Normalización de la característica de radio para el aprendizado en línea por identificación en radio cognitiva

Kenta Asakura *, Haruhisa Ichikawa *, and Yuusuke Kawakita **

Resumen: El identificador de la tecnología de acceso radio (TAR) del usuario principal por usuarios secundarios es importante para evitar interferencias al compartir la espectro con técnicas de tarificación usando radio cognitiva (RC). TARs han evolucionado con la diversificación de la espectro con la introducción de servicios que utilizan la comunicación inalámbrica. Por lo tanto, es deseable tener un identificador de TAR que pueda adaptarse a la diversificación de TARs, y se aplica a RC. El propósito de este estudio es utilizar el aprendizado en línea para identificar múltiples TARs en el mismo espectro. Para mejorar la precisión de identificación de TARs con características similares, evalúan una normalización de la característica de radio propuesta en nuestro trabajo anterior para la identificación de TARs. El uso del método propuesto es efectivo para varios métodos de aprendizaje supervisado. Además, los resultados de la identificación de la precisión muestran que el método propuesto mejora la precisión de identificación comparado con los métodos convencionales.

Claves: radio cognitiva, tecnología de acceso radio, compartición de espectro, aprendizaje en línea, aprendizaje supervisado.

1. Introducción

El radio cognitivo (RC) [1] es útil para superar la escasez de los servicios de radio. Analiza el entorno radio y opera optimizando la transmisión y la recepción, garantizando la eficiencia y la utilización efectiva de la espectro con una solución tales como espectro de acceso dinámico (DSA) [2]. Por ejemplo, el smart meter, que es un tipo de sensor, sufre de problemas relacionados con el tráfico. Los problemas se pueden resolver usando RC para garantizar que la espectro esté efectivamente utilizada. En RC, es importante evitar interferencias por el volumen de datos que se pueden transmitir. En la red CR, mientras se protege a los usuarios secundarios, es importante evitar la interferencia de los usuarios secundarios. En la red CR, el identificador de TAR es importante para evitar la interferencia en las técnicas de compartir espectro. RC puede utilizarse para mejorar la eficiencia de uso de la espectro. En la red CR, el identificador de TAR es importante para evitar la interferencia en las técnicas de compartir espectro. RC puede utilizarse para mejorar la eficiencia de uso de la espectro.

RC ha diversificado con el surgimiento de servicios que utilizan la comunicación inalámbrica. En general, las frecuencias 802.11b, 802.11g, y 802.11n son utilizadas para 2.4 GHz, lo que también se utiliza en el diseño de nuevos TARs, como Wi-SUN, ZigBee, RFID, etc. Se pueden utilizar estos banda para uso en CR. El estudio de la diversificación de TARs para uso en CR es posible. Con el aumento del tráfico de datos, se espera que la diversificación de TARs se haga en el futuro. Los resultados muestran que el método propuesto mejora la precisión de identificación comparado con los métodos convencionales.
tal sensors, and other radio-connected devices. RFMLS has four technical components for integration into future RF machine learning systems, namely, (i) feature learning, (ii) attention and saliency, (iii) autonomous RF sensor configuration, and (iv) waveform synthesis. Feature learning is a set of techniques to automatically learn the RFMLS learn the characteristics of the data sets of RF signals, used to identify and characterize signals in various civilian and military settings. The aim is an overall RAT identification accuracy of at least 90%. In addition, DARPA held the Spectrum Collaboration Challenge (SC2) [9] to develop a new wireless paradigm in which radio networks collaborate autonomously while avoiding interference and jointly exploit opportunities to achieve the most efficient use of the available spectrum. The SC2 teams used recent advances in artificial intelligence (AI) and machine learning technology and the expanding capacities of software-defined radio to develop these breakthrough capabilities. The existence of such competitions indicates that it is general to use machine learning in the field of wireless communication.

In the present paper, we evaluate our previously proposed method [10] for normalizing signal features for online learning to improve RAT identification accuracy. The proposed scheme involves implementing a RAT identification system that uses online learning methods to adapt easily to RAT diversification. Online learning methods learn faster than batch learning methods do and cope easily with new data by learning sequentially. However, online learning is less accurate than batch learning. We studied RAT identification in the sub-gigahertz bands and used online learning to improve the accuracy of identifying RATs with similar feature [10]. Additionally, we resolved the stability-plasticity dilemma [11] caused by using online learning and the difficulty in identifying RATs that exist in the same frequency bands [10].

