Codeword-Independent Representation Learning for Localization in Reconfigurable Intelligent Surface-Enhanced Environments

Luo, Xuanshu

Award date:
2022

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Department of Mathematics and Computer Science
Interconnected Resource-aware Intelligent Systems cluster

Codeword-Independent Representation Learning for Localization in Reconfigurable Intelligent Surface-Enhanced Environments

Master Thesis

Xuanshu Luo

Supervisors:
prof. dr. ir. Nirvana Meratnia

Eindhoven, October 2022

Abstract

Reconfigurable Intelligent Surfaces (RISs) are expected to be a key component in modern wireless communication systems due to their capability to manipulate signal propagation for software-defined radio. The deployment of RISs signifies additional adaptions for wireless communication applications. In this thesis, we focus on the localization problem in RIS-enhanced environments.

Many related works tried to calculate the locations of the device of interest (DOI) using sophisticated communication systems. However, these methods require channel matrices among the base stations (BSs) and DOIs based on accurate field measurements, which are rarely available in practice. Even worse, as the number of BSs and DOIs increases, the fast-growing complexity indicates that these methods could hardly provide reliable localization services. Some studies noticed this problem and designed data-driven methods using simple fingerprints. Despite their proposed solutions, these methods are only feasible under an unrealistic assumption that the RIS is deployed only for localization purposes instead of more important tasks, such as enlarging the signal coverage. Meanwhile, the model-based and data-driven methods have a common problem that is the RIS codewords are required for location inference, which would induce considerable communication burdens.

Therefore, we propose a new localization paradigm that a reasonable fingerprint-based localization system in RIS-enhanced environments should make no particular assumption about the RIS functions and not require RIS codewords for online inference. This requires the localization model to possess RIS codeword domain generalization capabilities. In this thesis, we address this by extracting codeword-independent representations of fingerprints by adversarial learning on the RIS codeword domain. Concretely, we adopt the framework of the domain adversarial neural network and build our system containing four crucial parts. Firstly, we design a preprocessing step to transform fingerprints with non-Euclidean features into graphs. Secondly, a graph neural network-based feature extractor encodes the graphs and outputs the fingerprint representations. Thirdly, the location estimator composed of a multi-layer perceptron (MLP) provides the estimation using fingerprint representations. Finally, the complex-value MLP-based codeword discriminator conducts adversarial learning on the RIS codeword domain by returning its reversed gradients to the feature extractor during backpropagation. Consequently, this training process guides the feature extractor to generate representations whose distributions are as similar as possible for different codewords. In this way, our method can extract codeword-independent representations of fingerprints, which supports accurate online inference without RIS codewords.

We evaluate our solution using the DeepMIMO dataset. Due to the lack of results from other studies, for fair comparisons, we also define oracle and baseline cases, which are the theoretical upper and lower bounds of our system, respectively. In all experiments, the proposed solution performs much closer to the oracle cases rather than the baseline cases, which demonstrates the effectiveness and robustness of our method.

Preface

I would like to first thank my family who fully supports my future choices, provides sound advice, and most importantly, always loves me without any expectation of rewards. I also appreciate my friends for their help when I was in trouble and for their company when I was depressed. Although unfortunately, I did not complete all preset goals during my master's life in TU/e, the experiences of more than two years still benefit me for a lifetime, which is not only because of academic knowledge but also something more precious.

Last but not least, I sincerely appreciate my supervisor, prof.dr.ir. Nirvana Meratnia for her sophisticated guidance and dedicated support. Especially when my explanation is ambiguous, she always shows her patience and helps me to straighten out the facts, combine them through the contexts of specific questions, and explore new possibilities. Without her support, it is hard for me to imagine what I could have achieved so far. Once again, my sincere thanks to her.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Modern wireless communication systems are expected to offer not only better quality of services (QoS) but also software-defined radio for diverse ubiquitous applications [1]. To realize these expectations, the (beyond) fifth-generation (5G) mobile telecommunication systems consider the reconfigurable intelligent surface (RIS) as a crucial component for its capability to customize the radio frequency (RF) propagation properties, enabling configurable wireless networks [2, 3]. Specifically, the RIS contains a two-dimensional array of discrete elements, whose electromagnetic impedance is individually tunable [4, 5], which theoretically supports artificial manipulation of signal reflection, diffraction, refraction, polarization, and absorption [6, 7, 8].

Leveraging higher-frequency RF for faster data rates is generally accepted for modern communication networks [9, 10], whereas results in more severe signal attenuation when there is no line-of-sight (LoS) link between the base station (BS) and the device of interest (DOI) due to weaker diffraction abilities [11, 12]. Therefore, among the mentioned RIS functionalities, controlling the signal reflection properties is the most influential one, since it prevents excessive received signal strength indicator (RSSI) drop by constructing a RIS-assisted link to reflect the incoming signal to the direction towards the DOI in non-line-of-sight (NLoS) areas [13, 14]. RISs achieve this by assigning suitable complex-valued (CV) RIS codewords to specify the beamforming configurations of all RIS elements [15, 16]. For instance, Figure 1.1 shows that using an appropriate codeword, a RIS-assisted link could be created, and improve the RSSI of the NLoS DOI_2. Therefore, with proper RIS deployment and codeword specification, various applications in (beyond) 5G networks such as smart sensing, imaging, and localization [17, 18], benefit from the improvement of the QoS in NLoS areas, thereby achieving comparable performances as the LoS areas [19, 20].

Such enhancement particularly promotes research on the localization problems in RIS-enhanced environments, since alternative technologies, such as the well-known Global Navigation Satellite System (GNSS) [21], usually cannot obtain satisfactory accuracy for all localization scenarios, especially in NLoS areas [22, 23]. Localization has assumed considerable importance as a foundational application in wireless communication systems, which supports a series of location-based services, such as navigation [24], smart factories [25], and virtual reality entertainment [26]. Conventional localization techniques in wireless communication networks utilize various auxiliary technologies, such as time of flight (ToF) [27], RSSI [28], etc., and manage to obtain high-accuracy location estimation in differ-

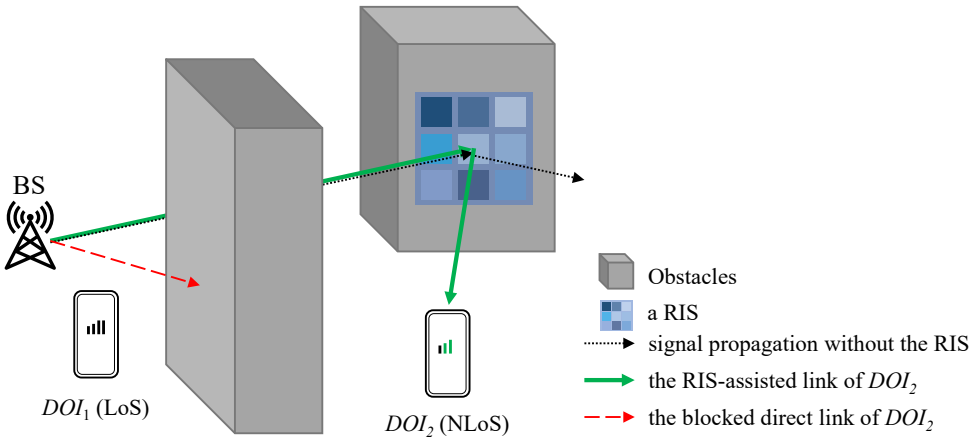

Figure 1.1: RIS-assisted signal propagation from a BS to a NLoS device in a RIS-enhanced environment using a codeword that controls the RIS to reflect the impinging signal to DOI_2. Here, both the RSSI and the time of flight (ToF) for DOI_2 would increase.

ent scenarios [29]. However, these methods cannot be directly adopted in RIS-enhanced environments, since the RIS changes the original signal propagation properties, inducing wrong inference on distances or locations. For example, both the RSSI and the ToF for DOI_2 in Figure 1.1 would increase when the RIS reflects the signal from the BS in the direction of DOI_2. Then conventional RSSI-based methods would underestimate the distance between the BS and DOI_2 since the RSSI is inversely proportional to distance according to signal attenuation models [30]. Whereas conventional ToF-based methods would overestimate this distance because the RIS-assisted link is longer than the blocked directed link in Figure 1.1. Such opposite estimation errors illustrate that the deployment of RIS leads to irregular estimation errors when using conventional localization technologies, which require non-trivial adaptations for RIS-enhanced environments.

Therefore, many research works focus on the localization problem in RIS-enhanced modern wireless communication systems. Some related studies proposed sophisticated communication models, then applied 3D geometric knowledge to either directly calculate DOI positions or analyze the estimation error bounds [31, 32, 15]. However, all these works require detailed CV channel gains/matrices among BSs and DOIs based on accurate field measurements (e.g., ray tracing and wave optics [7]), which are hardly commonly available in practice. Meanwhile, since multiple-input, multiple-output (MIMO) [33] orthogonal frequency-division multiplexing (OFDM) [34] is an accepted paradigm for modern wireless networks [35], as the number of BSs and DOIs increases, these models have to consider more and more data links on all sub-carriers between devices. Hence, the fast-growing complexity of these methods leads to long inference time [36], which severely limits their usability in real-world deployments. Considering these drawbacks, other solutions measure the RSSI from known BSs for data-driven fingerprint-based localization in RIS-enhanced environments, as the fingerprints are much easier to access than field measurements [37], and far simpler than CV channel matrices [38, 39]. However, these methods usually involve cooperation with optimized RIS codewords to provide high-accuracy localization, which unrealistically assumes that the RIS is solely deployed for localization rather than more pressing demands in modern high-frequency wireless networks, such as enlarging the signal coverage, increasing the overall network throughput [7, 40], etc. Even worse, these types of approaches share a common disadvantage that the CV RIS codewords

are indispensable during location inference, which introduces huge additional communication burdens to transmit large CV vectors to DOIs.

1.1 Problem statement

The above analysis reveals that a reasonable localization system in a RIS-enhanced environment should use simple fingerprint data to avoid the complexity issue without any particular assumptions on the RIS functions for more pragmatic real-world deployments. Meanwhile, the RIS codewords are expected to be unnecessary for online inference to reduce communication costs. To the best of our knowledge, we are the first to propose this paradigm. However, there are still problems to implement this idea.

Conventional fingerprint-based localization systems perform well since the fingerprints depend on only relative locations of DOIs and these methods have sufficient generalization capabilities to locations with the help of deep neural networks (DNNs) [29]. However, in RIS-enhanced environments, the fingerprints are affected by not only locations but also codewords, which means conventional methods could hardly provide high-accuracy predictions due to the lack of generalization ability to codewords. In our paradigm, when we lose control of the RIS functions, the RIS codewords could change at any time. Therefore, the dataset consisting of RSSI measurements for offline training is unlikely to contain fingerprints corresponding to all possible codewords. Then during online inference, the localization system will possibly make inferences using fingerprints of unknown codewords, leading to large estimation errors, because the location estimator only fits the codeword distribution of the training dataset. In the field of statistical learning, such performance degradation caused by differently distributed training and test data is known as a *domain generalization* (DG) problem [41], where the domain here contains all possible RIS codewords. Meanwhile, if the DG problem remains unsolved, regardless of the data-driven techniques used since the localization system is unaware of the codeword distribution for online inference, the codewords are still essential as auxiliaries to compensate for the performance drops caused by the lack of generalization ability to codewords.

To summarize, although the proposed localization paradigm in RIS-enhanced environments is more pragmatic than existing solutions, the introduced DG issue makes the localization system impossible to obtain high-accuracy online estimations without relying on corresponding codewords. Consequently, the location estimator either provides predictions with large errors or causes huge communication loads for codeword transmission.

1.2 Research questions

The analysis in the previous section points out that the barrier to implementing the novel paradigm is the DG problem. Therefore, the main goal of this thesis is to realize reliable fingerprint-based localization with codeword domain generalization capabilities in RIS-enhanced communication systems. To achieve this goal, we further decompose our goal into several specific research questions (RQs).

RQ1. How RIS codeword domain generalization can be achieved? What are the specific pros and cons of the candidate DG solutions for our problem?

RQ2. Which DNN architecture is proper to encode the fingerprints, especially in mod-

ern communication systems usually containing 4G and 5G operating on different frequency bands?

RQ3. How to compare the performance of our solution with respect to the state of the art, when no related works realize localization systems with codeword domain generalization capabilities in RIS-enhanced environments?

1.3 Our contribution

To solve the RQs and achieve the main goal of this thesis, we further analyze the DG problem. In RIS-enhanced environments, the fingerprints are affected by both locations and codewords, which means a qualified localization system conforming to the new paradigm is capable to generalize to both factors. Given the widely validated generalization abilities of conventional DNN-based systems to locations [29, 42, 43, 44], we consider operations on the codeword domain. Fortunately, the impact of codewords on fingerprints is not meaningless noise interference but depends on how the codewords manipulate signal propagation. This observation inspires us to learn from this correlation and try to decouple the dependency of fingerprints on codewords to obtain codeword-independent representations of fingerprints for localization in RIS-enhanced environments. Such a system will be able to generate fingerprint representations highly related to locations only, which kills two birds with one stone, i.e.,

1. solve the RIS codeword domain generalization problem and avoid large estimation errors for fingerprints of unknown codewords.
2. predict the locations without corresponding codewords during online inference.

We realize this idea by adversarial learning on the codeword domain using the Domain-Adversarial Neural Network (DANN) [45] framework consisting of three parts: a fingerprint feature extractor, a location estimator, and a codeword discriminator. Specifically, during the offline training stage, the feature extractor and the location estimator together behave like conventional fingerprint-based localization systems, i.e., minimizing the location estimation error. Simultaneously, the codeword discriminator performs adversarial learning on codewords by reversal gradients during backpropagation to guide the feature extractor to generate codeword-independent representations. In this way, the proposed system would solve the DG problem, thereby fulfilling the main goal of this thesis. Our contributions are as follows:

1. We analyzed the localization problem in RIS-enhanced networks in depth and proposed a novel paradigm that is more reasonable compared with prior studies, i.e., fingerprint-based solutions without additional assumptions on the RIS codewords, which also support online inference without codewords.
2. In this context, we proposed a localization solution based on codeword-independent representation learning using the DANN framework to solve the DG problem.
3. Our proposed solution is extensively evaluated by the *DeepMIMO* dataset [46]. We designed oracle and baseline cases for comparison, which convincingly demonstrate that our solution achieves accurate localization even for unknown RIS codewords. Additional experiments on the system parameters further demonstrate the rationality and robustness of the proposed solution.

1.4 Organization

The rest of this thesis is organized as follows. Chapter 2 discusses related works of conventional localization methods and their variants in RIS-enhanced environments. Chapter 3 provides fundamental information of RISs and briefly explains the domain generalization problem. Chapter 4 elaborates on the detailed design of the proposed codeword-independent representation learning-based localization system. In Chapter 5, extensive experiments and results are presented. Finally, we conclude this thesis in Chapter 6.

Chapter 2

Related Works

This chapter first discusses conventional localization techniques. Then we focus on the localization solutions in RIS-enhanced environments.

2.1 Conventional localization techniques

Though GNSSs provide universal localization services, the accuracy and coverage are restricted. For example, even under the open sky, most smartphones can only achieve a radius error of 4.9 m using Global Positioning System [21]. The accuracy would further drop in GNSS-denied areas (e.g., indoor) due to signal blockage [47]. Hence, localization is still an open issue.

Many related works utilize auxiliary technologies to realize localization with high accuracy, including ultrasonic [48], ultra-wideband [49], magnetic field [50], visible light communication [51], etc. Despite their compelling performances, all these techniques require additional installation of dedicated sensors on the areas for localization, which largely confines their application scopes.

In contrast, the widespread deployment of RF-based communication technologies, such as WiFi and Bluetooth, shows their potential for localization. A common pipeline for model-driven RF-based localization methods is first estimating the distances to known transmitters (e.g., BSs, routers), then calculating the locations of DOI using compatible geometry knowledge, such as triangulation [37, 52]. There are many different solutions to predict distances. Except for RSSI and ToF mentioned in Chapter 1, popular methods includes angle of arrival [53], phase of arrival [54], etc. These model-driven methods are straightforward to implement, but theoretical signal attenuation models cannot be aware of the real environmental information, such as the sizes and positions of obstacles [29]. Consequently, these methods fail to achieve desired performances in real-world deployment due to signal blockage, multi-path propagation, shadowing, [28] etc.

The limitation of model-based methods inspires research on data-driven localization techniques using the most accessible RSSI [55]. Instead of calculating locations directly from RSSI, data-driven methods first collect fingerprints, i.e., measurements of RSSI from all transmitters at some reference locations, to construct an offline fingerprint database for model training. Then for the query fingerprints, the trained model usually provides more

reliable predictions than model-driven methods, since a well-trained estimation model is aware of the on-site information, which is implicitly recorded by a good fingerprint database [56, 57].

