



## 关于频繁模式

徐华

清华大学 计算机系 智能技术与系统国家重点实验室

xuhua@tsinghua.edu.cn

1

### 频繁模式例子1



#### 超市消费者购物小票 说起

“购买啤酒的男士经常会购买小孩的纸尿裤”



2



### 频繁模式例子2



#### ● 谷歌预测 2009年 部分地区的流感

“2009年，Google通过分析5000万条美国人最频繁检索的词汇，将之和美国疾病中心在2003年到2008年间季节性流感传播时期的数据进行比较，并建立一个特定的数学模型。最终google成功预测了2009冬季流感的传播甚至可以具体到特定的地区和州。”



3

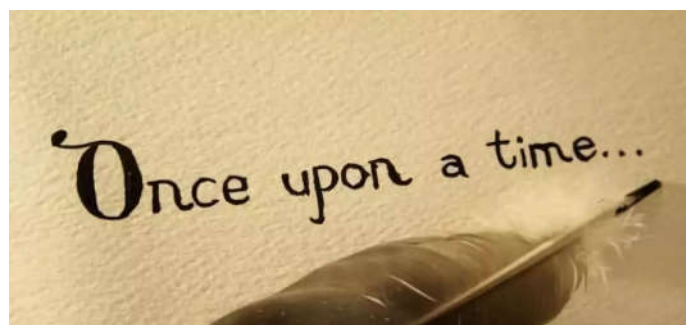


### 频繁模式例子3



#### ● 超市预知高中生顾客怀孕

明尼苏达州一家塔吉特（Target）门店被客户投诉，一位中年男子指控塔吉特将婴儿产品优惠券寄给他的女儿——一个高中生。但没多久他却来电道歉，因为女儿经他逼问后坦承自己真的怀孕了。塔吉特百货就是靠着分析用户所有的购物数据，然后通过相关关系分析得出事情的真实状况。



4



## 频繁模式例子4



### 电话通话规律的匹配

根据通话规律分析贩毒团伙成员。（视频）

5



## 频繁模式例子5



### 电表与空置房的匹配

**国家电网一查电表全国空置房就暴露出来，竟然达到了6540万套？**

消息称，国家电网最近对660个主要城市进行调查，其中显示总计有6540万套房子连续6个月的电表为0。足够两亿人居住。



6



## 频繁模式例子6



### 疫情的封控模式分析

有人看到大巴车、救护车或者民警到小区，就非常担心，会胡思乱想，担心小区里是不是有疫情了。下面简单整理汇总一下，大差不离，觉得有用的，欢迎点赞收藏：

- 1、大白+大巴+120+民警=阳了；
- 2、大白+大巴+综合执法=密接；
- 3、只有120=居民不舒服了；
- 4、警戒线围起来=有密接，捅完阴了就解；
- 5、直接上铁板=确诊；
- 6、铁板+警戒线=某楼重灾区，其他楼管控；
- 7、来车直接拉走=芭比Q了；
- 8、志愿者派出去的勇士有女的=大筛，半天活完事；
- 9、志愿者派出去的都是男的=阳了，14封控+14隔离，没商量；

7



## 关于频繁模式



### 频繁模式 = 规律

频繁项集挖掘是数据挖掘研究课题中一个很重要的研究内容分支，它可以告诉我们在数据集中经常**一起出现**的变量，为可能的决策提供一些支持。频繁项集挖掘是关联规则、相关性分析、因果关系、序列项集、局部周期性、情节片段等许多重要数据挖掘任务的基础。

8





**Thanks !**

