Hybrid Prediction of TVWS Channel Occupancy for Cognitive Radio Systems

Joy Lebese
Department of Electrical Engineering/FSATI
Tshwane University of Technology
Pretoria, South Africa
LebeseJ@tut.ac.za

Thomas O. Olwal
Department of Electrical Engineering/FSATI
Tshwane University of Technology
Pretoria, South Africa
Olwalto@tut.ac.za

Abstract— **The migration from analogue to digital broadcasting systems has created an opportunity for the utilization of the Television whitespace (TVWS) spectrum bands (that is, radio frequency spectrum bands between 470 MHz and 694 MHz). TVWS spectrum bands present opportunities for dynamic spectrum access in Cognitive Radio (CR) Communications. The four functions in CR systems generate processing time delays that reduce the throughput of the system. This research leveraged spectrum measurement campaigns to obtain real- time spectrum occupancy data for TVWS bands within a university campus. A hybrid spectrum occupancy prediction model constituting a combination of Hidden Markov and Bayesian models was developed. The developed hybrid spectrum occupancy prediction model showed a close correlation with a correlation factor of 89% and a probability prediction factor of 0.81, respectively in relations to the actual occupancy of the channel. The integrated solution had a better prediction accuracy than the Bayesian and the Hidden Markov spectrum occupancy prediction model that has a correlation factor of 65% and a probability prediction factor of 0.56. The Hidden Markov spectrum occupancy prediction had a correlation factor of 45% and a probability prediction factor of 0.71 with less computation and provided that all the input to the models is from the same real-time spectrum occupancy.**

Keywords— *Cognitive Radio, Frequency channel occupancy and Spectrum occupancy prediction.*

I. INTRODUCTION

In previous years, spectrum management was based on administrative registration, which allocated significant parts of spectrum to governments and licensed users. The goal of this strategy was to keep governments and licensed users safe from unwanted signals, which was a serious issue at the time [1]. According to recent studies, some of the designated Radio Frequency (RF) bands are being used inefficiently, resulting in a spectrum scarcity problem [2], [3]. However, inadequate spectrum utilization, has been identified as the major radio spectrum concern. Mitola and Maguire, in 1999, proposed Cognitive Radio (CR) technology as a solution to the problem of spectrum inadequate use [4]. Cognitive Radios (CRs) are transceivers that can intelligently recognize which spectrum channels are in use and which are not and move into vacant channels, while avoiding the ones that are occupied. CRs rely on spectrum detection, decision-making and sharing to be implemented successfully [4].

The behaviour and performance of a network of dynamic spectrum allocation/cognitive radios (DSA/CR) nodes are dependent on the primary spectrum occupancy pattern due to the opportunistic nature of the DSA/CR approach. In the sphere of DSA/CR research, accurate modelling of such patterns becomes vital and valuable [3]. The main functions of a CR are spectrum sensing, management, sharing and mobility [6]. Spectrum sensing allows CRs to scan the spectrum, detect when the primary or licenced user (PU) is not using a specific band and opportunistically enable a secondary (unlicensed) user (SU) to occupy the empty bands (white space), resulting in adequate utilisation of spectrum [2]. As a result, the majority of CR research to date has focused on spectrum sensing [3]. For CR to be fully implemented it relies mainly on four major procedures, that is, sensing, decision, sharing and mobility. However, the procedures cause delays in the network system.

Hence, there is a need to develop channel occupancy prediction models to alleviate the problem of delays caused in the sensing process [5]. The main purpose is to analyse current/past patterns that are likely to foster prediction of future occurrences or outcomes. Any prediction requires a history of previous observations as input to the prediction model. Spectrum prediction models have been widely used to reduce time delays caused by the effective implementation of CR [5]. However, such models are mainly theoretical with little practical value (data). Different methods for predicting spectrum occupancy and its relevance in CR systems have been studied and presented in literature. Binary time-series methods and non-binary time- series methods are the two primary types of these methods. Non-binary time series models employ the received signal power values for prediction, whereas binary time series models use the received signal power values for quantification [8].

This paper contributes to the existing knowledge by developing a hybrid spectrum occupancy prediction model which combines the Hidden Markov and Bayesian models. The resulting hybrid model is experimentally validated to determine suitability to alleviate the time delay caused in the CR spectrum occupancy measurements.

