A Minesweeper Algorithm for Improved Signal Area Estimation in Spectrum Aware Systems

Mohammed M. Alammar
Department of Electrical Engineering and Electronics
University of Liverpool
Liverpool, United Kingdom
King Khalid University, Abha, Saudi Arabia
M.M.Alammar@liverpool.ac.uk

Miguel López-Benítez
Department of Electrical Engineering and Electronics
University of Liverpool
Liverpool, United Kingdom
ARIES Research Centre, Antonio de Nebrija University, Spain
M.Lopez-Benitez@liverpool.ac.uk

***Abstract*—Signal area estimation is a critical component of spectrum aware systems. It entails determining the subsets of elements of a time-frequency matrix where a signal is present. This study proposes and assesses the potential of a minesweeper algorithm in estimating accurately the signal area. The proposed method can be employed in two ways: as a standalone signal area estimation technique and also as a pre/post-processing technique in combination with other signal area estimation methods in order to correct signal detection errors before applying the other estimation method (pre-processing) and/or errors introduced by the other estimation method itself (post-processing). The performance of the proposed minesweeper algorithm in both application approaches is evaluated by means of software simulations. The obtained results show that, when used as a standalone method, it can provide similar or even better accuracy than other methods at a much lower computational cost. However, the best performance is obtained when used as a pre/post-processing technique in combination with other existing methods, without increasing significantly the total computation time.**

***Index Terms*—Minesweeper algorithm, signal area estimation, spectrum awareness systems, smart spectrum systems.**

I. Introduction

Wireless communication relies on spectrum awareness systems (SAS) to allocate spectrum to users. The introduction of dynamic spectrum access (DSA) has helped secondary users (SU) be assigned the spectrum without interfering with the primary users (PU). Additionally, paid users' utilised spectrum leaves significant white space (WS), which can be assigned to secondary users [1]. It is necessary to have spectrum usage detection (SUD) methods that can attain high accuracy, low cost, and low latency to achieve an smart spectrum assignment to the SU [2], [3]. The idea of SUD can be useful in other application scenarios such as compliance verification and enforcement of spectrum regulations, and network planning and optimisation. Thus, there is a need to implement techniques that ensure high performance in SAS.

In this context, signal area (SA) estimation is an essential process in SAS. It entails determining the subsets of elements of a time-frequency matrix where a signal is present, where each element of such matrix corresponds to a measured frequency bin and time slot. A SA is a rectangularly-shaped cluster of tiles detected as occupied (i.e., where a signal component is present). Several SA estimation methods with varying performance have been proposed in the existing literature. As an illustration, [4] identifies fast Fourier transform (FFT) based energy detection (ED) and signal area estimation (SAE) as the standard SUD techniques. A previous study in [5] investigated FFT and ED's concepts, proposing their associated best methods. In contrast, this work focuses on the SAE approach. The study in [2] reveals that the SAE mechanism's application is associated with reduced costs and increased efficiency. However, [4] points out that the technique's potential in determining SA is jeopardised by practical limitations such as false alarms and signal missed detections.

The detection of the SA can be undertaken using several methods. FFT-ED techniques [2], [4], contour tracing SA (CT-SA) estimation methods [6] or the simple signal area (SSA) algorithm proposed in [7] are some of the existing SA detection procedures. The three methods (ED, CT-SA and SSA) have been discussed in [8] and their performances under different configuration parameters are investigated in [9].

In this work, a novel SA estimation technique based on a minesweeper algorithm (MA) is proposed to improve the performance of existing SAE methods. The proposed method processes each tile in the time-frequency matrix where the SA is to be detected to determine the most likely state (busy/idle) based on the number of busy tiles around it. This MA approach can be applied as a standalone SAE method but the best performance is obtained when combined with other SAE methods such as CT-SA and SSA. The proposed MA method has a low complexity and therefore does not increase significantly the computational cost of existing SAE methods while providing noticeable accuracy improvements.

