



Channel Clustering and QoS Level Identification Scheme for Multi-Channel Cognitive Radio Networks

Ali, A., Yaqoob, I., Ahmed, E., Imran, M., Kwak, K. S., Ahmad, A., Hussain, S. A., & Ali, Z. (2018). Channel Clustering and QoS Level Identification Scheme for Multi-Channel Cognitive Radio Networks. *IEEE Communications Magazine*, 56(4), 164-171. <https://doi.org/10.1109/MCOM.2018.1700752>

[Link to publication record in Ulster University Research Portal](#)

Published in:
IEEE Communications Magazine

Publication Status:
Published (in print/issue): 13/04/2018

DOI:
[10.1109/MCOM.2018.1700752](https://doi.org/10.1109/MCOM.2018.1700752)

Document Version
Author Accepted version

General rights

The copyright and moral rights to the output are retained by the output author(s), unless otherwise stated by the document licence.

Unless otherwise stated, users are permitted to download a copy of the output for personal study or non-commercial research and are permitted to freely distribute the URL of the output. They are not permitted to alter, reproduce, distribute or make any commercial use of the output without obtaining the permission of the author(s).

If the document is licenced under Creative Commons, the rights of users of the documents can be found at <https://creativecommons.org/share-your-work/licenses/>.

Take down policy

The Research Portal is Ulster University's institutional repository that provides access to Ulster's research outputs. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact pure-support@ulster.ac.uk

Channel Clustering and QoS Levels Identification Scheme for Multi-Channel Cognitive Radio Networks

Amjad Ali, Ibrar Yaqoob, Ejaz Ahmed, Muhammad Imran, Kyung Sup Kwak, Adnan Ahmad, Syed Asad Hussain, and Zulfiqar Ali

Abstract—The increasing popularity of wireless services and devices necessitates high bandwidth requirements; however, spectrum resources are not only limited but also heavily underutilized. Multiple license channels that support the same levels of quality of service (QoS) are desirable to resolve the problems posed by the scarcity and inefficient use of spectrum resources in multi-channel cognitive radio networks (MCRNs). One reason is that multimedia services and applications have distinct, stringent QoS requirements. However, due to a lack of coordination between primary and secondary users, identifying the QoS levels supported over available licensed channels has proven to be problematic and has yet to be attempted. This paper presents a novel, Bayesian, non-parametric channel clustering scheme, which identifies the QoS levels supported over available license channels. The proposed scheme employs the infinite Gaussian mixture model and collapsed Gibbs sampler to identify the QoS levels from the feature space of the bitrate, packet delivery ratio, and packet delay variation of licensed channels. Moreover, the real measurements of wireless data traces and comparisons with baseline clustering schemes are used to evaluate the performance of the proposed scheme.

Index Terms—Multi-channel cognitive radio networks, quality of service levels, multimedia transmissions, channel clustering.

I. INTRODUCTION

Recent advances in wireless access technology and the rapid proliferation of media-rich devices and applications have changed wireless network traffic on a fundamental level. In a recent study, Cisco predicted that wireless traffic would grow threefold from 2016 to 2021, with 82% of the increase being related to video¹. Supporting this increase in multimedia traffic on a severely limited spectrum will prove to be fairly challenging as it is critical in ensuring the future

A. Ali, A. Ahmad, and S.A. Hussain are with the Department of Computer Science, COMSATS Institute of Information Technology, Lahore, Pakistan. (Email: amjad.ali@ciitlahore.edu.pk; adnanahmad@ciitlahore.edu.pk; asad-hussain@ciitlahore.edu.pk).

I. Yaqoob and E. Ahmed are with Centre for Mobile Cloud Computing Research, Faculty of Computer Science and Information Technology, University of Malaya, Malaysia. (Email: ibraryaqoob@yahoo.com; ejazahmed@ieee.org)

M. Imran and Z. Ali is with the College of Computer and Information Sciences, King Saud University, Saudi Arabia. (Email: cimran@ksu.edu.sa; zuali@ksu.edu.sa)

A. Ali, K.S Kwak are with the UWB Wireless Communications Research Center, Inha University, Incheon 402-751. (Email: amjad.ali@ciitlahore.edu.pk; kskwak@inha.ac.kr)

competitiveness of wireless networks. However, the wireless spectrum is limited and heavily underutilized. According to reports by the Federal Communication Commission, between 15% and 85% of the wireless spectrum is currently going to waste. Hence, the limitations and inefficient use of the wireless spectrum cause the primary bottlenecks when accommodating the increasing needs of multimedia traffic. Multi-channel cognitive radio technology has emerged as a promising tool in addressing the inefficient utilization of the wireless spectrum and accommodates wireless multimedia applications across the wide range of licensed channels (LCs) [1].

Multi-channel cognitive radio networks (MCRNs) define the two types of spectrum users: (1) primary users (PUs) and (2) secondary users (SUs). PUs are privileged users who have the right to use LCs at any time without interruption. SUs are unlicensed users equipped with multiple interfaces, who are always looking to capitalize on any available LCs to enhance spectrum utilization. However, due to the lack of coordination between PUs and SUs, finding multiple LCs that support identical quality of service (QoS) levels has proven to be problematic and has yet to be attempted.

