Cho là một tập hữu hạn các mục; mỗi tập được gọi là một **tập mục**. Cho là một **cơ sở dữ liệu giao dịch**. Mỗi giao dịch , với định danh duy nhất , là một tập con của .

**Utility nội bộ** đại diện cho số lượng của mục trong giao dịch . **Utility bên ngoài** là giá trị lợi nhuận đơn vị của . **Utility** của trong được định nghĩa là:

**Utility** của tập mục trong được định nghĩa là:

**Utility** của trong được định nghĩa là:

**Utility giao dịch** (TU) của giao dịch được định nghĩa là:

**Ngưỡng utility tối thiểu** , được chỉ định bởi người dùng, là một tỷ lệ phần trăm của tổng giá trị TU của cơ sở dữ liệu, và giá trị **utility tối thiểu** được tính như sau:

Một tập mục được gọi là **tập mục utility cao** (HUI) nếu . Với một cơ sở dữ liệu giao dịch , nhiệm vụ của **khai phá HUI** (HUIM) là xác định tất cả các tập mục có utility không nhỏ hơn **min\_util**.

**Transaction-weighted utilization** (TWU) của tập mục là tổng của utility giao dịch của tất cả các giao dịch chứa , và được định nghĩa là:

là một **tập mục có transaction-weighted utilization cao** (HTWUI) nếu . Một HTWUI chứa mục được gọi là **k-HTWUI**.