The remainder of the present paper is organized as follows. In Section 1 we describe various studies associated with RAT classification, online learning, and normalization. We present our proposed method in detail in Section 2 and evaluate it experimentally in Section 4. Finally, we present our conclusions in Section 5.

2. Related Work

Before discussing our proposed methods for RAT classification in detail, we review related literature. In this section, we consider RAT classification, online learning algorithms, and normalization methods.

2.1 RAT Classification

Reference [12] proposes using spectrogram analysis to quantify radio features for RAT identification. The proposed radio features consist of the center frequency, bandwidth, and transmission interval, which are extracted from the signal’s spectrogram. The transmission interval is the time span between the reception of a signal and the reception of a consecutive one. By considering the feature of transmission time and the frequency distribution, the method can identify RATs even when multiple signals exist in the same band. However, these features do not discuss the application of machine learning methods. Therefore, it is necessary to consider a method for adapting the features to allow machine learning to be used.

Reference [13] focuses on the blind identification of standards for green communication and proposes a RAT detector that is configured using a bandwidth shape detector and a pilot-based detector. In this study [13], a radial-basis-function neural network was used in the bandwidth shape detector. Moreover, this study uses the power spectral density (PSD) and pilot signals as the radio features but does not discuss using online learning methods in the detector blocks. Thereby, this method is difficult to adapt to new RATs that might appear in the future.

Reference [5] proposes a method for RAT recognition in CR networks that achieves good performance at very low signal-to-noise levels. That study uses self-organizing maps and a support vector machine (SVM) to identify the RAT. Moreover, it uses PSDs as radio features. However, similar to [5],[13] does not discuss using online learning methods in the RAT recognition system. Therefore, it has a similar problem to that in [13].

The aforementioned studies show that using online learning to identify RATs that exist in the same frequency band is yet to be considered. To cope easily with the diversification of RATs, it is necessary to implement such a RAT classifier.

2.2 Online Learning Algorithm

Online learning is a form of supervised machine learning that learns the data sequentially. Unlike batch learning, online learning can learn to avoid relearning previously learned data when dealing with new data. An additional advantage of online learning is that it converges faster compared with batch learning and the computational cost is lower. However, online learning is less accurate than batch learning and suffers from being vulnerable to learning noise. In addition, because online learning basically performs binary classification, it is necessary to expand a classifier based on online learning when it is necessary to classify multiclass data. Here, we explain some well-known algorithms developed in recent years.

2.2.1 Passive aggressive learning

Passive aggressive learning (PA) [14] is a type of online learning that uses hinge loss as the loss function. It updates the weight vector when it makes an incorrect prediction. Moreover, it updates the weight vector even in certain cases in which it has performed correct classification, but the prediction is of low confidence, indicating that the amount of learning obtained is highly inadequate. However, PA suffers from poor classifier accuracy with regard to learning noise.

2.2.2 Confidence-weighted learning

Conventional online learning considers a single weight vector and updates it sequentially. However, confidence-weighted learning (CW) [15] follows the principle that each weight vector is generated with a normal distribution, changing the weight vector width that is updated for each feature. CW reduces the weight vector when learning high-confidence data and increases the weight vector width when learning low-confidence data. CW has a high learning efficiency, better than that of PA, but it requires more computations.

2.2.3 Adaptive regularization of weights

Adaptive regularization of weights (AROW) [16] overcomes the defect of CW being susceptible to learning noise. CW learns to identify the learned data correctly by updating the weight vector, but the actual data may contain noise. Instead, AROW optimizes the weight vector by considering three conditions while learning. First, AROW learns the data so that it can correctly identify which data has been learned already. Second,
it searches for distributions that are close to the previous distribution. Third, it updates the confidence parameter for each feature whenever it learns a new feature. AROW has the same learning efficiency as that of CW when it learns data without noise. However, it shows a higher learning efficiency when it learns data that contain noise.

