

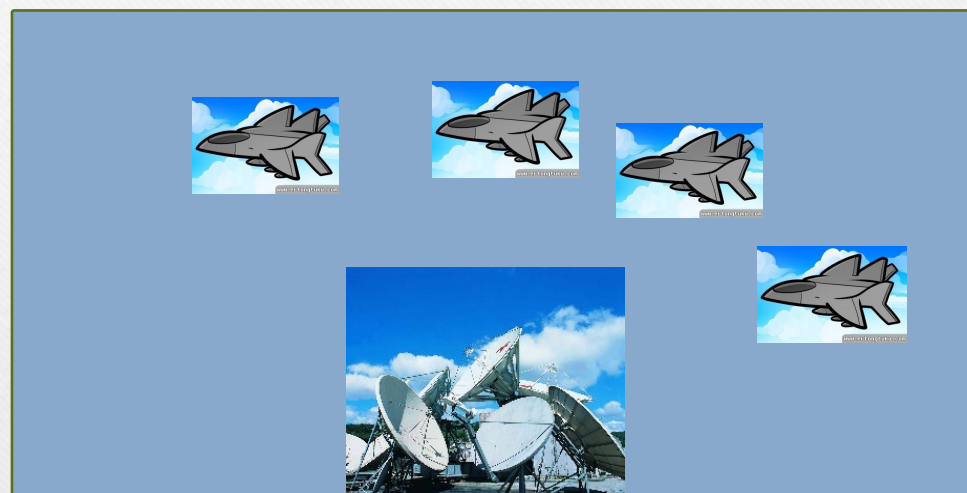
# 数据互联

金学波

北京工商大学



# 多传感器数据融合3大问题

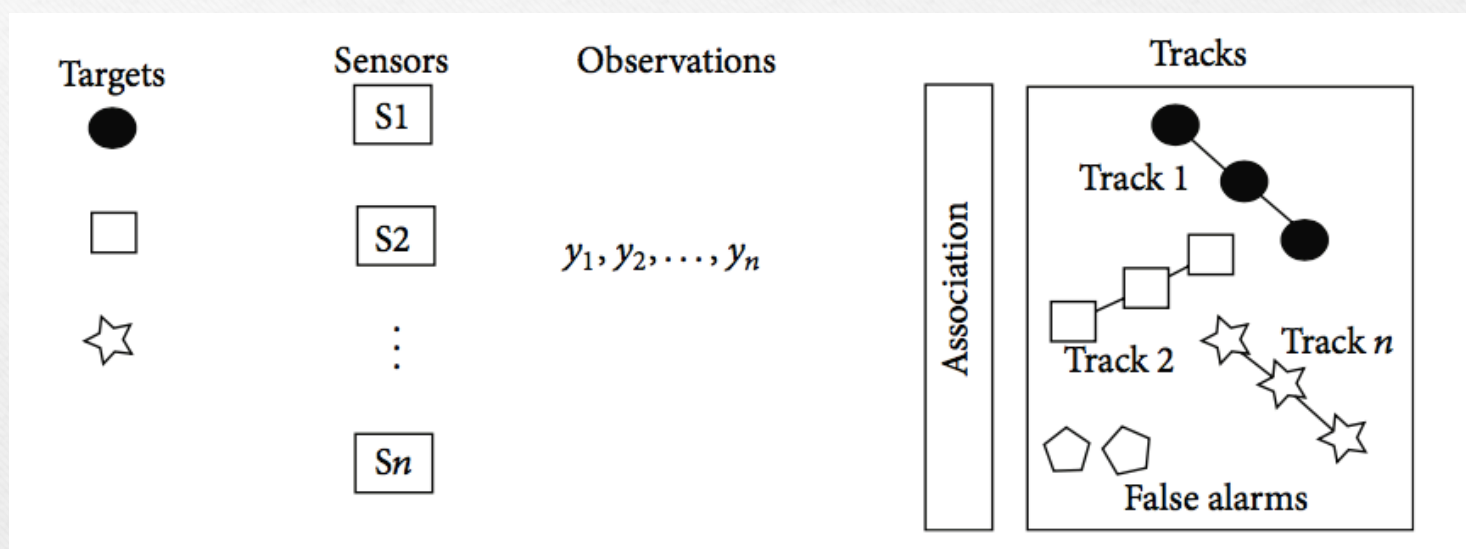


数据互联：是1架飞机还是多架？

状态估计：目标在哪？速度如何？往哪飞呐？

决策融合：谁家的飞机？干嘛来了？！！

# 数据互联问题的描述





# 数据互联问题的描述

- 假设多个目标、多个传感器，数据互联的问题是：  
在下列情况下，确定与每个目标相对应的测量集。
  - (i) 以离散的时间间隔在融合节点中接收每个传感器的观察结果；
