1, the compensation signal at time $k = t_0 + nT + i$ is denoted by $\mathbb{Y}^f(i)$. When $m = i$, $\mathbb{Y}^f(m) = \mathbb{Y}^f(i)$. Thus, the measurement value $y^e(k,m)$ after adding the compensation signal $\mathbb{Y}^f(m)$ can be derived as follows:

$$
\begin{aligned}
y^e(k,m) =& y^r(k) + \mathbb{Y}^f(m) \\
&= y^r(k) + \mathbb{Y}^f(t_0 + nT + i)C\sum_{p=1}^{3}\hbar_p(\varrho(k))(A_{dp}x^r(k-1) + B_{dp}(u(k-1) \\
&\quad + \Delta o(k-1)) + \theta_{di}(k-1) - A_{dp}x^f(k-1) - B_{dp}\Delta o(k-1) + v(k) \\
&= C\sum_{p=1}^{3}\hbar_p(\varrho(k))A_{dp}^k x(0) + \sum_{i=0}^{k-1}A_{dp}^i(B_{dp}u(k-1-i) + \theta_{di}(k-i-1)) + v(k) \\
&= Cx(k) + v(k) \\
&= y(k)
\end{aligned}
\tag{19}
$$

As established by (19), at time k, the measurement $y^e(k,m)$, which incorporates the compensatory signal $\mathbb{Y}^f(m)$, coincides with the system's original measurement $y(k)$. This concurrence indicates that the disturbances induced by the watermark signal in the measurement are effectively neutralized by the compensatory signal.

The prediction state estimate by the Kalman filter post-compensation is represented as $\hat{x}^e(k|k-1)$, and the original system's state prediction estimate is denoted as $\hat{x}(k|k-1)$. At the initial, it is affirmed that $\hat{x}^e(0|-1) = \hat{x}(0|-1)$. Following the integration with (19), it is derived that:

$$\begin{aligned}\hat{x}^e(k|k-1) &= \textstyle\sum_{n=1}^{3} \hbar_p(\varrho(k))\big(A_{dp}\hat{x}^e(k-1|k-1) + B_{dp}u(k-1)\big)\\ &= \textstyle\sum_{n=1}^{3} \hbar_p(\varrho(k))\big(A_{dp} + B_{dp}F\big)\hat{x}^e(k-1|k-1)\\ &= \textstyle\sum_{n=1}^{3} \hbar_p(\varrho(k))(\varepsilon^k \hat{x}(0|1) + \sum_{i=0}^{k-1} \varepsilon^i (A_{dp} + B_{dp}F)Ly(k-i-1))\\ &= \textstyle\sum_{n=1}^{3} \hbar_p(\varrho(k))\big(\hat{x}(k-1|k-2) + (A_{dp} + B_{dp}F)Ly(k-1)\big)\\ &=\hat{x}(k|k-1)\end{aligned} \tag{20}$$

Given the condition that $m = i$, it is found that the post-compensation state prediction estimate $\hat{x}^e(k|k-1)$ aligns with the state prediction estimate $\hat{x}(k|k-1)$ obtained when a control quantity watermark has not been injected. Hence, the residual $r^e(k,m)$ under the scheme presented in this paper is determined as:

$$\begin{aligned}r^e(k,m) &=y^e(k,m) - C\hat{x}^e(k|k-1)\\ &=y^r(k) + \mathbb{Y}^f(m) - C\hat{x}^e(k|k-1)\\ &=y(k) - C\hat{x}^e(k|k-1)\\ &=y(k) - C\hat{x}(k|k-1)\\ &=r(k)\end{aligned} \tag{21}$$

Consequently , the chi-squared detection function derived from the compensated residual is determined as follows:

$$\begin{aligned}g(k,m) &= (r^e(k,m))^T\Sigma_r^{-1}r^e(k,m)\\ &= (r(k))^T\Sigma_r^{-1}r(k)\\ &= g(k)\end{aligned} \tag{22}$$

Assume compensatory signal $\mathbb{Y}^f(m)$ is injected at time k. It has been established that the chi-squared detection function for compensated residual matches the $g(k)$ observed on the condition of no control quantity watermark is introduced. If at time $k-1$, a control quantity watermark signal $o(k-1) = o(i-1)$ was injected, and at time k, the compensatory signal $\mathbb{Y}^f(i)$ is introduced with $j \neq i$, then the measurement processed by the Kalman filter is as follows:

$$\begin{aligned}y^e(k,j) &= y^r(k) + \mathbb{Y}^f(j)\\ &= y^r(k) + \mathbb{Y}^f(k) + \mathbb{Y}^f(j) - \mathbb{Y}^f(k)\\ &= y^r(k) + \mathbb{Y}^f(i) + \mathbb{Y}^f(j) - \mathbb{Y}^f(i)\\ &= y(k) + \mathbb{Y}^f(j) - \mathbb{Y}^f(i)\end{aligned} \tag{23}$$

