Salami, Dariush; Ning, Wanru; Ruttik, Kalle; Jantti, Riku; Sigg, Stephan

A Joint Radar and Communication Approach for 5G NR using Reinforcement Learning

Published in:
IEEE Communications Magazine

DOI:
10.1109/MCOM.001.2200474

Published: 01/05/2023

Please cite the original version:
https://doi.org/10.1109/MCOM.001.2200474
A Joint Radar and Communication Approach for 5G NR using Reinforcement Learning

Dariush Salami, Wanru Ning, Kalle Ruttik, Riku Jäntti, and Stephan Sigg

Abstract—Radar operation partly overlaps and thus interferes with 5G spectrum bands. Examples are vehicular radar, long-range air traffic control, terminal air traffic control, marine radar, airport surveillance, or also operations in the mmWave band. We propose a mechanism to efficiently share the spectrum resources for communication and radar operation using a Reinforcement Learning (RL)-based approach. Unlike the state-of-the-art, our approach enables both systems to keep their own waveforms. Compared to the use of a single waveform for joint radar and communication, this results in less complex signal processing and improved sensing resolution. Our approach is compatible with existing radar systems and requires software modification only for the communication system. We demonstrate how both systems can work simultaneously, thereby eliminating the need for time sharing. The effectiveness of the approach is studied through a comprehensive set of experiments implemented in an open source simulation environment. It is shown that in the presence of interference, the radars can still achieve a high accuracy for range and velocity estimation of targets. The system achieves high spectrum utilization and is on-demand adjustable to realize any desired level of trade-off between communication and sensing.

Index Terms—Radio Frequency (RF) communication and sensing convergence, Frequency-Modulated Continuous Wave (FMCW) radar, reinforcement learning, machine learning

I. INTRODUCTION

RECENT advances in RF-sensing have demonstrated that communication systems (e.g. WiFi, cellular, LoRa, Bluetooth, etc.) may not only provide connectivity, but also sensing and environmental perception capabilities [1]. At the same time, researchers have investigated the transmission of communication symbols via radar waveforms [2]. RF convergence has gained attention as a potential solution to better utilize the available spectrum. Transmission of communication symbols via radar waveforms [2] or utilizing communication signals for sensing purposes [3] have been investigated. These designs target architectures where sensing and communication are co-designed at physical and Medium Access Control (MAC) layer. It is claimed that for a chirp signal with infinite slope, the baseband signal converges to that of Orthogonal Frequency-Division Multiplexing (OFDM) signal and hence they have the same performance. However, compared to Continuous Wave (CW) radars, they have higher complexity [4], perform worse in presence of interference [5], and convergence is achieved only in the case of infinite chirp slope [6].

An alternative approach is the co-existence and co-operation of radar and communication systems. Particularly, applications including vehicular scenarios (radar band w 75 GHz to 110 GHz), long-range air traffic control (radar band L 1 GHz to 2 GHz), marine radars (radar band s 2 GHz to 4 GHz), airport surveillance (radar band Ka 27 GHz to 40 GHz), and mmWave band (40 GHz to 300 GHz) are examples of radar systems that interfere with reserved 5G communication bands [7]. No existing communication and sensing system has mechanisms to inter-operate with one another to mitigate the interference problem.

We propose a mechanism to enable cellular networks and radar systems to co-operate so that the interference between the systems is minimized. Specifically, we consider Frequency Shift Keying (FSK) and FMCW radars interfering with 5G New Radio (NR) cellular networks to demonstrate the capability of the proposed Joint Radar and Communication (JRC) system (cf. Fig. 1). We suggest two approaches to implement the JRC system: (1) co-designed sensing and communication where the spectrum is scheduled for simultaneous operations, and (2) co-existing sensing and communication where the systems are independent from each other and the behaviour of either communication or sensing system is predicted by the other one to avoid interference. Our contributions are:

• a communication-aware sensing approach for CW radars to realize JRC with 5G NR
• a radar-enabled cellular system that is able to transmit sensing signals alongside the communication signal
• a study of the trade-off between communication and sensing aspects of the proposed systems

II. RELATED WORK

Device-free gesture and activity recognition, by analysing time or frequency domain patterns of electromagnetic signals achieves high recognition accuracy [8]. In particular, the Channel State Information (CSI), phase and Received Signal Strength (RSS) of multi-antenna wireless interfaces provide information to allow the recognition of motion [9]. Plus, via micro-Doppler variations, whole-body motion can be distinguished. Furthermore, detection of subtle movements, such as respiration and pulse may be obtained using FMCW radars. Distinct target motion is then separated by installing multiple Tx/Rx points.