Channel conditions in a wireless environment are spatiotemporal; that is, an SU experiences different QoS gains over different wireless channels in terms of bitrate, packet loss rate, and jitter with respect to time and location. This disparity in QoS gain is highly significant across a wide range of spectrum bands, such as from the TV spectrum bands (e.g., 30 MHz to 300 MHz) to the ISM spectrum bands (e.g., 2.4 MHz to 2.5 GHz). The situation exacerbates when multiple SUs experience different QoS gains over the same channel in a phenomenon known as multi-user diversity. As multimedia applications require guaranteed QoS, the identification of multiple LCs supporting the same levels of QoS is crucial for efficient multimedia transmission. In MCRNs, each SU should estimate the available LCs and identify the QoS levels they support. These endeavors can be done by collecting the features of LCs and by clustering them to identify their supported QoS. Hence, a SU can optimize its transmission according to demands of wireless multimedia applications. Moreover, if the SU adopts a set of LCs without the prior knowledge of their supported QoS levels, it will repeatedly hand-off to other channels for successful transmissions. Consequently, the SU will also experience poor multimedia transmission quality and the inefficient utilization of LCs.

Given that information on the number of LCs and their

¹ Accessed on: 17th August 2017 <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html>

supported QoS levels is not immediately available, we introduce a novel non-parametric Bayesian channel clustering scheme, which provides basic techniques for the specification of prior beliefs, reveals dependencies among the data sets, and performs efficient inferencing. This scheme is utilized to figure out the structures present in the provided data set (i.e., number of clusters or classes) that best describe the observations. The proposed scheme is called non-parametric because the number of unique traffic patterns over PU LCs is not provided. In the proposed scheme, the SU observes and extracts the channel estimation features (i.e., bitrate, packet delivery ratio (PDR), and packet delay variation (PDV)) from each available LC. Multivariate Gaussian distribution was utilized with unidentified parameters to cluster the single channel feature space. Similarly, for multi-channel clustering, we use the infinite Gaussian mixture model (IGMM) and the collapsed Gibbs sampler (CGS). The CGS classifies the observation points into different clusters or classes by using the Markov Chain Monte Carlo (MCMC) technique. The CGS then generates posterior samples by sweeping through each observation point, sampling from its conditional distribution while keeping the other observation points fixed at their current values. Hence, to address the problem of clustering an infinite number of LC feature spaces, we develop an unsupervised manner in the non-parametric Bayesian technique, a process which involves the CGS generating posterior samples and identifying the total number of QoS levels or channel clusters and, consequently, the number of channels supporting the same level of QoS. The unique characteristics of our proposed channel scheme are as follows:

- Our scheme is based on heterogeneous or different LC feature spaces and does not require any exchange of control messages between a SU and a PU.
- Our scheme is based on the unsupervised clustering mechanism.
- The proposed scheme is highly flexible and can easily be extended by incorporating numerous LC features for increased efficiency and accuracy.
- The scheme does not require any prior knowledge to identify the distinct channel clusters or their supported QoS levels.
- The proposed clustering scheme relaxes parameter assumptions, which are similar to the real situation of distributed cognitive radio networks (CRNs).
- The real data traces of WiMAX, WiBro, and wireless channels² measured at different locations (i.e., subway, residential, and campus) under different traffic variations were utilized to evaluate the performance of the proposed scheme. Moreover, we evaluate the accuracy of the proposed scheme by comparing its performance with K-means, expectation maximization, and spectral clustering algorithms.

The main contributions of the study can be summarized as follows:

- We propose a novel, non-parametric, Bayesian-based channel clustering and a QoS level identification scheme

that considers channel features (i.e., bitrate, PDR, and PDV) and identifies the QoS support over each channel.

- The proposed scheme focuses on the heterogeneity of LCs, which challenges the algorithm design. However, the accuracy of the channel clustering and the identification of the QoS levels of different spectrum channels range from 98% to 99.5%.

The rest of the paper is organized as follows: Section II discusses the existing literature; Section III presents the proposed non-parametric Bayesian channel clustering scheme for the identification of QoS levels by using the IGMM and the CGS; Section IV exhibits the performance analysis; and Section V tackles our conclusions and recommendations.

II. RELATED WORK

The CRN is a candidate blueprint for the enhancement of current spectrum utilization. Several recent studies have corroborated that the utilization of cognitive technologies can enhance performance by meeting the delay sensitive requirement for multimedia applications. For instance, in wireless sensor networks, license-free spectrums have been observed to frequently suffer from uncontrolled interference. Accordingly, the concept of a cognitive radio sensor network (CRSN) was introduced in [2]. This study analyzed the performance of CRSN for real-time applications with the help of *Periodic Switching* and *Triggered Switching* channel access techniques. Another study [3] proposed a solution for multi-user downlink video streaming for infrastructural CRN. This study aimed to maximize the quality of experience for SUs by considering channel sensing and assignment policies jointly.

The study [4] proposed a distributed QoS-aware medium access control (MAC) protocol for CRN to support multimedia applications. In this study, a priority spectrum access scheme is introduced, which assigns differentiated random channel sensing durations to SUs based on their traffic priority type (i.e., an SU with the highest priority obtains substantial chances to access the LC). Moreover, an analytical model is formulated to measure the delay performance of the proposed scheme. In another study [5], authors considered geometric characteristics (i.e., path loss and random node positions) to analyze the performance of the MAC protocol for MCRNs. Similarly, a decentralized MAC protocol for MCRNs is also proposed in [6]. The proposed scheme is intended for SUs who want to communicate their control information along with data over LCs. The proposed scheme improves the network performance in terms of throughput and communication time by reducing the number of handshakes over common control channel. The study [7] proposed a preemptive priority queuing model to estimate the performance of the spectrum allocation strategy under multi-channel and imperfect sensing. Moreover, a three-dimensional discrete Markov chain model is developed for measuring the stochastic behavior of the SU and PU data packets. In multi-hop cognitive radio ad hoc networks (CRAHNs), the optimal utilization of the channel resources has become a key concern. A joint rate adaptation, channel assignment and routing (J-RCR) scheme is proposed in [8] to cope with this scenario. The proposed scheme helps maximize social welfare

² Accessed on: 15th August 2017 <https://crawdad.cs.dartmouth.edu/>

through performing optimization in multi-channel CRAHNS. Moreover, it dynamically adapts the transmission rates of SUs based on the network conditions and rate requirements.