The rest of the paper is organised into four sections. First, Section II provides a formal description of the signal area estimation problem considered in this work along with an overview of SA estimation techniques proposed in the existing literature and a description of the MA-based SAE method proposed in this work. Section III describes the current study's methodological framework for assessing the performance of the proposed MA under various parameter configurations and operational conditions. The analysed research findings are presented and discussed under Section IV. Lastly, Section V summarises and concludes the study.

II. SIGNAL AREA ESTIMATION

A. Problem Description and Formulation

SA estimation originates from spectrum measurements based on two-dimensional time/frequency grids, which [10] notes are composed of tiles where every element of the grid corresponds to a single time/frequency tile. A rectangular-shaped area referred to as the SA is detected when a set of contiguous adjacent tiles are detected as occupied by a signal. Accordingly, the detection distinguishes between two types of tile sets: (a) H_0 (not occupied) and (b) H_1 (occupied). Thus, the concept of SA refers to a rectangular set of tiles observed in the occupied (H_1) state. The challenge of detecting a SA is similar to the classical problem of signal detection or spectrum sensing. However, there are considerable differences. First, the focus is not on deciding the instantaneous busy/idle state of a channel but on knowing how its users exploit spectrum to understand their usage patterns in the time and frequency domains. Consequently, accurate detection of the H_0/H_1 state of every individual tile is in general not relevant. The fact holds as long as the whole detected SA (the set of tiles) is an accurate representation of the original time/frequency grid occupied by the signal (even though some of the individual tiles may be incorrect). Moreover, the detection of H_0/H_1 states in real-time is not relevant in SA estimation, as it is in spectrum sensing. This is because the outcome information is usually not useful in the short term, but in the longer-term to optimise spectrum and radio resource management decisions.

However, SAE methods depend on spectrum sensing decisions. Consequently, they are significantly affected by the same two types of errors in the signal detection process described by [2] as missed detection (busy tiles detected as idle) and false alarms (idle tiles seen as busy). The two errors affect the specific shapes of the estimated SAs, consequently degrading the performance of the adopted estimation techniques.

The SAE's performance is determined using several parameters, which can impact its accuracy and functioning. Therefore, [10] observes that the application of SAE methods involves sampling the spectrum into a set of observed power levels in the frequency and time domains. The outcome is a set of power levels corresponding to each frequency and time bin or tile in the grid. The power levels are then compared to a predefined threshold value, which produces a binary matrix indicating the H_0 or H_1 states of every tile. This binary matrix of busy/idle tiles is the input information provided to the SA estimation method to extract the tiles' rectangular sets where one or more SAs are detected in the time/frequency grid. The overarching purpose of the SA estimation approach is to identify perfectly rectangular sets of busy tiles in the time/frequency grid, which can be a challenging task given the corruption introduced by sensing errors in individual tiles.

Furthermore, the procedure of SAE is affected by the employed energy decision threshold and the time/frequency resolutions of the data grid. Accordingly, [8] posits that the time domain's resolution can be adjusted by modifying the sensing period. Additionally, the resolution in the frequency domain can be adjusted by altering the employed FFT size. Therefore, threshold selection and resolution are important parameters that affect SA estimation techniques' performance and accuracy and their effects have been investigated in [9].

B. Existing SA Estimation Methods

Various SA estimation approaches have been proposed in the existing literature. An example is a simple tile-by-tile ED in which the individual idle/busy state of each tile in the time/frequency grid is determined. In this case, the idle/busy decisions of ED are taken without further processing and therefore there is no estimation of rectangular SA within the grid. This procedure provides a reference benchmark for the present study. Additionally, this work also considers the CT-SA estimation method from [6], in which a rectangular SA is approximated based on contour tracing techniques. Further, this study considers as well the SSA estimation method described in [2], [4], [7], which is a more sophisticated method that estimates every SA present in the time/frequency grid by following a multi-step procedure. First, it performs a raster scan to find the first corner of a potential SA, then undertakes a horizontal (frequency) scan using a unit-width window with height Δt to estimate the SA's width. Next, a coarse estimation of the SA height (in the time domain) is carried out with a window of the same width as approximated for the SA and height ΔT. Finally, a fine height estimation is carried out to obtain an accurate approximation of the SA's width and height. Further details about this procedure are provided in [7].