  - (ii) 传感器可能不会以特定间隔提供观测结果；
  - (iii) 观测结果的来源包括“探测到的目标”及“噪声”；
  - (iv) 对于任何特定目标和每个时间间隔，我们事先不知道该目标产生观测结果的具体数据。

# 数据关联问题

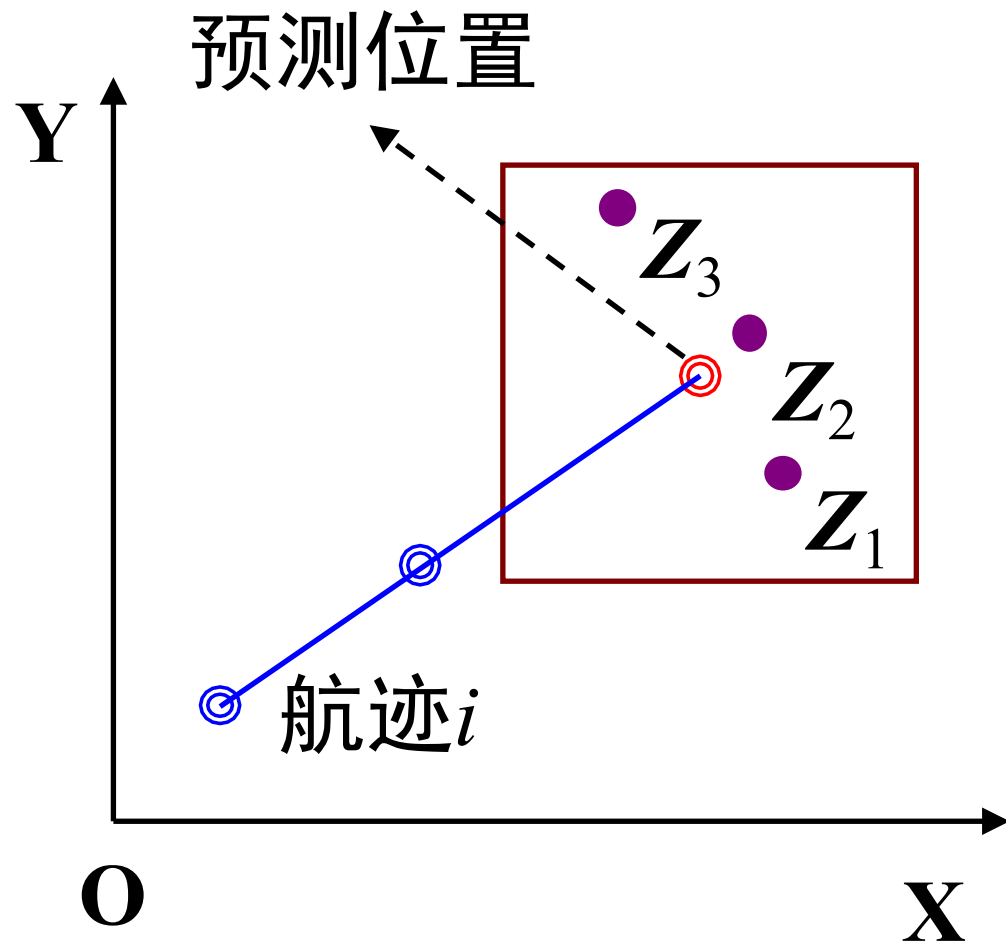
- 数据关联的目标：建立由同一目标随时间生成的一组测量数据。
- 通常在检测到的目标状态估计之前执行数据关联。
  - 如果数据关联错误，估计或分类性能将会受到影响。
- 数据关联过程也可以出现在所有融合级别中，但粒度根据每个级别的目标而变化。



