

Tato prezentace je součástí wiki-prezentace [Metoda GUHA a systém LISp-Miner](#)

Je dostupná z [této adresy](#)

Verse 29. 7. 2019

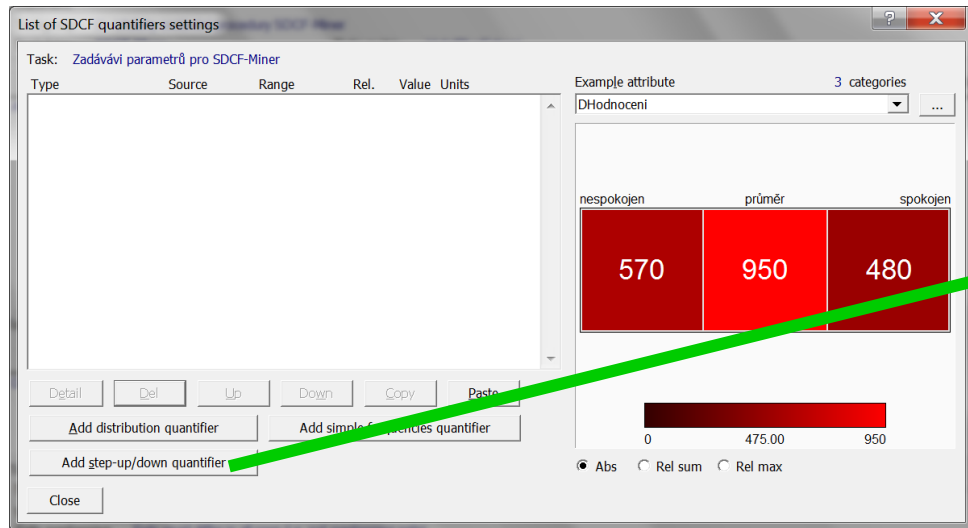
Zadávání step-up_down SDCF-kvantifikátorů pro proceduru SDCF-Miner

Jan Rauch

Katedra informačního a znalostního inženýrství

Vysoká škola ekonomická v Praze

Start tlačítkem Add simple frequencies quantifier



- Definice se provádí zadáním parametrů 1 - 6.
- Pro všechny parametry se nabízejí defaultní hodnoty.
- Podmínka kvantifikátoru je definována parametrem 6.

SDCF Step-up/down quantifier settings

Step type: Steps-up (6)

Number of steps-up in frequency of adjacant categories: 6 (6)

Relation: Greater than or equal

Step count threshold value: 80 (5) Consecutive steps (the longest sequence)

Step-count units: Relative [%] to act range (5)

Relative to number of categories in the currently selected range-1. Count value is multiplied by ('number of categories in current range minus 1')/100

Category Range: From 0 (1) To 100 (1)

Absolute category index

Relative range [%] to act number of categories

Step Properties: Minimal step size (frequency difference): 1 (4)

Minimal-step-size units: Absolute number (4)

Change in frequency as an absolute number. Minimal step size is left as given.

Primary IM Settings: Set as primary IM From 0 To 1 Normalize value range: 0 1

Note: -

Buttons: OK, Cancel

Výchozí SDCF-tabulka

Step-up_down SDCF-kvantifikátor se aplikuje na SDCF-tabulku $CF(A, \alpha, \beta, \chi, M)$

kategorie atributu A	a_1	...	a_p	...	a_q	...	a_K	Σ
absolutní četnosti v matici $M/\chi \wedge \alpha$	$n_{\alpha,1}$...	$n_{\alpha,p}$...	$n_{\alpha,q}$...	$n_{\alpha,K}$	n_α
absolutní četnosti v matici $M/\chi \wedge \beta$	$n_{\beta,1}$...	$n_{\beta,p}$...	$n_{\beta,q}$...	$n_{\beta,K}$	n_β
absolutní četnosti v matici M/χ	n_1	...	n_p	...	n_q	...	n_K	n
absolutní četnosti v matici M	m_1	...	m_p	...	m_q	...	m_K	n_T

1 - Category Range

- Category Range určuje výsek SDCF-tabulky, na který budou aplikovány ostatní parametry.
- Provádí se způsobem popsáním [zde](#).
- Výsek SDCF-tabulky týkající se K kategorií je dán dvojicí $\langle p, q \rangle$ celých čísel splňujících $1 \leq p \leq q \leq K$.
- Výsek $\langle p, q \rangle$ SDCF-tabulky se týká frekvencí pro kategorie a_p, \dots, a_q .

