



运动序列的时空结构特征表示模型

康文轩, 陈黎飞, 郭躬德

引用本文:

康文轩,陈黎飞,郭躬德. 运动序列的时空结构特征表示模型[J]. 智能系统学报, 2023, 18(2): 240–250.

KANG Wenxuan, CHEN Lifei, GUO Gongde. Spatiotemporal structure feature representation model for motion sequences[J]. *CAAI Transactions on Intelligent Systems*, 2023, 18(2): 240–250.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202203011>

您可能感兴趣的其他文章

基于同步频繁树的时间序列关联规则分析

Association rules analysis of time series based on synchronization frequent tree

智能系统学报. 2021, 16(3): 502–510 <https://dx.doi.org/10.11992/tis.202008012>

基于时空周期模式挖掘的活动语义识别方法

Active semantic recognition method based on spatial–temporal period pattern mining

智能系统学报. 2021, 16(1): 162–169 <https://dx.doi.org/10.11992/tis.202012035>

基于知识图谱和用户长短期偏好的个性化景点推荐

Personalized attraction recommendation based on the knowledge graph and users' long–term and short–term preferences

智能系统学报. 2020, 15(5): 990–997 <https://dx.doi.org/10.11992/tis.201904064>

分段聚合近似和数值导数的动态时间弯曲方法

Dynamic time warping based on piecewise aggregate approximation and data derivatives

智能系统学报. 2016, 11(2): 249–256 <https://dx.doi.org/10.11992/tis.201507064>

基于触觉序列的物体分类

Object classification based on the tactile sequence

智能系统学报. 2015(3): 362–368 <https://dx.doi.org/10.3969/j.issn.1673–4785.201408026>

基于特征矩阵的多元时间序列最小距离度量方法

A minimum distance measurement method for a multivariate time series based on the feature matrix

智能系统学报. 2015(3): 442–447 <https://dx.doi.org/10.3969/j.issn.1673–4785.201405047>

DOI: 10.11992/tis.202203011

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.tp.20221223.1550.002.html>

运动序列的时空结构特征表示模型

康文轩^{1,2}, 陈黎飞^{1,2,3}, 郭躬德^{1,2,3}

(1. 福建师范大学 计算机与网络空间安全学院, 福建 福州 350117; 2. 数字福建环境监测物联网实验室, 福建 福州 350117; 3. 福建省应用数学中心, 福建 福州 350117)

摘要: 运动序列是一种与运动信号相关的多维时间序列, 各个维度序列之间具有高耦合性的特点。现有的多维序列表征方法大多基于维度间相互独立的假设或缺乏可解释性, 为此, 提出一种适用于运动序列的时空结构特征表示模型及其两阶段构造方法。首先, 基于空间变化事件的转换方法, 将多维时间序列变换成一维事件序列, 以保存序列中的空间结构特性。接着, 定义了一种时空结构特征的无监督挖掘算法。基于新定义的代表度度量, 该算法从事件序列中提取一组具有代表性的低冗余变长事件元组为时空结构特征。在多个人类行为识别数据集上的实验结果表明, 与现有多维时间序列表示方法相比, 新模型的特征集更具代表性, 在运动序列模式识别领域可以有效提升分类精度。

关键词: 运动序列; 多维时间序列; 特征提取; 时空特征表示模型; 空间变化; 关键子序列挖掘; 事件序列; 人类行为识别

中图分类号: TP301.6 **文献标志码:** A **文章编号:** 1673-4785(2023)02-0240-11

中文引用格式: 康文轩, 陈黎飞, 郭躬德. 运动序列的时空结构特征表示模型 [J]. 智能系统学报, 2023, 18(2): 240-250.

英文引用格式: KANG Wenxuan, CHEN Lifei, GUO Gongde. Spatiotemporal structure feature representation model for motion sequences[J]. CAAI transactions on intelligent systems, 2023, 18(2): 240-250.

Spatiotemporal structure feature representation model for motion sequences

KANG Wenxuan^{1,2}, CHEN Lifei^{1,2,3}, GUO Gongde^{1,2,3}

(1. College of Computer and Cyber Security, Fujian Normal University, Fuzhou 350117, China; 2. Digital Fujian Internet-of-Things Laboratory of Environment Monitoring, Fuzhou 350117, China; 3. Center for Applied Mathematics of Fujian Province, Fuzhou 350117, China)

Abstract: The motion sequence is a multidimensional time series associated with motion signals, and there are high coupling characteristics between different dimensional sequences. Most of the existing multidimensional sequence representation methods are based on the assumption that the dimensions are independent or lack interpretability. As a result, we propose a spatiotemporal structure feature representation model for motion sequences and its two-stage construction method. First, the multidimensional time series are transformed into one-dimensional event sequences using the transformation method of spatially changing events, preserving the spatial structure characteristics of sequences. Second, an unsupervised mining algorithm for spatiotemporal structure features is defined. A collection of representative event grams of various durations and low redundancy are extracted from the event sequence as spatiotemporal structure features using the newly proposed representation measure. The experimental findings on various datasets for recognizing human behavior demonstrate that the new model's feature set is more representative than the methods currently used to represent multidimensional time series, and it can considerably increase classification accuracy in the domain of motion sequence pattern recognition.

Keywords: motion sequence; multidimensional time series; feature extraction; spatiotemporal feature representation model; spatial transformation; key subsequence mining; event sequence; human activity recognition

例如, 在手势识别系统中, 一串手势动作可以用人手在三维空间中连续移动的坐标序列来表示, 这里, 每个坐标序列即是一个一维时间序列; 在人体行为识别中, 从所佩戴传感器上采集的数据是一个相互耦合的三维时间序列。人们可以通过 MTS 数据挖掘自动地识别手势类型、运动姿态等, 在智能家居、病人意外风险防控等领域有重要的应用价值^[2]。

序列表示 (sequence representation) 旨在将各种类型序列数据转换为固定长度的数值向量, 从而可以利用支撑向量机 (support vector machine, SVM)^[3] 等成熟的算法进行序列挖掘^[1]。显然, 序列表示模型应能有效捕捉隐含在 MTS 中的结构特征, 保留序列的时空属性^[4]: 每个维度上数据序列的变化规律以及多个维度数据序列之间的耦合关系。本文称前者为“时间结构”特征, 后者为“空间结构”特征, 统称为时空结构特征 (spatio-temporal structure feature)。通常, 序列表示过程所提取的特征还应具有良好的可解释性。

现有序列结构特征提取方法大致可以分为两类: 全局特征及局部特征的提取方法。前者的代表是基于模型的方法, 先为每条序列学习隐马尔可夫模型 (hidden Markov model, HMM)^[5-6]、自回归移动平均模型 (auto regressive integrated average model, ARIMA) 等^[7], 再基于模型参数刻画不同序列的结构特征。此类方法的性能显然依赖于预设的先验分布 (包括具体的概率分布形式及其超参数等)。另一类方法则致力于提取重要子序列模式为序列的局部结构特征, 如用于数值序列的 Shapelets^[8] 和针对符号序列的 n -gram^[9-11] 等, 这类方法所构造的特征直观、可解释性强, 但多用于单维序列挖掘^[12], 作用于多维序列时要基于“各维序列间相互独立”这一朴素假设以保证挖掘效率^[2]。

近年来, 对于复杂数据的表征任务, 深度学习方法受到广泛关注。对于序列数据, 代表性方法包括卷积神经网络 (convolutional neural network, CNN)^[13]、长短时记忆网络 (long short-term memory, LSTM)^[14] 和注意力机制网络^[15] 等, 它们利用多层神经网络的非线性特征学习能力, 从 MTS 中发掘多维耦合的时空结构特征。然而, 这些方法所挖掘的特征缺乏可解释性, 且需要支撑以大量训练数据。已有研究表明^[16], 对于小规模序列数据 (样本数量少、序列长度短), 这些方法的性能并不尽如人意, 有时甚至逊于传统的 ARIMA 等统计方法。

