

文章编号: 2095-4980(2022)01-0008-08

基于电磁环境数据的信息挖掘与关联分析

李 爽¹, 刘海鹏², 郭兰图^{*3}

(1.哈尔滨工程大学 信息与通信工程学院, 黑龙江 哈尔滨 150000; 2.装备发展部驻济南地区军事代表室, 山东 济南 250100;
3.中国电波传播研究所, 山东 青岛 266107)

摘 要: 城市通信技术的改革换代和用频设备的逐渐增多使得电磁环境变得越来越复杂。充分了解频谱资源利用的特性是提高频谱管理效率的关键。为了更全面地探索频谱利用的特点, 提出一套完整的对复杂多样电磁环境大数据进行详细数据质量分析和处理的流程, 分别对处于同一服务的不同信道、处于不同服务的不同信道进行频谱相关性分析, 证明了频谱之间的相关性; 对电磁环境大数据进行属性构造, 构造了频率维占用度和时间维占用度属性。引入图像处理领域的多维混合高斯模型, 对电磁信号进行背景噪声的去除, 提取电磁信号, 为后续的信息挖掘和关联分析奠定基础。

关键词: 关联分析; 属性构造; 图像处理; 多维混合高斯模型; 电磁环境数据

中图分类号: TN97

文献标识码: A

doi: 10.11805/TKYDA2021168

Information mining and association analysis based on electromagnetic environment data

LI Shuang¹, LIU Haipeng², GUO Lantu^{*3}

(1.School of Information and Communication Engineering, Harbin Engineering University, Harbin Heilongjiang 150000, China;
2.Military Representative Office of Equipment Development Department in Jinan Area, Jinan Shandong 250100, China;
3.China Research Institute of Radio Wave Propagation, Qingdao Shandong 266107, China)

Abstract: With the development of urban communication technology and the increase of frequency equipment, the electromagnetic environment becomes more and more complex. Fully understanding the characteristics of spectrum resource utilization in the past is the key to improve the efficiency of spectrum management. A complete process about detailed data quality analysis for big data in complex and diverse electromagnetic environment is proposed, in order to explore the characteristics of spectrum utilization more comprehensively. The spectrum correlation for different channels in the same service, and for different channels in different services, is performed. Attribute construction is carried out for big data of electromagnetic environment, including the attributes of frequency dimension occupancy and time dimension occupancy. The multi-dimensional Gaussian mixture model in the field of image processing is introduced to remove the background noise of the electromagnetic signal and extract the electromagnetic signal, which can lay the foundation for the subsequent information mining and association analysis.

Keywords: association analysis; attribute construction; image processing; multi-dimensional Gaussian mixture model; electromagnetic environment data

无线电行业和监测类设备不断的更新迭代促进了频谱监测应用步入大数据时代。5G 的迅猛发展和各种电磁设备的增加, 使公众持续暴露于无线通信网络(如 GSM,Wifi,WiMAX 等)或广播网络(如 AM,FM,DAB,DVB-T, DVB-H)的天线所发射的射频电磁场中^[1], 因此对电磁环境大数据的信息挖掘与关联分析与人们密切相关。

与此同时, 电磁环境大数据采集设备和挖掘算法也在不断进步^[2]。Hamid Eltom 等^[3]提出频谱占用预测分析, 认为频谱感知测量中的时间相关性可用于预测主要用户活动模式, 他们所研究的基于软融合的协作频谱占用预测

收稿日期: 2021-04-21; 修回日期: 2021-05-01

*通信作者: 郭兰图 email:guolantu@163.com

的概念和方法在预测误差上有显著提高；文献[4]提出对认知无线电系统中主要信道的业务模式进行分类的方法，并对不同类型的业务应用不同的预测规则，更准确地预测主信道的空闲时间，利用智能信道筛选方案根据预测结果来寻找空闲时间最长的信道以供二次使用；文献[5]提出一种敏感的关联规则挖掘算法 EFO4ARH，该算法采用两个适应度函数，以达到副作用最小解，减少了运行时间，优化了算法性能；李斌等^[6]提出了一种快速检测异常频谱使用和异常电磁目标(出现在异常时空)的新方法。针对频谱数据的处理，提出了基于马氏距离的分段建模算法来发现频谱数据中的异常，实时检测出两个频谱之间的差异；艾小锋等^[7]在雷达系统电磁环境效应问题上，研究了电磁环境元的表示方法，建立了复杂雷达电磁环境的信号模型，初步结果验证了研究思路的正确性。当前，电磁数据的隐藏信息挖掘和综合利用难度仍然较大。一方面，电磁数据种类繁多，呈现多源异构特征^[8]，电磁设备和台站的丰富多样化，频谱资源利用技术的超前发展^[9]，导致非结构化和半结构化数据容量快速扩张^[10]；另一方面，电磁数据管理技术不够完善，缺少有效的对潜在关键信息的挖掘和处理方法^[11]。目前常用的数据库管理技术只处理单一类型的数据，无法实现多类型数据的融合分析和相关性分析^[12]。电磁环境数据涵盖多维域，如“能量域”、“频域”、“时域”和“空域”^[13]，电磁大数据具有时间尺度大、空间跨度大、环境信息动态变化等特点^[14]。因此，迫切需要一个能够从多角度、深层次挖掘电磁大数据、实现相关性分析的模型。