Category Range

From To Reset to All

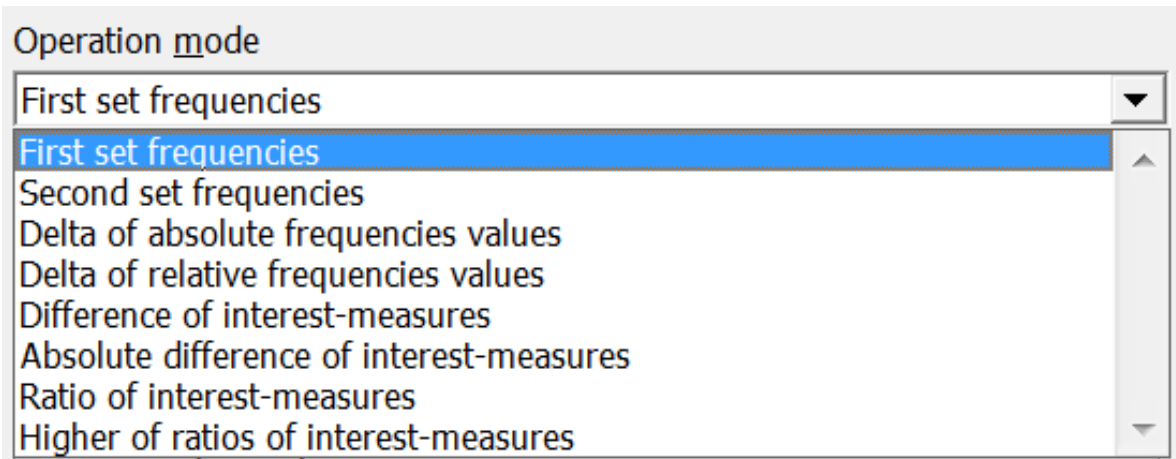
Absolute category index

Relative range [%] to act number of categories

kategorie atributu A	a_1	...	a_p	...	a_q	...	a_K	Σ
absolutní četnosti v matici $M/\chi \wedge \alpha$	$n_{\alpha,1}$...	$n_{\alpha,p}$...	$n_{\alpha,q}$...	$n_{\alpha,K}$	n_α
absolutní četnosti v matici $M/\chi \wedge \beta$	$n_{\beta,1}$...	$n_{\beta,p}$...	$n_{\beta,q}$...	$n_{\beta,K}$	n_β
absolutní četnosti v matici M/χ	n_1	...	n_p	...	n_q	...	n_K	n
absolutní četnosti v matici M	m_1	...	m_p	...	m_q	...	m_K	n_T

2 - Operation mode

Vybírá se jeden z operačních módů nabízených v menu *Operation mode*. Ten určuje, jakým způsobem bude aplikována vybraná míra zajímavosti na SDCF-tabulku.



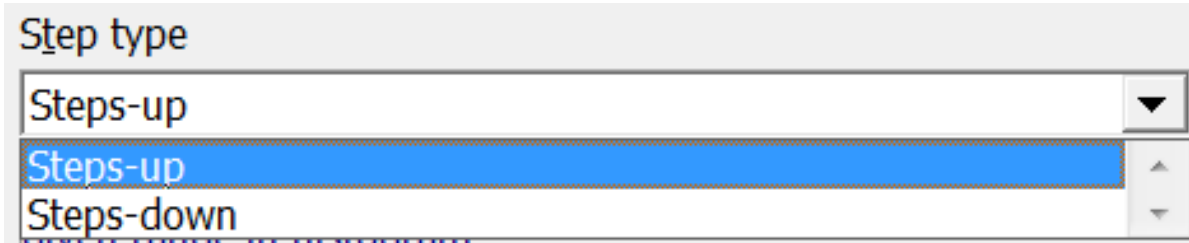
Pro první čtyři nabízené Operation mode se nejprve vytvoří vektor $\langle z_p, \dots, z_q \rangle$ takto:

- pro *First set frequencies* $\langle z_p, \dots, z_q \rangle = \langle n_{\alpha,p}, \dots, n_{\alpha,q} \rangle$
- pro *Second set frequencies* $\langle z_p, \dots, z_q \rangle = \langle n_{\beta,p}, \dots, n_{\beta,q} \rangle$
- pro *Delta of absolute frequencies values*
 $\langle z_p, \dots, z_q \rangle = \langle |n_{\alpha,p} - n_{\beta,p}|, \dots, |n_{\alpha,q} - n_{\beta,q}| \rangle$
- pro *Delta of relative frequencies values*
 $\langle z_p, \dots, z_q \rangle = \langle \left| \frac{n_{\alpha,p}}{n_\alpha} - \frac{n_{\beta,p}}{n_\beta} \right|, \dots, \left| \frac{n_{\alpha,q}}{n_\alpha} - \frac{n_{\beta,q}}{n_\beta} \right| \rangle$.