本文提出一种最优前缀子序列的时空结构表征新方法, 来高效提取运动序列时空特征。首先,

提出一种基于空间变化事件的变换方法, 将多维时间序列转换为一维事件序列, 用一系列事件符号表示序列中多维耦合的空间结构信息, 具有良好的可解释性; 其次, 提出变长元组挖掘方法从事件序列中提取无冗余的关键事件子序列, 作为运动序列的时空结构特征, 是一种不依赖于大规模训练样本的无监督方法。基于此, 定义了新的多维时间序列特征表示模型——时空特征提取模型 (spatio-temporal extraction model, STEM); 最后, 通过多个运动序列的模式分类实验验证了 STEM 的理想性能。

1 相关工作

本节介绍序列表示的相关背景知识, 分析若干代表性相关工作。首先约定全文使用的记号。设给定的 MTS 集为 $S = \{X^{(1)}, X^{(2)}, \dots, X^{(N)}\}$, 其中每个样本 $X^{(i)} (i = 1, 2, \dots, N)$ 由 D 个长度为 T 的序列组成, 即 $X^{(i)}$ 是一个 D 维序列; 对于运动序列, D 通常是 2 或 3。记 $X^{(i)}$ 的第 d 维 ($d = 1, 2, \dots, D$) 序列为 $x_{d1}^{(i)} x_{d2}^{(i)} \dots x_{dT}^{(i)}$, 若其中每个观测值 $x \in \mathbb{R}$, 称为数值序列; 若 $x \in E$, 则称为事件 (符号) 序列, 这里 E 是一个由有限个离散符号组成的集合, 其中每个符号代表一个事件 (event), 事件数目为 $|E|$ 。

现已提出多种直接在 S 上进行序列分类等模式挖掘的方法, 其实质是定义或学习序列间的距离 (相似性) 度量。例如, 为使用 K-近邻 (K-NN)、SVM^[3] 等基于相似度的传统分类器, 提出了针对事件序列的编辑距离和适用于多维数值序列的动态时间规整 (dynamic time warping, DTW)^[17-18], 以及基于 HMM、ARIMA 等统计模型的相似度等^[5-7]。新近的方法包括循环神经网络、LSTM^[14,19] 及任务自适应映射网络 (TapNet)^[20] 等, 基于多层神经网络将原始序列嵌入到新空间中进行距离度量和模式分类。从数据表征的角度, 这些方法构造新距离度量的过程可以看作对输入序列实施了一种“隐式”表征^[21]。

与之对应的是“显式”表征方法 (序列表示), 其目的是显式地构造每个输入样本的向量表示。以图 1 的卷积自动编码器 (convolutional auto-encoder, CAE)^[13,19] 为例, 每个序列样本 X 经过 3 层神经网络 (编码器) 转换成长度固定为 L 的数值向量 $Y \in \mathbb{R}^L$, 后续的解码器输出重构样本 X' , 使 X' 与 X 间损失最小的 Y 即为 X 的序列表示。序列表示的早期方法还包括主成分分析 (principle component analysis, PCA)、奇异值分解 (singular value decomposition, SVD)^[1,12] 等, 它们将每个序列

样本 X 视为一个 $D \times T$ 数据矩阵, 通过矩阵分解获得其低维表示, 但忽略了多维序列内在的时空结构特性; 此外, 与如图 1 所示的神经网络编码器一样, 这些方法构造的序列表示缺乏可解释性。

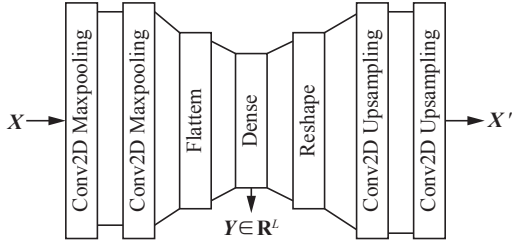


图 1 一个序列数据自动编码器的例子

Fig. 1 An example of the auto-encoder for sequences

对于一维序列, 已提出 Shapelets^[8]、 n -gram^[9-11] 等具有较好可解释性的表示方法。其原理是从序列集中提取重要的形态子序列 (shapelet)、长度为 n 的符号子序列 (称为 n -元组, 即 n -gram) 作为表示序列的结构特征, 再根据每个序列样本中这些特征的统计信息 (如元组的频度、反向文档频率 IDF^[22] 等) 构造其在新空间的表示向量。这类方法提取的特征直观、易于解释, 但面临特征规模庞大、冗余性高等问题^[23]; 更重要的是, 难以推广到多维序列。对于不满足“各维序列间相互独立”假设的多维序列 (如运动序列), 其结构特征必然是有序数值/符号间多维耦合的结果。本文提出的两阶段多维序列表示方法, 继承了可解释强这一优点的同时, 使得运动序列数据中的空间结构信息得以保存, 并有效降低了元组特征间的冗余性。

2 时空结构特征表示模型

本节详细阐述新表示模型 STEM 及其构造方法, 如图 2 所示。

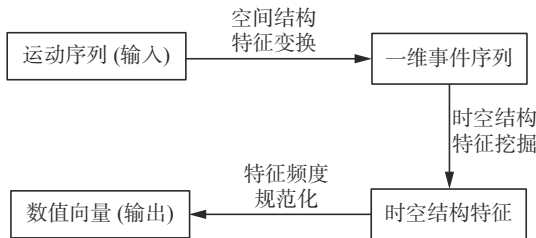


图 2 基于时空结构特征的多维序列表征方法示意图

Fig. 2 Schematic diagram of multivariate time-series representation method based on spatio-temporal structure features

2.1 时空结构特征变换

在运动序列集 S 中, 每个样本 $X^{(i)}$ 各维度 d 上都是数值序列 $x_{d1}^{(i)} x_{d2}^{(i)} \cdots x_{dt}^{(i)} \cdots x_{dT}^{(i)}, \forall t \in [1, T]: x_{dt}^{(i)} \in \mathbb{R}$ 。在一些应用场合, 各维序列可能长度不同, 此时, 令 T 为最长的序列长度, 长度短于 T 的序列追加

以其最后一个观测值 (即假设随后的观测值保持不变)。为解析多维序列之间的耦合关系, 本文提出基于空间变化事件 (spatially changing events, SCE) 表示多维序列结构的方法, 如图 3 所示。

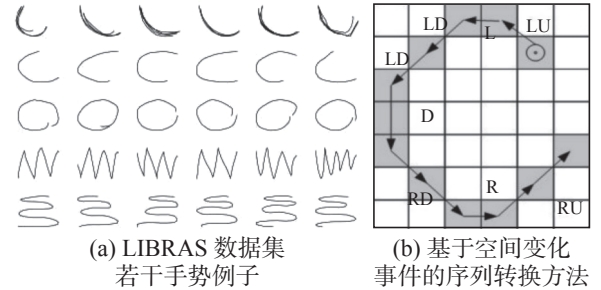


图 3 LIBRAS 数据集若干手势例子及基于 SCE 的序列转换示意图

Fig. 3 Some examples of the hand movement in the LIBRAS dataset and an illustration for the sequence transformation method based on SCE

图 3(a) 给出了来自 LIBRAS 手势轨迹序列数据集^[24] (详见 4.1 节) 的若干例子, 图 3(b) 是将二维序列 ($D=2$) 转换为 SCE 序列的示意图, 具体过程如下:

