Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 95 (2016) 193 – 200

Procedia Computer Science

Complex Adaptive Systems, Publication 6
Cihan H. Dagli, Editor in Chief
Conference Organized by Missouri University of Science and Technology
2016 - Los Angeles, CA

Detection of Fault Data Injection Attack on UAV Using Adaptive Neural Network

Alireza Abbaspour[a]*, Kang K. Yen[a], Shirin Noei[b], Arman Sargolzaei[c]

[a]*Department of Electrical and Computer Engineering at Florida International University, Miami, FL, USA*
[b]*Department of Civil and coastal Engineering at University of Florida, Gainesville,FL,USA*
[c] *Department of Electrical Engineering at Florida Polytechnique University, Lakeland,FL,USA*

Abstract

A resilient and secure control system should be designed to be as safe and robust as possible in face of different types of attacks such as fault data injection (FDI) attacks; thus, nowadays, the control designers should also consider the probable attacks in their control design from the beginning. For this reason, detection of intentional faults and cyber-attacks attracts a great concern among researchers. This issue plays a great role in the safety of unmanned aerial vehicles (UAVs) due to the need of continuous supervision and control of these systems. In order to have a cyber-attack tolerant (CAT) controller, the attack and the type of attack should be detected in the first step. This paper introduces a new algorithm to detect fault data injection attack in UAV. An adaptive neural network is used to detect the injected faults in sensors of an UAV. An embedded Kalman filter (EKF) is used for online tuning of neural networks weights; these online tuning makes the attack detection faster and more accurate. The simulation results show that the proposed method can successfully detect FDI attacks applied to an UAV.

Peer-review under responsibility of scientific committee of Missouri University of Science and Technology

Keywords: Cyber-attack; UAV; Sensors; Attacks and Faults Detection; Fault Data Injection; Adaptive Neural Network

1. Introduction

In recent years, application of unmanned aerial vehicles (UAVs) increased significantly [1]. This growth will be

* Corresponding author. Tel.: +1-316-516-6525
E-mail address: aabba014@fiu.edu

Peer-review under responsibility of scientific committee of Missouri University of Science and Technology
doi:10.1016/j.procs.2016.09.312

continued by technology improvement which will lead to cheaper and more intelligent UAV's systems. This wide spread application of UAV increases the safety and security concerns. Conventional autopilot designs are based on the human supervision and do not consider the cyber security threats, which makes them vulnerable to cyber-attacks. In 2009, a predator UAV stream hijacked which increased the attention to the UAV cyber-security considerably [1].

Generally, cyber-attacks can be divided in to three main categories: hardware attacks, wireless attacks, and sensor spoofing [1]. In hardware attacks the attacker has access to the autopilot hardware components; while, in wireless attacks, the attackers use wireless communication channels to penetrate the system. Several researches investigated strategies to prevent wireless communication attacks [2-4]. In these works, they introduce algorithms to detect attacks that want to penetrate and extract the cypher keys. By extracting the cypher keys, attackers can enter their false data within the same structure of correct data. In these kinds of attacks, attackers introduce errors in cryptographic algorithm in order to obtain the faulty cipher text so that they can decode the cipher key.

Sensor spoofing which is the main subject of this paper can be defined as passing false sensor information to the UAV's on-board autopilot. In these kinds of attacks, it is considered that the attacker has already entered in the system, so in order to have a secure system, it should be tried to detect and isolate the injected faults from the real data. Attackers can perform sensor spoofing with various kinds of procedures such as injecting faults to the sensor feedback link, malfunctioning sensors, or overloading the on-board processor which can be led to denial of service between the sensor and the controller. These kinds of attacks can be detected by the fault detection algorithms.

Various researches have been done to obtain robust controllers that can tolerate small fault and uncertainties [5-8]; however, in order to have a robust controller their designs had to be very conservative that limit the maneuverability of the control system. Therefore, the detection algorithms attract researcher's attention [9-16]. Several algorithms have been proposed for FDI attacks but only a few of them focused on detection of the FDI attacks on UAV [9-11]. Among these methods, neural network (NN) received most of attention due to its ability in online learning and estimation of the nonlinear systems [12-16]. A fault detection design which is based on a bank of observer with learning ability using diagnostic residuals for discrete-time nonlinear systems is introduced by Wu et al. [12]. In their design in order to update the observer parameters, for each sample a type of Proportional-Derivative learning algorithm was used.