# 方法

- 
- *Nearest Neighbors and K-Means.*
  - *Probabilistic Data Association.*
  - *Joint Probabilistic Data Association.*
  - *Multiple Hypothesis Test.*
  - *Distributed Joint Probabilistic Data Association.*
  - *Distributed Multiple Hypothesis Test.*
- 最邻近和K-Means
  - 概率数据互联
  - 联合概率数据互联
  - 多重假设检验
  - 分布式联合概率数据互联
  - 分布式多假设检验

# Nearest Neighbors. 最邻近



残差:

$$e_{ij}(k) = Z_j(k) - \hat{Z}_i(k | k-1)$$

统计距离:

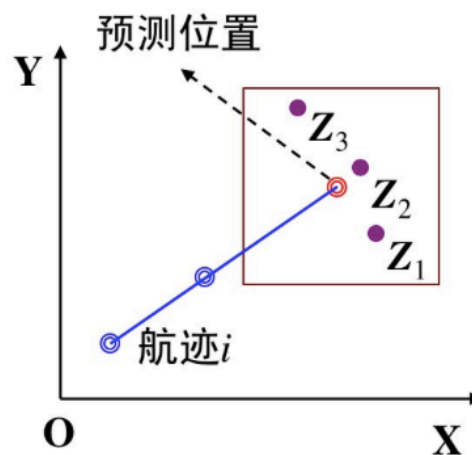
$$d_{ij}^2 = e_{ij}(k) S_{ij}^{-1}(k) e_{ij}^T(k)$$

似然函数:

$$g_{ij} = \frac{e^{-d_{ij}^2/2}}{(2\pi)^{\frac{M}{2}} \sqrt{|S_{ij}|}}$$

# Nearest Neighbors. 最邻近

- 最邻近：
  - Euclidean 最小
- 缺点：
  - 无法处理多目标的情况



残差:

$$e_{ij}(k) = Z_j(k) - \hat{Z}_i(k|k-1)$$

统计距离:

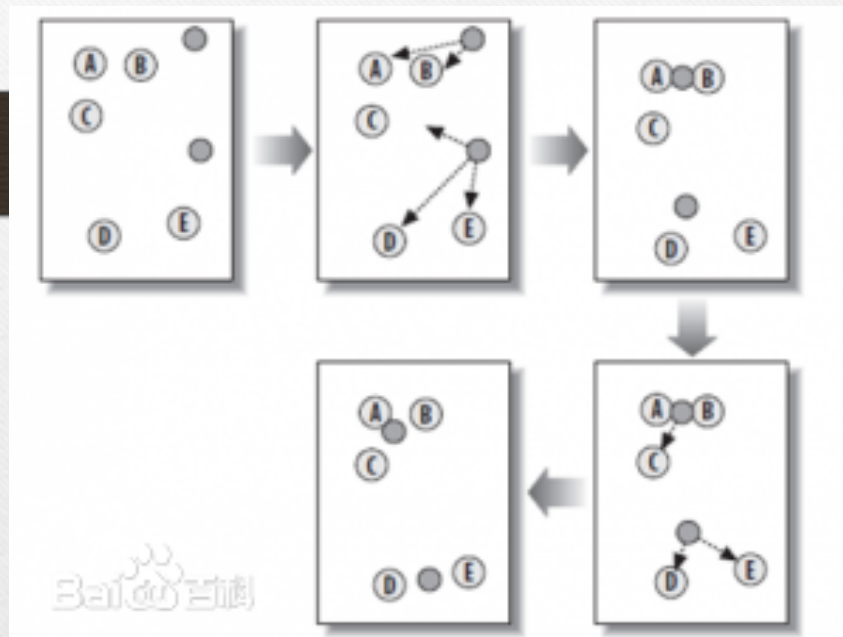
$$d_{ij}^2 = e_{ij}(k) S_{ij}^{-1}(k) e_{ij}^T(k)$$

似然函数:

$$g_{ij} = \frac{e^{-d_{ij}^2/2}}{(2\pi)^{\frac{M}{2}} \sqrt{|S_{ij}|}}$$



# *K-Means.*



- K-Means将数据集划分为 $K$ 个不同的簇。
- 步骤为：
  - (1) 获取输入数据和所需簇数 ( $K$ ) ;
  - (2) 随机分配每个簇的质心;
  - (3) 将每个数据点与每个簇的质心相匹配;
  - (4) 将聚类中心移动到聚类的质心;
  - (5) 如果算法不收敛, 则返回步骤 (3) 。

