**Automatic Discovery of Tactics in Spatio-Temporal Soccer Match Data**

**目标 :**

输入 ：用“事件”来描述的比赛的集合

输出 ：相应的刻画进攻策略的时空模式

**方法论：**

**第一步：将一场比赛的事件流分成多个阶段**

**第二步：将阶段按时空成分聚类**

**第三步：聚类排序**

**第四步：挖掘模式**

**第五步：模式排序**

**具体步骤描述：**

**第一步：将一场比赛的事件流分成多个阶段**

比赛用事件序列描述，一个事件可以用**ei**=(time,location,players involved,event type)来描述。

将比赛分成多个阶段P={ei}，每个阶段都由一系列事件组成，划分依据：

1.事件间相隔超过10秒

2.球权转换

只考虑阶段的事件数不少于3的阶段。

**第二步：将阶段按时空成分聚类**

用层次聚类(AHC)的方法进行阶段聚类

输入：{P}:一场比赛的所有阶段，k:聚类停止时的类总数

1. 初始时令每个阶段为一个聚类
2. while(所剩类总数>k)

合并距离最近的两个类；

输出：k个聚类

其中，两个类的距离计算采用DTW算法

分别包含m,n个事件的P1和P2两个阶段的DTW结果为D[m][n],采用动态规划计算，转移方程为：





**第三步：聚类排序**

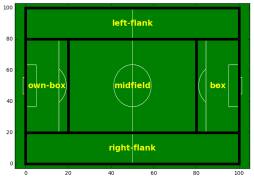
按聚类中包含的总射门数进行排序

**第四步：挖掘模式**

辨识出每个聚类中常见的序列模式（FS）

通过事件发生的位置和类型为事件建立itemset(无序集)用于数据挖掘

位置的处理：按下图对场上位置划分



事件类型处理：

细分不同类型的传球和射门

为使结果更容易理解，采取如下的类自然语言记号：

1. AT [X]: 在X位置发生事件A
2. FROM [Y] TO [Z]: 位置Y到位置Z的传球

**第五步：模式排序**

根据模式相关度进行排序，相关度分数计算函数为：

