

Sensorless Air Temperature Sensing using LoRa Link Characteristics

Aitian Ma, Jean Tonday Rodriguez, Mo Sha, Dongsheng Luo
Knight Foundation School of Computing and Information Sciences
Florida International University
{aima, jtond004, msha, dluo}@fiu.edu

Abstract—Air temperature monitoring is essential for many Internet of Things (IoT) applications. Many existing applications rely on the readings provided by the weather stations maintained by federal, regional, or local government agencies. While the accuracy of the data provided by those weather stations is high, the ability of such data to reflect the temperature variability experienced by urban populations is generally low because the measurements are collected at the mesoscale. In reality, the air temperature varies at the microscale and local scale, and the health risks associated with extreme weather are assumed to vary with the exposure. In this paper, we present LORATEMP, a novel solution that uses LoRa link characteristics and advanced machine learning techniques to predict air temperature. LORATEMP leverages a unique correlation map representation and a novel dual attention network to capture the complex dependency between LoRa link characteristics and air temperature, and employs adversarial domain adaptation to transfer the temperature prediction knowledge learned from one device to those without temperature sensors using a few temporal measurements. We implement LORATEMP and evaluate it in real-world environments. Experimental results show that LORATEMP significantly outperforms all baselines and reduces air temperature prediction errors by at least 30%.

Index Terms—Air temperature monitoring, LoRa link characteristics, domain adaptation, multivariate time series forecasting

I. INTRODUCTION

Air temperature monitoring is essential for many Internet of Things (IoT) applications. Many existing applications rely on the readings provided by the weather stations maintained by federal, regional, or local government agencies [1]–[4]. For example, Zhang et al. used the provided air temperature readings to conduct research on predicting heat-related mortality in urban environments [1]. While the accuracy of the data provided by those weather stations is high, the ability of such data to reflect the temperature variability experienced by urban populations is generally low because the measurements are collected at the mesoscale (3,000–100,000m). In addition, the weather stations are often located in open areas to ensure no interference from shading, therefore they do not reflect the distribution of populations, nor of built environments that can generate urban heat island effects. In reality, the air temperature varies at the microscale (<100m) and the local scale (100–3,000m), and the health risks associated with extremely hot weather are assumed to vary with the exposure. To overcome such limitations, recent studies have proposed to deploy new infrastructures to provide fine-scale

measurements [5]–[7]. However, industry practitioners have shown a marked reluctance to embrace such solutions due to the high deployment and maintenance costs.

Recent years have witnessed rapid deployments of LoRa networks in both urban and rural areas. As an emerging Low-Power Wide-Area Network (LPWAN) technology, LoRa provides a low-cost wireless solution that supports long-range data collection for low data rate applications. Over the past decade, LoRa networks have been deployed in 153 countries to support various applications. For example, a city-wide LoRa network that consists of 100 gateways and 19,821 end devices is deployed to support 12 kinds of smart city applications, such as gas meter monitoring, water meter monitoring, and gas alarming [8]. The widely deployed LoRa networks offer new opportunities to provide fine-scale temperature measurements. Unfortunately, most of the deployed LoRa devices do not have temperature sensors, and adding new hardware is costly or infeasible in many cases where the device owners do not allow a third party to access their hardware or make modifications due to their security and privacy concerns. This motivates us to explore *a new solution that only requires device owners to share link traces and allows us to collect a few temporary temperature measurements near their deployments, which are enough to turn those deployments into virtual weather stations.*

In this paper, we first present our empirical study that demonstrates the feasibility of using LoRa link characteristics to predict ambient air temperature. While our study shows promising results, it also highlights the important challenges posed by (1) the complex dependency between LoRa link characteristics and temperature measurements observed in the real world, (2) the lack of methods in capturing the non-static and complex correlation among link metrics, and (3) the fact that the prediction model generated on one device are not applicable in other devices. To address those challenges, we develop a novel air temperature monitoring solution, namely LORATEMP, that leverages a unique correlation map representation method and a novel dual attention network to capture the complex dependency between LoRa link characteristics and temperature measurements, and employs adversarial domain adaptation to transfer the temperature prediction knowledge learned from one device to those without temperature sensors using a few temporal measurements. We implement our solution and test it in real-world environments. Experimental results show that

LORATEMP outperforms all baselines and provides accurate air temperature predictions.

Our paper is organized as follows. Section II introduces our empirical study. Section III presents the design of LORATEMP. Section IV evaluates LORATEMP. Section V reviews the related work. Section VI concludes this paper.

II. EMPIRICAL STUDY

We perform an empirical study to explore the feasibility and identify the challenges of using LoRa link characteristics to predict air temperature. In this section, we first introduce our hardware deployment and collected data. We then present our problem formulation and findings.

A. Hardware Deployment and Collected Data

We perform our empirical study on a network consisting of one LoRa base station and six LoRa end devices, which are deployed in an urban area to support a smart city application. We add a temperature sensor to each LoRa end device to collect ground truth temperature readings. Each LoRa end device periodically measures ambient air temperature and sends the measurement to the LoRa base station together with the link characteristics, including the Received Signal Strength (RSS), the Signal-to-Noise Ratio (SNR), the Packet Receive Ratio (PRR), and the Bit Error Rate (BER), measured when receiving the last packet transmitted by the LoRa base station. After receiving each packet, the LoRa base station measures the link characteristics and sends back an acknowledgment packet. The data traces are gathered from six LoRa end devices, denoted as $d_0, d_1, d_2, d_3, d_4,$ and d_5 .