# K-Means

---

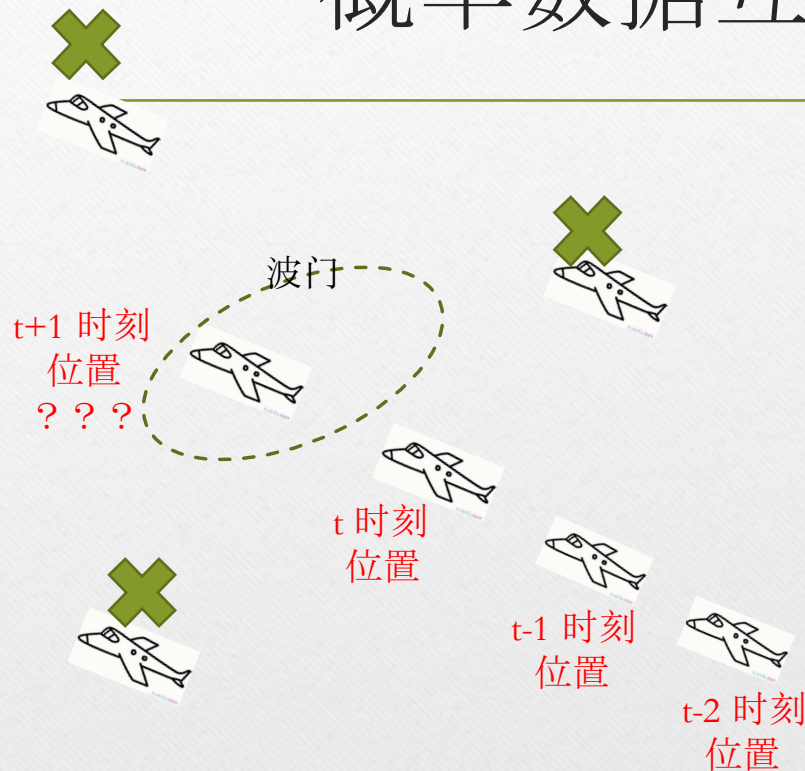
- 缺点

- (1) 在 Kmeans 算法中  $k$  需要事先确定，这个  $k$  值的选定有时候是比较难确定。
- (2) 初始 $k$ 个聚类中心的选择对聚类结果有较大的影响，一旦初始值选择的不好，可能无法得到有效的聚类结果。
- (3) 该算法需要不断地进行样本分类调整，不断地计算调整后的新的聚类中心，因此当数据量非常大时，算法的时间开销是非常大的。
- (4) 对离群点很敏感。