An NN based system modeling for fault detection process is introduced by Samy et al. [13]. In this design, the system identification was obtained through off-line learning process, so, the fault detection process does not need the model data during the operation. However, due to the fact that the model identification was performed off-line, the model cannot provide needed information for highly nonlinear system with uncertainties. A fault detection design based on NN is introduced for actuator faults in satellite by Talebi et al. [14]. However, this design was based on the assumption that all the detailed behavior of the system is available.

Chen et al. introduced a neural fault tolerant control design which was designed for faults in the actuators of a three degree of freedom helicopter [15]. In their design neural observer was designed based on radial basis function. However, this design cannot detect the faults in the helicopter sensors. Shen et al. presented a neural network fault detection technique that it considers the time delay between the fault detection and fault accommodation [16]. However, this technique assumes that fault occurrence time is larger than the system stabilization time ($t_s<t_f$) which make this design vulnerable to sudden faults.

This paper introduces a novel detection system for FDI in sensors of the UAV. This design is based on a three layer adaptive NN in which it's learning coefficients are updated by EKF that can be applied to the UAV systems to detect and isolate the cyber-attacks in sensors. Conventional NN detection techniques suffer from the slow learning rate that makes the system vulnerable to abrupt faults and attacks. Using EKF for updating the learning coefficients helps to improve the detection ability of neural network, increase the learning rate, and gives the ability to detect sudden FDI. The proposed algorithm has been implemented and evaluated on a six-degree-of-freedom (DoF) aircraft model, i.e., a nonlinear model of WVU YF-22 unmanned aircraft. The simulation results via MATLAB SIMULINK software show that the proposed method can detect FDI attack and the sensor faults successfully.

The paper is organized as follows: Section 2 provides the problem definition and explanation, while the proposed detection strategy is illustrated in Section 3. Then in Section 4, simulation results are presented to demonstrate the effectiveness of the introduced method. Finally, the conclusions and future work are provided in Section 5.

2. Problem Statement

The significant growth in wireless communication system brought convenience in aircraft design and management, but at the same time brought the security concerns for designers [17-22]. These wireless communications can be used as a potential gate for cyber attackers to enter and take the control of the system or interrupt it. Due to the nature of UAV's systems, they use more wireless communication system which makes them better target for attackers. Moreover, lack of human supervision makes them more vulnerable to the cyber-attacks. In recent years, the security and stability of UAVs have been at the centre stage for researchers, engineers, and governmental entities, since exploited security risks could have potential catastrophic consequences [1]. Although the security and stability schemes for UAVs have advanced in the past several years, there have been several acknowledged cyber-attacks and confirmations that these systems are not very safe and stable in the face of attacks and natural disturbances.

This paper deals with FDI attacks in the aircraft sensor systems. This class of attack injects faults into a device performing some type of computation. These faults can be anything from unusual environmental conditions (increased heat, for example), the injection of a laser beam at the appropriate frequency [23], or the injection of data packets that collide with legitimate packets [24].

An illustration of an FDI attack on a UAV is shown in Fig.1. This figure shows the gates that attacker can use for penetration in the system; like inertial measurement units (IMU), GPS satellites, and communication protocols (CP). In this paper we focused on the IMU sensors. IMU is an integrated circuit contains accelerometers and gyroscope which gives the information related to linear and angular motion of the aircraft. This information is used as a feedback for the control system to modify the aircraft attitude based on the desired manoeuvre.

