SUMMARY

As the role of wireless communication is becoming more important for realizing a future connected society for not only humans but also things, spectrum scarcity is becoming severe, because of the huge numbers of mobile terminals and many types of applications in use. In order to realize sustainable wireless connection under limited spectrum resources in a future wireless world, a new dynamic spectrum management scheme should be developed that considers the surrounding radio environment and user preferences. In this paper, we discuss a new spectrum utilization framework for a future wireless world called the “smart spectrum.” There are four main issues related to realizing the smart spectrum. First, in order to recognize the spectrum environment accurately, spectrum measurement is an important technology. Second, spectrum modeling for estimating the spectrum usage and the spectrum environment by using measurement results is required for designing wireless parameters for dynamic spectrum use in a shared spectrum environment. Third, in order to effectively gather the measurement results and provide the spectrum information to users, a measurement-based spectrum database can be used. Finally, smart spectrum management that operates in combination with a spectrum database is required for realizing efficient and organized dynamic spectrum utilization. In this paper, we discuss the concept of the smart spectrum, fundamental research studies of the smart spectrum, and the direction of development of the smart spectrum for targeting the future wireless world.

key words: smart spectrum, spectrum measurement, spectrum usage modeling, spectrum database, spectrum management

1. Introduction

More than 100 years after radio communication was first commercially used for intercommunication between ships, spectrum regulation was introduced to provide efficient spectrum use. In 1904, the Wireless Telegraphy Act was legislated in the UK for preserving the quality of radio station devices and deciding the policy of spectrum use. Then, in 1906 the International Radiotelegraph Convention (IRC) was held with the purpose of drafting the international law. In 1915, the Wireless Telegraphy Law was legislated in Japan for a future wireless world called the “smart spectrum.” There are four main issues related to realizing the smart spectrum. First, in order to recognize the spectrum environment accurately, spectrum measurement is an important technology. Second, spectrum modeling for estimating the spectrum usage and the spectrum environment by using measurement results is required for designing wireless parameters for dynamic spectrum use in a shared spectrum environment. Third, in order to effectively gather the measurement results and provide the spectrum information to users, a measurement-based spectrum database can be used. Finally, smart spectrum management that operates in combination with a spectrum database is required for realizing efficient and organized dynamic spectrum utilization. In this paper, we discuss the concept of the smart spectrum, fundamental research studies of the smart spectrum, and the direction of development of the smart spectrum for targeting the future wireless world.

key words: smart spectrum, spectrum measurement, spectrum usage modeling, spectrum database, spectrum management

1. Introduction

More than 100 years after radio communication was first commercially used for intercommunication between ships, spectrum regulation was introduced to provide efficient spectrum use. In 1904, the Wireless Telegraphy Act was legislated in the UK for preserving the quality of radio station devices and deciding the policy of spectrum use. Then, in 1906 the International Radiotelegraph Convention (IRC) was held with the purpose of drafting the international law. This was the first international regulatory spectrum allocation conference [1]. In Japan, the Wireless Telegraphy Law was legislated in 1915 and regulatory spectrum allocation was initiated. In a traditional spectrum regulatory policy, the spectrum is fixedly allocated to a system such that it neither suffers interference from nor causes interference to other systems. After World War II, the basic spectrum allocation policy was decided at the International Telecommunication Union Radiocommunication Sector (ITU-R) and the detailed spectrum allocation is now decided by a regulator of each country, such as the Federal Communications Commission (FCC) in the US, the Office of Communications (OFCOM) in the UK, and the Ministry of Internal Affairs and Communications (MIC) in Japan. This type of exclusive spectrum allocation has the advantage that inter-system interference can be avoided by ensuring a certain distance between multiple systems using the same spectrum band. In order to manage the allocated spectrum, radio licenses are issued to legal users (this spectrum allocation policy is called “command and control”). However, fixed spectrum allocation suffers the problem that the spectrum usage is not efficient. However, some spectrum sharing systems have been utilized, such as wireless LAN, ZigBee, and Bluetooth. In spectrum sharing systems, without a radio license many types of users can use wireless devices on the limited spectrum bands called unlicensed bands, such as the 2.4 GHz band (this spectrum allocation policy is called “spectrum commons”) [2]. In some spectrum bands, a spectrum auction is conducted for obtaining a radio license for commercial wireless services. In particular, the spectrums for cellular systems are allocated by auction in many countries. In these types of current spectrum allocation policies, the spectrum band is fixedly used without any real-time dynamic spectrum reallocation according to the demand of users and the radio environment.

In 1999, a new spectrum allocation concept called “cognitive radio” was considered and dynamic spectrum allocation according to the radio environment recognition was proposed [3]. The spectrum sharing system on the TV broadcasting band, called TV white space, was regulated and some services were launched in the US, UK, and other countries [4]. In a TV white space system, in order to manage the spectrum such that interference to the TV broadcasting receiver is avoided, a geolocation database storing a list of the available channels of each location is used [5]. With the current spectrum database, since a conservative propagation model is used for calculating the available channels on the database, spectrum efficiency by sharing the spectrum is not high.

