使用Wi-Fi感知连续行为动作的跨域身份认证

孔浩，俞嘉地

引用本文

孔浩，俞嘉地. 使用Wi-Fi感知连续行为动作的跨域身份认证[J]. 计算机科学，2023，50(10)：299-307.

KONG Hao, YU Jiadi. Cross-domain User Authentication via Wi-Fi Sensing of Continuous Activities[J]. Computer Science, 2023, 50(10): 299-307.

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.220900163

使用 Wi-Fi 感知连续行为动作的跨域身份认证

孔　浩　　俞嘉地
上海交通大学电子信息与电气工程学院　上海 200240
(haokong@shu.edu.cn)

摘　要　目前,面向智能物联网场景的用户身份认证方法正蓬勃发展。一些工作利用室内环境中广泛存在的 Wi-Fi 信号感知用户的行为动作,并提取用户行为动作中蕴含的个体行为的独特性来实现用户身份认证。然而,用户必须在已知域背景(环境、位置、方向)下执行独立的行为动作,系统才能有效地进行身份认证。为突破现有方法的限制,提出了使用 Wi-Fi 信号感知人体连续行为动作的跨域身份认证系统 CroAuth,其能够在用户执行连续行为动作时实现跨环境、位置、方向的用户身份认证。为突破执行独立行为动作的限制,提出了基于动态时间规整的连续行为动作分离算法,在用户多样化的连续行为中分离出特定的行为动作序列,以实现有效的行为信息提取。之后,提出了基于孪生神经网络的跨域身份认证方法,提取域无关的个体行为特征,并进一步利用知识蒸馏方法构建小样本学习的跨域身份认证模型,以实现在不同环境、位置和方向下的用户身份认证。实验结果表明,CroAuth 能够在用户执行多样化的连续行为动作时,在跨环境、位置、方向的场景下对用户身份进行认证。

关键词:Wi-Fi 感知;身份认证;连续行为;跨域场景;孪生神经网络;小样本学习
中图法分类号　TP391

Cross-domain User Authentication via Wi-Fi Sensing of Continuous Activities

KONG Hao and YU Jiadi
School of Electronic Information and Electrical Engineering,Shanghai Jiao Tong University,Shanghai 200240,China

Abstract　Nowadays,Internet of Things(IoT)-based user authentication has been gradually developed. Some works utilize widespread Wi-Fi signals to sense user activities and extract individual uniqueness for user authentication. However,users must perform an independent activity under a known domain(i.e.,environment,location,and orientation),before the system can conduct user authentication. In order to break through the limitation of existing methods,this paper proposes a cross-domain user authentication method based on Wi-Fi signals,CroAuth,to realize user authentication across environments,locations,and orientations when users perform continuous activities. To release the requirement of performing independent activities,this paper proposes a continuous activity separation algorithm based on dynamic time warping,which can separate specific activity sequences from diversified continuous activities. Then,this paper designs a cross-domain user authentication method based on siamese neural network to extract domain-independent features,which can characterize essential behavioral uniqueness of each user under various environments,locations,and orientations. Finally,a knowledge distillation method is utilized to construct a few-shot cross-domain user authentication model. Experimental results show that CroAuth can authenticate users under cross-environment,location,and orientation scenarios when users perform diversified continuous activities.

Keywords　Wi-Fi sensing,User authentication,Continuous activities,Cross-domain scenario,Siamese neural network,Few-short learning

1　引言

随着物联网与人工智能的发展,智能物联网技术应运而生。用户认证作为隐私和安全的第一道屏障,是智能物联网中不可或缺的安全保障措施。面向智能物联网场景的用户身份认证能以多种方式提供隐私和安全保护。例如,防止未经授权的用户访问仅允许指定人员访问的机密文件,或防止恶意攻击者在包含敏感信息的私人设备上操作。此外,智能家庭、智能办公等新兴领域也利用身份认证能力推出定制服务,例如禁止儿童使用高危电器(如烤箱和烘干机),根据使用者特性调整房间温度或照明条件,根据个人喜好智能推荐电视内容等。这些智能环境的需求使得用户身份认证的场景更加广泛。

为了提供面向智能物联网场景的用户认证服务,一些

到稿日期:2022-09-17　返修日期:2022-12-23
基金项目:国家自然科学基金(62172277)
This work was supported by the National Natural Science Foundation of China(62172277).
通信作者:俞嘉地(jiadiyu@sjtu.edu.cn)

工作探索利用 Wi-Fi 信号来感知用户日常活动，提取日常活动中隐含的用户行为独特性，以此为基础构建机器学习模型，以对用户身份进行分类，实现用户认证[1-8]。然而，这些方法通常对用户执行行为动作的场景有着严格的要求。首先，用户必须执行独立的预定义行为动作，在此动作前后必须保持静止。然而，在现实生活中，人体行为动作往往不是完全独立的，人们通常会连续执行日常行为动作，或者在执行特定行为动作的前后附带有非刻意的附加动作，因此严格要求用户执行独立的行为动作难以符合真实场景下的认证需求。其次，现有方法大多要求用户处在预定义的环境、位置、方向来执行行为动作。然而，在使用场景中，用户所处的环境、位置、方向等域场景复杂多变，严格要求用户处在特定的域场景下进行身份认证会影响系统的可用性和用户体验。因此，现有基于无线 Wi-Fi 感知的认证工作难以支持广泛的应用场景，无法提供良好的用户身份认证服务。

基于行为动作的身份认证方法，能够在用户执行多样化的连续行为动作的情形下，实现跨域（环境、位置、方向）的用户身份认证，以支持在真实场景中高效的隐私安全保护。基于此，本文的目标是在用户执行的多样化连续行为动作中，分离出有效的行为动作成分，并抑制环境、位置、方向对人体行为特征的影响，提取每个个体独特的行为特异性，实现基于 Wi-Fi 信号感知人体连续行为的跨域身份认证。本文设计了一种基于无线 Wi-Fi 信号的身份认证系统，即 CroAuth，其能够在用户执行多样化的连续行为动作情形下，实现跨环境、位置、方向的用户身份认证。首先，在感知人体行为动作的 Wi-Fi 信道状态信息（Channel State Information，CSI）中提取出 Doppler 频移序列。基于预处理的 Doppler 频移序列，本文提出基于动态时间规整（Dynamic Time Warping，DTW）的连续行为动作分离算法，在用户执行的多样化连续行为动作中分离出特定的行为动作序列，将其作为提取个体行为特征的基础。然后，为实现跨域（环境、位置、方向）身份认证，本文利用孪生神经网络（Siamese Neural Network）和特定的训练样本选择策略，来抑制环境、位置、方向的影响，提取域无关的个体行为特征，并进一步利用知识蒸馏的方法构建小样本学习的域无关身份认证模型。使用经训练的模型进行合法用户认证和非法用户检测，系统实现了基于无线 Wi-Fi 信号感知连续行为动作的跨域身份认证。