In [1], authors claimed that channel assignment schemes for CRN generally do not consider the temporal usage of LCs. Therefore, for providing high QoS to the SU, while protecting PUs, a PU temporal activity-based channel assignment scheme is proposed for MCRNs. The IEEE 802.22 is the new standard for CRNs, emphasizing the significance of interference-free and reliable SU communication over LCs. However, the channel sensing mechanisms provided by this standard have radically failed to satisfy the rigorous sensing requirements. A few other shortcomings of the standard are a lack of flexibility, unbalanced performance, instability, and performance non-uniformity. Hence, a distributed channel-sensing scheme known as the multi-channel learning-based distributed sensing fusion mechanism is proposed for MCRNs in [9]. The proposed scheme is highly stable and self-trained and even compensates for false detection.

In our proposed scheme, we used Gibbs sampling, which is based on the MCMC sampling algorithm. The problem with the MCMC scheme is that it suffers from slow convergence. In the literature, [2]–[4] authors performed video transmissions based on a single channel assignment scheme. However, none of the above mentioned works consider the actual supported QoS levels on LCs and PU traffic patterns for channel selection and transmitting multimedia over MCRN. In the future, it should consider the multi-media traffic specific heuristics on channel selection algorithms that can lead to the further improvement of traffic perception quality for MCRNs. Unlike most existing work, we proposed a PU channel feature (i.e., traffic patterns)-based channel clustering and the identification of QoS, which helps in identifying multiple LCs supporting the same QoS level for improved multimedia transmissions and efficient LCs utilization over MCRNs.

III. CHANNEL CLUSTERING AND QoS LEVELS IDENTIFICATIONS SCHEME BASED ON BAYESIAN NON-PARAMETRIC INFERENCE

In this section, we discuss our proposed channel clustering and QoS level identification scheme. The proposed scheme is based on the Gaussian mixture model (GMM), the finite Gaussian mixture model (FGMM), the IGMM, and the CGS. In a mixture model, observation points are used to generate the mixture distribution, in which each observation point has its own mixing weight. In the mixture model, the first step is to decide which type of statistical distribution is used for clustering the observation points. In the current study, we selected the Gaussian distribution, which is most widely adopted in the literature [10], [11]. The Gaussian distribution is chosen because of its good theoretical properties (i.e., mixtures of Gaussians are dense in the space of densities on an Euclidean space), closed form of some distances (i.e., Hellinger and Kullback-Leibler), and low computational effort (i.e., conjugacy) [11]. In second step, each mixture distribution is substituted by Gaussian distribution to obtain the GMM. The GMM-based clustering technique is highly flexible as it

accommodates clusters of different sizes and the correlation of structures within them. The GMM-based clustering scheme is also called a soft clustering scheme (i.e., all observation points are assigned to different clusters based on certain probability values), and we can also estimate through mixing the weight of how much data each observation point represents. Thus, on the basis of the nature of our clustering problem and data set, the GMM is a more appropriate choice than other distributions.

In the proposed scheme, a SU can cluster the available LCs based on their unique feature space. Features are the unique characteristics of LC usage, such as experiencing PDR, PDV, and bitrate. Given that these features depend on the time and geographical location of a PU and a SU, each LC may follow a unique behavior in terms of PDR, PDV, and bitrate. In the proposed scheme, the SU observes the LCs to extract their features, revealing the following definitions of our selected features:

- *Bitrate*: Bitrate for different LCs could be different. For example, a WiFi channel can support 54-600 Mbps under optimal channel conditions, whereas WiMAX channel supports 70 Mbps, and long-term evolution (LTE) channels can support up to 3 Gbps [12]. Moreover, the bitrate of wireless channels varies on uplink and downlink channels. Hence, we adopt the Bitrate as one feature point to identify the QoS support over different LCs.
- *Packet Delivery Ratio*: The PDR is derived from dividing the number of packets received successfully over the number of total packets sent. A great value of PDR represents a good quality of wireless channels and, consequently, less packet loss.
- *Packet Delay Variation*: Also referred to as jitter, the PDV measures the difference of the delay among successive packets. The PDV may change from channel to channel due to the PUs running different applications. Moreover, by inspecting the real data traces of different LCs, we have discovered that PDV varies drastically during communication. Thus, the PDV is another important parameter that severely affects the performance of multimedia applications.

For the observation of the number of QoS levels identified from the features mentioned above and the determination of which observation point belongs to which distinct channel cluster or QoS level, the posterior distribution (PD) for the set of observation points was generated. In this study, we use the terms observation points and data points interchangeably. The PD of the data points signifies the entire QoS levels or channel clusters and further classifies which observation point belongs to which QoS level. Currently, two generic models exist to generate the PD of observations: (1) the generative model and (2) the discriminant model:

Generative Model: This model utilizes probabilities to generate a data set for a specific phenomenon based on the model and provided (i.e., hidden) parameters.

Discriminant Model: This model is a probabilistic model that is also called conditional model and only provides model for data set. The data set is generated on the basis of analyzing the observation points.