The TV white space system can be interpreted as spectrum sharing in the space domain, and time domain spectrum sharing has been also investigated in dynamic spectrum access (DSA) [6], [7]. In DSA, an unlicensed user, called the secondary user (SU), can utilize a vacant spectrum originally owned by a licensed user (primary user: PU), provided
that the vacant spectrum utilization by the SU will not cause any harmful interference to the PU. To enable DSA for time domain spectrum sharing, a technique for understanding the state of the spectrum, i.e., whether it is vacant or occupied, is necessary. Thus, spectrum sensing [8], [9] was investigated in order to understand the spectrum state. Database-based techniques are suitable for managing spectrum sharing with time invariant wireless systems, such as a TV system, but spectrum sensing is rather more appropriate for wireless systems with dynamic PU traffic, since it attempts to recognize the instantaneous state of the spectrum. In addition, the DSA system does not need to rely on an external system to obtain the state information of the spectrum. However, spectrum sensing is typically required to be high accuracy, low latency, and low cost, which are very difficult to achieve simultaneously [8].

Currently, an advanced spectrum database in conjunction with spectrum measurement using huge numbers of mobile terminals has been considered for accurate spectrum environment prediction based on big data technologies [10]–[13]. In this concept, huge numbers of mobile terminals located worldwide, such as smart phones, laptop PCs, and vehicles, are used as measurement devices. The spectrum information is gathered in the spectrum database and the statistical information is registered in the database to provide efficient spectrum use according to the surrounding spectrum environment. By using this kind of database, exact interference prediction for dense spectrum sharing without harmful interference and on-demand spectrum use according to user and application demands can be achieved. This technology may drastically change future spectrum use, because the spectrum allocation policy can be dynamically changed according to the real-time spectrum environment and user demands.

In this paper, we summarize a possible future realization of a flexible spectrum utilization concept called the “smart spectrum” that can solve the spectrum scarcity problem. Thus, the sustainable development of wireless applications and technologies can be realized. In order to realize a smart spectrum world, the accurate construction of the spectrum database and the management of the spectrum are important. In this paper, we categorize the techniques required for realizing the smart spectrum as “spectrum measurement,” “spectrum modeling,” “spectrum database,” and “spectrum management.” By combining these techniques, a smart spectrum world for the future wireless era can be created.

2. Concept of the Smart Spectrum

In the smart spectrum, as compared to the typical framework of wireless communication, two different main points should be considered: spectrum allocation and management policy, and actual spectrum usage. The first point is that a user can obtain the necessary spectrum resources and achieve appropriate radio and network access based on software reconfigurations according to user demand. The former (spectrum resources) corresponds to flexible spectrum usage, in-
must be stored in the database appropriately. Specifically, statistical information and any information that is useful for spectrum sharing is extracted from the stored measurement data by computer resources (“Modeling” in Fig. 1). The role of modeling is to transform the measurement data into statistical information and the data must be stored efficiently. In this paper, two types of spectrum modeling; spectrum usage modeling and spectrum environment modeling are introduced. The spectrum usage modeling of spectrum usage based on measurements is discussed in Sect. 4 and typically this type of modeling is considered in time and frequency domains. The other type of spectrum modeling is spectrum environment modeling which is typically considered in space domain, such as propagation modeling and this is discussed in Sect. 5. In smart spectrum management, an appropriate database based on exhaustive spectrum measurements must be prepared [14]. Specifically, the database must provide site-specific information for each user in the local area. The spectrum usage based on spectrum sharing in the local area may cause interference to the neighboring local areas. Therefore, hierarchical architecture is suitable for the database in the smart spectrum (“Database” in Fig. 1). The issues related to the database are discussed in Sect. 5. In the smart spectrum, very flexible and efficient spectrum sharing and adaptability to the user’s environment and demands are required. A spectrum manager provides sufficient spectrum resources and an appropriate radio configuration and network based on the databases (“Management” in Fig. 1). The spectrum manager that manages the entire database, modeling, and measurements can be one of several entities: a third party, entrusted by a regulatory body, such as the FCC in USA and the MIC in Japan, an important stakeholder in spectrum sharing, such as an operator, and a company that has developed its business based on obtained information.

In the current framework of wireless communication based on the spectrum allocation and management policy, this concept is not applicable. Therefore, the development of a new framework for the smart spectrum management is required and this is discussed in Sect. 6.

3. Spectrum Measurement

3.1 Related Work

Early-stage spectrum measurement was reported in 1988 in [18] and this spectrum measurement provided a comparison of spectrum usage activities in different cities. Since 2000, many spectrum measurement campaigns have been conducted worldwide in the context of spectrum sharing and DSA [19]–[21]. For example, in TV broadcasting, the duty cycle (DC) can be very high: 70.9% was reported in [20] and 92.1% in [22]. The DC of cellular systems is also typically high. In [23], the measured DCs of cellular systems in indoor (a room in a building) and outdoor (roof of a building) environments were shown. In the case of an outdoor environment, the DC measured is approximately 90%; however, the DC measured in the case of an indoor environment is approximately 20%. This result implies that the measurement results also strongly depend on the situation/environment of the measurement sensors. Based on extensive spectrum measurements, several important aspects of statistical information, such as the DC, continuous occupied spectrum time, continuous vacant spectrum time, and spectrum usage state transition rates modeled by a Markov chain were discussed in [6], [46]. The details of the model of spectrum usage in the time domain are provided in Sect. 4.

3.2 Spectrum Measurement for Smart Spectrum and Issues

In the context of the smart spectrum, it is necessary to achieve wide-band, long-term, and wide-area spectrum usage measurement. The frequency range for the measurement is between several dozen of MHz and several GHz, the time range for the measurement can be a few years, and the space range for the measurement can cover one city.