本文的主要贡献如下：

1）提出了基于动态时间规整的连续行为动作分离算法，在用户多样化的连续行为中分离出特定的行为动作序列，实现连续行为动作下的有效行为信息提取。

2）提出了基于孪生神经网络的跨域身份认证方法，其能够提取域无关的个体行为特征，实现跨域身份认证。

3）在真实环境下进行了实验验证，结果表明其能够在用户执行多样化的连续行为动作时，在跨环境、位置、方向的场景下有效地对用户身份进行认证。

2 相关工作

近年来，Wi-Fi 设备被广泛部署在各种室内环境中，这促使了 Wi-Fi 感知应用的急速发展。一些研究利用 Wi-Fi 信号的独特感知能力实现各种无线感知应用，如人群计数[9]、动作识别[10-11]、室内定位[12-14]等。此外，另一些研究[15-17]利用机器学习或转换模型来实现更广泛场景的 Wi-Fi 感知，以提高这类应用的通用性。这些基于 Wi-Fi 感知的工作促使了无线感知应用在智能室内环境下的激增。

目前，基于行为特征的用户认证方法逐渐得到了人们的广泛关注。一些研究人员探索利用 Wi-Fi 信号来感知用户行为动作以进行身份认证。早期的研究[1-4]通过感知人类步态进行用户身份验证，随后的工作将用户认证扩展到日常活动[5-6]、交互手势[7-8]等中。然而，大部分基于 Wi-Fi 信号的方法只能在非跨域场景下实现用户身份认证，即只有用户在与获得训练数据的环境、位置、方向一致时执行动作，才能有效进行身份认证。近期，文献[18]提出了基于 Wi-Fi 信号感知的跨域身份认证。然而，此工作要求用户必须执行单一的预定义行为动作，即在动作前后必须保持静止，这极大地影响了方法的可用性。同时，此工作需要收集每个用户在各种域场景下的大量数据来训练神经网络模型，这引入了较高的训练成本，造成了用户体验的下降。因此，现有基于 Wi-Fi 信号的认证方法的场景通常是受限的，无法支持广泛的实际应用场景。

3 系统设计

本章详细介绍了基于 Wi-Fi 信号感知人体连续行为的跨域身份认证系统，其能够利用无线 Wi-Fi 信号感知人体日常的多样化连续行为动作，实现跨域（环境、位置、方向）场景下的用户身份认证。

3.1 系统概述

图 1 给出了基于 Wi-Fi 感知人体连续行为的跨域身份认证系统 CroAuth 的结构图，其分为注册阶段和认证阶段。

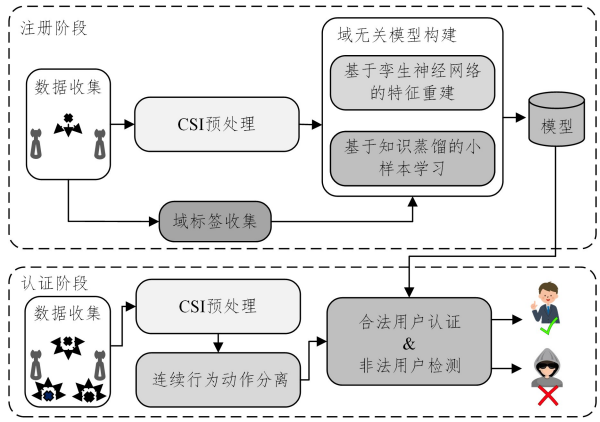


图 1 系统结构图

Fig. 1 System architecture

注册阶段收集用户在少量域场景下的行为动作数据，进行跨域身份认证模型的构建。首先，在设置有 Wi-Fi 设备收发端的环境中，用户执行行为动作进行身份注册。CroAuth 收集被用户影响的 Wi-Fi 信号并提取 CSI 数据。然后，CroAuth 对 CSI 进行预处理，在时域 CSI 数据中提取出 Doppler 频移序列。基于 Doppler 频移序列，CroAuth 利用孪生神经网络，并采用特定的训练样本选择策略，在 Doppler 频移序列

中抑制环境、位置、方向的影响，提取域无关的个体行为特征，并进一步利用知识蒸馏的方法，构建小样本学习的域无关身份认证模型。

在认证阶段，用户可在非预定的环境、位置、方向下，执行独立或连续的行为动作，以进行身份认证。CroAuth 首先获取被用户行为动作影响的 Wi-Fi 信号中的 CSI 数据；其次以与注册阶段相同的方式进行预处理，以提取 Doppler 频移序列；然后，CroAuth 通过基于动态时间规整的连续行为动作分离算法，对用户所执行的多样化连续行为动作进行分离，以在连续行为序列中提取出有效行为动作对应的 Doppler 频移序列；最后，基于提取的 Doppler 频移序列，CroAuth 利用经过训练的跨域身份认证模型进行身份识别，实现用户认证和非法用户检测。

3.2　数据预处理

人体行为特征在 Wi-Fi 信号 CSI 中以两种信息呈现，即持续时间和频率[19]。持续时间指用户完成一项活动所需的时间，而频率则揭示了用户的运动速度。由于不同用户的行为特征各不相同，CSI 中捕获的用户行为动作的持续时间和频率也不同。因此，通过时频分析，可以在 CSI 中提取用户在执行行为动作时的时间与速度变化信息，从而进一步提取用户的行为特征。Wi-Fi 信号 CSI 具有与 Doppler 雷达类似的感知能力，因此本文通过短时傅里叶变换（Short-Time Fourier Transform，STFT）[20]进行时频分析，以提取用户的行为动作特征。图 2 给出了 3 个用户执行手臂前伸时的 Doppler 频移序列。从图中可以观察到人体肢体运动引起的主频率带的变化，其中包括速度的变化情况、动作序列的间隔等行为特征。对比 3 个用户的行为动作 Doppler 序列可以看到，3 个用户的行为动作存在明显的不同，这说明通过 Wi-Fi 信号 CSI 感知用户行为动作并提取 Doppler 频移序列进行身份认证是可行的。

