



Universiteit
Leiden
The Netherlands

Progressive Indexes

Timbó Holanda, P.T.

Citation

Timbó Holanda, P. T. (2021, September 21). *Progressive Indexes*. *SIKS Dissertation Series*. Retrieved from <https://hdl.handle.net/1887/3212937>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3212937>

Note: To cite this publication please use the final published version (if applicable).

Samenvatting

Interactieve verkenning van grote hoeveelheden gegevens komt steeds vaker voor, omdat datawetenschappers proberen interessante informatie te extraheren uit grote complexe gegevenssets. Dit scenario vormt een uitdaging voor traditionele databasesystemen, aangezien (1) er van tevoren niets bekend is over de query-workload, (2) de query-workload voortdurend verandert, en (3) het systeem interactieve antwoorden moet geven op de uitgegeven queries. Deze omgeving is een uitdaging voor het maken van indexen, aangezien traditionele database-indexen vooraf moeten worden gemaakt, en dus a priori kennis van de werkbelasting nodig hebben, om efficiënt te zijn.

In dit werk introduceren we *Progressive Indexing*, een nieuwe prestatiegerichte indexeringstechniek die zich richt op het automatisch bouwen van indexen en tegelijkertijd interactieve reactietijden biedt op inkomende vragen. Dankzij het ontwerp kunnen zoekopdrachten een beperkt budget hebben om te besteden aan het maken van indexen. Het indexeringsbudget wordt automatisch afgestemd op elke query voordat de query wordt verwerkt. Hierdoor kunnen systemen interactieve antwoorden geven op vragen tijdens het maken van een index, terwijl ze robuust zijn tegen verschillende werkbelastingpatronen en gegevensverdelingen.

We ontwikkelen progressieve algoritmen om één en meerdere dimensies te indexeren. Daarnaast introduceren we *Progressive Merges*, een robuust algoritme dat toevoegingen in onze progressieve indexen samenvoegt zonder afzonderlijke zoekopdrachten te bestraffen.

