SemProtector: A Unified Framework for Semantic Protection in Deep Learning-based Semantic Communication Systems

Xinghan Liu, Guoshun Nan, Qimei Cui, Zeju Li, Peiyuan Liu, Zebin Xing, Hanqing Mu, Xiaofeng Tao, Tony Q.S. Quek, *Fellow, IEEE*

Abstract—Recently proliferated semantic communications (SC) aim at effectively transmitting the semantics conveyed by the source and accurately interpreting the meaning at the destination. While such a paradigm holds the promise of making wireless communications more intelligent, it also suffers from severe semantic security issues, such as eavesdropping, privacy leaking, and spoofing, due to the open nature of wireless channels and the fragility of neural modules. Previous works focus more on the robustness of SC via offline adversarial training of the whole system, while online semantic protection, a more practical setting in the real world, is still largely under-explored. To this end, we present SemProtector, a unified framework that aims to secure an online SC system with three hot-pluggable semantic protection modules. Specifically, these protection modules are able to encrypt semantics to be transmitted by an encryption method, mitigate privacy risks from wireless channels by a perturbation mechanism, and calibrate distorted semantics at the destination by a semantic signature generation method. Our framework enables an existing online SC system to dynamically assemble the above three pluggable modules to meet customized semantic protection requirements, facilitating the practical deployment in real-world SC systems. Experiments on two public datasets show the effectiveness of our proposed SemProtector, offering some insights of how we reach the goal of secrecy, privacy and integrity of an SC system. Finally, we discuss some future directions for the semantic protection.

Index Terms—Semantic communications, Semantic Protection, Inference Attacks, Privacy Leaking, Adversarial Attacks.

I. INTRODUCTION

Existing wireless communication systems, which are built on the classical Shannon information theory, primarily focused on how to accurately and effectively transmit raw bits over channels from a source to a destination. In the past decades, the achieved transmission rate has been improved tens of thousands of times to meet the increasing traffic demands, gradually driving a wireless system capacity to the Shannon limit. Meanwhile, various newly emerged mobile applications, such as virtual reality (VR) and Human-to-Machine (H2M) communications, tend to require more intelligence among different parties, as well as interpretations of the received information to make smart decisions. This motivates us to rethink a more intelligent and a more efficient wireless communication paradigm beyond the Shannon approach.

Recently proliferated semantic communications (SC) systems [1], [3], [5], [13] have shown great potential in transmitting the semantics conveyed by the source and accurately interpreting the meaning at the destination. Such a novel communication paradigm greatly facilitates the aforementioned mobile applications such as VR and H2M. Specifically, the transmitter of an SC system relies on neural networks to extract semantic information from an input and then sends compressed semantic symbols to noised wireless channels. At the receiver side, the system reconstructs the received semantics and interprets the semantics to make intelligent decisions.

Although promising, SC systems may suffer from more challenging security issues compared to traditional wireless communications, such as eavesdropping, tampering, and spoofing due to the broadcast nature of wireless communications and the fragility of deep neural networks [2], [4]. First, anyone within the physical communication range of a transmitter can receive the wireless signal and potentially decode the symbols. As an SC system primarily focuses on the transmission of the meaning conveyed at source rather than accurate raw bits, it facilitates eavesdroppers to carry out malicious attempts to semantics, such as inference attacks and adversarial attacks [8]. Second, as an SC system is learned from a large volume of data examples, which may consist of sensitive information from the personal knowledge base, such as the user's health status and service access history, the system may unintentionally reveal sensitive information in semantic representations to be sent to wireless channels, potentially leading to privacy leaking of the input semantics [6]. Third, an SC system is vulnerable to adversaries due to the blind spots of neural models. The adversarial perturbations are able to mislead the system to make incorrect semantic interpretations.