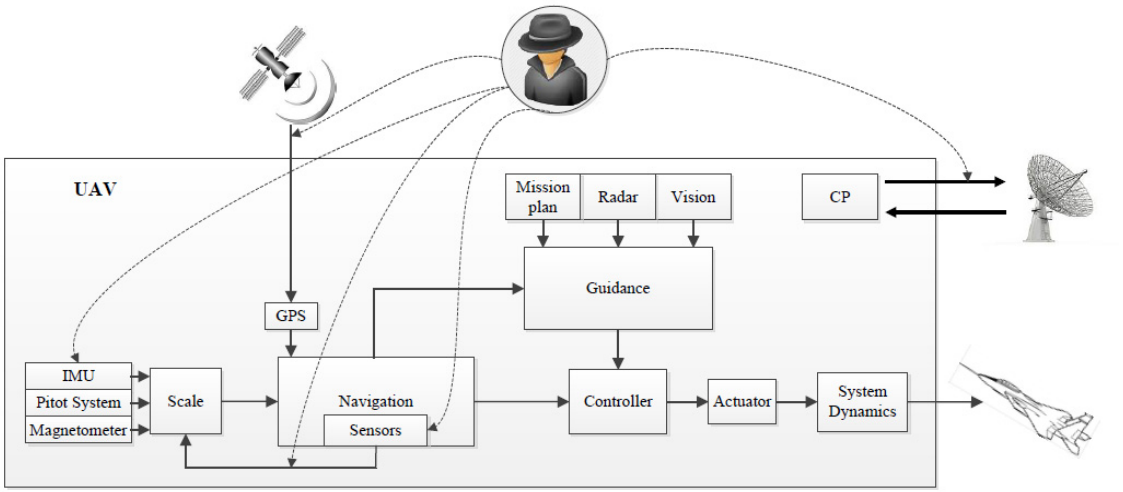

Fig. 1. Potential sub-systems that cyber attackers can penetrate in the UAV system

3. Fault Data Injection Detection Design

Faults can be injected in the sensors described by the following nonlinear system

$$\begin{aligned} \dot{x}(t) &= f(x(t)) + g(x(t))u(t) + D(x,t) \\ y(t) &= h(x(t)) + f_s(x,t) \end{aligned} \tag{1}$$

where $u(t) \in R^m$ is the input vector, $y(t) \in R^r$ is the output vector, $x(t) \in R^n$ is the state vector, $f : R^n \rightarrow R^n$ is the state function, g: $R^n \rightarrow R^{n \times m}$ is the input function, $D \in R^n$ is sensor uncertainties and disturbances, h: $R^n \rightarrow R^r$ is the output function, and $f_s(x,u)$ is the injected fault data in sensors, whose elements describe the faults in the system.

The $f_s(x,u)$ attack can be modelled as follows:

$$f_s(x,t) = \begin{cases} 0 & otherwise \\ z & attack \end{cases} \tag{2}$$

where z is an input signal intended by the attacker for the purpose of either misleading the control system, causing systems inefficiencies, or sabotaging it.

3.1. Neural network adaptive structure design

The proposed neural network adaptive structure (NNAS) which is used for FDI attack detection is developed in this subsection. Injected faults do not have any specific pattern and can be very complicated in their structure; hence, NN can be a suitable candidate for estimating them. Unlike the direct NN modelling procedures, NNAS estimates faults based on the nonlinear observer output and sensor output by using

$$\begin{aligned} \dot{\hat{x}} &= f(\hat{x}(t)) + g(\hat{x}(t))u(t) \\ \hat{y}(t) &= h(\hat{x}(t)) + M(t) \end{aligned} \tag{3}$$

where $\hat{x}(t)$ is the state vector of the nonlinear observer and $M(t)$ is the neural network observer at time t, that can be defined as follow [12]:

$$M_i(t) = W_i(t)\sigma(V_i(t)I_i(t)) \tag{4}$$

where, $M_i(t)$ is the i-th vector of observer input $M(t)$ for $i = 1, \ldots, n$. $W_i(t)$ and $V_i(t) = [V_{i,1}(t), \ldots, V_{i,m+n}(t)]$ are the i-th NNAS input parameters at time t. $\sigma(\bullet)$ is the sigmoid activation function, and it is called 'tansig' or 'logsig' in neural networks. Here, the activation function considered is $\sigma(x) = (1-e^{-x})/(1+e^{-x})$, and $I_i(t)$ can be defined by $I_i(t) = [M_i(t-\tau),...,M_i(t-m\tau),e_i(t-\tau),...,e_i(t-n\tau)]^T$. where τ indicate the sampling period or the step size, and $e_i(t)$ is the i-th estimated fault variable of the output $e_i(t) = h(x(t)) - h(\hat{x}(t))$. Here, m and n are chosen based on the needed speed response in the system. Large values of m and n guaranty the convergence of the training; however, they may increase the computation time and add unnecessary delays [25]. The input of observer $M(t)$ is recursively updated with the previous m samples of the observer inputs $M(k-i\tau)$ for $i = 1,2,\ldots,m$, and also previous n samples of the system output error $e_i(t-i\tau)$ for $i = 1,2,\ldots,n$.