The generative model was used in the characterization of data points due to following five reasons: (1) This model is flexible enough to express dependencies in a complex learning environment; (2) this model randomly generates hidden parameters and a joint distribution of data points; (3) this model uses data directly; (4) this model generates samples of two or more variables from their joint distributions, which is impossible through a discriminant model; and (5) the discriminant model is inherently supervised in nature and could not be used for unsupervised learning. The examples of generative models include the hidden Markov model and the GMM, which was used in this study due to its flexibility to extend for a case, in which QoS levels are unidentified. Furthermore, the proposed scheme is extendable beyond the above mentioned channel features. The identification of the QoS levels supported by available LCs is reflected by our mentioned features, which individually separates available LCs into a high dimensional space.

A. Finite Gaussian Mixture Model

We had primarily assumed that the number of channel clusters was given and that each QoS level follows a Gaussian distribution. Therefore, at the initial step, the model became the FGMM, which correctly clusters the feature points, utilizing the exact estimation of multi-modal probability density in cases where hidden variables are inferred as class labels. Moreover, the FGMM considers that the observations are produced through a finite number of Gaussians with unknown parameters. In this study, the number of dimensions is fixed at three, representing the feature space. In the FGMM, every observation is assigned to a distinct cluster among total K clusters based on a unique value or weight that is assigned to every cluster in the mixture. Each observation point must be associated with a cluster, revealing which observation point belongs to which cluster out of the total K clusters. The FGMM is an effective model when the total numbers of channel clusters are initially available. Whereas, the total number of clusters is based on total available distinct LCs. In MCRNs, channel availability depends on time varying PU activities and the geographical location of the SUs. Therefore, we did not know the exact number of clusters in the mixture, and a highly flexible model that does not require such information in advance was needed. Moreover, by adopting the FGMM, our channel clustering problem becomes considerably complex. In the IGMM, the term weight or probability of belonging is dependent on the K ; as the value of K increases, the weight is affected.

B. Infinite Gaussian Mixture Model

The IGMM is a highly advanced version of the FGMM that does not require any prior information. In the IGMM, the total number of clusters can be infinite and the term weight or probability of belonging is dependent on the K ; as the value of K increases, the weight is affected. The IGMM is profoundly suitable for adoption in a situation where the number of clusters is not provided in advance. The IGMM can incorporate an infinite dimensional model and uses the

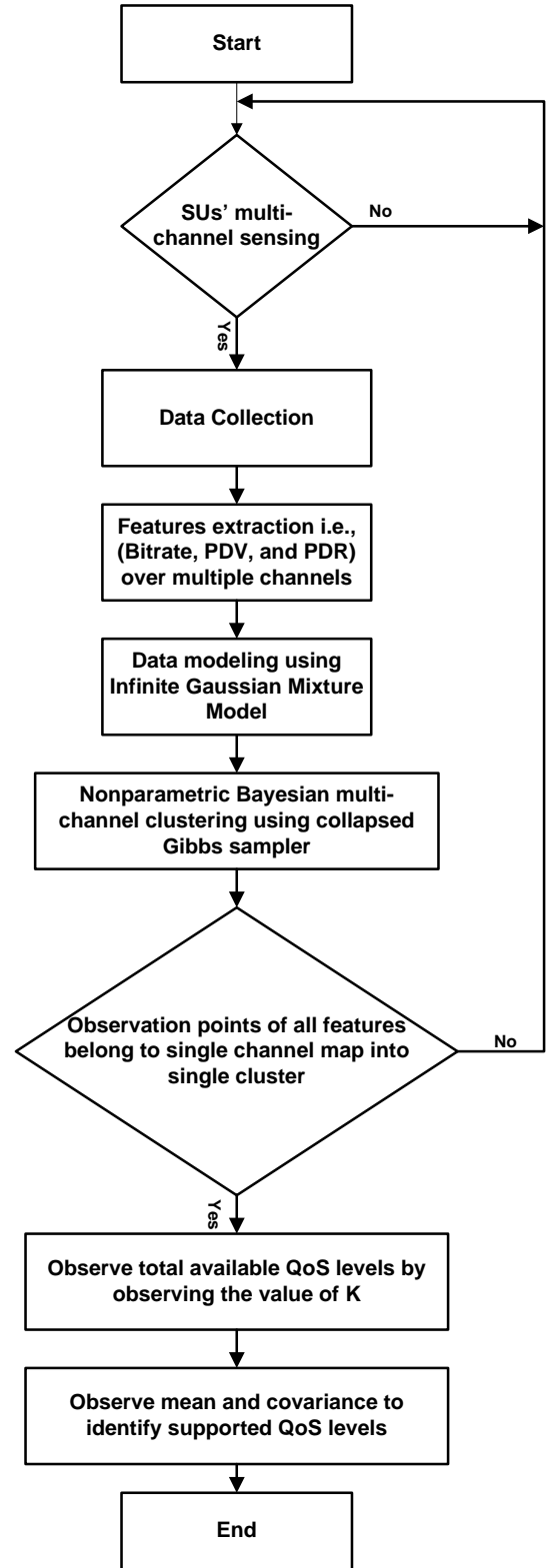


Fig. 1: QoS levels Identification.

stick breaking process (SBP) to model the variable number of clusters and the concept of countable infinity in the form of the SBP. The SBP is an elegant and direct way to represent the Dirichlet process mixture model. In the SBP, a stick of unit length (i.e., one) is divided into an infinite number of

sticks modeled as beta distribution. Whereas, the size of each piece is dependent on the α , which represents the total number of channel clusters observed from the given data set. In our channel clustering problem, each cluster specifies a certain level of QoS in terms of bitrate, PDR, and PDV. However, the IGMM is a generic model widely used for parameter estimation; therefore, for clustering accuracy, we used the MCMC algorithm.