Existing works focus more on the robustness of SC via offline adversarial training of the whole system, mitigating the blind spots of the neural models with adversary examples [14], [15]. These studies are able to harden the system against adversarial attacks. However, they may not be practically deployed in real-world scenarios as it is infeasible to improve the robustness of neural models by interrupting a communication system for re-training. Furthermore, these studies mainly aim

This work was supported by National Key R&D Program of China (Grant No.2022YFB2902200). Xinghan Liu, Guoshun Nan, Zeju Li, Qimei Cui, Peiyuan Liu, Zebin Xing, Hanqing Mu and Xiaofeng Tao are with National Engineering Research Center for Mobile Network Technologies, Beijing University of Posts and Telecommunications, Beijing 100876, China. (*Corresponding author: Guoshun Nan.*)

T. Q. S. Quek is with the Singapore University of Technology and Design, Singapore 487372, and also with the Yonsei Frontier Lab, Yonsei University, South Korea.

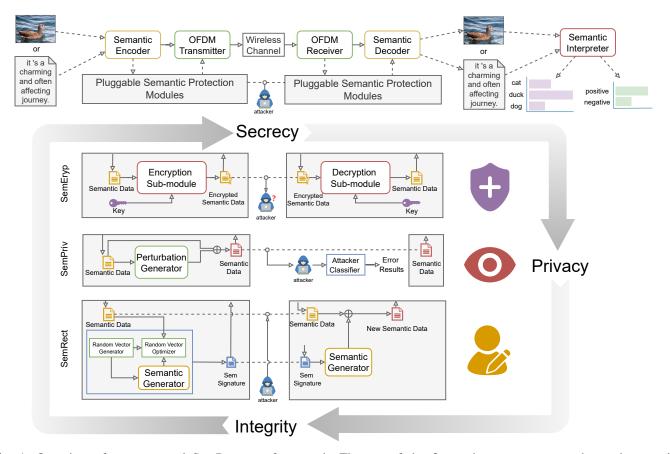


Fig. 1: Overview of our proposed SemProtector framework. The top of the figure demonstrates an end-to-end semantic communication system, equipped with our hot-pluggable semantic protection modules residing in the transmitter and receiver. At the bottom of the figure are hot-pluggable modules, including SemEryp, SemPriv, and SemRect, with the goal of secrecy, privacy, and integrity, respectively. The three modules can be flexibly assembled based on the customized protection requirements.

to defend against a single attack and hence are unable to protect an SC system against various threats. Therefore, it is still unclear how an online SC system can be simultaneously protected against various malicious attempts such as eavesdropping, tampering and spoofing.

To this end, we present SemProtector, a unified framework that aims to secure the semantic transmission and interpretation of SC systems with three hot-pluggable semantic protection modules, including SemEryp, SemPriv and Sem-Rect. Specifically, SemEryp is able to encrypt semantics to be transmitted by an encryption at the transmitter, securing the semantics over an open wireless channel, and SemPriv aims to mitigate privacy leaking via a perturbation generator that can craft adversaries to distort the malicious actors. SemRect is capable of ensuring the semantic interpretation by generating a signature at the transmitter and then calibrating the distorted semantics via the signature at the receiver. We conduct experiments on two benchmarks.

In conclusion, our main contributions can be summarized in three-fold. First, we propose a unified framework to secure the online deep learning-based semantic communication systems with the goal of secrecy, privacy and integrity. Second, we present three hot-pluggable semantic protection modules under the unified framework. Our framework enables an existing online SC system to dynamically assemble the three modules to meet customized semantic protection requirements, facilitating the practical deployment in real-world SC systems. Third, we show the superiority of our method in securing semantic communications against various threats, giving some helpful insights for online semantic protections. We also discuss some future directions in the field.

II. END-TO-END SEMANTIC COMMUNICATION SYSTEM

A. Overview

In this section, we describe the end-to-end semantic communication system used in our framework and outline each component. We refer to the previous work JSCC-OFDM [11] and DeepSC [5] to implement our SC system. Our system involves a semantic encoder, an Orthogonal Frequency Division Multiplexing (OFDM) transmitter, an OFDM receiver, and a semantic decoder. We additionally introduce a semantic interpretation module based on Mobilenetv2 [12]. We formulate our objective as the weighted sum of the data reconstruction loss and semantic interpretation loss, training the system under a multi-task learning paradigm. The top of Fig. 1 demonstrates the proposed end-to-end SC system.