At this point, the residual $r^e(k,j)$ is as follows:

$$\begin{aligned}r^e(k,j) &= y^e(k,j) - C\hat{x}^e(k|k-1)\\ &= y(k) + \mathbb{Y}^s(j) - \mathbb{Y}^s(i) - C\hat{x}^e(k,j)\\ &= y(k) - C\hat{x}(k|k-1) + \mathbb{Y}^s(j) - \mathbb{Y}^s(i)\\ &= r(k) + \mathbb{Y}^s(j) - \mathbb{Y}^s(i)\end{aligned} \tag{24}$$

The chi-squared detection signal based on the residual is defined as:

$$\begin{aligned}g(k,j) &=(r^e(k,m))^T\Sigma_r^{-1}(r^e(k,j))\\ &=(r(k) + \mathbb{Y}^f(j) - \mathbb{Y}^f(i))^T\Sigma_r^{-1}(r(k) + \mathbb{Y}^f(j) - \mathbb{Y}^f(i))\\ &=r^T(k)\Sigma_r^{-1}r(k) + \Big(\mathbb{Y}^f(j) - \mathbb{Y}^f(i)\Big)^T\Sigma_r^{-1}(\mathbb{Y}^f(j) - \mathbb{Y}^f(i))\\ &=g(k) + \left|\Sigma_r^{-1}\right|\left\|\mathbb{Y}^f(j) - \mathbb{Y}^f(i)\right\|_2^2\end{aligned} \tag{25}$$

Given that $j \neq i$ and the compensatory signals within the same period are distinct, it invariably holds that $\left\|\mathbb{Y}^f(j) - \mathbb{Y}^f(i)\right\|_2^2 > 0$. Therefore, it is deduced that $g(k,j) > g(k) = g(k,i)$. □

Based on Theorem 2, the following replay attack detection method can be devised. Commencing from moment t_0, a periodic watermark signal is continuously injected into the control variable. At moment $k-1$, this watermark signal is given by $o(k-1) = o(i-1)$, where $i = (k - t_0) \mod T$. At moment k, where $k = t_0 + nT + i$, and $i \in [0, T-1], n \geq n_0$, the minimum value of the chi-squared detection function after incorporating different compensatory signals is denoted by:

$$a(k) = \min_m g(k,m), \quad m = 0, \ldots, T-1 \tag{26}$$

A decision function is defined as:

$$G(k) = a(k) - i \tag{27}$$

In accordance with (27), the following hypothesis test is proposed: If $G(k) = 0$, UMVs are considered to be operating normally. Otherwise, UMVs are presumed to be under a replay attack.

Theorem 3. *For a UMV system as depicted by (7), replay attacks can be detected to assume the delay of the replayed data is not an integer multiple of the watermark signal period, utilizing the approach described in this paper.*

Proof. It is established from Theorem 2 that during normal system operation at moment $k = t_0 + nT + i$, the compensatory signal $Y^f(m)$ that minimizes $g(k)$ satisfies $m = i$, with (26) yielding $a(k) = m$. Therefore, under the scheme proposed in this paper, the detection signal during normal operation is $G(k) = a(k) - i = m - i = 0$.

When a replay attack occurs at time $k = t_0 + nT + i$, where data from a normal state at time $k_1 = t_0 + n_1T + m_1$ is replayed, the corresponding control quantity watermark of the replayed measurement signal is $\Delta o(k_1 - 1) = \Delta o(m_1 - 1)$, with $m_1 = (k_1 - t_0) \mod T$. Given that the system operates normally at time $k-1$, it follows that $y^e(k-1) = y(k-1)$, and $\hat{x}^e(k|k-1) = \hat{x}(k|k-1)$. According to Theorem 2, for the detection function to achieve a minimal value, the injected compensatory signal should be $y^f(k_1) = Y^f(m_1)$, resulting in $a(k) = m_1$. If the delay of the replayed data $\Delta k = k - k_1 = (n - n_1)T + i - m_1$ is not an integer multiple of the watermark signal period, i.e., $i \neq m_1$, as indicated by (27), the detection signal $G(k) = \alpha(k) - i = m_1 - i \neq 0$ at this point, indicating the presence of a replay attack. □

Remark 3. *Employing the method proposed in this paper, a replay attack can be detected whenever the condition* $|\Sigma_r^{-1}|\left\|\mathbb{Y}^f(j) - \mathbb{Y}^f(i)\right\|_2^2$ *is satisfied in (25). Therefore, in an ideal scenario, the detection rate of replay attacks using this scheme reaches 1. However, due to the presence of random noise disturbances and computational errors, especially assuming* Σ_o *is small, there is a possibility of false detection occurring.*

Remark 4. *To compromise system stability, an attacker must record and store system signals for an extended period prior to launching a replay attack. Conversely, control watermarks can be integrated into the control variables at the onset of system operations. It is established that* n_0 *is sufficient to ensure that* $A_{di}^{(n_0-1)T} \approx 0$. *Consequently, this paper assumes that no replay attack occurs within the initial* n_0 *periods.*

Remark 5. *If the delay of the replayed data* $\Delta k = k - k_1 = (n - n_1)T + i - m_1$ *is an integer multiple of the watermark signal period, i.e.,* $i = m_1$*, as indicated by (27), the detection signal*

$G(k) = \alpha(k) - i = m_1 - i = 0$ at this point, which can result in a false positive. However, this method has certain limitations on the condition that the length of the chosen watermark period is sufficiently large that the occurrence of such cases can be minimized. Future work will focus on addressing this issue in more depth.

