

Charakteristika výkonu podobnostného vyhľadávania v neštruktúrovaných databázach závisí od použitého podobnostného modelu. Vlastnosti metrických priestorov nám umožňujú efektívne indexovať dáta pomocou tzv. metrických prístupových metód. Ale pre prípad nemetrických priestorov, ktoré sú typické pre multimediálne, medicínske a vedecké databázy, a v ktorých neplatia axiomy metrických priestorov, zatiaľ nepoznáme všeobecné riešenie.

Na základe úspešnej aplikácie ptolemaického modelu indexovania, predstavujeme SIMDEX Framework, univerzálny nástroj, ktorý dokáže objaviť alternatívne metódy indexácie dát za účelom efektívneho podobnostného vyhľadávania pre ľubovoľný podobnostný model. Na pozadí prehľadáva priestor platných axióm tak, aby našiel nové techniky určené pre indexovanie databáz. Preskúmame všetky existujúce varianty (prostý I-SIMDEX; GP-SIMDEX a PGP-SIMDEX využívajúce genetické programovanie) a zhodnotíme ich prínos a použitie v praxi pre profesionálov v rozličných doménach.

Nakoniec opíšeme konkrétnu aplikáciu SIMDEX Framework-u v praxi na vytvorenie indexu Smart Pivot Table s pokročilým filtrovaním pre metrické priestory (Triangle+ filtering) spoločne s technikou na zlepšovanie kvality filtrovania (LowerBound Tightening). Vo všetkých prípadoch uvádzame aj experimentálne vyhodnotenie a porovnanie spomínaných metód.