

```
graph TD; Start(( )) --> A[计算与每个聚类中心的距离]; A --> B{距离最小值超过了聚类时的阈值?}; B -- yes --> C[新模板出现, 暂存网页]; C --> D(( )); B -- no --> E[选择距离最近的聚类对应的模板]; E --> F[将序列S同模板对齐, 根据标注抽取我们感兴趣的部分]; F --> G[输出为XML]; G --> H(( )];
```

计算与每个聚类中心的距离

距离最小值超过了聚类时的阈值？

yes no

新模板出现，
暂存网页

选择距离最近的聚类对应的模板

将序列S同模板对齐，
根据标注抽取我们
感兴趣的部分

输出为XML