3.2. Neural network weight update law

In order to have fast fault detection, neural network weights should be tuned [26]. Here an adaptive parameter tuning algorithm based on Extended Kalman Filter (EKF) is introduced. The EKF helps to update the NN weighting parameters online, so, fast convergence rate of the NN learning will be guaranteed. Through the updating process, if we consider the i-th element of NAS, then the EKF updating parameter can be described by [25]:

$$\theta_i(k) = [W_i(k), V_{i,1}(k), ..., V_{i,p+q}(k)]^T \tag{5}$$

where k is the k-th sampling instant, and the relation between variable k and time variable t is $k = \lfloor t/\tau \rfloor$. The parameters will be calculated in each sampling time with the following rules [12]

$$\begin{aligned} \theta_i(k) &= \theta_i(k-1) + \eta_i K_i(k)[y_i(k) - \hat{y}_i(k)] \\ K_i(k) &= P_i(k)H_i(k)[H_i(k)^T P_i(k)H_i(k) + R_i(k)]^{-1} \\ P_i(k+1) &= P_i(k) - K_i(k)H_i(k)^T P_i(k) \end{aligned} \tag{6}$$

where η_i is the learning coefficient, $P_i(k)$ is the covariance matrix of the state estimation error, $K_i(k)$ is the Kalman gain, and $R_i(k)$ is the covariance matrix of the estimated noise, which is computed recursively by [27]:

$$R_i(k) = R_i(k-1) + [e_i(k)^2 - R_i(k-1)]/k \tag{7}$$

Here $H_i(k)$ is the derivative of $e_i(k)$ with respect to θ_i. Based on the observer input in equation (4), $H_i(k)$ can be calculated as follow:

$$H_i(k) = \left.\frac{\partial e_i(k)}{\partial \theta_i}\right|_{\theta_i=\theta_i(k-1)} = \begin{cases} \sigma(Z_i(k)) & \theta_i = W_i \\ W_i(k) M_i(k-j) \sigma'(Z_i(k)) & \theta_i = V_{i,j} \\ W_i(k) e_i(k-j) \sigma'(Z_i(k)) & \theta_i = V_{i,p+j} \end{cases} \tag{8}$$

where

$$Z_i(k) = \sum_{j=1}^{p} V_{i,j}(k) M_i(k-j) + \sum_{j=1}^{q} V_{i,p+j}(k) e_i(k-j) \tag{9}$$

In this paper, we consider that attacker injects faults to angular rate sensors, i.e., *p*, *q* and *r*. Typical motion equations of the airplane angular rate with faults can described as

$$\begin{aligned} \dot{x} &= f(x) + g(x)u + D(x,t) \\ y &= [p \quad q \quad r]^T + N_s + f_s(t) \\ x &= [p \quad q \quad r]^T \\ f(x) &= [f_p \quad f_q \quad f_r]^T \\ g(x) &= \begin{bmatrix} L_{\delta A} & 0 & L_{\delta R} \\ 0 & M_{\delta E} & 0 \\ N_{\delta A} & 0 & N_{\delta R} \end{bmatrix} \end{aligned} \tag{10}$$

where $D(x,u)$ is the system uncertainties, N_S is the sensor noise; $f_s(t)$ is the fault function which occurs in the sensors; $L_{\delta A}$, $L_{\delta R}$, $M_{\delta E}$, $N_{\delta A}$, $N_{\delta R}$ are rolling, pitching and yawing moment about the control input deflection. f_p, f_q and f_r can be obtained from the aircraft nonlinear equation of the motion [28, 29].

$$\begin{aligned} f_p &= I_z l_{aero} + I_{xz} n_{aero} + I_{xz}(I_x - I_y + I_z) pq \\ &\quad + \frac{I_z(I_y - I_z) - I_{xz}^2) qr}{I_x I_z - I_{xz}^2} \\ f_q &= m_{aero} + pr(I_z - I_x) + \frac{I_{xz}(r^2 - p^2)}{I_y} \\ f_r &= I_{xz} l_{aero} + I_x n_{aero} + (I_x(I_x - I_y) + I_{xz}^2) pq \\ &\quad - \frac{I_{xz}(I_x - I_y + I_z) qr}{I_x I_{xz} - I_{xz}^2} \end{aligned} \tag{11}$$

Based on the above description, an NNAS is designed by

$$\begin{aligned} \hat{\dot{x}} &= f(\hat{x}) + g(\hat{x})u \\ \hat{y} &= [\hat{p} \quad \hat{q} \quad \hat{r}]^T + M_i(t) \end{aligned} \tag{12}$$

where $\underline{M_i(t)}$ is can be calculated using Equation (4-9)

$$M_i(t) = W_i(t) \sigma\left(\sum_{j=1}^{3} V_{i,j}(t) M_i(t - j\tau) + V_{i,4}(t) \tilde{y}_i(t-\tau)\right) \tag{13}$$

The overall structure of the proposed NNAS method is depicted in Fig.2.