There are three issues in the spectrum measurements. The first is the accuracy of the spectrum measurement. For the conventional/spectral measurements, a simple technique, such as one utilizing energy detection (ED) [24], has been used. As mentioned in [19], [25], the achievable detection performance of ED significantly depends on the accuracy of the threshold setting. In the case of wide-band spectrum measurement, a single threshold is not suitable, but frequency selectivity has to be considered. In addition, the performance level of ED is not very high [26]. Any advanced signal processing, such as feature detection, can improve the detection performance; however, it may increase the computational cost. The second issue is the cost of the spectrum measurement. For achieving extensive spectrum measurements in a huge measurement space, a single sensor is not sufficient, and multiple sensors have to be deployed in the measurement area. Therefore, an inexpensive sensor is preferred to an expensive sensor having a high performance level. For achieving accurate measurement, cooperation among multiple sensors is a reasonable approach and was investigated in a cooperative spectrum sensing study [27], [28]. There are also different types of cost, such as the amount of data treated in the spectrum measurement process. Specifically, each sensor has to send observed data to a fusion center/database Raw data, such as I-Q complex samples, can be very large and the communication cost to send the observed data is significant. The communication cost is related to the data updating period required in the database and depends on the location, time, environment of the site, and interesting statistical information. In a measurement campaign, the empirical results of which are shown in Fig. 6, the observed bandwidth was set to 40 MHz, one continuous measurement time duration was 0.1 sec, and the period for data transmission from a sensor to the database was set to 60 sec.

The size of the raw data in one continuous measurement was approximately 0.5 G bits. In the case of n sensors, the data size is 0.5 × n G bits. A considerable amount of information is necessary to achieve accurate statistical information estimation and this typically increases the communication
cost. The observed raw data can be compressed by a preprocessing procedure and this adds to the computational cost at the sensor. To resolve this issue, efficient data fusion was investigated in the context of cooperative spectrum sensing [29] and this technique may be also useful for the spectrum usage measurements. Spectrum usage measurement based on a cooperative strategy is available if the measurement sensor network consists of a central control station and a sensor network. In this case, the central control station fills the role of a fusion center. The third issue is the specification of the spectrum measurement. In fact, an appropriate specification of the spectrum measurement signal processing, such as the resolutions of time and frequency, strongly depends on actual spectrum usage parameters, for example, continuous signal length and signal bandwidth. Past measurements have not considered the appropriate measurement specifications carefully, but a single specification is typically used for the wide-band spectrum measurement. However, this led to the spectrum measurements lacking reliability.

In the case of the smart spectrum, not only the above three issues, but also a few more important issues must be considered. In the context of the smart spectrum, prior information, such as statistical information obtained by spectrum usage measurements, is utilized to enhance the efficiency of spectrum usage in DSA and spectrum sharing. In this case, it is necessary to find prior information that is useful for managing the smart spectrum. For example, spectrum utilization rates, which can be represented by the DC and the transition rate in the Markov chain, are utilized in spectrum sensing, and it has been shown that this information can enhance the spectrum sensing performance [30], [31]. As mentioned in [19], [30], [31], some prior information is useful for efficient spectrum sharing, such as both the spectrum sensing and channel access [32], [33]. However, it is still not clear what type of prior information or statistical information is very useful for managing the smart spectrum and this issue has not been investigated in-depth. Although the spectrum utilization rates, such as the DC, may be useful for spectrum sensing, as confirmed in [30], [31], a situation exists where the spectrum utilization rates will provide no benefits. Specifically, in the case described in [30], if the rate is time-invariant, there is no gain in the spectrum sensing. This implies that aspects of statistical information and the effect of statistical information on the smart spectrum should be clarified. Errors may exist in the statistical information estimation and these errors affect the spectrum sharing performance [34]. The statistics of the errors (accuracy of the measurements) are determined by using the spectrum usage measurements. Achieving more accurate measurements requires a higher cost, such as a greater number of samples in the spectrum usage measurement. For efficient spectrum sharing in the smart spectrum, it is necessary to consider the interdependent relationships among the issues in spectrum usage measurements, modeling, and spectrum sharing.

3.3 Investigations of Efficient Spectrum Measurement

As mentioned in the previous subsection, several challenges and issues remain to be resolved for the smart spectrum. Several studies have addressed these challenges and issues and some examples of these studies are introduced in the following.

In the study reported in [35], we investigated an efficient spectrum usage detection method. The entire process consists of three parts: Welch FFT-based power spectrum estimation [36]–[38], ED with noise power estimation [25], and simple signal area estimation [35], as shown in Fig. 2. We also conducted an actual spectrum usage measurement campaign, the empirical results of which are shown in Fig. 6. The main target of spectrum usage measurement was a 2.4 GHz wireless local area network and the observed spectrum bandwidth was set to 40 MHz. A Real-Time Spectrum Analyzer RSA611A (Tektronix) was used to obtain I-Q samples for 0.1 sec continuously. The remaining processes, Welch FFT, ED, noise floor estimation, and signal area estimation, were executed by a computer. The results shown in Figs. 3, 5 were obtained based on the above assumption by computer simulations. This approach can recognize spectrum usage in the time frequency domain. Therefore, it is suitable for statistical information estimation.

In the Welch FFT, averaging in the time domain is used for suppressing the effect of the noise and randomness of data symbols. The randomness causes frequency selectivity in the estimated power spectrum and leads to significant degradation of the signal detection [38]. However, averaging leads to the expense of spectrum resolution, and therefore, excessive averaging may lead to overestimation of the signal bandwidth. This corresponds to setting an appropriate time and frequency resolution in ED based on the Welch FFT. An appropriate averaging size corresponds to an appropriate segment size in the Welch FFT and depends on the signal-to-noise ratio (SNR) in power. Therefore, in fact it is difficult to set the appropriate segment size. To resolve this issue, a method to set the appropriate segment size based on the
noise floor (NF) estimation was proposed in [37]. In this method, NF estimation is used, and it has been shown that the appropriate segment size can be achieved without SNR information. This investigation corresponds to appropriate time and frequency resolution setting for accurate spectrum usage measurements.