B. Problem Formulation

We approach the task of air temperature prediction as a regression problem, where the input is a sequence of measurements on LoRa link characteristics, and the output is the corresponding air temperature. Mathematically, we define the sequence of link measurements as a matrix $\mathbf{X} \in \mathbb{R}^{T \times L}$, where T is the length of the sequence and L is the number of link metrics considered. Our goal is to learn a nonlinear mapping $f: \mathbb{R}^{T \times L} \rightarrow \mathbb{R}$, which translates the sequence of link measurements into an accurate temperature prediction. This mapping is anticipated to encapsulate the underlying dynamics and dependency between the link metrics and the temperature, offering a viable solution to the challenge of direct temperature computation from link metrics in variable outdoor conditions.

C. Feasibility of Using Link Characteristics to Predict Air Temperature

Although early efforts have been made to theoretically model the effects of air temperature on wireless communication medium and wireless transceiver, there exist significant challenges in using the link characteristics measured by a wireless receiver to predict air temperature because the link measurements are largely affected by other factors, such as nearby wireless transmitters, external events like moving objects and falling trees, and environmental variables including

TABLE I: Pearson correlation coefficients and MI test results between air temperature and different LoRa link metrics.

Pearson Correlation Test			Mutual Information Test	
	Coefficient	P Value	Results	
RSS	-0.057	2.564×10^{-16}	RSS	0.454
SNR	0.049	2.167×10^{-12}	SNR	0.529
BER	0.034	1.406×10^{-6}	BER	0.008
PRR	0.028	5.973×10^{-5}	PRR	0.322

TABLE II: MAE of temperature predictions provided the MLP models trained with different features over five days. Single: MLP trained with RSS as input. Multiple: MLP trained with RSS, SNR, PRR, and BER as input. Multi-TS: MLP trained with RSS, SNR, PRR, BER, and Timestamp as input.

Day	Single	Multiple	Multi-TS
First	0.064	0.061	0.058
Second	0.060	0.055	0.048
Third	0.122	0.109	0.083
Forth	0.093	0.086	0.056
Fifth	0.090	0.082	0.053
Avg. MAE	0.086	0.079	0.060

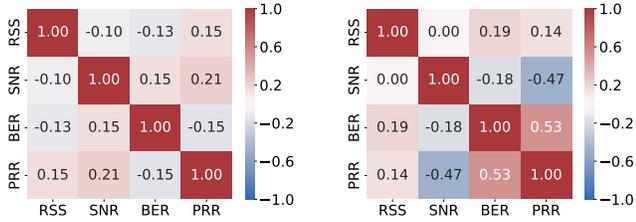
precipitation and humidity. The challenges are exaggerated in the context of LoRa, which operates at low power in the license-free sub-gigahertz radio frequency bands.

We begin our empirical study by analyzing the linear correlation between air temperature measurements and various LoRa link parameters. Table I lists the Pearson correlation coefficients between air temperature and each LoRa link metric. As Table I lists, the linear correlation between temperature and LoRa link characteristics is weak and fragile. For example, the correlation coefficient between air temperature and RSS measurements is -0.057 and the P value is 2.564×10^{-16} . Due to the weak and fragile correlation, it would be very challenging for any linear models to make good temperature predictions based on link characteristics.

We further perform the Mutual Information (MI) test to explore the nonlinear dependence between temperature and link characteristics. The test results listed in Table I show some significant dependencies (0.454 for RSS, 0.529 for SNR, and 0.322 for PRR). While the MI test shows promising results, it also highlights the important challenges posed by the complex nonlinear dependency, which motivates us to develop a new deep-learning based method to capture it.

Observation 1: *There exists a complex nonlinear dependence between air temperature and LoRa link characteristics.*

To explore the use of LoRa link characteristics to predict air temperature, we employ the Multi-Layer Perceptron (MLP) to train the prediction models and examine their prediction performance. We preprocess the data before training by applying a min-max scaler with a maximum temperature difference of 16°C and train each MLP model with 200 epochs and a learning rate of 0.01. The MLP model consists of four hidden layers with 128, 256, 128, and 64 neurons, respectively. Table II lists the Mean Absolute Error (MAE) of temperature predictions



(a) In low temperature range. (b) In high temperature range.

Fig. 1: Pearson correlation between link metrics in two temperature ranges.

over a 5-day period. As Table II lists, the MLP model with RSS, SNR, PRR, BER, and Timestamp as input provides the best performance with the smallest averaged MAE (0.06). The results demonstrate the benefit of treating the LoRa link measurements as the time series data when training the prediction model and leveraging the temporal dependency in the link measurements to produce good temperature predictions.

Observation 2: *The temporal dependency in the LoRa link measurements carries additional knowledge on air temperature predictions.*

To further understand the temperature knowledge carried by the link characteristics, we analyze the correlation between those link metrics in different temperature ranges. Figure 1 plots the Pearson correlation coefficients in two ranges: high-temperature range (no less than 28°C) and low-temperature range. Figure 1 shows that the correlation coefficients differ a lot in different temperature ranges. For example, the correlation coefficient between RSS and BER in the low-temperature range is -0.13, while the coefficient is 0.19 in the high-temperature range. Similarly, the correlation coefficients between PRR and SNR are 0.21 in the low-temperature range and -0.47 in the high-temperature range. The results indicate that temperature plays an important role in modulating the correlation between different link metrics, which carries additional knowledge on air temperature predictions. Unfortunately, the existing machine learning models, such as Linear Regression (LR) and MLP, cannot encode such inter-correlation (correlation between different link metrics) between different pairs of time series [9].