The performances of fingerprint-based localization methods depend on both the quality of the fingerprint datasets and the efficiency of the data models adopted. For the former, a crucial problem is the localization environments may change over time due to variations of obstacles and device attrition [58, 59]. As a result, the fingerprint datasets are outdated and invalid for the current environments. Hence, many studies rely on crowd-sourcing to update the dataset to optimize the model parameters online for reliable localization services [60, 61]. However, the crowd-sourcing fingerprints can be from anyone, including the imposters who intend to attack the localization systems using falsified data [62]. To address this issue, Sun et al. [63] proposed a transformer-based [64] adversarial sample discriminator to identify and filter out the fake fingerprints. Whereas in [65], a generative adversarial network-based [66] anomaly detection model is designed to distinguish abnormal unlabeled fingerprints. Besides, some studies noticed that fingerprints of different locations may be similar, inducing large errors for some RSSI samples [38, 67, 68]. Therefore, ViViPlus [58] designed RSSI spatial gradient consisting of the RSSI differences of selected adjacent locations to reduce the possibility of fingerprint similarity by considering more ambient data. There are also some studies trying to compress the fingerprints or the databases for lightweight models, quick responses, and efficient deployment. For instance, MonoFi [69] proposed a localization system using fingerprints containing only one RSSI value, and Arya et al. proposed Block-based Weighted Clustering [70] to simplify the fingerprint databases. These methods all focused on the optimization of fingerprint datasets for better localization performances.

Except for the improvement of fingerprint dataset quality, many related studies were more concerned with the data models. Some methods analyzed their application scenarios carefully and chose appropriate algorithm frameworks or auxiliary techniques. In [71], an extreme learning machine [72] is adopted for extremely fast inference speed in real-time systems. Subedi et al. [73], on the other hand, made full use of smartphones and proposed a fingerprint-based localization method with the help of inertial sensors. Whereas more works were concerned with proper machine learning/DNN models to encode fingerprints. Some related works focused on availability using small fingerprint dataset, and adopt simple k nearest neighbors (k-NN) or Naive Bayes estimators [74, 75]. Another prior studies utilized relatively complex models such as support vector machine [76] and multi-layer perceptron (MLP) [42] to predict the locations. There are also solutions pursuing ultimate performances by more advanced DNN models. These methods usually involve additional preprocessing steps for compatibility. For instance, the WiDeep [57] manually generates noisy data and trains a denoising autoencoder to enhance the robustness of the localization system to random noise. Ibrahim et al. [43] organized multiple fingerprints as time series and adopted a convolutional neural network (CNN) model for localization. Chen et al. [77] proposed a similar method to generate time series but utilized a long short-term memory (LSTM) [78] model to encode the fingerprints. In [79], extensive experiments on various recurrent neural networks (RNNs) [80] were conducted, including Naïve RNNs [81], gated recurrent unit [82], and LSTM.

Although these solutions achieve competitive results, they are still far from ideal performances for data-driven approaches, since the mentioned models are not particularly designed for fingerprints with non-Euclidean properties [83]. Such lack of concordance

significantly downgrades the localization accuracy, especially when the fingerprints involve multiple RF technologies or frequency bands [84]. By contrast, graph neural network (GNN)-based models, which are designed for non-Euclidean data, achieve state-of-the-art performances in fingerprint-based localization [84, 85, 86]. Hence, in this thesis, we would extend a heterogeneous GNN-based model first proposed in one of our prior works [84] to encode the fingerprints, which keeps the heterogeneity of fingerprints measured in modern communication networks operating on multiple frequency bands. We will provide further explanation in Chapter 3.2.

2.2 Localization in RIS-enhanced environments

The RIS is capable to manipulate signal propagation, thus all applications relying on wireless communication networks need to be updated when the RIS is deployed. Similar to conventional methods, there are both model-driven and data-driven localization techniques for RIS-enhanced environments.

Model-driven methods usually have specific optimization goals to directly or indirectly facilitate localization. Wymeersch et al. [87] utilized the Fisher information analysis [88, 89] to select the best RIS codewords for localization. Elzanaty et al. [31] designed a RIS codeword optimization scheme to maximize the signal-to-noise ratio at the DOI to facilitate localization, and provided analysis on the Cramér-Rao lower bound [90] of the localization error. Despite their contributions, these works are only feasible when the channel gains/-matrices among BSs, DOIs, and the RIS are known, which requires accurate field measurements (e.g., ray tracing [7]), thereby greatly restricting their applicability. Meanwhile, modern communication networks usually adopt MIMO antennas using OFDM. Consequently, as the number of BSs and DOI increases, the complexity of these methods increases fast. Hence, many studies explicitly restrict the number of devices in the environment. For instance, in the analysis of Wymeersch et al. [87], the number of BSs is set to 1 for simplicity. Such fast-growing complexity would induce unacceptable time delay during location inference when there is a vast number of requests for localization services.

Unfortunately, most research works on localization in RIS-enhanced environments could be categorized as model-driven ones. Only a few works noticed the mentioned drawbacks, utilized fingerprints, and proposed data-driven methods. Zhang et al. [38] designed a codeword selection method that tries to enlarge the differences of fingerprints of adjacent positions for high-accuracy localization. They further refined this work and proposed an integrated localization system in [91] realizing a centimeter-level error. Huang et al. [39] also picks the best available RIS codeword, but simply for signal strength improvement, which indirectly enhances localization performance. However, the prerequisite for these methods is the RIS is deployed solely for localization rather than more important tasks like enlarging the coverage, increasing the network throughput [40], etc. Such assumptions are very unlikely to happen in real-world deployments. Even if realized, these data-driven methods and aforementioned model-driven methods share a common disadvantage in that they all require RIS codewords for location inference, which results in inescapable additional communication burdens on the networks to continuously transmit large CV vectors as requested by DOI. As the number of DOIs grows, the whole localization system would eventually collapse and affect the operation of RIS-enhanced communication systems. Hence, this thesis pursues a localization system in RIS-enhanced environments that does not require RIS codewords during online inference.

2.3 Domain generalization

The domain generalization (DG) problems [41] usually occur in deployments of statistical learning techniques due to distinct distributions of offline training data and online data. This phenomenon is very common because collecting a training dataset perfectly representing real-world scenarios is tend to be impossible. Although statistical learning methods, especially for DNNs, have recognized generalization capabilities, this is only true for online data of the same distribution as the training data [92]. Even for computer vision tasks involving millions of images for training, the DG issue still significantly affects online accuracy. The most famous example is [93], where the classification accuracy drops up to 15% for re-selected online data without adding any new images. Figure 2.1 shows an example of the DG problem using digits caused by distinct fonts, backgrounds, and colors. In some studies, the DG problem is also termed as domain shift [94] or out-of-distribution problem [95, 41].

Figure 2.1: Digits from different datasets. The digits are samples from MNIST [96], MNIST-M [45], and SVNM [97].

There are many different research topics addressing the DG problem. Multi-task [98] learning reuses the data representations for training on different related tasks to help the model to perform better for the original task. However, multi-task learning models obviously only work for the domains they have already seen, which hardly fully solves the DG problem. Transfer learning [99, 100] pre-trains the models using data from classes with a large number of samples, then fine-tunes the model parameters with data from other classes. This canonical pipeline could be used for solving the DG problem by fine-tuning the models using data from different domains. The main drawback of transfer learning methods is they require data from other domains for training, which is not always possible [101], such as the localization problem in this thesis. Besides, meta-learning techniques [102, 103] are popular recently as theoretically promising solutions to DG problems. Meta-learning tries to learn a relatively universal rule on different datasets for different tasks, i.e., learning to learn [104], which is close to human learning habits. Although meta-learning methods are attractive for DG problems, the computation time is always extremely long [105]. Meanwhile, their DG performances are sometimes unsatisfactory due to incidental overfitting on training datasets [105].

Considering the disadvantages of the discussed methods of DG issues, we focus on more widely adopted solutions here for our localization problem in this thesis. Data augmenta-

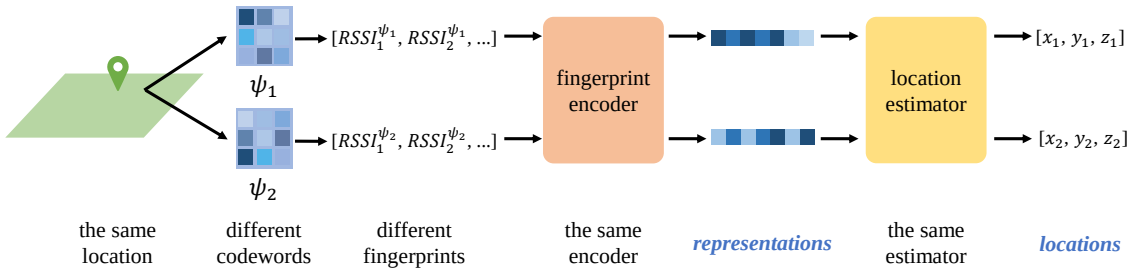

Figure 2.2: The pipeline of representation learning-based solutions solving the DG problem in this thesis. For the data of the same location with different codewords, feasible solutions would minimize either the differences in fingerprint representations or the final location estimations or both of them.

tion [106] aims to generate more data according to the existing training dataset. As a result, the augmented data is of greater diversity to fulfill DG. However, the manually generated data should be meaningful for the concerned tasks, otherwise, this process is just creating noisy data. In our problem, data augmentation only works if we manage to generate fingerprints of unknown codewords, which is impossible when the channel matrices among BSs, DOIs, and the RIS are unknown. Hence, methods based on data augmentation are hardly feasible for our problem. In contrast, representation learning [107] methods are particularly suitable in this thesis. The basic idea of representation learning methods to solve the DG problems is trying to unify either the data representations or the inference results, or both of them [41, 108], where the domain labels for training are required. When the extracted data representations are invariant for different domains, only relevant to the concerned tasks, domain generalization is achieved. Taking the localization problem in RIS-enhanced environments as an example illustrated in Figure 2.2, our goal is to make the fingerprint representations and the estimated locations consistent for fingerprints at the same locations regardless of the codewords.

In this context, contrastive learning [109] is a straightforward solution. A feasible pipeline for contrastive learning for our question is first generating fingerprint pairs whose corresponding locations are the same, whereas codewords are different. Then we try to minimize either the contrastive loss [110] between estimated locations or the Kullback–Leibler divergence [111] between the distributions of fingerprint representations for different codewords. However, training a high-accuracy regression model preferably traverses data pairs as much as possible, leading to heavy training overheads [112]. Assuming we have D domains and N samples for each domain, then there will be $\frac{1}{2}D^2N(N-1)$ data pairs available for training. In comparison, as another effective solution for the DG problems, DANN [45, 113] only uses DN data points to perform adversarial learning on the domain labels to generate data representations that are irrelevant to the concerned domain [114, 115]. DANN implements this by reversing the gradient of the domain discriminator during backpropagation [116]. Although the DANN requires explicit domain labels for training [41], which limits its versatility, this disadvantage does not hold in this case, since the domain labels are just the codewords. Therefore, the DANN is the more suitable technique for our problem. Section 3.3 further provides details of DANN frameworks.

Chapter 3

Preliminaries

This chapter will first provide fundamental information about RISs, including their working principles and the method for calculating path loss in RIS-enhanced environments. Next, we will review our prior work describing the way to transform fingerprints into graphs. Finally, the training process of DANN frameworks is elaborated.

3.1 Reconfigurable Intelligent Surfaces

This section will first explain how the RISs work. Then we will elaborate on the path loss calculation in RIS-enhanced environments.

3.1.1 The working principle of RISs

In communication systems equipped with MIMO antennas, the directions of beamforming depend on the constructive interference of all RF signals from radiators [117]. Figure 3.1 illustrates that the signal directions of two scenarios are different due to different time delays of two radiators. Such time delays manifest as phase shifts on the frequency domain. Hence, the two radiators together in Figure 3.1 (a) or (b) can be considered as a phase shifter.

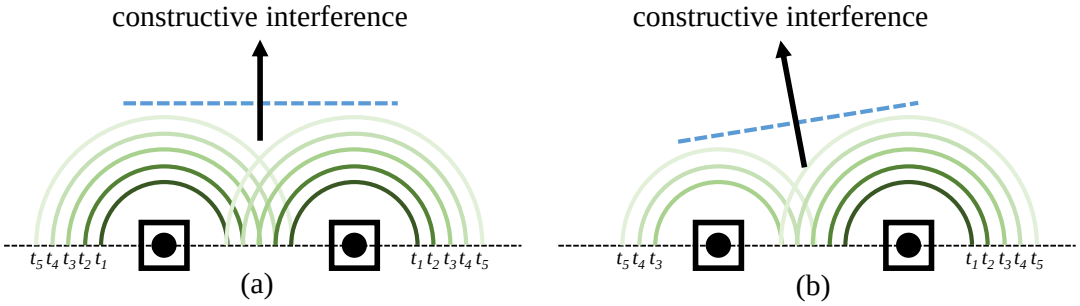

Figure 3.1: Different beamforming directions for two radiators with different time delays. In (a), two radiators emit the same signal at the same time; In (b), two radiators also emit the same signal, but the left one starts from t_3, inducing a different constructive interference compared with (a), thereby a left direction signal.

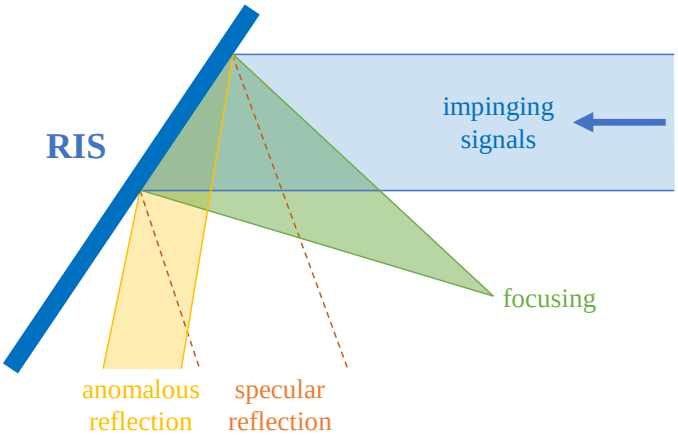

Figure 3.2: RIS behaviors: anomalous reflection (including specular reflection) and focusing.

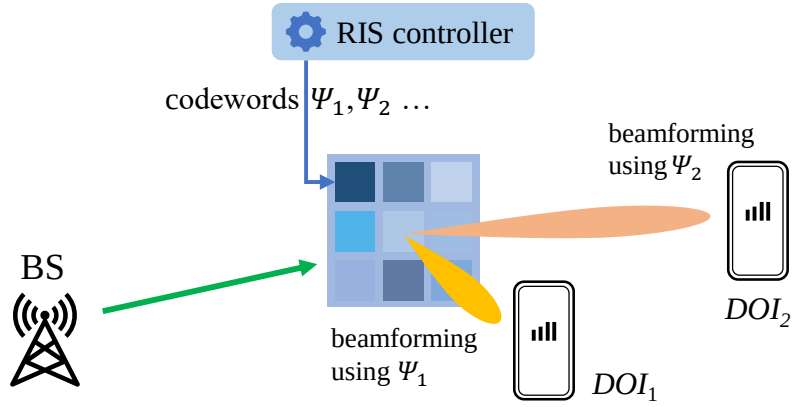

Figure 3.3: Different beamforming directions given different codewords.

RISs are capable to reflect the incident signals in the desired directions, as a RIS contains a 2D array of such phase shifters whose phases are individually reconfigurable and controlled by assigned impedance [4]. Codewords are CV vectors characterizing phase shifts of all RIS elements. Given a RIS with M elements, its codeword $\psi \in \mathbb{C}^{M\times 1}$, where for the element m with a phase shift ϕ, $\psi_m = e^{j\phi}$ [118]. Codewords determine the RIS behaviors, which generally include two categories: anomalous reflection and focusing [119, 120], as shown in Figure 3.2. Concretely, anomalous reflection means that the RIS reflects the impinging signals towards arbitrary directions in parallel, thus also including the common specular reflection (the angle of incidence equals the angle of reflection). Whereas focusing would converge the reflected signals to one point, which is very effective for a single DOI, but at the same time means low versatility. Hence, in this thesis, we only consider the RIS as a pure anomalous reflector, whose function is merely the manipulation of phase shifts.

In this context, assigning different codewords to the RIS can help obtain different reflection angles for different devices/functions, as shown in Figure 3.3. Hence, when the codewords change, shifts in RSSI values occur even for the same locations, which consequently downgrade the estimation accuracy of conventional RSSI fingerprint-based localization solutions in RIS-enhanced wireless communication systems.

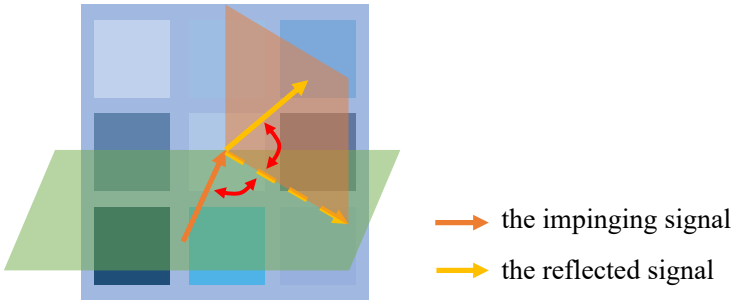

Figure 3.4: A sketch map illustrating a RIS manipulating reflection angles in two dimensions.