C. Proposed Minesweeper Algorithm

As mentioned earlier, the performance of existing SAE methods can be severely degraded by the presence of signal detection errors (false alarms and missed detections). For instance, false alarms can lead the SSA method to detect SAs where they do not exist, which motivated the false alarm cancellation variant FC-SSA proposed in [2]. In many cases these errors are uncorrelated and occur in isolated random tiles rather than in clusters of tiles (e.g., false alarms are uncorrelated as they are typically caused by increased noise, which is essentially an uncorrelated random process). As a result, many false alarms tend to occur in isolated (busy) tiles surrounded by several idle tiles while many missed detections tend to occur in isolated (idle) tiles surrounded by several busy tiles. Therefore, the state of neighbouring tiles can be used as an indication to infer the potential occurrence of errors in certain tiles and take corrective actions. This is the main principle exploited by the proposed MA method.

The proposed MA method follows a two-step process. The first step entails assigning to every tile in the time/frequency grid a number that equals the number of tiles in the busy state observed in the group of 3×3 tiles composed by the tile under evaluation and the eight neighbouring tiles immediately surrounding such tile as illustrated in Figure 1. Notice that the state of the tile under evaluation (i.e., the central tile in the set of 3×3 tiles) is also counted. This counting procedure is similar to that of the popular *minesweeper* game, after which

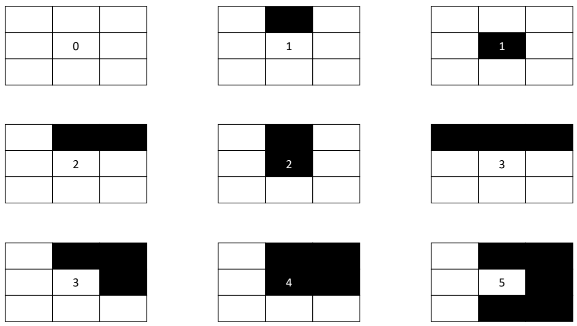
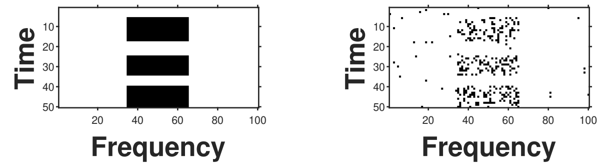


Fig. 1: An example of a tile and its eight surrounding neighbours.



(a) (b)

Fig. 2: Example of a randomly generated time/frequency test grid: (a) Clean test grid, (b) Test grid with noise (SNR = –8 dB).

the proposed method is named. The second step decides the final idle/busy state of every tile in the time/frequency grid by comparing the number assigned to each tile in the first step to a properly set threshold, which can be found by simulation as it will be shown. Tiles whose values are equal to or greater than the selected threshold are decided to be in busy state, while those with lower values are decided to be in idle state. By following this procedure, the final state of every tile is decided based on its own state and the state of the neighbouring tiles. With a properly set threshold, this method should be able to reduce false alarms and missed detections while leaving most of the other tiles detected in the correct state unchanged.

It is worth noting that two more complex formulations of the proposed method were analysed in the context of this work. The first one was a double threshold approach where tiles with values below a low threshold are decided as idle tiles, tiles with values above a high threshold are decided as busy tiles, and tiles whose values are between both thresholds are left unchanged (i.e., in their originally detected state). The second approach was based on an exhaustive analysis of all possible combinations of 0, 1, 2, . . . 9 busy tiles in a set of 3×3 tiles and the corresponding most likely idle/busy state applicable to each case. In both approaches the resulting accuracy of the estimated SA was below that attained by the single threshold approach described above and therefore only the single threshold approach will be considered in the remainder of this paper, which also has the advantage of having a more simple formulation and therefore being less computationally costly in practical implementations.

Notice that the proposed method can be applied as a standalone SAE method directly applied to the idle/busy outcomes of ED as well as a pre- and/or post-processing method for existing SAE methods (e.g., CT-SA and SSA) in order to correct spectrum sensing errors before applying a SAE method (pre-processing) and/or errors introduced by the SAE method itself (post-processing). All these cases will be analysed.

