# PRÍLOHA G – Detailné vyhodnotenie úspešnosti odporúčania pre dataset SACBEE

V tejto časti uvádzame niektoré podrobnejšie grafy a tabuľky z celkového vyhodnotenia pre dataset SACBEE.

## Vyhodnotenie úspešnosti odporúčania pre skupinu parametrov hľadania frekventovaných množín

V tejto časti sa budeme venovať vyhodnoteniu úspešnosti odporúčania s rôznymi konfiguráciami skupiny parametrov pre algoritmus hľadania frekventovaných množín. Ide o parametre (bližšie sú opísané v časti 7.1.6) a ich hodnoty:

* **Minimálna podpora (ms):** {0.005, 0.01, 0.02, 0.03, 0.04, 0.05, 0.1}
* **Miera uvoľnenia (rr):** {0.1, 0.5, 0.9}
* **Dĺžka segmentu (fsl):** {25, 50, 100, 150, 200, 500}
* **Maximálna dĺžka množiny (mil):** {10}
* **Veľkosť okna (ws)**: {5,10,15}

Spolu je to 270 rôznych konfigurácií. Čo sa týka hodnôt minimalnej podpory tá vyjadruje minimálnu podporu v rámci jedného segmentu. Teda ak dĺžka segmentu je napr. 50 a hodnota minimálnej podpory je 0.02 a menšia tak je každá množina považovaná za frekventovanú. Kombinácie parametrov kde bude teda hodnota budeme ignorovať.

## Vyhodnotenie úspešnosti odporúčania pre skupinu parametrov zhlukovania

V tejto časti sa budeme venovať vyhodnoteniu úspešnosti odporúčania s rôznymi konfiguráciami skupiny parametrov pre algoritmus zhlukovania *CluStream*. Ide o parametre (bližšie sú opísané v časti 5.1.6) a ich hodnoty:

* **Počet zhlukov** *(skr. gc)*: {2, 4, 6, 8}
* **Hraničný počet zmien v modeli používateľa** *(skr. tuc)*: {5,10,15}
* **Hraničný počet zmien v mikrozhlukoch***(skr. tcm)*: {50, 100, 200, 400, 800}
* **Maximálny počet mikrozhlukov***(skr. mmc****)*:** {100,1000}

Spolu je to 120 rôznych konfigurácií. Nastavenie ostatných parametrov sme vybrali z jednej z najlepších konfigurácií z predchádzajúcej časti (konfigurácia 2 v tabuľke 8), ktorá dosahovala najlepšie výsledky pre menšie počty odporúčaných položiek a zároveň bola veľmi rýchla.

## Vyhodnotenie úspešnosti odporúčania pre skupinu parametrov odporúčania

V tejto časti sa budeme venovať vyhodnoteniu odporúčania v spojitosti s rôznymi konfiguráciami skupiny parametrov pre odporúčanie. Táto skupina je zvláštna, keďže pozorujeme len jeden parameter hoci tu patrí aj parameter *rc* teda počet odporúčaných položiek, ktorého rôzne hodnoty ale sledujeme implicitne pri každom behu aj v ostatných skupinách parametrov. Ide teda o jediný parameter:

* **Veľkosť okna vyhodnocovania** *(skr. ews)*: {1,2,3,4,5,6,7, 8, 9, 10}

Nastavenie hodnôt ostatných parametrov sme zobrali z najlepšej konfigurácie čo sa týka presnosti vybranej z predchádzajúcej skupiny parametrov zhlukovania (konfigurácia 1 v tabuľke Tabuľka 6).

Zmena hodnoty parametra *ews* mení aj množinu sedení, ktoré sú vyhodnocované. Čím dlhšie je vyhodnocované okno tým menej sedení má dostatočnú dĺžku aby mohli byť vyhodnotené. To je závažný fakt, ktorý má vplyv na relevantnosť porovnávania jednotlivých konfigurácií.

## Vyhodnotenie úspešnosti odporúčania pre ostatné parametre metódy

V tejto časti sa budeme venovať vyhodnoteniu odporúčania v spojitosti s rôznymi konfiguráciami skupiny osatných parametrov metódy. Ide o parametre (bližšie sú opísané v časti 5.3) a ich hodnoty:

* **Hraničná hodnota rozdielu identifikátorov zhlukovania** *(skr. tcdiff)*: {1,2,3,4, 5,6,7,8,9,10}
* **Minimálna rýchlosť** *(skr. mts)*: {15}

V tejto skupine sme sa rozhodli testovať rôzne hodnoty len prvého parametra *tcdiff*. Rôzne hodnoty parametra pre obmedzenie minimálnej rýchlosti budeme testovať až vo vyhodnotení rýchlosti spracovania (7.3.2), kde sledujeme k akému zhoršeniu metriky *presnosť* bude dochádzať pri zrýchľovaní spracovania transakcií na už vybranej celkovo najlepšej konfigurácií. Nastavenie hodnôt ostatných parametrov sme zobrali z  konfigurácie vybranej z predchádzajúcej skupiny parametrov odporúčania, kde sme pozorovali len jeden parameter určujúci veľkosť vyhodnocovacieho okna. Vybrali sme nastavenie vyhodnocovacieho okna s veľkosťou 2, ktoré príliš nezmenšuje množinu vyhodnocovaných sedení a zároveň dosahuje jedny z najlepších výsledkov metriky *presnosť*.

## Celkové vyhodnotenie úspešnosti odporúčania pre najlepšie konfigurácie

Na základe vybraných najlepších konfigurácií zo všetkých skupín parametrov ako sme ich uvádzali v predchádzajúcich častiach sme skonštruovali priestor prehľadávania v ktorom budeme hľadať celkovo najlepšiu konfiguráciu. Ide o parametre a ich hodnoty:

* **Minimálna podpora** *(skr. ms):* {0.005, 0.03, 0.05}
* **Miera uvoľnenia** *(skr. rr)*:{0.1, 0.5}
* **Dĺžka segmentu** *(skr. sl)*:{25, 50, 500}
* **Maximálna dĺžka množiny** *(skr. mil)*:{10}
* **Veľkosť okna** *(ws)*: {15}
* **Počet zhlukov** *(skr. gc)*: {6, 8}
* **Hraničný počet zmien v modeli používateľa** *(skr. tcu)*: {5}
* **Hraničný počet zmien v mikrozhlukoch***(skr. tcm)*: {400, 800}
* **Maximálny počet mikrozhlukov***(skr. mmc****)*:** {100,1000}
* **Veľkosť okna vyhodnocovania** *(skr. ews)*: {1,2,3,4,5}
* **Hraničná hodnota rozdielu identifikátorov zhlukovania** *(skr. tcdiff)*: {15}
* **Minimálna rýchlosť** *(skr. mts)*: {15}