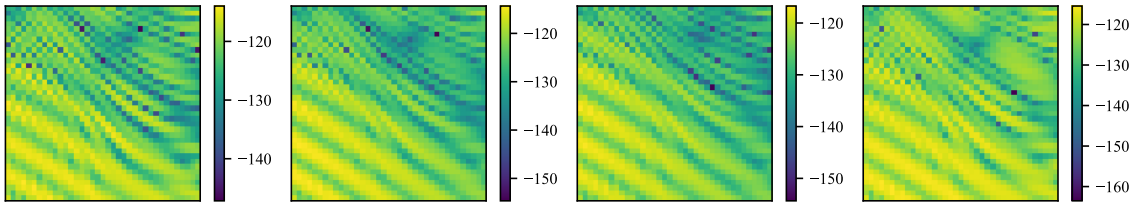

Figure 3.5: Shifted RSSI measurements (dB) using four different codewords for the same area. Each pixel denotes a RSSI value.

3.1.2 RIS codewords

Limited by material and production costs, the number of codewords supported by RIS is usually restricted [121]. The number of available RIS codewords C depends on the RIS resolution [122]. For instance, if a RIS supports a phase shift range of 180° with a 30° resolution, then $C = (180/30)^2 = 36$, because the RIS can manipulate reflection angles in two planes as shown in Figure 3.4. An alternative way to denote the resolution is the number of bits, i.e., a resolution of b bits means $C = 2^b$ [123]. Most current RIS prototypes support only 1 bit ($C = 2$) or 2 bits ($C = 4$) [124, 125], and a few could achieve up to 6 bits ($C = 64$) [126].

In subsequent evaluation, we need to generate all possible RIS codewords (codebook) for experiments given $C = C_l^2$ (resolution counted in degrees) and the phase shift range R for RSSI calculation and domain adversarial learning. We apply the method in [127]. Suppose the RIS contains $M = M_H \times M_V$ elements. Then for the m_H-*th* ($m_H \in [1, M_H]$) column of the Discrete Fourier Transform (DFT)-based codebook [128] $\mathscr{C}_H \in \mathbb{C}^{M_H \times C_l}$ for the horizontal dimension is

$$\frac{1}{\sqrt{M_H}}[1, e^{-j\frac{m_H}{M_H}R\cdot 1}, e^{-j\frac{m_H}{M_H}R\cdot 2}, ..., e^{-j\frac{m_H}{M_H}R\cdot(M_H-1)}]^T$$

Then we can use a similar way to calculate the DFT-based codebook $\mathscr{C}_V \in \mathbb{C}^{M_V \times C_l}$ for the vertical dimension. Finally, the codebook $\mathscr{C} = \sqrt{M_H M_V}\mathscr{C}_H \bigotimes \mathscr{C}_V$, where $\mathscr{C} \in \mathbb{C}^{M \times C}$, and $\bigotimes$ denotes Kronecker product [129]. Each column in $\mathscr{C}$ is a legitimate codeword.

Codewords have a great impact on fingerprints. As shown in Figure 3.5, the RSSI values are different at the same locations when using different RIS codewords, which means as long as the fingerprint dataset for offline training does not contain RSSI measurements corresponding to all possible codewords, the fingerprints for online inference would be

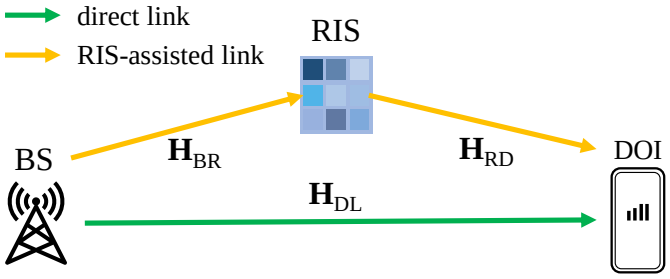

Figure 3.6: Path loss model in RIS-enhanced environments where the channel matrices **H** among the BS, the DOI, and the RIS are known.

of different distribution from the training dataset. We will provide our solution to the induced DG problem in the next chapter.

3.1.3 RSSI calculation in RIS-enhanced environments

The DeepMIMO dataset [46] for evaluation in this thesis does not provide RSSI values, thereby we cannot directly build fingerprints for experiments. Hence, this section elaborates on the RSSI calculation method in RIS-enhanced environments using channel matrices among BSs, DOI, and the RIS given by the DeepMIMO.

To calculate the RSSI measured by DOI, a common practice is to apply a link budget model, which describes all power gains and losses during the whole signal transmission process [130]. In general, the power gains (dB) G_p come from transmitter output power, transmitter/receiver antenna gain, whereas the power losses mainly include transmitter/receiver losses (dB) L_p and path loss (dB) PL [131]. Then applying a similar link budget model in [132], we have

$$RSSI\ (dB) = G_p - L_p - PL. \tag{3.1}$$

The values of G_p and L_p are usually given in the device documentation, so the remaining question is how to calculate PL. First, we assume that the RIS would only affect the phase shifts, which means there is no loss in the amplitude of the incident signals. Next, our path loss model works in a MIMO-OFDM system with K sub-carriers, containing a BS with P antennas, a DOI with Q antennas, and a RIS with M elements. For the sub-carrier k, the channel matrices of the direct link $\mathbf{H}_{DL,k} \in \mathbb{C}^{Q\times P}$ (from the BS to the DOI) and the RIS-assisted link, including $\mathbf{H}_{BR,k} \in \mathbb{C}^{M\times P}$ (from the BS to the RIS) and $\mathbf{H}_{RD,k} \in \mathbb{C}^{Q\times M}$ (from the RIS to the DOI), are provided by the DeepMIMO dataset. Figure 3.6 depicts the scenario we consider for path loss calculation. According to [133], for all K sub-carriers, the channel matrix of the direct link $\mathbf{H}_{DL} \in \mathbb{C}^{Q\times P}$ is just the sum of all sub-carrier components, i.e.,

$$\mathbf{H}_{DL} = \sum_{k=1}^{K} \mathbf{H}_{DL,k} \tag{3.2}$$

As defined by [127], given the codeword $\psi \in \mathbb{C}^{M\times 1}$, the channel matrix of the RIS-assisted link $\mathbf{H}_{RL} \in \mathbb{C}^{Q\times P}$ over all K sub-carriers is calculated as:

$$\mathbf{H}_{RL} = \sum_{k=1}^{K} \mathbf{H}_{RD,k} diag(\psi) \mathbf{H}_{BR,k}, \tag{3.3}$$

where $diag(\psi)$ is the diagonal matrix with the entries of ψ on its diagonal. Note that the codeword ψ is the same for all K sub-carriers. Finally, using the method described in [38, 91], the path loss *PL* in decibels from the BS to the DOI in this model is

$$PL(dB) = 10\log_{10}|\mathbf{H}_{DL} + \mathbf{H}_{RL}|^2 = 20\log_{10}|\mathbf{H}_{DL} + \mathbf{H}_{RL}|, \tag{3.4}$$

where $|\cdot|$ means the complex magnitude. Finally, combining these formulas together, we can calculate the RSSI of BSs for DOI using the DeepMIMO dataset.

Note that we do not need this RSSI calculation step in real-world deployments, since we could directly access the RSSI values using user devices, such as mobile phones. We emphasize that our fingerprint-based method has no dependencies on accurate field measurements. The experiments in this thesis rely on simulation using the DeepMIMO because RIS hardware is currently rare to be found. There are only several prototypes all over the world [124, 125, 126]. Hence, collecting real-world data through experiments is hardly possible by 2022.

3.2 Fingerprint-graph Transformation

Section 2.1 argued that most data-driven fingerprint-based methods utilize conventional machine learning techniques or deep learning models, which only work for Euclidean data whereas ineffective for non-Euclidean data (e.g., fingerprints) [84]. In comparison, GNNs are particularly designed for non-Euclidean data and have proven to be efficient for various downstream tasks [134]. Hence, the fingerprints should be first transformed into graphs to adapt GNN-based models. In our prior work [84], we proposed a two-stage preprocessing method to achieve this.

We consider an environment with $|T|$ RF technologies, where $T = \{t_1, t_2, ...\}$ denotes the set of all RF types. Then, as illustrated in Figure 3.7, the transformation method is as follows.

> **Step 1. Abstraction:** First, we gather information of all transmitters, including their types and locations. In the example shown in Figure 3.7, the transmitters are access points (APs), and $T = \{t_1, t_2\}$, which means there are two types of RF signals. We consider transmitters as vertices in a graph. The green vertex is the DOI, whose location is unknown. Then we assign the vertex features for transmitters by the combination of their locations and RSSI. Note that vertices of different types should be considered per type, which would result in *heterogeneous* graphs.

We decided to generate heterogeneous graphs for keeping the heterogeneity of the whole environment. Different RF technologies usually have different propagation laws, which means that encoding with the same set of model parameters is insufficient. Next, for connectivity, we first make two assumptions to decide the adjacency between vertices.

> **Assumption I**: Edges between vertices denote all possible signal propagations and interferences.
>
> **Assumption II**: A transmitter will only affect other transmitters of the same type.

Then we could follow the assumptions and link the nodes.

> **Step 2. Connection:** Considering Assumption I, since the DOI measures RSSI from all transmitters, there must be edges between the DOI and all transmitters. Note that

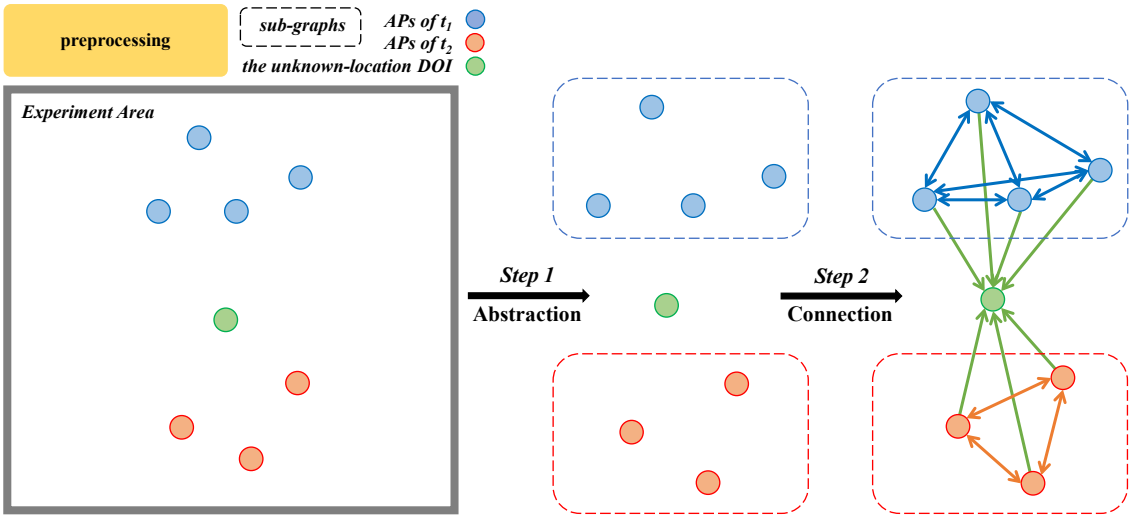

Figure 3.7: The fingerprint-graph transformation method proposed in our prior work [84], instantiated for $|T| = 2$ as an example.

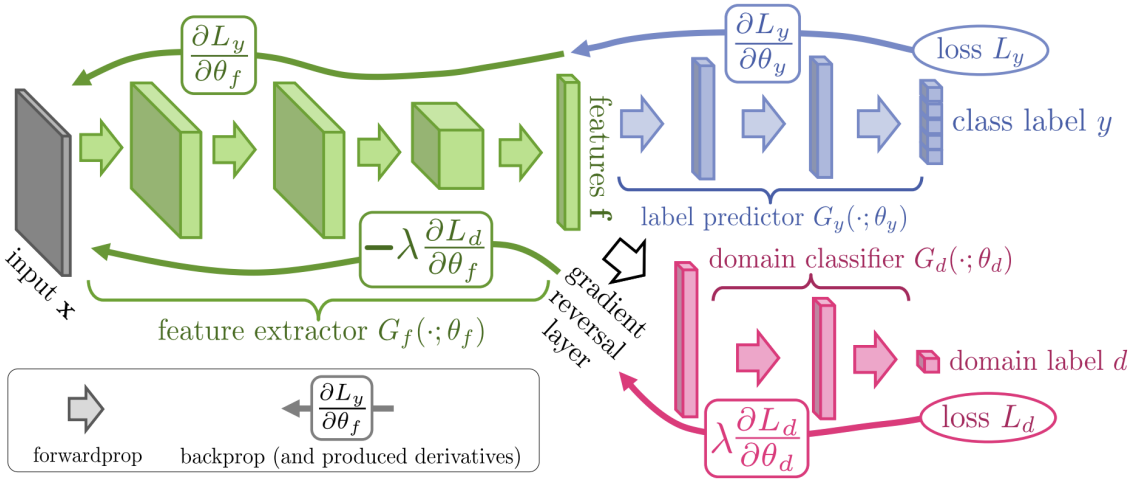

Figure 3.8: The architecture of DANN first proposed in [45].

these edges are *unidirectional* because the DOI is only a measuring device. Assumption II implies that the RSSI from a transmitter measured by the DOI is a combined result of all transmitters of the same type. Hence, transmitters of the same type will be fully connected. Since transmitters of the same type will affect each other, the edges within each *sub*-graph are *bidirectional*.

We can also assign edge features by theoretical signal attenuation models. In [84], the Log-Normal Shadowing Model (LNSM) is adopted [30]. In this way, given an arbitrary fingerprint, this transformation method could generate a corresponding graph for GNN models. In the next chapter, we would modify this preprocessing method to generate graphs for fingerprints gathered in RIS-enhanced environments.

3.3 Domain Adversarial Neural Network

In Section 2.3, we discussed the DG problem and concluded that the DANN framework fits the problem in this thesis. This section would briefly elaborate on the theory of DANN. Figure 3.8 illustrates the DANN framework, containing three main parts: a feature extractor, a domain discriminator (classifier), and a label predictor, which collaborates during offline training to achieve domain generalization [45] by extracting features that are irreverent to the concerned domain. Specifically, except for minimizing the errors on the label estimation, the DANN also reverses the gradient with the help of a gradient reversal layer (GRL) [113]. Concretely, the GRL only works during backpropagation (BP). Its behaviors include first obtaining the gradient from the first layer of the domain discriminator, reversing the gradients from the domain discriminator by multiplying a negative number $-\lambda$ ($\lambda > 0$), and finally passing the reversed gradient to the subsequent layer. Consequently, the feature extractor would simultaneously consider gradients from both the label predictor to minimize the estimation errors and the domain discriminator to minimize the divergence of feature distributions for different domain labels.

There are some additional concerns on the negative constant $-\lambda$ ($\lambda > 0$). During the initial phase of training, the feature extractor is not well-trained, so at this stage, the reversed gradient from the GRL should be suppressed, i.e., λ closing to 0. Then, as the training proceeds, the feature extractor could gradually extract meaningful representation vectors, thus the importance of adversarial learning on the domain is growing. Therefore, λ should gradually grow from 0. Supposing $p \in [0, 1]$ denotes the training progress, then instead of a fixed λ, we can define a more flexible version λ_p [113]:

$$\lambda_p = \frac{2}{1 + exp(-\gamma * p)} - 1,$$

where we set $\gamma = 10$ by default in the following experiments. λ_p would gradually grow from 0 to 1 as the training proceeded.

In this way, a well-trained feature extractor can ensure the distributions of the extracted features over different domains are as similar as possible. In this way, no matter what the domain labels are, the generated data features are ideally of the same distribution, which means we manage to solve the DG problem according to the analysis in Section 2.3. In the next chapter, the building blocks in the DANN framework for our problem are explained in detail.

Chapter 4

System Design

In this chapter, the details of our proposed localization solution are given, including the overall design, the offline training, and online inference pipelines, building blocks, and their corresponding motivations.

4.1 System overview

In this thesis, the proposed paradigm advises against making any assumptions on the RIS functions for more pragmatic deployments but leading to codeword domain generalization issues. However, conventional fingerprint-based localization systems would suffer from significant performance drops, because their underlying algorithms depend on an assumption that is invalid here, i.e., the fingerprints for training and testing are of the same distribution.

Therefore, we propose our DANN-based solution to solve the DG problem for localization in RIS-enhanced environments. This section would dive into the details of our proposed localization system. Section 3.3 mentioned that using DANN, a well-trained feature extractor can obtain domain-independent data representations. In order words, during online inference, the domain labels are no longer necessary. Hence, the offline training and online inference pipeline are different in our system. This section will describe these two phases separately. The structures of the building blocks in the proposed system would be explained in Section 4.2.

4.1.1 The offline training pipeline

In Section 2.3, we discussed that domain labels are indispensable for DANN model training. Therefore, except for the fingerprints and the corresponding locations L^*, the domain labels, i.e., the codewords ψ^* are also required in the training dataset. Before model training, our system first transforms the fingerprints into graphs, since recent studies argued that for non-Euclidean data like fingerprints, graph neural networks could extract more effective encodings for various downstream tasks than other models designed for Euclidean data [83, 84]. The specific fingerprint-graph transformation method would be elaborated on later in Chapter 4.2.1. After we obtain the fingerprint graphs, the training process is ready.