3.5. Analysis of System Performance Degradation

This paper focuses on analyzing the impact of this method on system control. Assuming no additional signals are present, the true state variables of the UMVs system are expressed by $x(k)$. The state error is then described as $\epsilon(k) = x(k) - \hat{x}(k)$. Employing the methodology presented herein, the state variables for the system during normal operation are given by $\hat{x}^r(k)$, and the undisturbed steady-state error as defined within this section is $\hat{e}(k)$, thus yielding $\delta(k) = \hat{x}^r(k) - \hat{x}(k)$. Upon constructing the matrix $\tilde{x}(k) = [x(k) \;\; \epsilon(k) \;\; \delta(k)]^T$, the following relation is established:

$$\tilde{x}(k+1) = \sum_{p=1}^{3} \begin{bmatrix} A_{dp} + B_{dp}F & -B_{dp}F & 0 \\ 0 & A_{dp} - LCA_{dp} & 0 \\ 0 & 0 & A_{dp} \end{bmatrix} \tilde{x}(k) + \begin{bmatrix} w(k) \\ Hw(k) - Lv(k+1) \\ \sum_{p=1}^{3} B_{dp}u(k) \end{bmatrix} \tag{28}$$

where $H = I_n - LC$, Θ and $\gamma(k)$ can be achieved as:

$$\Theta = \sum_{p=1}^{3} \begin{bmatrix} A_{dp} + B_{dp}F & -B_{dp}F & 0 \\ 0 & A_{dp} - LCA_{dp} & 0 \\ 0 & 0 & A_{dp} \end{bmatrix}, \gamma(k) \quad = \begin{bmatrix} w(k) \\ Hw(k) - Lv(k+1) \\ \sum_{p=1}^{3} B_{dp}u(k) \end{bmatrix}$$

Consequently , the covariance of the system is expressed as:

$$cov(\tilde{x}) = \Theta cov(\tilde{x})\Theta^T + \iota \tag{29}$$

Wherein , ι is characterized by:

$$\iota = \begin{bmatrix} W & WH^T & 0 \\ HW & HW(I_n - LC)^T & 0 \\ LCW & 0 & \sum_{p=1}^{3}(B_{dp}\Sigma_o B_{dp}^T) \end{bmatrix}$$

In the scheme proposed in this paper, the system control performance index can be represented as:

$$\begin{aligned} J = &\lim_{\ell\to\infty} \frac{1}{\ell+1} \sum_{k=1}^{\ell} \mathbb{E}\Big\{x^{*T}(k)Qx^*(k) + (u(k) + \Delta o(k))^T \\ &\times R(u(k) + \Delta o(k))\} \\ &= \lim_{\ell\to\infty} \frac{1}{\ell+1} \sum_{k=1}^{\ell} \mathbb{E}\Big\{\tilde{x}(k)\Omega\tilde{x}^T(k) + \Delta o^T(k)R\Delta o(k)\Big\} \\ &= tr(\mathrm{cov}(\tilde{x}(k))\Omega + \Sigma_o R) \end{aligned} \tag{30}$$

where $\Omega = \begin{bmatrix} Q + F^TRF & Q & -F^TRF \\ -F^TRF & 0 & F^TRF \\ Q & Q & 0 \end{bmatrix}$.

In the scheme proposed in this paper, the system performance degradation is expressed as:

$$\Delta J = J - J^\star \tag{31}$$

wherein, the optimal system control performance index $J^\star$ is derived using (10).

Remark 6. *According to (31), it can be observed that an increase in the covariance of the watermark signal Σ_o results in a corresponding rise in the control performance function J. The methodology* *adopted in this paper, which utilizes the relative positions of watermark signals for attack detection, enables the application of smaller Σ_o values, effectively minimizing control costs in UMV systems.*

4. Numerical Simulation

This section validates the efficacy of replay attack detection. The UMVs model selected for this study, borrowed from Bian et al. [30] and Wang et al. [33], is presented below:

$$\mathbb{M} = \begin{bmatrix} 1.0852 & 0 & 0 \\ 0 & 2.0575 & -0.4087 \\ 0 & -0.4087 & 0.2153 \end{bmatrix}, \mathbb{N} = \begin{bmatrix} 0.0865 & 0 & 0 \\ 0 & 0.0762 & 0.1510 \\ 0 & 0.1510 & 0.0031 \end{bmatrix}, \mathbb{O} = \begin{bmatrix} 1.0389 & 0 & 0 \\ 0 & 0.0266 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