The rest of paper is organized as follows. Section II presents related work, section III provides the proposed model, section IV outlines the Algorithm, Section V discusses results and lastly section VI concludes the paper.

II. RELATED WORK

Frequency spectrum occupancy prediction in CR networks is a complex task with several subtopics, including channel status prediction, radio environment prediction and primary user (PU) activity prediction. For brevity of this paper, we describe a few commonly applied prediction methods in CR networks.

A. Markov models based spectrum occupancy prediction

The authors in [8] offer the N-order Markov model, which incorporates more historical data and enhances the prediction effect over the first-order Markov model. However, as the order of the model increases, the prediction impact decreases. The model's prediction latency increases exponentially as the model's complexity increases [8]. As a result, a Hidden Markov Model (HMM) based spectrum state prediction approach is suggested, considering the absence of history information in first-order Markov models and the rapid growth complications in the N-order Markov models. It is commonly utilised in the study of CR networks. Improving the hardware platform latency problem has been crucial. In [9], HMM's local prediction and cooperative prediction have been investigated. Improvements have been made primarily to the HMM collaborative prediction method and the high-order prediction algorithm. Moreover, the use of an HMM-based prediction model to increase the effectiveness of CR communication is investigated. A three-state HMM prediction model is developed to enhance spectrum use rate. Even so, all of these techniques assume that channel occupancy duration and idle time follow an exponential decay. This implies that channel previous knowledge is complete, whereas the reality of these conditions is frequently incomplete or unidentified, limiting the suitability of the HMM prediction method.

In [12], POMDP was proposed to anticipate channel state that may improve spectral efficiency and determine the optimal channel access. It was demonstrated that POMDP can forecast the spectrum under suboptimal conditions. The researchers created another variable order prediction model, the variable length Markov model, to boost the effect of Markov model prediction even more. It is feasible to enhance the classic Markov prediction models by incorporating the benefits of machine learning, especially with the continual advancement of machine learning. Recent improvements in deep learning models have been established as leading models compared to standard statistical models for spectrum prediction, capturing the intricate relationships among the observed spectrum data. Long short-term memory (LSTM) model [11], convolutional neural network (CNN) [10], convolutional LSTM model [11], deep transfer learning algorithm [12], and a hierarchical learning system of others, have already been developed for spectrum prediction, with the goal of exploiting the spatial, temporal, or spectral dependence of the spectrum.

B. Bayesian inference based spectrum occupancy prediction

In situations, where large pragmatic data is acquired, Bayesian Inference Function (BIF) is a method of inference that uses Bayes' principles to adjust the probability distribution of a proposition [7]. In CR networks, each system component can have a preceding probability distribution (known as preceding) computed by a CR user θ, defined by $P(\theta)$, prior data is evaluated. Through experimental judgements some data is observed using n time-slot spectrum sensing, $X = \{x_1, x_2, \ldots x_n\}$, this data is gathered from observations. The CR user then computes a parameter likelihood function θ, defined by $L(\theta|X)$ given that parameter, as the likelihood of the measured data. That is, $L(\theta\ X) = P(X|\theta)$. Bayesian inference may be utilised to calculate the subsequent probability distribution of the system parameter after obtaining the prior probability distribution and the probability function θ, based on the information $X = \{x_1, x_2, \ldots x_n\}$. The CR user determines the posterior likelihood function first in Bayesian inference-based prediction, $P(\theta|X)$, in accordance with Bayes' rule [8]:

C. Moving average based spectrum occupancy prediction

The Moving Average (MA) based prediction [13] is a popular method for predicting a trend in a series of numbers. A k-order moving average predictor predicts the next value of a history series of length N as the average of the past k values in the sequence. An improved version of the moving average-based prediction, Exponential Moving Average (EMA) based forecasting, can be implemented to increase the influence of the most recent findings on the prediction result, where exponentially decaying weighting factors are implemented to older findings. In order to improve on spectrum sensing performance, an EMA-based prediction is applied [14]. Every CR user records the historical energy levels of the channels as observations and uses an EMA-based predictor to forecast future energy levels. The CR user then avoids the sensing task on those channels where the projected energy level exceeds a current threshold (seen as occupied by the PUs).