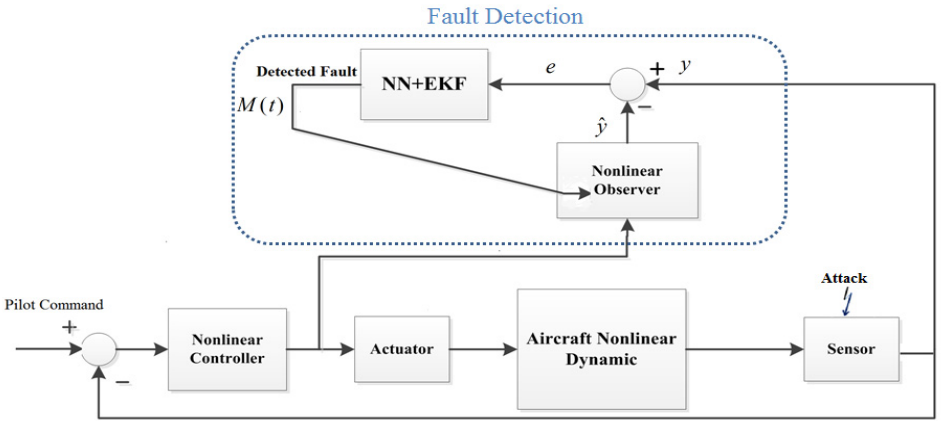

Fig. 2. Overall diagram of proposed NNAS design for FDI detection.

4. Numerical Simulation

In this section two different FDI attacks are simulated to examine the performance of the proposed NNAS technique in detection of the FDI attack in UAV's sensors. The proposed fault detection algorithm has been implemented and evaluated on a six-degree-of-freedom (DoF) aircraft model, i.e., a nonlinear model of WVU YF-22 unmanned aircraft [29]. Two scenarios have been considered in the simulated attacks. In the first scenario a step shape attack is inserted in the IMU sensors. In the second scenario, a Gaussian shape attack is inserted to the IMU sensors. In both scenarios a random signal with the amplitude of 0.08 radian is also added as a noise to the considered attacks. Simulations are done separately for the attacks in *p*, *q*, and *r* which are the rolling, pitching, and the yawing rate of the aircraft, respectively. The results of the simulations are presented in Fig.3 and Fig.4. Fig.3 shows that the proposed NNAS design can successfully detect the FDI attack with sudden changes. The step FDI attack has a sharp edge which is difficult to be detected by conventional algorithms, while, as it can be seen the proposed NNAS design detected it fast. Fig. 4 shows the proposed design is able to detect smooth attacks in the sensors as well. The chattering in Fig.3 and Fig.4 is related to the considered noise and the attack scenarios. In Fig.3 and Fig.4 solid line is the FDI attack and dashed line is the estimated attack using proposed detection technique.

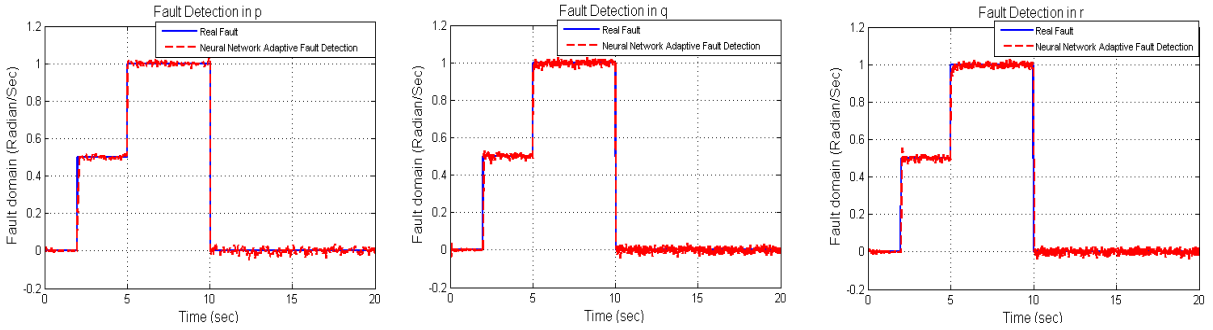

Fig.3. FDI attack detection using proposed technique in presence of a step shaped attack