For the UMVs to effectively track the reference system's output through the yaw angle, consider the following matrix $C = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$. The disturbance $\theta(k)$ is presented as:

$$\begin{cases} \theta_1(k) = 0.27F_1(s)\mathbb{N}_1(k)\mathbb{N}_2(k) \\ \theta_2(k) = -0.6\cos(1.6k)e^{-0.12k} \\ \theta_3(k) = 0.58F_2(s)\mathbb{N}_3(k)\mathbb{N}_4(s) \end{cases}$$

where $F_1(s) \triangleq [K_{\theta 1}s/(s^2 + 2h_1\varsigma_1 s + \varsigma_1^2)]$ and $F_2(s) \triangleq [K_{\theta 2}s/(s^2 + 2h_2\varsigma_2 s + \varsigma_2^2)]$. Designated wave strength coefficients are $K_{\theta 1} = 0.26$ and $K_{\theta 2} = 0.8$. The damping factors, $h_1 = 0.2$ and $h_2 = 1.7$. The frequencies of the waves are $\varsigma_1 = 1.3$ and $\varsigma_2 = 0.9$. Levels of white noise power are $\mathbb{N}_1(k) = 2.69$ and $\mathbb{N}_3(k) = 1.56$. Where

$$\mathbb{N}_2(\mathrm{t}) = \begin{cases} 1, t \in [0s, 20s] \\ 0, t > 20s \end{cases}, \quad \mathbb{N}_4(\mathrm{t}) = \begin{cases} 1, t \in [0s, 15s] \\ 0, t > 15s \end{cases}$$

At $t_0 = 50$ s, assuming the injected watermark signal has a periodicity of $T = 10$ s with $\Sigma_o = 0.5763$, there exist $A_{dp}^{10T} \approx 0$. After 10 periods of watermark signal injection, the injected watermark and compensation signals are illustrated in Figure 7. As depicted in the following figures, both the compensation and watermark signals exhibit periodic behavior with distinct values within each period.

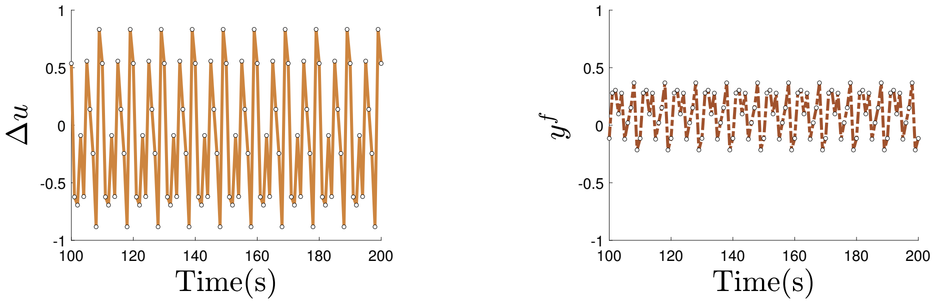

Figure 7. The periodic watermarking signal and the periodic compensation signal with a period of $T = 10$.

In the following, an evaluation of watermark signal periods is conducted, detailing detection rates and false alarm rates under two types of replay attacks at various watermark intervals, as shown in Table 2. For the same watermark signal value, $\Sigma_o = 0.5763$, β values for both continuous and discontinuous replay attacks exhibit an increase for $T < 10$. Conversely, β values decrease as the watermark period increases beyond $T > 10$. Based on these observations, a period of $T = 10$ has been selected.

In scenarios devoid of process and measurement noise interference, with $\Sigma_o = 0.015$ and $\Sigma_o = 0.4353$, the detection results under normal system operation are as illustrated in Figure 8. The values of the discriminant function are consistently zero, indicating the absence of false positives. With noise interference present in the system, the watermark signal variances are set to $\Sigma_o = 0.015$ and $\Sigma_o = 0.4353$, with the corresponding normal operational detection results displayed in Figure 9, respectively. The state noise and measurement noise in the system can lead to false negatives and false positives. As the variance Σ_o of the watermark signal increases, the incidence of false positives decreases.

Table 2. Detection Rates of Replay Attacks under Different Watermark Periods

T	Σ_o	β of Continuous Replay Attacks	β of Discontinuous Replay Attacks
6	0.5763	92.5%	87.3%
8	0.5763	92.7%	88.4%
10	0.5763	93.8%	90.1%
12	0.5763	93.3%	89.2%
14	0.5763	89.6%	87.5%

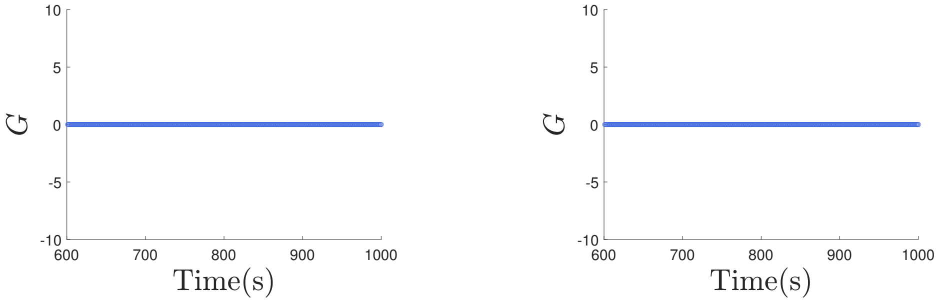

Figure 8. $G(k)$ for UMVs with different watermark signal variances in the absence of noise.