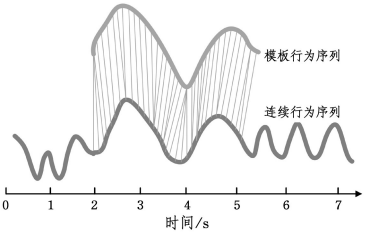


图 3　基于 DTW 的序列匹配示意图

Fig. 3　Illustration of sequence matching based on DTW

3.3　基于动态时间规整的连续行为动作分离

基于行为动作的身份认证需要提取特定行为动作下的个体行为的独特性，来进行用户身份认证。然而，在现实生活场景下，人体行为动作往往不是完全独立的。人们通常会连续执行日常行为动作，或者在执行特定行为动作的前后附带有非刻意的附加行为。因此，对人体连续行为动作进行分离，提取出可用于身份认证的特定行为动作序列，是实现基于连续行为动作的身份认证的关键。

为实现连续行为动作分离，本文借鉴动态时间规整（Dynamic Time Warping，DTW）[21]的思想，在连续行为序列中实现对特定行为动作的分离。DTW 在语音识别方面具有广泛的应用，可以计算两个语音序列的相似度，特别是对于长度不同的序列，解决了发音长短不一的模板匹配问题。本文发现，由行为动作产生的 Doppler 频移序列具有与语音序列相似的特点，其都是与用户具体行为对应的时序信号。同时，由于不同用户以不同的速度和节奏执行行为动作，因此行为动作的 Doppler 频移序列具有类似于语音序列的长度不同的特性。综上，利用 DTW 可以有效衡量两个时间序列的相似度，这为连续行为动作的分离提供了解决思路。然而，不同于计算两个确定序列之间的相似度，连续行为分离任务需要通过将模板序列与连续行为序列进行匹配，来找到连续行为序列中与模板行为序列相似性最高的部分，从而在连续行为动作中分离出特定的行为动作。图 3 为基于 DTW 的序列匹配任务示意图，基于模板行为动作序列，本文的目标是在连续行为动作序列中匹配到与模板序列一致的行为动作段，从而将特定的行为动作从连续行为动作中分离。

为实现连续行为动作中特定行为动作的分离，本文提出了基于动态时间规整的局部最短路径匹配方法。首先，为预定义的行为动作构建 Doppler 频移模板序列，其是在注册阶段收集的由多个用户执行此行为动作而产生的 Doppler 频移序列的平均值。将每个行为动作的模板序列记为 $\boldsymbol{C}=[c_1, c_2, \cdots, c_m]$，将待分离的连续行为序列记为 $\boldsymbol{Q}=[q_1, q_2, \cdots, q_n]$。之后，构建一个 $m\times n$ 的距离矩阵 $\boldsymbol{d}$，矩阵元素 (i,j) 表示 q_i 与 c_j 两点之间的欧氏距离，即 $\boldsymbol{d}(q_i,c_j)=(q_i-c_j)^2$。之后，构建累积距离矩阵 $\boldsymbol{D}$，矩阵元素 (i,j) 为当前格点距离 $\boldsymbol{d}(i,j)$ 与可以到达该点的最小的邻近元素的累积距离之和。

$$\boldsymbol{D}(i,j)=\boldsymbol{d}(q_i,c_j)+\min(\boldsymbol{D}(i-1,j-1),\boldsymbol{D}(i-1,j),\boldsymbol{D}(i,j-1)) \tag{1}$$

进一步地，构建路径长度矩阵 $\boldsymbol{l}$，矩阵元素 (i,j) 为从原点 $(0,0)$ 以最小累积距离到达此格点所经由的路径长度。基于累积距离矩阵 $\boldsymbol{D}$ 与路径长度矩阵 $\boldsymbol{l}$，构建累积距离密度矩阵 $\boldsymbol{\gamma}$，矩阵元素 (i,j) 为到达该格点的累积距离与路径长度的比值。

$$\boldsymbol{\gamma}(i,j)=\boldsymbol{D}(i,j)\boldsymbol{l}(i,j)$$

其中，累计距离密度表示从原点到达当前格点的平均累计距离。

基于累积距离密度矩阵，本文通过动态规划算法，计算在累积距离密度矩阵中的最短路径，其终点是累积距离密度矩阵最后一行中最小值对应的格点，其起点是对应于此终点的最短路径的起始点。路径终点的累积距离密度值，表示模板序列与连续行为的部分序列之间的相似度。因此，每一个模板序列通过此算法与连续行为序列进行匹配，都将产生一个累积距离密度终点。具有最小累积距离密度终点的模板序列对应于当前连续行为中的特定行为动作，并且此路径的起点和终点即为待分离的行为动作的起始点和结束点。图 4 给出了在累积距离密度矩阵中计算得到的最佳路径。可以看出，通过此算法，连续行为序列的部分与模板序列高度匹配，其匹配起点和终点即为连续行为动作中与模板序列对应的行为动作的起始点和结束点。至此，基于 DTW 的连续行为动作分离算法能够在连续行为动作序列中精确地提取出特定的行为动作序列。基于特定的行为动作序列，CroAuth 进一步提取

用户行为特征进行身份认证。

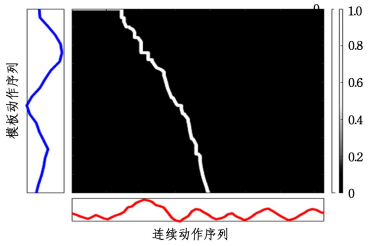


图 4 累积距离密度矩阵最佳路径示意图

Fig. 4 Optimal path in cumulative distance density matrix

3.4 基于孪生神经网络的域无关特征重建

基于感知特定行为动作的无线信号 Doppler 频移序列，CroAuth 提取用户行为动作中的个体行为特征，从而实现用户的身份认证。然而，在感知用户行为动作的 Wi-Fi 信号中，不仅包含用户的行为特征，还包含由用户所处环境、位置、方向所带来的干扰信息。由于用户身份认证依赖于细粒度的行为动作特征，域信息的干扰会影响对细粒度行为特征的提取，进而影响身份认证的表现。因此，本文拟提取域无关的行为特征，以消除域信息对行为动作特征的影响，实现跨域（环境、位置、方向）场景下的用户身份认证。