III. Methodology

A. Simulation Procedure

The current study adopts a simulation-based evaluation approach aimed at testing SAs under defined transmission constraints. The grids include channelized SAs with occupancies randomly generated under well-defined constraints. The generated test grids are then corrupted by adding noise before being fed to the evaluated SA estimation methods (ED, CT-SA and SSA) in combination with the proposed MA method. The process is executed in four main steps as follows:

Step 1. *Create clean time/frequency test grids:* A rectangular time/frequency grid is randomly generated with a resolution of 100×50 tiles, which corresponds to a medium resolution (see [9] for details). The number of tiles in the test grid's horizontal dimension (100) is given by the frequency resolution (i.e., frequency bins) while the number of tiles in the vertical dimension (50) is determined by the considered time resolution. Channelized signal areas are then generated in the test grid based on specified transmission constraints. A known number of channels is set in the frequency domain, and random on/off transmission durations are generated in the time domain from exponential distributions with rate parameters $\lambda_{on}/\lambda_{off}$ until the total height of the grid is completed for every channel. Furthermore, minimum on/off durations are specified, and guard bands between channels (i.e., idle tiles on each channel's sides) as a fraction of the channel width. Fig. 2a shows an example of a randomly generated test grid with three channels, where only the central one is in use, with rate parameters $\lambda_{on} = \lambda_{off}$ = 0.5 time units (t.u.), minimum on/off duration of 10 and 5 t.u., respectively, and guard bands of 5% of the channel bandwidth.

Step 2. *Add noise to the test grids:* Noise is added to the test grids generated in the previous step. The noise affects both types of tiles (idle and busy). In this process, the test grid's idle tiles may change to a busy state according to a predefined probability of false alarm, while busy tiles may change to an idle state according to a given probability of missed detection. The false alarm and missed detection probabilities are calculated according to an ED set for a Constant False Alarm Rate (CFAR) [9]. Fig. 2b shows, as an example, the test grid of Fig. 2a, as detected when the decision threshold is set based on the CFAR method with a target false alarm probability of 10% and an SNR of –8 dB.

Step 3. *Estimate the SAs:* In this step, one of the SA detection methods described in Section II-B is applied to the noisy test grid to estimate the SAs present in the original clear test grid. We used the MA method in combination with the CT-SA and SSA algorithms as pre- and post-processing techniques (i.e., before, after, or before & after). Examples of the SAs

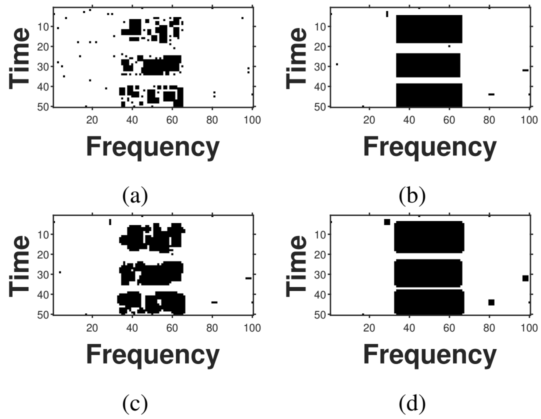


Fig. 3: An example of SAs estimated by (a) the CT-SA method, (b) MA method before CT-SA, (c) MA method after CT-SA, (d) MA method before & after CT-SA [SNR = –8 dB, $\gamma_{threshold} = 2$].

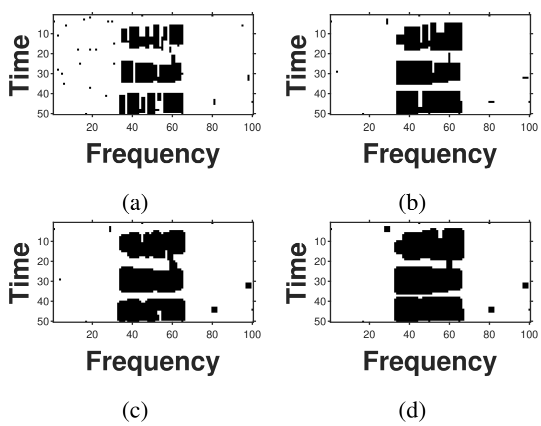


Fig. 4: An example of SAs estimated by (a) the SSA method, (b) MA method before SSA, (c) MA method after SSA, (d) MA method before & after SSA [SNR = –8 dB, $\gamma_{threshold} = 2$].