2.2.4 Soft confidence-weighted learning

Soft confidence-weighted learning (SCW) [17] incorporates the features of CW and AROW but improves on them by adding adaptive margins and the ability to handle nonseparable data. Moreover, SCW avoids learning noise by allowing a margin for failure prediction. Table 1 lists the properties of each of the aforementioned online learning algorithms, wherein SCW is the only one to exhibit all possible properties. Furthermore, SCW has the highest identification accuracy [17] and is an excellent algorithm regarding most aspects of online learning. In addition, SCW contains two algorithms, namely, SCW-I and SCW-II, with different updating rules. The difference between SCW-I and SCW-II exists in update rules. SCW-I is better than SCW-II regarding accuracy and learning speed.

2.2.5 Deep learning

Deep learning is a method of machine learning based on a multi-layered neural network. A standard neural network comprises several simple connected processors that can be referred to as neurons, and each neuron produces a sequence of real-valued activations. Further, the input neurons are activated by sensors that perceive the environment, and other neurons are activated via weighted connections from the previously activated neurons. An extensively used method of deep learning is to divide the training data into several small batches and update the weights for each. There are also deep learning algorithms that can sequentially learn from the input data such as online learning [18]. When a sufficient amount of data is available, deep learning can be used to automatically learn by extracting the feature quantities from the data. However, this process is associated with high computational complexity and requires a lot of memory when compared to that required by general machine learning algorithms.

In our study, the feature quantities that are required for identification are obvious; thus, there is no need to collect a large amount of data and search for features. Therefore, deep learning was considered to be excessive for this application. Furthermore, for RAT identification, it is better to have fast speeds for identification; thus, deep learning is not suitable because it has high computational complexity. Hence, the verification of deep learning was not conducted in this study. However, it is possible to use deep learning for RAT identification if these drawbacks can be overcome.

2.3 Normalization Methods

Reference [19] proposes a novel normalization method for dramatically accelerating the training of deep networks. Normalization is performed for each training minibatch enabling stochastic optimization methods commonly used in deep network training. A higher learning rate can be achieved in deep learning using the method proposed in [19], meaning that the network training converges faster. However, the method can only be applied to the neural network method. Reference [20] proposes a method that downconverts a high-dimensional feature vector to a constant two-dimensional feature vector for machine learning techniques while maintaining the same spectrum sensing performance. That method reduces the training and classification time because of its lower feature dimension and can be applied to RAT identification systems, but it cannot improve the classification accuracy. Normalization methods are commonly applied to machine learning, but applications to RAT identification is yet to be considered.

3. Radio Feature Normalization Method

In this section, we present the method proposed in [10] for normalizing radio features for online learning to identify RATs that exist in the same frequency band. Through sorting, dividing, and digit aligning as shown in Fig. 1 the method normalizes those radio features that can identify RATs in the same frequency band [12]. Normalizing the feature quantities in these processes improve the machine learning efficiency of wireless information. Moreover, the identification accuracy can be improved by suppressing the occurrence of the stability-plasticity dilemma. We assume that the radio features in Fig. 1, are used instead of the original radio features when learning and predicting with machine learning. Figure 2 shows an example of using the radio feature normalization method. To improve the identification accuracy, the proposed method is applied to the learned information to generate the RAT classifier. In addition, the proposed method is applied to the signal information.
Algorithm 1: Sorting Radio Features

Require: features, label
Ensure: features, label

// length(data) is the function which calculates the length
// of the array data.
for i ← 0 to length(data)/2 do
    // Rand(0,length(features)) is the function calculates a random
    // integer value from 0 to length(features).
    pivot1 ← Rand(0, length(features));
    pivot2 ← Rand(0, length(features));
    // Swap(x, y) means it is a process of swapping the values of x and y.
    Swap(features[pivot1], features[pivot2]);
    Swap(label[pivot1], label[pivot2]);
end for

Algorithm 2: Dividing Radio Features

Require: frequency, bandwidth, interval
Ensure: ret

for i ← 0 to length(features) do
    // log10(feature[i]) is the function which return the
    // base 10 logarithm of feature[i].
    digit ← log10(features[i]);
    tmp[0] ← feature[i] * 10^digit;
    tmp[1] ← feature[i] − tmp[0];
    tmp[2] ← digit;
    tmp[3] ← bandwidth[i]
    tmp[4] ← log10(bandwidth[i]);
    tmp[5] ← interval[i];
    ret[i] ← tmp;
end for

Algorithm 3: Digit Aligning of Radio Features

Require: features
Ensure: ret

for i ← 0 to length(features) do
    for j ← 0 to length(features[i]) do
        digit ← log10(features[i][j]);
        if digit > 2 then
            ret[i][j] ← features[i][j] / 10^digit;
        end if
    end for
end for

3.1 Sorting Radio Features

The radio features are sorted using the R-Random method
in the permutation data structure method (PDSM) [21]. The
process of the R-Random method randomly shuffles the list
order of the label and radio feature data to control the occurrence
of the stability-plasticity dilemma. The pseudocode of the
algorithm is given in Algorithm 1, and Fig. 3 shows a concrete example
of sorting radio features.