首先, 将各维坐标序列统一归一化到 $[0, 1]$ 区间, 然后从第 2 个序列点开始追踪每个点相对于前一个点的移动方向, 根据点所在的相对空间关系来检测空间变化的方向。如图 3(b) 所示, 数据空间每个维度上都有 3 种可能的空间变化, 分别是 0 (不变)、1 (正向移动) 和 -1 (反向移动)。由此, $X^{(i)}$ 的第 d 维序列观测值 $x_{dt}^{(i)} (1 \leq t \leq T)$ 可转换为符号 $m_{dt}^{(i)}$:

$$m_{dt}^{(i)} = \begin{cases} 0, & |x_{d(t-1)}^{(i)} - x_{dt}^{(i)}| \leq \delta \\ 1, & x_{d(t-1)}^{(i)} - x_{dt}^{(i)} > \delta \\ -1, & x_{d(t-1)}^{(i)} - x_{dt}^{(i)} < -\delta \end{cases} \quad (1)$$

式中: δ 是一个很小的正实数, 表示对空间变化的敏感程度。

其次, 构造序列空间变化事件 SCE 及事件集合 E 。式 (1) 所示符号的 D 维组合直观地反映了 D 维序列观测值间的耦合关系, 鉴于此, 可以通过式 (1) 中 3 种符号的排列组合来构造 SCE 集合, 即 $E = \otimes_{d=1}^D \{0, 1, -1\}$, 这里 \otimes 表示笛卡尔乘积; 每种组合对应一个 SCE。实际应用中, 为方便起见可用一些可视符号表示这些事件。以图 3(b) 中的二维序列为例, 共产生 9 个 SCE ($|E| = 3^2 = 9$), 分别用符号表示为: S (静止)、U (向上)、D (向下)、L (向左)、R (向右)、LU (左上)、LD (左下)、RU (右上) 和 RD (右下)。

最后, 将长度为 T 的 D 维序列 $X^{(i)}$ 转换为一条长度为 $T-1$ 的事件序列 $X^{(i)} = e_1^{(i)} e_2^{(i)} \cdots e_{T-1}^{(i)}$, 其

中 $e_i^{(i)} \in E$ 是空间变化组合 $(m_{1i}^{(i)}, m_{2i}^{(i)}, \dots, m_{Di}^{(i)})$ 对应的事件。该转换可以视为一种空间离散化过程, 但转换后的序列与原始序列有实质差异: 首先, 原始序列为多维序列, 而转换后的序列是单维的, 这简化了下一步时空结构特征提取操作; 其次, 更重要的, 新序列是一种离散符号序列, 其每个符号反映了各维序列数据间的耦合关系 (人体或物体在运动空间中的位置变化), 直观刻画了多维序列中的空间结构特征, 具有良好的可解释性。接下来, 用 $S = \{X^{(1)}, X^{(2)}, \dots, X^{(N)}\}$ 表示转换后的运动序列集。

2.2 表示模型 STEM

经 2.1 节的变换, 原始序列 $X^{(i)}$ 中的空间结构特征已经用 $X^{(i)}$ 中的一系列符号来表示。因此, 为运动序列 $X^{(i)}$ 提取时空结构特征相当于从 $X^{(i)}$ 中抽取若干具有代表性的符号片段 (对应运动轨迹中的动作片段, 一些例子见图 3(a))。下面, 用 $g = e_1 e_2 \dots e_n$ 表示长度为 n 的 SCE 元组 (简称 n -元组并约定 $n \geq 1$)。注意到, 1 个长片段包含了比它短的诸多动作片段; 定义 1 用前缀子序列描述此现象。

定义 1 (前缀子序列) 令 $g^{(u)} = e_1^{(u)} e_2^{(u)} \dots e_n^{(u)}$ 和 $g^{(v)} = e_1^{(v)} e_2^{(v)} \dots e_m^{(v)}$ 分别表示 1 个 n -元组和 m -元组, 且 $n < m$ 。若 $\forall t \in [1, n]: e_t^{(u)} = e_t^{(v)}$, 则称 $g^{(u)}$ 是 $g^{(v)}$ 的前缀子序列, 记作 $g^{(u)} g^{(v)}$ 。相应地, 称 $g^{(v)}$ 为 $g^{(u)}$ 的扩展子序列 (元组)。

由于实际应用场景下的动作序列或多或少地含有一些噪声, 并非所有 SCE 元组都会对应有意义 (具有代表性) 的动作片段。定义 2 给出了代表性时序结构 (representative temporal structure, RTS) 特征需要满足的条件。

定义 2 (RTS 特征) 若 SCE 元组 g 满足:

1) $\mathcal{R}_S(g) > \Delta_d$ 且 $\nexists g' : g \sqsubset g', \mathcal{R}_S(g') > \Delta_d$, 其中 $\mathcal{R}_S(g)$ 是衡量 g 在事件序列集 S 中代表能力 (表示度) 的函数, Δ_d 为表示度的最低阈值;

2) $\mathcal{F}_S(g) > \Delta_f$, 其中 $\mathcal{F}_S(g)$ 表示 g 在事件序列集 S 中出现的次数 (频度), Δ_f 为最低频度阈值,

则 g 是 S 的一个 RTS 特征, 并记由所有 RTS 特征组成的集合为 \mathcal{G} , 特征总数为 $L = |\mathcal{G}|$ 。

由定义 2 可知, RTS 特征是一系列具有代表性的、频繁的最长元组的集合。作为一个推论, RTS 特征不一定具有相同的元组长度, 这与传统的 n -gram 特征^[9]不同, 后者是一些长度固定为 n 的符号元组, 而 RTS 特征的长度可能随元组频度和表示度等统计特征的变化而发生改变 (详见第 3 节)。基于此, 定义运动序列的新表示模型 STEM 如定

义 3。

定义 3 (STEM 向量表示模型) 运动序列 $X^{(i)}$ 的 STEM 表示是一个长度为 L 的单位向量 $\mathbf{Y}^{(i)}$:

$$\mathbf{Y}^{(i)} = [y_i(g_1) \ y_i(g_2) \ \dots \ y_i(g_l) \ \dots \ y_i(g_L)]$$

这里 $\forall g \in \mathcal{G}$,

$$y_i(g_l) = \frac{\mathcal{F}_{\{X^{(i)}\}}(g_l)}{\sqrt{\sum_{j=1}^L [\mathcal{F}_{\{X^{(i)}\}}(g_l)]^2}} \quad (2)$$

根据定义 3, 原始 D 维时间序列 $X^{(i)}$ 被投影到一个由 RTS 特征构成的向量空间中, 表示为一个规范化的特征频度向量。需要说明的是, 虽然这里的特征频度是在 $X^{(i)}$ 上计算的, 但 RTS 特征本身是通过挖掘 (事件序列集 S 中的) 所有序列而得到的; 故此, RTS 特征并非均匀分布于每个输入样本中, 定义 3 将表示向量规范化为单位向量, 可以抵消由此分布不均现象带来的影响。

3 RTS 特征挖掘及 STEM 算法

根据定义 2 挖掘代表性时序结构表征, 输入是经空间结构特征变换生成的符号序列集 S , 输出是 RTS 特征集合 \mathcal{G} ; 接着, 给出构造 STEM 表示模型的算法 (STEM 算法)。

3.1 RTS 特征表示度量

考虑如下简单例子: $S = \{X^{(1)}, X^{(2)}, X^{(3)}\}$, 其中 $X^{(1)} = \text{SDUSDSDUUSDS}$, $X^{(2)} = \text{SUSDSUUSSSUS}$ 和 $X^{(3)} = \text{UDSDUUSSSUSUDSDUSDDUSSD}$; 序列包含 3 个 SCE: S (静止)、D (向下) 和 U (向上)。图 4 以前缀树 (prefix tree) 方式显示其中以 S 为前缀的 3-元组分布情况。