B. Semantic Transmitter

Semantic Encoder: We feed the image or text to the semantic encoder to obtain the semantic representations. We refer to JSCC-OFDM/DeepSC to jointly consider the source encoding and channel encoding in our semantic encoder.

OFDM Transmitter: To make efficient use of the spectrum and reduce the computational overhead, we employ OFDM as our wireless transmission scheme, which is able to encode the semantic representations into multiple carrier frequencies without inter-symbol interference (ISI). It should be noted that the transmitter also supports the single-carrier OFDM mode with configurable parameters.

C. Wireless Channel

As discussed in end-to-end communications, deep neural networks are able to model the wireless physical channels, including additive white Gaussian noise (AWGN), the erasure channel, and the Rayleigh fading channel. In this article, we mainly consider the Rayleigh fading channel as it can better model the effect of a propagation environment for semantic communication systems.

D. Semantic Receiver

OFDM Receiver: The OFDM receiver takes received symbols as the input and performs inverse operations that have been done in the OFDM transmitter.

Semantic Decoder: We feed semantic representations to the semantic decoder to reconstruct the input data, which jointly considers the channel and source decoding.

Semantic Interpreter: This module takes the output of the semantic decoder as an input, and interprets semantics to make the corresponding decisions.

III. SEMPROTECTOR FRAMEWORK

A. Overview

In this section, we introduce our proposed SemProtector framework, which is illustrated in Fig.1. Our framework consists of three key modules including SemEryp, SemPriv, and SemRect. The goal of these three modules is to obtain the secrecy, privacy, and integrity of our SC system, respectively. We first outline the three modules as follows.

1) SemEryp: The module aims to encrypt the semantic representations at the sender side and decrypt the information at the receiver side. Equipped with our SemEryp, an illegal user who can receive the wireless signal is unable to reconstruct the semantic data and hence confidentiality of semantics can be properly protected, while legitimate users can normally decode the signal.

2) SemPriv: This module takes the semantic representation as an input and then generates tiny perturbations that can mislead an eavesdropper to make an incorrect semantic interpretation. Meanwhile, such an operation will be transparent to legitimate users for interpretations. Hence the privacy leaking risks over the wireless channels can be properly mitigated. 3) SemRect: The goal of this module is to ensure the accurate semantic interpretation of an SC system under the destructive physical layer adversarial attacks. It generates a high-dimension random vector based on semantic representation at the transmitter to calibrate the semantic data against adversarial attacks at the receiver. We term such a vector as a semantic signature. By doing so, we are able to protect the semantic integrity of the source to accurately understand the meaning for making decisions.

We train these modules independently and then plug them into the above end-to-end SC system for semantic protection. Our framework allows an existing online SC system to flexibly assemble the three modules to meet customized semantic protection requirements, greatly facilitating practical deployment in real-world SC systems. Fig.2 shows the training process of the three modules, and we detail them in the following parts.

B. SemEryp

The primary goal of SemEryp is to protect the confidentiality of the semantics underlying the input data. A malicious user can potentially reconstruct the input data based on semantics to carry out model inversion attacks [7], [9], in case the eavesdropper acquires the knowledge of semantic decoder or trains a surrogate semantic decoder. This type of attack may lead to privacy leaking issues. We present a semantic encryption module to protect semantics over the air in our SemEryp.

Our SemEryp is composed of two components, i.e., semantic encryption (SEC) and semantic decryption (SDC). SEC and SDC share the same neural structure, and both of them use multiple convolutional layers for encryption and decryption. As shown in the top of Fig 2 colored in yellow, we feed the semantic representation of an input data and a semantic key to output encrypted semantic data, where the key is generated by a negotiation between the transmitter and the receiver. SDC takes the encrypted data and the semantic key as inputs to obtain the original semantics.

