Journal of Electrical Engineering and Automation (EEA) (2021)
Vol.03/ No.01
Pages: 24-33
https://www.irojournals.com/iroeea
DOI: https://doi.org/10.36548/jeea.2021.1.003

False Data Detection in Smart Grid using Artificial Intelligence

S. R. Mugunthan,
Associate Professor,
Department of Computer Science and Engineering,
Sriindu College of Engineering and Technology,
Hyderabad, India.
srmugunth@gmail.com

Dr. T. Vijayakumar,
Professor,
Department of ECE,
Guru Nanak Institute of Technology,
Hyderabad, India.
vishal_16278@yahoo.co.in

Abstract: In order to increase the utilization of artificial intelligence in smart grids, it is necessary to have an accurate state estimation. This criterion is an essential aspect, along with other functionalities for successful control and monitoring. As the internet and utility network form an increasing interconnectivity, it leaves the state estimators in a state of vulnerability to various attacks like bad data detection and false data injection. Though there are many research-works done on detectors for false data detection, depending on the contingencies, the counter measure will also vary. A sudden change physically will have a high impact on the available data, resulting in incorrect classification of the future instances. As a means of addressing this issue, we have analyzed the differences between data manipulation change and physical grid change for better understanding. Focusing on distribution change, we used outage and have introduced analysis of historical data. The goal is to determine the important aspects thereby identifying the scope. We have also used statistical hypothesis and dimensionality reduction for testing purpose. We have used IEEE 14 bus system for evaluation based on the scenario of attack: under concept drift and without concept drift. The result shows a more accurate output when compared with the other previously existing methodologies using concept drift.

ISSN: 2582-3051
Submitted: 23.03.2021
Revised: 5.04.2021
Accepted: 30.04.2021
Published: 12.05.2021

Journal of Electrical Engineering and Automation (EEA) (2021)
Vol.03/ No.01
Pages: 24-33
https://www.irojournals.com/iroeea
DOI: https://doi.org/10.36548/jeea.2021.1.003

Keywords: Machine Learning; Line outage; Data Integrity attacks; Smart Grid; False data injection

1. Introduction

Smart grid depends on the ability to communicate to ensure proper control and operation of the power system. However, this dependency has created a strong reliance making the grid open to a number of malicious attacks which decreases the credibility of smart grids and might even debilitate the crucial infrastructure. This will result in severe substantial financial loss, social unrest and operational failures [1]. A good example is the Ukrainian electrical grid cyber attack that took place in December 2015. It resulted in a power outage for over 200,000 customers through circuit breakers. One of the most crucial type of cyber-attacks is that of false data injection (FDI) that makes the results of SE inaccurate. Based on the attacker's knowledge about the power grid connectivity, the effectiveness of a typical FDI [2] attack will vary. Once these attackers enter the power grid system, they have the ability to exploit multiple channels to obtain information illegally like cyber-public channels, physical channels and cyber channels. Any wrong estimate will lead the system in a different direction, putting the security of the power system at risk of drastic consequences like blackouts. This is primarily because the information about the network is implemented in other parts of the management like load shedding and transmission stability analysis, etc [3]. In particular, an attack can be initiated by changing the readings of phasor measurement units and multiple sensors to inject malicious values such that it will result in arbitrary errors that remain undetected. This paper is organised such that the second session give the brief outline of the related work. The proposed methodology is prescribed in section 3 and the evaluation results are observed and graphically represented in section 4. Based on the examination conclusion is drawn in section 5 identifying the effectiveness of the proposed methodology.

2. Related Works

To prevent FDI attacks [4] during control and operation of the system a number of mitigation methods and several detectors have been developed. There are two categories which vitamin the countermeasures namely detection based and protection based. The

ISSN: 2582-3051
Submitted: 23.03.2021
Revised: 5.04.2021
Accepted: 30.04.2021
Published: 12.05.2021

Journal of Electrical Engineering and Automation (EEA) (2021)
Vol.03/ No.01
Pages: 24-33
https://www.irojournals.com/iroeea
DOI: https://doi.org/10.36548/jeea.2021.1.003

detection based methods depend on the type of anomaly detected in order to determine modification of the measurements in a malicious manner such that it does not comply with the distribution of historic readings such as statistical threshold testing, classification algorithms, Kalman filter and graph theory etc. On the other hand protection based methodology will attempt to identify the FDI attack by protecting critical metres and identifying the attacks [5]-[6]. However the drawback with this methodology deals with unassured effectiveness of protection, decrease in measurement redundancy and increased cost of implementation when used in large scale power systems. In particular the detection based methodology will evaluate the underlying data distribution using historical data and further utilise it to determine future attacks that might result in high deviation with respect to the reference distribution [7].

