

BigData Question:Item-based vs User-based in filtering

姓名：于文明 学号：2011762 专业：信息安全

Q:In practice, it has been observed that item-item often works better than user-user

A:

- 数据稀疏性：用户-用户的矩阵通常是稀疏的，用户可能只评价了几个项目。在这种情况下，用户-用户协同过滤可能不是很有效，因为没有足够的数据来找到相似的用户。另一方面，项目-项目协同过滤仍然有效，因为它依赖于项目之间的相似性，而这些相似性通常比用户-用户相似性更密集。

- 冷启动：对于用户-用户协方法来说，新用户几乎没有关于他们的信息可以与其他用户进行比较。但是 item-item 方法来说，它受到较小的冷启动影响

- 稳定性：在很多情况下，物品比用户更稳定。例如，电影的特征可能会随着时间的推移保持不变，而用户的偏好可能会发生变化。这意味着项目之间的相似性可能比用户之间的相似性更稳定，从而使项目-项目协同过滤更加有效。

- 可扩展性：对于 user-user 方法来说，系统中的用户越多，寻找最近的 K 个邻居的成本就越大。另一方面，item-item filtering 只需要计算 item 之间的相似度，通常计算量较小。

总体而言，项目-项目过滤在实践中通常比用户-用户过滤效果更好，因为它具有可扩展性、处理稀疏性的能力以及更好的稳定性。