Radio Frequency Interference Detection in Microwave Radiometry Using Bayesian Detection

Imara Mohamed Nazar and Mustafa Aksoy
Electrical and Computer Engineering
University at Albany, SUNY
Albany, NY, USA
Emails:{imohamednazar, maksoy}@albany.edu

***Abstract*—Herein, a probabilistic approach is proposed to detect and mitigate radio frequency interference (RFI) in microwave radiometer measurements. Simulated RFI-free and RFI-contaminated radiometer measurements are analyzed at the feature level while considering their probability density functions. For each measurement, the features are reviewed sequentially. The final decision is made based on maximum a posteriori (MAP) estimation. The proposed framework is evaluated using monte–carlo simulations. The comparison with the existing RFI detection techniques through numerical simulations demonstrates the ability of the proposed approach to detect RFI even when the interference to noise (INR) of the RFI is as low as -15 dB.**

I. Introduction

Radio spectrum is a limited resource that has been allocated for passive services is contaminated by RFI [1]. Thus far, multiple studies have been conducted to detect RFI in microwave remote sensing measurements in which researchers have mainly focused on a single domain (e.g., time, frequency) to identify RFI. Since the nature of the RFI is more dynamic, the detection algorithms which utilize multiple features demonstrate the ability to detect more RFI than relying on a single feature [2]. Considering the drawbacks, we have previously introduced techniques such as support vector machines (SVM) [3] and density–based spatial clustering of applications with noise (DBSCAN) [4]. To be noted, these algorithms are capable of analyzing multiple features together. In this study, we propose a statistical method [5] to detect RFI in microwave radiometry. Here, the problem is formulated as a binary hypothesis testing that utilizes Bayes theorem along with class conditional independence assumption of the features.

II. Problem Formulation

Consider a set $\mathcal{R}$ of radiometer measurements where each measurement $\mathbf{r} \in \mathcal{R}$ is described by a d–dimensional feature vector $\mathbf{r} \triangleq [X_1, X_2, \ldots, X_d]^\top$. Furthermore, each radiometer measurement r belongs to one of two hypotheses namely, RFI–contaminated ($H_\mathcal{C}$) and RFI–free ($H_\mathcal{N}$). For each feature value $X_n = x_n$, the probability $P(x_n|H_\mathcal{C})$ (similarly $P(x_n|H_\mathcal{N})$) of the evaluation of nth feature is when the true hypothesis is $H_\mathcal{C}$ (similarly for true hypothesis $H_\mathcal{N}$) is empirically calculated from the training data. Further, the *a priori* probability $P(H_i)$ of measurement r being an hypothesis of $i = \{\mathcal{C}, \mathcal{N}\}$ is also estimated empirically. For given radiometer measurement r, the posterior probability is denoted by $P(H_i|\mathbf{r})$. In Bayesian detection method, the prediction output is estimated using maximum a posteriori (MAP) method as follows:

$$\hat{y} = \underset{i=\{\mathcal{C},\mathcal{N}\}}{\arg\max}\, P(H_i|\mathbf{r}) \tag{1}$$

III. Experimental setup

The RFI–free radiometer measurements are modeled as white Gaussian noise. On the other hand, the RFI is assumed as a pulsed sinusoidal signal. This assumption, when the duty cycle (DC) and the INR values are allowed to vary, covers wide range of RFI observed in microwave radiometry. The interference source is assumed to be 12 MHz sinusoidal source with 1 V amplitude. The pulse repetition interval of the source is set to 300 μs. To study the performance of the proposed approach against the interference source properties such as INR and DC the INR is varied from -20 dB to 10 dB in steps of 1 dB and DC is varied from 1 % to 100 % in steps of 1 %. Pulse–sinusoidal interference source and the white Gaussian noise is added to create the RFI-contaminated data. Generated measurements are sampled at a rate of 96 MSPS. The features such as power and kurtosis are calculated within 300 μs radiometer integration periods. Note that the values for the sampling rate and integration period are set similar to the SMAP radiometer [6].

IV. Bayesian Detection Implementation

RFI–free and the RFI–contaminated measurements are generated as discussed in section III. Three thousand one hundred datasets are created by varying the interference source properties such as INR and DC. Each dataset consists of power and the kurtosis feature values of 400 RFI–free measurements and 400 RFI–contaminated measurements. To avoid overfitting, we have used a widely used technique in machine learning applications called cross–validation. Here, we have implemented the five–fold cross–validation where each dataset is divided into five parts and each part is tested by the model trained using the remaining four parts.

We have observed that the power and kurtosis values are continuous random variables. To calculate the class conditional probabilities of these continuous random variables, we have used the binning approach. Specifically, the probability of the occurrence of the nth feature, i.e., $Xn = x_n, n =$

$\{1, 2, \ldots, d\}$ when the true hypothesis is H_i is estimated as $P(Xn = x_n|H_i) = \frac{N_{n,i}+1}{N_i+K}$. Here, $N_{n,i}$, N_i, and K denote the number measurements with feature value of nth feature is $Xn = x_n$ and belong to the hypothesis H_i, the number of training measurements with the hypothesis H_i, and number of bins, respectively. Further, the empirical class prior probability is derived as maximum likelihood estimates (MLE) of training data as $P(H_i) = \frac{N_i}{\sum_{i=\{\mathcal{C},\mathcal{N}\}} N_i}$.