To qualitatively [8] compare the proposed approach with relevant data, four characteristics have been analysed. They are as follows:

- Required external devices: This attribute will indicate if the method used to detect attacks or to protect the system depends on external devices.
- Applicable to large scale power systems: this attribute indicates if the method of implementation is costly and computationally complex.
- Attack localisation: this characteristic shows if the method proposed will be able to identify the location of attack.
- Dealing with contingencies: this character portrays the capacity of the approach to determine attacks under the conditions when a line outage contingency occur in the system.

However lines on these characteristics have a negative impact in a compromise of the device will not prove effective [9]-[10]. However the positive aspect of the existing attack detection methodologies is there dependency on assumption of statistic discovered patterns and time invariant historical data that makes it apt for stationary information. This indicates that this approach has been introduced for a specific system configuration and has not taken into account the effect of topology reconfigurations [11]. But in practical circumstances the

ISSN: 2582-3051
Submitted: 23.03.2021
Revised: 5.04.2021
Accepted: 30.04.2021
Published: 12.05.2021

Journal of Electrical Engineering and Automation (EEA) (2021)
Vol.03/ No.01
Pages: 24-33
https://www.irojournals.com/iroeea
DOI: https://doi.org/10.36548/jeea.2021.1.003

data involved will vary which time and the distribution of data is not stationery. This indicates that the data set developed cannot be used when topology changes. Hence the predictions that are made using the historical data and its subsequent trained or developed model will not result in an accurate value as there is a shift in distribution of data leading to two unrelated old findings that will not have any impact on the new findings [12].

3. Development of Use Case Scenario

There are number of existing detection methods that are used to handle concept drift caused because of branch outage contingencies. However, these methodologies will not be able to the check the attacks in a stealthy manner. Hence we have proposed a noble approach that will complement the existing detection methodology insuring that they are able to identify the attacks after concept drift. In particular K-Nearest Neighbour (KNN) [13] proves to be highly efficient in detecting FDI attacks and has an outstanding performance in adapting to variation in data distribution when compared with other methodologies. The proposed architecture illustrated is in figure 1. Initially the baseline of the concept is built on system topology without contingency. Further depending on the branch outage contingency the network can be designed. This gives a possibility for a drift from Cbase to Ci concept such that a line outage contingency exists. However, the existing methods are not stable against topology changes and will not be able to differentiate between a normal sample and one that is under attack. Hence in this proposed work we have used not only the baseline concept historical data but also that of crucial concepts which are caused due to contingency from line outage resulting in noticeable variation of data distribution.

ISSN: 2582-3051
Submitted: 23.03.2021
Revised: 5.04.2021
Accepted: 30.04.2021
Published: 12.05.2021

Journal of Electrical Engineering and Automation (EEA) (2021)
Vol.03/ No.01
Pages: 24-33
https://www.irojournals.com/iroeea
DOI: https://doi.org/10.36548/jeea.2021.1.003