Algorithm 1. Sorting Radio Features

Require: features, label
Ensure: features, label

3.2 Dividing Radio Features

The radio features are divided by dividing the three-vector
radio features into six vectors. The process divides the center
frequency feature data into the first two digits and the remaining
digits of its number, and the number of digits center frequency
feature is added to the radio features. In addition, this process
adds the number of digits in the bandwidth feature to the radio
features, but it does not perform any processing for the trans-
mittance interval. Finally, six radio feature vectors are obtained,
as shown in Fig. 4. The pseudocode of the algorithm appears in
Algorithm 2. The number of digits is added to distinguish
two numbers that are shifted by one digit, (e.g., 240 MHz and
2.4 GHz). This process aims to expand the margin between the
data and draw a comfortable dividing line between them.

3.3 Digit Aligning of Radio Features

The digit-aligning process aligns the radio feature data to
three- and four-digit numbers containing one decimal place, as
shown in Fig. 5. The pseudocode of the algorithm is given in
Algorithm 3. However, the information regarding the number
of digits remains unchanged. The process aims to improve the
identification accuracy and hasten convergence by reducing the
distance between features. By reducing this distance, the ma-
chine learning algorithm can write a gentle slope for a boundary
line of identification. Therefore, it is possible to draw a boundary
line with sufficient margin without learning a large amount
of information.

4. Evaluation

In this section, the radio feature normalization method is
evaluated by comparing cases in which the proposed method
is and is not applied to the radio features. The comparison
is performed by calculating the curve of the receiver operat-
ing characteristics (ROC) for RAT classifiers and determining
the accuracy of RAT identification with multiple learning data.
The RATs of the radio features used as the identification targets
are ARIB STD-T107 and STD-T108, these RATs are assigned to the sub-gigahertz bands and existing in the same frequency band. The SCW-I algorithm was elected for online learning to generate the RAT classifier because it has a higher identification accuracy when compared with the other online learning algorithms. The RAT classifier obtained by the SCW-I algorithm using the proposed method was evaluated, and the results were compared with those of a RAT classifier obtained using an SVM and a neural network algorithm, which were implemented in scikit-learn [22].

4.1 Experimental Setup

The radio features allow three features to be extracted from the spectrogram, as shown in Fig. 6. Data whose signal strength exceeds a threshold value are regarded as a signal. The center frequency (labeled as \(x_1\)) is extracted from the frequency data of the strongest signal strength in the detected signal. The bandwidth (labeled as \(x_2\)) is extracted from the frequency range of the data whose signal strength exceeds the set threshold. The transmission interval (labeled as \(x_3\)) is extracted from the interval between consecutive transmissions regarded as signals. A signal that is continuously received without intervals of 0.1 ms is considered to be a single signal.

The radio features used in the evaluation were those of actual data extracted from the EnOcean battery-powered vibration sensor ETM502J and the RFID reader/writer Impinj Speedway R420. The STD-T108 radio features were extracted using ETM502J and those of STD-T107 were extracted from the Speedway R420 using LabVIEW and USRP-N210. LabVIEW is a software radio toolkit that works on a PC, and USRP-N210 is a type of signal input/output module usually used for developing software for radio. The extracted center frequency features do not exist in the same frequency band. Therefore, their center frequency values were adjusted to 922 MHz and 922.2 MHz to reproduce a situation in which they exist in the same frequency band. In addition, for evaluation, the extracted radio features were applied with the proposed methods. Finally, there were 300 (row) \(\times\) 20 (set) features for use as training data, 30 (row) \(\times\) 20 (set) features for use as test data, and 60 features for use as validation data per one RAT.