III. SYSTEM MODEL

In this model, the state of a cognitive radio (CR) system is described by the parameter θ. The spectrum occupancy status of a primary user (PU) is represented by the variable P(θ). Let the CR sensed spectrum data take the format: X = $[x_1, x_2 \ldots\ldots x_n]$. Bayesian inference may be used to estimate the posterior probability distribution of the system parameter conditioned on the data after obtaining the prior probability distribution and the likelihood function; X = $[x_1, x_2 \ldots\ldots x_n]$. Then, the Bayesian Inference for the CR sensing decision is defined by

$$P(\theta/X) = P(X/\theta).\frac{P(\theta)}{P(X)} \tag{1}$$

The probability of busy or idle priori states for a busy or idle current state is represented by the likelihood function. From the PU's spectrum occupancy state, let A be the number of busy slots and B be the number of idle slots. Let the number of occurrences in which both the current and preceding states are busy is denoted by C, while D represents the number of times both the current and previous states are idle. If the priori condition was busy, then the likelihood that the current state would be idle is computed as

$$P\left(\frac{A}{C}\right) = P\left(\frac{C}{A}\right).\frac{P(A)}{PC}$$

where

$$P(C) = P\left(\frac{C}{A}\right)P(A) + P\left(\frac{C}{B}\right).P(B) \quad (2)$$

Alternatively, the likelihood of the next state being busy may be computed as follows

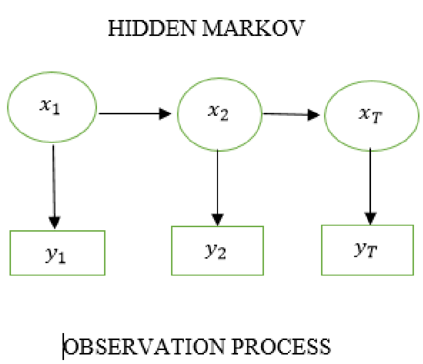
$$P\left(\frac{A}{D}\right) = P\left(\frac{D}{A}\right).\frac{P(A)}{PD}$$

where

$$P(D) = P\left(\frac{D}{A}\right)P(A) + P\left(\frac{D}{B}\right).P(B) \quad (3)$$

Applying Hidden Markov Model (HMM) into equation (1), the data sequence from channel occupancy history can be examined. The related mathematical formulation for HMM, in which empirical data is provided as an input to the algorithm to help the model forecast spectrum occupancy for TV whitespace better can be described as follows.

In the case of a concealed process, $\{X_t\}$, as well as a process that may be seen, $\{Y_t\}$. A double embedded process is what an HMM is $\{X_t, Y_t\}$, t = 0; 1; ……T as it is illustrated in the procedure below $\{X_t\}$. Within the state space, there are N number of concealed states. S of $\{S_{1,\ldots\ldots\ldots} S_N\}$ fulfils Markov requirements as,



$$P(x_t = S_i | x_{t-1} = S_{j;} x_{t-2} = S_N) = P(x_t = S_i | x_{t-1} = S_{j;} x_{t-1}) \quad (4)$$

where

x_t represents hidden condition at instant duration of t.

Therefore, the X_t transition probabilities characteristic is a_{ij}, with the original allocation of states π_i. These probabilities reflect the relationship between the current process and previous states, whereas the starting state distribution indicates the chances that the process will begin in a certain state, officially.

$$a_{ij} = P(x_t = S_i | x_{t-1} = S_{j;} x_{t-1}); \; 1 \le i, j \le N$$
$$\pi_i = P(x_i = S_i), 1 \le i, \le N$$

With the transition probabilities, $a_{ij} \ge 0$ and

$$\sum_j^N a_{ij} = 1 \quad (5)$$

The goal of spectrum occupancy prediction is to first ascertain the channel occupancy status up to a time instant T before predicting the occupancy status $T + d$. Here, T and d represent the duration of spectrum sensing record and prediction step, accordingly.