estimated by CT-SA, SSA, with and without the MA method (before, after, and before & after) based on the test grid of Fig. 2b are shown in Figs. 3 and 4 with an MA threshold value of $\gamma_{threshold} = 2$ busy tiles. As it can be appreciated, the application of the MA method removes most false alarms and fills several gaps created in the SAs by missed detections.

Step 4. *Assess the accuracy of the estimated SAs:* The final step involves evaluating the accuracy of the estimated SAs by comparing the set of SAs estimated by the considered method with those present in the original test grid. The accuracy is evaluated based on the performance metrics explained below.

B. Performance Metrics

The probabilities of detection and false alarm are commonly used to assess the performance of signal detection methods. However, these metrics are of little use in the context of SAE since the focus is not the accuracy of the detection on every individual tile of the time/frequency signal grid but the set of SA present. Accordingly, they result from some reconstruction processes in which subsets of tiles are associated and recognized together as a SA. The analysis of these two probabilities individually does not provide a complete characterization of the reconstruction's efficacy in a SAE method. Therefore, these probabilities are not considered separately in this work. Instead, other metrics that take into account the combined impact of these metrics are factored in.

The accuracy of the studied SAE methods could be assessed based on a simple accuracy metric defined as the percentage of tiles (in either state, idle or busy) that are correctly detected in their real condition, which can be obtained as the sum of true positive and true negative detection rates. However, the tiles in one of the states (idle/busy) may outnumber the tiles in the other state, thus biasing the value of such accuracy metric. This motivates using the F1 score as the main performance metric, which considers the possible imbalance between the number of tiles in idle and busy states in the original test grid. The F1 score metric is defined by [11] as follows:

$$F1 = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{1}$$

where TP is the number of tiles that are a true positive, FP is the number of tiles that are a false positive, and FN is the number of tiles that are a false negative. If the number of TP and TN (true negative) tiles is the same, the F1 score metric is equivalent to the accuracy metric defined previously. In general, the F1 score is a better metric when the ratio of actual idle/busy tiles is imbalanced, which is usually the case.

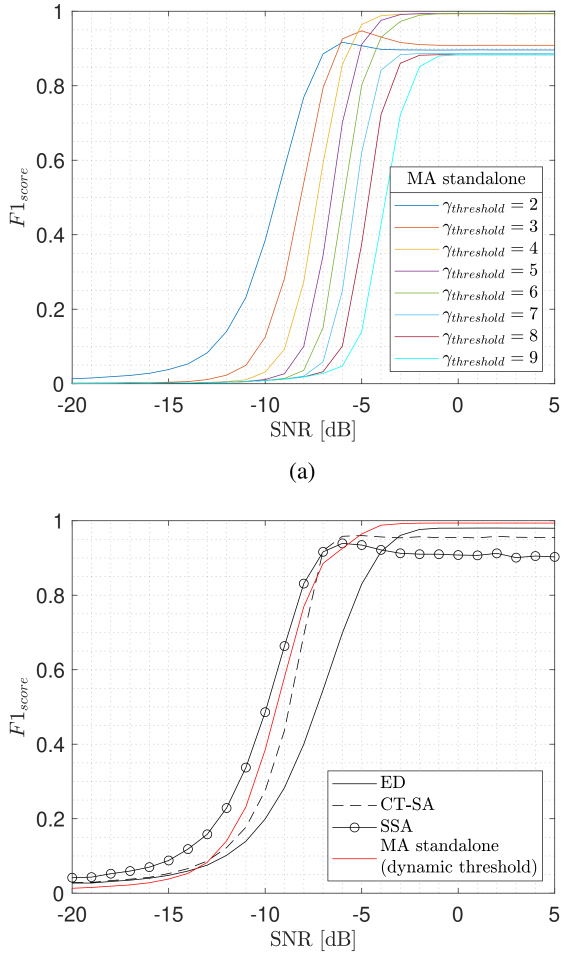
The computation time of a SAE method will affect the overall performance in practical system implementations and is evaluated as part of this study. Thus, given the existence of a direct correlation between computation time and the overall cost of implementation, as observed by [8], finding the method that attains the lowest computation time is preferable.