本文对复杂多样电磁环境大数据进行数据质量分析和处理的工作流程进行了总结，对处理后的去量纲频谱数据进行最大类间方差的二值化，避免引入人为经验选取阈值方法的主观因素；分别分析了处于同一服务频段的信道的相关性和处于不同服务频段的信道相关性，并对电磁环境大数据进行属性构造，从而加快大规模电磁环境数据的处理速度，高效提取占用度和相关性等信息。运用多维混合高斯模型，对前景电磁信号进行提取，为后续的电磁环境数据挖掘提供了思路。

1 数据处理技术

1.1 数据质量分析

数据质量分析是整个电磁环境数据分析过程中不可缺少的一个环节，质量分析结果直接关系到模型效果和最终结论。数据清洗是利用一定的规则和策略对数据中的“脏”数据进行处理，生成符合要求的数据，不同类型的数据需不同的处理方法。数据质量分析的整体框图如图 1 所示。

第一步：缺失值分析。对由于设备故障、随机缺失或条件缺失导致的缺失数据要进行缺失数据的填充，填充方法有 3 种：a) 通过专业知识或经验推测进行缺失数据的填充；b) 缺失值使用相同指标的计算结果，如平均值、中位数、众数等进行填充；c) 缺失值也可用不同指标的计算结果进行填充，如数据拟合。

第二步：异常值检索。对于数据中的局部离群因子即异常数据值，先分析异常值出现的原因，再判断异常值是否该舍弃：若是正确的数值，则对离群值的数据集进行挖掘、分析；若为错误的数值，则采用平均值修正法，舍弃原来的数据，然后用前后两个观测值修正该异常值。

第三步：一致性分析。对目标数据集进行数据挖掘过程中，数据集成环节会产生不一致的数据，这是由于数据挖掘算法对重复存放或不同源数据没有进行一致性分析。数据来自不同的数据源，并不能对重复存储的数据进行一致的更新。

针对上述的异常值检索环节，本文采用箱型图分析法。箱型图可起到对异常值的定位、识别功能，异常值的表达式为：

$$Q_L - 1.5IQR \leq W \leq Q_U + 1.5IQR \quad (1)$$

$$IQR = Q_U - Q_L \quad (2)$$

式中： W 为异常值； Q_L 为下四分位数，代表样本数据中有 1/4 的数据小于该段数据； Q_U 为上四分位数，代表样本数据中有 1/4 的数据大于该段数据； IQR 为全部电磁数据样本值的一半。

箱型图通过对实际输入数据的处理来检测异常数据的分布，对数据的类型没有任何要求。箱型图确定离群值的标准是基于四分位数和四分位数间的距离，四分位数具有高鲁棒性，是一种优越的具有客观性的方法。

1.2 数据变换

数据预处理标准化是数据挖掘工作的首要环节，对基于样本间距离的挖掘算法具有关键性的意义。不同的输入数据或评价指标通常具有不同的数值范围和量纲，若不进行预处理工作，会对后续数据的处理分析造成困难。数据预处理主要是对整体所有数据进行比例缩放、量纲归一化，消除数据各项指标之间不同的类型差别，以便于

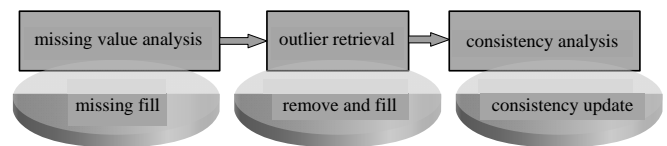


Fig.1 Block diagram of data quality analysis

图 1 数据质量分析框图

后续的数据处理流程。本文采用最小—最大规范化,也称为离差标准化,是一种最常见的标准化算法。其主要原理是将输入数值通过计算其最小值和最大值来映射到[0,1]区间内。转换公式如下:

$$x' = \frac{x - \min}{\max - \min} \quad (3)$$

式中: \max 为频谱样本数据的最大值; \min 为其最小值; $\max - \min$ 为极差。最小—最大规范化在一定程度上保留了原始数据原有的数值关系,同时也消除了数据取值范围和量纲对整体数据分析的影响。

2 数据挖掘技术

主要从三方面对电磁环境频谱监测数据进行浅层数据挖掘:第一,将频谱数据值转换为灰度图像的灰度值,然后采用最大类间方差法对图像阈值进行处理,对频谱数据进行二值化处理;第二,对同一服务的信道进行相关性分析和不同服务的信道进行相关性分析,分别运用 Euclidean 距离和 Kendal's tau-b 相关系数来表征相关性关系;第三,对电磁环境数据进行属性构造,分别构造频域维占用度和时域维占用度,分析多域的占用度问题。浅层数据挖掘的主要方案如图 2 所示。