Pro zbývající čtyři nabízené módy není výpočet k dispozici.

3 - Step type

Vybírá se jedna z voleb nabízených v menu *Interest measure type*.



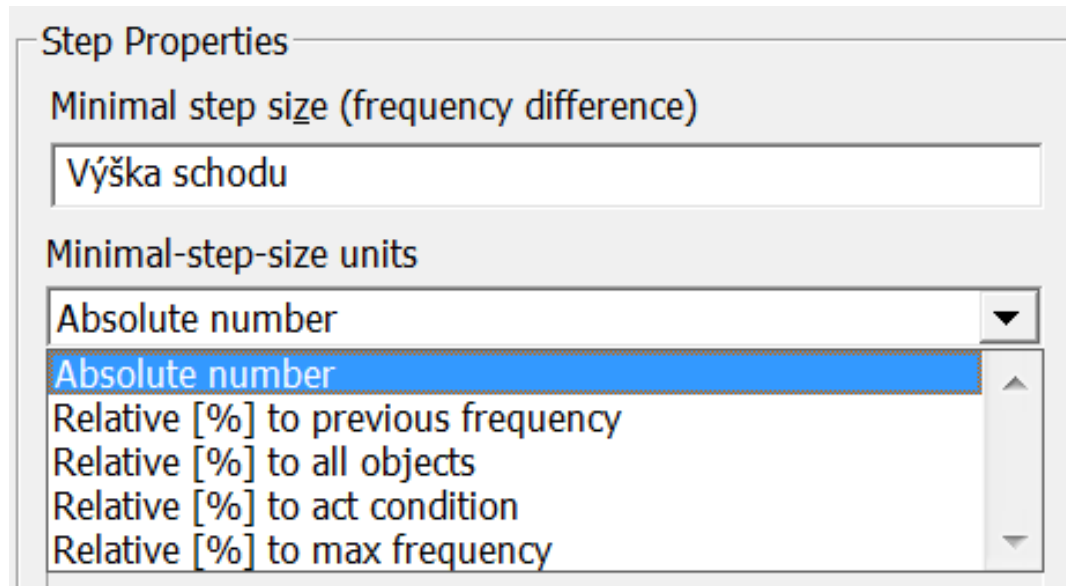
Schody se aplikují na vektor $\langle z_p, \dots, z_q \rangle$.

Podrobnosti o nabízených typech schodů jsou zde:

- [Steps-up](#)
- [Step-down](#) .

4 - Step properties

Zadáva se minimální výška schodu v poli *Minimal step size (frequency difference)*
a vybírá se jedna z voleb nabízených v menu *Minimal-step-size units*

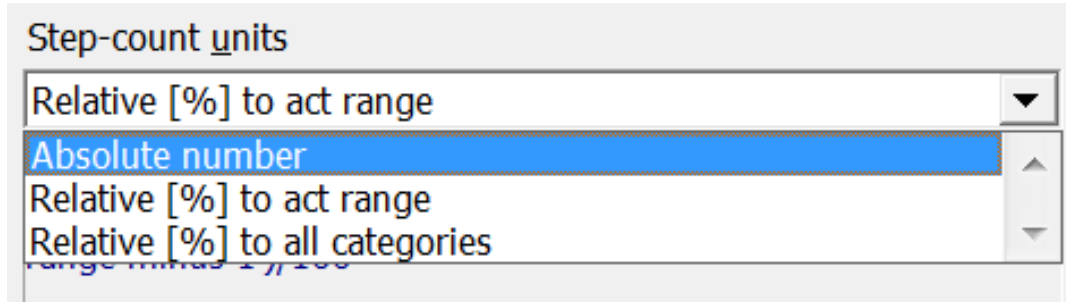


Parametr se uplatňuje na vektor $\langle z_p, \dots, z_q \rangle$.

Podrobnosti jsou [zde](#).