We demonstrate how we build our SemEryp with an adversarial training method. Specifically, we first introduce the semantic reconstruction component (SRC), which simply consists of multiple convolutional layers. We train SRC to recover the encrypted semantic data without the semantic key. Such a module can be regarded as a malicious user. Then we train SEC and SDC to defend against the attacks generated by SRC. We repeat the above procedure to achieve better confidentiality protections.

C. SemPriv

The primary goal of SemPiv is to protect the semantic privacy of inputs over an end-to-end SC system. We adapt Adversarial Transformation Networks (ATN) to our semantic communication system to defend against attribute inference attacks (AIA), a destructive privacy attack that is able to infer hidden information of the semantics. An eavesdropper within the physical communication range of the transmitter can receive the wireless signal and potentially decode the semantics, leading to the privacy leakage of legitimate users.

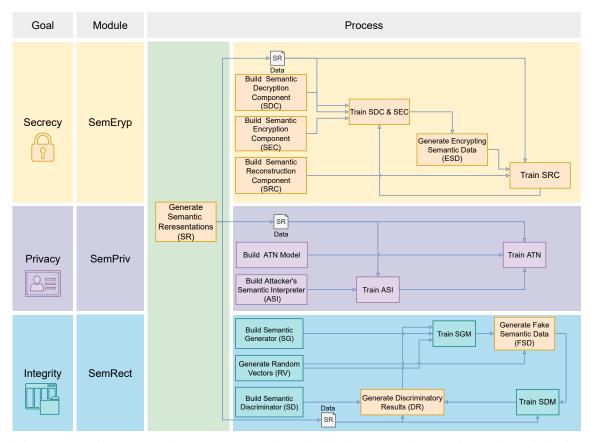


Fig. 2: Training process of the proposed SemProtector, equipped with three semantic protection modules, including SemEryp, SemPriv, and SemRect.

We demonstrate how we train our SemPriv as follows, as shown in the middle of Fig.2 colored in purple. Specifically, we introduce an attacker's semantic interpreter (ASI) module. The module is a multilayer perceptron (MLP) and can be regarded as an eavesdropper. Then we feed the semantic representations to ATN to learn to craft the adversarial privacy perturbations by maximizing the cross-entropy loss of the ASI, so that the perturbation can mislead the eavesdropper to make an incorrect prediction. However, such a privacy perturbation will not affect the semantic interpretation of legitimate users, and hence the privacy leaking issue can be properly mitigated.

D. SemRect

Our SemRect aims to protect the integrity of semantics for accurate semantic interpretation at the receiver side. Here, an attacker relies on FGSM [10] to generate an adversarial perturbation against the semantic information to mislead the interpretations, where the perturbation will be added to the semantic data. Meanwhile, we also produce a random vector (RV), which is termed as a semantic signature, to calibrate such a perturbation, so that the impact of the attacks can be properly mitigated for legitimate users.

The detailed training process of SemRect is described as follows. We introduce a semantic generator (SG), which is adapted from the previous Defense-GAN. First, we train generative adversarial networks (GAN) to produce semantic representations based on a random vector. Such a process is illustrated at the bottom of Fig.2 colored in blue. We deploy the generator RV to both the sender and receiver. At the sender, RV learns to generate and then optimize semantic signatures. At the receiver, RV learns to calibrate the semantics with the signature, subtracting the adversarial perturbations for accurate semantic interpretations. By doing so, our SemRect is able to protect the integrity of semantics against adversarial perturbations over the end-to-end SC system.

E. Summary

In the above SemEryp, SemPriv and SemRect, we only select some attacks to demonstrate how each single module defends against a certain type of threat. By dynamically assembling these hot-pluggable modules, our SemProtector framework is able to defend against various attacks and meet the customized protection requirements, such as secrecy, privacy and integrity. Our unified framework is flexible for a further extension and can be generalized to defend against other attacks, facilitating semantic protection in a real-world SC environment. Due to the space limitation, we will not give more discussions for such an extension.