Figure 1 presents the step-by-step procedure used to identify the total supported QoS levels. In the proposed scheme, the SU first generates observation points that are being collected from different LCs before extracting features, such as bitrate, PDR, and PDV from each LC. Subsequently, the SU models the data and performs channel clustering using the IGMM and the CGS, respectively. The SU then observes whether the observation points of all features belong to a single channel map into unique clusters before the total number of available QoS levels is identified by observing the value of K . Each QoS level is differentiated by observing the mean and covariance; otherwise, the data collection step is performed again.

C. Collapsed Gibbs Sampler

We used the CGS that generates inferences about the set of observations based on the joint probability distribution of a set of random sequences. The CGS is the realization of the MCMC algorithm, and these random samples are used to estimate the unknown parameters. The CGS classifies the clustering results into two categories: (1) when the data points belong to one of the distinct clusters from the total K available clusters and (2) the data point belongs to a new cluster. Moreover, we use it to estimate the PD for the IGMM.

IV. PERFORMANCE ANALYSIS

In this section, we first discuss the experimental setup and then we present the simulation results.

A. Experimental Setup

To evaluate the performance of the proposed scheme, we used the MATLAB tool for simulations. We performed simulations using the measurements of bitrate, PDR, and PDV from the traces of different wireless channels of WiMAX and WiBro networks. Traffic on wireless channels was captured on the bus, subway, and campus using TCPdump. In simulations, we followed the following channel notations: (1) Channel 1 refers to the WiBro wireless channel measured for the World of Warcraft (WoW) gaming application; (2) Channel 2 refers to the WiMAX wireless channel measured for the BitTorrent application; (3) Channel 3 refers to the WiBro wireless channel measured for a VoIP application; and (4) Channel 4 refers to the WiBro wireless channel measured for a UDP application.

TABLE I: Hyper parameters setting

Setting name	Λ_0	μ_0	k_0	v_0	$\Gamma(a, b)$
Setting H1	0.3	0	0.1	4	1.21
Setting H2	0.3	0	0.01	4	3.5
Setting H3	0.3	0	0.001	4	4.79

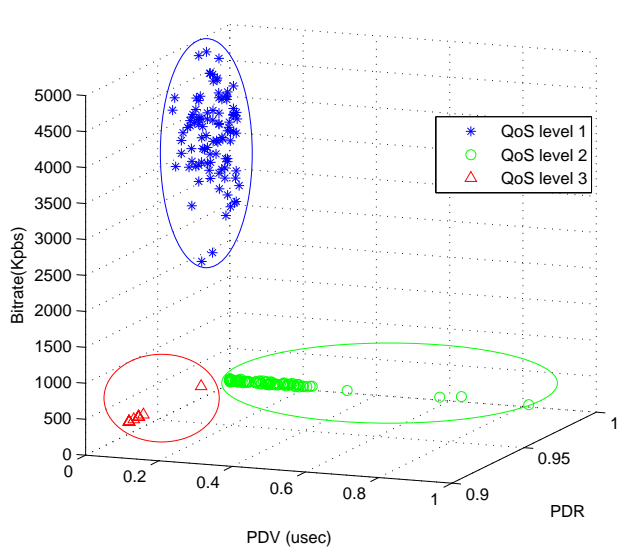
The hyper-parameters presented in Table I are used to evaluate the performance of the proposed scheme. In our proposed scheme, we used the IGMM; therefore, hyper-parameters should be set to reflect the actual true data set. The complexity of the CGS increases considerably with the increase in data points. Therefore, hyper-parameters play an important role in the fast attainment of the true data set and the reduction of the complexity of the Gibbs sampler. The IGMM uses the generative model; hence, we must set the hyper-parameters $H = \{\mu_0, k_0, \Lambda_0, v_0\}$, and Γ carefully to obtain the clustering results with minimum errors. Hence, we initialize $\Lambda_0 = \text{diag}(0.3)$, $\mu_0 = 0$, $k_0 = 0.01$, $v_0 = 4$, and $\Gamma = (3, 2)$. The vector of a cluster follows a Gaussian distribution with mean μ_0 and covariance Σ_0/k_0 . The μ_0 represents the mean of a cluster and k_0 is the dispersion measure of a cluster. A high value of k_0 depicts that the clusters are close to one another, and vice versa. The hyper-parameter v_0 represents the degree of freedom where Λ_0 represents the variability around the mean of the feature space. The covariance matrix of a cluster is highly dependent on these two parameters. For example, the feature point bitrate shows a wide range of variability from its mean value μ_0 , which is specific to a particular QoS cluster. Therefore, Λ_0 quantifies this variability, and v_0 represents our confidence about Λ_0 . We assumed that the mean values of the features are available for all QoS levels (channel clusters). After clustering, we compared and then mapped the clusters to the QoS levels.