IV. Hybrid Prediction Algorithm

Spectrum sensing history (O_t) defines the PU activity and sensing output at the SU. PU activity, O_t is referred to the SU, as a result, it is thought to be the concealed condition of an underlying process. At the CR node, spectrum sensing is determined by measuring the received signal strength and comparing it to a specified detection threshold to ascertain the PU occupancy status, O_t , as active or inactive. The PU occupancy condition is computed and utilised for spectrum occupancy forecasting utilising prior spectrum sensing outputs without any information of the present sensing output. After finding the best HMM parameters, one moves on to the next step. $O_t = (logic = o \; or \; logic = 1)$ is the channel occupancy state at a future time instant following the training procedure, $T + d$. Based on the following selection method, the duration of a forecast time span d was calculated.

$$x_{t+d} = \begin{cases} 0, P(O_{t+d}, 0|\lambda) \ge (O_{t+d} 1|\lambda) \\ 1, P(O_{t+d}, 1|\lambda) < (O_{t+d} 1|\lambda) \end{cases} \quad (6)$$

Here, $P(O_{t+d}, 0|\lambda)$ and $P(O_{t+d}, 1|\lambda)$ are a way to represent the combined probability of a set of observations O_t of a vacant channel, followed by an occupied channel in a later time slot, $T + d$. The suggested predictive spectrum occupancy algorithm is shown here, with input from a real-time spectrum measurement experiment TV band O_{t+d}. The below algorithm shows a technique for storing the logic of the output in the hybrid prediction model's memory. O_t determines the amount of time it spends idle or unoccupied and then calculates the maximum time argument, $P(O_{t+d}, 0|\lambda)$, and return $\boldsymbol{x_{t+d}}$. depicts the core algorithm the following algorithm is function to determine the occupancy once the determination process is done it will return its calculation back to the main algorithm. Algorithms 1 and 2 respectively, illustrate the generic and hybrid channel occupancy prediction function.

Algorithm 1: Spectrum occupancy prediction function

Require: $\boldsymbol{\lambda_{ed}}$ **,** O_t
Ensure: x_{t+d}
Determine prediction span d
Calculate $P(O_{t+d}, 0|\lambda)$
Calculate argmax $P(O_{t+d}, 0|\lambda)$
Return $\boldsymbol{x_{t+d}}$

Algorithm 2: Hybrid spectrum occupancy prediction

Require: Spectrum sensing history O_t
Ensure: Channel state prediction x_{t+d}
Algorithm starts
IF SU senses for the first time, **then**
Initialise λ_{ed} with measured parameters
ELSE
Initialise λ_{ed} with random parameters
ENDIF
x_{t+d} = Prediction (λ_{ed}, o_T)
Energy Detection = $[x_{t+d} = 0]$

The proposed algorithm 2 utilises the input of the spectrum occupancy results obtained from the pragmatic spectrum measurement campaigns carried out in [3], where the frequency spectrum occupancy of TVWS is conducted using the energy level detection (ED) to distinguish between the occupied channel and idle. The data collected from the measurement campaigns is then applied to the hybrid prediction model as an input [3]. On the algorithm we then initialise the algorithm to learn the patterns we then take the advantage of applying a combination of HMM and BIF to produce predicted states.

V. Results And Discussion

In this section, we provide the findings of the hybrid spectrum occupancy prediction model that was developed. Next, we compare the results of the prediction models to determine which one best predicts the occupancy of the TV frequency spectrum. We accomplish this by comparing the actually measured spectrum occupancy to the model-based predicted results, evaluating the probability and correlation for each and then comparing all three to the actual spectrum occupancy.

The following results were simulated using MATLAB platform by applying first order HMM method then applied BIF prediction to predict the occupancy patterns. Fig. 1 depicts a comparison between the HMM occupancy probability and the anticipated probability using a Bayesian approach the actual serves as a reference for prediction preciseness. The anticipated outcome for Bayesian has a correlation factor of 0.65 (65%) and the probability of correctly anticipating the same logic state is 0.56, while for HMM correlation factor of 0.74 and the probability of correctly predicting the correct logic is very low 0.47.

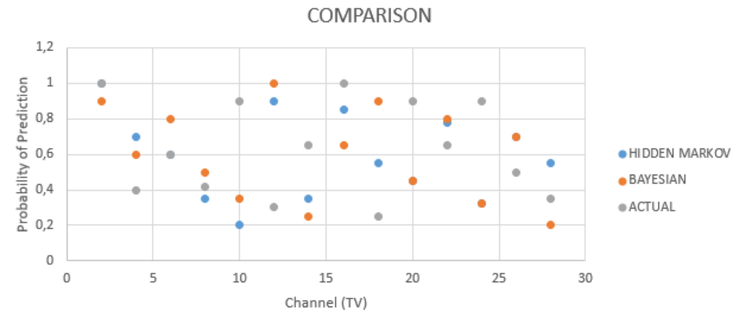


Fig. 1: Bayesian and Hidden Markov spectrum occupancy predictions

Fig. 2 depicts a comparison between the HMM occupancy probability and the anticipated probability using a Hybrid approach the actual serves as a reference for prediction preciseness. The anticipated outcome for HMM has a correlation factor of 0.45 (45%) and the probability of correctly anticipating the same logic state is 0.59 while for Hybrid correlation factor of 0.86 (86%) and the probability of correctly predicting the correct logic is 0.67.