本文构建孪生神经网络（Siamese Neural Network）[22]，用于进行域无关特征的提取。孪生神经网络的基本思想是，将成对的经特殊选择的数据输入到一对具有相同结构和权值的神经网络中进行模型训练，通过优化两类输入数据的距离度量，学习自定义特征提取的能力。因此，通过孪生神经网络对各种域下输入数据的自定义学习，可抑制数据中域场景的影响，放大域无关的个体特异性，实现域无关的行为特征提取。

基于孪生神经网络的特性，本文采取特殊的训练样本选择方案。除了 Doppler 频移特征外，每个行为动作的样本还包含当前域的信息。Doppler 频移特征被用作孪生神经网络的训练输入，而域信息被用于训练样本的选择。具体来说，对于每一对样本，根据它们的域信息和个体标签将其分为 4 类。对于同一个体在不同域下的数据，其包含了此个体与域无关的一致性个体特异性；而对于相同域下不同个体的数据，其包含了不同个体与域无关的差异性个体的特异性。因此，这两类数据能够整合跨域场景下域无关的个体特异性。如果个体标签是相同的，CroAuth 选择域信息不同的训练样本，因此模型可以学习到抑制域信息的能力；而如果个体标签是不同的，CroAuth 选择域信息相同的样本，因此模型可以学习到提取个体本质行为特征的能力。基于以上策略，CroAuth 选择上述训练样本组合作为孪生神经网络的输入数据，以学习提取域无关特征的能力，从而构建跨域身份认证模型。

图 5(a)给出了所设计的孪生神经网络的体系结构。在给定一对 Doppler 频移输入时，孪生神经网络通过两个相同的子网络从输入中提取出个体行为特征，并计算个体行为特征的距离作为输入样本的相似性。子网络由一个 6 层的递归时延神经网络（Time-delay Neural Network，TDNN）组成，如图 5(b)所示。具体来说，子网络包括 6 个一维卷积块，用于在 Doppler 频移中提取行为特征，由一个全连接层将提取到的特征合并。此子网络的输入为一维 Doppler 频移序列特征，经归一化处理，大小为 1×300。每个卷积块中包含多个大小为 1×50 的一维卷积核。同时，为防止过拟合，包含一个 batch-normalization（BN）层以及 ReLU 层来对特征进行归一化。子网络的输出为 20×1 的特征向量。在多层卷积网络架构下，子网络能够从输入数据中学习到行为特征表示。

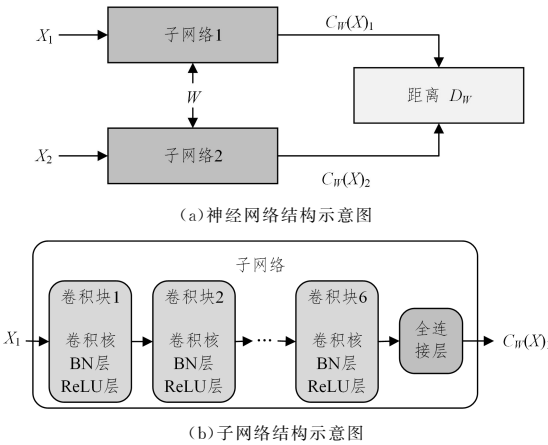


图 5 孪生神经网络示意图

Fig. 5 Illustration of siamese neural network

通过将子网络的参数表示为 W，本文设计了如下的损失函数 $L(W)$：

$$L(W)=\sum_{i=1}^{N}Y(D_W^i)^2+(1-Y)\max(M-D_W^i,0)^2 \tag{2}$$

其中，Y 是显示两个输入样本是否来自同一用户的标志变量，如果两个样本来自同一用户，则 $Y=1$，否则 $Y=0$。D_W^i 是第 i 个样本的欧氏距离，W 是代表下降间隔的边距。构造式(2)作为损失函数的原因是，当两个输入样本来自同一个用户时，损失 $L(W)$ 由距离 $D(W)$ 单调递增；而当两个输入样本来自不同的用户时，$L(W)$ 由距离 $D(W)$ 单调递减。通过最小化损失 $L(W)$，本文设计的孪生神经网络能够最小化同一用户样本之间的距离，并最大化不同用户样本之间的距离。经由多个个体的数据进行训练，所设计的孪生神经网络可以在 Doppler 频移输入中提取出域无关的个体行为特征，从而实现跨域身份认证。

3.5 基于知识蒸馏的小样本学习模型

在真实场景中，用户所处的背景环境、位置、方向等域场景复杂多变，因此收集各种场景下的大量数据来构造域无关的身份认证模型将带来较高的训练成本。为了提高身份认证系统的可用性，基于知识蒸馏的思想[23]，本文构建了一个小样本学习模型，将特征提取能力从一个庞大而繁琐的模型转移到一个更适合部署的小模型，以实现低成本的跨域身份认证模型的构建。知识蒸馏技术的特点是将复杂模型和大量数据进行训练提取的特征信息“蒸馏”给只具有少量数据的简单模型，使之在训练后也能达到与复杂模型大量数据训练后接近的效果。

本文采用图 6 所示的基于知识蒸馏的小样本训练方法。具体来说，在构建跨域身份认证模型的过程中，拟事先选取一定数量的用户，让其在不同的环境、位置、方向下采集一定量的数据，并通过一个具有多层结构的孪生网络（图 6 中的教师

网络）进行训练，学习到提取域无关个体特征的方法，将此方法作为知识“蒸馏”给学生网络。而对于学生网络，只需要少量的数据进行训练，通过结合“蒸馏”得到的知识，即可提取包含个体行为特征且与域无关的个体特异性，从而满足跨域身份认证的需求。

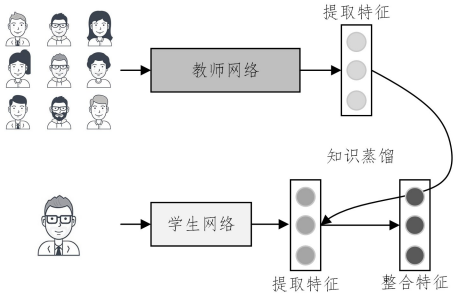


图 6　基于知识蒸馏的小样本训练

Fig. 6　Knowledge distillation-based few-shot training

具体来说，6 层的递归时延网络作为教师网络，将域无关特征提取的知识传递到一个只有两个卷积块的较小的学生网络。学生网络的卷积块为两个大小为 1×30 的卷积核，同时也包含将特征进行归一化的 BN 层以及 ReLU 层。学生网络的输入和输出特征尺寸与教师网络一致。学生网络的培训是一个双重过程，经由两个阶段进行训练。在第一个训练阶段，教师网络输出的特征（记为F_T）将作为学生网络在相同训练数据下的输出（记为F_S）的训练目标。其损失函数L_{TS}是F_T和F_S的交叉熵。

$$L_{TS}=H_{F_T}(F_S)=-\sum_i F_T\times\log(F_S)$$

在第二个训练阶段，学生网络利用特定个体的训练数据来更新网络参数，以学习提取域无关的行为特征的能力。当两训练阶段完成后，学生网络便能够在较少的训练数据情况下，实现与教师网络同样有效的域无关特征提取。