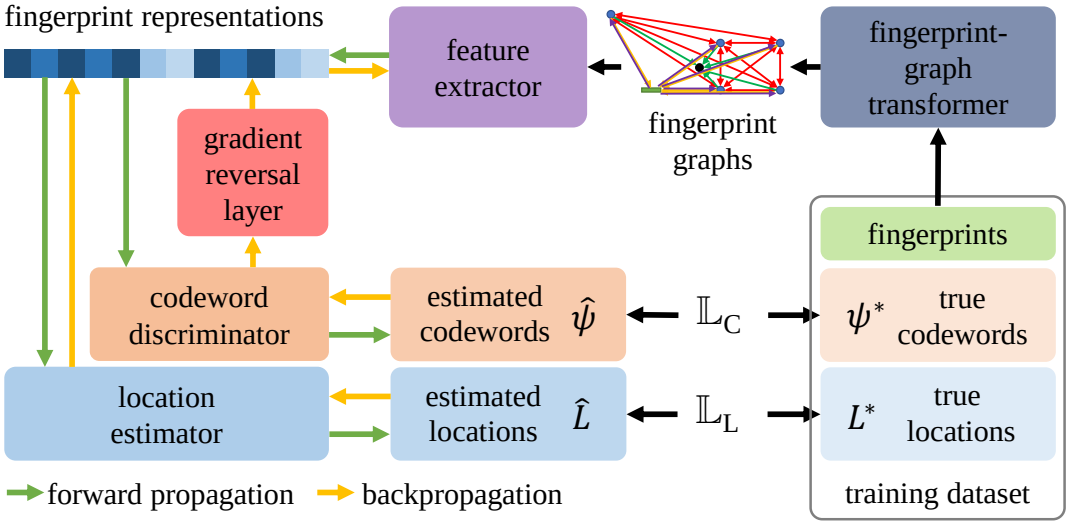

Figure 4.1: The framework of our proposed localization system (for offline training).

Figure 4.1 illustrates the whole training pipeline. During the forward propagation, the feature extractor first encodes the fingerprint graphs and obtains their representation vectors, which would be fed to both the location estimator and the codeword discriminator for location estimation $\hat{L}$ and codeword estimation $\hat{\psi}$, respectively. To measure the errors between the estimated results and the true labels, we utilize two loss functions separately for locations and codewords. Concretely, the mean squared error (MSE) [135] measures the localization error $\mathbb{L}_L$ between estimated locations $\hat{L}$ and true locations L^*, i.e., $\mathbb{L}_L = MSE(\hat{L}, L^*)$. Otherwise, measuring the codeword errors is relatively more complex, since the codewords are complex-valued vectors, whereas conventional loss functions only support real numbers. Hence, here we adopt the complex-valued version MSE (CV-MSE) for codewords [136, 137, 138], i.e., given $\overrightarrow{c_1} = \overrightarrow{a_1} + \overrightarrow{b_1}i$ and $\overrightarrow{c_2} = \overrightarrow{a_2} + \overrightarrow{b_2}i$, where $c_1, c_2 \in \mathbb{C}^{N\times 1}$,

$$CV\text{-}MSE(\overrightarrow{c_1}, \overrightarrow{c_2}) = \frac{1}{N}\sum_{i=1}^{N}(|a_1[i]-a_2[i]|^2 + |b_1[i]-b_2[i]|^2).$$

In this way, we can obtain a real-number loss to describe the estimation error for complex values. Then, for estimated codewords $\hat{\psi}$ and true codewords ψ^*, the codeword estimation error $\mathbb{L}_C = CV\text{-}MSE(\hat{\psi}, \psi^*)$. The whole forward propagation process is denoted by green arrows in Figure 4.1. On the other hand, the backpropagation (BP) process is already discussed in Section 3.3, and denoted by yellow arrows in Figure 4.1.

4.1.2 The online inference pipeline

The feature extractor is theoretically capable to obtain codeword-independent representations after a proper training process. Hence, thanks to the adversarial learning on the RIS codeword domain by the DANN framework, during online inference, the codewords are unnecessary for location estimation, which meets our requirement theoretically. Figure 4.2 depicts the online inference pipeline of the proposed system. Compared with the training pipeline, the codeword discriminator branch is non-essential, thereby being removed.

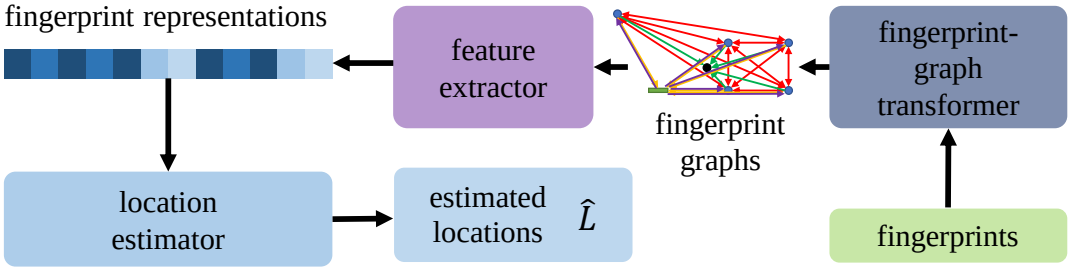

Figure 4.2: The online inference pipeline of our proposed system.

4.2 Building blocks

This section focuses on the implementation details of each part of the proposed system as shown in Figure 4.1.

4.2.1 Fingerprint-graph transformer

We propose an extended version of the fingerprint-graph transformer in our prior work [84] given in Section 3.2.

We first define the localization scenario. Suppose a communication system operates on frequencies $F = \{f_1, f_2, ...\}$. For each $f \in F$, there are several BSs $B_f = \{b_f^1, b_f^2, ...\}$ whose locations $\mathscr{L}(\cdot)$ are known. Besides, the DOI can measure the RSSI for all BSs operating in all frequencies. Finally, there is a RIS in this system whose function is unknown, whereas we know its location. Now we could start to build graphs.

We first consider all the BS, the RIS, and the DOI as vertices in a graph. Then we can construct a sub-graph for every $f \in F$ as shown in the left part of Figure 4.3, including B_f, the RIS, and the DOI. Note that the DOI and the RIS are shared among sub-graphs. The BSs working in different f should be categorized as different kinds of vertices, inducing heterogeneous graphs when $|F| > 1$. Actually, the type of RF signals T in Section 3.2 and the operating frequency F here play the same role, i.e., as discrimination between different sub-graphs. This is because, for different f, the signal propagation laws are still different,

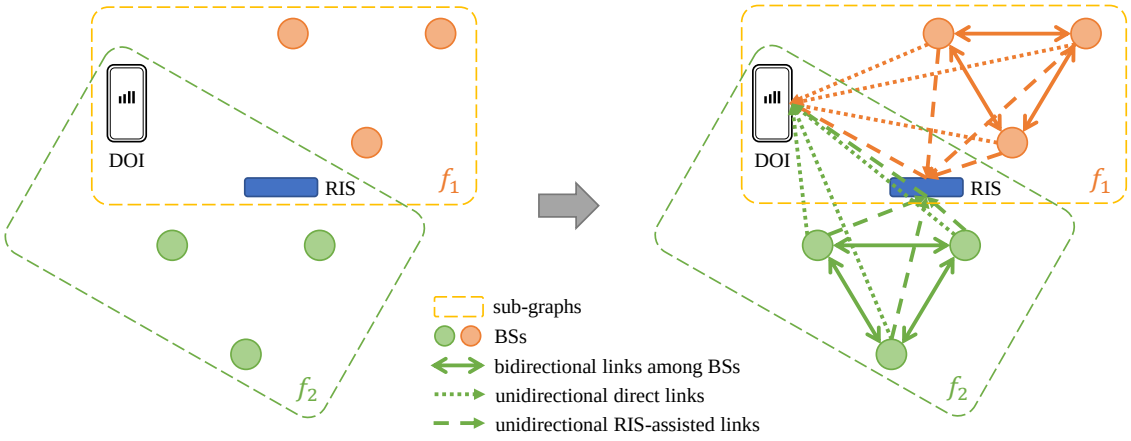

Figure 4.3: A fingerprint-graph transformation example for $F = \{f_1, f_2\}$.

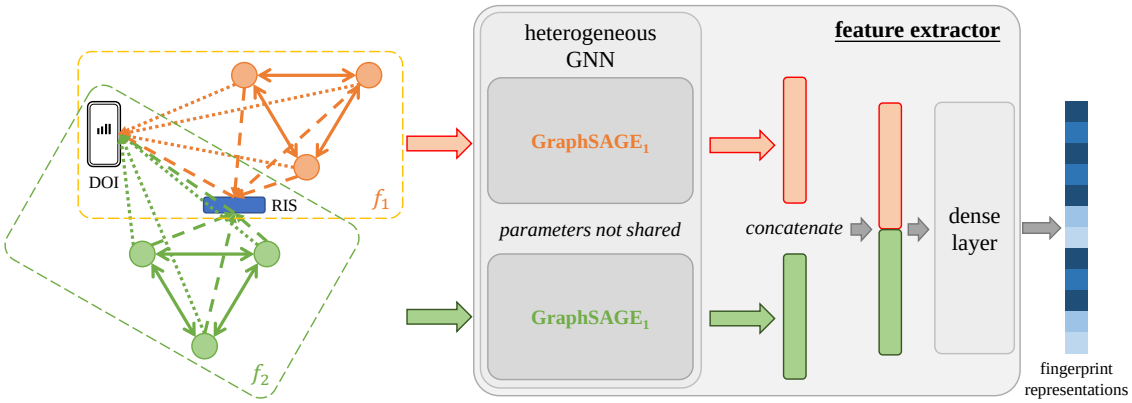

Figure 4.4: The feature extractor for the fingerprint graphs from Figure 4.3.

which also require different sets of model parameters to encode. The vertex features for every BS b is $[RSSI_b, \mathscr{L}(b)]$, where $RSSI_b$ is the RSSI value of b measured by the DOI. We can also assign vertex features of the RIS by $\mathscr{L}(RIS)$. Obviously, we extend the idea in [84] by also considering the RIS as vertices in graphs.

Next, we consider the connectivity between vertices. The right part of Figure 4.3 illustrates the edges among all vertices. Concretely, we adopt the same assumptions in Section 3.2. Therefore, we fully connect B_f within each sub-graph by bidirectional edges. For the direct links, the edges are unidirectional from BSs to the DOI, as the DOI is a measuring device here. Similarly, the RIS is also a (nearly) passive device [139], so the RIS-assisted links are also unidirectional from the BSs to the RIS and from the RIS to the DOI. In this way, these edges represent all possible signal propagations and interferences in this communication system. Additionally, for the edges among BSs and edges from BSs to the RIS, we also assign edge features by LNSM, indicating the theoretical relative strength of signal attenuation in between.

In this way, given arbitrary fingerprints, we can apply the transformation method to generate a corresponding heterogeneous graph containing $|F|$ sub-graphs, which is ready to be fed into the GNN-based feature extractor. The heterogeneous graphs here essentially correspond to the heterogeneity of communication networks coming from different kinds of devices, multiple frequency bands of carriers, etc., which requires many sets of model parameters to encode. Some may argue that this preprocessing method involves locations of BSs and DOI, which still results in additional communication burdens compared with transmitting RIS codewords. However, we can easily represent the locations by 3-axis coordinates, which are much simpler than large CV vectors of codewords. Meanwhile, the locations of BSs and RIS are unlikely to change very often, thus infrequent updates are sufficient. By contrast, we cannot make sure of how often the codeword changes. Hence, the extra transmission of locations is almost negligible.

4.2.2 Feature extractor

The fingerprint graphs generated by the transformer are heterogeneous, thereby we need a heterogeneous GNN-based model to encode these graphs. Specifically, as shown in Figure 4.4, for each sub-graph, we first respectively assign a GraphSAGE [140] model using pooling aggregators by default to obtain the corresponding sub-graph-wise readout by the mean

of latent features of all vertices. Then all sub-graph-wise readouts are concatenated to a vector, which would be fed into a dense layer for the fingerprint representations. We utilize leaky ReLU [141] as the activation functions for the feature extractors, where the negative slope coefficient is -0.02 for all the following experiments.

The reason for choosing the GraphSAGE to encode each sub-graph is we can manually set the aggregation depth for neural message passing [142] and randomly select several paths among all possible ones to reduce the complexity due to potential numerous vertices (BSs) and edges (communication links) in our fingerprint graphs [140].

4.2.3 Location estimator

The fingerprint-graph transformer and the feature extractor cooperate to generate fingerprint representations. To further obtain the location estimation, we need to apply the location estimator to read out the fingerprint representations. In our system, we simply adopt a three-layer perceptron as the location estimator activated by the leaky ReLU [141].

4.2.4 Codeword discriminator

In comparison, we cannot similarly apply an MLP for the codeword discriminator, since ψ^* are CV vectors, which means that the codeword discriminator should support CV outputs. Hence, we implement dense layers that accept both real-value (RV) and CV inputs and outputs CV vectors with the help of the *cplxmodule* library [143]. In this way, we can build a CV-MLP for the codeword discriminator. Additionally, the activation function should also support these CV-adapted layers. Therefore, we utilize the *modReLU* [144] to activate the codeword discriminator in our system, which is a variation of the ReLU designed for pointwise nonlinearity and only manipulates the magnitude of the CV inputs. For $c \in \mathbb{C}$,

$$modReLU(c) = ReLU(|c| + b) \times \frac{c}{|c|}$$

where $b \in \mathbb{R}$ is a bias parameter of the nonlinearity. In other words, b is a threshold to decide whether to make the activated c equal to zero. We set $b = 0.5$ by default.

In Section 3.1.1, we discussed that the codewords describe the phase shifts of all RIS elements. Instead of using complex numbers to represent the codewords, we can alternatively use the RV phase shifts of all RIS elements by radians. Then additional modifications for CV supports are unnecessary for the codeword discriminator. However, the complex numbers are widely used in signal processing and electrical engineering, as they provide convenient representations for the phases and amplitudes of periodic signals [145, 146]. Whereas real numbers are not straightforward to represent this information. Hence, using real numbers to represent codewords would make it difficult for the codeword discriminator to learn from the data and provide reliable adversarial gradients for the DANN framework [147, 148]. Subsequent experiments in Chapter 5 would demonstrate that the CV version of our proposed solution performs better than RV version one.

Chapter 5

Evaluation

This chapter provides the experiment results of our proposed localization system in RIS-enhanced environments. Our experiments fully rely on the DeepMIMO dataset [46]. We would first introduce the settings of the DeepMIMO for our evaluation. Next, the performances of our proposed system are given, including the experiments on the impacts of crucial system parameters to demonstrate the robustness of our method.

5.1 Experimental setup

5.1.1 Configurations of the DeepMIMO dataset

For our experiments, we use the dataset generated by the DeepMIMO [46]. The dataset is a generative dataset based on ray-tracing measurements, which is semi-customizable by user specifications on the system parameters [46]. Our experiments are fully conducted in the O1 (Outdoor 1) scenario of the DeepMIMO as shown in Figure 5.1. Concretely, we are interested in a modern wireless communication network containing both 4G and 5G. Hence, among all available operating frequency choices in the DeepMIMO, we consider 3.4 GHz, 3.5 GHz (4G LTE Band 42 [149]), and 28 GHz (5G NR FR2 Band n257 [150]). We activate six BSs in the DeepMIMO simulation tool, where the `BS5` plays the role of RIS using the same method discussed in [127]. We specify the BSs, RIS, and DOI, all equipped with MIMO antennas. Detailed settings are listed in Table 5.1.