These results, together with the realization that electromagnetic signals are omnipresent through cellular system deployment, explain the interest in integrating sensing capabilities into communication systems. A good overview on RF communications and sensing convergence research is given in [10]. The authors argue that the wireless medium is limited
Finally, a voting scheme is adopted to converge to a joint decision about the intended classification task. However, simultaneous radar operation and communication is not possible in this scheme since the communication happens after the sensing phase and not at the same time.

In contrast to this previous work, we suggest to aggregate communication symbols tightly with FMCW or FSK radar signals, thus enabling simultaneous sensing and communication without modifying the respective systems waveforms. Hence, spectrum sharing would be maximized by allowing one or more non-synchronized radar-communication system that would otherwise interfere, to simultaneously operate.

III. 5G NR Cellular Communication System

5G NR is an OFDM-based Radio Access Technology (RAT) developed by the 3rd Generation Partnership Project (3GPP) for the fifth generation mobile network. The 5G spectrum consists of sub-6 GHz and mmWave bands that are 25 GHz and above. In OFDM, the number of subcarriers that can be packed into a specific frequency range directly affects spectrum efficiency. In other words, narrower spacing translates to more subcarriers, leading to an increased data rate. Narrow subcarrier spacing results in a longer OFDM symbol length. Consequently, leaving more room for Cyclic Prefix (CP) which in turn results in better resilience to fading [15]. In sub-6 GHz, subcarriers are closely spaced due to the limited bandwidth, while in higher frequency bands the available bandwidth allows OFDM with wider spacing (flexible spacing in 5G NR).

NR features different types of subcarrier spacing (flexible numerology). Five subcarrier spacings are defined in the 5G NR standard: 15 kHz, 30 kHz, 60 kHz, 120 kHz and 240 kHz.

To maintain orthogonality between subcarriers in OFDM, no spacing below 15 kHz is allowed. A wider spacing benefits higher frequency carriers which experience more significant Doppler shift and also, to perform beamforming, the phase of the signal is easier to control for wider subcarrier spacing [15].

While the radio frame structure depends on the numerology type, frame and subframe length are 10 ms and 1 ms, respectively. The number of slots per subframe changes with
In this work, we use a specific family of FSK radars called MFSK which has two frequency tones in the transmitted signal. Although we use FMCW and MFSK radars in our experiments as examples, other types of radars can also be used as long as we train the proposed RL model in an environment in which that radar is also operating. Range resolution in FSK radars is also inversely proportional to the bandwidth $f_{sweep}$.

V. PROPOSED SYSTEM

We propose two approaches for JRC, specifically for co-existing and for co-designed communication and sensing. In the co-existing approach, we assume that the systems operate independently. To realize JRC, the systems must therefore be able to predict the operation of the other to avoid interference. In the second approach, where the systems are co-designed, sensing and communication signals are transmitted using a single infrastructure, thereby eliminating the need for prediction.

A. Co-existing JRC

Assume a setting in which systems operate independently without a feedback channel between them. To minimize interference, they should be able to predict the operation of the other system. A schematic view of this approach is shown in Fig.1.a. Two different scenarios can be regarded:

1) Communication-first co-existing JRC: In communication-first co-existing JRC, a communication system, i.e. 5G NR is co-present. In this case, the radar frontend needs to predict the behavior of the communication system and avoid the channels which are used for communication. The sensing signal shall utilize only those channels which are not simultaneously used for communication. We use an MFSK radar for the experiments, however any type of radar that does not use the whole spectrum at once can be used. Spectrum sensing techniques combined with RL models have been successfully applied to predict the operation of communication systems even under unstable conditions utilizing online-learning. Recent advancements in this cognitive radio field are studied by Gaurav et al. [17].