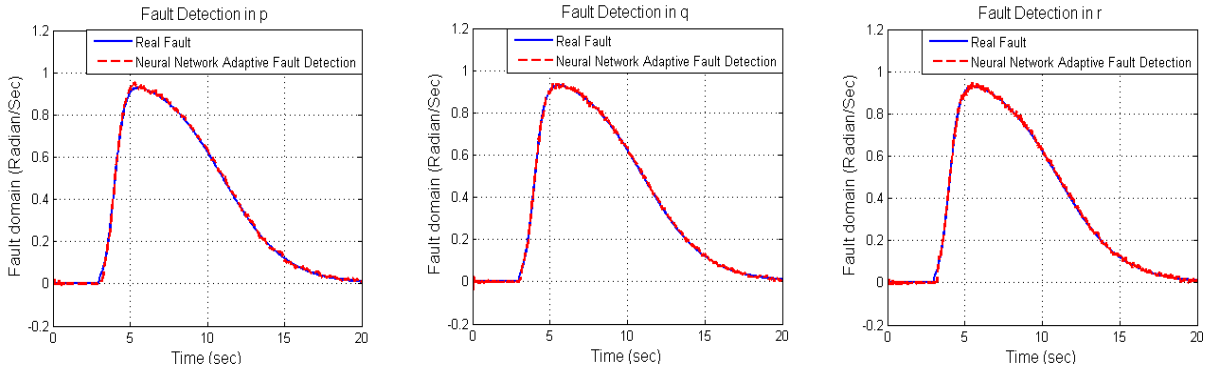

Fig.4. FDI attack detection using proposed technique in presence of a Gaussian shaped attack

5. Conclusion and future works

In this work the possible cyber-attacks to the aircraft attitude sensors have been investigated. A new approach based on neural network observer is introduced which is capable of online detection of possible attacks on the UAV sensors in the IMU. The proposed design uses EKF to tune the NN weights which increases the learning speed of the NN, and subsequently, improves the ability of the detection of the sudden attacks. The simulation results show that the developed method can successfully and accurately detect the sudden and smooth attacks in the sensors. This detection can be further used to help the system to correct itself and to be robust against cyber-attacks.