Considering the positions of BSs and obstacles, we specify the test areas of DOI containing both LoS and NLoS regions for `BS18` as shown in Figure 5.2. Taking the testing point (`ROW 1268, COL 91`) as the center, we can equidistantly expand the boundaries to obtain test

Table 5.1: DeepMIMO settings for our experiments

Operating frequency	3.4 GHz, 3.5 GHz, 28 GHz
Activated BSs	1, 2, 3, 4, 5 (RIS), 18
Antennas of BSs and RIS	4×4
Antennas of DOI	2×2
Bandwidth	200 MHz
The number of OFDM sub-carriers	512

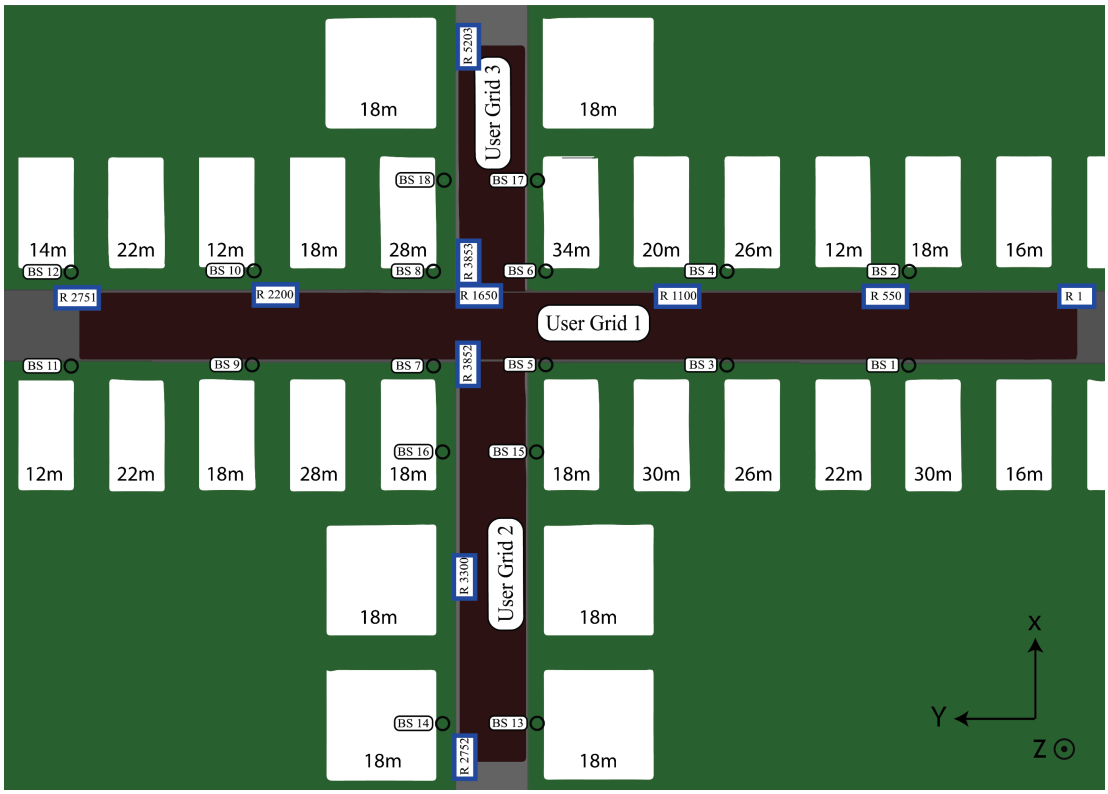

Figure 5.1: The O1 scenario in the DeepMIMO [46].

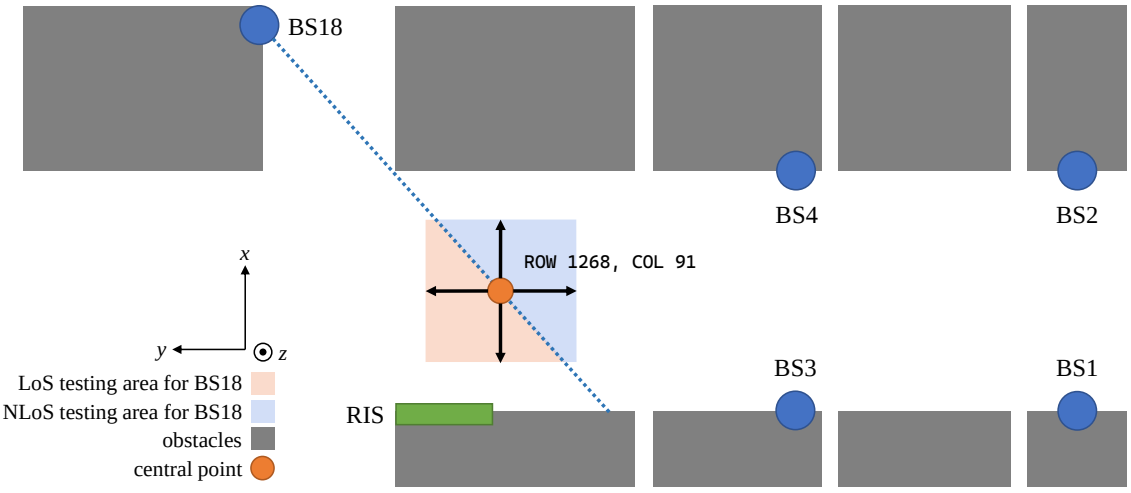

Figure 5.2: An top-view sketch map of the experimental area using the O1 scenario in the DeepMIMO.

areas, where half are LoS whereas the other half are NLoS for BS18. We test the localization performances in the testing area.

5.1.2 Experimental parameters

With the help of the DeepMIMO dataset, we set up the localization scenario for evaluation. In this section, we further specify the experimental parameters for the following data generation. Concretely, we consider the number of codewords C and the size of test areas A. Additionally, we test the robustness of our localization system by manually adding extra additive white Gaussian noise (AWGN) $\mathbf{N} \sim \mathcal{N}(0, \sigma^2)$ [151] to the calculated RSSI. To summarize, we can use a parameter set $\{C, A, \sigma\}$ to describe the experiments.

In Section 3.1.2, we mentioned that most RIS prototypes only support $C = 2/4$ [124, 125], several could realize $C = 64$ [126]. For forward-looking results in this thesis, we set an initial resolution of 15°, and $R = 180°$ here. Then we have $C = 144$. Besides, we set $A = 51.84$ and $\sigma = 0$ by default. We will first apply the default experimental parameter set $\{C = 144, A = 51.84, \sigma = 0\}$ for evaluation to check the feasibility of our solution. Then we respectively change one parameter and keep the other two unchanged to separately investigate the impact of these three experimental parameters on the proposed solution.

5.1.3 Dataset generation and model implementation

Given the experimental parameters $\{C, A, \sigma\}$, we first apply the formulas in Section 3.1.2 to generate the codebook $\mathscr{C}$ using C. Next, we input the setting in Table 5.1 and A to the DeepMIMO dataset generator by modifying the `parameters.m` file, which is one of the supporting scripts provided by the DeepMIMO [46]. Then, after running another script `DeepMIMO_Dataset_Generator.m`, we would obtain the channel matrices among BSs, DOI, and the RIS, and corresponding DOI locations, which are the labels in our problem. Finally, applying the RSSI calculation pipeline given in Section 3.1.3 using $\mathscr{C}$ and σ, we could calculate the RSSI values, thereby constructing fingerprint datasets for evaluation. Here, we can simply assume that for all BSs we activate, $G_p = 40\ dB$ and $L_p = 20\ dB$.

For the train/test split, we first randomly sample 80% available codewords, then randomly sample 80% locations in the testing area. Then for a data point, if both its location and codeword are sampled, it belongs to training datasets. Conversely, if neither its location nor codeword is sampled, it would be assigned to test datasets. In other words, we test the estimation model using fingerprints of unknown locations and unknown codewords. In this way, the test data for online inference challenge the generalization capabilities for both locations and codewords.

We implemented the proposed system using PyTorch [152] with the help of DGL library [153] to build graph models for the feature extractor. All the neural network layers contain 64-dimensional latent features. The whole system is trained using Adam optimizer [154] with an initial learning rate of 0.01.

5.1.4 Oracle and baseline cases for evaluation

One of the RQs in this thesis is to design a reasonable comparison scheme to evaluate our system, as few previous works have noticed the problem we discussed, much less corre-

sponding localization performances using the DeepMIMO dataset. Hence, an alternative evaluation method is required.

The goal of this thesis is to solve the codeword domain generalization problem. Ideally, the best possible model completely eliminates the negative effects of the DG issue. Hence, we could simulate this case by manually setting $C = 1$, which means there is only one possible RIS codeword. Then there would be no fingerprint shifts because the codeword domains for both training and testing are always identical. We call this case the *oracle case*, because this is the theoretical upper bound of our localization system.

On the other hand, the worst possible model completely fails to decouple the correlation between fingerprints and corresponding codewords, which is equivalent to applying no adversarial learning. In this case, for the same location, there will be multiple different fingerprints corresponding to it due to different RIS codewords. The localization system will have difficulties learning effective fingerprint representations from such data. We call this case the *baseline case*, since this is the theoretical lower bound of the proposed system.

We implemented these two cases in our solution simply by removing the codeword discriminator component because neither of them requires adversarial learning on the codewords. In the following experiments, all experiments parameters $\{C, A, \sigma\}$ in the *baseline case* remain the same as the under-test scenario. Whereas for the *oracle case*, C always equals 1, so keeping only A and σ the same would be sufficient.

If the performances of our solution are closer to the *oracle case* rather than the *baseline case*, then we tend to conclude that our model has successfully learned codeword-independent representations of fingerprints. This marks the accomplishment of the main goal in this thesis, i.e., designing a localization system with RIS codeword domain generalization capabilities for modern RIS-enhanced networks.

5.2 Performance evaluation

In this section, we apply the default experimental parameter set $\{C = 144, A = 51.84, \sigma = 0\}$ for evaluation to check the feasibility of our solutions. First, Figure 5.3 visualizes fingerprint shifts in the testing area. RSSI values at the same location can differ by up to 52.1 dB, which illustrates that the fingerprints are sensitive to codeword change, thereby demonstrating the feasibility of using the DeepMIMO for our evaluation.

To visualize the location estimation errors, we apply bar charts, where each bar's middle line is the mean squared loss (MSE) [135], and its height equals two times the standard

Table 5.2: The localization errors of LoS (orange) and NLoS (blue) areas for the oracle/baseline cases, and the CV/RV versions of our solution when $\{C = 144, A = 51.84, \sigma = 0\}$

	LoS		*NLoS*	
	MSE	Var	MSE	Var
Oracle case	0.045	0.002	0.047	0.003
Our solution (CV ver.)	0.050	0.007	0.090	0.018
Our solution (RV ver.)	0.125	0.010	0.199	0.013
Baseline case	2.053	0.958	2.956	1.033

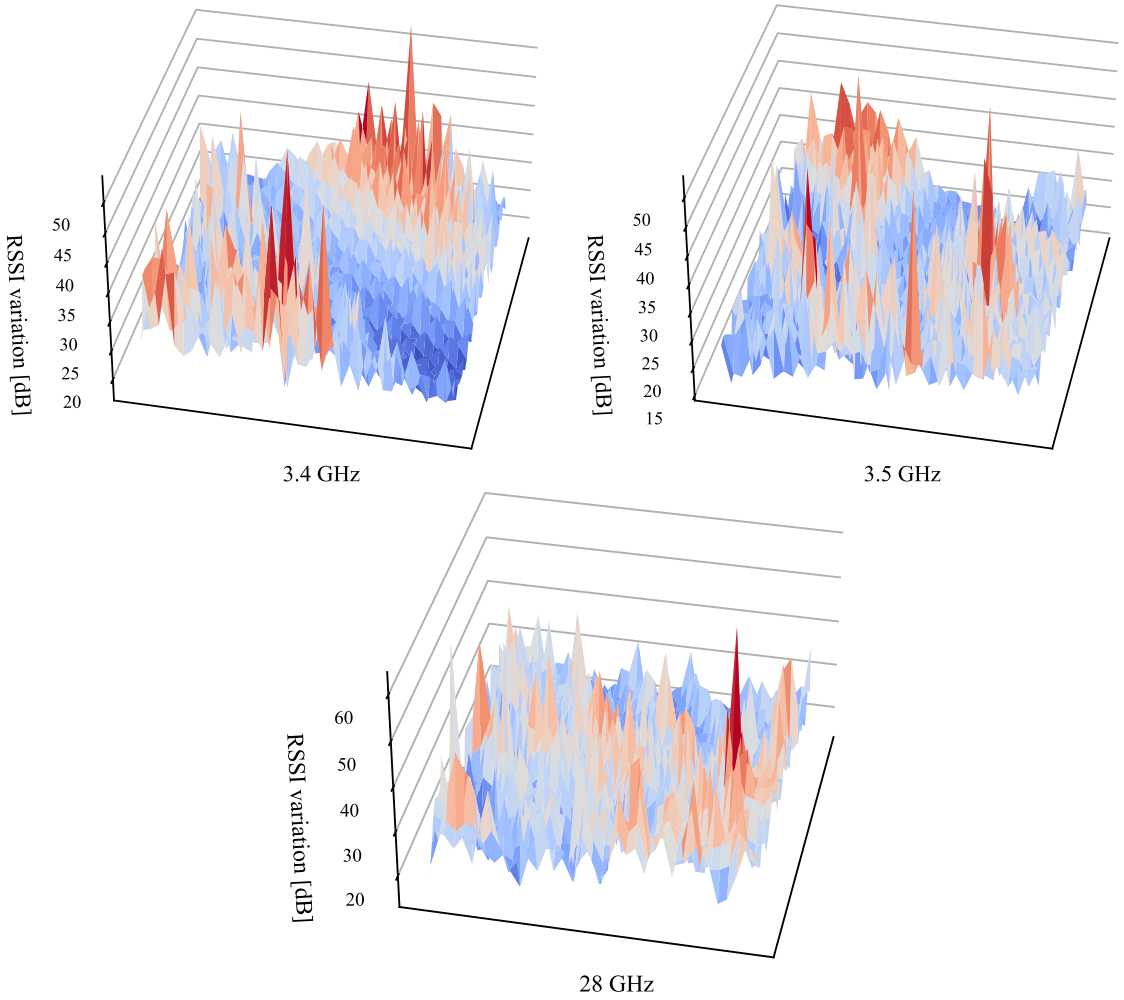

Figure 5.3: Fingerprint shifts (dB) in the testing area for three frequencies. The height means the maximum RSSI shift at that position when $\{C = 144, A = 51.84, \sigma = 0\}$.

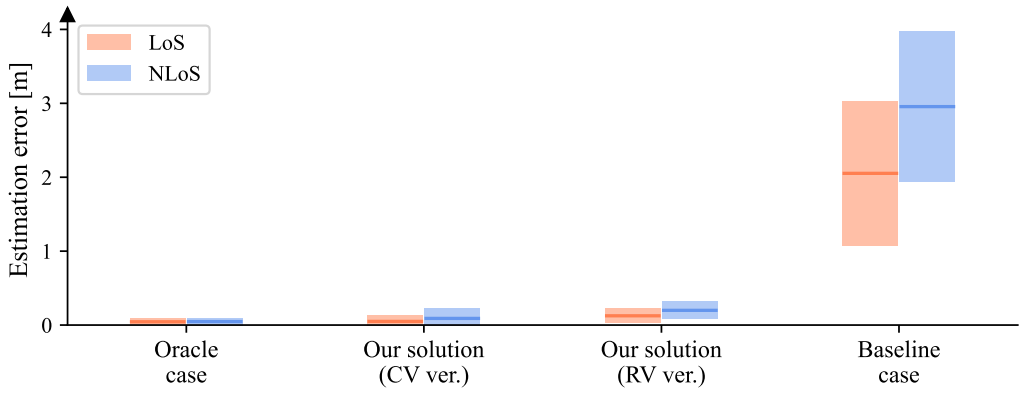

Figure 5.4: The location estimation errors of both LoS and NLoS areas for the oracle/baseline cases, and the CV/RV versions of our solution when $\{C = 144, A = 51.84, \sigma = 0\}$. Given the squared error set e, for each bar, the middle line is $mean(e)$, i.e., MSE, and its height ranges from MSE-$std(e)$ to MSE+$std(e)$.

deviation over the squared loss of all concerned testing points. Most of the errors are within this interval, which could be used to observe the prediction stability of each method. Besides, the numerical results are also provided in tables.

In Section 4.2.4, we claimed that CV codewords and codeword discriminators perform better than the RV version. Hence, both CV and RV versions of the proposed system are considered here. Then, for the experimental parameter set $\{C = 144, A = 51.84, \sigma = 0\}$, the results of the oracle/baseline cases and our solution are depicted in Figure 5.4. First, the location estimation errors of our solution are much closer to the oracle case rather than to the baseline case, which illustrates the representations extracted by our method accommodate different codewords. Therefore, we can conclude that the feature extractor in our system can obtain codeword-independent representations of fingerprints. Our solution achieves our main goal, i.e., realizing a localization system with RIS codeword domain generalization capabilities. Then, compared with the RV version, the desired CV version of our method obtains smaller errors, which demonstrates our analysis in Section 4.2.4. Finally, the detailed numerical results are listed in Table 5.2. We manage to obtain centimeter-level errors, which meets the expectation of the future 6G [10].

5.3 Impact of experimental parameters

In this section, we investigate the impact of experiment parameters on the model performance by respectively changing one parameter and keeping another two fixed.

5.3.1 The number of codewords

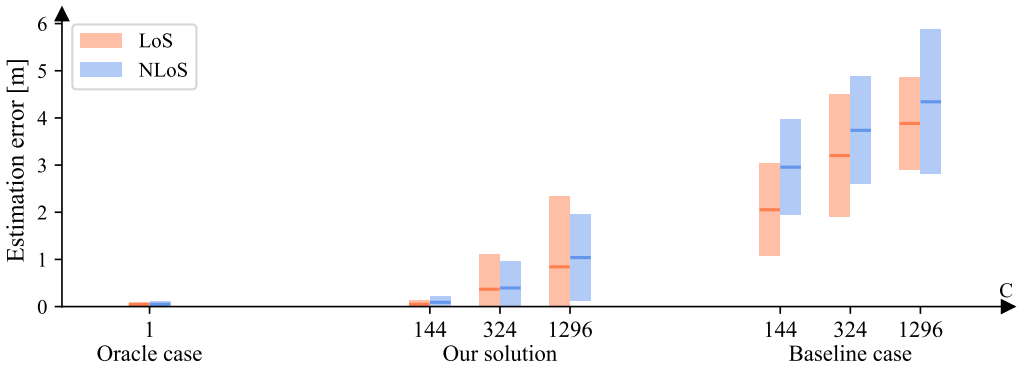

Figure 5.5: The location estimation errors of both LoS and NLoS areas for the oracle/baseline cases and our solution when $\{C = 144/324/1296, A = 51.84, \sigma = 0\}$. Given the squared error set e, for each bar, the middle line is $mean(e)$, i.e., MSE, and its height ranges from MSE-$std(e)$ to MSE+$std(e)$. Note that for the oracle case, C always equals to 1 for reference.