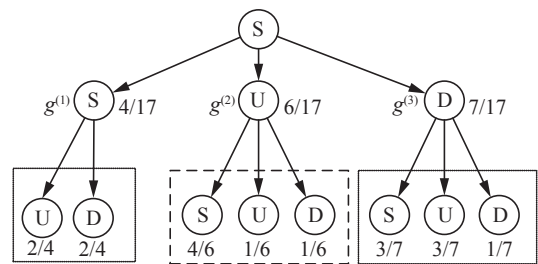


图 4 一个由 3-元组构成的前缀树例子

Fig. 4 An example of a prefix tree consisting of 3-grams

图 4 中, 树上每条从根到叶子的路径上所有节点组成 1 个 3-元组, 父节点对应其前缀子序列 (2-元组), 所附带数字表示元组的频度及兄弟元组的总频度。例如, 图上 $g^{(2)}$ 表示 2-元组 “SU”, 其频度 $\mathcal{F}_S(g^{(2)}) = 6$, 3 个 2-元组的总频度为 $\mathcal{F}_S(g^{(1)}) + \mathcal{F}_S(g^{(2)}) + \mathcal{F}_S(g^{(3)}) = 17$ 。根据定义 2, 若 $g^{(2)} = \text{SU}$ 是一个 RTS 特征, 那么以其为前缀子序列的扩展元组 (粗虚线框内的 3-元组 “SUS”、“SUU” 和 “SUD”)

就应给予删除,也就是用“SU”来代表其所有的扩展元组。这种代表性成立与否和其扩展元组的分布有关。如图所示, SUS、SUU 和 SUD 的分布相对集中(集中在 SUS 上,频率为 4/6)。此时,可视 SUU 和 SUD 元组为噪声,也因此 SU 元组可以代表比其更长的扩展元组,满足了 RTS 特征的条件。反观 $g^{(1)} = SS$ 和 $g^{(3)} = SD$, 其扩展元组分布均匀(如 $g^{(1)}$ 两个扩展元组 SSU 和 SSD 的频率均为 2/4)或接近均匀,故而缺乏类似的代表性。

上例为衡量元组 g 的表示能力(表示度) $\mathcal{R}_S(g)$ 提供了直观依据: 设 g 的所有扩展子序列组成集合 G , G 中的元组越是集中分布, $\mathcal{R}_S(g)$ 就应越大。为此, 定义如下表示度量:

$$\mathcal{R}_S(g) = \frac{|G|}{|G|-1} \times \left(1 - \sum_{g_+ \in G} (\Pr[g_+; G])^2 \right) \quad (3)$$

$$\text{s.t. } G = \{g_e | e \in E, \mathcal{F}_S(g_e) > 0\}$$

式中: $\Pr[g_+; G]$ 表示 G 中任一元组 g_+ 的概率, 按式 (4) 用其相对频度来估计:

$$\Pr[g_+; G] \triangleq \frac{\mathcal{F}_S(g_+)}{\sum_{g_+ \in G} \mathcal{F}_S(g_+)} \quad (4)$$

在极端情况 $|G|=1$ 时, 式 (3) 中乘号右侧项的值为 0, 此时 $\mathcal{R}_S(g)=0$ 。式 (3) 中乘号左侧项是一个规范化系数, 因此, 该式定义的代表度实际上是一种扩展元组概率分布散度的规范化度量, 有如下性质:

性质 1 元组 g 的表示度 $\mathcal{R}_S(g) \in [0, 1]$, 当 G 中所有元组均匀分布时取得最大值 1。

证明 首先, 引入函数 $\varphi(P) = \sum_{g_+ \in G} [p(g_+)]^2$, 其中: $p(g_+) = \Pr[g_+; G]$, $P = \{p(g_+) | g_+ \in G\}$ 。由式 (3) 可知, 若 $\varphi(P) \in \left[\frac{1}{|G|}, 1 \right]$, 则性质 1 得证。接下来, 证明 $\varphi(P)$ 的值域范围是该区间。从式 (4) 可得约束条件 $\sum_{g_+ \in G} p(g_+) = 1$, 故 φ 的极值可通过优化如下目标函数加以求解:

$$J(P) = \varphi(P) + \lambda \left(1 - \sum_{g_+ \in G} p(g_+) \right) \quad (5)$$

其中 λ 为拉格朗日乘子。

对式 (5) 所有参数求偏导得: $\forall g_+ \in G: \frac{\partial J}{\partial p(g_+)} = 2p(g_+) - \lambda = 0$ 和 $\frac{\partial J}{\partial \lambda} = 1 - \sum_{g_+ \in G} p(g_+) = 0$, 易知 $p(g_+) \equiv \frac{\lambda}{2} = \frac{1}{|G|}$ 时 J 取得极值 $|G| \times \frac{1}{|G|^2} = \frac{1}{|G|}$; 此时 Hessian 矩阵正定, 故 $\frac{1}{|G|}$ 为极小值。又因前述极端情况 $|G|=1$ 时 φ 取得极大值 1, 故 $\frac{1}{|G|} \leq \varphi(P) \leq 1$ 。

从以上证明过程可知, $\mathcal{R}_S(g)$ 越大, 表明 g 的所有扩展元组分布得越均匀, g 也因此获得越高的表示度。仍以图 4 为例, 根据式 (3) 计算得 $\mathcal{R}_S(g^{(1)}) = 1$ 、 $\mathcal{R}_S(g^{(2)}) = 0.75$ 和 $\mathcal{R}_S(g^{(3)}) \approx 0.92$, 有 $\mathcal{R}_S(g^{(1)}) > \mathcal{R}_S(g^{(3)}) > \mathcal{R}_S(g^{(2)})$, 与预想相吻合。

根据定义 2, \mathcal{R}_S 值超过阈值 Δ_d 的最长元组即为一个 RTS 特征。理想的, 所提取的 RTS 特征应能区分数据中隐含的类别模式, 譬如图 3(a) 所示的不同手势运动模式。为此, 下面以 \mathcal{R}_S 取值与预期表示度之间的差距为依据估计阈值 Δ_d 。设数据中类别模式数目为 K (实际应用中 K 可能未知, 此时根据文献 [24] 的建议估计其上界, 即 $K \triangleq \sqrt{N}$)。从式 (3) 可知预期表示度与 $1 - K/K^2 = (K-1)/K$ 成正比, 此时 \mathcal{R}_S 取值 $1 - |E| \times |E|^2 = (|E|-1)/|E|$ 。综上, 引入规范化系数后得:

$$\Delta_d = \frac{|E|}{|E|-1} \left(\frac{|E|-1}{|E|} - \frac{K-1}{K} \right) = \frac{|E|}{|E|-1} \left(\frac{1}{K} - \frac{1}{|E|} \right) \quad (6)$$

3.2 STEM 挖掘算法

由于序列集 S 的所有 n -元组可以组织成 $|E|$ 棵如图 4 所示的前缀树 (每棵树以 E 中的一个 SCE 为根节点, 每棵树的高度为 n), 同时, 由定义 2 可知所要挖掘的 RTS 特征必定分布在某棵树从根到叶节点的路径上, 故可采用类似于决策树的自下而上剪枝策略^[17] 从这些前缀树上挖掘 RTS 特征。算法 1 描述了基于此策略的挖掘算法。

算法 1 候选 RTS 特征挖掘算法 (CRTS)

输入 前缀树根节点 R

输出 无

// s_1, s_2, \dots, s_k 表示 R 的子节点, $k \geq 0$ 为子节点数。

1) 若 $k > 0$, 则遍历子节点 $s_i, i=1, 2, \dots, k$:

CRTS(s_i) // 递归调用

2) 若 R 的高度为 2, 则:

① 按式 (3) 计算 R 所代表元组 g 的表示度 $\mathcal{R}_S(g)$;

② 若 $\mathcal{R}_S(g) \leq \Delta_d$, 则将 R 置为叶节点。

算法 1 是一种递归算法。该算法递归地计算高度为 2 的节点 (即该节点有子节点且所有子节点均为叶节点) 所对应的元组的表示度, 若表示度小于阈值, 则删除其子节点; 此时, 其自身变成叶节点。算法运行结束后, 给定的前缀树就仅含有满足定义 2 中条件 1) 的元组, 但这些元组不一定满足定义中的条件 2), 故为候选 RTS (candidate RTS, CRTS) 特征, 这也是算法 1 命名为 CRTS 的原因。