IV. Results

The performance of the proposed MA method, both as a standalone SAE method and as a pre/post-processing method for other SAE methods, is evaluated by simulations based on 100 different randomly generated test grids. The performance is compared with the CT-SA and SSA methods. For the SSA method, the parameters are configured as recommended in [4].

A. MA as Standalone SAE Method

The proposed MA method can be employed as a standalone SAE method directly applied to the ED outcomes only. The performance for different thresholds is illustrated in Fig. 5a. These results indicate that no single value of the threshold can provide the best performance over the whole range of SNR values. This outcome suggests that a dynamic threshold approach that dynamically selects the best performing threshold based on the current SNR could improve the method's performance. According to the results in Fig. 5a, threshold values of 2, 3 and 4 should be used for SNR values in the ranges SNR $<$ –6 dB, SNR $\in$ [–6, –5] dB, and SNR $>$ –5 dB, respectively. As shown in Fig. 5b, this dynamic MA approach clearly outperforms ED as an SAE technique over the whole range of SNR values and CT-SA over most of the SNR range (with the exception of the



(a)

(b)

Fig. 5: F1 score as a function of the SNR for the different SAE methods (ED, CT-SA, SSA) and the standalone MA method with: (a) static threshold approach, and (b) dynamic threshold approach.

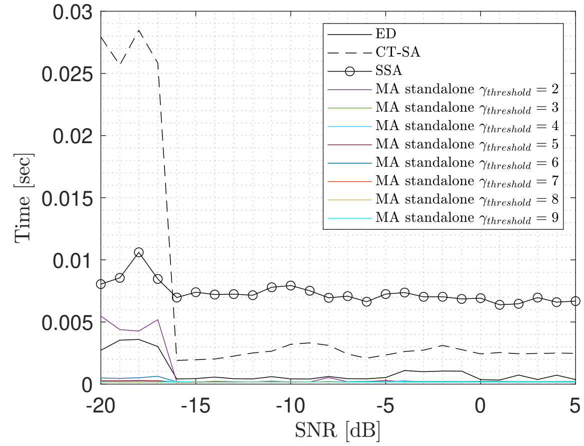


Fig. 6: Computation time as a function of the SNR for the different SA estimation methods (ED, CT-SA, SSA) and the proposed MA method as a standalone SAE technique.

range from –7 dB to –5 dB, where the accuracy is slightly lower). When compared to the SSA method, the standalone MA method performs worse below –6 dB and better above –6 dB. As mentioned earlier, false alarms can lead the SSA method to detect SAs where they do not exist (this problem has been addressed in other works such as [2], which are out of the scope of this work and therefore not considered here). At high SNR values (above –5 dB), the proposed MA method as a standalone SAE technique can provide improved F1 scores, reaching values close to 100% (the only SAE method that achieves this high accuracy).

When compared to the CT-SA and SSA methods, the standalone MA method is associated with a much lower computational cost (see Fig. 6) due to the more straightforward processing, which leads to a computation time similar to that of a simple ED strategy regardless of the threshold used for MA. Consequently, the proposed MA technique with an SNR-based dynamic threshold selection can deliver a performance better than CT-SA and comparable to that of SSA (and better at high SNR values) at a much lower computational cost.

B. MA as Pre/Post-Processing for Other SAE Methods

The proposed MA method can also be employed as a pre/post-processing technique for other SAE methods such as CT-SA and SSA. In this case, the MA method is combined with an existing SAE method in order to correct spectrum sensing errors before applying another SAE method (pre-processing) and/or errors introduced by the other SAE method itself (post-processing). The results in Fig. 7 show the performance of the CT-SA and SSA methods used as standalone SAE methods and also combined with the proposed MA method used as a pre/post-processing technique. In general, one can state that applying the MA method before the other SAE method results in a slightly better accuracy than using it afterwards. However, with some exceptions, the highest improvement is in general achieved by employing the proposed MA mechanism both before and after the SA estimation method (i.e., as pre- and post-processing simultaneously).