B. Simulation Results

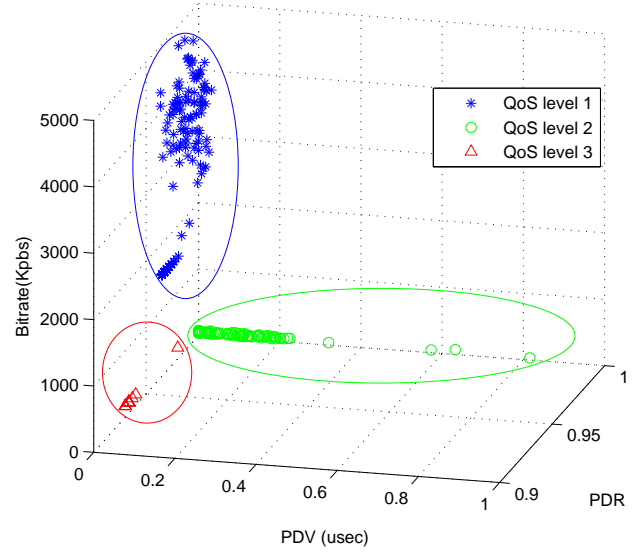
In this subsection, the simulation results of the proposed scheme are compared with the following three most common and widely used clustering schemes: (1) K-means clustering [13], 2) EM clustering [14], and 3) SC clustering [15]. The QoS levels of LCs may vary over time and location, and encountering new QoS level is anticipated. Furthermore, our proposed clustering scheme is robust in terms of QoS level identification.

Figure 2(a) clusters the three LC features and identifies their supported QoS levels without any error. In this figure, the three distinct QoS levels corresponding to the three LCs means that each channel is providing a distinct QoS level. QoS levels 1, 2, and 3 are depicted by channels 2, 1, and 3, respectively. Similarly, Figure 2(b) presents unforeseen QoS level identification, in which the features of another channel (no. 4) are included. The proposed channel clustering and the QoS level identification scheme show that its supported QoS level is similar to the already existing QoS level. Therefore, channels 2 and 4 are clustered together and form QoS level 1, which is a significant feature of the proposed clustering scheme that helps us identify multiple channels supporting the same QoS levels for assigning over multiple interfaces of MCRNs. However, Figures 2(c), 3(d), and 3(e) show the clustering errors in the EM, K-mean, and SC clustering algorithms.

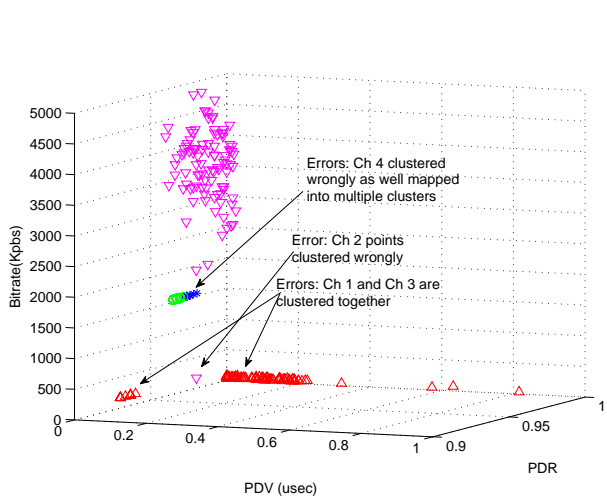
Figure 3(a) presents the receiver operating characteristic curves of all the clustering algorithms and exhibits that the proposed clustering scheme outperforms the other clustering



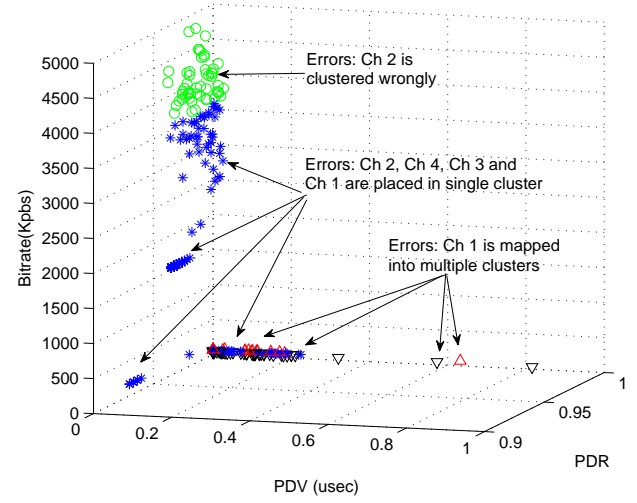
(a) QoS levels identification for three channels features



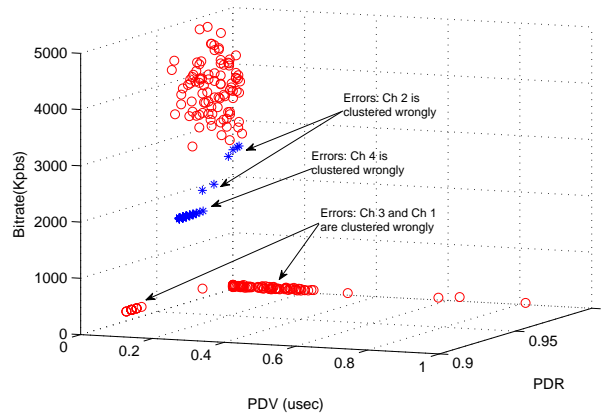
(b) Unforeseen QoS levels identification for four channels features