2) Radar-first co-existing JRC: In radar-first co-existing JRC, the primary application is sensing, while one or multiple radars are operating in the environment. To minimize the interference, the communication infrastructure (5G NR) must predict the pattern of the radar operation, then utilize the unused part of the spectrum to transmit communication signals. For this purpose, given the fact that radar signals have

IV. RF SENSING USING FMCW/MFSK RADARS

CW radars emit a continuous signal to estimate range, velocity, and angle of the target by processing its reflections and Doppler effects. Reflected and transmitted signals are mixed into an Intermediate Frequency (IF) signal and peaks in the Fast Fourier Transform (FFT) domain of the IF signal that are identified by the Constant False Alarm Rate (CFAR) algorithm indicating distance to distinct targets. Phase differences are used to estimate velocity of targets. Although, using antenna configuration any type of radar can estimate Angle of Arrival (AoA), these radars offer robust range, velocity, and angle estimation particularly in the presence of large stationary or slow-moving objects in cluttered environments. CW radars are especially useful to search targets against a background reflector since the strong reflection off the background can be filtered out using Doppler analysis. In this way, the subtle reflection from a target can be isolated.

FMCW radars transmit a sinusoidal signal with increasing frequency (chirp; cf. Fig. 2). The chirp is characterized by bandwidth $B_c$, chirp time $T_c$, and chirp slope $S_c$. By processing the reflections, the radar estimates range, velocity, and angle of a target. The performance of the radar depends on the chirp configuration. Range resolution is defined as the minimum distance between two targets, such that the radar can still resolve them as separate objects. It is proportional to the reciprocal of the bandwidth $B_c$. Although FMCW radars might utilize the available bandwidth during a chirp time, they occupy only a fraction of the bandwidth. Therefore, a considerable part of the spectrum remains unused at any point in time (cf. Fig. 2). In Section V, we explain how the unused part of the spectrum can be used for communication to realize JRC.

As shown in Fig. 2, FSK radars transmit multiple discrete carrier frequencies and estimate the Doppler shift to determine range and velocity of the target. Although there is a problem of range-Doppler coupling [16], they require less bandwidth to sense absolute distance and subtle movement of the target compared to FMCW radars. Similar to FMCW radars, MFSK radars can achieve better range and range resolution by using more bandwidth. Likewise, at each time-instant they use a narrow part of the spectrum and leave the rest unused.

In this work, we use a specific family of FSK radars called MFSK which has two frequency tones in the transmitted signal. Although we use FMCW and MFSK radars in our experiments as examples, other types of radars can also be used as long as we train the proposed RL model in an environment in which that radar is also operating. Range resolution in FSK radars is also inversely proportional to the bandwidth $f_{sweep}$. 

![Fig. 2. Schematic view of the MFSK radar transmitting two frequency tones in each step compared to an FMCW radar transmitting a chirp. Note that the effect of Doppler shift is not shown for simplicity.](image-url)
various parameters to be configured and there can be more than one radar in the environment, we propose an RL-based model to predict the operation of the radars enabling gNb to identify the unused channels in each time-instant that can be used for communication. We assume that the radars are of FMCW type.

a) Problem formulation: We assume a total available bandwidth $B_c$ and $c$ communication channels, each of which has a bandwidth of $C_b$. We define $x^k_t$ as a variable indicating whether the channel $k$ in time-instant $t$ is occupied by the radar signal (1) or not (0). Moreover, we define $a^k_t$ as a variable which should be predicted by the algorithm indicating which communication channels are free to use (1) and which are occupied by a radar (0). Naturally, some channels will be incorrectly predicted as free by the radar (collision) and some channels will be incorrectly predicted as occupied (missed opportunity).

b) RL for radar operation prediction: The next state $(s_{t+1})$ not only depends on the current state $(s_t)$ and the action $(a^k_t)$, but also the previous states. Consequently, the problem does not comply with the Markov process. To make it a Markov process, we apply a window of size $s_w$ indicating the width of the state variable the model should use to predict the next time-instant. Note that the state is a binary matrix of size $s_w \times c$ and the action space is a binary array of length $c$. We choose two on-policy algorithms namely Proximal Policy Optimization Algorithms (PPO) and Trust Region Policy Optimization (TRPO) and an off-policy algorithm e.g. Advantage Actor-Critic (A2C) to estimate the operation of radars. The main issue with policy-gradient methods is that they are unstable in terms of parameter updates since the parameters can drastically change from one update to another. TRPO adds a KL-divergence term to the optimization problem to keep the parameters in a trust region resulting in updates that are not too different from the previous iteration.

