Adversarial Jamming for a More Effective Constellation Attack

Haidong Xie
Qian Xuesen Laboratory
China Academy of Space Technology

Yizhou Xu
School of Aerospace Engineering
Tsinghua University

Yuanqing Chen
Qian Xuesen Laboratory
China Academy of Space Technology

Nan Ji
Qian Xuesen Laboratory
China Academy of Space Technology

Shuai Yuan
Qian Xuesen Laboratory
China Academy of Space Technology

Naijin Liu*
Qian Xuesen Laboratory
China Academy of Space Technology

Xueshuang Xiang*
Qian Xuesen Laboratory
China Academy of Space Technology

Abstract—The common jamming mode in wireless communication is band barrage jamming, which is controllable and difficult to resist. Although this method is simple to implement, it is obviously not the best jamming waveform. Therefore, based on the idea of adversarial examples, we propose the adversarial jamming waveform, which can independently optimize and find the best jamming waveform. We attack QAM with adversarial jamming and find that the optimal jamming waveform is equivalent to the amplitude and phase between the nearest constellation points. Furthermore, by verifying the jamming performance on a hardware platform, it is shown that our method significantly improves the bit error rate compared to other methods.

I. INTRODUCTION

With the development of wireless communication technology, the corresponding jamming technology has become a popular research topic [1], [2]. Wireless signals are allocated with different frequency bands, and different frequencies do not interfere with each other, which makes it possible to barrage regional targets. For example, near important meeting locations or college entrance examination halls, jammers can be used to barrage mobile phone frequencies and prevent illegal communication.

For a target frequency band, the most commonly used method is noise barrage. It is characterized by adding noise to a signal, and the amplitude distribution obeys a Gaussian distribution; this noise is commonly known as additive white Gaussian noise (AWGN) [3], [4]. This noise can mask the information contained in a signal. The larger the amount of noise is, the smaller the signal-to-noise ratio (SNR) or signal-to-jamming ratio (SJR) are and the worse the bit error rate (BER) or symbol error rate (SER) is, which causes poor communication performance. When nothing is known about the transmitted signal characteristics, adding noise is almost the only choice. However, intuitively, AWGN is obviously not the best jamming waveform. Therefore, we are concerned about finding and constructing a waveform with the best barrage effect based on some signal feature knowledge.

An adversarial example (AE) is a new concept recently proposed in the field of deep learning that aims to induce a model to output incorrect results that are similar to the real data [5], [6]. As deep learning technology has been successfully applied to the field of wireless communication [7], we believe that the signal receiving and demodulating process can be realized by a deep learning model, and the construction of a jamming waveform can correspond to the generation of AEs in deep learning.

Thus, we apply the idea of AEs to the jamming problem and propose the adversarial jamming (AJ) waveform. Our method first transforms communication signal demodulation and jamming into a deep learning model and then generates the corresponding AEs to obtain the jamming waveform, which can attack a demodulation model with minimum power. In terms of effect, our method can optimize and find the best jamming waveform amplitude and phase, which can ensure that the information contained in the interfered signal is completely covered up so that the communication party cannot accurately obtain effective information. When the constrained jamming power is low, our method can concentrate the jamming energy in a few waveforms to improve the jamming effect.

A numerical simulation and hardware verification show that our method has significant advantages over noise barrage. The variation curve of SER with SJR shows that our method can achieve the best jamming attack under any condition. The hardware test shows that our method is fully practical, and the jamming effect will be further improved considering factors such as clock synchronization.

II. ADVERSARIAL JAMMING

To use AE technology, we first need to transform the signal demodulation process into a deep learning process [8]. For-
The waveform to the coordinate points in the constellation. Therefore, finding the best jamming waveform essentially becomes the geometric problem of finding the nearest constellation point. After the numerical experiments, we found that the jamming waveform numerically solved by the AE method is often the displacement vector corresponding to the nearest constellation point, which is the minimum necessary for jamming with only a difference in numerical precision, as shown in Fig. 1(c). If the desired SJR is low, i.e., the allowable jamming is stronger than the generated AEs, there is no need to change or enhance the perturbation amplitude to ensure the jamming effect. On the other hand, if the allowable average jamming power is weaker than AEs, we cannot simply reduce the size of the perturbation because it will invalidate the jamming outcome. Here, our strategy is to use time intermittent jamming to ensure that a single jamming waveform still has enough perturbation, but the occurrence time of jamming decreases accordingly to meet the target average jamming power, \( t^{%} \leftarrow P_{jamming}/P_{Signal} \). Finally, we summarize the AJ algorithm in Alg. 1.