(c) EM clustering errors for three channels

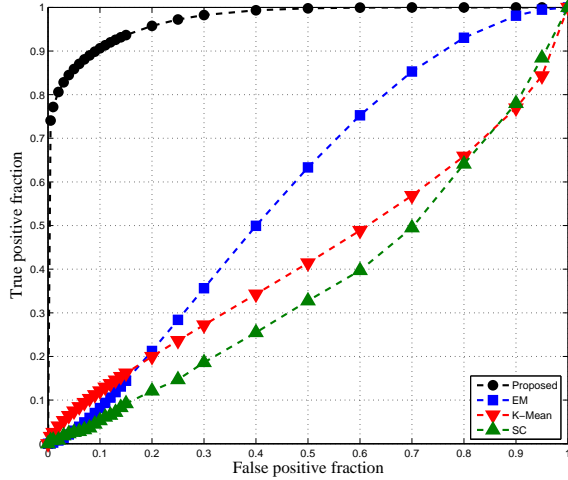


(d) K-mean clustering errors for three channels

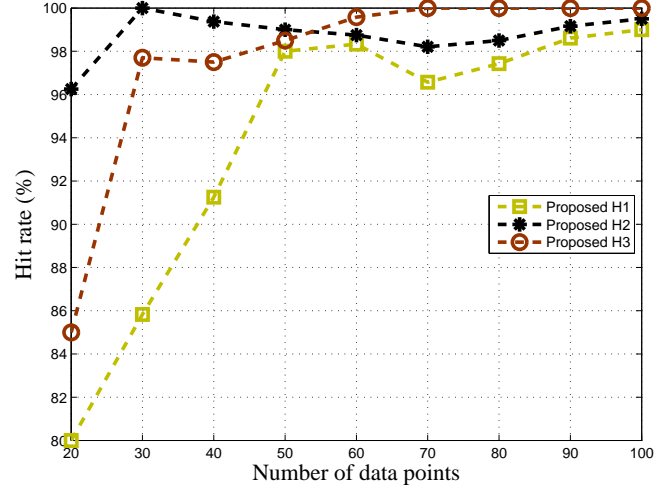


(e) Spectral clustering errors for three channels

Fig. 2: Channel clustering for QoS levels identification and error detection.



(a) Receiver operating curve



(b) Hit accuracy

Fig. 3: Channel clustering errors in EM and K-Mean clustering algorithms.

schemes. The receiver operating characteristic helps analyze the sensitivity and accuracy of a clustering algorithm and illustrates the performance of a system as it has a varied discrimination threshold.

We measured the accuracy of the proposed channel clustering and the QoS identification scheme under the different settings of hyper-parameters (Table 1). Figure 3(b) demonstrates the clustering accuracy in terms of hit rate. Hyper-parameters setting H2 is accurate when the number of data points is in the range of 20-50. However, when the data points exceed 50, hyper-parameter setting H3 becomes accurate. The figure affirms that the settings of the hyper-parameters significantly affect the clustering accuracy and that, to achieve clustering accuracy, the hyper-parameter settings should be dynamic.

V. CONCLUSION AND FUTURE RESEARCH DIRECTIONS

In this study, we proposed a novel non-parametric Bayesian based channel clustering scheme to identify the supported QoS levels over multiple available LCs. The proposed scheme exploits the PU channel usage features (i.e., bitrate, PDR, and PDV) and modeled them by using the infinite Gaussian mixture model and the collapsed Gibbs sampling method. Thus, with the help of our proposed scheme, SUs can identify an appropriate cluster or set of LCs that fulfills its stringent QoS requirements for maintaining a certain QoS threshold, without any need for communicating with the PU. Through extensive simulation results, the proposed scheme significantly outperforms baseline clustering algorithms, such as K-mean, EM, and SC. Moreover, the unsupervised nature of our proposed channel clustering and QoS identification scheme makes it highly desirable in practice, as the number of channels and their features may vary with time and the geographical location of SUs. The proposed scheme can also be extended to many other applications, in which the numbers of underlying clusters are unknown. The possible research directions of our proposed scheme are as follows:

- 1) Design the new protocols for multi-channel CRNs for video transmission and optimization. Moreover, investigating the clustering-based QoS level identification scheme by using other distributions (i.e., multinomial, binomial, and student-t) would also be interesting.
- 2) As the traffic patterns correspond with different PUs, SUs may obtain different spectral access and energy harvesting opportunities in multi-channel CRNs. Therefore, designing the novel approaches for multi-channel CRNs to maximize the CRN capacity based on the estimations of energy harvested from LCs would be interesting.
- 3) PU activities (i.e., traffic patterns) based on channel sensing may help in categorizing the available spectrum opportunities in different classes (i.e., long term, medium term, and short term). Therefore, new mechanisms must be investigated to optimize the multimedia SU transmissions based on its sensing edifice for the improved protection of PU transmissions.

VI. ACKNOWLEDGEMENT

This work is supported by the Deanship of Scientific Research at King Saud University through Research group No. (RG # 1435-051) and National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT)-NRF-2017R1A2B2012337.

REFERENCES

- [1] G. S. Uyanik and S. Oktug, "Cognitive channel selection and scheduling for multi-channel dynamic spectrum access networks considering qos levels," *Ad Hoc Networks*, vol. 62, pp. 22–34, 2017.
- [2] Z. Liang, S. Feng, D. Zhao, and X. S. Shen, "Delay performance analysis for supporting real-time traffic in a cognitive radio sensor network," *IEEE Transactions on Wireless Communications*, vol. 10, no. 1, pp. 325–335, 2011.
- [3] Z. He, S. Mao, and S. Kompella, "Qoe driven video streaming in cognitive radio networks: The case of single channel access," in *Proc. IEEE Global Communications Conference (GLOBECOM'14), Austin, TX, Dec. 2014*, pp. 1388–1393.