The results shown in Fig. 8 indicate that the total computational time of the CT-SA and SSA methods is not significantly affected regardless of whether the proposed MA method is applied. This can be explained by the fact that the MA method is based on a straightforward and much simpler process than CT-SA and SSA and therefore its contribution to the total computation time is very low compared to that of the CT-SA and SSA methods. Consequently, the introduction of the MA method as a pre/post-processing technique does not affect significantly the total computational time of the other SAE methods. Interestingly, we find that, in some cases, the MA mechanism can in fact reduce the total computational time. This phenomenon may be due to the MA method's potential to eliminate some errors in the original signal grid, thus simplifying the other SAE methods' reconstruction procedure and therefore reducing the total computation time. This explanation seems to hold true mostly in the region of low SNR values.

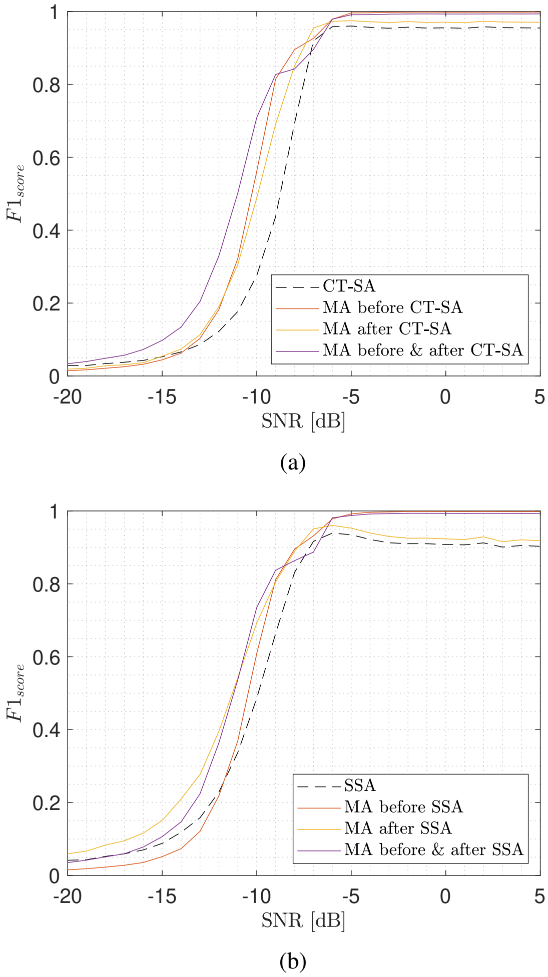


Fig. 7: F1 score as a function of the SNR for different combinations of the proposed MA method and: (a) CT-SA, (b) SSA.