Observation 3: *Air temperature plays an important role in modulating the correlation between different link metrics, which carries additional knowledge on air temperature predictions.*

D. Feasibility of Sharing Temperature Prediction Models Between Different Devices

One of our main objectives is to enable accurate air temperature predictions for those devices without temperature sensors. However, obtaining temperature labels for the LoRa end device not equipping a temperature sensor is labor-intensive and time-consuming, primarily due to the requirement for manual data collection and labeling. This process involves directly measuring ambient air temperature around devices without built-in sensors, accurately associating these measurements

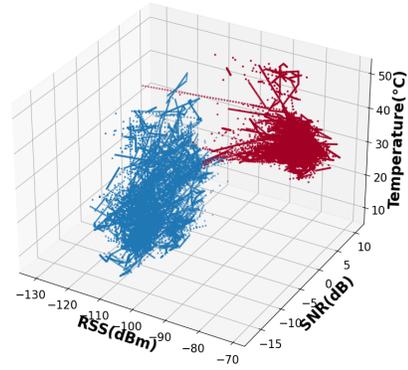


Fig. 2: Air temperature distributions under various RSS and SNR combinations on two different LoRa end devices: one marked in blue and the other marked in red.

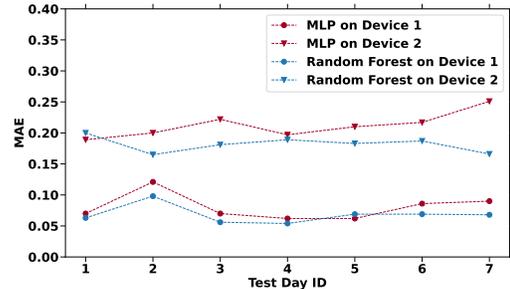


Fig. 3: MAE of temperature predictions when using MLP and Random Forest models.

with the correct devices and time stamps, and integrating this information into the existing dataset. We explore the feasibility of applying the temperature prediction model trained with the data collected from one device to the data gathered by other devices. Figure 2 plots the air temperature distributions under various RSS and SNR combinations on two different LoRa end devices. As Figure 2 shows, there exist significant differences in temperature values measured by those two devices, even under the same RSS and SNR conditions. The centers and shapes of those two distributions differ a lot. Therefore, directly applying the prediction model trained on one device to others is very unlikely to succeed. However, the similarities in the distribution patterns indicate the feasibility of transferring temperature prediction knowledge between different devices.

To intuitively quantify the discrepancy between two temperature distributions, we train the temperature prediction model on one LoRa end device and apply it on another. Figure 3 plots MAE of temperature predictions when we apply MLP and the random forest models trained on one device to the new data collected from the same device and another device. As Figure 3 shows, the MAE values are small and always no more than 0.121 when we test the models on the same device. However, the MAE values increase up to 0.251 under MLP and up to 0.200 under the random forest model when we apply the model on the other device. The results show that directly sharing the temperature prediction model between devices may not be able to provide good prediction performance and there is a critical need for a good solution to transfer the temperature prediction

knowledge between devices.

Observation 4: Directly applying the temperature prediction model trained on one device to others does not work well.

III. LORATEMP

Guided by the insights gathered from our empirical study, we develop a novel air temperature prediction solution, namely LORATEMP. Figure 4 shows the framework of LORATEMP, which consists of three components: **Correlation Map Representation** (preprocessing), **Dual Attention Network**, and **Adversarial Domain Adaptation**. The preprocessing aims to capture the dynamic correlation between link metrics over time by transforming the multivariate link characteristics into a sequence of correlation maps (see Section III-A). Dual Attention Network takes the correlation maps generated by the preprocessing component as input and leverages a dual attention mechanism to capture the complex temporal and link-wise dependency within the data (see Section III-B). Adversarial Domain Adaptation offers a label-efficient methodology for model training, which robustly transfers the temperature prediction knowledge learned from one device to others (see Section III-C).

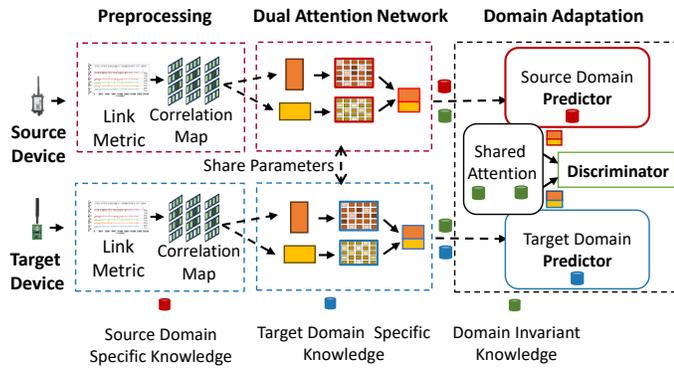


Fig. 4: Overview of LORATEMP.