5 - Step-count units

Vybírá se jedna z voleb nabízených v menu *Step-count units*



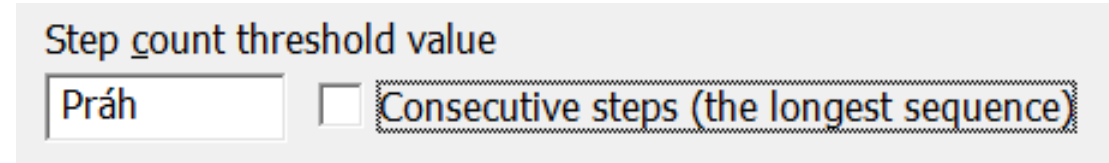
Step-count units

Relative [%] to act range

Absolute number

Relative [%] to act range

Relative [%] to all categories



Step count threshold value

Práh Consecutive steps (the longest sequence)

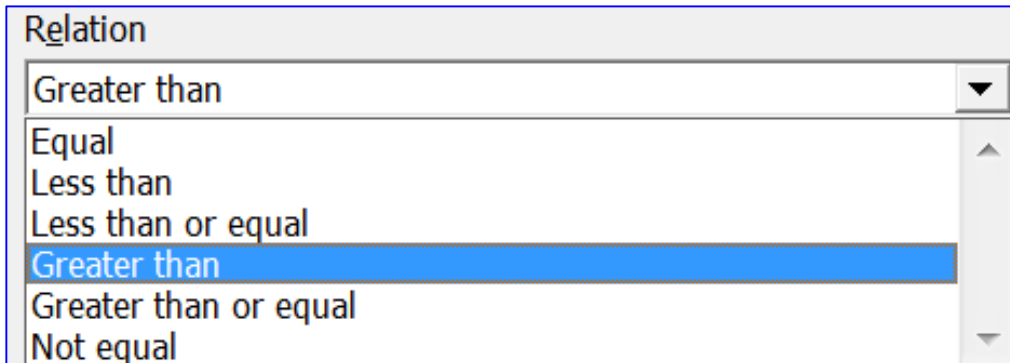
Na základě této volby se určí jednotky, ve kterých je hodnota *Práh* zadaná v poli *Step-count threshold value*. To znamená, že se před porovnáváním dle parametru 5 zadaná hodnota *Práh* vynásobí vhodným číslem dle volby v menu. Výslednou hodnotu pro porovnání značíme *PráhAkt*. Platí:

- pro *Absolute number* $PráhAkt = Práh$
- pro *Relative [%] to act range* $PráhAkt = \frac{100}{p-q} * Práh$
- pro *Relative [%] to all categories* $PráhAkt = \frac{100}{K-1} * Práh$

6 - Relation x Threshold value

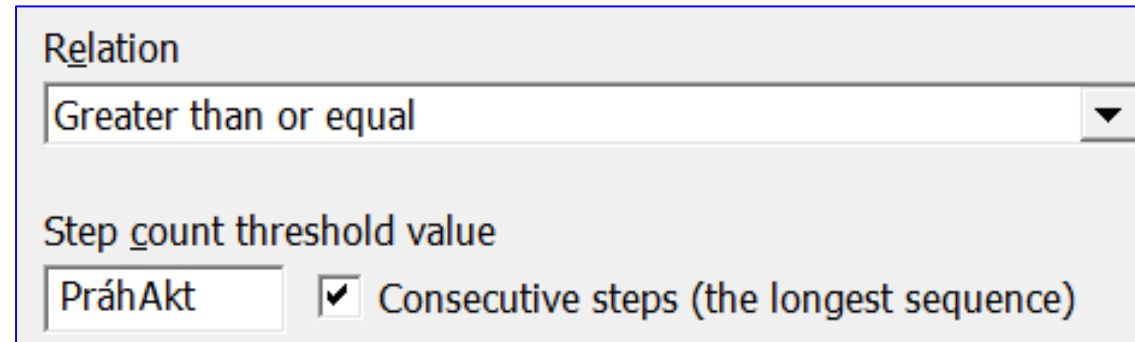
Na základě volby v nabídce *Relation* se vybere relace, která se použije pro porovnání počtu schodů zadaného typu s hodnotou *PráhAkt* vypočtenou dle parametru 5 - [Step-count units](#).

Zaškrtnutím volby *Consecutive steps (the longest sequence)* je možno požadovat, aby schody vybraného typu následovaly bezprostředně za sebou.



Relation

- Greater than
- Equal
- Less than
- Less than or equal
- Greater than
- Greater than or equal
- Not equal



Relation

Greater than or equal

Step count threshold value

PráhAkt Consecutive steps (the longest sequence)

Platnost vybrané relace je považována za podmínku definující SDCF-kvantifikátor zadaný parametry 1 až 6.