3.6　合法用户识别与非法用户检测

基于孪生神经网络和知识蒸馏模型，在小样本数据训练下，CroAuth 可以实现域无关的用户认证。具体说来，假设用户认证系统有 n 个注册用户，则对应有 n 个特征轮廓模板。当对某一用户进行身份认证时，神经网络模型将从此用户的感知数据中提取当前用户的特征轮廓。然后，CroAuth 将此特征轮廓与 n 个已注册用户的特征模板进行比较。距离最近的特征轮廓模板对应的身份，即被认定为当前用户的身份。

除了识别合法用户，登录用户还可以是试图进入系统的非法用户。因此，CroAuth 需要在用户身份认证阶段有效地检测非法用户的登录。由于认证模型是为了最大限度地扩大不同输入之间的距离而训练的，因此非法用户的数据往往与任何注册用户有较大的距离。因此，如果当前用户的特征轮廓与特征轮廓模板的最小距离大于预先通过实验验证确定的阈值，CroAuth 则会将此用户检测为非法用户。

4　实验评估

本文在真实环境下进行实验，以验证 CroAuth 的可行性和有效性。

4.1　实验设置与方法

CroAuth 以软件原型的形式被部署在装配有 Intel5300 无线网卡以及 Linux 80211n CSI tool 的笔记本电脑上。两台设备以 Monitor 模式持续不断地在 5 GHz 频段发送和接收 Wi-Fi 信号并提取 CSI。分别在 3 个室内环境下进行实验，即会议室、实验室和公寓，其环境布局如图 7 所示。接收端和发送端处在固定的位置，用户处于两者之间的感知区域中。实验预定义多个方位设定，如图 8 所示。每个感知区域有 6 个位置，每个位置有 5 个方向，以评估 CroAuth 的跨域身份认证能力。实验预定义 6 个行为动作，分别是手臂前伸、手臂上下挥动、手臂左右挥动、手臂画 Z 字、手臂画圈、双手拍动。

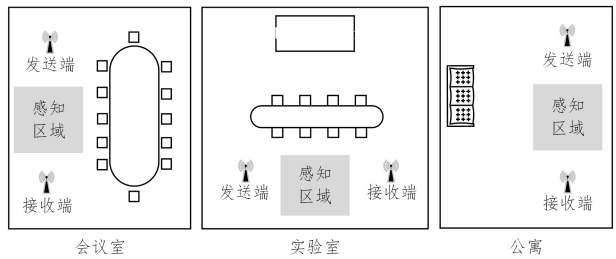


图 7　实验环境示意图

Fig. 7　Illustration of experimental environments

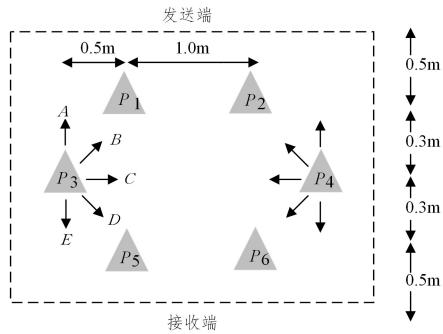


图 8　感知区域示意图

Fig. 8　Illustration of sensing area

实验募集 35 名志愿者参与实验，其年龄在 20～55 岁之间，包括 22 名男性和 13 名女性。其中，15 名志愿者的数据被作为先验数据，用来训练 6 层结构的孪生神经网络（图 6 中的教师网络）。在此训练过程中，每个用户将在 3 个环境、6 个位置、5 个方位分别执行每个行为动作 10 次，以提供教师网络所需的训练数据。其余 20 名志愿者将分别作为合法用户和非法用户使用该系统，以评估 CroAuth 的身份认证表现，其中 10 名合法用户在系统中注册其身份信息，10 名非法用户不注册其身份信息，而是尝试以合法用户的身份侵入系统。由于教师网络已经由先验数据构建，因此合法用户仅需提供少量训练数据对学生网络进行构建。合法用户在会议室内的 2 个位置 P_3 和 P_4 以及 3 个朝向 A,C,E 上提供训练数据，每个动作执行 4 次。在身份认证阶段，每个志愿者将在 3 个环境中分别以 6 个位置、5 个方位执行行为动作，并在执行此动作前后可连续地执行任意其他动作，以评估连续行为动作下进行跨域身份认证的表现。

本文实现了两个相关工作的系统原型并将其作为对比，即 WiHF[18] 和 Smart[5]。其中，WiHF 是针对跨域场景下人机交互姿势设计的身份认证系统，Smart 是针对非跨域场景下的广泛日常活动而设计的身份认证系统。WiHF 和 Smart

均要求用户执行独立的行为动作，即在预定义动作的前后必须保持静止。Smart 和 WiHF 均由常规模型训练方式实现，在本实验中分别采用 10 名合法用户提供的训练数据进行模型训练。

4.2 整体性能

本文首先评估 CroAuth 的整体用户认证精度，其代表每个用户在所有域中的认证精度的平均值。图 9 给出了 10 个合法用户（表示为 $U_1, U_2, \cdots, U_{10}$）和 10 个非法用户（表示为 Sp）的认证精度混淆矩阵。从图中可以看出，当考虑所有域下的平均结果时，本系统能够以 90.3% 的精度实现合法用户认证，以及以 88.7% 的精度实现非法用户检测。此结果表明，CroAuth 能够有效地对用户身份进行认证。此外，分析混淆矩阵中每个用户的认证精度的差距，可以计算得到 10 名合法用户认证精度的方差仅为 3.2%。此结果表明，本系统对不同年龄、性别的用户具有较好的认证鲁棒性。

