考量時間機率之循序樣式探勘方法

循序樣式探勘是一門研究如何從序列資料庫裡找出頻繁循序樣式的資料探勘方法，過去的序列資料探勘方法大致可分成兩大類[2]：Apriori-like methods[14][16][17][18]和Pattern-growth methods [10][11][12][13][23]。無論是哪一類的方法，基本上頻繁樣式都是經由通過使用者設定的門檻值產生，而門檻值的設定往往只有項目出現次數的最小支持度而已。然而，在現實的情況下，決策者需要透過探勘的結果了解時間序列事件間的關聯，甚至期望在探勘過程中能一併計算得到事件發生的可能性，作為決策判斷時的參考。本研究因考量到上述需要，以PrefixSpan 演算法為基礎，加入時間機率的考量，發展出PCTP 演算法。其探勘得到的樣式不旦包括時間機率的訊息，且能依使用者給定的機率作為樣式過濾的準則，以減少PrefixSpan 執行過程中映射資料庫的產生，提升整體的效能。