A. Correlation Map Representation

As Observation 1 in Section II-C states, there exists complex nonlinear dependence between air temperature and LoRa link characteristics. Observation 3 in Section II-C further suggests that the knowledge of the correlation between different link metrics can help characterize the system status and produce good air temperature predictions. Inspired by Zhang et al. [9], we employ a correlation map to explicitly model the dynamic correlation. Specifically, each time step t is characterized by a correlation map, where the relationship between link metrics is quantified using the inner product of their respective measurements over the past T_w time steps. With $\mathbf{X} \in \mathbb{R}^{T_w \times L}$ encapsulating the preceding link metrics, where L represents the number of link characteristics, the correlation map is constructed as follows:

$$\mathbf{M}_t = \mathbf{X}_t^\top \mathbf{X}_t, \quad (1)$$

with \mathbf{X}_t^\top being the transpose of \mathbf{X}_t . Here, \mathbf{M}_t is the correlation map, and its element at position (i, j) reveals the correlation

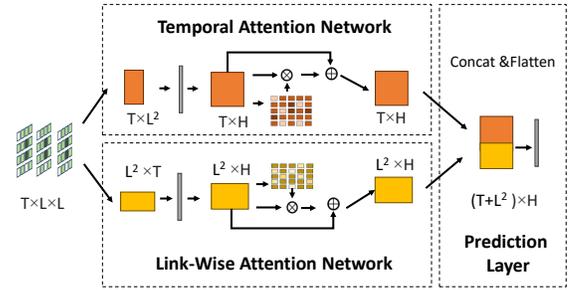


Fig. 5: Dual Attention Network with three components: a temporal attention network, a link-wise attention network, and a prediction layer.

between the i -th and j -th features at time t . In addition to explicit correlation modeling, this preprocessing design is resilient to input noise, as turbulence at certain time steps has a negligible effect on the overall correlation patterns.

B. Dual Attention Network

Observations 2 and 3 in Section II-C highlight the critical role of temporal-level and link metric dependency in providing accurate temperature predictions. Nonetheless, extracting such a dependency from the complex, multivariate characteristics of wireless links is a formidable challenge. The existing wireless sensing techniques designed for device tracking, gesture recognition, and activity recognition rely on capturing multipath channel variations, such as amplitude attenuation and phase shifts, for prediction purposes. However, acquiring such detailed data often necessitates hardware modifications, which are impractical in our application scenario.

To overcome such obstacles, we turn to the self-attention mechanism, a concept that has demonstrated remarkable flexibility and effectiveness in identifying correlations within sequence data across a range of fields, including natural language processing, computer vision, and audio processing [10]. This mechanism draws inspiration from human cognitive processes, enabling the explicit modeling of dependency between each pair of input units without requiring domain-specific knowledge. By training the attention mechanism alongside other model parameters using prediction loss, we can capture complex temporal dynamics and link metric dependency directly from the data.

To this end, we develop a novel Dual Attention Network, which is tailored for air temperature prediction from multivariate link features. This network leverages the self-attention mechanism to address two key challenges: first, it captures the intricate temporal patterns that are essential for understanding temperature fluctuations over time; and second, it identifies the subtle dependency between different link metrics that are indicative of temperature changes. By integrating these dual aspects of attention, our network offers a comprehensive and data-driven approach to temperature prediction. As Figure 5 shows, our Dual Attention Network consists of three components: a temporal attention network, a link-wise attention network, and a prediction layer.

The temporal attention network aims to capture the temporal dependency within the input LoRa link sequences. We represent the input multivariate link features as a three-dimensional tensor, $\mathbf{M} \in \mathbb{R}^{T \times L \times L}$, where T represents the window size and L denotes the number of link features. At each time step, we flatten the signature matrix to produce the input, denoted by $\mathbf{X}_T \in \mathbb{R}^{T \times L^2}$, for the temporal attention process. We employ a single-head self-attention network to facilitate the temporal collaboration among the inputs. The temporal attention is defined as:

$$\mathbf{A}_T = \text{softmax}\left(\frac{\mathbf{Q}_T \mathbf{K}_T^\top}{\sqrt{d_{k_T}}}\right) \mathbf{V}_T, \quad (2)$$

where the query vectors \mathbf{Q}_T , the key vectors \mathbf{K}_T , and the value vectors \mathbf{V}_T are computed within the attention mechanism using the input $\mathbf{X}_T \in \mathbb{R}^{T \times L^2}$, d_{k_T} is the dimension of \mathbf{K}_T , and the element at the i, j -th position quantifies the relevance of time step j for i .

The link-wise attention network is adept at unraveling temporal patterns within the data and engineered to extract and analyze the intricate dependency existing among the various link features. This analysis is pivotal for understanding the nuanced interplay between different link metrics and their collective effects on air temperature prediction. Given the nature of our data and the specific focus on link features, the input for the link-wise attention network is a transposed version of \mathbf{X}_T , denoted as $\mathbf{X}_L \in \mathbb{R}^{L^2 \times T}$. This transposition facilitates a shift in focus from temporal dynamics to the relationships between link features across the same time frame. Similar to the temporal domain, we employ a single-head self-attention mechanism tailored for the link-wise domain. The objective is to enable the network to discern and leverage the dependency between different link metrics that might influence temperature readings. The link-wise attention is defined as:

$$\mathbf{A}_L = \text{softmax}\left(\frac{\mathbf{Q}_L \mathbf{K}_L^\top}{\sqrt{d_{k_L}}}\right) \mathbf{V}_L, \quad (3)$$

where the query vectors \mathbf{Q}_L , the key vectors \mathbf{K}_L , and the value vectors \mathbf{V}_L are computed within the attention mechanism using the input $\mathbf{X}_L \in \mathbb{R}^{L^2 \times T}$, d_{k_L} is the dimension of \mathbf{K}_L , and the element at the i, j -th position quantifies the relevance of time step j for i .

To synthesize the insights garnered from both the temporal and link-wise attention networks, the prediction layer concatenates their outputs and subsequently flattens the resulting matrix to form a unified representation vector. Formally, we have

$$\mathbf{h} = \text{flattening}([\mathbf{A}_T; \mathbf{A}_L]), \quad (4)$$

where $[\cdot; \cdot]$ denotes the concatenation operator, and the flattening operation transforms a matrix into a vector by sequentially aligning its elements along the first dimension.