In all these methods, we need a reward function that encourages the model to learn an effective policy to estimate the operation of the sensing system. To reduce both missed opportunities and collisions, we consider their $\alpha$ weighted aggregation as the reward function. Reducing $\alpha$ maximizes the utilization of the free channels, while increasing $\alpha$ reduces collisions between communication and sensing system.

B. Co-designed JRC

As shown in Fig.1.b, for co-designed JRC, a single infrastructure including a single antenna array is used to transmit both radar and communication signals. Thus, there is no need for operation prediction module making it less complex from a machine learning perspective compared to the co-existing approach. Moreover, since the sensing and communication functionalities are co-designed, the integrated mechanism is adjustable to different scenarios to reach a trade-off between the systems. For example, if the communication aspects are more important than sensing, most of the channels can be dedicated for communication, or vice versa. Each channel, frame or slot can be assigned to either sensing or communication. Thus, both of the applications operate simultaneously without time-sharing. Depending on the state of the slots and their availability, any type of radars such as FMCW or FSK, can be used in conjunction with a 5G NR cellular system. In Section VI, the performance of the system is analyzed.

VI. RESULTS

First, we discuss the effect of interference imposed by sensing on the communication system. Then we evaluate the performance of the RL models on an environment with a single radar and multiple radars for co-existing scenario where the communication system needs to predict the operation of the radar(s). Finally, the best model is used to do the rest of the experiments on the same environment with different parameters.

A. Implementation details

The JRC environment and the models are implemented using Gym and TensorFlow frameworks, respectively. We use AdamOptimizer to perform gradient decent and the models are trained for 2.5 M timesteps. Additionally, we use a two-layer Multi-Layer Perceptron (MLP) with 64 neurons in each layer as the function approximator. The server used for training and inference is equipped with 64 GB of RAM and a Tesla V100 16 GB GPU. The simulation environment for the sensing and communication systems are implemented in Matlab using Radar Toolbox and 5G Toolbox, respectively.

We analyze the accuracy of sensing in a simulation environment involving a car traveling at 120 km/h at 50 m distance away from the radar, as well as a truck moving towards the radar at 60 km/h and starting from 55 m away. In the simulation environment, the communication users or User Equipments (UEs) are independent from radar targets meaning that the two sensing targets are not the UEs of the communication system.

We assume that FMCW and MFSK radars as well as the communication system operate at 40 GHz which is the same frequency in mmWave range for 5G NR. The radars have 4000 MHz bandwidth and the channel bandwidth for communication system is 50 MHz resulting in 80 channels each of which uses a multi-career system with the flexible spacing mechanism discussed in Section III. Note that according to the 3GPP standard channel-width in mmWave range could be 50 MHz, 100 MHz, 200 MHz and 400 MHz. Although the numerical results presented in this paper are based on 50 MHz...
channel-width, the same developed simulation framework can be used for different bandwidths available for channels. The complete set of the configuration parameters for the simulation purposes is shown in Table I. Moreover, to identify the blocked channels in real-world test-beds, spectrum sensing techniques can be implemented. However, in the simulation environment, having access to the historical data about spectrum usage, we assume there exists a perfect feedback mechanism for this purpose.

B. Effect of interference on communication throughput

Fig. 3a depicts the throughput of a communication system when various channels are affected by the operation of a radar system and conditioned on the signal power of the radar symbols. For weak radar signals, no matter the ratio of the affected channels, the throughput does not decrease. However, the throughput decreases dramatically for radar signal power of $-20$ dBW or higher. The reason is that, orthogonality in OFDM systems is of high importance. As the interferer signal becomes stronger, the drift in the frequency gives rise to non-orthogonal subcarriers resulting in low throughput.