4.1.1 Two-class classification

In this section, we evaluate the RAT classifier made with data of two classes, namely, STD-T108 922 MHz and STD-T107 922.2 MHz. To compare with popular machine learning algorithms, we performed a similar experiment with an SVM implemented as batch learning. The RAT classifier ROC curve results are shown in Fig. 7. We generated the RAT classifier learning training data using the SCW algorithm, whose parameters were optimized using the validation data. We calculated the RAT classifier ROC curve by calculating the true positive rate and false positive rate from prediction using the test data. In addition, we calculated the ROC curve similarly for the radio features without using the proposed method.

Meanwhile, we calculated the accuracy of two-class RAT identification in terms of its training data, as shown in Fig. 8. Each point in Fig. 8 was obtained as the average identification accuracy calculated by predicting 200 times using 10 sets of training data and 20 sets of test data. These results were evaluated for each part of the proposed method (e.g., sorting only and sorting with division). Moreover, the data marked as original training data in Fig. 8 are the training data acquired using no learning method. The accuracy of two-class RAT identification was calculated for SCW only and not for SVM.

4.1.2 Three-class classification

In this section, we evaluate the RAT classifier made with data of three classes, namely, STD-T108 922 MHz, STD-T108 922.2 MHz, and STD-T107 922 MHz. We generated the RAT classifier learning training data using the SCW algorithm in the same way as calculating the two-class RAT accuracy. To investigate the effect of increasing the number of classification classes on the accuracy of RAT identification, we calculated the accuracy of three-class RAT identification in terms of the training data, as shown in Fig. 9. Each point in Fig. 9 was obtained in the same way as for two-class classification.

4.1.3 Relearning classification

In this section, we evaluate the RAT classifier made by relearning data of one class, namely, STD-T108 922.2 MHz, after learning data of two classes, namely, STD-T108 922 MHz and STD-T107 922 MHz. We generated the RAT classifier learning training data using the SCW algorithm in the same way as we calculated the two-class RAT accuracy. To evaluate the effect of relearning on identification accuracy, we calculated the accuracy of the relearned RAT classifier and compared it with the results marked Proposal in Fig. 9 and the results of [23] for the accuracy of three-class RAT that exist in separate frequency bands. In [23], the RAT classifier was made by a neural network algorithm, and thus we generate the same one to compare
Fig. 8 Two-class classification accuracy of RAT identification for numbers of training data.

Fig. 9 Three-class classification accuracy of RAT identification for numbers of training data.

with Proposal. The parameters of the neural network used in the present experiment and [23] are listed in Table 2, and the comparison results are shown in Fig. 10.

4.2 Results

4.2.1 Two-class classification

The area under the curve (AUC) increased from 0.61 to 0.98 using the proposed method in the case of SCW and increased from 0.50 to 0.98 in the case of SVM, as shown in Fig. 7. This shows that the proposed method improves the identification performance as can be deduced by comparing the AUC of the solid line with that of the dotted line. Additionally, the proposed method is shown to be effective for both SCW and SVM. Comparing the ROC results for SCW and SVM for RAT identification within the same frequency band shows that the online learning algorithms have the same identification performance as that of batch learning. Incidentally, the AUC of SVM classification is 0.50, and it looks like a reference curve. However, it is the correct AUC of SVM classification curve implemented in scikit-learn. It is difficult to classify using SVM when the distance between the two classes is close as shown in Fig. 11.

In Fig. 8 the accuracy of two-class RAT identification using the proposed method is around 90% higher than that of any other process in nearly all aspects, whereas the accuracy using the original training data converges only around 50%. The proposed method is shown to improve the identification accuracy for multiple RATs in the same frequency band by around 40% compared with the original learning data. Moreover, based on comparing learning with and without sorting, online learning clearly requires sorting to learn correctly. In addition, the sort-
ing process alone requires at least 190 pieces of data per RAT for RAT identification to converge accurately. By contrast, digit alignment requires at least 10 pieces of data per RAT for it to converge. By comparing the numbers of training data for each process and the identification accuracies, digit alignment was shown to be important in the proposed method for achieving convergence accuracy early. Dividing appears to be unnecessary, but it is required to compare the proposed method with the other methods regarding RAT identification. There is a minimum difference of around 5% between the proposed method and processing without dividing. The effectiveness of dividing was shown to be small, but it was effective in the RAT identification system.