# Probabilistic Data Association.

## 概率数据互联 (PDA)



- 缺点：

(i) 轨道丢失：由于PDA忽略了对其他目标的干扰，它有时会错误地分类最近的轨道。因此，当目标彼此接近或交叉时，它提供的性能很差；

(ii) 只适用于一个目标；

(iii) 轨道管理：由于PDA假设轨道已经建立，因此必须提供算法以进行轨道初始化和轨道删除。

(iv) 需要使用阈值，正确的阈值很难选择。

# 关联概率（波门）的计算

两个集合：

$Z(k) = \{z_i(k)\}_{i=1}^{m_k}$  :  $k$ 时刻的确认量测集

$Z^k = \{Z(j)\}_{j=1}^k$  : 直到 $k$ 时刻的累积确认量测集

两个事件：

$\theta_i(k)$  :  $z_i(k)$ 是源于目标的量测

$\theta_0(k)$  :  $k$ 时刻没有源于目标的量测

关联概率：  $\beta_i(k) = P\{\theta_i(k) | Z^k\}$   $\sum_{i=0}^{m_k} \beta_i(k) = 1$



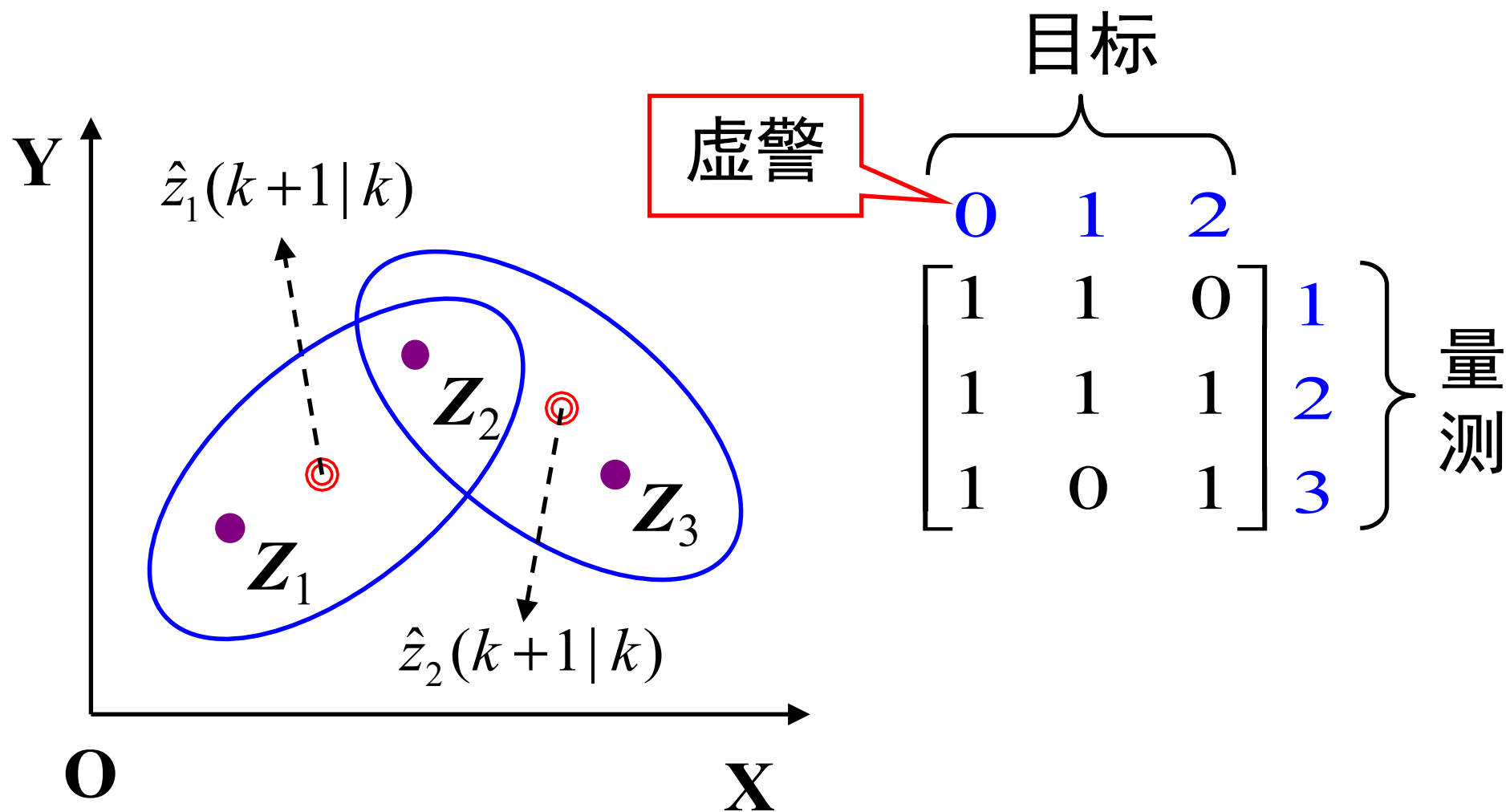
# *Joint Probabilistic Data Association.*