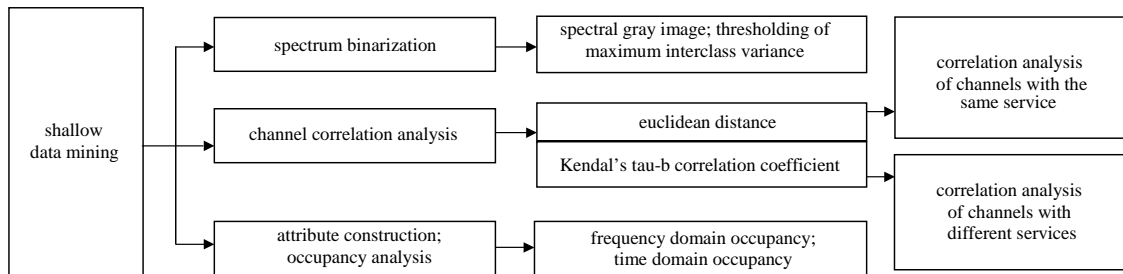


Fig.2 Shallow data mining
图2 浅层数据挖掘框图

2.1 频谱二值化

采用荷兰马斯特里赫特某住宅区地点 day1 的频谱数据。将频谱数据值转换为灰度图像的灰度值,得到频谱数据的灰度图像。本文不选用彩色图像,主要是对于图像处理的许多应用,颜色信息无助于识别边缘或其他特征,反而会增加代码的复杂性。且对于电磁环境大数据,数据本身的繁复要求信息挖掘的手段更加高效,彩色图像的三通道数据处理比灰色图像耗费更多时间。因此选用灰度图像,不仅能快速地对信道工作状态直观感知,同时能加快信息处理的速度。

频谱占用状态只有占用和空闲两种,需进行二值化处理来表征信道占用状态。通常情况下,如果信道此刻的功率谱密度强度大于阈值,则判定此刻信道处于占用状态,用“1”表示;反之,则认为信道此刻处于空闲状态,用“0”表示:

$$CS = \begin{cases} 1, P_c \geq P_0 \\ 0, P_c < P_0 \end{cases} \quad (4)$$

式中: CS 表示信道状态; P_c 表示信号的电平值; P_0 表示预设的门限值。由于人工选择阈值进行二值化处理具有一定的主观性,因此本文通过最大类间方差法计算图像二值化阈值,实现对频谱图像的二值化:

$$g(t) = w_0 \times (u_0 - t)^2 + w_1 \times (u_1 - t)^2 \quad (5)$$

式中: $g(t)$ 为分割阈值为 t 情况下所求的类间方差值,最大类间方差算法使 $g(t)$ 取得全局最大值, $g(t)$ 最大时对应的 t 称为最佳阈值; w_0 和 w_1 、 u_0 和 u_1 分别为灰色像素和白色像素的权重和均值; u 为整个图像的均值, $u = w_0 u_0 + w_1 u_1$ 。

2.2 信道相关性分析

对同一服务内的不同信道进行相关分析,由于信道变量众多,本文采用基于距离的统计学相关分析,选取 Euclidean 距离,计算各信道频谱数据之间平方差之和的平方根。计算公式为:

$$EUCLID = \sqrt{\sum_{i=1}^k (x_i - y_i)^2} \quad (6)$$

式中: k 表示每个样本中有 k 个变量; x_i 为第一个电磁信号样本在第 i 个变量上的频谱值; y_i 为第二个电磁信号样本在第 i 个变量上的频谱值。

本文选用 Kendall's tau-b 等级相关系数对不同的无线服务的信道进行相关分析。Kendal's tau-b 等级相关系数是一种非参数检验指标,通常用于对定序变量的相关程度的估计。一致对数目 U 以及非一致对数目 V 是通过变量秩所计算出用以判断两个变量是否较强的正相关关系。如果具有正相关关系,则一致对数目 U 较大,而非一致对数目 V 较小;当变量间呈现负相关关系时,则一致对数目 U 较小,非一致对数目 V 较大;如果存在相关关系较弱的情况,则一致对数目 U 和非一致对数目 V 会很相似。Kendal's tau-b 相关系数 τ 定义为:

$$\tau = \frac{2(U - V)}{n(n-1)} \quad (7)$$

式中 n 为样本容量。

2.3 电磁环境属性构造

信道频域占用度定义为:监测频段的扫描检测过程中信号阈值高于其最大类间方差阈值的时间长度(T_0)与总测试时间(T_1)的比值:

$$FCO = \frac{T_0}{T_1} \quad (8)$$

时间维占用度定义为:频段占用度随时间维的变化,是在扫描监测频段过程中高于阈值频谱的繁忙信道的频段(N_0)占总监测频段(N_1)的比值:

$$TCO = \frac{N_0}{N_1} \quad (9)$$

2.4 多维混合高斯模型进行前景电磁信号提取

引用图像处理领域的混合高斯背景模型理论,将采集到的大规模电磁环境频谱数据进行预处理后转换为一系列灰度图像,然后通过软件将图像合成一系列频谱视频数据。通过多维混合高斯模型构建背景电磁环境噪声,从而实现前景电磁信号的提取,背景建模流程主体具有三个步骤:

第一步:高斯分布模型匹配。构建混合高斯背景模型,模型中一般包含 3~5 个高斯分布,用来近似估计所有像素点的值,计算出频谱视频帧图像中该像素值的出现概率,概率密度函数如式(10)所示。

$$p(X_t) = \sum_{i=1}^K w_{i,t} p_i(x) = \sum_{i=1}^K w_{i,t} \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (10)$$

式中: X_t 表示 t 时刻的像素值; $w_{i,t}$ 为权重; $\mu_{i,t}$ 为均值向量; $\Sigma_{i,t}$ 为协方差矩阵。处于 t 时刻的第 i 个高斯分布的密度函数如式(11)所示。

$$\eta(X_t, \mu_{i,t}, \Sigma_{i,t}) = \frac{1}{\sqrt{(2\pi)^n |\Sigma_{i,t}|}} e^{-\frac{1}{2}(X_t - \mu_{i,t})^T \Sigma_{i,t}^{-1} (X_t - \mu_{i,t})} \quad (11)$$

在匹配高斯分布模型的同时,需将其中每个像素与对应的 K 个高斯分布作对比:

$$|X_t - \mu_{i,t}| \leq D \times \delta_i \quad (12)$$

若该像素点能够满足该判别式,则像素点与高斯分布匹配;若不满足,则该像素点与高斯分布不匹配。

第二步:混合高斯模型的更新。在进行权值更新时:

$$w_{i,t} = (1 - \alpha) w_{i,t-1} + \alpha M_{i,t} \quad (13)$$

若像素点与高斯分布匹配,则 $M_{i,t}=1$;否则, $M_{i,t}=0$,更新所匹配的高斯分布方差和均值。当该像素点与所有 K 个高斯分布均不匹配,则通过对高斯分布进行重建,同时设置对应位置的像素值为均值。将权重最小的高斯分布替换,对各高斯分布数据的权重值进行归一化处理。

第三步:生成背景模型。判定像素点是否为目标点并不是将所有的高斯模型与该点进行匹配,由于权重越大,越有可能用来描述背景,在权重排名中取前 B 个高斯模型作为背景模型, B 取值为:

$$B = \operatorname{argmin}_b \left(\sum_{i=1}^b w_{i,t} > T \right) \quad (14)$$

式中 T 为背景所占比例。若 B 个高斯分布中任意一个像素与当前像素匹配，则判定该频谱像素为电磁环境背景噪声点；否则，判定该频谱像素为前景电磁环境信号点。多维混合高斯模型对前景电磁信号的提取整体流程如图 3 所示。

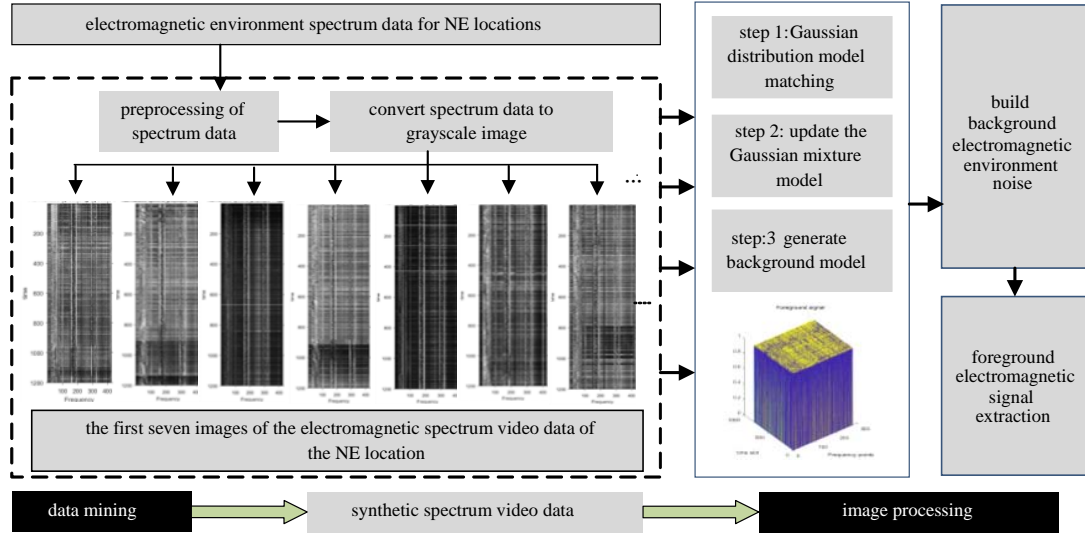


Fig.3 Multidimensional Gaussian mixture model for electromagnetic signal extraction
图 3 多维混合高斯模型提取电磁信号