Based on the power spectrum estimation, ED has been widely used for spectrum usage detection in spectrum measurements. In this case, threshold setting is an important issue and the NF has to be estimated appropriately. One difficult issue in NF estimation is that the state of a spectrum can be either occupied or vacant and the state is usually unknown. The method proposed in [25], median filter-forward consecutive mean excision-β(MED-FCME-β) can achieve an appropriate threshold setting, since this technique can approximately recognize noise samples and they are used for noise power estimation. This study was extended to MED-FCME-β-based NF estimation based on the Welch FFT, as reported in [38], and it was confirmed that the noise power estimation performance can be also improved by the averaging in the Welch FFT. The NF estimation ratio, which is defined as the estimated NF to the true NF ratio in dB, is shown in Fig. 3. An NF estimation ratio of 0 dB indicates no error; the poorest NF estimation ratio of MED-FCME-β is about 1 dB. The median filter in MED-FCME-β-based NF estimation performs well where the DC is less than 50%. However, Welch FFT can achieve a lower estimation error performance for any DC and SNR, as shown in Fig. 3. This investigation corresponds to setting the appropriate threshold for accurate spectrum usage measurements.

DC is a metric that represents the spectrum usage ratio or the probability of spectrum occupancy. Let $D_{n_T,n_F}$ denote the state of spectrum usage at time $n_T$ and frequency $n_F$ as

$$D_{n_T,n_F} = \begin{cases} 
1 & \text{(occupancy)} \\
0 & \text{(vacancy)} 
\end{cases}$$

Now, DC is defined as

$$DC = \frac{1}{N_T} \sum_{n_T} N_F \left( 1 - \prod_{n_F} \left( 1 - D_{n_T,n_F} \right) \right),$$

(2)

where $N_T$ and $N_F$ indicate the observed time duration and frequency range, respectively. Typically, $n_T$ and $n_F$ correspond to index numbers for the time slot and frequency bin, respectively. In spectrum measurement, $D_{n_T,n_F}$ denotes the spectrum usage detection at the $n_T$-th time slot and the $n_F$-th frequency bin.

One significant difference between spectrum measurement and spectrum sensing is the required latency\(^1\). In the case of spectrum sensing, latency can be an issue, since it requires that the instantaneous state of the spectrum be recognized [7]. In this case, the latency due to spectrum sensing has to be less than one continuous spectrum usage (either occupancy or vacancy) time and it can be a few ms. However, spectrum measurement attempts to understand the behavior of spectrum use, and therefore, the latency requirement is very relaxed as compared to that of the spectrum sensing. The required latency in the spectrum measurement can be a few hours and depends on the type of statistical information that is used in spectrum sharing. An efficient spectrum measurement, i.e., one that is accurate and of low complexity, can be achieved by employing post-processing, as shown in Fig. 2. An investigation of signal area estimation as post-processing was reported in [35]. The outputs of FFT-based ED are shown in Fig. 4(a). In the case of digital communication, typically the shape of the area occupied by one time continuous signal, such as packet data, is rectangular, as shown in Fig. 4(b). The idea of signal area estimation is to estimate the rectangular shape based on the ED outputs. In FFT-based ED, the states of spectrum usage among neighboring frequency bins/time slots may be correlated, and combining them can provide a gain in the spectrum usage detection. Approaches similar to the signal area estimation approach were presented in [39]–[41].

In [35], a simple signal area (S-SA) estimation method, which consists of two estimations, width and height, as shown in Fig. 4(b), was presented. The process of S-SA estimation can enhance the detection performance, but it increases false alarms inherently. To resolve this issue, we also proposed a false alarm cancellation (FC) technique. We showed that, based on numerical evaluations, our proposal, simple signal area estimation with false alarm cancellation (S-SA+FC), has advantages in terms of not only detection performance, but also computational cost [35]. The advantage of the detection performance of S-SA+FC is visualized in Fig. 2. Specifically, the output of ED (Fig. 2(b)) includes several miss detections (the detection result argues the spec-

\(^1\)In this study, spectrum sensing is used for detecting instantaneous spectrum usage and the detection result is either “occupancy” or “vacancy” and it is typically formulated as a binary hypothesis. Spectrum measurement is used to obtain statistical information regarding spectrum usage. Therefore, it typically consists of spectrum usage detection and the estimation of statistical information. This indicates that a spectrum sensing technique is used in the spectrum measurement.
trum is vacant, whereas in fact it is occupied) and false alarms (the detection result argues that the spectrum is occupied, whereas in fact it is vacant). However, the output of the signal area estimation may include very few errors (miss detections and false alarms). Figure 5 shows an evaluation of the detection probability for several methods: S-SA+FC, S-SA, contour tracing (CT) based signal area estimation, and ED. It can be seen that S-SA+FC can achieve the best detection performance. CT is a typical approach in image processing [42]. It is designed to extract an arbitrary signal area, whereas the S-SA estimation is designed to estimate the rectangle shape. In comparison with ED, S-SA+FC can achieve more than a 4 dB gain in SNR. In addition, a comparison in terms of computational cost showed that S-SA-based signal area estimation can achieve the lowest computational cost [35]. This investigation addressed the issues of accuracy and computational cost.