4.2.2 Three-class classification

In Fig. 9, the accuracy of three-class RAT identification using the proposed method is around 82% higher than that of any other process in nearly all aspects. The accuracy using the original training data converges around 33%, and the accuracy of sorting plus digit alignment and dividing only converge less than 33%. Moreover, the accuracy of sorting only converge around 48%, and sorting only improves the accuracy by around 15% compared with the original training data. By contrast, the accuracy of sorting plus dividing converges around 70%, and sorting plus dividing improves the accuracy by around 37% compared with the original training data. Combining sorting and dividing is shown to be important for improving the identification accuracy for multiple RATs in the same frequency band.

4.2.3 Relearning classification

In Fig. 10, three-class learning is more accurate than two-class learning and one-class relearning for numbers of training data per RAT between 10 and 150. However, three-class learning and relearning results are almost equally accurate when there are more than 160 pieces of training data. The results show that learning a large amount of learning data does not affect the identification accuracy. The accuracy of three-class RAT identification converges around 82% even if it relearns the training data, whereas in Fig. 10 the accuracy of three-class RAT identification in the same frequency band using the neural network algorithm of [23] converges around 33%. However, the accuracy of three-class RAT identification in separate frequency bands in [23] exceeds 90%. This is because the radio features of identification target RATs in [23] are not so much similar to each other as the features adopted in our research. The proposed method is shown to improve the identification accuracy for multiple RATs in the same frequency band.

5. Conclusion

In this paper, we discussed the implementation of a RAT identification system that can adapt easily to the diversification of RATs. Moreover, we improved the accuracy of identifying RATs with similar feature quantities using the radio feature normalization method for online learning. In addition, we evaluated the proposed method by applying it to a situation in which multiple RATs existed in the same sub-gigahertz frequency band. In the evaluation, the proposed method was evaluated by the ROC curve and the average accuracies of two- and three-class RAT identification for generating the RAT classifier using the SCW online learning algorithm. Moreover, we evaluated the proposed method in comparison with relearning and neural network used in [23].

The results for the ROC curve revealed that the proposed method is effective for certain supervised learning. In addition, the results for the accuracy of two class RAT identification showed that the proposed method improves the identification accuracy for multiple RATs existing in the same frequency band by nearly 40% compared with the bottom one, and the results for the accuracy of three class RAT identification showed that combining sorting and dividing is important for improving the accuracy of identifying multiple RATs in the same frequency band. Moreover, the results for the relearning RAT identification showed that the proposed method improves the accuracy of identifying multiple RATs in the same frequency band by nearly 49% compared with the methods in [23]. Based on these results, the proposed method does not satisfy the goal of RFMLS, which is an overall RAT identification accuracy of over 90%. However, the proposed method improves the identification performance and gets the property of easily adapting for new RAT compared to that of conventional methods.

As future research, we intend to classify the three-class dataset, which has the same center frequency. Identifying three or more classes in the same frequency band reduces misidentification.

Acknowledgments

This work was supported by JSPS KAKENHI grant no. 16K16042.

References

[1] J. Mitola and G.Q. Maguire: Cognitive radio: Making software radios more personal, IEEE Personal Communications, Vol. 6, No. 4, pp. 13–18, 1999.
[2] S. Haykin: Cognitive radio: Brain-empowered wireless communications, IEEE Journal on Selected Areas in Communications, Vol. 23, No. 2, pp. 201–220, 2005.
[3] S.S.S.R. Depuru, L. Wang, V. Devabhaktuni, and N. Gudi: Smart meters for power grid: Challenges, issues, advantages and status, 2011 IEEE/PES Power Systems Conference and Exposition, pp. 1–7, 2011.
[4] Y. Li, L. Zhou, H. Zhu, and L. Sun: Privacy-preserving location proof for securing large-scale database-driven cognitive radio networks, IEEE Internet of Things Journal, Vol. 3, No. 4, pp. 563–571, 2016.
[5] S. Baban, O. Holland, and H. Aghvami: Wireless standard classification in cognitive radio networks using self-organizing maps, ISWCS 2013: The Tenth International Symposium on Wireless Communication Systems, pp. 1–5, 2013.
[6] H. Fuji, S. Miura, and H. Kayama: Novel cognitive radio technique for using white space in public cellular networks, 2012 18th Asia-Pacific Conference on Communications (APCC), pp. 266–271, 2012.
[7] T. Dudda and T. Ihrlich: Capacity of cellular networks deployed in TV white space, 2012 IEEE International Symposium on Dynamic Spectrum Access Networks, pp. 254–265, 2012.
[8] RF machine learning systems (RFMLS), https://www.darpa.mil/attachments/RFMLSIndustryDayPublicReleaseApproved.pdf [Accessed 6 September 2018]
[9] DARPA: spectrum collaboration challenge, https://spectrumcollaborationchallenge.com/ [Accessed 6 September 2018]
[10] K. Asakura, H. Ichikawa, and Y. Kawakita: Normalization method for online learning on radio access technology identification in cognitive radio, 2017 23rd Asia-Pacific Conference on Communications (APCC), pp. 682–687, 2017.
[11] S. Grossberg: Nonlinear neural networks: Principles, mechanisms, and architectures, *Neural Networks*, Vol. 1, No. 1, pp. 17–61, 1988.