C. Adversarial Domain Adaptation

One of our design goals is to enable accurate air temperature predictions for those devices without temperature sensors. The

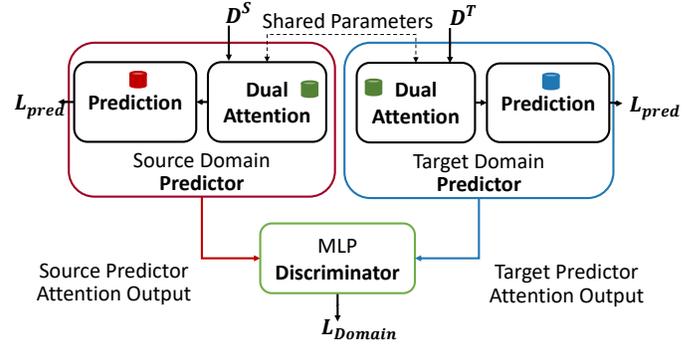


Fig. 6: Adversarial Domain Adaptation. The red cylinder represents the source domain-specific knowledge, the blue one represents the target domain-specific knowledge, and the green one represents the domain-invariant knowledge shared between source and target domains.

availability and quality of labeled data play an important role in the performance of deep learning models. However, obtaining temperature labels for the LoRa end device that does not have a temperature sensor is labor intensive and time consuming, which emphasizes the critical need for domain adaptation techniques, which enable the transfer of knowledge from a well-labeled source domain to a less-labeled target domain, thus alleviating the dependency on extensive labeled datasets and reducing the overall cost and effort required for model training. Traditional approaches, such as the Teacher Student Model (TSM), presuppose the availability of parallel data from both the source and the target domains, a condition that is rarely met in our scenario due to the disparities in wireless conditions and environmental variables between the source and target devices [11].

To address the domain discrepancy issue, we develop the Adversarial Domain Adaptation framework tailored for our air temperature prediction task. As Figure 6 shows, the framework consists of three components: the source domain predictor, the target domain predictor, and the MLP discriminator. The source domain predictor extracts knowledge from the source domain, while the target domain predictor extracts knowledge from the target domain. The MLP discriminator identifies whether the predictor's attention output originates from the source or target domain, aiding in the distinction between domain-specific and domain-invariant knowledge. Furthermore, the attention network within the predictor is shared across both domains, facilitating the transfer of domain-invariant knowledge from the source domain to the target domain.

Specifically, we employ a domain-specific predictor in each domain to accommodate domain-specific characteristics: f_{θ_S} in the source domain and f_{θ_T} in the target domain. Here, θ_S and θ_T represent the learnable parameters of the domain-specific predictors. The prediction loss minimizes MAE between the predicted air temperature values and the ground truth. Formally,

we have

$$\begin{aligned} \mathcal{L}_{\text{pred}} = & \frac{1}{|\mathcal{D}^{(S)}|} \sum_{(x,y) \in \mathcal{D}^{(S)}} |f_{\theta_S}(\mathbf{h}_x) - y| \\ & + \frac{1}{|\mathcal{D}^{(T)}|} \sum_{(x,y) \in \mathcal{D}^{(T)}} |f_{\theta_T}(\mathbf{h}_x) - y|, \end{aligned} \quad (5)$$

where $\mathcal{D}^{(S)}$ and $\mathcal{D}^{(T)}$ are sets of labeled samples from the source domain and the target domain, respectively. The variable \mathbf{h}_x represents the output from the dual attention network.

In addition, our framework incorporates a parametric discriminator network, denoted as $g_{\theta_D}(\cdot)$, to introduce an adversarial component into the training regimen. Here, θ_D represents the learnable parameters. This discriminator is designed to challenge the shared dual attention network, compelling it to generate embeddings that are indistinguishable between the source and target domains. The discriminator, implemented as a MLP, takes the representation embedding \mathbf{h}_x as input and outputs a prediction regarding its domain of origin. The objective during training is to optimize the discriminator’s ability to distinguish between the source and target domain embeddings, leading to the following domain loss function:

$$\begin{aligned} \mathcal{L}_{\text{domain}} = & -\frac{1}{|\mathcal{D}^{(S)}|} \sum_{(x,y) \in \mathcal{D}^{(S)}} \log(g_{\theta_D}(\mathbf{h}_x)) \\ & - \frac{1}{|\mathcal{D}^{(T)}|} \sum_{(x,y) \in \mathcal{D}^{(T)}} \log(1 - g_{\theta_D}(\mathbf{h}_x)). \end{aligned} \quad (6)$$

We employ the pretrain-and-fine-tune paradigm [12] to enhance domain-invariant feature learning and model generalization. We first train the shared dual attention network and the source domain predictor and then initialize the target predictor with $\theta_T = \theta_S$ for adversarial fine-tuning. The predictors and the discriminator are alternately trained: the predictors are optimized using $\mathcal{L}_{\text{pred}} + \lambda \mathcal{L}_{\text{domain}}$, with λ balancing prediction and domain losses, while freezing θ_D . The discriminator is then trained with $\mathcal{L}_{\text{domain}}$, keeping the predictor’s parameters static. This process repeats until convergence.