4. Spectrum Usage Modeling

4.1 Related Work

In spectrum usage modeling, several points of view, such as power, time, space, and frequency, are considered [19], [43]. In the time domain, the DC, which indicates the probability of spectrum occupancy, and the Markov chain, which shows the burst characteristics of traffic, are used as metrics for the modeling. In the space domain, mainly propagation is considered, which is discussed in Sect. 5. In the frequency domain, the distribution of the DC in the frequency domain was investigated [44]. In [45], an investigation of the distribution of the number of idle channels was presented. This corresponds to the spectrum usage model in the time and frequency domains.

As mentioned in [43], there are two main approaches for modeling: empirical and theoretical. In both approaches, statistics, such as the probability density function (PDF) and cumulative distribution function (CDF) of the observed power and the DC, and a Markov model representing the behavior of spectrum usage [19] are employed. In previous investigations of the modeling, two main requirements were determined. The first is that a model must describe the aspect of the spectrum usage appropriately. This means not only that the statistical information (e.g., PDF) fits the measurement results (e.g., empirical PDF) well, but also that the model regenerates the aspect appropriately. The second requirement is that a model must be tractable for analysis. In terms of the PDF of the DC, typically beta distribution/Kumaraswamy distribution has been used and it is well known that Kumaraswamy distribution is considerably more tractable than beta distribution [43]. The spectrum usage is strongly related to the behavior of the users. For example, the traffic of cellular phone users is varied according to time. e.g., the traffic is high during daytime, but low during nighttime. The deterministic model proposed in [46] can capture this type of behavior. An example of spectrum usage measurement, which can be expressed by the deterministic model, measured at Tokyo University of Agriculture and Technology, Tokyo, Japan (35°41′55.8″N 139°31′00.6″E) is shown in Fig. 6. The number of days for the measurements was 29 and the sixth channel of a 2.4 GHz IEEE 802 wireless local area network was used.

However, there is also traffic that has no deterministic
aspect, such as that related to public safety. In addition, the estimated DC at a given time, $DC(t)$, is also random. In this case, the stochastic model presented in [46] is suitable for modeling the aspect of spectrum usage. An example of a stochastic model in terms of $DC(t)$ during $t = 14:00$ to $t = 16:00$ is shown in Fig. 7. There are two models, one is beta distribution and the second is mixed beta distribution. In the beta distribution model, there are two parameters, whereas in the mixed beta distribution model, there are five parameters. Obviously, a greater number of parameters (five parameters) can achieve more accurate fitting than two parameters, but this requires that more information be stored in the database.

4.2 Modeling for Smart Spectrum and Issues

As mentioned above, the main purpose of the modeling in previous studies was to capture the behavior of spectrum usage. In the smart spectrum, information obtained by measurements is utilized for achieving efficient spectrum sharing. Therefore, it is necessary to know how to utilize the information and which information facilitates efficient spectrum sharing. In the case of channel access in the context of spectrum sharing between a PU and an SU, information that indicates the probability of a busy state in the channel, such as the DC, may enhance the spectrum usage [32]. Specifically, an SU may access a channel in which the probability of a busy state, denoted by $Pr(H_1)$, where $H_1$ indicates a busy channel state, is the lowest. This type of information can also enhance the spectrum sensing performance [30], [31].

However, there are several types of DC. If the traffic has a deterministic aspect [46], the DC has to be varied according to time and is denoted by $DC_i(t)$. The DC averaged over one day is denoted by $DC$. Clearly, channel access based on $DC(t)$ can achieve a better performance than that based on $DC_i(t)$. If the traffic is bursty, it can be captured by a two-state Markov model [46] and the two-state probabilities for $H_1$ are $Pr(H_1|H_1)$ and $Pr(H_1|H_0)$, where $H_0$ indicates that the channel state is idle. In this case, $DC(t)$ is calculated by averaging the two-state probabilities, i.e.,

$$DC(t) = Pr(H_1)Pr(H_1|H_1) + Pr(H_0)Pr(H_1|H_0).$$

The information of the two-state probabilities and $DC(t)$ can provide a gain in spectrum sensing. Spectrum sensing based on the two-state probabilities ($Pr(H_1|H_1)$ and $Pr(H_1|H_0)$) can achieve a better performance than that based on $DC(t)$ [30].

In spectrum sensing, the test statistics for sensing depend on the probability of $H_1$ and the gain is obtained by time-varying the probability. The former spectrum sensing yields more detailed time-varying statistical information than the latter, which corresponds to averaged two-state probabilities, and this leads to the gain in the former. However, estimating the two-state probabilities increases the cost of the spectrum usage measurement. Specifically, the required time resolution of the measurement to obtain the two-state probability is smaller than the minimum time duration of the state, such as the data packet. However, for the time variant DC and the averaged DC, the time resolution requirements are considerably relaxed. Finally, we have to consider not only the expected gain in the spectrum sharing, but also the expected cost of the selection of information used for enhancing the spectrum sharing.

The estimated information has to involve error and this error leads to an unexpected performance. In spectrum sharing, the first priority is to protect other users whose priority is higher, such as the PU and incumbent user. Therefore, statistical information-based spectrum sharing has to ensure the protection of other users by considering the error in the estimated information. One straightforward means to achieve this is to design a margin based on the statistics of the errors. Obviously, a smaller margin can achieve more effective spectrum sharing. However, accurate measurement is necessary for a smaller margin, and this increases the cost of the measurement. This indicates that, in the modeling for the smart spectrum, it is also necessary to consider the trade-off between the possible cost of the measurement and the achievable measurement performance in the spectrum sharing.