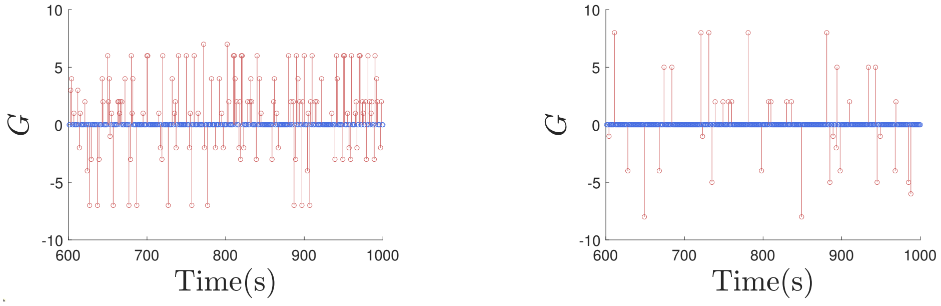

Figure 9. $G(k)$ for UMVs with different watermark signal variances in the presence of noise.

Case 1: The UMVs system is subjected to continuous replay attacks. Setting the watermark signal at $\Sigma_o = 0.5763$, data are recorded by the attacker from 30 to 40 s. As illustrated in Figure 10 (left figure), the UMVs undergo replay attacks from 650 to 950 s. Similarly, as shown in Figure 10 (right figure), replay attacks occur from 750 to 1000 s. This case study aims to test the strategy's effectiveness in detecting prolonged continuous replay attacks.

Case 2: The UMVs system is subjected to discontinuous replay attacks. Setting the watermark signal at $\Sigma_o = 0.5763$, data are recorded by the attacker from 30 to 40 s. As illustrated in Figure 11 (left figure), the UMVs undergo replay attacks from 620 to 660 s, 700 to 740 s, 780 to 820 s, 860 to 900 s, and 940 to 980 s. Similarly, as shown in Figure 11 (right figure), replay attacks occur from 620 to 710 s, 750 to 840 s, and 880 to 970 s. This case study aims to test the strategy's effectiveness in detecting prolonged continuous replay attacks.

Detection effectiveness of the UMVs system under replay attacks, as demonstrated for both continuous and discontinuous types of replay attacks using the method proposed, is showcased in **Case 1** and **Case 2**. In **Case 1**, the UMVs underwent continuous replay attacks of two different durations during regular operations. As Figure 10 indicates, during normal system operation, $G(k) = 0$. However, when the UMVs are undergoing a replay attack, significant changes occur in the values of the detection function. Between 650 and 750 s, $G(k) \neq 0$, indicating that the UMVs are experiencing a replay attack. In **Case 2**, the UMVs underwent several brief intermittent replay attacks. As displayed in Figure 11, the system faced five short, discontinuous attacks and two additional quick, intermittent ones. Similar to **Case 1**, with the application of the proposed strategy, the detection function maintains $G(k) = 0$ during normal operations. Notable shifts in detection function values when under attack suggest a replay attack is occurring. After repeating experiments 1000 times, the detection rates for continuous and discontinuous replay attacks are $\beta = 93.8\%$ and $\beta = 90.1\%$, respectively, with false positive rates of $\alpha = 4.5\%$ and $\alpha = 6.7\%$. Thus, the method proposed in this article effectively detects the occurrence of continuous with discontinuous replay attacks of UMV systems.

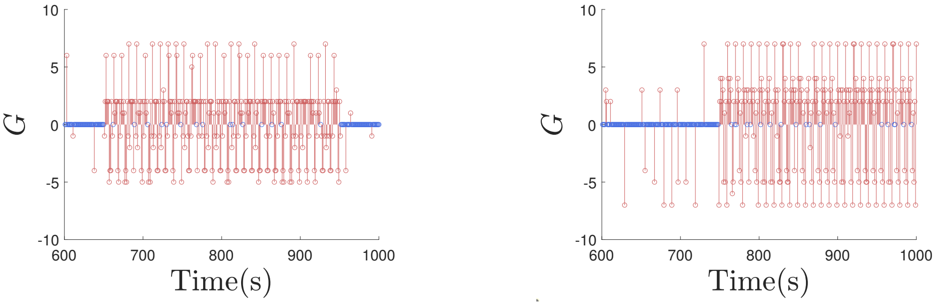

Figure 10. The $G(k)$ of the UMVs during continuous replay attacks in **Case 1**.

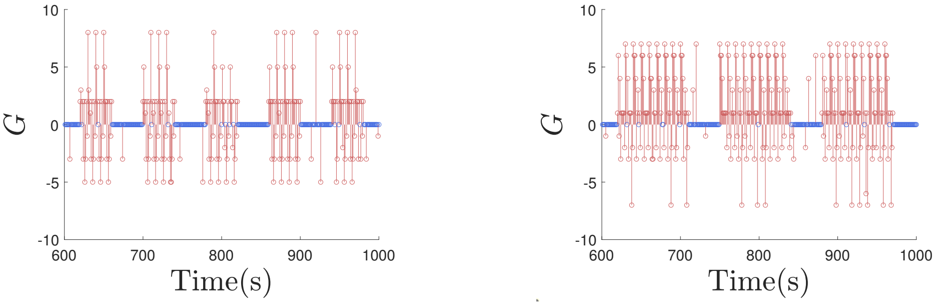

Figure 11. The $G(k)$ of the UMVs during discontinuous replay attacks in **Case 2**.