References

1. Kim, A., et al., *Cyber attack vulnerabilities analysis for unmanned aerial vehicles.* Infotech@ Aerospace, 2012.
2. Dofe, J., et al., *Strengthening SIMON Implementation Against Intelligent Fault Attacks.* Embedded Systems Letters, IEEE, 2015. **7**(4): p. 113-116.
3. Zhang, F., et al., *A Framework for the Analysis and Evaluation of Algebraic Fault Attacks on Lightweight Block Ciphers.* 2013.
4. Wang, B., et al., *Against Double Fault Attacks: Injection Effort Model, Space and Time Randomization Based Countermeasures for Reconfigurable Array Architecture.* 2013.
5. Gholami, H., A. Khalilnejad, and G.B. Gharehpetian, *Electrothermal performance and environmental effects of optimal photovoltaic–thermal system.* Energy Conversion and Management, 2015. **95**: p. 326-333.
6. Khalilnejad, A., A. Sundararajan, and A. Sarwat. *Performance evaluation of optimal photovoltaic-electrolyzer system with the purpose of maximum Hydrogen storage.* in *2016 IEEE/IAS 52nd Industrial and Commercial Power Systems Technical Conference (I&CPS)*. 2016. IEEE.
7. Khalghani, M.R., et al., *Modifying power quality's indices of load by presenting an adaptive method based on Hebb learning algorithm for controlling DVR.* AUTOMATIKA: časopis za automatiku, mjerenje, elektroniku, računarstvo i komunikacije, 2014. **55**(2): p. 153-161.
8. Khalghani, M.R. and M.H. Khooban, *A novel self-tuning control method based on regulated bi-objective emotional learning controller's structure with TLBO algorithm to control DVR compensator.* Applied Soft Computing, 2014. **24**: p. 912-922.
9. Chaojun, G., P. Jirutitijaroen, and M. Motani, *Detecting False Data Injection Attacks in AC State Estimation.* Smart Grid, IEEE Transactions on, 2015. **6**(5): p. 2476-2483.
10. Rawat, D.B. and C. Bajracharya, *Detection of false data injection attacks in smart grid communication systems.* Signal Processing Letters, IEEE, 2015. **22**(10): p. 1652-1656.
11. Sargolzaei, A., K.K. Yen, and M.N. Abdelghani, *Preventing Time-Delay Switch Attack on Load Frequency Control in Distributed Power Systems.*
12. Wu, Q. and M. Saif. *Repetitive learning observer based actuator fault detection, isolation, and estimation with application to a satellite attitude control system.* in *American Control Conference, 2007. ACC'07*. 2007. IEEE.
13. Samy, I., I. Postlethwaite, and D.-W. Gu, *Survey and application of sensor fault detection and isolation schemes.* Control Engineering Practice, 2011. **19**(7): p. 658-674.
14. Talebi, H. and R. Patel. *An intelligent fault detection and recovery scheme for reaction wheel actuator of satellite attitude control systems*. in *Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control, 2006 IEEE*. 2006. IEEE.
15. Chen, M., P. Shi, and C.-C. Lim, *Adaptive Neural Fault-Tolerant Control of a 3-DOF Model Helicopter System.* 2015.
16. Shen, Q., et al., *Novel neural networks-based fault tolerant control scheme with fault alarm.* Cybernetics, IEEE Transactions on, 2014. **44**(11): p. 2190-2201.
17. Yang, X., N. Wu, and J.H. Andrian, *A novel bus transfer mode (AS transfer) and a performance evaluation methodology.* Integration,

the VLSI Journal, 2016. **52**: p. 23-33.
18. Ghanavati, M., S. Mobayen, and V.J. Majd. *A new robust model predictive control strategy for rotational inverted pendulum system*. in *Control and Communications (SIBCON), 2011 International Siberian Conference on*. 2011. IEEE.
19. Ghanavati, M., V.J. Majd, and M. Ghanavati. *Control of inverted pendulum system by using a new robust model predictive control strategy*. in *In: Proceedings of International Siberian Conference on Control and Communications*. 2011.
20. Ghanavati, M. and A. Chakravarthy, *Demand-side energy management using an adaptive control strategy for aggregate thermostatic loads*, in *AIAA Infotech@ Aerospace*. 2015. p. 0121.
21. Zeng, K., et al., *Local Visual Feature Detection and Description for Non-Rigid 3D Objects.* Advances in Image and Video Processing, 2016. **4**(2): p. 01.
22. Ghanavati, M. and A. Chakravarthy. *Demand-side energy management by use of a Design-then-Approximate controller for aggregated thermostatic loads*. in *2015 American Control Conference (ACC)*. 2015. IEEE.
23. Di-Battista, J., et al., *When failure analysis meets side-channel attacks*, in *Cryptographic Hardware and Embedded Systems, CHES 2010*. 2010, Springer. p. 188-202.
24. Moradi, A., et al., *On the power of fault sensitivity analysis and collision side-channel attacks in a combined setting*, in *Cryptographic Hardware and Embedded Systems–CHES 2011*. 2011, Springer. p. 292-311.
25. Wu, Q. and M. Saif. *Neural adaptive observer based fault detection and identification for satellite attitude control systems*. in *American Control Conference, 2005. Proceedings of the 2005*. 2005. IEEE.
26. Abaspour, A., S.H. Sadati, and M. Sadeghi, *Nonlinear optimized adaptive trajectory control of helicopter*. Control Theory and Technology, 2015. **13**(4): p. 297-310.
27. Ljung, L., *S oderstr om, T.(1983). Theory and Practice of Recursive Identification*. MIT Press, Cambridge, MA.
28. Abaspour, A., M. Sadeghi, and H. Sadati. *Using fuzzy logic in dynamic inversion flight controller with considering uncertainties.* in *13th Iranian Conference on Fuzzy Systems (IFSC)*. 2013.
29. Sadeghi, M., A. Abaspour, and S.H. Sadati, *A Novel Integrated Guidance and Control System Design in Formation Flight.* Journal of Aerospace Technology and Management, 2015. **7**(4): p. 432-442.