IV. EXPERIMENTS

We conducted experiments on three popular datasets, MNIST, CIFAR10 and SST-2. The first two datasets are used for image classification, and the last one is used for text classification. We trained the SC system and the pluggable semantic protection modules on these three datasets and performed attacks on the SC system before and after adding the pluggable semantic protection modules to evaluate the effectiveness of SemProtector against various attacks. Although we employ the Rayleigh fading channel in our SC system, our SemProtector can also be applied in other wireless environments, and we will not give the details due to the space limitation. We then detail the attack methods and evaluation metrics, and give some insightful conclusions based on our observations.

A. Attack Methods

To evaluate the effectiveness of our framework, we reimplement three attack methods and outline them as follows.

- Model Inversion Attacks: In this attack, the attacker needs to intercept the semantic data and convert it to the original information. There are two ways for the attacker to get the original information by semantic data, one is to train the surrogate model and the other is to steal the semantic decoder of the receiver. Here we re-implement the second one, as it is more destructive.
- Attribute Inference Attacks: For this attack, the goal of the attacker is to infer privacy information from the semantic data. Specifically, the attacker trains a semantic classifier that is targeted to interpret semantic data. In our implementation, the classifier trained by the attacker is an MLP. We use the accuracy of the attacker's classification of semantic data to evaluate the protection effectiveness of the attack.
- Adversarial Attacks: The goal of an adversarial attacker is to mislead the receiver by adversarial perturbations. Here we integrate FGSM to our SemRect module. We train the FGSM with a substitute model to mimic the SC system. The FGSM attacker is able to generate perturbations under a black-box setting and hence can be directly used in an online deep learning-based semantic communication system.

B. Evaluation Metrics

We use Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) to measure the impact of the optional pluggable semantic protection modules on the SC system.

- **PSNR** is a measure of the quality of an image or video. It is typically used to assess the fidelity of a compressed image or video file. The higher the PSNR, the better the quality of the image or video.
- SSIM is used to show how closely two images resemble one another. In contrast to PSNR, SSIM is more in accordance with how the human eye naturally perceives an image.

C. Implementation Details

We implement our SemProtector based on Pytorch, one of the most popular machine-learning libraries. We adapt JSCC-OFDM [11] and DeepSC [5] to our end-to-end SC system. Considering the computing resources and energy efficiency of

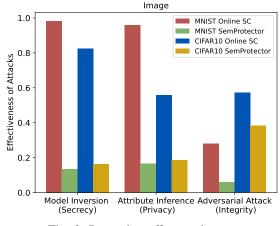


Fig. 3: Protection effect on image.

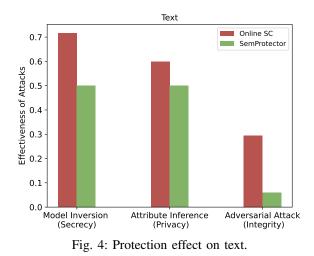
mobile devices, we use Mobilenetv2 [12] and TextCNN as the semantic interpreter of the SC system. For SemEryp, we use a simple neural network with five convolutional layers, which can encrypt and decrypt the semantic data very well at the same time taking into account the efficiency. For SemPriv, we use the modified ATN, which aims to craft semantic perturbations based on semantic data with six linear layers. For SemRect, we use the modified Defense-GAN, which can produce semantic signatures based on semantic data. These three hot-pluggable protection modules can be flexibly assembled based on customized protection requirements.

TABLE I: Parameters in experiments.