As shown in Figure 12, in order to further validate the effectiveness of the proposed method, a comparative analysis is performed to compare the control performance of the proposed method with that of the method outlined in [30,36,37]. Simulations were conducted for 200 iterations on four approaches to obtain the average performance degradation and detection rates under replay attacks. It can be observed from Figure 12 that as the detection rate increases, the performance degradation of different methods gradually increases.

This analysis underscores the superiority of the proposed method by comparing various metrics, including the continuous replay attack detection rate β_1, discontinuous replay attack detection rate β_2, continuous replay attack false alarm rate α_1, discontinuous replay attack false alarm rate α_2, and the average control cost under multiple attacks J_{ave}, with those reported in previous studies [30–32,36,37]. The results of this comparative analysis are presented in Tables 3 and 4.

It is revealed in Table 3 that when the detection requirements are specified as follows: continuous replay attack detection rate β_1 between 88% and 92%, discontinuous replay attack detection rate β_2 between 87% and 91%, continuous replay attack false alarm rate α_1 between 4% and 6%, and discontinuous replay attack false alarm rate α_2 between 6% and 10%, the lowest control cost of 24 is achieved by the proposed method.

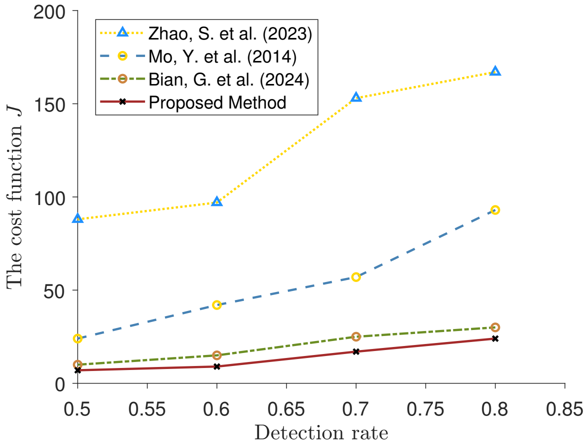

Figure 12. Comparison between the performance index of the method proposed in this paper and others [30,36,37].

Table 3. Comparison of Detection Efficiency and Control Cost Between Methods in This Paper and Other Papers under the Same Detection Efficiency

	J_{ave}	β_1	α_1	β_2	α_2
Method in This Paper	24	88–92%	4–6%	87–91%	6–10%
Method in Bian et al. [30]	30	88–92%	4–6%	87–91%	6–10%
Method in Zhao et al. [37]	55	88–92%	4–6%	87–91%	6–10%
Method in Fang et al. [32]	93	88–92%	4–6%	87–91%	6–10%
Method in Mo et al. [31]	167	88–92%	4–6%	87–91%	6–10%
Method in Mo et al. [36]	175	88–92%	4–6%	87–91%	6–10%

It is shown in Table 4 that when the control cost J_{ave} for various methods is fixed at 80, other methods are surpassed by the proposed method in both continuous replay attack detection rate β_1 and discontinuous replay attack detection rate β_2. Additionally, the continuous replay attack false alarm rate α_1 and discontinuous replay attack false alarm rate α_2 are both lower than those reported in previously published methods.

Table 4. Comparison of Detection Efficiency and Control Cost Between Methods in This Paper and Other Papers under the Same Control Performance Loss

	J_{ave}	β_1	α_1	β_2	α_2
Method in This Paper	80	95–96%	1–2%	95–96%	1–2%
Method in Bian et al. [30]	80	92–94%	1–2%	92–94%	1–3%
Method in Zhao et al. [37]	80	90–92%	3–5%	89–92%	3–4%
Method in Fang et al. [32]	80	82–84%	9–12%	79–80%	11–13%
Method in Mo et al. [31]	80	67–70%	16–19%	57–66%	22–27%
Method in Mo et al. [36]	80	65–70%	17–23%	53–62%	28–35%

The superiority of the proposed method is evident from the data and analysis presented. When the detection accuracy requirement is held constant, the proposed method results in reduced control costs. Conversely, when the control cost requirement is fixed, the proposed method achieves higher detection accuracy and a lower false alarm rate.

5. Conclusions

This research has presented a method for detecting replay attacks on UMVs by employing periodic watermark and compensatory signals that adhere to a Gaussian signal. Unlike conventional approaches that use fixed thresholds to compare residual detection function magnitudes, this method has improved detection efficiency by evaluating the relative position across compensatory signals. The technique has not only achieved high detection rates but also reduced the variance of watermark signals, thus preserving system performance. However, when the attack delay aligns with an integer multiple of the watermark signal period, the method may produce some false positives. In the method used by Bian et al. [30], although the method incurs a greater loss in control system performance compared to the paper, it does not produce false positives under these circumstances. Addressing this issue is planned for future research. Future research will also focus on designing new watermark signals to minimize their effect on system control performance further.