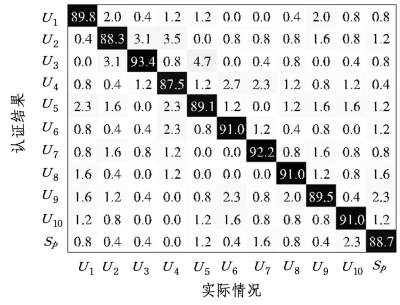


图 9 认证精度混淆矩阵

Fig. 9 Confusion matrix of authentication accuracy

本文进一步评估了 CroAuth 在检测非法用户时的误识率和识别合法用户时的拒识率。图 10 给出了在跨域和非跨域场景下，CroAuth 进行用户认证时的误识率与拒识率。从图中可以观察到，CroAuth 实现了平均 7.4% 的误识率和 10.9% 的拒识率，并且在跨域场景下的表现要略差于非跨域场景下的表现。然而，即便在跨域场景下也可以实现 8.7% 和11.9% 的误识率与拒识率。此结果表明，本系统具有良好的跨域用户认证能力，较少将非法用户识别为合法用户，并较少拒绝合法用户的请求。

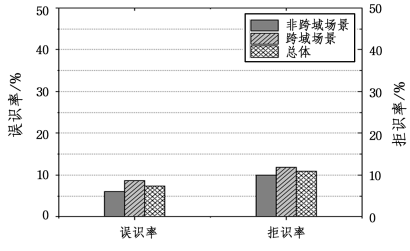


图 10 用户认证误识率与拒识率

Fig. 10 FAR and FRR of user authentication

由于 Wi-Fi 信号的传播易受非视距（None-line-of-sight，NLOS）的影响，因此本文展现在视距（Line-of-sight，LOS）和非视距情况下系统的整体用户精度表现，以探究系统在不同信号传播场景下的能力。本实验利用两种不同的材料作为阻碍物，分别为木质隔板和不锈钢板，在实验室环境中进行实验。表 1 列出了系统在不同材质的障碍物下的平均认证精度。从表中可以看出，当木质隔板作为障碍物时，系统实现了 88.6% 的认证精度，与无障碍物的 LOS 场景下的系统表现相当。这说明 Wi-Fi 信号在存在木质障碍物时依然具有一定的感知能力。然而，在不锈钢板作为障碍物时，系统仅仅取得了 77.4% 的认证精度，相比 LOS 场景下的表现大大降低。这表明不锈钢板显著地影响了信号的感知能力。尽管 Wi-Fi 信号的衍射能力较强，能够以衍射的方式绕开障碍物，但是在感知场景下，信号的衍射会导致采集的信号模式产生一定的变化，影响基于模式识别的用户认证表现。因此，NLOS 场景对系统表现具有一定的影响。

表 1 系统在不同材质障碍物下的平均认证精度

Table 1 Average authentication accuracy of system with different obstacles

（单位：%）

障碍物	无障碍物（LOS）	木质隔板	不锈钢板
认证精度	90.3	88.6	77.4

4.3 跨域性能

本文展示了 CroAuth 与两个基准系统 WiHF 和 Smart 在不同背景环境、用户位置和用户方向下的身份认证精度，以验证本系统的跨域认证能力。参考系统 WiHF 和 Smart 均采用独立的行为动作数据。

本文首先探究了 3 个系统在跨环境场景下的身份认证表现。图 11 给出了在会议室、实验室和公寓中，CroAuth，WiHF 和 Smart 这 3 个系统的认证精度比较。从图中可以看出，CrossAuth 能够在 3 个环境中分别以 91.0%，89.1%，88.5% 的平均精度实现身份认证。相比用户注册时所在的会议室，在其他两个环境中本文方法身份认证的表现并没有明显降低。然而，从其他两个基准系统的跨环境表现可以看出，WiHF 的认证精度方差较高，而 Smart 的平均认证精度有较明显的下降。这是由于，虽然 WiHF 具有跨域认证能力，但由于其采用常规的训练方法，在小样本训练前提下，无法有效提取用户本质的行为特征，造成认证模型的泛化能力减弱；而 Smart 系统并没有实现跨域认证的能力，因此在注册阶段环境变化的情况下，其行为特征提取会受到背景环境的影响，从而导致身份认证能力减弱。以上结果表明，在跨环境场景下，本系统能够有效弥补现有方法的不足，实现小样本学习下的跨环境身份认证。

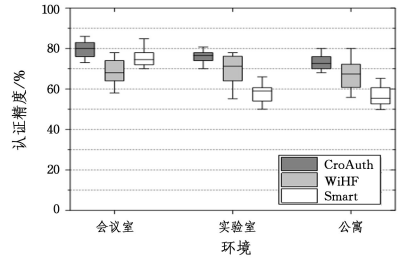


图 11 不同环境下 3 个系统的认证精度

Fig. 11 Authentication accuracy of 3 systems in different environments

本文给出了用户处在不同位置时，CroAuth，WiHF 和

Smart 的身份认证精度，以验证不同系统的跨位置认证能力。图 12 给出了在感知区域的 6 个位置下，所有用户进行身份认证的平均精确度。从图中可以看出，当用户处在收集训练数据的位置时（即 P_3 和 P_4），3 个系统都能分别以接近 90% 的精度实现身份认证。当用户处在其他位置时（即 P_1，P_2，P_5 和 P_6），CroAuth 依然实现了较高的身份认证精度（即 87.8%，88.0%，89.1% 和 88.5%）。然而，WiHF 和 Smart 在跨位置的场景下身份认证精度都出现了下降，其中 Smart 的认证精度下降极为明显，达到了平均 13.7% 的速度下降。两个对比系统精度下降的原因分别是学习样本过少和跨域能力弱。此结果表明，相比现有的系统，CroAuth 能够有效地在小样本学习下进行跨位置身份认证。

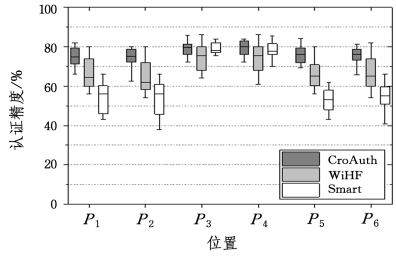


图 12　不同用户位置下 3 个系统的认证精度

Fig. 12　Authentication accuracy of 3 systems in different user locations

本文展示用户朝向不同方向时，CroAuth，WiHF 和 Smart 的身份认证精度，以验证本系统的跨方向认证能力。图 13 给出了用户分别处于 5 个方向时的身份认证精度，其中每个方向的数据是用户在所有位置的平均值。从结果中可以看出，CroAuth 不仅在提供训练数据的方向（即 A，C 和 E）上取得了较高的认证精度，在跨方向的场景下（即 B 和 D）依然取得了超过 88% 的认证精度，而两个对比系统的认证精度下降较为明显。此实验结果表明，CroAuth 相比现有系统能够在小样本学习下实现有效的跨方向身份认证。