Range and velocity estimation errors for different interferer signal power and interference ratio are shown in Fig. 3b and Fig. 3c, respectively. To calculate the error we consider two ground truth vectors of size three $[55, 50, 0]$ and $[120, -60, 0]$ for range and velocity, respectively based on the experiment setting we described in Section VI-A. Range ground truth vector, contains three elements which are the ground truth distances for the two targets plus one ghost target (a target that does not exist in the environment but is detected by the radar). The same principle goes with the velocity ground truth vector. Since the value for the ghost target is zero for both vectors, in case we have only two targets detected by the radar the last element is ignored, otherwise the range and the velocity of the ghost target will contribute to the error.

We calculate the normalized Root Mean Square Error (RMSE) between the estimated values and the ground truth vector resulting in values between 0 and 1 for minimum and maximum observed estimation error. Moreover, the different number of detected targets is shown in Fig. 3d. For both velocity and range, when the interferer signal power is low or the interference ratio is low, a perfect estimation is achieved. With increasing impact on the communication signal power or interference ratio, the accuracy of sensing decreases. For extreme conditions, the radar fails to identify even the number of the targets. There are few cases where a ghost target appears in the sensing as shown in Fig. 3d as a peak appeared in the middle of the figure. The reason is that, under heavy interference, CFAR algorithm fails to correctly identify the peaks leading to ghost targets.

C. Radar operation prediction results

In this section, we evaluate the performance of the proposed RL-based system to estimate the operation of the radar(s). We use FMCW radars for experiments, but the same simulation framework and models can be applied to other types of radars as well. For the evaluation purposes, we define environments with one and two radars (operating in the same frequency with different chirp slopes).

1) Model comparison for radar operation estimation: As it is shown in Fig. 4a, TRPO outperforms PPO and A2C when the models are trained for 2.5 M timesteps. The average reward for a random policy where each channel is used for communication with a probability of 50% is $-31800$, which suggests that the models manage to learn a meaningful strategy. However, TRPO converges to a higher reward, while keeping a stable rewards trend. It also outperforms the other models in a two-radar environment.

2) Communication-sensing trade-off: $\alpha$ controls the accuracy of sensing and communication. For $\alpha = 0$, the collision penalty is removed from the reward function, favoring communication. For $\alpha = 1$, the model avoids collisions at the cost of reduced data-rate and spectrum utilization. Fig. 4b shows the average reward for TRPO with different values for $\alpha$. For $\alpha$ equal to 0 and 1, the rewards are close to zero since the reward function degrades to a single term (collision penalty or missed opportunity) which is easy to optimize. In other cases, the best performance is achieved with $\alpha = 0.8$.

We quantify the normalized number of channels in different cases: clean channels for sensing, clean channels for communication, collision channels, and unused channels. In Fig. 5, at each time step, a small number of channels are designated to the radars since they use a small portion of the spectrum. So, the number of channels are in a different scale for communication and sensing systems. To have a better comparison, we divide the number of channels used for solely
communication and sensing (clean channels) to the maximum possible number of channels for each purpose. The same goes with unused and collision channels.

For the case of single radar, as $\alpha$ increases to 1, the number of clean channels for radar and the number of unused channels increase while the number of clean channels for communication and the number of collision channels decrease. Specifically, the model avoids collision at the cost of having a considerable part of the spectrum unused. However, for lower values of $\alpha$, the model increases the number of channels used for communication at the cost of having many collisions with the radar system. For $\alpha = 0.8$, the model can efficiently reduce both collision channels and unused channels while retaining a considerable number of collision-free channels for sensing and communication. As shown in the bottom chart of Fig. 5, a similar trend holds for a two-radar environment. Note that for $\alpha = 0$, we have maximum possible collision meaning that all the radar channels are contamined by the communication system. However, the communication system still uses the maximum possible capacity for transmitting communication signal resulting in using all the channels that are not being used by the radar. That is why the normalized value for both collision channels and clean communication channels is 1.

3) State-width: As discussed in Section V-A2b, to make the problem a Markov process we apply a windowing technique. The window-length is called state-width. The effect of window-length on the average reward is shown in Fig. 4.c. For window-length of 1, the reward is low suggesting that it is difficult to predict the operation of the radar using information from a single timestep. However, for larger windows, the average reward increases showing that the longer the window-length, the easier the radar operation prediction. The same trend applies for a two-radar environment.