3 仿真分析

3.1 数据集介绍

本文使用的数据来自德国亚琛大学在 2006~2007 年期间收集的数据集。该数据集是从 3 个不同的位置进行数据监测与收集，收集频率范围大约为 20 MHz~6 GHz，并在该频率范围内分为 4 个子带。表 1 给出了收集位置信息，表 2 给出了一些重要的技术参数。

表 1 收集地点信息

Table1 Collect location information

location number	short name	short description
location 1	IN	office building, Aachen, Germany
location 2	NE	residential area, Maastricht, Netherlands
location 3	AB	third floor balcony, Aachen, Germany

表 2 数据集的技术参数

Table2 Technical parameters of the data set

parameter	value
center frequency/MHz	band 1: $f_c=770$; band 2: $f_c=2\ 250$; band 3: $f_c=3\ 750$; band 4: $f_c=5\ 250$
frequency span/MHz	1 500
resolution bandwidth/kHz	2 00
number of measurement points	8 192
sweep time/s	1
measurement duration	about 7 days per sub-band

该数据集包含的频段有：电视频段(614~698 MHz), LSA 频段(2 300~2 400 MHz), ISM 频段(2 400~2 485 MHz), GSM1800u 频段(1 710~1 785 MHz), GSM1800d 频段(1 820~1 875 MHz)。

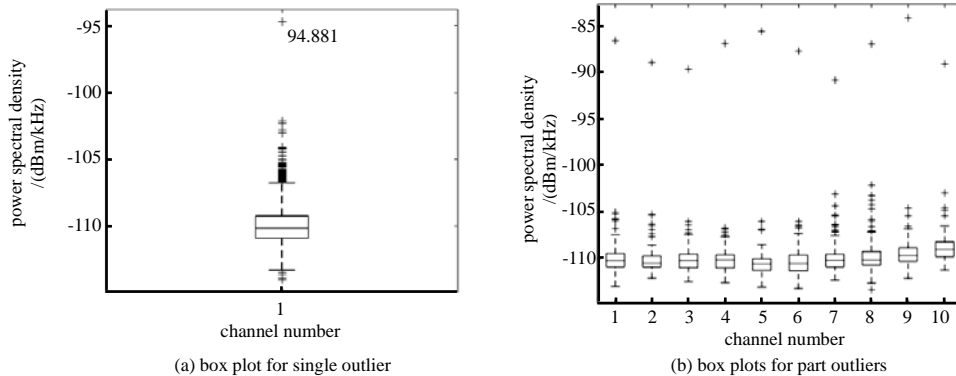


Fig.4 Outlier box chart
图 4 异常数据箱型图

3.2 数据处理结果

对亚琛数据集进行初步数据质量分析,对于电磁频谱数据的缺失值采用数据拟合进行填充。主要工作量集中在进行数据异常值检索,采用箱线图对电磁环境大数据进行异常值的直观检索。以 NE 地点 day1 某频段数据进行箱线图分析为例,得到箱线图如图 4(a)所示。从图 4 中可以看出,原始频谱数据中 NE 地点 day1 的 94.881 为频谱数据的异常值之一,其他信道频点异常值如图 4(b)所示。从图 4 中可以看出,基于四分位数的异常检测算法是有效的,从图中即可直观看出该段频谱数据是否存在异常,以及异常存在的具体位置。基于数据统计特性对数据挖掘分析,直接简单地确定了数据异常情况,计算复杂度低,具有物理可解释性,易于工程实现。

3.3 数据挖掘结果

3.3.1 频谱二值化结果

对 NE 地点 TV 频段频谱数据处理后,得到的灰度图像如图 5(a)所示。对该灰度图像进行最大类间方差阈值处理后如图 5(b)所示,得到最大类间方差标准化灰度阈值为 0.466 7。将上述频谱数据转化为 0 和 1 的二进制序列。本文中信道是带宽为 200 kHz 的射频间隔,信道被顺序索引,由于 200 kHz 是测量设备的分辨力,因此信道是区分能量的最小单位。

图 5 中,黑色像素点表示繁忙的频道,白色像素点表示空闲的信道。可以看出,TV 段存在高频率繁忙使用频段,也存在一些长时间空闲频段。在频谱灰度图中,每个信道的黑色像素和白色像素,分别表示该信道在大约 80 s 的时间分辨力内为信道繁忙状态或信道空闲状态。频谱灰度图相比彩色图像,能够在电磁大数据条件下,提高数据处理的实效性和实时性,减少彩色图像理解和分辨的时间,占用存储空间更小,数据转换速度快。