## 联合概率数据互联 (JPDA)

---

- 适用于杂波环境中跟踪**多个目标**
- JPDA与PDA类似，区别在于使用所有观察和所有目标计算关联概率。

# 联合概率数据关联 (JPDA)





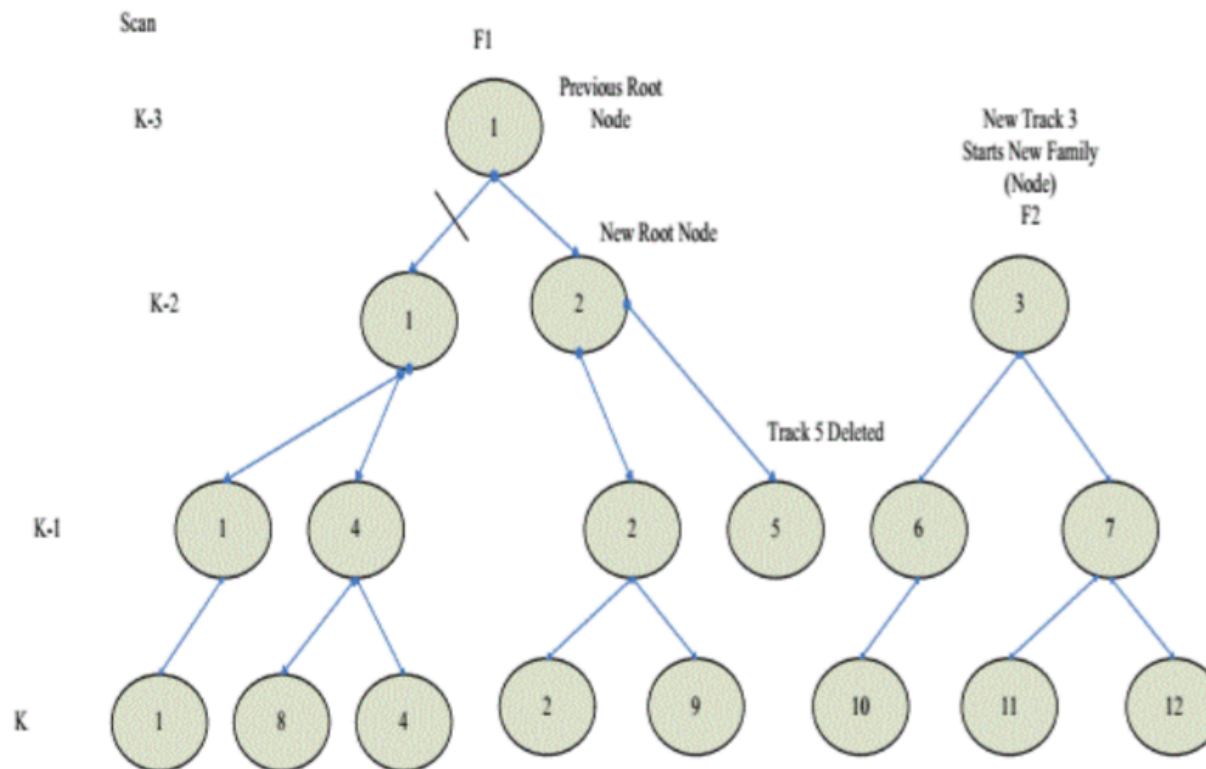
# *Multiple Hypothesis Test.*

## 多重假设检验 (MHT)

- MHT算法把多个假设继续传递，让后续的观测数据来解决这种不确定的问题。
- 举个例子来说，PDA对所有假设的概率进行加权平均，然后再对航迹进行更新。
  - 也就是如果有10个假设，PDA会将这是个假设有效的合并最后留下一个假设的行迹。
  - 而MTH会将这是个假设继续保留着，也就是保留着十条假设的航迹再继续观察，直到确认某些航迹是假的，再把它们删除。

# *Multiple Hypothesis Test.*

## 多重假设检验 (MHT)





## *Multiple Hypothesis Test.*

### 多重假设检验 (MHT)

- 通常使用贝叶斯规则或贝叶斯网络来计算MHT假设。
- 一般来说，研究人员声称MHT的误报率低于JPDA。
- 然而，MHT的主要缺点是当轨道数或误报数增加时的计算成本。
- 使用窗口修剪假设树可以解决这个缺陷。
- 也是针对一个目标的。

*Distributed Joint Probabilistic Data Association.*  
分布式联合概率数据互联 (D-JPDA)

---

- 适用于**多个**传感器跟踪**单个**目标
- 每个传感器都在计算目标的航迹，然后在中心合在一起
- 每个传感器都以JPDA来计算航迹



## *Distributed Multiple Hypothesis Test.* 分布式多假设检验 (D-MHT)

---

- 适用于**多个**传感器跟踪**多个**目标
- **每个**传感器都在计算**每个**目标的航迹，然后在中心合在一起
- 每个传感器都以MHT来计算航迹

# 总结

---

- 数据互联适用于具有干扰的测量数据
- 多目标系统更加复杂
- 单目标首选MHT
- 多目标首选D-MHT
- 其原理涉及估计理论，比较复杂，本节是概要介绍