Author Contributions: Conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, supervision, project administration, funding acquisition, G.B. and X.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported in part by the National Natural Science Foundation of China under Grants 52301418; Sichuan Science and Technology Program under Grants 2024NSFSC0878; Joint Fund of Science and Technology Department of Liaoning Province and State Key Laboratory of Robotics under Grant 2022-KF-22-08.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Wang, Y.; Yang, X.; Hao, L.; Li, T.; Chen, C.L.P. Integral Sliding Mode Output Feedback Control for Unmanned Marine Vehicles Using T–S Fuzzy Model with Unknown Premise Variables and Actuator Faults. *Appl. Sci.* **2024**, *12*, 920. [CrossRef]
2. Li, W.; Zhou, H.; Zhang, J. Quasi-Infinite Horizon Model Predictive Control with Fixed-Time Disturbance Observer for Underactuated Surface Vessel Path Following. *Appl. Sci.* **2024**, *12*, 967. [CrossRef]
3. Wu, Y.; Wang, T.; Liu, S. A Review of Path Planning Methods for Marine Autonomous Surface Vehicles. *Appl. Sci.* **2024**, *12*, 833. [CrossRef]
4. Li, Z.; Lei, K. Robust Fixed-Time Fault-Tolerant Control for USV with Prescribed Tracking Performance. *Appl. Sci.* **2024**, *12*, 799. [CrossRef]
5. Pandey, P.K.; Kansal, V.; Swaroop, A. PKI-SMR: PKI based secure multipath routing for unmanned military vehicles (UMV) in VANETs. *Wirel. Netw.* **2024**, *30*, 595–615. [CrossRef]
6. Zhang, P.; Li, H.; Wang, Y.; Zhao, X. Optimal search path planning of UUV in battlefield ambush scene. *Def. Technol.* **2024**, *19*, 101–111.
7. Wang, H.; Wang, T.; Lv, H.; Liu, S. An adjustable pendulum mechanism for in-situ wave energy harvesting in an unmanned marine vehicle. *Ocean Eng.* **2024**, *297*, 117116. [CrossRef]
8. Song, W.; Tong, S. Event-triggered fuzzy finite-time reliable control for dynamic positioning of nonlinear unmanned marine vehicles. *Ocean Eng.* **2022**, 266, 113139. [CrossRef]
9. Wang, W.; Liang, H.; Pan, Y.; Li, T. Prescribed Performance Adaptive Fuzzy Containment Control for Nonlinear Multiagent Systems Using Disturbance Observer. *IEEE Trans. Cybern.* **2020**, *50*, 3879–3891. [CrossRef]
10. Hao, L.; Zhang, H.; Li, H.; Li, T. Sliding mode fault-tolerant control for unmanned marine vehicles with signal quantization and time-delay. *Ocean Eng.* **2020**, *215*, 107882. [CrossRef]
11. Chwa, D. Global Tracking Control of Underactuated Ships With Input and Velocity Constraints Using Dynamic Surface Control Method. *IEEE Trans. Control Syst. Technol.* **2011**, *19*, 1357–1370. [CrossRef]
12. Ge, S.S.; Zhang, J. Neural-network control of nonaffine nonlinear system with zero dynamics by state and output feedback. *IEEE Trans. Neural Netw.* **2003**, *14*, 900–918. [PubMed]
13. Liu, M.; Zhao, C.; Xia, J.; Deng, R.; Cheng, P.; Chen, J. PDDL: Proactive Distributed Detection and Localization against Stealthy Deception Attacks in DC Microgrids. *IEEE Trans. Smart Grid* **2023**, *14*, 714–731. [CrossRef]
14. Alfaro-Cid, E.; McGookin, E.W.; Murray-Smith, D.J.; Fossen, T.I. Genetic Programming for the Automatic Design of Controllers for a Surface Ship. *IEEE Trans. Intell. Transp. Syst.* **2008**, *9*, 311–321. [CrossRef]