在此基础上, 可以定义 STEM 算法, 如算法 2 所示。该算法按照图 1 描述的流程, 在运动序列空间结构特征变换的基础上 (步骤 1), 构造

n 元 SCE 元组前缀树(步骤 2), 调用候选 RTS 特征挖掘算法生成候选特征集 \mathcal{G} (步骤 3), 再剔除其中非频繁候选特征以满足定义 2 约定条件 2)(步骤 4)。这里, 根据文献 [25] 的建议, 设置频度阈值条件为

$$\Delta_f = \sqrt{N}$$

最后, 步骤 5 使用 \mathcal{G} 中的 RTS 特征, 根据定义 3 构造出每条序列的 STEM 表示。

算法 2 构造 STEM 表示模型的算法 (STEM 算法)

输入 运动序列集 $S = \{X^{(1)}, X^{(2)}, \dots, X^{(N)}\}$, 最大特征长度 n 。

输出 序列的 STEM 表示 $\{Y^{(1)}, Y^{(2)}, \dots, Y^{(N)}\}$

1) 用 2.1 节方法变换 S 为 \mathcal{S} , 其中 $\mathcal{S} = \{X^{(1)}, X^{(2)}, \dots, X^{(N)}\}$, 初始化 RTS 特征集 $\mathcal{G} = \emptyset$;

2) 生成 SCE 集合 E , 构造 $|E|$ 棵前缀树;

3) 对每棵 $e \in E$ 为根的前缀树:

① CRTS(树的根节点)//调用算法 1

② $\mathcal{G} = \mathcal{G} \cup \{\text{从根到叶节点每条路径代表的元组}\}$

4) 检查候选特征 $\forall g \in \mathcal{G}$ 并剔除其中 $\mathcal{F}_s(g) \leq \Delta_f$ 者;

5) 据定义 3 生成各序列样本 $X^{(i)}$ 的 STEM 表示 $Y^{(i)}$ 。

STEM 算法步骤 1) 的时间复杂度为 $O(DNT)$ 。步骤 2) 从 N 条长度为 $T-1$ 的 SCE 序列中摘取重叠式 (overlapping) n -元组, 数量为 $N \times (T-n)$, 再从中构造 n 层前缀树并统计每个唯一元组 (设总数为 M) 的频度 \mathcal{F}_s ; 因 $n \ll T$ 和 $M \ll NT$, 故时间复杂度为 $O(nNT)$ 。步骤 3) 调用的 CRTS 算法其最坏时间复杂度 $O(nM)$, 最后两个步骤的复杂度分别为 $O(M)$ 和 $O(NL)$ 。通常, 算法最终挖掘到 RTS 特征数 $L < nT$, 且运动序列的 D 相对较小, 故 STEM 算法总的时间复杂度可表示为 $O(nNT)$ 。

4 实验与分析结果

本节主要验证所提出的多维时间序列时空结构表征模型 STEM, 在实际应用中的多个运动序列数据集上进行了实验, 并与若干代表性的序列表示模型和序列分类挖掘方法相比较。所有实验在配置为 2.90 GHz i7 CPU, 16GB RAM, 操作系统为 Windows10 的个人计算机上运行。

4.1 实验数据

实验使用了来自 3 个应用领域的 4 个公开数据集, 有关信息参见表 1。第 1 个数据集 HAR (human activity recognition) 是人类活动时间序列

数据集^[26], 数据采集自智能手机设备中嵌入式 3 轴线加速度和 3 轴角速度计, 实验使用了其中 tBodyAcc_mean_x、tBodyAcc_mean_y 和 tBodyAcc_mean_z 记录的三维时间序列, 长度从 9~47 不等, 区分为站、坐、走、躺 4 种活动类型。

表 1 实验数据集的详细信息

Table 1 Details of the data sets for experiments

| 数据集 | 样本数 N | 维度 D | 类数 K | 序列长度 T (范围) | 平均 长度 |
|---------|---------|--------|--------|------------------|----------|
| HAR | 242 | 3 | 4 | [10,48] | 30 |
| LIBRAS1 | 120 | 2 | 5 | [45,45] | 45 |
| LIBRAS2 | 120 | 2 | 5 | [45,45] | 45 |
| WISDM | 51 | 3 | 4 | [557,34642] | 6319 |

接下来的 2 个数据集均为 LIBRAS 手势轨迹数据^[24], 但包含不同的手势类型。该数据集可从 UCI Machine Learning Repository 获得, 一些例子参见图 3(a)。每个手势动作其运动轨迹上 45 个二维坐标序列表示, 因此是一个定长的二维时间序列集。LIBRAS1 包含 curved-swing、anti-clockwise-arc、vertical-swing、horizontal-swing 及 clockwise-arc 5 种手势类型, LIBRAS2 的 5 种手势类型分别是 vertical-zigzag、horizontal-wavy、vertical-wavy、circle 和 face-up-curve。

第 4 个实验数据是 WISDM 数据集 (<https://www.cis.fordham.edu/wisdm/>), 来自 Wireless Sensor Data Mining(WISDM) 实验室^[27]。每个样本都是记录人体活动的三维时间序列, 对应慢跑、走路、坐和站立 4 种活动类型。该数据集的特点是样本数少但序列长度的跨度大, 如表 1 所示, 最短长度为 557, 最长的达到 34 642, 而样本数仅有 51 个; 以此验证 STEM 在不同数量和长度的序列集上的性能。

4.2 实验设置

如前所述, 序列表示方法可大致分为“隐式”和“显式”表征 2 种类型; 因此, 实验分别从中选取若干代表性工作为对比方法。为进行性能比较, 将各种方法应用于序列分类任务, 依据各数据集上取得的分类准确率来评价不同方法的性能。

实验选取了基于多层神经网络的序列分类器作为第一种隐式序列表示对比方法, 包括 LSTM^[13]和新近出版的 TapNet^[20], 二者都内嵌了多维序列数据表征学习和序列分类的功能, 可以直接输出测试样本的分类结果。实验中, LSTM 模型设置为 3 层, 分别是 LSTM 层、Dropout 层和全连接 (Dense) 层; 第 1 层的神经元个数设为 32, Dro-

pout 层的阈值为 0.5, Dense 层的神经元数目设定为数据集的真实类别数(表 1 中的 K)。TapNet 模型由三层卷积神经网络组成, 各层通道数分别设置为 64、64、32, 卷积核尺寸分别是 3、4、5。训练 LSTM 和 TapNet 时, Epochs 统一设置为 500。

第 2 种隐式序列表示方法基于序列相似性度量学习, 选取了 OWBC^[28] 和经典的 DTW 距离度量^[17] 做实验对比。其中, OWBC 基于变阶 Markov 模型衡量序列间的相似度并内置贝叶斯分类器输出分类结果; 由于 OWBC 仅适用于符号序列, 实验中, OWBC 的输入设置为各原始序列集经 2.1 节提出的时空结构特征变换方法转换而成的事件序列集 S 。

选取的显式序列表示对比方法也有两种类型。第 1 种是传统的 n -gram 模型^[21]。但该模型无法直接运用于表 1 所示的多维时间序列表示任务, 与 OWBC 一样, 其输入设置为变换后的序列集 S , 从中构造基于 n -gram 特征频度(TF)的序列向量表示模型。在此基础上, 另选取了 TF-IDF 模型^[22] 用于实验对比, 以检验不同表示模型的性能。

第 2 种显式序列表示对比方法为近年提出的神经网络自动编码器 CAE^[13,19]。如图 1 所示, CAE 具有 7 层编码器-解码器结构。实验中, 其时间步长设置为数据集中序列的最大长度 T (表 1), 长度小于 T 的序列用空值填充至长度 T ; 卷积窗口大小设置为序列的维度 D 。为公平比较, 将 CAE 全连接层的神经元数目(编码长度)设置为 L , 与 STEM 构造的 RTS 特征数相同(见定义 3)。