For each measurement in the test data, the MAP is estimated using (1). Finally, the Bayesian detection model is evaluated using three performance metrics, namely (i) accuracy, (ii) precision, and (iii) recall. The accuracy gives the fraction of correctly predicted measurements in the test data. Precision gives the fraction of correctly predicted RFI–contaminated measurements out of the total number of the measurements that are predicted as RFI–contaminated. The recall denotes the fraction of correctly predicted RFI–contaminated measurements out of the RFI–contaminated measurements in the dataset.

V. Results

Fig. 1 shows the performance of the proposed approach. As seen in the figure, the performance of the proposed approach is increasing with increasing INR and DC values.

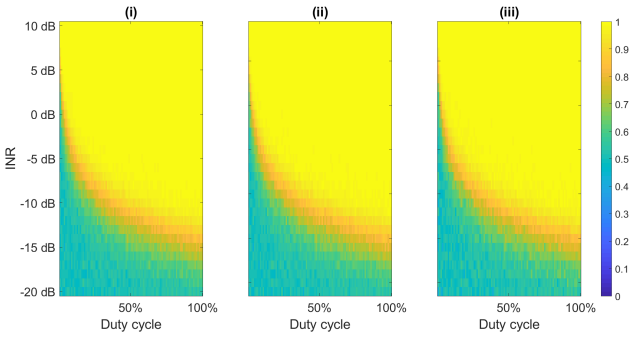

Fig. 1. Accuracy (i), precision (ii), and recall (iii) values of the proposed approach as a function of INR and DC.

Further, we have compared the proposed approach with traditional RFI detection approaches such as pulse blanking, kurtosis detection, and pulse blanking or kurtosis detection (OR method). Here, both pulse blanking and kurtosis detection are implemented by thresholding such that the power or kurtosis values that fall three standard deviations or further apart from the mean are considered as RFI–contaminated. OR method is implemented by combining the labels derived from both pulse blanking and kurtosis detection using a logical OR operator. The details of the traditional RFI detection algorithms are explained in our previous work [3]. In that study, we have observed that OR method performs better than other two. Therefore, here, we have compared the performance of the proposed approach with OR method. Fig. 2 shows the performance difference between the proposed approach and OR method as a function of INR and DC. Comparing with

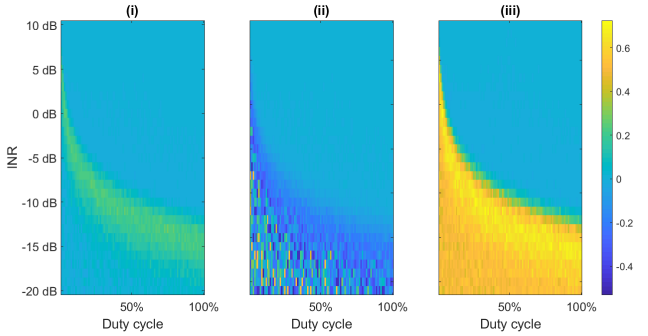

Fig. 2. Accuracy (i), precision (ii), and recall (iii) difference between Bayesian detection and the OR method as a function of INR and DC.

OR method, our approach has high accuracy and recall for lower INR cases i.e., INR ≤ -15 dB. Higher recall difference in lower INR cases implies that our approach is better than OR method in detecting more RFI measurements that are in the dataset.

VI. Conclusions and Future Work

In this paper, the RFI detection in microwave radiometry is formulated as a binary hypothesis testing problem. Features that are extracted from the simulated dataset are sequentially reviewed. The posterior probabilities for each measurement are computed using the Bayes rule, likelihood, and prior distributions of the classes. The optimal class label is derived by maximizing the posterior probability for the reviewed features. The Bayesian detection approach uses the information embedded in the features to make evident–based decisions. Even though the proposed method requires prior information, this can be collected using the available radiometers. As a part of our future work, we plan to realize this data collection, as well as increase the number of features and extend this study to detect different types of RFI. We also plan to implement the proposed approach on real radiometer data.

References

[1] M. Aksoy, J. T. Johnson, S. Misra, A. Colliander, andI. O'Dwyer, " L-band Radio-frequency Interference Observations During the SMAP Validation Experiment 2012," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, 3, March 2016, pp. 1323–1335.

[2] J. Querol, A. Perez, and A. Camps, "A Review of RFI Mitigation Techniques in Microwave Radiometry," *Remote Sensing*, vol. 11, 24, December 2019, pp. 3042.

[3] I. M. Nazar, M. Aksoy, "Radio Frequency Interference Detection in Microwave Radiometry Using Support Vector Machines," *Radio Science Letters*, vol. 2, 2020, doi: 10.46620/20-0034.

[4] I. M. Nazar, M. Aksoy, "Radio frequency interference detectionin microwave radiometry using density based spatial clustering,"*XXXIVth URSI General Assembly and Scientific Symposium* , 2021.

[5] K.P. Murphy, "Machine learning: a probabilistic perspective", *MIT press*, 2012.

[6] J. R. Piepmeier, J. T. Johnson, P. N. Mohammed, D. Bradley, C. Ruf, M. Aksoy, R. Garcia, D. Hudson, L. Miles, and M. Wong., "Radio-Frequency Interference Mitigation for the Soil Moisture Active Passive Microwave Radiometer", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 761-775, Jan. 2014, doi: 10.1109/TGRS.2013.2281266.