Spectrum usage modeling in the smart spectrum has to capture the huge space consisting of time, frequency, and space domains, and this huge space is denoted by the observation space. It is impossible to observe the entire space, but rather spectrum usage measurement has to select sampling points in the observation space. In this case, interpolation between observation points is required. To resolve this issue, the correlation between the observed points has to be clarified. The appropriate interpolation can reduce the required amount of information in the database and decrease the number of observed points. This leads to efficient spectrum usage measurement and the database.

5. Measurement-Based Spectrum Database

5.1 Spectrum Database for Smart Spectrum

A spectrum database stores spectrum information for dynamically allocating the spectrum to wireless systems. A typical spectrum database is that utilized for spectrum sharing in TV white space [4]. TV white space is a shared...
spectrum system on a primary TV broadcasting system; secondary wireless devices are utilized on the spatially unused spectrum allocated to the TV broadcasting system. In order to avoid interference to the primary TV broadcasting system by the secondary wireless devices, the area permitting transmission by the secondary system is controlled by a geolocation spectrum database. However, since the spectrum sharing area is designed by using a conservative propagation model with a large margin, the spectrum sharing efficiency remains low. In order to realize smart spectrum management with highly efficient spectrum utilization, we considered an advanced hierarchical spectrum database based on spectrum measurement results [12]. The architecture of the hierarchical spectrum database is summarized in Fig. 8.

The upper layer of the hierarchical database manages the global area and the lower layer of the hierarchical database manages the local area. The highest layer of the database is a regulation database, in which the spectrum regulators define the spectrum utilization policies for coordinated spectrum access. The unit of the size of the supported area in this layer is a country or state. This layer can be operated by a regulator, such as the FCC in the US and the MIC in Japan. If the regulator permits dynamic spectrum allocation, the database can be used for spectrum reallocation, dynamic spectrum policy change, and dynamic spectrum sharing. The global database can be used to manage the spectrum in each region. For example, the interference management of the border area of a country and region is managed by considering the spectrum sharing condition of multiple regions. This type of long-term span spectrum utilization achieved by considering the spectrum environment can be supported in the global database. On the other hand, the lower layer database can manage the spectrum in a local area. The lowest layer can collect spectrum information measured by mobile terminals and the collected measurement information can be used to realize efficient spectrum utilization. The area database can store the statistical modeling spectrum information collected at the local database for estimating the propagation in a surrounding area. The database of each layer is connected to the spectrum managers and the modeling servers for managing the spectrum utilization and for modeling the spectrum usage and the spectrum environment, respectively. In Fig. 8, we show one example of the connection between databases and servers. In order to store the statistical information to the upper layer database, multiple layer databases are connected to the same server and manager. Therefore, the hierarchical spectrum database can dynamically manage the spectrum by considering the radio regulator policy and the measurement results collected from mobile terminals in the local area. In this subsection, we show only an example of the hierarchical spectrum database structure; however, many models can be considered according to the progress of the future wireless communication systems. Detailed examples of the hierarchical spectrum database are explained for exploring the future wireless world in the following subsection.

5.2 Radio Environment Collection with Crowdsourcing

One of the advantages of the hierarchical spectrum database concept is that the measurement data collected from mobile terminals are utilized. Currently, a huge number of personal mobile terminals, such as smart phones, laptop PCs, and wireless sensor nodes, distributed worldwide can be used. The spectrum database can be smartly constructed by using measured radio environment big data. An illustration of the system is shown in Fig. 9. Here, distributed mobile nodes take measurements of the spectrum, such as the received signal power of each node. These measurement data are collected at the local database with a terminal location and statistical information is calculated for estimating the surrounding information. Examples of statistical information are the average power, the variance at a certain location, the distribution of the received signal, bit and packet error probability, and so on. The statistical information can be selected according to the application and communication service. It is registered at the spectrum database and provided to mobile users for improving spectrum sharing efficiency and selecting a suitable spectrum for each system. The registered information is updated on a periodic basis for supporting radio environment changes.

5.3 Radio Environment Map

In this paper, we show some examples of spectrum databases, such as a radio environment map, measurement-based radio
propagation modeling using crowdsourcing, and a spectrum database containing topology information. The basic concept of the radio environment map was proposed by a research group at the Virginia Polytechnic Institute and State University [10]. The radio environment map provides the signal power of each location calculated by using the radio propagation model or measurement results. The radio environment map based on the measurement results collected from the mobile nodes can provide accurate signal power. The simplest stored data are the average received power collected from multiple nodes at the same location for removing small-scale fading and the measurement error in each measurement device, provided by

\[ P_i = \frac{1}{M} \sum_{r=0}^{M-1} |h_i[r]s_i[r] + w_i[r]|^2, \]

where \( h_i[r] \) represents the channel coefficient at the \( i \)th node location, including propagation loss, shadowing, and fading. \( s_i[r] \) is the signal transmitted from the primary user. \( w_i[r] \) is additive white Gaussian noise (AWGN). \( M \) is the number of averaged samples of each sensor. The mobile node stores the averaged received power \( P_i \) with its location as observed by a GPS device. In order to reduce the amount of data registered in the database, quantization in the spatial domain called mesh is important. For example, the region is divided using a 10 m by 10 m mesh and the measurement data observed in the same mesh are averaged for storing statistical data. The mesh size can be changed for supporting a small area in a lower layer spectrum database or a large area in a higher layer spectrum database. The required accuracy of the statistical data should also be considered when selecting the size. The interpolation among meshes is also an important issue for creating a spectrum map with a reasonable database size.