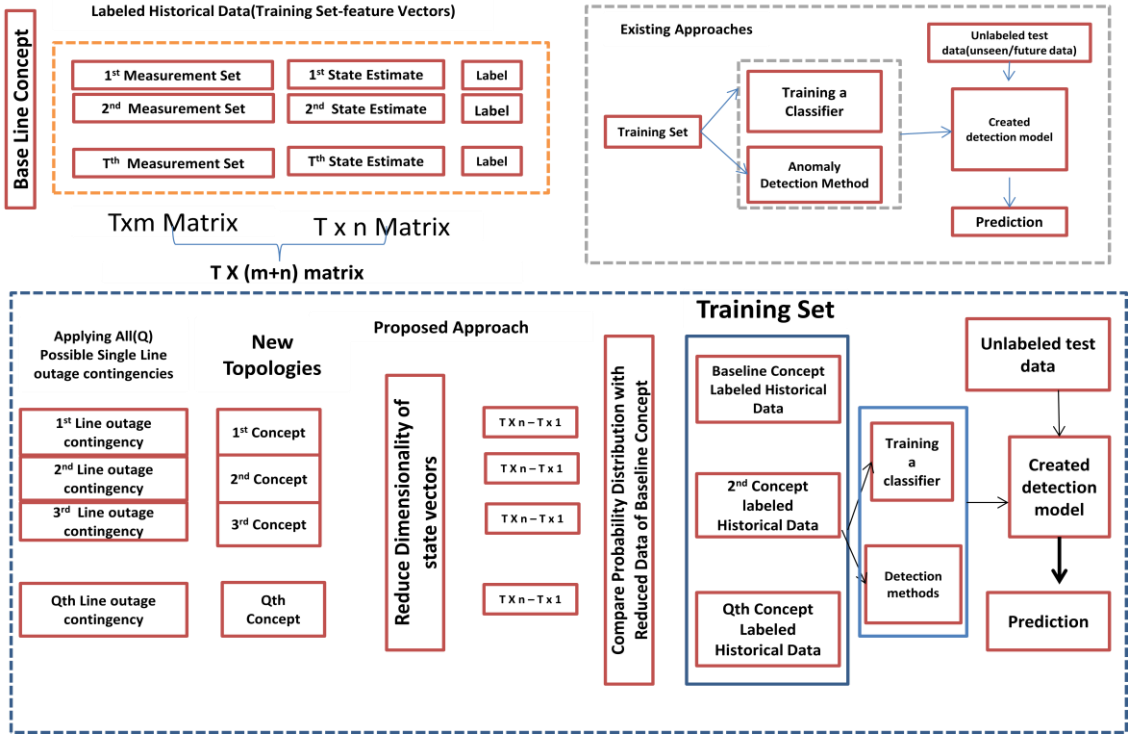


Fig.1. Architecture of Proposed Model

A contingency analysis is used to evaluate the outage events especially in the case of transmission power systems. This analysis plays a crucial role in assessing the security of the system. The purpose of the choice in line outage contingency analyse is to determine the contingencies that are not dependable. In general the choice of contingency methods depends on the performance index calculated using fast decoupled or DC load flow solution. The choice is based on identification of line outages that cause a dramatic change in the underlying data distribution. Determining the critical concept is vital because the choice of concept will change the data distribution drastically, enabling the classifier to understand unidentified the drift of the incoming samples even before the drift. According to the shift in distribution of underlying data the choice of critical concept will differ in this proposed work the first step involves determining all the concepts that could be drafted by network topology from the baseline because of branch failure contingency. However, this is exclusive of the data obtained from load flow solution and system divergence.

ISSN: 2582-3051
Submitted: 23.03.2021
Revised: 5.04.2021
Accepted: 30.04.2021
Published: 12.05.2021

Journal of Electrical Engineering and Automation (EEA) (2021)
Vol.03/ No.01
Pages: 24-33
https://www.irojournals.com/iroeea
DOI: https://doi.org/10.36548/jeea.2021.1.003

