**[1]Spectral clustering for sensing urban land use using Twitter activity**

**内容：**本文采用一种非监督学习方法根据Twitter数据对地理区域进行聚类，以自动识别城市土地利用类型。

**使用地理定位信息进行土地利用建模分析的方法分类（以数据来源为依据）：**

1. 基于定位的社交网络 (Location-based social networks (LBSN) traces from Foursquare of Twitter)
2. 移动电话的通话细节记录 (CDRs)
3. GPS记录

**先前研究所限之处：**

1. 对结果缺乏正规的检验
2. 研究只针对某一特定城市，对研究方法有所限制
3. 某些数据具有较强的隐私限制性，如电话记录
4. 部分监督分类方法要求提供城市相关的基本信息

**利用Twitter识别土地利用类型：**

分割+识别

1. 使用地理定位数据进行分割

采用方法——[Self-Organizing Maps (SOM)](https://en.wikipedia.org/wiki/Self-organizing_map)，将研究区域分割为与不同类型tweets相关的分区。

由于可以选择多种初始格网大小，因此分割结果会包含多个不同大小的映射，通过最小化DB系数来选择最优的分割方案来构建拓扑关系。最后获得一个映射，其中每个神经元都指向一个具有高密度tweets的区域。集中大量tweets的区域其附近会有大量神经元。最终，采用Voronoi tessellation方法对研究区域进行划分。

As a result of the process, we obtain a map where each neuron represents a pointer to a region with a high density of tweets. Additionally, areas with larger concentrations of tweets will have larger numbers of neurons geographically located nearby. Finally, Voronoi tessellation (Aurenhammer, 1991) is applied over the location of the neurons in order to compute the land segments that each neuron represents. These land segments are used as the elements for the characterization of land use.

1. 非监督学习法检测土地利用类型

计算每日tweets流量向量，包含weekday和weekend两个，以20分钟为间隔，计算得到两个包含72个元素的向量。