4) Generalizability: A model that is trained on a single-radar environment, fails to predict the operation of the second radar in a two-radar environment. Furthermore, a model trained on a two-radar environment fails to efficiently utilize the spectrum for a single-radar system meaning that a considerable part of the spectrum remains unused. Thus, the number of radars in the training phase should match that of the inference phase to have an efficient spectrum utilization. However, in applications where communication throughput or sensing accuracy can be compromised, the number of radars in train and inference phases can be different.

5) Effect of bandwidth on radars: Another important aspect that affects the performance of the radars is the availability of bandwidth. Range resolution in both radar systems, namely FMCW and FSK, is inversely proportional to their operational bandwidth (cf. Section IV). For outdoor scenarios, for instance in automotive settings, a range resolution of 60 cm is achievable with only 250 MHz bandwidth. However, in a pulse radar, to achieve the same range resolution, the pulse time needs to be 4 ns resulting in a considerable technical complexity.

VII. CONCLUSION

We proposed two mechanisms to use the RF spectrum for communication and sensing simultaneously. In a simulation framework, we showed that it is possible to efficiently utilize, for a radar sensing system, the unused part of the spectrum for communication purposes. We further proposed a flexible reward function for RL algorithms that can be adapted to different situations where a specific degree of reliability is needed to be guaranteed for either sensing or communication system. We also demonstrated how a radar sensing mechanism can be used to perform sensing in 5G NR cellular networks meaning that the same infrastructure is used for transmitting...
both sensing and communication signals. Moreover, we evaluated throughout the communication system as well as range and velocity estimation accuracy of the radar in the presence of interference caused by one another. The results showed that CW radars are able to operate with a high accuracy in the presence of interference imposed by a communication system.

**ACKNOWLEDGMENT**

Part of the calculations presented above were performed using computer resources within the Aalto University School of Science “Science-IT” project.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant agreement No. 813999, HPY Research Foundation and Nokia Foundation.

**REFERENCES**


**Dariush Salami** received his BSc and MSc degrees from Shahid Beheshti University and Amirkabir University of Technology in Software Engineering in 2016 and 2019, respectively. He is a former Marie Skłodowska Curie fellow in ITN-WindMill project and a PhD researcher at the department of communications and networking at Aalto University. He is now a radio research scientist in Nokia Bell Labs focused on Machine Learning for Wireless Communications.

**Wanru Ning** received the B.E. degree in electronic information engineering from Hebei Beihesi University, Baoding, China, in 2017. She is currently pursuing the Ph.D. degree in circuits and systems with the School of Microelectronics, Tianjin University, Tianjin, China. Her current research interests include radio frequency identification (RFID) and indoor localization.

**Kalle Ruttik** is currently a teaching researcher in the Department of Communications and Networking at the School of Electrical Engineering, Aalto University. He received a Dipl.Eng. degree from Tallinn Technical University in 1993, and a Lic. Tech. degree from TKK in 1999. His research interests include radio physical layer algorithms, software radio, base station implementation, NB-IoT, and cloud-RAN.

**Riku Jantti** is an associate professor (tenured) in communications engineering and the head of the Department of Communications and Networking at the Aalto University School of Electrical Engineering. He received his M.Sc (with distinction) in electrical engineering in 1997 and D.Sc (with distinction) in automation and systems technology in 2001, both from Helsinki University of Technology (TKK). Prior to joining Aalto (formerly TKK) in August 2006, he was professor pro tem in the Department of Computer Science, University of Vaasa. He is an Associate Editor of IEEE Transactions on Vehicular Technology. He is also an IEEE Vehicular Technology Society Distinguished Lecturer (Class of 2016). His research interests include radio resource control and optimization for MTC, cloud-RANs, spectrum and coexistence management, and RF Inference.

**Stephan Sigg** received his M.Sc. degree in computer science from TU Dortmund, in 2004 and his Ph.D. degree from Kassel University, in 2008. Since 2015 he is an assistant professor at Aalto University, Finland. He is a member of the editorial board of the Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies as well as of the Elsevier journal of Computer Communications. He has served as a TPC member of renowned conferences including IEEE PerCom, IEEE ICDCS, etc. His research interests include Ambient Intelligence, in particular, Pervasive sensing, activity recognition, usable security algorithms for mobile distributed systems.