The resolution of the RIS in previous experiments is set to 15°, inducing $C = 144$, which is already far more than current RIS prototypes. Here, we further enlarge C for more harsh scenarios by assuming smaller resolutions, i.e., 10° and 5°, obtaining $C = 324$ and $C = 1296$, respectively. Note that for the oracle case, C always equals 1 for the purpose of benchmarking. Other parameters remain unchanged in this section, i.e., $A = 51.84$ and $\sigma = 0$. Figure 5.5 illustrates the location estimation errors of the oracle/baseline cases and our solution. As C increases, both the MSE and the standard deviations become larger, as the feature extractor has to adapt to more codewords at the same time. Nonetheless, the performance

Table 5.3: The localization errors of LoS (orange) and NLoS (blue) areas for the oracle/baseline cases and our solution when $\{C = 144/324/1296, A = 51.84, \sigma = 0\}$. Note that for the oracle case, C always equals to 1 for reference.

Oracle case ($C = 1$)					**Our solution**				**Baseline case**			
LoS		*NLoS*		C	*LoS*		*NLoS*		*LoS*		*NLoS*	
MSE	Var	MSE	Var		MSE	Var	MSE	Var	MSE	Var	MSE	Var
0.045	0.002	0.047	0.003	144	0.050	0.007	0.090	0.018	2.053	0.958	2.956	1.033
				324	0.368	0.554	0.394	0.312	3.202	1.692	3.737	1.296
				1296	0.843	2.231	1.039	0.845	3.882	0.958	4.342	2.333

Table 5.4: The localization errors of LoS (orange) and NLoS (blue) areas for the oracle/baseline cases and our solution when $\{C = 144, A = 51.84/92.16/144.00, \sigma = 0\}$.

	Oracle case				**Our solution**				**Baseline case**			
$A\,[m^2]$	*LoS*		*NLoS*		*LoS*		*NLoS*		*LoS*		*NLoS*	
	MSE	Var	MSE	Var	MSE	Var	MSE	Var	MSE	Var	MSE	Var
51.84	0.045	0.002	0.047	0.003	0.050	0.007	0.090	0.018	2.053	0.958	2.956	1.033
92.16	0.045	0.002	0.047	0.002	0.113	0.048	0.119	0.056	2.268	1.086	3.002	0.991
144.00	0.046	0.002	0.050	0.003	0.205	0.081	0.236	0.170	2.410	0.879	3.013	1.211

of our solution is still closer to the oracle case than the baseline case. When $C = 1296$, our solution even performs better than the baseline case when $C = 144$. These performances all demonstrate that our solution maintains its capability of being codeword-independent when C becomes much larger than current prototypes, which proves the robustness to the number of available codewords. Table. 5.3 presents the corresponding numerical results.

5.3.2 Size of testing areas

The sizes of the testing area A in previous experiments are all 51.84 m^2. In this section, we set it to $A = 51.84/92.16/144.00$ and investigate its impact on the system performance. We keep $C = 144$ and $\sigma = 0$.

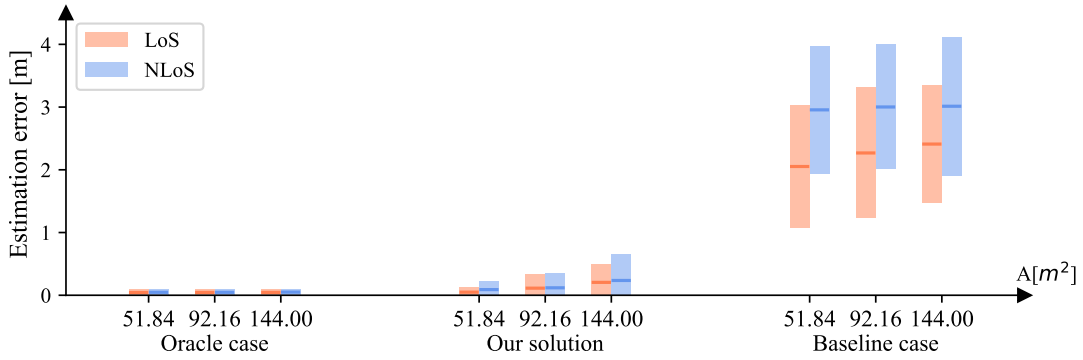

Figure 5.6: The location estimation errors of both LoS and NLoS areas for the oracle/baseline cases and our solution when $\{C = 144, A = 51.84/92.16/144.00, \sigma = 0\}$. Given the squared error set e, for each bar, the middle line is $mean(e)$, i.e., MSE, and its height ranges from MSE-$std(e)$ to MSE+$std(e)$.

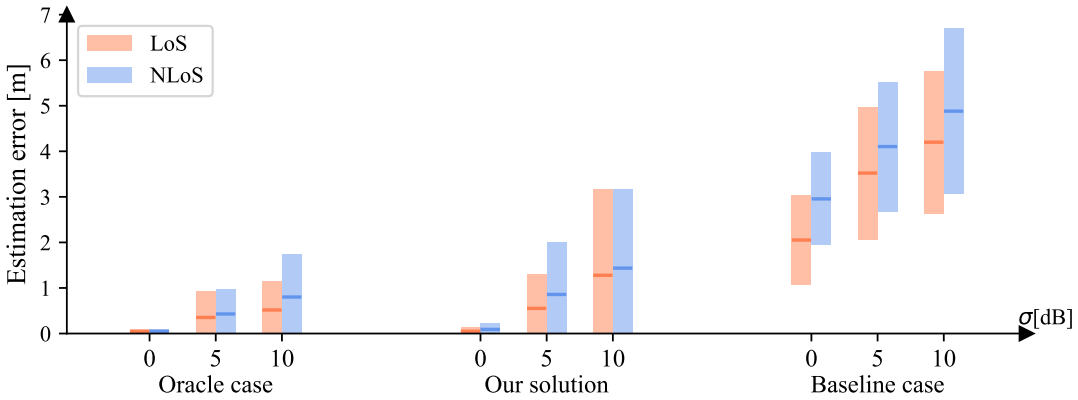

Figure 5.7: The location estimation errors of both LoS and NLoS areas for the oracle/baseline cases and our solution when $\{C = 144, A = 51.84, \sigma = 0/5/10\}$. Given the squared error set e, for each bar, the middle line is $mean(e)$, i.e., MSE, and its height ranges from MSE-$std(e)$ to MSE+$std(e)$.

Table 5.5: The localization errors of LoS (orange) and NLoS (blue) areas for the oracle/baseline cases and our solution when $\{C = 144, A = 51.84, \sigma = 0/5/10\}$.

	Oracle case				**Our solution**				**Baseline case**			
$\sigma[dB]$	*LoS*		*NLoS*		*LoS*		*NLoS*		*LoS*		*NLoS*	
	MSE	Var	MSE	Var	MSE	Var	MSE	Var	MSE	Var	MSE	Var
0	0.045	0.002	0.047	0.003	0.050	0.007	0.090	0.018	2.053	0.958	2.956	1.033
5	0.353	0.315	0.430	0.298	0.552	0.556	0.858	1.280	3.522	2.118	4.104	2.022
10	0.517	0.402	0.802	0.910	1.279	3.569	1.435	2.994	4.200	2.430	4.881	3.291

Figure 5.6 depicts the performance of the oracle/baseline cases and our solution. The performance of both the oracle and baseline cases almost remain the same as A increase, whereas the location estimation error of our solution grows faster than the reference cases. However, this phenomenon is totally reasonable because our solution needs to perform adversarial learning on more testing points, which means that it becomes harder for the feature extractor to obtain codeword-independent representations as the number of fingerprints grows. In comparison, both the oracle and baseline cases perform no adversarial learning, hence their performance remains nearly unchanged. Nevertheless, our solution still achieves significantly better results than the baseline cases and closer performances to the oracle cases. Therefore, we conclude that our solution is robust against changes in the size of the testing areas. All corresponding numerical results are presented in Table. 5.4.

5.3.3 Strength of the AWGN

In this section, we manually add extra AWGN to the calculated RSSI to validate the stability of our solution. The AWGN here refers to some miscellaneous noises, including fading, and polarization mismatch [155, 156], which are not discussed in our link budget model and may be modeled by Gaussian processes [157, 158]. The strength of the AWGN $\mathbf{N}$ (dB) is controlled by the standard deviation σ for $\mathbf{N} \sim \mathcal{N}(0, \sigma^2)$. We investigate the performances when $\sigma = 0/5/10$. We keep $C = 144$ and $A = 51.84$.

Figure 5.7 shows the location estimation errors of the oracle/baseline cases and our solution for different AWGN strength levels. Similarly, our solution performs much better than the baseline cases and is closer to the performances of the oracle cases. Therefore, our solution is also robust against extra AWGN, which indicates that the model can maintain certain stability of location inference. Note that the extra AWGN here is not caused by obstacles in the environments, which are already recorded by the channel matrices in the DeepMIMO dataset, and implicitly represented by the calculated RSSI. We investigate AWGN here for simulation on the device noises to further refine our results. Numerical results are presented in Table 5.5.

5.4 Discussion

In this section, we first introduced the DeepMIMO dataset and its settings for our evaluation. We designed oracle and baseline cases for reasonable comparison with our solution. The method to calculate the RSSI using DeepMIMO was elaborated in detail. In the first experiments, we compared the performance of oracle/baseline cases and our solution (both CV and RV versions). We concluded that our solution managed to extract codeword-independent features. Experiment results show that CV codeword discriminators perform better than the RV ones, which demonstrates that the CV-MLP is the better choice in our system. We further investigated the impacts of different experimental parameters, respectively. All comparison results indicated that our solution is robust against the number of codewords, size of the testing area, and extra AWGN. Our solution in all experiments performs much closer to the oracle cases instead of the baseline cases, which verifies that our solution achieves the main goal of this thesis, i.e., realizing localization with codeword domain generalization capability in RIS-enhanced environments.

Chapter 6

Conclusion

We conclude the thesis in this chapter. Except for a brief summary of this thesis, we also provide our reflections on the experiments.

6.1 Summary

In this thesis, we first investigated the localization problem in modern RIS-enhanced wireless communication networks. Model-driven methods usually rely on sophisticated communication models and set specific optimization goals to realize accurate localization. However, the complexity of these models grows fast as the number of BSs and DOI increase, which restricts their application scopes. Although several data-driven methods have realized this problem and performed fingerprint-based localization, they unrealistically assumed that the RIS is sorely deployed for localization rather than more important tasks, such as enlarging the signal coverage. Even worse, both model- and data-driven methods share the same drawbacks that the RIS codewords are required for location inference, which induce huge additional communication burdens.

These observations inspired us to design a localization solution in RIS-enhanced environments, which can decouple the correlation between fingerprints and codewords. In this way, regardless of actual RIS functions, the localization system could provide reliable estimations without RIS codewords. In other words, the goal of this thesis is to design a localization solution with the capability of RIS codeword domain generalization. The discussion on the techniques for domain generalization revealed that we can achieve this by applying domain adversarial learning on the codewords using the DANN framework [113] to extract codeword-independent representations of fingerprints.

We proposed a localization solution based on DANN that can adopt RIS codeword changes. Concretely, we first designed a pre-processing step to transform the fingerprints into graphs for the heterogeneous GNN-based feature extractor to make full use of the non-Euclidean features of the fingerprints, especially when they are collected in environments with multiple operating frequencies. Next, during the offline training stage, the feature extractor generates the representations of the fingerprint graphs and feeds them into both the MLP-based location estimator and CV-MLP codeword discriminator for location and codeword estimation, respectively. Then, the gradients from the codeword discriminator will be reversed by a gradient reversal layer to perform adversarial learning on the codeword

domain during backpropagation to ensure the distributions of the representations over different codewords are as similar as possible. In this way, the feature extractor can theoretically extract codeword-independent features for location inference, thereby the codewords will not be necessary for the online inference stage.

We evaluated our system using the O1 scenario in the DeepMIMO dataset [46]. We defined oracle and baseline cases for a fair comparison with our solution and elaborated the method to calculate the RSSI using the channel matrices given in the DeepMIMO simulation tool. Our evaluation results firstly prove that CV codeword discriminators perform better than the RV ones, which validates our system design. Then we investigated the impacts of the experiment parameters, including the number of codewords, size of the testing area, and strength of additional AWGN. All comparison results indicate that our solution performs much closer to the oracle cases rather than the baseline cases, which demonstrates the robustness of our method against the change of the concerned parameters. Additionally, all these results prove that the proposed solution achieves the goal of this thesis, i.e., realizing a localization system in RIS-enhanced environments with the capabilities of RIS codeword domain generalization.

Here, we explain how our research questions were answered.

RQ1. How RIS codeword domain generalization can be achieved? What are the specific pros and cons of the candidate DG solutions for our problem?

We adopted the DANN framework to perform adversarial learning on the RIS codeword domain to obtain codeword-independent representations of the fingerprints to solve the DG problem. We fully discussed this topic in Section 2.3 and 3.3.

To achieve codeword domain generalization, both contrastive learning and DANN would be sufficient as discussed in Section 2.3. Contrastive learning needs to generate more data pairs for training compared with the DANN. Whereas the DANN requires explicit domain labels to perform adversarial learning, but not for contrastive learning. In our problem, the domain labels are just codewords. Therefore, the disadvantage of the DANN is invalid here. Therefore, we utilize the DANN as the solution in this thesis. All evaluation results shown in Chapter 5 indicate that the proposed solution using DANN to perform adversarial learning on the codeword domain manages to solve the DG problem since we obtain performances much closer to the oracle cases rather than baseline cases.

RQ2. Which DNN architecture is proper to encode the fingerprints, especially in modern communication systems usually containing 4G and 5G operating on different frequency bands?

The GNN-based models perform better than other commonly used data-driven models, such as k-NN, MLP, and CNN. This is because the fingerprints collected in environments with multiple operating frequencies usually show non-Euclidean features, which could be efficiently utilized by GNN-based models [84]. Hence, we designed a fingerprint-graph transformation method in Section 4.2.1, and a heterogeneous GNN-based feature extractor in Section 4.2.2 to make full use of the fingerprints.

RQ3. How to compare the performance of our solution with respect to the state of the art, when no related works realize localization systems with codeword domain generalization capabilities in RIS-enhanced environments?

We defined oracle and baseline cases for fair comparisons in Chapter 5.1.4. Concretely, the oracle case denotes the best possible model we can obtain, i.e, the model could completely eliminate the RIS codeword domain shift problem. Hence, we always set the number of codewords to 1 for the oracle cases. In this way, the fingerprints would not change due to different codewords. The baseline case is the worst model we may obtain, i.e., the model fails to perform any adversarial learning on the RIS codeword domain. For both oracle and baseline cases, we can implement them by removing the codeword discriminator component in our solution because neither of them needs adversarial learning. Then during the evaluation, except for the number of codewords for the oracle cases, we always apply the same set of experimental parameters to the baseline/oracle cases and our solution for comparison. If our solution performs closer to the oracle cases rather than the baseline cases, we can conclude that the proposed method achieves RIS codeword domain generalization.

6.2 Reflection

An apparent issue in this thesis is that our evaluation fully relies on simulations using the DeepMIMO dataset. We could not obtain the RIS hardware, since currently, it is still under development, even the number of RIS prototypes [124, 125, 126] is small. Hence, in future works, if a RIS is available, we can design experiments, collect real-world data, and perform the evaluation. Such results will be more convincing.

Another issue is the number of codewords. Although we set a far larger number of the codewords compared to the current RIS prototypes [124, 125, 126] and achieved performances close to the oracle case, there is still a chance that our model fails to deal with more codewords when the resolution is extremely small. If it happens, then we may have to require the codewords for location inference. Nevertheless, considering the fact that most prototypes only support $C = 2/4$ [124, 125], we are confident that our solution has strong practicality in modern wireless communication systems.

Bibliography

[1] Allen Welkie, Longfei Shangguan, Jeremy Gummeson, Wenjun Hu, and Kyle Jamieson. Programmable radio environments for smart spaces. In *Proceedings of the 16th ACM Workshop on Hot Topics in Networks*, pages 36–42, 2017.

[2] Ying-Chang Liang, Jie Chen, Ruizhe Long, Zhen-Qing He, Xianqi Lin, Chenlu Huang, Shilin Liu, Xuemin Sherman Shen, and Marco Di Renzo. Reconfigurable intelligent surfaces for smart wireless environments: channel estimation, system design and applications in 6g networks. *Science China Information Sciences*, 64(10):1–21, 2021.