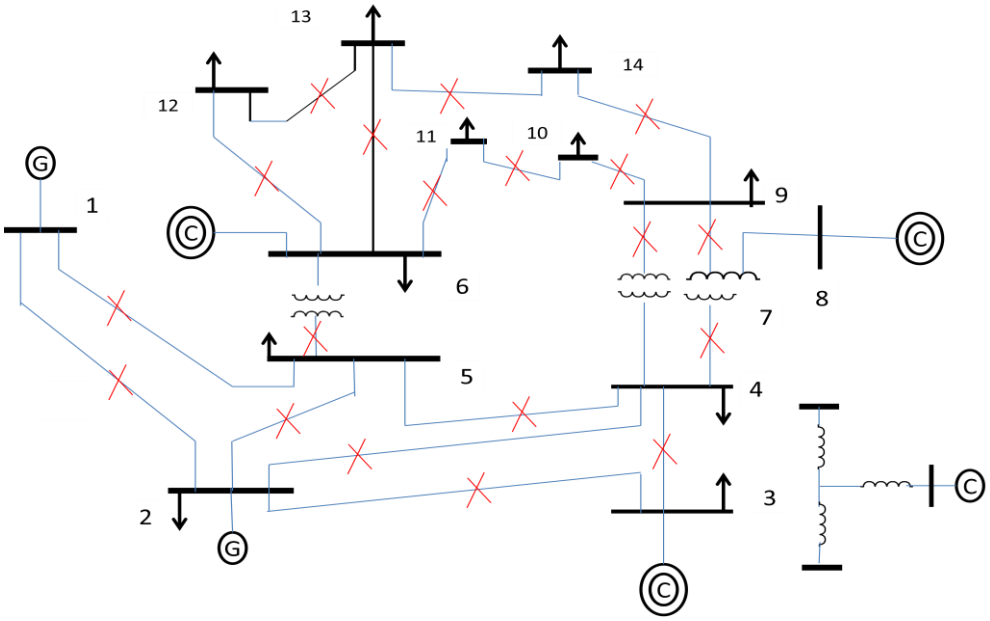


Fig.2. IEEE Bus Test System

By applying the line outage contingencies for all possible solutions we generate data follow various concepts, $C_{i==1,2,\ldots Q}$ as shown in Fig.2. This indicates that we eliminate a line and repeatedly perform power flow in order to determine the normal sequence of $a_m = a_1; a_2; \ldots . a_w; \ldots ; a_T$, such that T represents the total number of data points collected and w is the time index. Accordingly, the output of this matrix function is the collection of new derived concepts and normal vectors of the baseline concept:

$$\begin{bmatrix} a_{0,1} & a_{0,2} & \ldots & a_{0,T} \\ a_{1,1} & a_{1,2} & \ldots & a_{1,T} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,T} \\ a_{Q,1} & a_{Q,2} & \ldots & a_{Q,T} \end{bmatrix} \rightarrow \begin{bmatrix} am_{Cbase} \\ am_{c1} \\ am_{c2} \\ . \\ . \\ am_{CQ} \end{bmatrix}$$

Principal Component Analysis (PCA) is used to decrease computational complexities and to decrease the dimension of the vectors used. Hence the new vector space can be constructed as shown below:

ISSN: 2582-3051
Submitted: 23.03.2021
Revised: 5.04.2021
Accepted: 30.04.2021
Published: 12.05.2021

Journal of Electrical Engineering and Automation (EEA) (2021)
Vol.03/ No.01
Pages: 24-33
https://www.irojournals.com/iroeea
DOI: https://doi.org/10.36548/jeea.2021.1.003

$$\begin{bmatrix} b_{0,1} & b_{0,2} & \dots & b_{0,T} \\ b_{1,1} & b_{1,2} & \dots & b_{1,T} \\ b_{2,1} & b_{2,2} & \dots & b_{2,T} \\ b_{Q,1} & b_{Q,2} & \dots & b_{Q,T} \end{bmatrix} \rightarrow \begin{bmatrix} PC_{1,1} \\ PC_{2,1} \\ PC_{3,1} \\ . \\ . \\ PC_{T,1} \end{bmatrix}$$

4. Results and Discussion

An IEEE 14 bus system is used to simulate the results for the proposed work. It has been observed that for a particular bus, in central region, the distribution of normalized load value will represent over the profile power flow as shown in Fig.3.