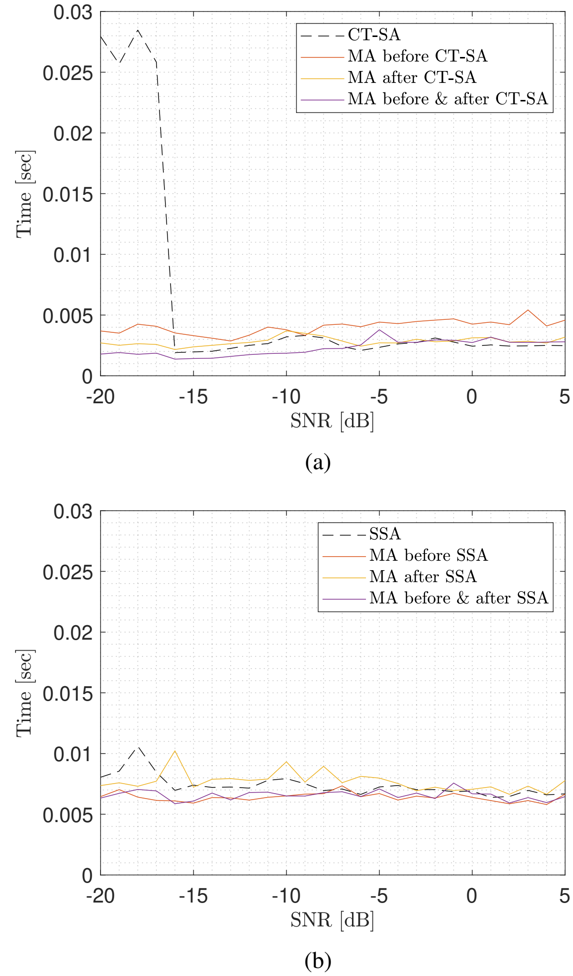


Fig. 8: Computation time as a function of the SNR for different combinations of the MA method and: (a) CT-SA, (b) SSA.

V. Conclusion

A minesweeper-based algorithm (MA) has been proposed as an effective approach for improved Signal Area Estimation (SAE). The method has a low computational cost and can be employed both as a standalone SAE method as well as a pre/post-processing technique for other SAE methods (to correct sensing errors or errors introduced by the other method itself). In both cases the proposed MA helps detect the signal area more accurately at a low computational cost.

References

[1] F. Hessar and S. Roy, "Capacity considerations for secondary networks in TV white space," *IEEE Transactions on Mobile Computing*, vol. 14, no. 9, pp. 1780–1793, Sep. 2015.

[2] R. Mizuchi, K. Umebayashi, J. J. Lehtomaki, and M. López-Benítez, "A study on false alarm cancellation for spectrum usage measurements," in *2017 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. IEEE, Mar. 2017, pp. 1–6.

[3] T. Fujii and K. Umebayashi, "Smart spectrum for future wireless world," *IEICE Trans. on Comms.*, vol. 100, no. 9, pp. 1661–1673, Mar. 2017.

[4] K. Umebayashi, H. Iwata, J. J. Lehtomäki, and M. López-Benítez, "Study on simple signal area estimation for efficient spectrum measurements," in *European Conf. Netw. & Comms. (EuCNC)*, Jun. 2017.

[5] M. López-Benítez and J. Lehtomäki, "Energy detection based estimation of primary channel occupancy rate in cognitive radio," in *2016 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, 2016, pp. 355–360.

[6] J. Kokkoniemi and J. Lehtomäki, "Spectrum occupancy measurements and analysis methods on the 2.45 GHz ISM band," in *2012 7th international ICST conference on cognitive radio oriented wireless networks and communications (CROWNCOM)*. IEEE, Jun. 2012, pp. 285–290.

[7] K. Umebayashi, K. Moriwaki, R. Mizuchi, H. Iwata, S. Tiiro, J. J. Lehtomäki, M. Lopez-Benitez, and Y. Suzuki, "Simple primary user signal area estimation for spectrum measurement," *IEICE Transactions on Communications*, vol. 99, no. 2, pp. 523–532, Feb. 2016.

[8] R. Mizuchi, K. Umebayashi, J. J. Lehtomaki, and M. López-Benítez, "A study on FFT-ED based signal area estimation for spectrum awareness (short range wireless communications)," *IEICE technical report*, vol. 116, no. 30, pp. 27–34, May 2016. [Online]. Available: https://ci.nii.ac.jp/naid/40020849217/en/

[9] M. M. Alammar and M. López-Benítez, "Evaluation of the impact of thresholding and frequency/time resolution on signal area estimation methods," in *IEEE 93rd Vehic. Tech. Conf. (VTC 2021-Spring), 7th IEEE Int'l. Works. on Smart Spectrum (IWSS 2021)*, Apr. 2021, pp. 1–7.

[10] K. Y. Sohn and Y. O. Park, "Method for transmitting and receiving random access channel signal in wireless communication system," Jun. 30 2016, uS Patent App. 14/949,168.

[11] D. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," *J. Mach. Learn. Technol*, vol. 2, pp. 2229–3981, Jan. 2011.