在性能对比实验中(4.3 节), 本文提出的 STEM 其参数 n 取值同 n -gram、TF-IDF 和 OWBC, 取 $n=\min\{15, \min_seq_len\}$, 这里 \min_seq_len

表示数据集中最短序列的长度(如表 1 所示, HAR 数据集包括长度仅为 9 的序列, 故此时设 n 为 9); 4.4 节通过实验分析不同的 n 对 STEM 性能的影响。

与 n -gram、TF-IDF 和 CAE 一样, STEM 输出所有序列的向量表示模型(忽略表 1 数据集中样本的类别标号), 通过分类算法来评价性能。实验使用了 K-NN、SVM^[3] 两种分类算法; K-NN 的近邻数固定为 1(即 1-NN), SVM 的核函数设置为线性核函数。此外, 由于 DTW 仅是一种距离度量, 其性能也采用这两种分类算法来评价(通过替换 1-NN 使用的距离度量及 SVM 中的样本相似度矩阵来实现)。

在分类实验中(包括 LSTM、TapNet 和 OWBC 内置的分类及 1-NN、SVM 分类), 每个输入序列集都随机分割为训练集(占 70%)和测试集(余下 30%), 在测试集上根据式(7)计算分类准确度(CA):

$$CA = \frac{\sum_{i=1}^{N'} I(\hat{z}_i = z_i)}{N'} \quad (7)$$

式中: N' 表示测试样本数; z_i 是第 i 个测试样本的类别标签; \hat{z}_i 为相应样本的分类结果; $I(\cdot)$ 是指示函数, 满足 $I(\text{True})=1$ 和 $I(\text{False})=0$ 。对每种对比方法和每个数据集, 上述训练-测试实验进行 100 次, 最后以“平均值±标准差”形式报告平均分类准确度。

4.3 性能对比实验

性能对比包括分类准确率和时间效率两方面。表 2 列出了不同方法在各序列集上取得的平均分类准确率, 表 3 汇总了各方法消耗的 CPU 时间(s)。

表 2 不同序列表示方法及分类器的平均分类准确率对比

Table 2 Comparison of classification accuracy obtained by different classifiers using different representation methods

| 分类器 | 表征模型 | HAR | LIBRAS1 | LIBRAS2 | WISDM |
|--------|--------|------------------|------------------|------------------|------------------|
| 1-NN | STEM | 0.70±0.04 | 0.86±0.09 | 0.89±0.05 | 0.96±0.03 |
| | CAE | 0.51±0.05 | 0.88±0.06 | 0.83±0.06 | — |
| | TF-IDF | 0.30±0.05 | 0.61±0.10 | 0.88±0.05 | 0.81±0.06 |
| | N-gram | 0.20±0.03 | 0.40±0.05 | 0.50±0.09 | 0.64±0.02 |
| | DTW | 0.47±0.05 | 0.92±0.03 | 0.78±0.07 | 0.41±0.05 |
| SVM | STEM | 0.78±0.04 | 0.90±0.06 | 0.94±0.05 | 0.97±0.04 |
| | CAE | 0.20±0.05 | 0.78±0.09 | 0.59±0.07 | — |
| | TF-IDF | 0.19±0.02 | 0.34±0.11 | 0.49±0.10 | 0.39±0.05 |
| | N-gram | 0.20±0.03 | 0.40±0.05 | 0.50±0.09 | 0.64±0.02 |
| LSTM | 内置 | 0.24±0.10 | 0.19±0.09 | 0.21±0.06 | 0.45±0.08 |
| OWBC | 内置 | 0.77±0.04 | 0.81±0.08 | 0.93±0.04 | 0.92±0.06 |
| TapNet | 内置 | 0.78±0.03 | 0.93±0.01 | 0.81±0.06 | — |

表3 不同方法完成序列分类所消耗的CPU时间

Table 3 CPU time in seconds consumed by different methods for sequence classification

s

| 分类器 | 表征模型 | HAR | LIBRAS1 | LIBRAS2 | WISDM |
|-----------|--------|-------------|-------------|-------------|--------------|
| KNN(1-NN) | STEM | 5.20 | 0.89 | 0.64 | 40.32 |
| | CAE | 15.02 | 5.64 | 5.44 | — |
| | TF-IDF | 0.58 | 0.03 | 0.17 | 37.80 |
| | N-gram | 0.47 | 0.01 | 0.12 | 33.88 |
| | DTW | 17.59 | 2.85 | 2.85 | 894.50 |
| SVM | STEM | 5.97 | 0.85 | 0.62 | 41.14 |
| | CAE | 15.97 | 5.45 | 6.01 | — |
| | TF-IDF | 1.63 | 0.18 | 0.97 | 43.23 |
| | N-gram | 0.54 | 0.05 | 0.11 | 40.97 |
| LSTM | 内置 | 8.39 | 3.74 | 4.31 | 661.14 |
| OWBC | 内置 | 4.55 | 0.51 | 0.94 | 58.09 |
| TapNet | 内置 | 38.22 | 42.02 | 42.03 | — |

根据表2的结果,基于STEM的SVM分类器在3个数据集上都取得了最高的分类准确率,在LIBRAS1上略逊于TapNet。使用1-NN分类器时,除该数据集外,STEM的准确率大幅高于其他表示模型。4种隐式表示方法(DTW、LSTM、OWBC和TapNet)的性能总体上高于除STEM外的其他显式表示方法,这是因为这些方法内置了与STEM类似的特征优化过程。基于 n -gram特征的两种方法虽使用了经本文的时空结构特征变换方法转换而来的事件序列,但缺乏STEM的特征挖掘过程,大量的冗余特征降低了其有效表示序列的性能。

基于神经网络的方法CAE和TapNet,在序列长度较短的3个数据集(HAR、LIBRAS1和LIBRAS2)上取得了较高的分类准确率,但因样本数较少,LSTM仅达到与 n -gram相当的性能。然而,在拥有长序列的WISDM数据集上,CAE和TapNet均无法在可接受的时间内完成神经网络模型的训练,故而表2(和表3)未报告其相应结果。受序列长度大幅波动的影响(见表1),基于

序列“对齐”的DTW在该数据集上的分类准确率甚至低于 n -gram方法。

时间效率方面,如表3所示,神经网络方法和DTW需要的运行时间明显高于其他方法,且随序列长度增加而大幅增长。与 n -gram和TF-IDF相比,同样是显式序列表示方法的STEM效率略低,这是因为STEM增加了递归剪枝步骤(见算法2)挖掘具有高代表度的低冗余特征,换取了分类准确率大幅提升的效果,是值得的。

4.4 RTS特征挖掘实验

本部分实验的目的是进一步分析STEM算法提取的RTS特征及算法参数 n 的影响。根据定义2,RTS特征是STEM算法提取的一种变长 n -元组。表4列出了 n 为3、6、9、12、15时算法为4个数据集挖掘的RTS特征数量,并与定长 n -元组特征(STEM算法剪枝前的元组特征)相比较。因HAR数据集中的最短序列长度为9,表中未报告该数据集上 $n>9$ 的实验结果。实验数据为4个原始序列集经算法2的步骤1生成的事件序列集。