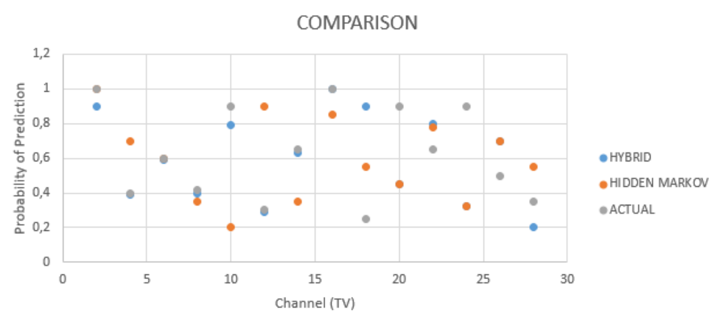


Fig. 2: Hybrid versus HMM spectrum occupancy predictions

Fig. 3 depicts a comparison between the Bayesian occupancy probability and the anticipated probability using a Hybrid approach the actual serves as a reference for prediction preciseness. The anticipated outcome for Bayesian has a correlation factor of 0.38 (38%) and the probability of correctly anticipating the same logic state is 0.59 while for Hybrid correlation factor of 0.97 (97%) and the probability of correctly predicting the correct logic is 0.88.

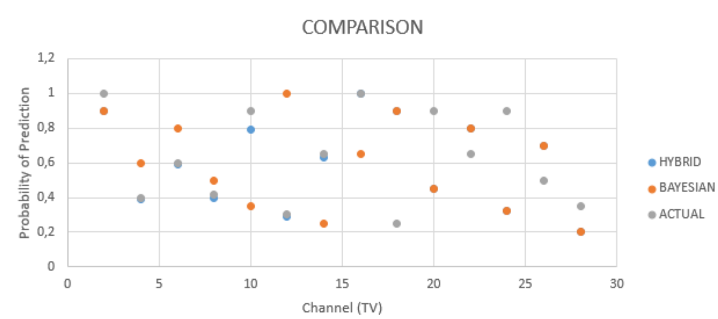


Fig. 3: Hybrid and Bayesian spectrum occupancy predictions

Fig. 4 illustrates a comparison of all the spectrum occupancy prediction models versus the real time spectrum occupancy learned from the measurements. Bayesian is depicted by an orange, while Hidden Markov is in grey plot and the developed Hybrid prediction method in yellow. The results on the graph show that the Hybrid which is a combination of both HMM and Bayesian has a better probability of predicting the correct channel occupancy state in many instances, except in instances, where channel numbers, where selections of priors influencing posterior distributions are skewed and also where spectrum occupancy observations are conditionally independent given the hidden states.

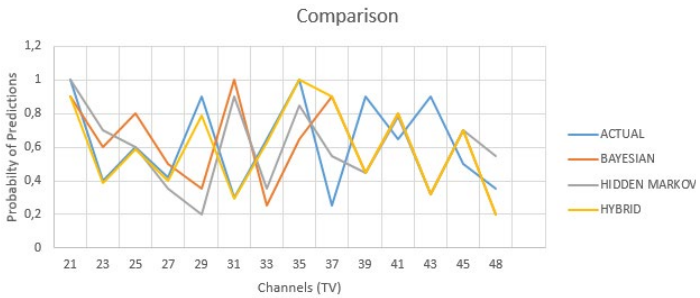


Fig. 4: Actual spectrum occupancy versus Prediction Models

Table 1 illustrates the comparison of Bayesian, Hidden Markov and Hybrid spectrum occupancy prediction (SOP) models when the same input data is utilised as the input to all models so as to compare the prediction accuracy that replicates the measured/ observed spectrum occupancy [3]. The comparison is done in terms of correlation of predicting the correct frequency channel and probability of predicting the occupancy status (i.e., occupied or unoccupied). It also shows the numbers determining which model best replicates real time spectrum occupancy.