IV. EVALUATION

We conduct a series of experiments to evaluate the performance of LORATEMP. We first perform five experiments to examine the ability of LORATEMP to provide accurate air temperature predictions and compare its performance against five baselines: Weather Station Data (WS)¹, LR, MLP, LSTM, and TSM (Section IV-B). We then perform a 79-day measurement to evaluate the long-term performance of LORATEMP (Section IV-C). Finally, we examine the generalizability of LORATEMP by applying it to a publicly accessible dataset collected from a vineyard in Italy (Section IV-D).

¹This baseline uses the temperature readings provided by the nearest weather station as predictions.

A. Implementation of LORATEMP and Baselines

We implement LORATEMP and all baselines using Python 3.8 with the libraries including PyTorch, TensorFlow, NumPy, and Scikit-learn. We have performed an ablation study to select 40 as the window size and 0.01 as the flexible factor λ for our LORATEMP implementation. We use the Stochastic Gradient Descent (SGD) optimizer with a learning rate of 0.005² to train the temperature prediction model. We implement the baseline MLP with four hidden layers (128, 256, 128, and 64 neurons) and ReLU activation and set the learning rate to 0.0025 with the SGD optimizer. We implement the baseline LSTM with two layers, each of which has 128 neurons, and the SGD optimizer, set the learning rate to 0.0001, and use 0.1 as the dropout to prevent overfitting. The baseline TSM employs the knowledge distillation-based domain adaptation with a MLP model for both teacher and student neural networks. The student neural network is trained using the SGD optimizer with a learning rate of 0.0001. All training and testing are performed on a server equipped with eight NVIDIA A100 GPUs, each of which boasts a memory capacity of 40GB.

B. Performance of LORATEMP

We first examine the capability of LORATEMP to provide good temperature predictions for five LoRa end devices deployed in different locations. We deploy a standalone thermometer close to each LoRa end device to collect ground truth temperature readings. Our training data has the link and temperature measurements collected from the source LoRa device d_0 equipped with a temperature sensor over half a month and a single day of measurements gathered from each of those five target devices (d_1 , d_2 , d_3 , d_4 , and d_5). We preprocess the data before training by applying a min-max scaler with a maximum temperature difference of 16°C. We use the new data collected from the target devices as our testing data³. Table III lists MAE and the Mean Squared Error (MSE) of the temperature predictions provided by different methods. We make four observations from Table III. First, WS exhibits the least accurate performance, with an average MAE of 0.186 and an average MSE of 0.045. This discrepancy underscores the significant variation between the sensor-generated temperature readings and those obtained from weather stations, highlighting the necessity for fine-grained measurement capabilities. Second, LR and MLP demonstrate averaged MAE values of 0.139 and 0.119, respectively. Despite the Universal Approximation Theorem suggesting MLP’s potential to approximate any function [13], the suboptimal results show that the direct computation of air temperature from link metrics does not work well in real-world environments. Third, TSM, a domain adaptation method, surpasses LSTM by

²The learning rate of LORATEMP or each baseline is set to achieve the best performance.

³Please note that the MAE values reported in this section is not comparable to the ones listed in Table II. Both the training and testing data used to produce Table II are collected from the same device, which represents a much easier problem without the need of prediction knowledge transfer.

TABLE III: MAE and MSE of temperature predictions provided by LORATEMP and baselines on five LoRa end devices.

Device	WS		LR		MLP		LSTM		TSM		LORATEMP	
	MAE	MSE	MAE	MSE								
d_1	0.187	0.045	0.089	0.010	0.059	0.006	0.051	0.005	0.050	0.005	0.042	0.003
d_2	0.180	0.042	0.181	0.043	0.118	0.025	0.111	0.023	0.143	0.035	0.106	0.020
d_3	0.162	0.035	0.168	0.044	0.144	0.032	0.149	0.036	0.150	0.035	0.118	0.023
d_4	0.215	0.056	0.110	0.019	0.102	0.017	0.113	0.021	0.097	0.015	0.069	0.007
d_5	0.186	0.045	0.147	0.033	0.130	0.027	0.154	0.037	0.128	0.025	0.102	0.016
Avg.	0.186	0.045	0.139	0.030	0.119	0.025	0.135	0.032	0.114	0.023	0.088	0.014

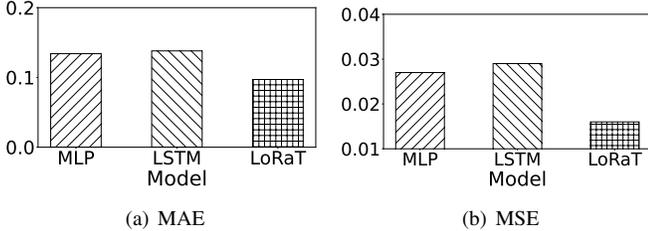


Fig. 7: MAE and MSE of temperature predictions over 78 days.

15.96% in MAE and 28.12% in MSE on average, indicating the critical role of domain adaptation in addressing discrepancies between different devices for temperature predictions. Last but not least, LORATEMP consistently outperforms all baselines, achieving MAE values of 0.042 (0.67°C), 0.106 (1.70°C), 0.118 (1.89°C), 0.069 (1.10°C), and 0.102 (1.63°C) on those five LoRa end devices. Such a superior performance, with an over 30% reduction in MAE compared to all baselines, is attributed to our unique designs presented in Section III.