3.3.2 信道相关性结果

表 3 为 NE 地点的部分频谱数据的信道非相似性矩阵计算结果,数值越大,表明相似强度越弱。从表中可以看出,各个信道的非相似性取值最大为 11.475,且大部分非相似性仅为个位数,表明信道的频谱数据存在很强的相关性,体现了同一服务内信道的频谱使用在统计上是相似的,即具有信道相关性。

表 3 信道非相似性矩阵

Table3 Matrix for channel dissimilarity

	C1	C2	C3	C4	C5	C6	C7	...	C 426
C1	0	2.101	4.761	8.758	8.792	5.822	6.005		10.769
C2	2.101	0	5.383	9.423	9.487	6.422	6.638		11.475
C3	4.761	5.383	0	5.493	5.524	3.638	3.611		7.352
C4	8.758	9.423	5.493	0	3.638	3.611	4.038	...	4.839
C5	8.792	9.478	5.524	3.66	0	4.955	4.475		4.547
C6	5.822	6.442	3.638	4.958	4.998	0	3.66		6.781
C7	6.005	6.638	3.611	4.955	4.94	3.66	0		6.835
C426	10.769	11.475	7.352	4.839	4.547	6.781	6.835		0

表 4 为不同服务下信道相关性分析的结果,信道间的相关性用 Kendal's tau-b 等级相关系数来表征。数值越大,表明相似强度越强,表中标有“*”和“**”的为相关性非常显著的信道。从表中可以看出,GSM1800d 和 GSM1800u 两个频段下的信道的相关性最大,表明信道的频谱数据存在很强的相关性,体现了不同信道的频谱使用在统计上是相似的。

表 4 地点 NE 不同频谱相关系数

Table4 Spectral correlation coefficients of location NE

	TV	GSM1800d	GSM1800u	ISM	LSA
TV	1.000	0.080**	0.085**	0.019	0.046
GSM1800d	0.080**	1.000	0.088**	0.048	0.073**
GSM1800u	0.085**	0.088**	1.000	0.026	0.045
ISM	0.019	0.048	0.026	1.000	0.046
LSA	0.046	0.073*	0.045	0.046	1.000

3.3.3 属性分析结果

对属性构造得到的频谱占用度和时间占用度进行可视化分析。从图 6 中可以看出,信道的时域维占用度存在一定的周期效果;在频域维中,有些频段始终处于高使用状态,有些则处于空闲状态。这种多维的直观分析,有

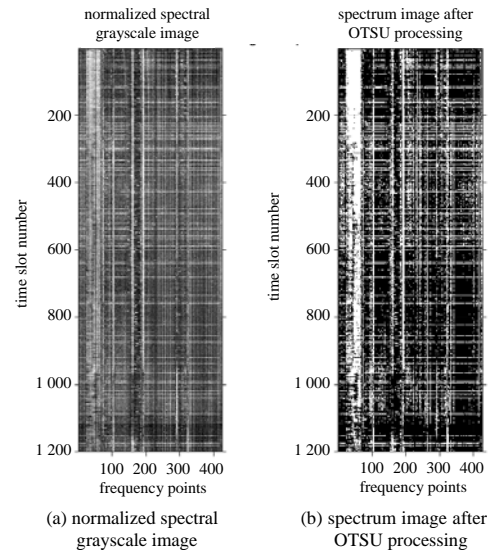


Fig.5 Channel busy and idle

图 5 信道繁忙与空闲情况

利于频谱资源的正确分配和高效利用。

在研究的所有信道和位置中,经验结果显示,信道空置持续时间(CVD)分布均具有指数拖尾,但不完全是指数分布。这一经验结果具有非常高的统计意义。在某种程度上,这一证据支持了一些广泛使用的渠道模型,如0-1 Gilbert-Eliot模型,在这种模型下,这种空缺持续时间呈指数分布。另一方面,对数据的统计分析表明,这些空缺持续时间并非像通常所假设的那样随时间独立分布。这一发现表明,频谱占用度具有可预测性,且更好、更复杂的模型可以促进这种可预测性的利用。

