過去高校用項目集挖掘模式（high utility itemset pattern）已經被廣泛的研究，但是並沒有考慮到有可能會在儲存設備而不是主要的設備上，因此，會有

儘管過去有很多相似的模型，基於 high utility itemset pattern

\textit{高效用模式挖掘}比\textit{頻繁模式挖掘}困難得多，因為缺乏反單調性~\cite{??}。

因此，之前的工作傾向於掃描保存在輔助存儲設備而不是主內存上的整個數據集，而且可能是巨大的數據集。

考慮到存儲設備的性能遠低於主存儲器，減少為高效用模式挖掘產生的 IO 流量至關重要。

為了實現這一目標，我們提出了一種

高效用模式探勘近似隨機演算法（PAHUPMA），它只需要掃描整個數據級的一小部分即可以發現高校用模式。

概率近似高效用模式挖掘算法 (PAHUPMA)}，它只需要掃描整個數據集的一小部分即可發現高效用模式。我們與過去演算法進行了一系列分析和實驗結果，結果令人滿意。

頻繁模式挖掘是一種常被拿來發現不同項目之間關聯關係的方式，但是頻繁挖掘模式並沒有考慮到每一個項目之間有可能會有不同的效用值，與同一筆交易不相同的項目有可能會有出現次數不同。因此提出了高校用挖掘模式，解決了上訴的問題，

頻繁項目挖掘模式是資料探勘中關聯規則常見的一種方法，其中主要目的在於尋找數據庫中頻率最高的項目集，而頻繁項目挖掘並沒有考慮到每一個項目之間有可能會有不同的效用值，與同一筆交易不相同的項目有可能會有出現次數不同。為了克服了上述的問題，提出了高校用挖掘模式。藉由每一個項目都會

過去研究 high

因此，工作傾向於掃描保存在輔助存儲設備而不是主內存上的整個數據集，而且可能是巨大的數據集

我們考慮到整個數據集可能在輔助儲存設備上面而不是主記憶體上面，通常輔助儲存設備會性能會遠低於主記憶體，為了減少輔助儲存設備所造成的性能瓶頸，因此我們提出了一種高效用模式探勘近似隨機演算法框架（PAHUPMA），它只需要掃描整個數據級的一小部分即可以發現高校用模式。我們與其他演算法進行了一系列分析和實驗結果，結果令人滿意。

高效用模式挖掘是資料探勘中常見的一種方法，其主要目的是在高校用項目集中尋找數據庫中最高效用的關聯規則。考慮項目的數量和不同項目的重要性，高效用模式挖掘方法在處理現實世界的大規模資料庫時相對於傳統的模式挖掘方法提供更高效、更準確的結果。而過去的演算法並沒有考慮到整個數據集可能在輔助儲存設備上面而不是主記憶體上面，通常輔助儲存設備會性能會遠低於主記憶體，為了減少輔助儲存設備所造成的性能瓶頸，因此我們提出了一種高效用模式探勘近似隨機演算法框架（PAHUPMA），它只需要掃描整個數據級的一小部分即可以發現高校用模式。我們與其他演算法進行了一系列分析和實驗結果，結果令人滿意。

每一個項目之間有可能會有不同的效用值，與同一筆交易不相同的項目有可能會有出現次數不同。為了解決了上訴的問題，提出了高校用挖掘模式，

高效用模式挖掘是資料探勘中常見的一種方法，其主要目的在於在項目集中尋找數據庫中最高效用的關聯規則。該方法考慮項目的數量和不同項目的重要性，並相對於傳統頻繁模式挖掘方法，在處理現實世界的大規模資料庫時提供更高效、更準確的結果。而過去的演算法並沒有考慮到整個數據集可能在輔助儲存設備上面而不是主記憶體上面，通常輔助儲存設備會性能會遠低於主記憶體，為了減少輔助儲存設備所造成的性能瓶頸，因此我們提出了一種高效用模式探勘近似隨機演算法框架（PAHUPMA），它只需要掃描整個數據級的一小部分即可以發現高校用模式。我們與其他演算法進行了一系列分析和性能評估，所提出的方式更能符合現實環境中的需求。

我們對高效用模式挖掘方法與其他演算法進行了一系列的分析和性能評估，結果顯示我們提出的方法更能夠符合現實環境中的需求。這些分析和評估的基礎是對不同演算法在處理現實世界大規模資料庫時的效能進行全面的比較和評測。我們的方法在考慮項目數量和不同項目的重要性的同時，能夠提供更高效、更準確的結果，從而更好地滿足現實環境中的需求。

相較於過去的演算法，它們未考慮到整個數據集可能存儲於輔助儲存設備，而非主記憶體。這一點非常重要，因為輔助儲存設備（例如硬碟或快閃記憶體）的存取速度較慢，且相對主記憶體而言，I/O操作的成本更高。在我們提出的高效用模式挖掘方法中，我們特別考慮到了這個問題，並採用了相應的優化策略，以最小化對輔助儲存設備的存取次數，從而提升整體效率和性能。這使得我們的方法更適合處理實際環境中存儲在輔助儲存設備上的大規模數據集，並能夠有效處理存取速度較慢的情況。

在一個淺在的數據集中，發現重要的模式往往非常有用

在非常龐大的數據集中，發現重要模式通常很有用。

在一個潛在龐大的資料集中，探索重要的模式通常具有相當大的價值。此類問題中最基本的形式是頻繁模式的挖掘，它的目的是發現在資料集中出現次數最多的項目集合。雖然已經有很多演算法來解決頻繁模式挖掘的問題，但是現實的環境通常更加複雜。例如：假設在數據集中，每一個項目都可能具有不同的價值，因此有了高效用模式挖掘，目的在於發現數據集中具有最高效用的項目集合。由於高效用模式不具有反單調的特性，因此高效用模式的子集不一定保證也是高效用模式。缺乏反單調性特性，使得在龐大的數據集中挖掘高效用模式比頻繁模式更加困難。

在這項研究中，我們觀察到在實踐中通常不需要進行確定性和精確性的高效用模式挖掘。相反，我們通常只需發現一些具有高概率是高效用模式的項目集，而不一定是真正具有最高效用的項目集。透過放寬這些要求，我們提出一種近似高效用模式挖掘演算法（PAHUPMA），他可以在給定的數據集中發現高效用模式，而無須掃描整個數據集。因此，高效用模式挖掘所產生的 IO 流量大幅減少，而顯著提升挖掘的性能。根據我們初步的實驗結果，PAHUPMA 的延遲只是現有演算法（如 HUI-Miner 和 HUMP）的一小部分。我們相信，PAHUPMA的優異性能能夠有效擴展高效用模式挖掘的應用範疇。