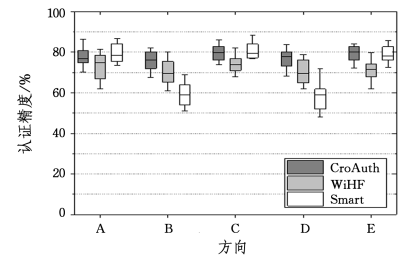


图 13　不同用户方向下 3 个系统的认证精度

Fig. 13　Authentication accuracy of 3 systems in different user orientations

4.4　连续行为动作下的用户认证性能

为展现 CroAuth 在连续行为动作下进行用户身份认证的有效性，本文分别探究了 CroAuth，WiHF 和 Smart 这 3 个系统在用户执行独立以及连续行为动作下的认证性能并对其进行了比较分析。其中，独立行为动作指用户执行独立的行为动作，并在动作前后保持静止；连续行为指用户在执行行为动作的前后，不间断地执行任意动作。由于 Smart 没有实现跨域能力，为了更好地控制变量进行验证，此实验在非跨域场景下进行。图 14 给出了 3 个系统在用户执行独立以及连续行为下的身份认证表现。从图中可以看出，CroAuth 在用户执行独立行为动作和连续行为动作时的认证精度分别是 91.2% 和 90.5%。相比之下，WiHF 和 Smart 在面对连续行为动作时，认证精度分别下降到 65.5% 和 73.3%。这表明本文所提的系统能够有效地对人体执行的连续行为动作进行分离。因此，即使用户执行连续行为动作，即在特定动作的前后任意运动，本系统依然能够有效地对用户进行身份认证。

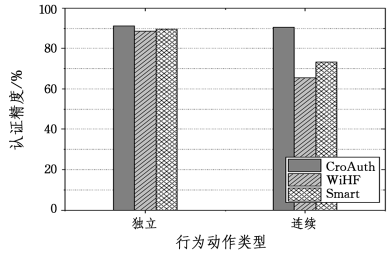


图 14　独立和连续动作下的认证精度

Fig. 14　Authentication accuracy with individual gestures and continuous gestures

4.5　训练方式的影响

基于知识蒸馏的小样本学习是 CroAuth 实现低成本身份认证系统的基础。为了衡量训练方式对本系统的影响，本文另实现了一个常规训练方式的模型。具体来说，无须训练教师神经网络以向学生网络蒸馏知识，而是直接基于 10 名合法用户的数据训练学生网络，并用经训练的学生网络模型进行身份认证。

图 15 给出了基于 4.1 节中的训练策略，在同样的训练样本大小下（即每个用户每个动作 4 个样本），小样本训练方式和常规方式分别构建的系统的认证精度。具体来说，在非跨域和跨域场景下，基于小样本训练方式的系统实现了 91.7% 和 88.9% 的认证精度，而基于常规训练方式的系统仅实现了 76.4% 和 60.1% 的认证精度。可以看出，对于跨域场景和非跨域场景，基于小样本训练方式的系统比基于常规训练的系统精度更高，尤其在跨域场景下，二者的差异更加明显。此结果表明，在训练数据有限的情况下，本文提出的基于知识蒸馏的小样本学习方法能够有效地减小有限训练数据对模型泛化能力的影响。

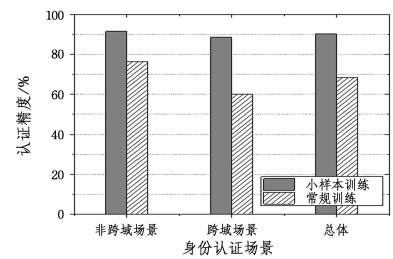


图 15　不同训练方式对用户身份认证的影响

Fig. 15　Authentication accuracy with different training manners

为了更进一步地展示基于知识蒸馏的小样本学习方式对模型训练的提升，本文进行了实验，以衡量训练样本大小对不同训练方式的影响。具体来说，本文分别将每个用户在每个动作下的训练样本数量设置为 2～18，并分别在不同的训练样本数下训练模型并验证其身份认证表现。图 16 给出了

小样本训练和常规训练两种方式在不同训练样本大小下的身份认证精度。从图中可以观察到，对于小样本训练方式，在训练样本较少的情况下(即小于等于 4)，CroAuth 便能够达到接近 90%的身份认证精度。更多的训练数据并不会显著地提升模型的表现。而对于常规训练方式，可以看出较少的训练样本极大地影响了身份认证的表现。当训练样本大小增大到 15 时，身份认证表现才逐渐与基于小样本训练方式模型接近。相比常规训练方式，基于知识蒸馏的小样本学习方式能够在降低 73.3%的训练数据成本的前提下实现同样的认证表现。此结果充分证明了本文构建的基于知识蒸馏的小样本学习方式在较少训练数据下的优势，能够在保证良好的用户体验的同时，实现传统训练方式需要大量训练数据才能实现的表现。

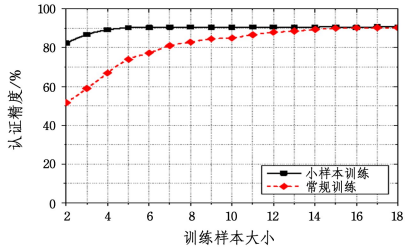


图 16 不同训练样本大小下的认证精度

Fig. 16 Authentication accuracy with different training sample sizes

5 讨论

本章讨论了本文的一些实际问题和局限性。本文实验采用了基于 Intel5300 无线网卡的笔记本电脑作为硬件。为减少硬件因素对系统表现的影响，两台设备均处于 1.3 m 的同一水平面上，构建了相对稳定且有效的信号感知区域。在实际使用场景中，用户通常会在固定的高度位置执行各类行为动作，因此本文系统在固定的高度设置下进行了实验。当设备在物理空间中的高度发生改变时，无线信号对用户行为动作的感知和刻画能力将会降低，用户认证的表现将会下降。

本文在无线网卡的收发设置中设定了 2000 每秒的采样率，从而能够以较高的采样频率获得 CSI 数据，更精细地刻画人体的行为动作。然而，较高的采样率会造成一定的计算负担，这不仅增加了对 CSI 数据提取的延迟，同时影响了后续信号处理的速度，使得系统的即时性降低。在低采样率下如何有效地获取感知数据，是仍可进行深入研究的一个方向。