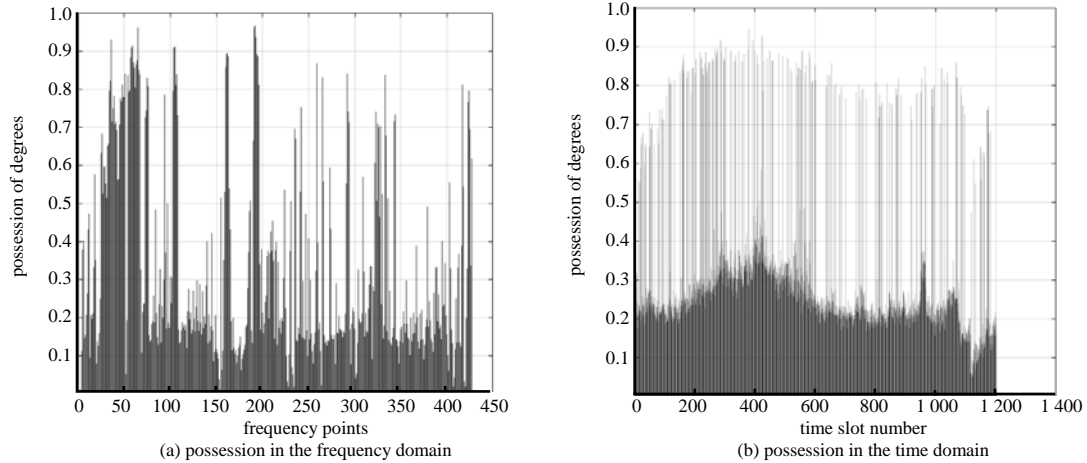


Fig. 6 The duty cycle distribution of each frequency point and time slot in TV segment
图 6 TV 段各个频点和时隙的占空比分布情况

3.3.4 多维混合高斯模型分析结果

图 7 为 NE 地点电磁频谱视频数据的前 7 帧图像,图中颜色较深的频点为频谱占用度情况,颜色较浅的频点为频道空闲状态。

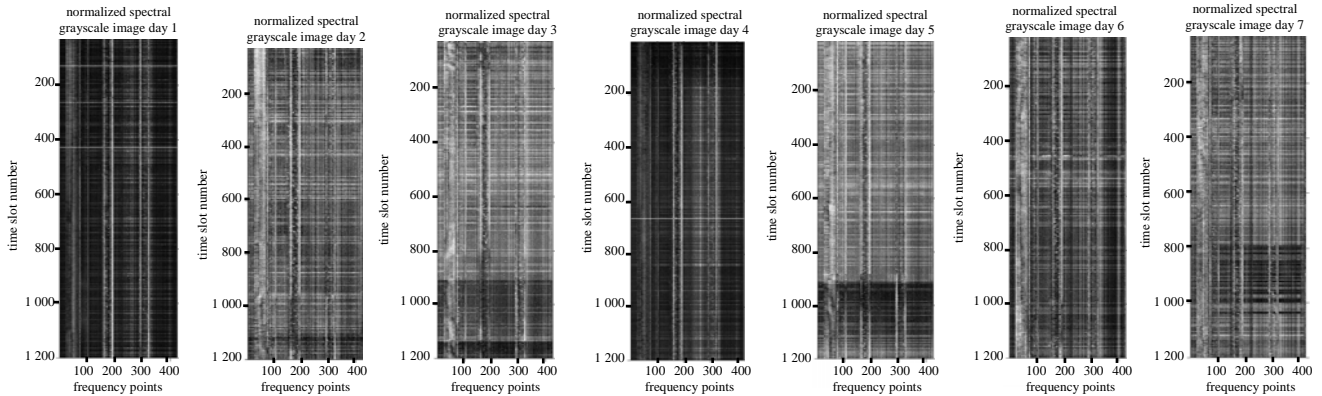


Fig.7 The first seven frames of electromagnetic spectrum video data
图 7 电磁频谱视频数据的前 7 帧图像

图 8 为对合成的电磁环境频谱数据视频进行多维高斯混合模型噪声背景建模后,提取出的前景电磁信号的灰度图像和三维图像。从图中可以看出,所提取出的电磁信号主要集中在频谱繁忙段,对灰度较浅的频段即较为空闲频段进行了滤除,电磁信号的去噪可视化效果得到了加强。

4 结论

本文通过对亚探电磁环境的大数据进行数据集的质量分析、数据整理、数据变换后,进行了数据的一系列浅层分析:包括对处于同一服务的不同信道进行

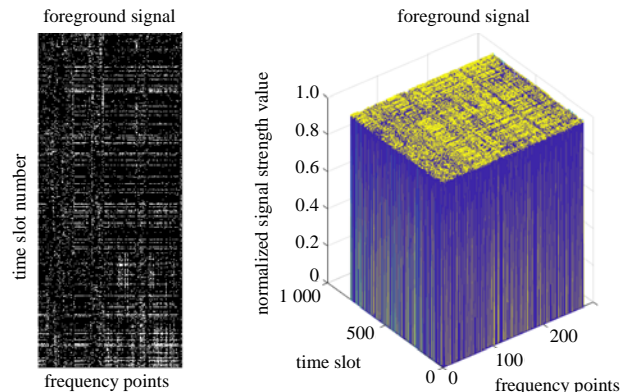


Fig.8 Extraction results of foreground electromagnetic signal
图 8 前景电磁信号的提取结果图

频谱相关性分析,结果表明不同信道间存在频谱相关性;对处于不同服务的不同信道进行频谱相关性分析,结果表明尽管所处的服务信道不同,但仍存在频谱间相关性;然后从频域维度和时域维度分别对频谱占用度和时域占用度进行计算和分析,通过条形图对占用度情况进行可视化,结果表明占用度存在一定的潮汐效应,与人类的活动和频谱使用密切相关;运用图像处理领域的算法对电磁频谱数据进行多角度多维度的数据挖掘,通过多维度的混合高斯背景建模,实现了对电磁信号背景噪声的建模和去除,提取出前景电磁信号。尽管电磁辐射是不可见的,而且频谱资源非常抽象,但电磁环境可以从多角度、多层次、多粒度、多维、多尺度来进行大数据挖掘和分析。这不仅可以为有效利用频谱资源提供帮助,而且可以快速关联异常事件,维护频谱秩序,保护电磁生态环境。