- [4] L. X. Cai, Y. Liu, X. Shen, J. W. Mark, and D. Zhao, "Distributed qos-aware mac for multimedia over cognitive radio networks," in *Proc. Global Telecommunications Conference (GLOBECOM'10), Miami, FL, USA, Dec. 2010*, pp. 1–5.
- [5] C. Rattaro, P. Bermolen, F. Larroca, and P. Belzarena, "A stochastic geometry analysis of multichannel cognitive radio networks," in *Proc. ACM Latin America Networking Conference (LANC'16), Valparaiso, Chile, Oct. 2016*, pp. 32–38.
- [6] W. Alhakami, A. Mansour, and G. A. Safdar, "Performance analysis of a novel decentralised mac protocol for cognitive radio networks," in *Proc. IEEE Wireless and Mobile Conference (APWiMob'16), Bandung, Indonesia, Sept. 2016*, pp. 85–91.
- [7] S. Jin and Y. Sun, "System model and performance estimation of dynamic spectrum allocation strategy with multi-channel and imperfect sensing," *International Journal of Computer Mathematics*, vol. 94, no. 9, pp. 1727–1737, 2017.
- [8] F. Tang and J. Li, "Joint rate adaptation, channel assignment and routing to maximize social welfare in multi-hop cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 4, pp. 2097–2110, 2017.
- [9] N. Tadayon and S. Aissa, "A multichannel spectrum sensing fusion mechanism for cognitive radio networks: Design and application to ieee 802.22 wrans," *IEEE Transactions on Cognitive Communications and Networking*, vol. 1, no. 4, pp. 359–371, 2015.
- [10] K. P. Murphy, "Conjugate bayesian analysis of the gaussian distribution," *Univ. Brit. Columbia, Vancouver, BC, Canada, Canada, Tech. Rep.*, 2007 [Online]. Available: <http://www.cs.ubc.ca/~murphyk/Papers/bayesGauss.pdf>.
- [11] R. Argiento, A. Cremaschi, and A. Guglielmi, "A density-based algorithm for cluster analysis using species sampling gaussian mixture models," *Journal of Computational and Graphical Statistics*, vol. 23, no. 4, pp. 1126–1142, 2014.
- [12] E. Borgia, "The internet of things vision: Key features, applications and open issues," *Computer Communications*, vol. 54, pp. 1–31, 2014.
- [13] S. Anand, S. Mittal, O. Tuzel, and P. Meer, "Semi-supervised kernel mean shift clustering," *IEEE transactions on pattern analysis and machine intelligence*, vol. 36, no. 6, pp. 1201–1215, 2014.
- [14] X. Jin and J. Han, "Expectation maximization clustering," *analysis*, vol. 41, no. 8, pp. 578–588, 2016.
- [15] L. Ding, "Research and development of spectral clustering algorithms," *International Journal of Collaborative Intelligence*, vol. 1, no. 4, pp. 275–286, 2016.

Amjad Ali is currently working as an assistant professor at COMSATS Institute of Information Technology Pakistan. He received his Ph.D. degree from Electronics and Radio Engineering Department at Kyung Hee University of South Korea in 2015. He has published number of research papers in refereed international conferences and journals. His main research interests include Cognitive Radio Networks, 5G Cellular Networks, Multimedia Cloud Computing, Internet of Things, and Vehicular Networks.

Ibrar Yaqoob received his Ph.D. (Computer Science) from the University of Malaya, Malaysia, in 2017. He is an associate editor of IEEE Access Journal. He has published a number of research articles in refereed international journals and magazines. His numerous research articles are very famous and among the most downloaded in top journals. His research interests include big data, mobile cloud, the Internet of Things, cloud computing, and wireless networks.

Ejaz Ahmed (S'13, M'17) received his Ph.D. (Computer Science) from University of Malaya, Kuala Lumpur, Malaysia in 2016. He is an associate editor of *IEEE Communication Magazine*, *IEEE Access*, *KSII TIS*, and *Elsevier JNCA*. His areas of interest include Big Data, Mobile Cloud Computing, Mobile Edge Computing, Internet of Things, and Cognitive Radio Networks. He has also received a number of awards during his research career.

Muhammad Imran is currently working at King Saud University and visiting scientist with Iowa State University. His research interest includes MANET, WSNs, WBANs, M2M/IoT, SDN, Security and privacy. He has published number of research papers in refereed international conferences and journals. He serves as a Co-Editor in Chief for EAI Transactions and Associate/Guest editor for IEEE (Access, Communications Magazine), Computer Networks, Sensors, IJDSN, JIT, WCMC, AHSWN, IET WSS, IJAACS and IJITTE.

Kyung Sup Kwak (M'17) is working as a professor and dean School of Information and Communication Engineering, Inha University, South Korea. He received the Ph.D. degree from the University of California at San Diego. He was with Hughes Network Systems and the IBM Network Analysis Center, USA. He has authored over 200 peer-reviewed journal papers. His research interests include multiple access communication.

Adnan Ahmad holds a BS (Hons), MS, and a PhD in Computer Science. He has worked for five years in industry, getting hands on experience with cutting edge software and hardware technologies. His PhD at Massey University, New Zealand, was on a formal model of distributed rights allocation in online social interaction. He has published in well-known conferences like Worldcomp, IFIP SEC, IAS, IAIT and Trustcom. He has also co-authored a book on socio-technical design.

Syed Asad Hussain is working as a professor and dean faculty of information science and technology at COMSATS Institute of Information Technology Pakistan. He obtained his Master's degree from Cardiff University, UK and PhD from the Queen's University Belfast, UK. His area of research and teaching is communication networks. He has authored and co-authored of more than 45 journal and conference papers.

Zulfiqar Ali earned his PhD degree in Electrical and Electronic Engineering from Center for Intelligent Signal and Imaging Research (CISIR), Universiti Teknologi PETRONAS (UTP), Malaysia in 2017 and MS degree from the University of Engineering and Technology (UET), Pakistan in 2007. Since 2010, he has been working as a researcher in the Digital Speech Processing, King Saud University, Saudi Arabia. His research interests include digital signal processing, privacy, and security in health-care, and multimedia forensics.