III. EXPERIMENTAL RESULT

A. Jamming Analysis

We take 16QAM as an example to show the impact of multiple jamming on SER under different SJRs, as shown in Fig. 2(a). Each data averages a sequence of 500000 bits. On the whole, AJ shows the best jamming effect due to its waveform jamming ability with the optimal amplitude, phase, and time.

In traditional noise jamming, the SER curve shows a gentle downward trend as a whole. Hence, if the direction idea in AJ is introduced, that is, the phase of the jamming waveform is fixed according to the constellation diagram, but the amplitude is still arranged according to a Gaussian distribution, then the phase jamming curve can be obtained. Compared with that of a noise attack, the overall curve of a phase attack is slightly improved.

In contrast, if we keep the phase uniform distribution unchanged and use a jamming waveform with a fixed amplitude instead of a Gaussian distribution, we can obtain fixed power jamming. Obviously, the curve of this jamming is no longer a
the jamming is strong, synchronization becomes relatively
time of each clock cycle of a waveform. Therefore, when
synchronization because a receiver cannot obtain the starting
results. The most significant difference between hardware
to two jamming waveform are found during the optimization process. The best waveform jamming is realized, and the outstanding performance of the method is tested on numerical and hardware platforms. Our method has practical value and can realize deception attacks. Finally, the purpose of this work is to jam an illegal spectrum, but there is the possibility of abuse, which requires the academic community to jointly improve the legal system.

IV. CONCLUSION AND DISCUSSION

Aiming at the problem of optimal jamming, an AJ waveform is proposed in this paper. The method is based on the idea of AEs, and the optimal amplitude, phase, and time of the jamming waveform are found during the optimization process. The best waveform jamming is realized, and the outstanding performance of the method is tested on numerical and hardware platforms. Our method has practical value and can realize deception attacks. Finally, the purpose of this work is to jam an illegal spectrum, but there is the possibility of abuse, which requires the academic community to jointly improve the legal system.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (Grant No. 12004422) and by Beijing Nova Program of Science and Technology (Grant No. Z191100001119129).

REFERENCES

[1] W. Xu, K. Ma, W. Trappe, and Y. Zhang, “Jamming sensor networks: attack and defense strategies,” IEEE Network, vol. 20, no. 3, pp. 41–47, 2006.
[2] Y.-S. Shiu, S. Y. Chang, H.-C. Wu, S. C.-H. Huang, and H.-H. Chen, “Physical layer security in wireless networks: a tutorial,” IEEE Wireless Communications, vol. 18, no. 2, pp. 66–74, 2011.
[3] D. Pauluzzi and N. Beaulieu, “A comparison of snr estimation techniques for the awgn channel,” IEEE Transactions on Communications, vol. 48, no. 10, pp. 1681–1691, 2000.
[4] K. Baddour and N. Beaulieu, “Autoregressive modeling for fading channel simulation,” IEEE Transactions on Wireless Communications, vol. 4, no. 4, pp. 1650–1662, 2005.
[5] C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, and R. Fergus, “Intriguing properties of neural networks,” in ICLR, 2014.
[6] I. J. Goodfellow, J. Shlens, and C. Szegedy, “Explaining and harnessing adversarial examples,” in ICLR, 2015.
[7] J. Jagannath, N. Polosky, A. Jagannath, F. Restuccia, and T. Melodia, “Machine learning for wireless communications in the internet of things: A comprehensive survey,” Ad Hoc Networks, vol. 93, p. 101913, 2020.
[8] C. de Vrieze, S. Barratt, D. Tsai, and A. Sahai, “Cooperative multi-agent reinforcement learning for low-level wireless communication,” arXiv:1801.04541, 2018.
[9] A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu, “Towards deep learning models robust to adversarial attacks,” arXiv:1706.06083, 2017.
[10] N. Carlini and D. Wagner, “Towards evaluating the robustness of neural networks,” in 2017 IEEE Symposium on Security and Privacy (SP), pp. 39–57, May 2017.
[11] “USRP X310.” https://kb.ettus.com/X300/X310.
[12] “UHD.” https://kb.ettus.com/UHD.