表4 4个事件序列集的 n -元组特征与RTS特征比较Table 4 Comparison between the n -grams and the RTS features on the four event-sequence sets

| 数据集 | $n=3$ | | $n=6$ | | $n=9$ | | $n=12$ | | $n=15$ | |
|---------|--------|-------|---------|--------|---------|--------|---------|--------|---------|--------|
| | SCE | RTS | SCE | RTS | SCE | RTS | SCE | RTS | SCE | RTS |
| HAR | 3 563 | 134 | 5 531 | 120 | 4 922 | 101 | — | — | — | — |
| LIBRAS1 | 117 | 322 | 1 200 | 235 | 1 626 | 253 | 1 725 | 243 | 1 679 | 212 |
| LIBRAS2 | 275 | 125 | 1 394 | 345 | 2 668 | 367 | 3 087 | 344 | 2 954 | 304 |
| WISDM | 14 627 | 3 351 | 282 989 | 34 747 | 663 428 | 38 579 | 738 462 | 38 605 | 757 230 | 38 586 |

从表 4 的对比结果可以看出, 经过 RTS 特征挖掘, 特征数量较之于 n -元组大幅减少, 随着 n 取值增大, 降幅更加明显。例如, 在 HAR 数据集上, 当 $n=3$ 时, 特征数从 3563 降为 134, 约简幅度约 96%。通过约简, 绝大部分冗余特征被剔除; 从序列表示的角度看, RTS 特征挖掘方法使得 STEM 的特征集变得更具有代表性。

图 5 给出不同长度的 RTS 特征分布情况, 其结果表明短元组特征占有较大比例: RTS 特征多集中在长度为 2 (HAR 数据集)、3 (LIBRAS1 及 LIBRAS2) 5 或 6 (WISDM) 的元组上。需要说明的是, 根据算法 2, 长度越短的特征代表了越多的 (以其为前缀的) 其他元组, 这从另一个侧面反映了 STEM 特征代表性高的事实。

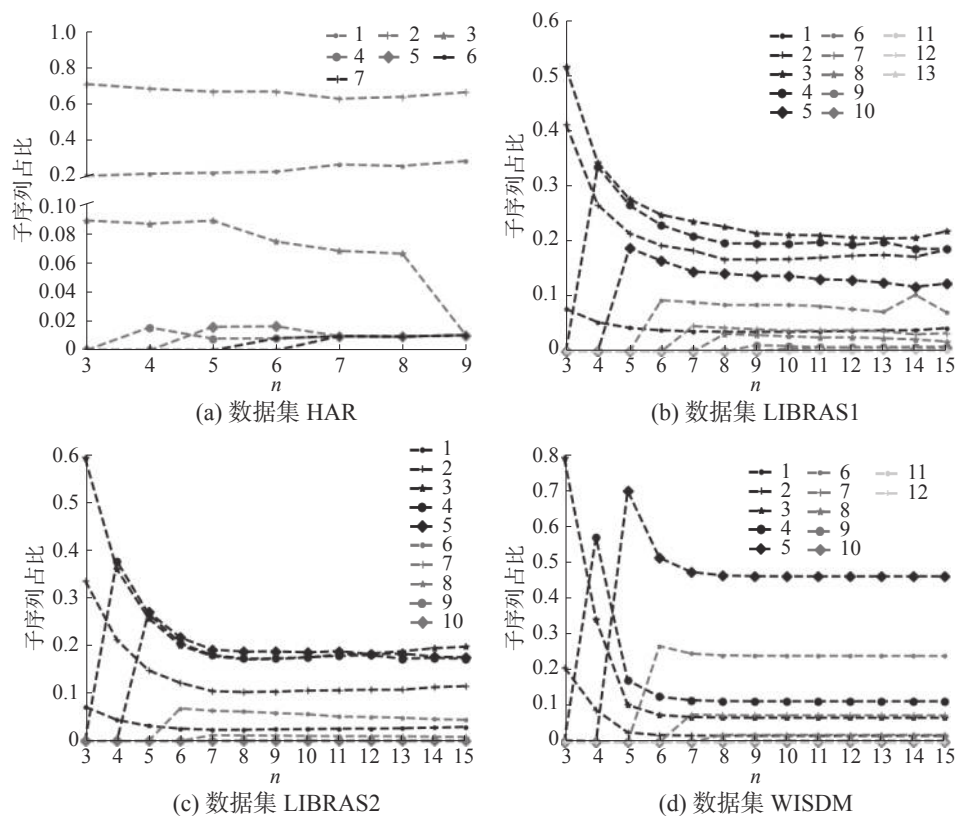
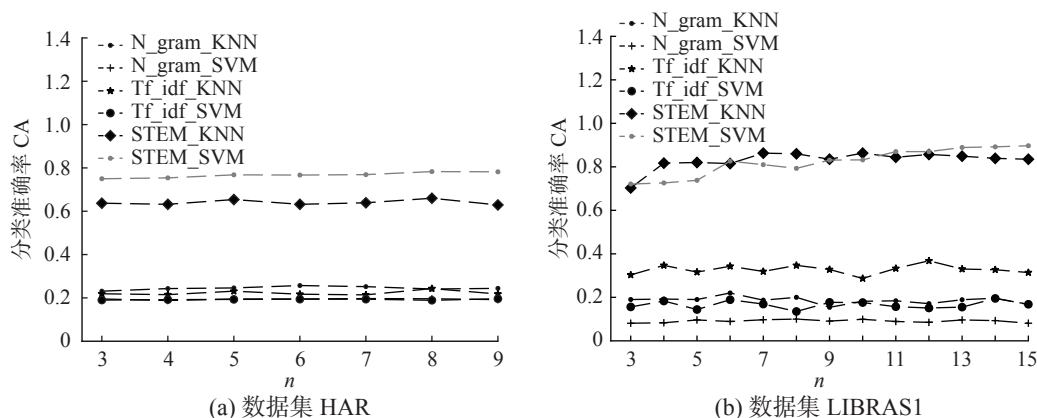


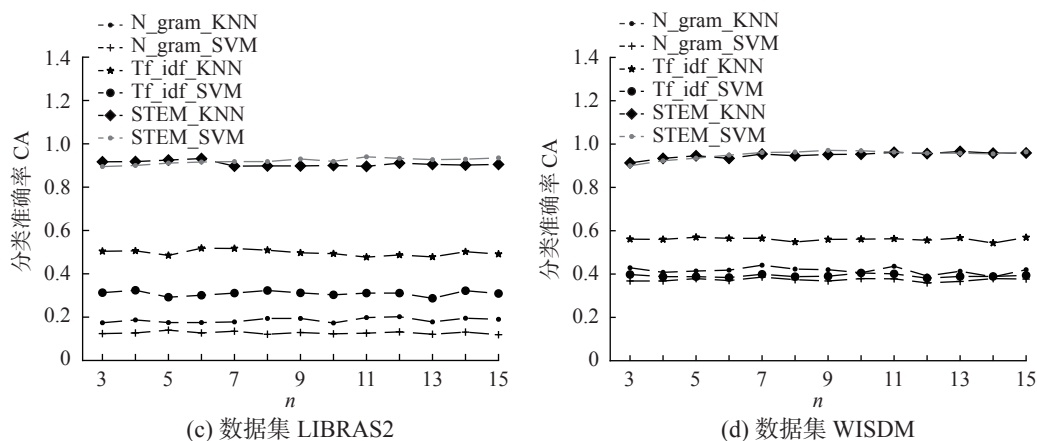
图 5 各数据集上不同长度 RTS 特征分布情况

Fig. 5 Distribution of the variable-length RTS features in different data sets

图 6 给出了 n -gram、TF-IDF 和 STEM 3 个基于元组的序列表示模型其分类性能随参数 n 变化的情况。图上显示了每个模型应用于 1-NN 和 SVM 分类时取得的平均准确率。如图所示, 随

着 n 增大, STEM 的分类准确率趋于平稳, 且明显高于其他两个模型。上述结果表明, 本文提出的 RTS 特征挖掘方法具有较高的鲁棒性, 可以有效挖掘多维时间序列中的时空结构特征。



图6 基于元组的序列表示方法的分类准确率随参数 n 变化情况Fig. 6 Change in classification accuracy obtained by the gram-based representation methods with different parameters (n)