C. Long-Term Evaluation

We perform a 79-day measurement to evaluate the long-term performance of LORATEMP. The training data includes the link and temperature measurements collected from the LoRa end device d_0 over half a month and from another LoRa end device d_6 over one day. The testing data is the measurements collected from d_6 over 78 consecutive days, encompassing a wide range of environmental variations. Figure 7 presents MAE and MSE of temperature predictions on the testing data made by LORATEMP, MLP, and LSTM. As Figure 7 shows, the MAE value under LORATEMP is 0.097, which is significantly lower than the ones under MLP (0.134) and LSTM (0.138). Similarly, LORATEMP significantly outperforms MLP and LSTM in terms of MSE, with a value of 0.016 compared to 0.029 under MLP and 0.041 under LSTM. The results show that LORATEMP excels in both minimizing absolute prediction errors and reducing the effects of larger deviations, highlighting its robustness in long-term forecasting.

D. Performance with a Publicly Accessible Dataset

To examine the generalizability of LORATEMP, we apply it to the publicly accessible dataset provided by Goldoni et al. [14]. The dataset consists of the data collected from eight Tinovi PM-IO-5-SM LoRaWAN end devices and a MikroTik gateway, which are deployed to monitor a 400m×30m vineyard

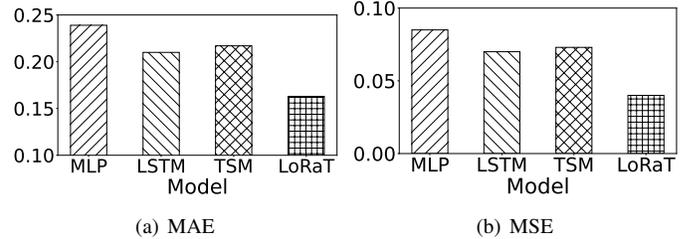


Fig. 8: Average MAE and MSE performance of temperature predictions on vineyard dataset.

in Rio Saliceto, Reggio Emilia, Italy. The dataset has the LoRa link characteristic measurements and temperature readings collected over 85 days (from November 16, 2020, to February 9, 2021). Each of the eight LoRaWAN end devices transmits every 300s and the temperature is collected from a nearby weather station every 600s. The distance between each LoRa end device and the gateway varies from 16m to 340m.

We use the data collected by one LoRa end device *tinovi-01* from November 16, 2020, to January 10, 2021, and the data gathered by another LoRa end device *tinovi-02* during a single day (January 11, 2021) as the training date. We employ LORATEMP to generate the prediction model and test it on the date measured at *tinovi-02* from January 11, 2021, to February 9, 2021. We repeat the experiments on six other end devices (*tinovi-03*, *tinovi-04*, *tinovi-05*, *tinovi-06*, *tinovi-07*, and *tinovi-08*). Figure 8 plots the averaged MAE and MSE of the temperature predictions when we employ LORATEMP and our baselines. As Figure 8 shows, the MAE value under LORATEMP is 0.163, much smaller than the ones provided by MLP (0.239), LSTM (0.210), and TSM (0.217). As Figure 8 shows, LORATEMP achieves the smallest MSE (0.04) among all solutions. The results confirm the effectiveness of LORATEMP in providing accurate air temperature predictions and transferring the prediction knowledge from one device to another.

V. RELATED WORKS

There has been increasing interest in using wireless link characteristics to perform environmental sensing. In recent years, significant efforts have been made to measure soil properties using wireless signals. For instance, Chang et al. [15] developed machine learning-based methods for soil moisture sensing by measuring the signal strength of the underground LoRa devices, which complement the solutions that use Wi-Fi [16], RFID [17], or LTE [18] signals for sensing in the soil. Chen

et al. [19] and Xie et al. [20] proposed to use LoRa signals for localization and activity recognition. However, such solutions rely on application-specific hardware, such as multiple antennas or Universal Software Radio Peripheral (USRP), and require modifications to the existing infrastructure, which is costly and even infeasible due to security and privacy concerns in many cases. In contrast to the existing solutions, this paper explores the feasibility of equipping the existing LoRa devices with air temperature sensing capability without hardware modifications.

Multivariate time series forecasting plays a pivotal role across various sectors by enabling the prediction of future trends from interrelated variables. Significant efforts have been made in the literature to explore the use of various models, such as Vector Autoregression (VAR) [21], Random Forests, Support Vector Machines (SVM), and LSTM. Recently, there has been increasing interest in applying the attention models, which address the inherent challenges of multivariate time series forecasting by enabling the decoder to access the complete encoded input sequence. Unfortunately, the attention models' applications to time series data forecasting remain relatively unexplored. Moreover, the existing solutions for time series forecasting with attention models involve complex and time-consuming training processes that do not fully leverage the temporal and feature correlations present in the time series data. In contrast, this paper proposes a dual attention network that enhances the performance of current attention-based time series prediction models, offering improved efficiency and effectiveness.

The pursuit of domain adaptation has been explored from various angles, encompassing both linear and non-linear hypotheses [22], [23]. In the realm of unsupervised domain adaptation, multiple methods have emerged to align the feature distributions between the source and target domains. Some approaches achieve this by reweighing or selecting samples from the source domain [24], while others seek explicit transformations of the feature space to map the source distribution onto the target distribution. A critical aspect of distribution matching approaches lies in measuring the (dis)similarity between the distributions. Our solution aims to encapsulate both target domain-specific and domain-invariant knowledge to ensure its robust performance on the target domain rather than employing reweighing or geometric transformations.

VI. CONCLUSIONS

This paper first demonstrates the feasibility and presents the challenges of using LoRa link characteristics to predict ambient air temperature and then presents novel deep learning-based methods that predict air temperature using LoRa link characteristics and transfer the temperature prediction knowledge between devices. Experimental results show that LORATEMP significantly outperforms all baselines and reduces air temperature prediction errors by at least 30%.

ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation under grant CNS-2150010.