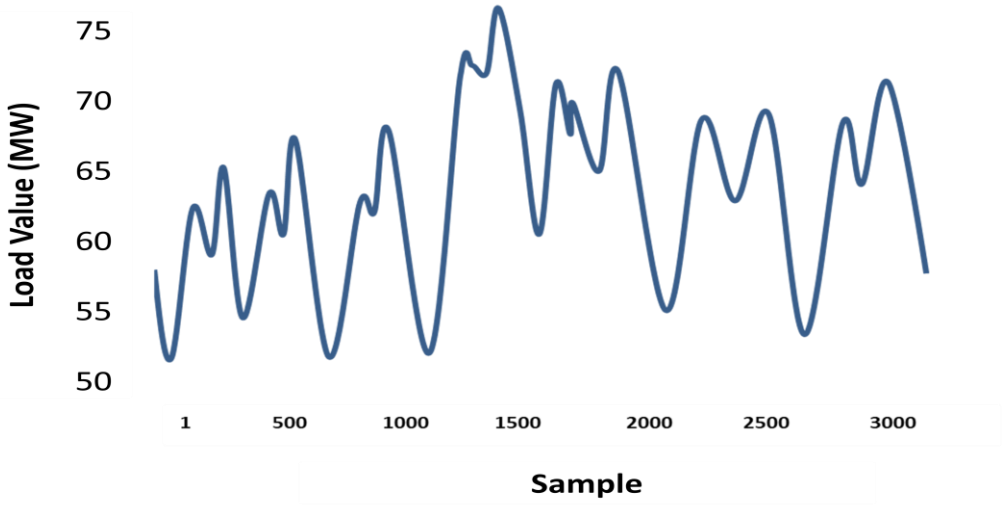
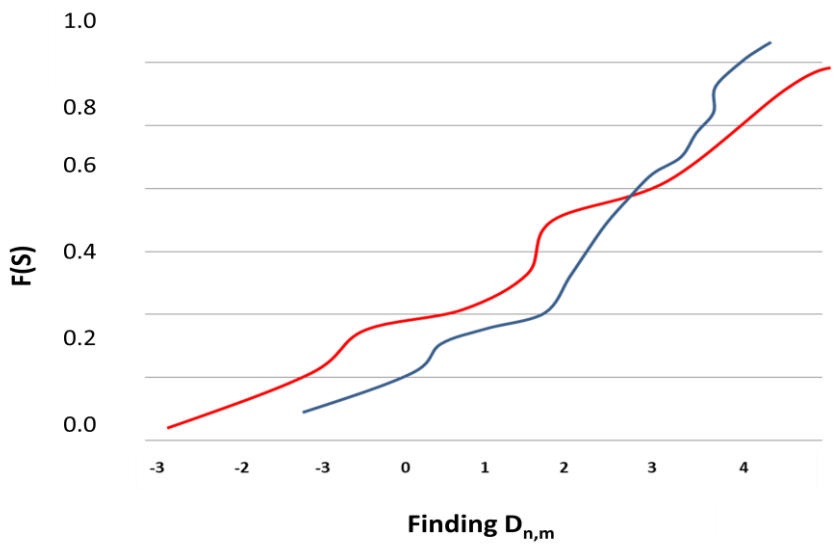


Fig.3. Normalized load data

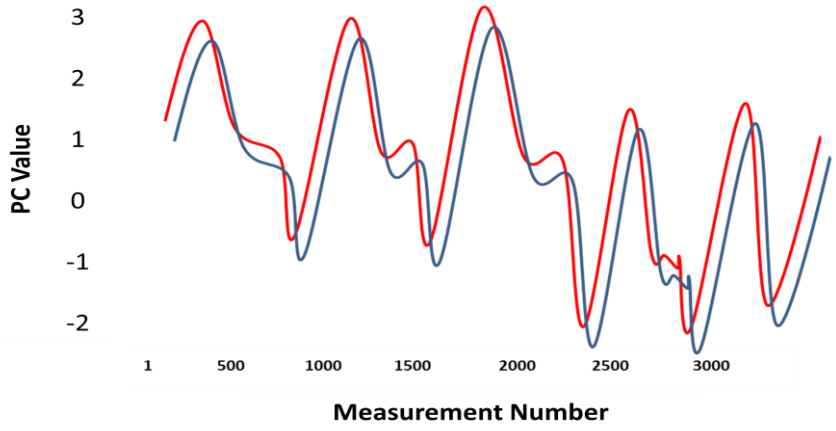
With respect to drift that occurs in baseline concept, Fig.4 shows an example of the methodology wherein a new concept is rejected hypothetically representing the highest vertical distance. Moreover, it is also used to determine the variation of PV values with respect to reduced xm and determining the distance with respect to the F(s) movement.