5.4 Measurement-Based Propagation Modeling

A radio environment map can achieve highly accurate received power estimation by using the data observed at the sensor nodes. However, it suffers a problem in that a huge amount of meshes is required and it is difficult to support the mobility of the transmitter in a large area. As a solution to the problem, a novel propagation modeling process based on the measurement results of huge numbers of mobile nodes can be considered. Classical propagation modeling provides an empirical formula created by researchers based on measurement experiments. In the near future, propagation modeling is expected to become a self-organized process based on huge amounts of measurement data from sensor nodes distributed worldwide. An example of the propagation prediction of distributed nodes was presented in [47]. In the proposed method, the received signal power data measured in the receivers of peer-to-peer communication in different locations in a certain area are collected at the server and a highly accurate received signal power is predicted in any location in a certain field. Figure 10 shows the distributed measurement-based propagation modeling realized by using the collected measurement data. It is important to consider the spatial domain correlation of shadowing and the selection of the effective data in the field. In the future, radio propagation modeling using artificial intelligence (AI) and learning technologies using collected measurement data from the distributed mobile nodes may become widespread.

5.5 Spectrum Database for Narrow Area Wireless System

In a distributed wireless system, such as wireless LAN on a 2.4 GHz band, it is difficult to use a meshed database because of the small communication range and the sharing spectrum policy using a carrier sense. Currently, since 2.4 GHz bands are congested because of the huge number of devices, such as smart phones, tablets, and laptop PCs, many instances of packet loss and reduction in throughput occur because of a hidden node problem among many devices and systems at train stations, airports, shopping malls, and so on. In order to provide information for supporting efficient spectrum sharing, the radio environment database for a narrow area that includes the link topology and the interference situation can be considered [48]. An example of a wireless link topology database is shown in Fig. 11. Here, the nodes check the packet header of the received packet, even if the destination address of the packet is not its own. The link information, such as the MAC address of the transmitter node and the received signal strength indication (RSSI) of the received signal can be stored to the database. The link topology and the statistical information, such as the averaged RSSI and the channel occupation ratio, are calculated at the database. Wireless systems using a shared band can receive the information from the spectrum database and adaptively select the channel and parameters for improving the spectrum sharing performance.

5.6 Field Test of Measurement-Based Radio Environment Database for TV White Space

To demonstrate an example of the implementation of the radio environment database, we present a summary of a field test of a measurement-based radio environment database for TV White Space. In this field test, the received signal power...
of the TV broadcasting signal at Kumagaya City, Saitama Prefecture, collected by using measurement devices installed in five vehicles, and collected data were stored in the spectrum database. To observe the signals from the TV broadcast station, we divided the area by a 10 m by 10 m mesh based on the geographic information system (GIS) format. The five vehicles moved in a 35 km (north-south) by 40 km (east-west) area. We implemented a spectrum sensing function on the software defined radio platform, USRP N210, supporting a 50 MHz to 2.2 GHz band. Two sets of measurement devices and a GPS receiver were installed on the vehicles, as shown in Fig. 12. The TV broadcast signals transmitted at 473.142857 MHz were measured by means of the five vehicles using a 200 kHz sampling rate. In total, the data of 2048 samples were averaged and the averaged data were stored in the hard disk drive (HDD) of a laptop PC. The data recorded at the HDD were uploaded to the radio environment database server implemented by MySQL. The measurement campaign was conducted for two periods, five days in October 2013 and five days in February 2014. The spectrum maps constructed based on the radio environment database in October 2013 and February 2014 are shown in Figs. 13(a) and (b), respectively. Since the vehicle routes were different in the two periods, the effective mesh locations are different in the two figures in some meshes. From these figures, we can understand almost the same color is painted in the same location mesh. An analysis of the detailed results derived in this measurement campaign were presented in [47]. This paper shows that we can obtain that the residual error level of the received power estimation is less than 7.0 dB in the 90th percentile of CDF.

6. Smart Spectrum Management

6.1 Spectrum Management for Future Wireless Systems

Spectrum scarcity is becoming increasingly severe every year, because large numbers of mobile terminals and many types of wireless applications are used for supporting a world in which not only humans but also things are connected. More flexible spectrum utilization is required for realizing a sustainable information society based on wireless communication. The smart spectrum is an important concept for explaining the future wireless world where dynamic spectrum use according to the surrounding radio environment and user demands will be realized. Currently, the radio spectrum is used for communication, broadcasting, positioning, radio navigation, radio astronomy, and so on. Communication and broadcasting are the main stakeholders, consuming a large amount of the spectrum. In these systems, future spectrum utilization directions will be aggregated and classified into two types [49]. In one type, multiple wireless systems are unified to form a common broadband mobile system (CBS) and the other type constitutes dynamic spectrum allocation and sharing based on spectrum demand and spectrum environment. The unified system and dynamic spectrum allocation are illustrated in Fig. 14(a) and (b).

A common broadband wireless system, shown in Fig. 14(a), can support multiple wireless applications, such as cellular systems and public safety systems. In this case, multiple operators operate the same CBS and multiple users with various applications access the same system for communication. Broadcasting systems may also share the system on a common broadband wireless system in the future. Currently, a long term evolution (LTE) system can be used for CBSs, such as cellular systems and public safety systems (PS-LTE). However, some users want to use their own dedicated systems because of security issues, short range communication, service quality, and so on. In current spectrum allocation, these dedicated systems reduce the spectrum efficiency because of its fixed spectrum allocation policy, causing spectrum division loss, spatial spectrum reallocation loss, and time domain activity loss. Dynamic spectrum allocation by considering the user spectrum demand and the spectrum environment, as shown in Fig. 14(b), would solve...
these types of problem. By using this type of dynamic spectrum allocation technique, the spectrum can be shared by multiple systems having minimum spectrum resources. To increase spectrum efficiency, spectrum management that includes highly accurate estimation of the desired channel and the interference channel is important.

