摘要

随着4G无线技术普及，无线基站流量持续增加。作为承载无线网络的基础设施，分析基站流量的静态和动态特征、挖掘基站流量演变模式对无线基站的运营方案制定、参数合理配置至关重要。目前已有的测量和模型工作，聚焦在短时间粒度（如分钟、小时、天）的流量变化规律，至今尚缺少一个长时间尺度（如一年）的城市基站流量变化的准确测量和模式挖掘结果。

为此，本文基于一个中国大型无线网络运营商在一个大型城市的基站网络的流量测量数据，对超过7千个基站、以月度为单位、持续一年的基站流量的静态和动态特征进行了观察与测量，并对基站流量的时间变化模式进行了聚类、分析和预测。本文贡献如下：

(1) 提出了一种新的基站流量演变模式的聚类方法。该方法基于基站月度总流量值在一年内的排序序列进行聚类。在我们数据集上的实验结果表明：该方法对短时间序列（不存在周期性）的波形涨跌特点有很好的描述，能得到比传统方法更容易理解的聚类结果。

(2) 基于该聚类方法，我们对运营商的7千多个基站的流量演变模式进行了大规模聚类分析，获得了6种典型基站流量演变模式，最主要的一种流量模式涵盖了38.6%的基站，特点为流量总体呈上升趋势，11月份达到高峰，次年2月降到低谷。其它模式包括：“春节返乡”、 “双11”电商购物模式等。结合该城市的特点，我们对各模式的形成原因做出了解释，这些发现为运营商掌握其基站的流量演变的规律提供了有益的指导。

(3) 提出了一种基于基站地理位置和地址语义信息的基站模式预测方法，对一个新建基站的流量模式进行预测。因为新建基站的初始信息很少，预测困难，因此我们创新性地将基站语义标签信息引入到基站流量模式预测中。实验结果表明：通过加入基站词向量表征，预测模型的F1-score提升了5%，其中两种模式的预测准确度较高。

本文对大规模实际基站网络流量的分析结果对于网络运营商有重要的实际价值，本文提出的长期流量变化模式模型、聚类算法、和预测算法，具有通用性，能够被应用于类似的各种应用场景，具有重要的理论价值和实际意义。

图20幅，表8个，参考文献43篇。

**关键词：**网络测量；模式挖掘；模式预测；机器学习