ISSN: 2582-3051
Submitted: 23.03.2021
Revised: 5.04.2021
Accepted: 30.04.2021
Published: 12.05.2021

Journal of Electrical Engineering and Automation (EEA) (2021)
Vol.03/ No.01
Pages: 24-33
https://www.irojournals.com/iroeea
DOI: https://doi.org/10.36548/jeea.2021.1.003



a) Finding $D_{n,m}$



b) Reduced xm

Fig.4. Example of Proposed work

5. Conclusion

FDI attacks will prove to be a severe threat to control and operate the smart grid. Though a number of mitigation methods and detectors exist, they are based on the assumption that the future unseen data and training is from one distribution and do not have

ISSN: 2582-3051
Submitted: 23.03.2021
Revised: 5.04.2021
Accepted: 30.04.2021
Published: 12.05.2021

Journal of Electrical Engineering and Automation (EEA) (2021)
Vol.03/ No.01
Pages: 24-33
https://www.irojournals.com/iroeea
DOI: https://doi.org/10.36548/jeea.2021.1.003

the capacity to manage concept drift. Hence, in this work, we have incorporated a novel methodology that ensures robustness of the system. This indicates that we have also taken into consideration, the critical concept set apart from the baseline concept historical data which will hold an impact on data distribution, accordingly influencing the training set. KS and PCA tests are used to determine these sets. We have also incorporated the KNN algorithm in order to improve the effectiveness of the work. This work is capable of identifying the critical concepts depending on the intensity with which the data distribution of changes, instead of power system indices that indicate the contingency set. Evaluation results indicate that the proposed work will be able to attain a higher rate of accuracy and is also known for its robustness.

References

[1] Niu, X., Li, J., Sun, J., & Tomsovic, K. (2019, February). Dynamic detection of false data injection attack in smart grid using deep learning. In *2019 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)* (pp. 1-6). IEEE.

[2] Manandhar, K., Cao, X., Hu, F., & Liu, Y. (2014). Detection of faults and attacks including false data injection attack in smart grid using Kalman filter. *IEEE transactions on control of network systems*, *1*(4), 370-379.

[3] He, Y., Mendis, G. J., & Wei, J. (2017). Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism. *IEEE Transactions on Smart Grid*, *8*(5), 2505-2516.

[4] Bhalaji, N. (2020). EL DAPP–An Electrıcıty Meter Trackıng Decentralızed Applıcatıon. *Journal of Electronics*, *2*(01), 49-71.

[5] Huang, Y., Tang, J., Cheng, Y., Li, H., Campbell, K. A., & Han, Z. (2014). Real-time detection of false data injection in smart grid networks: An adaptive CUSUM method and analysis. *IEEE Systems Journal*, *10*(2), 532-543.

[6] Shirley, D. R. A. (2014, July). Systematic diagnosis of power switches. In *2014 International Conference on Embedded Systems (ICES)* (pp. 32-34). IEEE.

[7] Esmalifalak, M., Liu, L., Nguyen, N., Zheng, R., & Han, Z. (2014). Detecting stealthy false data injection using machine learning in smart grid. *IEEE Systems Journal*, *11*(3), 1644-1652.

ISSN: 2582-3051
Submitted: 23.03.2021
Revised: 5.04.2021
Accepted: 30.04.2021
Published: 12.05.2021

Journal of Electrical Engineering and Automation (EEA) (2021)
Vol.03/ No.01
Pages: 24-33
https://www.irojournals.com/iroeea
DOI: https://doi.org/10.36548/jeea.2021.1.003

[8] Wei, L., Gao, D., & Luo, C. (2018, November). False data injection attacks detection with deep belief networks in smart grid. In *2018 Chinese Automation Congress (CAC)* (pp. 2621-2625). IEEE.

[9] Chekired, D. A., Khoukhi, L., & Mouftah, H. T. (2019, May). Fog-based distributed intrusion detection system against false metering attacks in smart grid. In *ICC 2019-2019 IEEE International Conference on Communications (ICC)* (pp. 1-6). IEEE.

[10] Bestak, R., & Smys, S. (2019). Big data analytics for smart cloud-fog based Applications. *Journal of trends in Computer Science and Smart technology (TCSST)*, *1*(02), 74-83.

[11] Smys, S. (2020). A Survey on Internet of Things (IoT) based Smart Systems. *Journal of ISMAC*, *2*(04), 181-189.

[12] Karthiban, M. K., & Raj, J. S. (2019). Big data analytics for developing secure internet of everything. *Journal of ISMAC*, *1*(02), 129-136.

[13] Stephens, J. C., Wilson, E. J., & Peterson, T. R. (2015). *Smart grid (R) evolution*. Cambridge University Press.

ISSN: 2582-3051
Submitted: 23.03.2021
Revised: 5.04.2021
Accepted: 30.04.2021
Published: 12.05.2021