Table 1**:** Comparison of SOP models

Models	**Correlation**	**Probability**
Bayesian	0.65 (65%),	0.56
Hidden Markov	0.45 (45)	0.71
Hybrid	0.89 (89%)	0.85

VI. Conclusion

We observed that the Hybrid prediction model is the best prediction model to closely replicate observed spectrum occupancy, since it has a high correlation to actual spectrum occupancy data, a high probability of prediction and requires less processing than other prediction models. We, therefore, infer that while the Bayesian spectrum occupancy prediction model has a greater correlation than Hidden Markov spectrum occupancy prediction model, Hidden Markov has a higher chance of prediction than Bayesian prediction data, assuming that the input to all models is not random but derived from real-time spectrum occupancy data.

Therefore, from the observed results, we can conclude that the developed Hybrid spectrum occupancy prediction model is indeed a better prediction than Hidden Markov and Bayesian spectrum occupancy prediction models when applied individually.

Acknowledgment

The authors would like to acknowledge the Department of Electrical Engineering/F'SATI at Tshwane University of Technology for their financial support.

References

[1] H. Sun, A. Nallthan, C. X Wang and Y. Chen. "Wideband Spectrum Sensing for Cognitive Radio Networks". The IEEE Wireless Communication, vol. 20, pp74-81, April. 2013.

[2] S.K. Dhurandher and B. Kumar, "An efficient radio frequency spectrum utilisation technique for cognitive radio networks," Internet of Things, vol. 20, Nov. 2022, 100586.

[3] J. Lebese, T.O. Olwal and L. Mfupe, "Realtime spectrum occupancy measurements based on the energy detection method in cognitive radios", 2021 IEEE Africon conference, pp. 1-5, 13-15 Sep. 2021, Arusha, Tanzania.

[4] J Mitola."Making Software Radios More Personal." IEEE Personal Communication, Vol. 6, pp. 13-18, Apr.1999.

[5] A. Ghosh, S. Kasera, and J. van der Merwe, "Spectrum usage analysis and prediction using long short-term memory networks," in Proc. 24th International Conf. on Distributed Computing and Networking, pp. 270-279, Jan. 2023

[6] M.R. Manesh and N. Kaabouch, "Security threats and countermeasures of MAC layer in cognitiv radio networs," Ad Hoc Networks, vol. 70, pp. 85-102, 2018.

[7] H. Kobayashi, B. L. Mark and W. Turin, "Probability, Random Process and Statistical Analysis," Cambridge University press. New York. June.2011.

[8] X. Xing, T. Jing, Y. Huo, H. Li and X. Cheng, "Channel quality prediction based on Bayesian inference in cognitive radio networks," Proceedings IEEE INFOCOM, pp. 1465-1473, 2013.

[9] G. Lin, Y. Cheng and H. Jiang, "Performance analysis of three - state HMM in shortwave channel estimation," Communication Technology, vol. 49, p. 3, 2016.

[10] Kai Zhang, Artist, *Hidden Markov Model for spectrum prediction and perception method.* [Art]. Master 's Thesis, Nanjing University of Posts and Telecommunications, 2015.

[11] M.G. Sumithra and M. Suriya, "Improved spectrum prediction model for CR networks using hybrid deep learning technique," International Journal of Intelligent Networks, vol. 5, pp. 286-292, 2024.

[12] D. C. Karia, B. K. Lande and R. D. Daruwala, "Performance analysis of HMM–and ANN–based spectrum vacancy predictor behaviour for cognitive radios," *International Journal of Ad Hoc and Ubiquitous Computing,* vol. 11, no. 4, pp. 206-213, 2012.

[13] P. Huang, C.-J. Liu, X. Yang, L. Xiao and J. Chen, "Wireless spectrum occupancy prediction based on partial periodic pattern mining," *IEEE Transactions on Parallel and Distributed Systems,* vol. 25, no. 7, pp. 1925-1934, 2013.

[14] J. Wu and Y. Li, "A survey of spectrum prediction methods in cognitive radio networks," *AIP Conference Proceedings,* vol. 1834, no. 1, pp. 2-18, 2017.