[12] M. Sato, Y. Mizutani, Y. Kawakita, and H. Ichikawa: Feature quantity for universal receivers to identify protocols of radio services and devices, *2013 IEEE 11th International Conference on Dependable, Autonomic and Secure Computing*, pp. 433–437, 2013.

[13] B. Aziz, A. Nafkha, J. Palicot, and H. Zhang: Blind wireless standard identification for green radio communications, *2013 International Conference on Advanced Technologies for Communications (ATC 2013)*, pp. 295–300, 2013.

[14] K. Crammer, O. Dekel, J. Keshet, S. Shalev-Shwartz, and Y. Singer: Online passive-aggressive algorithms, *Journal of Machine Learning Research*, Vol. 7, pp. 551–585, 2006.

[15] M. Dredze, K. Crammer, and F. Pereira: Confidence-weighted linear classification, *Proceedings of the 25th International Conference on Machine Learning*, pp. 264–271, 2008.

[16] K. Crammer, A. Kulesza, and M. Dredze: Adaptive regularization of weight vectors, *Advances in Neural Information Processing Systems*, pp. 414–422, 2009.

[17] J. Wang, P. Zhao, and S.C. Hoi: Exact soft confidence-weighted learning, *The 29th International Conference on Machine Learning*, pp. 121–128, 2012.

[18] D. Sahoo, Q. Pham, J. Lu, and S.C.H. Hoi: Online deep learning: Learning deep neural networks on the fly, CoRR, abs/1711.03705, 2017. [Online]. Available: http://arxiv.org/abs/1711.03705

[19] S. Ioffe and C. Szegedy: Batch normalization: Accelerating deep network training by reducing internal covariate shift, *International Conference on Machine Learning*, pp. 448–456, 2015.

[20] Y. Lu, P. Zhu, D. Wang, and M. Fattouche: Machine learning techniques with probability vector for cooperative spectrum sensing in cognitive radio networks, *2016 IEEE Wireless Communications and Networking Conference*, pp. 1–6, 2016.

[21] I. Hayashi and H. Miyauchi: A proposal of permutation data structure method, *31st Fuzzy System Symposium*, 2015 (in Japanese).

[22] Github - scikit-learn/scikit-learn: scikit-learn: machine learning in python, https://github.com/scikit-learn/scikit-learn [Accessed 6 November 2018].

[23] S. Baban, D. Denkoviski, O. Holland, L. Gavrilovska, and H. Aghvami: Radio access technology classification for cognitive radio networks, *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 2718–2722, 2013.

---

**Kenta ASAKURA**

He received a B.E. degree from the University of Electro-Communications in 2017. He is a Master’s student at the University of Electro-Communications. His current research interests are cognitive radio and information technology.

**Haruhisa ICHIKAWA**

He received B.S., M.S. and Dr. Eng. degrees in electrical engineering from the University of Tokyo in 1974, 1976, and 1989, respectively. He joined NTT Laboratories in 1976, where he was engaged in fundamental research on communications software and distributed computing. He created and conducted many R&D projects for software, Internet, information sharing platform, and ubiquitous networks, including business incubation. He has been with the University of Electro-Communications, Tokyo, since 2007. His current research interests focus on IoT platforms particularly for ubiquitous sensing, and electric power supply.

**Yuusuke KAWAKITA (Member)**

He received a B.A., an M.A. degree and Ph.D. from Keio University in 2000, 2002 and 2008, respectively. He is an associate professor at Kanagawa Institute of Technology. His present research interests focus on the ubiquitous sensing and its platform architecture.