通常情况下，一些环境中的物理因素，例如温度、湿度、大气压，不会对 Wi-Fi 信号的传播造成较大影响。本文在持续一段时间的实验过程中，也未发现环境中物理因素对系统表现的显著影响。因此，本系统能够在较为广泛的环境物理因素影响下正常工作。

结束语 本文提出了基于 Wi-Fi 感知人体连续行为动作的跨域身份认证系统，其能够在用户执行连续行为动作的场景下，实现跨环境、位置、方向的用户身份认证。本文的主要贡献在于，提出了基于动态时间规整的连续行为分离算法，弥补了执行单一行为动作的缺陷。其次，提出了基于孪生神经网络的特征提取方法，将用户行为特征构建为与环境、位置、方向无关的特征，从而实现跨域身份认证。此外，利用知识蒸馏方法构建小样本学习的跨域身份认证模型。在未来的研究中，实现基于 Wi-Fi 感知的人体姿态重建与用户身份认证将是一个可探索的方向，其能够同时展现人体行为姿态和对应的身份信息，促成用户在物理世界和网络世界的融合与映射。

参 考 文 献

[1] ZENG Y Z, PATHAK P H, MOHAPATRA P. WiWho: Wi-Fi-based person identification in smart spaces[C]// 2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks(IPSN). IEEE, 2016: 1-12.

[2] ZHANG J, WEI B, HU W, et al. Wi-Fi-id: Human identification using Wi-Fi signal[C]// 2016 International Conference on Distributed Computing in Sensor Systems(DCOSS). IEEE, 2016: 75-82.

[3] WANG W, LIU A X, SHAHZAD M. Gait recognition using Wi-Fi signals[C]// Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2016: 363-373.

[4] HONG F, WANG X, YANG Y, et al. WFID: Passive device-free human identification using Wi-Fi signal[C]// Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services. 2016: 47-56.

[5] SHI C, LIU J, LIU H B, et al. Smart user authentication through actuation of daily activities leveraging Wi-Fi-enabled IoT[C]// Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing. 2017: 1-10.

[6] ZHENG R Y, ZHAO Y C, CHEN B. Device-Free and Robust-User Identification in Smart Environment Using Wi-Fi Signal[C]// 2017 IEEE International Symposium on Parallel and Distributed Processing with Applications and 2017 IEEE International Conference on Ubiquitous Computing and Communications(ISPA/IUCC). IEEE, 2017: 1039-1046.

[7] SHAHZAD M, ZHANG S H. Augmenting user identification with Wi-Fi based gesture recognition [J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2018, 2(3): 1-27.

[8] KONG H, LU L, YU J D, et al. FingerPass: Finger gesture-based continuous user authentication for smart homes using commodity Wi-Fi[C]// Proceedings of the Twentieth ACM International Symposium on Mobile Ad Hoc Networking and Computing. 2019: 201-210.

[9] ZOU H, ZHOU Y X, YANG J F, et al. Freecount: Device-free crowd counting with commodity Wi-Fi[C]// 2017 IEEE Global Communications Conference(GLOBECOM 2017). IEEE, 2017: 1-6.

[10] WANG X, TANAKA J. GesID: 3D gesture authentication based on depth camera and one-class classification [J]. Sensors, 2018, 18(10): 3265.

[11] WANG C, CHANG J. CSI Cross-domain Gesture Recognition Method Based on 3D Convolutional Neural Network[J]. Computer Science, 2021, 48(8): 322-327.

[12] QIAN K, WU C S, YANG Z, et al. Widar: Decimeter-level passive tracking via velocity monitoring with commodity Wi-Fi

[C]// Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing. 2017：1-10.

[13] QIAN K，WU C S，ZHANG Y，et al. Widar2. 0：Passive human tracking with a single wi-fi link[C]// Proceedings of the 16th Annual International Conference on Mobile Systems，Applications，and Services. 2018：350-361.

[14] ZHOU C L，CHEN J D，HUANG F. WiFi-PDR Fusion Indoor Positioning Technology Based on Unscented Particle Filter[J]. Computer Science，2022，49(6A)：606-611.

[15] JIANG W，MIAO C L，MA F L，et al. Towards environment independent device free human activity recognition[C]// Proceedings of the 24th Annual International Conference on Mobile Computing and Networking. 2018：289-304.

[16] ZHANG J，TANG Z Y，LI M，et al. CrossSense：Towards cross-site and large-scale Wi-Fi sensing[C]// Proceedings of the 24th Annual International Conference on Mobile Computing and Networking. 2018：305-320.

[17] ZHENG Y，ZHANG Y，QIAN K，et al. Zero-effort cross-domain gesture recognition with Wi-Fi[C]// Proceedings of the 17th Annual International Conference on Mobile Systems，Applications，and Services. 2019：313-325.

[18] LI C N，LIU M N，CAO Z C. WiHF：Enable User Identified Gesture Recognition with Wi-Fi[C]// IEEE Conference on Computer Communications(INFOCOM 2020). IEEE，2020：586-595.

[19] WANG W，LIU A X，SHAHZAD M，et al. Understanding and modeling of Wi-Fi signal based human activity recognition[C]// Proceedings of the 21st Annual International Conference on Mobile Computing and Networking. 2015：65-76.

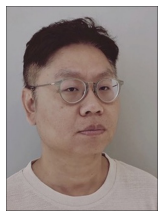
[20] GRIFFIN D，LIM J. Signal estimation from modified short-time Fourier transform[J]. IEEE Transactions on Acoustics，Speech，and Signal Processing，1984，32(2)：236-243.

[21] BERNDT D J，CLIFFORD J. Using dynamic time warping to find patterns in time series[C]// KDD Workshop. 1994，10(16)：359-370.

[22] CHICCO D. Siamese neural networks：An overview[J]. Artificial Neural Networks，2021，2190：73-94.

[23] HINTON G，VINYALS O，DEAN J. Distilling the knowledge in a neural network [J]. arXiv：1503. 02531，2015.

KONG Hao，born in 1996，Ph.D，assistant professor，is a member of China Computer Federation. His main research interests include mobile computing and wireless sensing.



YU Jiadi，born in 1975，Ph.D，associate professor，Ph.D supervisor，is a member of China Computer Federation. His main research interests include mobile computing，IOT，and wireless sensing.

（责任编辑：喻藜）