Category	Parameter	Value
	Shape(MNIST, CIFAR10,SST-2)	(1,28,27),(1,32,32),(1,64)
	Num of class type (MNIST, CIFAR10,SST-2)	10,10,2
Input	Num of train/test(MNIST)	60,000/10,000
	Num of train/test (CIFAR10)	50,000/10,000
	Num of train/test (SST-2)	67.350/1,821
ESC system	Num of residual block(image)	4
	Num of transformer block(text)	2
OFDM Module	Num of pilot symbols	1
	Num of subcarriers per symbol	64
	Length of Cyclic Prefix	16
	Exponential decay	4
	SNR (dB)	10
Classifier	Types of Classifiers	Mobilenetv2, TextCNN
SE Module	Num of Conv layers of the Encryption Module	4
	Num of Conv layers of the Decryption Module	4
	Num of Conv layers of the Reduction Module	4
PAJ Module	Num of linear layers of ATN	6
SR Module	Num of ConvTranspose layers of SemGenerators	3
	Num of Conv layers of SemDiscriminator	4
	The dimension of the random vector	100
Training	Batch size	256
	Learning rate	1.00E-04
	Optimizer	Adam

D. Experimental Settings

We conduct experiments on a server with Ubuntu 18.04 operating system, equipped with 90GB RAM and an RTX3090 GPU card. The Pytorch and Python versions are 1.7.0 and 3.7, respectively. For drivers, the CUDA and cuDNN versions installed in the server are 11.0 and 8004, respectively. The parameters of neural networks are randomly initialized and then iteratively updated by the Adam optimizer. Table I shows the detailed parameters for models used in SemProtector.



E. Main Results

Fig.3 and Fig.4 demonstrate the effectiveness of our Sem-Protector on both image and text datasets, respectively. We assemble the three modules, i.e., SemErvp, SemPriv, and SemRect, to harden the online SC system against various attacks. The results show that the proposed SemProtector can significantly reduce the injuries under various destructive threats on both text and image modalities. Fig.3 shows that an eavesdropper can use the model inversion attacks to successfully infer the semantic data of an MNIST image, achieving as high as 98.3% accuracy based on its semantic symbols over the air, while the attacker can only obtain a 13.4% accuracy for the online SC system under the protection of our SemProtector. Fig. 4 shows that our framework can effectively protect the text-based SC system against various attacks, e.g., an adversarial attacker can only achieve a 6% successful rate equipped with our SemProtector. These results confirm the effectiveness of our framework in protecting semantic confidentiality. We also observe that the protections on MNIST are better than the ones on CIFAR10. One possible reason is that the semantics in the images of CIFAR10 are much more complex than the ones in MNIST.

We also quantify the side effects of our SemProtector to the image reconstruction for legitimate users. Table II reports the performance of our SemProtector under various settings. For example, "Online SC + SemPriv" indicates that the SemPriv module is plugged into our online SC system, and "Online SC + SemEryp + SemPriv + SemRect" represents that we assemble all three modules to harden the communication system. The results show that our SemProtector can lead to slight performance decreases in terms of SSIM and PSNR on the two datasets. It is reasonable that there will be a larger performance decrease with more protection modules involved. For customized protection requirements, one may need a tradeoff between the security level of semantic secrecy, privacy, integrity, and transmission quality.

V. FUTURE DIRECTIONS

Our SemProtector addresses some fundamental challenges of securing semantics for online SC systems over open

	MNIST		CIFAR10	
	PSNR	SSIM	PSNR	SSIM
Online SC	27.88	0.99	25.49	0.87
Online SC + SemEryp	24.09	0.94	22.91	0.78
Online SC + SemPriv	27.75	0.97	25.04	0.86
Online SC + SemRect	26.69	0.93	23.19	0.77
Online SC + SemPriv + SemRect	26.67	0.93	23.08	0.76
Online SC + SemEryp + SemPriv	24.06	0.93	22.94	0.79
Online SC + SemEryp + SemRect	23.87	0.92	22.90	0.76
Online SC + SemEryp + SemPriv + SemRect	23.77	0.92	22.85	0.76

wireless channels. We further discuss two important research directions for the security of semantic communications.