[3] Ertugrul Basar. Reconfigurable intelligent surface-based index modulation: A new beyond mimo paradigm for 6g. *IEEE Transactions on Communications*, 68(5):3187–3196, 2020.

[4] Bo O Zhu, Junming Zhao, and Yijun Feng. Active impedance metasurface with full 360 reflection phase tuning. *Scientific reports*, 3(1):1–6, 2013.

[5] Emil Björnson, Henk Wymeersch, Bho Matthiesen, Petar Popovski, Luca Sanguinetti, and Elisabeth de Carvalho. Reconfigurable intelligent surfaces: A signal processing perspective with wireless applications. *IEEE Signal Processing Magazine*, 39(2):135–158, 2022.

[6] Chongwen Huang, Sha Hu, George C Alexandropoulos, Alessio Zappone, Chau Yuen, Rui Zhang, Marco Di Renzo, and Merouane Debbah. Holographic mimo surfaces for 6g wireless networks: Opportunities, challenges, and trends. *IEEE Wireless Communications*, 27(5):118–125, 2020.

[7] Yuanwei Liu, Xiao Liu, Xidong Mu, Tianwei Hou, Jiaqi Xu, Marco Di Renzo, and Naofal Al-Dhahir. Reconfigurable intelligent surfaces: Principles and opportunities. *IEEE Communications Surveys & Tutorials*, 23(3):1546–1577, 2021.

[8] Marco Di Renzo, Merouane Debbah, Dinh-Thuy Phan-Huy, Alessio Zappone, Mohamed-Slim Alouini, Chau Yuen, Vincenzo Sciancalepore, George C Alexandropoulos, Jakob Hoydis, Haris Gacanin, et al. Smart radio environments empowered by reconfigurable ai meta-surfaces: An idea whose time has come. *EURASIP Journal on Wireless Communications and Networking*, 2019(1):1–20, 2019.

[9] Hadeel Elayan, Osama Amin, Raed M Shubair, and Mohamed-Slim Alouini. Terahertz communication: The opportunities of wireless technology beyond 5g. In *2018 International Conference on Advanced Communication Technologies and Networking (CommNet)*, pages 1–5. IEEE, 2018.

[10] Mostafa Zaman Chowdhury, Md Shahjalal, Shakil Ahmed, and Yeong Min Jang. 6g wireless communication systems: Applications, requirements, technologies, challenges, and research directions. *IEEE Open Journal of the Communications Society*, 1:957–975, 2020.

[11] Philipp Hillger, Marcel van Delden, Udaya Sampath Miriya Thanthrige, Aya Mostafa Ahmed, Jonathan Wittemeier, Khaled Arzi, Marcel Andree, Benedikt Sievert, Werner Prost, Andreas Rennings, et al. Toward mobile integrated electronic systems at thz frequencies. *Journal of Infrared, Millimeter, and Terahertz Waves*, 41(7):846–869, 2020.

[12] Anthony Ngozichukwuka Uwaechia and Nor Muzlifah Mahyuddin. A comprehensive survey on millimeter wave communications for fifth-generation wireless networks: Feasibility and challenges. *IEEE Access*, 8:62367–62414, 2020.

[13] Ahmed Alkhateeb, Omar El Ayach, Geert Leus, and Robert W Heath. Channel estimation and hybrid precoding for millimeter wave cellular systems. *IEEE journal of selected topics in signal processing*, 8(5):831–846, 2014.

[14] Abdelrahman Taha, Muhammad Alrabeiah, and Ahmed Alkhateeb. Deep learning for large intelligent surfaces in millimeter wave and massive mimo systems. In *2019 IEEE Global communications conference (GLOBECOM)*, pages 1–6. IEEE, 2019.

[15] Jiguang He, Henk Wymeersch, Tachporn Sanguanpuak, Olli Silvén, and Markku Juntti. Adaptive beamforming design for mmwave ris-aided joint localization and communication. In *2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pages 1–6. IEEE, 2020.

[16] Roy Karasik, Osvaldo Simeone, Marco Di Renzo, and Shlomo Shamai Shitz. Beyond max-snr: Joint encoding for reconfigurable intelligent surfaces. In *2020 IEEE International Symposium on Information Theory (ISIT)*, pages 2965–2970. IEEE, 2020.

[17] Mohammed Aladsani, Ahmed Alkhateeb, and Georgios C Trichopoulos. Leveraging mmwave imaging and communications for simultaneous localization and mapping. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4539–4543. IEEE, 2019.

[18] Hadi Sarieddeen, Nasir Saeed, Tareq Y Al-Naffouri, and Mohamed-Slim Alouini. Next generation terahertz communications: A rendezvous of sensing, imaging, and localization. *IEEE Communications Magazine*, 58(5):69–75, 2020.

[19] Theodore S Rappaport, Yunchou Xing, Ojas Kanhere, Shihao Ju, Arjuna Madanayake, Soumyajit Mandal, Ahmed Alkhateeb, and Georgios C Trichopoulos. Wireless communications and applications above 100 ghz: Opportunities and challenges for 6g and beyond. *IEEE access*, 7:78729–78757, 2019.

[20] Demos Serghiou, Mohsen Khalily, Tim WC Brown, and Rahim Tafazolli. Terahertz channel propagation phenomena, measurement techniques and modeling for 6g wireless communication applications: a survey, open challenges and future research directions. *IEEE Communications Surveys & Tutorials*, 2022.

[21] Frank Van Diggelen and Per Enge. The world's first gps mooc and worldwide laboratory using smartphones. In *Proceedings of the 28th international technical meeting of the satellite division of the institute of navigation (ION GNSS+ 2015)*, pages 361–369, 2015.

[22] Teng Ma, Yue Xiao, Xia Lei, Wenhui Xiong, and Yuan Ding. Indoor localization with reconfigurable intelligent surface. *IEEE Communications Letters*, 25(1):161–165, 2020.

[23] Qianru Cheng, Liyan Li, Ming-Min Zhao, and Min-Jian Zhao. Cooperative localization for reconfigurable intelligent surface-aided mmwave systems. In *2022 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1051–1056. IEEE, 2022.

[24] Arlind Billa, Ibraheem Shayea, Abdulraqeb Alhammadi, Qazwan Abdullah, and Mardeni Roslee. An overview of indoor localization technologies: Toward iot navigation services. In *2020 IEEE 5th International Symposium on Telecommunication Technologies (ISTT)*, pages 76–81. IEEE, 2020.

[25] Luca Barbieri, Mattia Brambilla, Andrea Trabattoni, Stefano Mervic, and Monica Nicoli. Uwb localization in a smart factory: Augmentation methods and experimental assessment. *IEEE Transactions on Instrumentation and Measurement*, 70:1–18, 2021.

[26] Timothy Liu, Matthew Carlberg, George Chen, Jacky Chen, John Kua, and Avideh Zakhor. Indoor localization and visualization using a human-operated backpack system. In *2010 International Conference on Indoor Positioning and Indoor Navigation*, pages 1–10. IEEE, 2010.

[27] Waltenegus Dargie and Christian Poellabauer. *Fundamentals of wireless sensor networks: theory and practice*. John Wiley & Sons, 2010.

[28] Zheng Yang, Zimu Zhou, and Yunhao Liu. From rssi to csi: Indoor localization via channel response. *ACM Computing Surveys (CSUR)*, 46(2):1–32, 2013.

[29] Faheem Zafari, Athanasios Gkelias, and Kin K Leung. A survey of indoor localization systems and technologies. *IEEE Communications Surveys & Tutorials*, 21(3):2568–2599, 2019.

[30] John S Seybold. *Introduction to RF propagation*. John Wiley & Sons, 2005.

[31] Ahmed Elzanaty, Anna Guerra, Francesco Guidi, and Mohamed-Slim Alouini. Reconfigurable intelligent surfaces for localization: Position and orientation error bounds. *IEEE Transactions on Signal Processing*, 69:5386–5402, 2021.

[32] Davide Dardari, Nicoló Decarli, Anna Guerra, and Francesco Guidi. Los/nlos nearfield localization with a large reconfigurable intelligent surface. *IEEE Transactions on Wireless Communications*, 2021.

[33] Gregory G Raleigh and John M Cioffi. Spatio-temporal coding for wireless communication. *IEEE Transactions on communications*, 46(3):357–366, 1998.

[34] Arogyaswami J Paulraj, Dhananjay A Gore, Rohit U Nabar, and Helmut Bolcskei. An overview of mimo communications-a key to gigabit wireless. *Proceedings of the IEEE*, 92(2):198–218, 2004.

[35] Gordon L Stuber, John R Barry, Steven W Mclaughlin, Ye Li, Mary Ann Ingram, and Thomas G Pratt. Broadband mimo-ofdm wireless communications. *Proceedings of the IEEE*, 92(2):271–294, 2004.

[36] Derrick Wing Kwan Ng, Ernest S Lo, and Robert Schober. Energy-efficient resource allocation in ofdma systems with large numbers of base station antennas. *IEEE Transactions on Wireless Communications*, 11(9):3292–3304, 2012.

[37] Martin Sauter. *From GSM to LTE: An Introduction to Mobile Networks and Mobile Broadband*. John Wiley & Sons, 2010.

[38] Haobo Zhang, Hongliang Zhang, Boya Di, Kaigui Bian, Zhu Han, and Lingyang Song. Towards ubiquitous positioning by leveraging reconfigurable intelligent surface. *IEEE Communications Letters*, 25(1):284–288, 2020.

[39] Shubo Huang, Bo Wang, Yanping Zhao, and Mingan Luan. Near-field rss-based localization algorithms using reconfigurable intelligent surface. *IEEE Sensors Journal*, 22(4):3493–3505, 2022.

[40] Cunhua Pan, Hong Ren, Kezhi Wang, Jonas Florentin Kolb, Maged Elkashlan, Ming Chen, Marco Di Renzo, Yang Hao, Jiangzhou Wang, A Lee Swindlehurst, et al. Reconfigurable intelligent surfaces for 6g systems: Principles, applications, and research directions. *IEEE Communications Magazine*, 59(6):14–20, 2021.

[41] Kaiyang Zhou, Ziwei Liu, Yu Qiao, Tao Xiang, and Chen Change Loy. Domain generalization: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[42] Abebe Belay Adege, Hsin-Piao Lin, Getaneh Berie Tarekegn, and Shiann-Shiun Jeng. Applying deep neural network (dnn) for robust indoor localization in multi-building environment. *Applied Sciences*, 8(7):1062, 2018.

[43] Mai Ibrahim, Marwan Torki, and Mustafa ElNainay. Cnn based indoor localization using rss time-series. In *2018 IEEE symposium on computers and communications (ISCC)*, pages 01044–01049. IEEE, 2018.

[44] Mahdi Abid, Paul Compagnon, and Grégoire Lefebvre. Improved cnn-based magnetic indoor positioning system using attention mechanism. In *2021 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–8. IEEE, 2021.

[45] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. Domain-adversarial training of neural networks. *The journal of machine learning research*, 17(1):2096–2030, 2016.

[46] A. Alkhateeb. DeepMIMO: A generic deep learning dataset for millimeter wave and massive MIMO applications. In *Proc. of Information Theory and Applications Workshop (ITA)*, pages 1–8, San Diego, CA, Feb 2019.

[47] Ali Yassin, Youssef Nasser, Mariette Awad, Ahmed Al-Dubai, Ran Liu, Chau Yuen, Ronald Raulefs, and Elias Aboutanios. Recent advances in indoor localization: A survey on theoretical approaches and applications. *IEEE Communications Surveys & Tutorials*, 19(2):1327–1346, 2016.

[48] Nissanka B Priyantha, Anit Chakraborty, and Hari Balakrishnan. The cricket location-support system. In *Proceedings of the 6th annual international conference on Mobile computing and networking*, pages 32–43, 2000.

[49] Lukasz Zwirello, Tom Schipper, Marlene Harter, and Thomas Zwick. Uwb localization system for indoor applications: Concept, realization and analysis. *Journal of Electrical and Computer Engineering*, 2012, 2012.

[50] Binghao Li, Thomas Gallagher, Andrew G Dempster, and Chris Rizos. How feasible is the use of magnetic field alone for indoor positioning? In *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–9. IEEE, 2012.

[51] Ye-Sheng Kuo, Pat Pannuto, Ko-Jen Hsiao, and Prabal Dutta. Luxapose: Indoor positioning with mobile phones and visible light. In *Proceedings of the 20th annual international conference on Mobile computing and networking*, pages 447–458, 2014.

[52] Paramvir Bahl and Venkata N Padmanabhan. Radar: An in-building rf-based user location and tracking system. In *Proceedings IEEE INFOCOM 2000. Conference on computer communications. Nineteenth annual joint conference of the IEEE computer and communications societies (Cat. No. 00CH37064)*, volume 2, pages 775–784. Ieee, 2000.

[53] Jie Xiong and Kyle Jamieson. {ArrayTrack}: A {Fine-Grained} indoor location system. In *10th USENIX Symposium on Networked Systems Design and Implementation (NSDI 13)*, pages 71–84, 2013.

[54] Ales Povalac and Jirí Sebesta. Phase of arrival ranging method for uhf rfid tags using instantaneous frequency measurement. In *2010 Conference Proceedings ICECom, 20th International Conference on Applied Electromagnetics and Communications*, pages 1–4. IEEE, 2010.

[55] Fekher Khelifi, Abbas Bradai, Abderrahim Benslimane, Priyanka Rawat, and Mohamed Atri. A survey of localization systems in internet of things. *Mobile Networks and Applications*, 24(3):761–785, 2019.

[56] Chenshu Wu, Zheng Yang, Yunhao Liu, and Wei Xi. Will: Wireless indoor localization without site survey. *IEEE Transactions on Parallel and Distributed systems*, 24(4):839–848, 2012.

[57] Moustafa Abbas, Moustafa Elhamshary, Hamada Rizk, Marwan Torki, and Moustafa Youssef. Wideep: Wifi-based accurate and robust indoor localization system using deep learning. In *2019 IEEE International Conference on Pervasive Computing and Communications (PerCom*, pages 1–10. IEEE, 2019.

[58] Danyang Li, Jingao Xu, Zheng Yang, Chenshu Wu, Jianbo Li, and Nicholas D Lane. Wireless localization with spatial-temporal robust fingerprints. *ACM Transactions on Sensor Networks (TOSN)*, 18(1):1–23, 2021.

[59] Rathin Chandra Shit, Suraj Sharma, Kumar Yelamarthi, and Deepak Puthal. Ai-enabled fingerprinting and crowdsource-based vehicle localization for resilient and safe transportation systems. *IEEE Transactions on Intelligent Transportation Systems*, 22(7):4660–4669, 2021.

[60] Jindan Zhu, Kai Zeng, Kyu-Han Kim, and Prasant Mohapatra. Improving crowdsourced wi-fi localization systems using bluetooth beacons. In *2012 9th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, pages 290–298. IEEE, 2012.

[61] Boyuan Wang, Xuelin Liu, Baoguo Yu, Ruicai Jia, and Xingli Gan. An improved wifi positioning method based on fingerprint clustering and signal weighted euclidean distance. *Sensors*, 19(10):2300, 2019.

[62] Chenlu Xiang, Shunqing Zhang, Shugong Xu, and George C Alexandropoulos. Self-calibrating indoor localization with crowdsourcing fingerprints and transfer learning. In *ICC 2021-IEEE International Conference on Communications*, pages 1–6. IEEE, 2021.

[63] Xu Sun, Haojun Ai, Jingjie Tao, Tan Hu, and Yusong Cheng. Bert-adloc: A secure crowdsourced indoor localization system based on ble fingerprints. *Applied Soft Computing*, 104:107237, 2021.

[64] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[65] Haojun Ai, Tan Hu, and Tianshui Xu. Rad-gan: Radio map anomaly detection for fingerprint indoor positioning with gan. In *2021 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–8. IEEE, 2021.

[66] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *Communications of the ACM*, 63(11):139–144, 2020.

[67] Hongbo Liu, Yu Gan, Jie Yang, Simon Sidhom, Yan Wang, Yingying Chen, and Fan Ye. Push the limit of wifi based localization for smartphones. In *Proceedings of the 18th annual international conference on Mobile computing and networking*, pages 305–316, 2012.

[68] Han Xu, Zheng Yang, Zimu Zhou, Longfei Shangguan, Ke Yi, and Yunhao Liu. Enhancing wifi-based localization with visual clues. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, pages 963–974, 2015.

[69] Israa Fahmy, Samah Ayman, Hamada Rizk, and Moustafa Youssef. Monofi: Efficient indoor localization based on single radio source and minimal fingerprinting. In *Proceedings of the 29th International Conference on Advances in Geographic Information Systems*, pages 674–675, 2021.

[70] Azin Arya, Philippe Godlewski, Marine Campedel, and Ghislain Du Chéne. Radio database compression for accurate energy-efficient localization in fingerprinting systems. *IEEE Transactions on Knowledge and Data Engineering*, 25(6):1368–1379, 2011.

[71] Jun Yan, Guowen Qi, Bin Kang, Xiaohuan Wu, and Huaping Liu. Extreme learning machine for accurate indoor localization using rssi fingerprints in multifloor environments. *IEEE Internet of Things Journal*, 8(19):14623–14637, 2021.