参考文献:

- [1] 罗勇,石丹,王冠,等. 5G 基站电磁辐射[J]. 太赫兹科学与电子信息学报, 2021. doi:10.11805/TKYDA2020700. (LUO Yong,SHI Dan,WANG Guan,et al. Electromagnetic radiation of 5G base station[J]. Journal of Terahertz Science and Electronic Information Technology, 2021.) doi:10.11805/TKYDA2020700.
- [2] 王玉峰,邹积岩,廖敏夫. 基于数据挖掘技术预测与分析变电站中电磁环境[J]. 高压电器, 2009,45(5):77-79. (WANG Yufeng,ZOU Jiyan,LIAO Minfu. Prediction and analysis of electromagnetic environments in substation with data mining technology[J]. High Voltage Apparatus, 2009,45(5):77-79.)
- [3] ELTOM H,KANDEEPAN S,LIANG Y,et al. Cooperative soft fusion for HMM-based spectrum occupancy prediction[J]. IEEE Communications Letters, 2018,22(10):2144-2147.
- [4] HÖYHTYÄ M,POLLIN S,MÄMMELÄ A. Classification-based predictive channel selection for cognitive radios[C]// 2010 IEEE International Conference on Communications. Cape Town,South Africa:IEEE, 2010:1-6.
- [5] TALEBI B,DEHKORDI M N. Sensitive association rules hiding using electromagnetic field optimization algorithm[J]. Expert Systems with Applications, 2018(114):155-172.
- [6] 李斌,李高云,陈亮. 电磁大数据挖掘应用研究[J]. 电子信息对抗技术, 2019,34(4):50-54,69. (LI Bin,LI Gaoyun,CHEN Liang. Research of electromagnetic space big data mining application[J]. Electronic Information Warfare Technology, 2019,34(4):50-54,69.)
- [7] 艾小锋,赵锋,王满喜. 典型雷达电磁环境效应特征数据模拟系统[J]. 太赫兹科学与电子信息学报, 2019,17(4):636-640. (AI Xiaofeng,ZHAO Feng,WANG Manxi. Data simulation system of electromagnetic environment effects on typical radar[J]. Journal of Terahertz Science and Electronic Information Technology, 2019,17(4):636-640.)
- [8] HOU Changbo,ZHANG Xiao,CHEN Xiang. Electromagnetic signal feature fusion and recognition based on multi-modal deep learning [J]. International Journal of Performability Engineering, 2020,16(6):941.
- [9] PRZYSTUPA Krzysztof,VASILKIVSKIJ Igor,ISHCHENKO Vitalii,et al. Electromagnetic pollution:case study of energy transmission lines and radio transmission equipment[J]. Przegląd Elektrotechniczny, 2020(2):52-55.
- [10] HABER E,SCHWARZBACH C. Parallel inversion of large-scale airborne time-domain electromagnetic data with multiple OcTree meshes[J]. Inverse Problems, 2014,30(5):055011.
- [11] ZHAO Ying,LUO Xiaobo,LIN Xiaoru,et al. Visual analytics for electromagnetic situation awareness in radio monitoring and management[J]. IEEE Transactions on Visualization and Computer Graphics, 2019,26(1):590-600.
- [12] CORREA N M,ADALI T,LI Y O,et al. Canonical correlation analysis for data fusion and group inferences[J]. IEEE Signal Processing Magazine, 2010,27(4):39-50.
- [13] DING Guoru,JIAO Yutao,WANG Jinlong,et al. Spectrum inference in cognitive radio networks:algorithms and applications[J]. IEEE Communications Surveys & Tutorials, 2017,20(1):150-182.
- [14] RYABOV A,BAUM P. Electron microscopy of electromagnetic waveforms[J]. Science, 2016,353(6297):374-377.

作者简介:

李爽(1997-),女,黑龙江省绥化市人,在读硕士研究生,主要研究方向为电磁数据分析挖掘、知识图谱。
email:shuang97@hrbeu.edu.cn.

刘海鹏(1981-),男,济南市人,主要研究方向为通信工程。

郭兰图(1982-),男,山东省青岛市人,硕士,主要研究方向为频谱管理、复杂电磁环境建模等。