A. Certified robustness of semantic communications

Although our SemProtector secures online SC with a unified framework and significantly hardens the system against various attacks, we still lack a theoretical analysis under what conditions the system keeps stable. A further step we can take is certified robustness, which provides the lower bound of robust accuracy against attacks under certain conditions. Robustness verification holds the promise of building reliable the endto-end SC systems with a theoretical guarantee. Compared to traditional certified robustness, we need to additionally consider the physical layer adversarial attacks in fluctuated wireless channels, as well as the adversarial perturbations to the source data to derive a tight bound. These studies are still at an early stage, and developing certifiably robust approaches for an online SC system is still an open research problem.

B. Security of distributed semantic communications

Massive machine-type communications (mMTC) proposed in 5G is one of the key enablers for future cellular-based factory automation, using smart sensors and actuators to enhance manufacturing and industrial processes. Distributed semantic communications facilitate the above industrial IoT applications with a highly reliable and minimal latency communication paradigm. However, we will face challenging security issues due to the massive connected devices that can sense, communicate and store information about themselves. These devices are deployed in heterogeneous wireless networks and need to make intelligent decisions, and hence they are more fragile to various attacks. Securing distributed semantic communications in industrial IoT remains an open research problem.

VI. CONCLUSION

This paper introduces SemProtector, a unified framework that aims to secure an online SC system with three hotpluggable semantic protection modules. These protection modules are able to encrypt semantics, mitigate privacy risks from wireless channels and calibrate distorted semantics at the destination. Our framework can also enable an existing online SC system to dynamically assemble the three modules to meet customized semantic protection requirements, greatly facilitating the practical deployment in real-world SC systems. Although SemProtector is tested on an image-based semantic communication system, we believe that our framework can also be further extended to video-based SC systems.

REFERENCES

- G. Nan, X. Liu, X. Lyu, Q. Cui, X. Xu, and P. Zhang, "Udsem: A unified distributed learning framework for semantic communications over wireless networks," *IEEE Network*, pp. 1–8, 2023.
- [2] G. Nan, Z. Li, J. Zhai, Q. Cui, G. Chen, X. Du, X. Zhang, X. Tao, Z. Han, and T. Q. Quek, "Physical-layer adversarial robustness for deep learning-based semantic communications," *IEEE Journal on Selected Areas in Communications*, pp. 1–1, 2023.
- [3] Y. Xiao, Z. Sun, G. Shi, and D. Niyato, "Imitation learning-based implicit semantic-aware communication networks: Multi-layer representation and collaborative reasoning," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 3, pp. 639–658, 2022.
- [4] Q. Qin, Y. Rong, G. Nan, S. Wu, X. Zhang, Q. Cui, and X. Tao, "Securing semantic communications with physical-layer semantic encryption and obfuscation," arXiv preprint arXiv:2304.10147, 2023.
- [5] H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, "Deep learning based semantic communications: An initial investigation," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, 2020, pp. 1–6.
- [6] R. Shokri, M. Stronati, C. Song, and V. Shmatikov, "Membership inference attacks against machine learning models," in 2017 IEEE symposium on security and privacy (SP). IEEE, 2017, pp. 3–18.
- [7] M. Fredrikson, S. Jha, and T. Ristenpart, "Model inversion attacks that exploit confidence information and basic countermeasures," in *Proceedings of the 22nd ACM SIGSAC conference on computer and communications security*, 2015, pp. 1322–1333.
- [8] M. Sadeghi and E. G. Larsson, "Physical adversarial attacks against endto-end autoencoder communication systems," *IEEE Communications Letters*, vol. 23, no. 5, pp. 847–850, 2019.
- [9] X. Luo, Z. Chen, M. Tao, and F. Yang, "Encrypted semantic communication using adversarial training for privacy preserving," *IEEE Communications Letters*, vol. 27, no. 6, pp. 1486–1490, 2023.
- [10] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," arXiv preprint arXiv:1412.6572, 2014.
- [11] M. Yang, C. Bian, and H.-S. Kim, "Ofdm-guided deep joint source channel coding for wireless multipath fading channels," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 2, pp. 584–599, 2022.
- [12] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings* of the IEEE conference on computer vision and pattern recognition, 2018, pp. 4510–4520.
- [13] T. Han, Q. Yang, Z. Shi, S. He, and Z. Zhang, "Semantic-preserved communication system for highly efficient speech transmission," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 1, pp. 245–259, 2022.
- [14] Q. Hu, G. Zhang, Z. Qin, Y. Cai, G. Yu, and G. Y. Li, "Robust semantic communications against semantic noise," in 2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall). IEEE, 2022, pp. 1–6.
- [15] X. Peng, Z. Qin, D. Huang, X. Tao, J. Lu, G. Liu, and C. Pan, "A robust deep learning enabled semantic communication system for text," in *GLOBECOM 2022-2022 IEEE Global Communications Conference*. IEEE, 2022, pp. 2704–2709.