15. Liu, Q.; Long, Y.; Li, T.; Chen, C.L.P. Attack-resilient fault detection for interconnected systems under DoS attack. *ISA Trans.* **2024**, *148*, 201–211. [CrossRef] [PubMed]
16. Zhang, D.; Ye, Z.; Feng, G.; Li, H. Intelligent Event-Based Fuzzy Dynamic Positioning Control of Nonlinear Unmanned Marine Vehicles Under DoS Attack. *IEEE Trans. Cybern.* **2022**, *52*, 13486–13499. [CrossRef]
17. Fei, Z.; Wang, X.; Wang, Z. Event-Based Fault Detection for Unmanned Surface Vehicles Subject to Denial-of-Service Attacks. *IEEE Trans. Syst. Man Cybern. Syst.* **2022**, *52*, 3326–3336. [CrossRef]
18. Zhang, Z.; Cheng, P.; Wu, J.; Chen, J. Secure State Estimation Using Hybrid Homomorphic Encryption Scheme. *IEEE Trans. Control Syst. Technol.* **2021**, *29*, 1704–1720. [CrossRef]
19. Steiner, J.; Neuman, C.; Schiller, J. Kerberos: An authentication service for open networks. In Proceedings of the USENIX Annual Tech Conference, Dallas, TX, USA, 9–12 February 1988.
20. Greene, K.; Rodgers, D.; Dykhuizen, H.; Niyaz, Q.; Shamaileh, K.A.; Devabhaktuni, V. A Defense Mechanism Against Replay Attack in Remote Keyless Entry Systems Using Timestamping and XOR Logic. *IEEE Consum. Electron. Mag.* **2021**, *10*, 101–108. [CrossRef]
21. Farha, F.; Ning, H.; Yang, S.; Xu, J.; Zhang, W.; Choo, K.-K.R. Timestamp Scheme to Mitigate Replay Attacks in Secure ZigBee Networks. *IEEE Trans. Mob. Comput.* **2022**, *21*, 342–351. [CrossRef]
22. Xiao, M.; Song, W.; Yang, K.; OuYang, R.; Zhao, H. Formal Analysis of the Security Protocol with Timestamp Using SPIN. *Comput. Intell. Neurosci.* **2022**, *2022*, 2420590. [CrossRef] [PubMed]
23. Yang, C.; Chu, Z.; Ma, L.; Wang, G.; Dai, W. Joint Watermarking-Based Replay Attack Detection for Industrial Process Operation Optimization Cyber-Physical Systems. *IEEE Trans. Ind. Inform.* **2023**, *19*, 8910–8922. [CrossRef]
24. Zhu, H.; Liu, M.; Fang, C.; Deng, R.; Cheng, P. Detection-Performance Tradeoff for Watermarking in Industrial Control Systems. *IEEE Trans. Inf. Forensics Secur.* **2023**, *18*, 2780–2793. [CrossRef]
25. Du, D.; Zhang, C.; Li, X.; Fei, M.; Zhou, H. Attack Detection for Networked Control Systems Using Event-Triggered Dynamic Watermarking. *IEEE Trans. Ind. Inform.***2023**, *19*, 351–361. [CrossRef]
26. Liu, H.; Mo, Y.; Yan, J.; Xie, L.; Johansson, K.H. An Online Approach to Physical Watermark Design. *IEEE Trans. Autom. Control* **2020**, *65*, 3895–3902. [CrossRef]
27. Ma, L.; Chu, Z.; Yang, C.; Wang, G.; Dai, W. Recursive Watermarking-Based Transient Covert Attack Detection for the Industrial CPS. *IEEE Trans. Inf. Forensics Secur.* **2023**, *18*, 1709–1719. [CrossRef]
28. Naha, A.; Teixeira, A.M.H.; Ahlén, A.; Dey, S. Quickest detection of deception attacks on cyber–physical systems with a parsimonious watermarking policy. *Automatica* **2023**, *155*, 111147. [CrossRef]
29. Song, Z.; Skuric, A.; Ji, K. A Recursive Watermark Method for Hard Real-Time Industrial Control System Cyber-Resilience Enhancement. *IEEE Trans. Autom. Sci. Eng.* **2020**, *17*, 1030–1043. [CrossRef]
30. Bian, G.; Long, Y.; Li, T.; Park, J.H.; Chen, C.L.P. Watermark-Based Replay Attack Detection for Unmanned Marine Vehicles. *IEEE Trans. Intell. Veh.* 2024, *early access*. [CrossRef]
31. Mo, Y.; Chabukswar, R.; Sinopoli, B. Detecting Integrity Attacks on SCADA Systems. *IEEE Trans. Control Syst. Technol.* **2014**, *22*, 1396–1407. [CrossRef]
32. Fang, C.; Qi, Y.; Cheng, P.; Zheng, W.X. Optimal periodic watermarking schedule for replay attack detection in cyber–physical systems. *Automatica* **2020**, *112*, 108698. [CrossRef]
33. Wang, Y.-L.; Han, Q.-L.; Fei, M.-R.; Peng, C. Network-Based T–S Fuzzy Dynamic Positioning Controller Design for Unmanned Marine Vehicles. *IEEE Trans. Cybern.* **2018**, *48*, 2750–2763. [CrossRef] [PubMed]
34. Peng, Z.; Wang, J.; Wang, D.; Han, Q.-L. An Overview of Recent Advances in Coordinated Control of Multiple Autonomous Surface Vehicles. *IEEE Trans. Ind. Inform.***2021**, *17*, 732–745. [CrossRef]
35. Hao, L.-Y.; Zhang, H.; Li, T.S.; Lin, B.; Chen, C.L.P. Fault Tolerant Control for Dynamic Positioning of Unmanned Marine Vehicles Based on T-S Fuzzy Model with Unknown Membership Functions. *IEEE Trans. Veh. Technol.* **2021**, 70, 146–157. [CrossRef]
36. Mo Y.; Weerakkody, S.; Sinopoli, B. Physical Authentication of Control Systems: Designing Watermarked Control Inputs to Detect Counterfeit Sensor Outputs. *IEEE Control Syst. Mag.* **2015**, *35*, 93–109.
37. Zhao, S.; Li, Q.; Cao, H. Improved Smooth Watermarking Methods for Detecting Replay Attacks in Process Control Systems. *Electronics* **2023**, *12*, 3812. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.