5 结束语

针对现有多维时间序列的特征表示方法存在未充分考虑维度间关联性或可解释性差的问题,提出了时空结构特征表示模型 STEM。通过空间转换的方法,将多维时间序列变换成由空间变化事件组成的一维序列;接着定义了一个新的时空特征挖掘方法,从序列中提取一系列非冗余的变长元组作为时空特征,挖掘序列中蕴含的时空属性。新方法不仅克服了基于维度间独立性假设的特征表示模型忽略空间结构信息的不足,还可以自动地选择具有代表性的时空结构特征集。经过运动序列数据集的实验验证,与现有其他模型相比,STEM 模型在实验数据上获取的特征质量有明显的提升。通过一维事件序列构建的特征可直接观察到,有效提高了模型的可解释性。

相较于运动序列,高维时间序列数据维度间关联程度更加复杂。因此下一步的工作将集中于研究高维时间序列的特征表示模型,在确保时空结构特性得到保留的同时,探索基于聚类方法的序列表示模型构造方法。

参考文献:

- [1] 谭海龙. 多维时间序列的分类技术研究 [D]. 杭州: 浙江大学, 2015: 1–20.
TAN Hailong. Multivariate time series classification technology research [D]. Hanzhou: Zhejiang University, 2015: 1–20.
- [2] 任守纲, 张景旭, 顾兴健, 等. 时间序列特征提取方法研究综述 [J]. 小型微型计算机系统, 2021, 42(2): 271–278.
REN Shougang, ZHANG Jingxu, GU Xingjian. Overview of feature extraction algorithm for time series [J]. Journal of Chinese mini-micro computer systems, 2021, 42(2): 271–278.
- [3] CORTES C, VAPNIK V. Support-vector networks [J]. Machine learning, 1995, 20: 273–297.
- [4] 曹翰林, 唐海娜, 王飞, 等. 轨迹表示学习技术研究进展 [J]. 软件学报, 2021, 32(5): 1461–1479.
CAO Hanlin, TANG Haina, WANG Fei. Survey on trajectory representation learning techniques [J]. Journal of software, 2021, 32(5): 1461–1479.
- [5] RABINER L R. A tutorial on hidden Markov models and selected applications in speech recognition [J]. Proceedings of IEEE, 1989, 77: 257–286.
- [6] 陈炳鑫, 陈黎飞. 符号序列的预训练 HMM 分类方法 [J]. 南京大学学报 (自然科学版), 2021, 57(1): 52–58.
CHEN Bingxin, CHEN Lifei. A pre-training HMM classification method for symbolic sequences [J]. Journal of Nanjing University (natural science edition). 2021, 57(1): 52–58.
- [7] HYNDMAN R, KHANDAKAR Y. Automatic time series forecasting: the forecast package for R [J]. Journal of statistical software, 2008, 26(3): 1–22.
- [8] YE L, KEOGH E J. Time series shapelets: a novel technique that allows accurate, interpretable and fast classification [J]. Data mining and knowledge discovery, 2010, 22(1-2): 149–182.
- [9] XING Z, PEI J, KEOGH E. A brief survey on sequence classification [J]. SIGKDD exploration newsletter, 2010, 12(1): 40–48.
- [10] GUO Gongde, CHEN Lifei, YE Yanfang, et al. Cluster validation method for determining the number of clusters in categorical sequences [J]. IEEE transactions on neural networks and learning systems, 2017, 28(12): 2936–2948.
- [11] FINK G. Markov models for pattern recognition: from theory to applications [M]. New York: Springer, 2014: 119–201.

- [12] 钱爱玲. 复杂结构的时间序列数据挖掘与预测方法研究. [D]. 武汉: 华中科技大学, 2011: 13–17.
QIAN Ailing. Research on data mining and forecasting methods over time series with complex structure[D]. Wuhan: Huazhong University, 2011: 13–17.
- [13] KASHIPAREKH K, NARWARIYA J, MALHOTRA P, et al. ConvTimeNet: a pre-trained deep convolutional neural network for time series classification[C] //Proc of the 2019 International Joint Conference on Neural Networks. Budapest: IEEE, 2019: 1–8.
- [14] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. *Neural computation*, 1997, 9(8): 1735–80.
- [15] ASHISD V, SHASZEER N, PARMAR N, et al. Attention is all you need[C]//Proc of the 31st International Conference on Neural Information Processing Systems. Long beach: Curran Associates Inc., 2017: 5998–6008.
- [16] MAKRIDAKIS S, SPILOTIS E, ASSIMAKOPOULOS V. Statistical and machine learning forecasting methods: Concerns and ways forward[J]. *PLoS ONE*, 2018, 13(3): e0194889.
- [17] BERNDT D, CLIFFORD J. Using dynamic time warping to find patterns in time series[C]//Proc of the 3rd International Conference on Knowledge Discovery and Data Mining. Seattle: AAAI Press, 1994: 359–370.
- [18] WANG Jichen, ZHU Weiguo, SUN Yongqi, et al. An effective dynamic spatiotemporal framework with external features information for traffic prediction[J]. *Applied Intelligence*, 2021, 51: 3159–3173.
- [19] SHI Xingjian, CHEN Zhourong, WANG Hao, et al. Convolutional LSTM network: a machine learning approach for precipitation nowcasting[C]//Proc of the 28th Annual Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2015: 802–810.
- [20] SUNG W, JUN S, JAEKYUN M. TapNet: neural network augmented with task-adaptive projection for few-shot learning[C]//Proc. of the 36th International Conference on Machine Learning. Long Beach: Curran Associates Inc., 2019, 97: 7115–7123.
- [21] XIONG Tengke, WANG Shengrui, JIANG Qingshan, et al. A novel variable-order Markov model for clustering categorical sequences[J]. *IEEE transactions on knowledge and data engineering*, 2014, 26(10): 2339–2353.
- [22] BASU T, MURTHY C. A supervised term selection technique for effective test categorization[J]. *International journal of machine learning and cybernetics*, 2016, 7: 877–892.
- [23] YUAN Liang, WANG Wenjian, CHEN Lifei. Two-stage pruning method for gram-based categorical sequence clustering[J]. *International journal machine learning and cybernetics*, 2019, 10: 631–640.
- [24] DIAS D, MADEO R, ROCHA T, et al. Hand movement recognition for Brazilian Sign Language: a study using distance-based neural networks[C]//Proc of the 2009 International Joint Conference on Neural Networks. Atlanta: IEEE Computer Society, 2009: 697–704.
- [25] BEZDEK J. Pattern recognition in handbook of fuzzy computation[M]. Bristol: IOP publishing Ltd., 1998: 1–30.
- [26] DAVIDE A, ALESSANDRO G, LUCA O, et al. A public domain dataset for human activity recognition using smartphones[C]//Proc of 21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges: Louvain-la-Neuve, 2013: 24–26.
- [27] JENNIFER K, GARY W, SAMUEL M. Activity recognition using cell phone accelerometers[J]. *ACM SIGKDD explorations newsletter*, 2010, 12(2): 74–82.
- [28] 程铃纺, 郭躬德, 陈黎飞. 符号序列多阶 Markov 分类[J]. *计算机应用*, 2017, 37(7): 1977–1982
CHENG Lingfang, GUO Gongde, CHEN Lifei. Classification of symbolic sequences with multi-order Markov model[J]. *Journal of computer applications*. 2017, 37(7): 1977–1982.

作者简介:



康文轩, 硕士研究生, 主要研究方向为序机器学习、数据挖掘。



陈黎飞, 教授, 博士, 中国人工智能学会机器学习专业委员会委员, 主要研究方向为机器学习、数据挖掘和模式识别。发表学术论文 120 余篇。



郭躬德, 教授, 博士, 中国人工智能学会机器学习专业委员会常务委员, 主要研究方向为人工智能、量子计算、机器学习和数据挖掘技术及应用。发表学术论文 100 余篇。