BIOGRAPHIES

Xinghan Liu (liuxinghan_2022@bupt.edu.cn) is currently pursuing a master's degree in National Engineering Research Center for Mobile Network Technologies, Beijing University of Posts and Telecommunications, Beijing 100876, China. His research interests include semantic communication security and privacy preservation.

Guoshun Nan (nanguo2021@bupt.edu.cn) is a professor of National Engineering Research Center for Mobile Network Technologies, Beijing University of Posts and Telecommunications, Beijing 100876, China. He has broad interest in semantic communications, machine learning and wireless communications. He has published over 30 papers in top-tier conferences and journals, including ACL, CVPR, KDD, IEEE Network, and IEEE JSAC. **Qimei Cui** (cuiqimei@bupt.edu.cn) received the B.E. and M.S. degrees in electronic engineering from Hunan University, Changsha, China, in 2000 and 2003, respectively, and the Ph.D. degree in information and communications engineering from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2006. She has been a Full Professor with the School of Information and Communication Engineering, BUPT, since 2014. Her research interests include B5G/6G wireless communications, edge computing and IoT.

Zeju Li (lizeju@bupt.edu.cn) is an undergraduate student at Beijing University of Posts and Telecommunications (BUPT), Beijing, China. His research interests include AI safety, NLP and semantic communications.

Peiyuan Liu (pyuanL@bupt.edu.cn) is currently pursuing a master's degree in Beijing University of Posts and Telecommunications (BUPT). Her research interests include federated learning security and semantic communications.

Zebin Xing (xzebin@bupt.edu.cn) is an undergraduate student at Beijing University of Posts and Telecommunications (BUPT), Beijing, China. His research interests include Federal Learning and Differential privacy

Hanqing Mu (muhanqing@emails.bjut.edu.cn) is currently pursuing a master's degree in Beijing University of Posts and Telecommunications (BUPT). His research interests include AI safety and privacy preservation.

Xiaofeng Tao (taoxf@bupt.edu.cn) received the B.S. degree in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 1993, and the M.S. and Ph.D. degrees in telecommunication engineering from Beijing University of Posts and Telecommunications(BUPT), Beijing, China, in 1999 and 2002, respectively. He is a Professor in BUPT, a Fellow of the Institution of Engineering and Technology, and Chair of the IEEE ComSoc Beijing Chapter. He has authored or co-authored over 200 papers and three books in wireless communication areas. He focuses on 5G/B5G research.

Tony Q.S. Quek (tonyquek@sutd.edu.sg) received the B.E. and M.E. degrees in Electrical and Electronics Engineering from Tokyo Institute of Technology, respectively. At MIT, he earned the Ph.D. in Electrical Engineering and Computer Science. Currently, he is the Cheng Tsang Man Chair Professor and Full Professor with Singapore University of Technology and Design (SUTD). He is currently serving as an Editor for the IEEE Transactions on Wireless Communications as well as an elected member of the IEEE Signal Processing Society SPCOM Technical Committee. He was an Executive Editorial Committee Member of the IEEE Transactions on Wireless Communications, an Editor of the IEEE Transactions on Communications. He is a Fellow of IEEE and a Fellow of the Academy of Engineering Singapore.