Hybrid: A Large-Scale In-memory Image Analytics Engine

Michael Gubanov
Department of Computer Science, University of Texas at San Antonio
mikhail.gubanov@utsa.edu

ABSTRACT
Analytical image/video processing tasks such as scene/face/activity recognition are historically performed externally of most relational database management systems. Relational engines are optimized for relational data and therefore, have weaker support for non-relational data such as images or video. We have been working on Hybrid, a high-performance in-memory analytics engine, which supports the advanced access capabilities for both relational/data and structured data via SQL or JSON. This allows the user to query both relational (rows and columns of a table) and video/image contents (objects, activities, scene attributes) in a single SQL or HybridSQL/JSON statement [Gubanov and Pyayt, 2013, Gubanov and Shapiro, 2012]. Analytical tasks may then be performed on both types of data, without the expensive ETL (extract-transform-load) process. An example of the Hybrid query to a relational and image data source (queried in JSON) is illustrated below. The user manually specifies an SQL prefix in the query portion which accesses the relational data and JSON: prefix in the part querying image data. A query then fuses [Gubanov et al., 2009, Gubanov et al., 2014] information from images (cast to JSON) and Web tables (relational), before outputting the best deal for a popular Broadway show [Gubanov and Stonebraker, 2014, Abedjan et al., 2014].

```
SELECT MIN(S.Price) AS Price, S.Show, S.Theater, S.ShowTime
FROM Shows AS S
WHERE S.Show IN
(SELECT * FROM FlyersImage
WHERE date > "Sunday,27 November"
GROUP BY S.Show, S.Theater, S.ShowTime
ORDER BY Price ASC)
```

To accomplish this, it would first find all shows by Roald Dahl mentioned in the show flyers (pre-processed by OCR software and stored in a JSON database WebImage), after Sunday, November 27. Next, it fuses it with relational table Shows having prices, theater addresses, and show times to find the best deal. For homogeneous queries in one language, the language can be detected automatically by the Query Processor using a machine learning classifier [Speed, 2010, Lin and Lin, 2003, Mitchell, 1997] trained to detect the query language.

In addition to the ability store extracted image metadata in JSON, like in the query above, Hybrid represents and stores images and videos as matrices in a relational table column and supports a subset of linear algebra operators to perform more complex analytics on them. This approach differs from the classical, relational storage techniques which store images/videos as BLOBS with limited access through SQL. Because they are stored as unstructured binary objects, the meaningful use of BLOB data requires the user to develop an external, user-defined function (e.g. in C++). Hybrid provides access to both video/image and structured data from the same SQL statement, which enables Hybrid analytics, analytics on both relational and image data. Our current prototype performs OCR, interactive activity, and pattern recognition in SQL using both large-scale matrix and relational data.

1. REFERENCES