[72] Guang-Bin Huang, Qin-Yu Zhu, and Chee-Kheong Siew. Extreme learning machine: theory and applications. *Neurocomputing*, 70(1-3):489–501, 2006.

[73] Santosh Subedi, Dae-Ho Kim, Beom-Hun Kim, and Jae-Young Pyun. Improved smartphone-based indoor localization system using lightweight fingerprinting and inertial sensors. *IEEE Access*, 9:53343–53357, 2021.

[74] Antoni Perez-Navarro. Accuracy of a single point in knn applying error propagation theory. In *2021 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–7. IEEE, 2021.

[75] Sebastian Sadowski, Petros Spachos, and Konstantinos N Plataniotis. Memoryless techniques and wireless technologies for indoor localization with the internet of things. *IEEE Internet of Things Journal*, 7(11):10996–11005, 2020.

[76] Ahmed H Salamah, Mohamed Tamazin, Maha A Sharkas, and Mohamed Khedr. An enhanced wifi indoor localization system based on machine learning. In *2016 International conference on indoor positioning and indoor navigation (IPIN)*, pages 1–8. IEEE, 2016.

[77] Zhenghua Chen, Han Zou, JianFei Yang, Hao Jiang, and Lihua Xie. Wifi fingerprinting indoor localization using local feature-based deep lstm. *IEEE Systems Journal*, 14(2):3001–3010, 2019.

[78] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[79] Junxiang Wang, Canyang Guo, and Ling Wu. Gated recurrent unit with rssis from heterogeneous network for mobile positioning. *Mobile Information Systems*, 2021, 2021.

[80] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning internal representations by error propagation. Technical report, California Univ San Diego La Jolla Inst for Cognitive Science, 1985.

[81] Michael I Jordan. Serial order: A parallel distributed processing approach. In *Advances in psychology*, volume 121, pages 471–495. Elsevier, 1997.

[82] Wenhui Wang, Nan Yang, Furu Wei, Baobao Chang, and Ming Zhou. Gated self-matching networks for reading comprehension and question answering. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 189–198, 2017.

[83] Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

[84] Xuanshu Luo and Nirvana Meratnia. A geometric deep learning framework for accurate indoor localization. In *12th International conference on indoor positioning and indoor navigation (IPIN 2022)*, 2022.

[85] Zheshun Wu, Xiaoping Wu, and Yunliang Long. Multi-level federated graph learning and self-attention based personalized wi-fi indoor fingerprint localization. *IEEE Communications Letters*, 2022.

[86] Han Zheng, Yan Zhang, Lan Zhang, Hao Xia, Shaojie Bai, Guobin Shen, Tian He, and Xiangyang Li. Grafin: An applicable graph-based fingerprinting approach for robust indoor localization. In *2021 IEEE 27th International Conference on Parallel and Distributed Systems (ICPADS)*, pages 747–754. IEEE, 2021.

[87] Henk Wymeersch and Benoît Denis. Beyond 5g wireless localization with reconfigurable intelligent surfaces. In *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2020.

[88] Jorma J Rissanen. Fisher information and stochastic complexity. *IEEE transactions on information theory*, 42(1):40–47, 1996.

[89] Jiguang He, Henk Wymeersch, Long Kong, Olli Silvén, and Markku Juntti. Large intelligent surface for positioning in millimeter wave mimo systems. In *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, pages 1–5. IEEE, 2020.

[90] S.T. Smith. Covariance, subspace, and intrinsic crame/spl acute/r-rao bounds. *IEEE Transactions on Signal Processing*, 53(5):1610–1630, 2005.

[91] Haobo Zhang, Jingzhi Hu, Hongliang Zhang, Boya Di, Kaigui Bian, Zhu Han, and Lingyang Song. Metaradar: Indoor localization by reconfigurable metamaterials. *IEEE Transactions on Mobile Computing*, 2020.

[92] Zhilin Yang, Zihang Dai, Yiming Yang, Jaime Carbonell, Russ R Salakhutdinov, and Quoc V Le. Xlnet: Generalized autoregressive pretraining for language understanding. *Advances in neural information processing systems*, 32, 2019.

[93] Benjamin Recht, Rebecca Roelofs, Ludwig Schmidt, and Vaishaal Shankar. Do imagenet classifiers generalize to imagenet? In *International Conference on Machine Learning*, pages 5389–5400. PMLR, 2019.

[94] Yawei Luo, Liang Zheng, Tao Guan, Junqing Yu, and Yi Yang. Taking a closer look at domain shift: Category-level adversaries for semantics consistent domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2507–2516, 2019.

[95] Alexandre Rame, Corentin Dancette, and Matthieu Cord. Fishr: Invariant gradient variances for out-of-distribution generalization. In *International Conference on Machine Learning*, pages 18347–18377. PMLR, 2022.

[96] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[97] Yuval Netzer, Tao Wang, Adam Coates, Alessandro Bissacco, Bo Wu, and Andrew Y Ng. Reading digits in natural images with unsupervised feature learning. 2011.

[98] Rich Caruana. Multitask learning. *Machine learning*, 28(1):41–75, 1997.

[99] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359, 2009.

[100] Karl Weiss, Taghi M Khoshgoftaar, and DingDing Wang. A survey of transfer learning. *Journal of Big data*, 3(1):1–40, 2016.

[101] Jindong Wang, Cuiling Lan, Chang Liu, Yidong Ouyang, Tao Qin, Wang Lu, Yiqiang Chen, Wenjun Zeng, and Philip Yu. Generalizing to unseen domains: A survey on domain generalization. *IEEE Transactions on Knowledge and Data Engineering*, 2022.

[102] Ricardo Vilalta and Youssef Drissi. A perspective view and survey of meta-learning. *Artificial intelligence review*, 18(2):77–95, 2002.

[103] Timothy Hospedales, Antreas Antoniou, Paul Micaelli, and Amos Storkey. Meta-learning in neural networks: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 44(9):5149–5169, 2021.

[104] Chelsea B Finn. *Learning to learn with gradients*. University of California, Berkeley, 2018.

[105] Mike Huisman, Jan N Van Rijn, and Aske Plaat. A survey of deep meta-learning. *Artificial Intelligence Review*, 54(6):4483–4541, 2021.

[106] Connor Shorten and Taghi M Khoshgoftaar. A survey on image data augmentation for deep learning. *Journal of big data*, 6(1):1–48, 2019.

[107] Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, 35(8):1798–1828, 2013.

[108] Krikamol Muandet, David Balduzzi, and Bernhard Schölkopf. Domain generalization via invariant feature representation. In *International Conference on Machine Learning*, pages 10–18. PMLR, 2013.

[109] Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, and Dilip Krishnan. Supervised contrastive learning. *Advances in Neural Information Processing Systems*, 33:18661–18673, 2020.

[110] Raia Hadsell, Sumit Chopra, and Yann LeCun. Dimensionality reduction by learning an invariant mapping. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, volume 2, pages 1735–1742. IEEE, 2006.

[111] Solomon Kullback and Richard A Leibler. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86, 1951.

[112] Yanqiao Zhu, Yichen Xu, Qiang Liu, and Shu Wu. An empirical study of graph contrastive learning. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*, 2021.

[113] Yaroslav Ganin and Victor Lempitsky. Unsupervised domain adaptation by backpropagation. In *International conference on machine learning*, pages 1180–1189. PMLR, 2015.

[114] Ya Li, Xinmei Tian, Mingming Gong, Yajing Liu, Tongliang Liu, Kun Zhang, and Dacheng Tao. Deep domain generalization via conditional invariant adversarial networks. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 624–639, 2018.

[115] Rui Shao, Xiangyuan Lan, Jiawei Li, and Pong C Yuen. Multi-adversarial discriminative deep domain generalization for face presentation attack detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10023–10031, 2019.

[116] Yunpei Jia, Jie Zhang, Shiguang Shan, and Xilin Chen. Single-side domain generalization for face anti-spoofing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8484–8493, 2020.

[117] Emil Björnson, Luca Sanguinetti, Henk Wymeersch, Jakob Hoydis, and Thomas L. Marzetta. Massive mimo is a reality—what is next?: Five promising research directions for antenna arrays. *Digital Signal Processing*, 94:3–20, 2019. Special Issue on Source Localization in Massive MIMO.

[118] Muhammad Alrabeiah, Yu Zhang, and Ahmed Alkhateeb. Neural networks based beam codebooks: Learning mmwave massive mimo beams that adapt to deployment and hardware. *IEEE Transactions on Communications*, 70(6):3818–3833, 2022.

[119] Marco Di Renzo, Alessio Zappone, Merouane Debbah, Mohamed-Slim Alouini, Chau Yuen, Julien De Rosny, and Sergei Tretyakov. Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead. *IEEE journal on selected areas in communications*, 38(11):2450–2525, 2020.

[120] Marco Di Renzo, Fadil Habibi Danufane, Xiaojun Xi, Julien De Rosny, and Sergei Tretyakov. Analytical modeling of the path-loss for reconfigurable intelligent surfaces–anomalous mirror or scatterer? In *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5. IEEE, 2020.

[121] Abla Kammoun, Anas Chaaban, Mérouane Debbah, Mohamed-Slim Alouini, et al. Asymptotic max-min sinr analysis of reconfigurable intelligent surface assisted miso systems. *IEEE Transactions on Wireless Communications*, 19(12):7748–7764, 2020.

[122] Chongwen Huang, Alessio Zappone, George C Alexandropoulos, Mérouane Debbah, and Chau Yuen. Reconfigurable intelligent surfaces for energy efficiency in wireless communication. *IEEE Transactions on Wireless Communications*, 18(8):4157–4170, 2019.

[123] Chongwen Huang, George C Alexandropoulos, Alessio Zappone, Mérouane Debbah, and Chau Yuen. Energy efficient multi-user miso communication using low resolution large intelligent surfaces. In *2018 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6. IEEE, 2018.

[124] Xilong Pei, Haifan Yin, Li Tan, Lin Cao, Zhanpeng Li, Kai Wang, Kun Zhang, and Emil Björnson. Ris-aided wireless communications: Prototyping, adaptive beamforming, and indoor/outdoor field trials. *IEEE Transactions on Communications*, 69(12):8627–8640, 2021.

[125] Linglong Dai, Bichai Wang, Min Wang, Xue Yang, Jingbo Tan, Shuangkaisheng Bi, Shenheng Xu, Fan Yang, Zhi Chen, Marco Di Renzo, et al. Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results. *IEEE access*, 8:45913–45923, 2020.

[126] Roi Méndez-Rial, Cristian Rusu, Nuria González-Prelcic, Ahmed Alkhateeb, and Robert W Heath. Hybrid mimo architectures for millimeter wave communications: Phase shifters or switches? *IEEE access*, 4:247–267, 2016.

[127] Abdelrahman Taha, Muhammad Alrabeiah, and Ahmed Alkhateeb. Enabling large intelligent surfaces with compressive sensing and deep learning. *IEEE access*, 9:44304–44321, 2021.

[128] Junyeub Suh, Changhyeon Kim, Wonjin Sung, Jaewoo So, and Seo Weon Heo. Construction of a generalized dft codebook using channel-adaptive parameters. *IEEE Communications Letters*, 21(1):196–199, 2016.

[129] Bobbi Jo Broxson. The kronecker product. 2006.

[130] Ibrahim A Hemadeh, Katla Satyanarayana, Mohammed El-Hajjar, and Lajos Hanzo. Millimeter-wave communications: Physical channel models, design considerations,

antenna constructions, and link-budget. *IEEE Communications Surveys & Tutorials*, 20(2):870–913, 2017.

[131] Thomas Schneider, Andrzej Wiatrek, Stefan Preußler, Michael Grigat, and Ralf-Peter Braun. Link budget analysis for terahertz fixed wireless links. *IEEE Transactions on Terahertz Science and Technology*, 2(2):250–256, 2012.

[132] Jim Zyren and Al Petrick. Tutorial on basic link budget analysis. *Application Note AN9804, Harris Semiconductor*, 31, 1998.

[133] A. Zelst, van. *MIMO OFDM for wireless LANs*. PhD thesis, Agere Systems, 2004.

[134] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and S Yu Philip. A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*, 32(1):4–24, 2020.

[135] Maxim Vladimirovich Shcherbakov, Adriaan Brebels, Nataliya Lvovna Shcherbakova, Anton Pavlovich Tyukov, Timur Alexandrovich Janovsky, Valeriy Anatol'evich Kamaev, et al. A survey of forecast error measures. *World applied sciences journal*, 24(24):171–176, 2013.

[136] Zhimian Zhang, Haipeng Wang, Feng Xu, and Ya-Qiu Jin. Complex-valued convolutional neural network and its application in polarimetric sar image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(12):7177–7188, 2017.

[137] Yice Cao, Yan Wu, Peng Zhang, Wenkai Liang, and Ming Li. Pixel-wise polsar image classification via a novel complex-valued deep fully convolutional network. *Remote Sensing*, 11(22):2653, 2019.

[138] J Agustin Barrachina. Complex-valued neural networks (cvnn), January 2021.

[139] George C Alexandropoulos and Evangelos Vlachos. A hardware architecture for reconfigurable intelligent surfaces with minimal active elements for explicit channel estimation. In *ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pages 9175–9179. IEEE, 2020.

[140] Will Hamilton, Zhitao Ying, and Jure Leskovec. Inductive representation learning on large graphs. *Advances in neural information processing systems*, 30, 2017.

[141] Andrew L Maas, Awni Y Hannun, Andrew Y Ng, et al. Rectifier nonlinearities improve neural network acoustic models. In *Proc. icml*, volume 30, page 3. Citeseer, 2013.

[142] Justin Gilmer, Samuel S Schoenholz, Patrick F Riley, Oriol Vinyals, and George E Dahl. Neural message passing for quantum chemistry. In *International conference on machine learning*, pages 1263–1272. PMLR, 2017.

[143] Ivan Nazarov and Evgeny Burnaev. Bayesian Sparsification of Deep C-valued Networks. In *International Conference on Machine Learning*, volume 119, pages 7230–7242. PMLR, November 2020. ISSN: 2640-3498.

[144] Martin Arjovsky, Amar Shah, and Yoshua Bengio. Unitary evolution recurrent neural networks. In *International conference on machine learning*, pages 1120–1128. PMLR, 2016.

[145] Vijay Venkateswaran and Alle-Jan van der Veen. Analog beamforming in mimo communications with phase shift networks and online channel estimation. *IEEE Transactions on Signal Processing*, 58(8):4131–4143, 2010.

[146] Ian S. Grant and William Robert Phillips. *Electromagnetism*. John Wiley & Sons, 2013.

[147] Akira Hirose. *Complex-valued neural networks: theories and applications*, volume 5. World Scientific, 2003.

[148] Jose Agustin Barrachina, Chenfang Ren, Christele Morisseau, Gilles Vieillard, and J-P Ovarlez. Complex-valued vs. real-valued neural networks for classification perspectives: An example on non-circular data. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2990–2994. IEEE, 2021.

[149] Yixin Li, Yong Luo, Guangli Yang, et al. 12-port 5g massive mimo antenna array in sub-6ghz mobile handset for lte bands 42/43/46 applications. *IEEE access*, 6:344–354, 2017.

[150] Jian Pang, Zheng Li, Ryo Kubozoe, Xueting Luo, Rui Wu, Yun Wang, Dongwon You, Ashbir Aviat Fadila, Rattanan Saengchan, Takeshi Nakamura, et al. 21.1 a 28ghz cmos phased-array beamformer utilizing neutralized bi-directional technique supporting dual-polarized mimo for 5g nr. In *2019 IEEE International Solid-State Circuits Conference-(ISSCC)*, pages 344–346. IEEE, 2019.

[151] David R Pauluzzi and Norman C Beaulieu. A comparison of snr estimation techniques for the awgn channel. *IEEE Transactions on communications*, 48(10):1681–1691, 2000.

[152] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.

[153] Minjie Wang, Da Zheng, Zihao Ye, Quan Gan, Mufei Li, Xiang Song, Jinjing Zhou, Chao Ma, Lingfan Yu, Yu Gai, Tianjun Xiao, Tong He, George Karypis, Jinyang Li, and Zheng Zhang. Deep graph library: A graph-centric, highly-performant package for graph neural networks. 2019.

[154] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015.

[155] Yi Jiang, Keren Li, Jing Gao, and Hiroshi Harada. Antenna space diversity and polarization mismatch in wideband 60ghz-millimeter-wave wireless system. In *2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1781–1785. IEEE, 2009.

[156] Carl B Dietrich, Kai Dietze, J Randall Nealy, and Warren L Stutzman. Spatial, polarization, and pattern diversity for wireless handheld terminals. *IEEE transactions on antennas and propagation*, 49(9):1271–1281, 2001.

[157] Seok-Chul Kwon and Gordon L Stüber. Polarization division multiple access on nlos wide-band wireless fading channels. *IEEE transactions on wireless communications*, 13(7):3726–3737, 2014.

[158] David Tse and Pramod Viswanath. *Fundamentals of wireless communication*. Cambridge university press, 2005.