REFERENCES

- [1] K. Zhang, Y. Li, J. D. Schwartz, and M. S. O'Neill, "What Weather Variables Are Important in Predicting Heat-related Mortality? A New Application of Statistical Learning Methods," *Environmental Research*, vol. 132, pp. 350–359, 2014.
- [2] J. M. Colston, T. Ahmed, C. Mahopo, G. Kang, M. Kosek, F. de Sousa Junior, P. S. Shrestha, E. Svensen, A. Turab, and B. Zaitchik, "Evaluating Meteorological Data from Weather Stations, and from Satellites and Global Models for a Multi-site Epidemiological Study," *Environmental Research*, vol. 165, pp. 91–109, 2018.
- [3] J. Wu, Y. Lu, S. Zhou, L. Chen, and B. Xu, "Impact of Climate Change on Human Infectious Diseases: Empirical Evidence and Human Adaptation," *Environment International*, vol. 86, pp. 14–23, 2016.
- [4] R. Vega-Rodríguez, S. Sendra, J. Lloret, P. Romero-Díaz, and J. L. Garcia-Navas, "Low Cost LoRa based Network for Forest Fire Detection," in *IOTSMS*. IEEE, 2019.
- [5] J. B. Cannon, L. T. Warren, G. Ohlson, J. K. Hiers, M. Shrestha, C. Mitra, E. M. Hill, S. J. Bradfield, and T. W. Ocheltree, "Applications of Low-cost Environmental Monitoring Systems for Fine-scale Abiotic Measurements in Forest Ecology," *Agricultural and Forest Meteorology*, vol. 321, p. 108973, 2022.
- [6] A. Pagano, D. Croce, I. Tinnirello, and G. Vitale, "A Survey on LoRa for Smart Agriculture: Current Trends and Future Perspectives," *IEEE Internet of Things Journal*, vol. 10, no. 4, pp. 3664–3679, 2023.
- [7] L.-Y. Chen, H.-S. Huang, C.-J. Wu, Y.-T. Tsai, and Y.-S. Chang, "A LoRa-Based Air Quality Monitor on Unmanned Aerial Vehicle for Smart City," in *ICSSE*, 2018.
- [8] S. Tong and J. Wang, "Citywide LoRa Network Deployment and Operation: Measurements, Analysis, and Implications," in *SenSys*, 2023.
- [9] C. Zhang, D. Song, Y. Chen, X. Feng, C. Lumezanu, W. Cheng, J. Ni, B. Zong, H. Chen, and N. V. Chawla, "A Deep Neural Network for Unsupervised Anomaly Detection and Diagnosis in Multivariate Time Series Data," in *AAAI*, 2019.
- [10] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Łukasz Kaiser, and I. Polosukhin, "Attention is All You Need," in *NIPS*, 2017.
- [11] J. Li, M. L. Seltzer, X. Wang, R. Zhao, and Y. Gong, "Large-scale Domain Adaptation via Teacher-student Learning," in *INTERSPEECH*, 2017.
- [12] S. Huang, D. Xu, I. Yen, Y. Wang, S.-E. Chang, B. Li, S. Chen, M. Xie, S. Rajasekaran, H. Liu, and C. Ding, "Sparse Progressive Distillation: Resolving Overfitting under Pretrain-and-finetune Paradigm," in *ACL*, 2022.
- [13] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [14] E. Goldoni, P. Savazzi, L. Favalli, and A. Vizziello, "Correlation between Weather and Signal Strength in LoRaWAN Networks: An Extensive Dataset," *Computer Networks*, vol. 202, 2022.
- [15] Z. Chang, F. Zhang, J. Xiong, J. Ma, B. Jin, and D. Zhang, "Sensor-Free Soil Moisture Sensing Using LoRa Signals," in *IMWUT*, 2022.
- [16] J. Ding and R. Chandra, "Towards Low Cost Soil Sensing Using Wi-Fi," in *MobiCom*, 2019.
- [17] J. Wang, L. Chang, S. Aggarwal, O. Abari, and S. Keshav, "Soil Moisture Sensing with Commodity RFID Systems," in *MobiSys*, 2020.
- [18] Y. Feng, Y. Xie, D. Ganesan, and J. Xiong, "LTE-Based Low-Cost and Low-Power Soil Moisture Sensing," in *SenSys*, 2022.
- [19] L. Chen, J. Xiong, X. Chen, S. I. Lee, K. Chen, D. Han, D. Fang, Z. Tang, and Z. Wang, "WideSee: Towards Wide-area Contactless Wireless Sensing," in *SenSys*, 2019.
- [20] B. Xie and J. Xiong, "Combating Interference for Long Range LoRa Sensing," in *SenSys*, 2020.
- [21] C. A. Sims, "Macroeconomics and Reality," *Econometrica: Journal of the Econometric Society*, pp. 1–48, 1980.
- [22] B. Y. Glorot, Xavier, "Domain Adaptation for Large-scale Sentiment Classification: A Deep Learning Approach," in *ICML*, 2011.
- [23] J. Shi, A. Ma, X. Cheng, M. Sha, and X. Peng, "Adapting wireless network configuration from simulation to reality via deep learning based domain adaptation," *IEEE/ACM Transactions on Networking*, vol. 32, no. 3, pp. 1983–1998, 2024.
- [24] K. M. Borgwardt, A. Gretton, M. J. Rasch, H.-P. Kriegel, B. Schölkopf, and A. J. Smola, "Integrating Structured Biological Data by Kernel Maximum Mean Discrepancy," *Bioinformatics*, vol. 22, pp. e49–e57, 2006.