6.2 Smart Spectrum Management Operating in Conjunction with Spectrum Database

In order to effectively use the spectrum by considering the above two types of spectrum utilization policy, smart spectrum management operating in conjunction with the hierarchical spectrum database, as explained in the previous section, can be considered, as shown in Fig. 15. The smart spectrum manager can be used for bridging the user communication demand and the spectrum environment by operating in conjunction with the spectrum database. In this model, the requests and preferences of users and applications are gathered at the spectrum manager and the spectrum resource is dynamically allocated to the users, taking into consideration the spectrum information stored in the spectrum database. Examples of the preferences of users and applications are the required data rate, communication interval, distance for communication, mobility of nodes, and so on. In this scheme, since real-time environment recognition can be obtained by using the spectrum manager, real-time dynamic resource allocation can be realized. Moreover, by using the information of the spectrum map, the spectrum sharing performance among multiple systems can be significantly improved, because exact interference management can be achieved by using the statistical information stored in the measurement-based spectrum database. Initially, the entire spectrum can be managed by using the spectrum manager together with the spectrum database, as explained in the previous section. For communication and broadcasting services, we described two types of future spectrum utilization directions, a unified CBS and a dynamic spectrum allocation method, in the previous subsection. In the CBS, users having different priority, such as emergency systems, public safety systems, and general users, are mixed in the same system. In order to manage priority while considering the spectrum environment and characteristics, a smart spectrum manager can be used. On the other hand, in a dynamic spectrum allocation, the allocated spectrum and spectrum sharing users can be controlled by using a smart spectrum manager. According to the demand and preference of the users, the optimal spectrum can be allocated without the users causing each other harmful interference. In Sect. 5.4, we explained the possibility of using AI and learning technologies for modeling radio propagation. In the more distant future, the spectrum may be directly managed by AI and learning technologies, and require no human decisions. However, in such a spectrum management architecture, datasets for learning are required in order to establish environment-aware intelligent spectrum management. Our proposed hierarchical spectrum database is a candidate structure of spectrum management based on the learning approach. From these points of view, smart spectrum management with a smart spectrum database is a key technology for future spectrum utilization.

7. Conclusion

In this paper, we discussed dynamic spectrum utilization that considers the wireless environment and user preferences called the smart spectrum for the future wireless world. In the smart spectrum, measurement results collected at measurement devices distributed worldwide are gathered in the spectrum database and statistical spectrum modeling for estimating the spectrum usage and the spectrum environment in the time, frequency, and space domains is used for designing wireless communication parameters for the shared spectrum. In order to realize the smart spectrum concept, mutual cooperation among spectrum measurement, spectrum modeling, spectrum database, and spectrum management is utilized. We expect that it will be difficult to change all the spectrum allocation policies in a short time, because current existing wireless systems are optimized in their own allocated spectrum. However, in the long-term view, the smart spectrum concept can drastically improve spectrum efficiency while satisfying user preferences. In a future study, we should consider the process of changing the management rule from the current fixed spectrum allocation to a dynamic environment-aware spectrum utilization based on a smart spectrum policy.

Acknowledgment

A part of this research and development work was sup-
ported by the MIC R&D for Network Establishment Utilizing Vehicle-to-Vehicle Communications, and MIC/SCOPE #135003005, #165003006, #165003007.

References


[38] K. Umebayashi, R. Takagi, N. Iorio, Y. Suzuki, and J.J. Lehtomäki,


Takeo Fujii was born in Tokyo, Japan, in 1974. He received the B.E., M.E. and Ph.D. degrees in electrical engineering from Keio University, Yokohama, Japan, in 1997, 1999, and 2002, respectively. From 2000 to 2002, he was a research associate in the Department of Information and Computer Science, Keio University. From 2002 to 2006, he was an assistant professor in the Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology. From 2006 to 2014, he was an associate professor at the Advanced Wireless Communication Research Center, the University of Electro-Communications. Currently, he is a professor at the Advanced Wireless and Communication Research Center, the University of Electro-Communications. His current research interests are cognitive radio and ad-hoc wireless networks. He received the Best Paper Award at IEEE VTC 1999-Fall, the 2001 Active Research Award in Radio Communication Systems from the IEICE Technical Committee of RCS, the 2001 Ericsson Young Scientist Award, the Young Researcher’s Award from the IEICE in 2004, the Young Researcher Study Encouragement Award from the IEICE Technical Committee of AN in 2009, the Best Paper Award at IEEE CCNC 2013, and the IEICE Communication Society Best Paper Award in 2016. He is a member of IEEE.

Kenta Umebayashi was born in Tokyo, Japan, in 1974. He received the LL.B. degree from Ritsumeikan University in 1996 and the B.E., M.E., and Ph.D. degrees from the Yokohama National University in 1999, 2001, and 2004, respectively. From 2004 to 2006, he was a research scientist at the University of Oulu, Centre for Wireless Communications (CWC). He is currently an associate professor at the Tokyo University of Agriculture and Technology. In 2010 and 2016, he was a visiting professor at the University of Oulu. His research interests lie in the areas of signal detection and estimation theories for wireless communication, and cognitive radio. He received the Best Paper Award at IEEE WCNC 2012 for a paper he authored. He is a Member of IEEE.