

GREDIA Space Filing Curves Component

Motivation and Challenges

The GREDIA Space Filing Curves (SFC) component is exploited in the Gredia platform to provide an efficient and scalable solution for searching multimedia content based on their annotations. The search functionality of the platform depends highly on the effectiveness of the indexing method for multiple attributes. Since the Kademia protocol supports natively only simple lookups for keys, an indexing method for multiple attributes is critical so as to enable query processing of point and range queries in one or more attributes.

In all the usage scenarios of the Gredia platform, it is of great importance, that queries on the indexed attributes with any combination of values can be proposed quickly. Nevertheless, it is necessary that the indexing of the metadata files will be compatible with the structure of the Kademia Distributed Hash Tables (DHT), so as queries to be handled automatically by the inherent mechanisms of the metadata overlay without influencing its features, such as scalability and fault tolerance.



Solution Overview

The GREDIA SFC component implements an indexing method based on Space Filing Curves algorithm, which can be integrated with the enhanced Kademia protocol to facilitate point and range queries on any combination. The main problem is to take into account the values of all these attributes during the placement of the metadata files in the overlay. The SFC method achieves to enable the indexing of multiple attributes, which can be queried either independently or combinatorially. It also aims to place relative descriptions in close nodes so as to reduce the queried to be served by a relatively limited number of nodes.

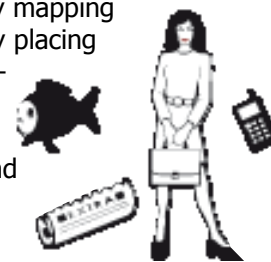
The IDs during insertion and search of metadata files are generated by the SFC component. According to the indexing method based on SFCs, each attribute corresponds to a dimension of a multidimensional space. Afterwards, a combination of values in the multidimensional space is mapped down to a single SFC index to be used as key in the Metadata overlay.



Moreover, the SFC module accomplishes all the needed operations, so as the communication with the Metadata overlay to be achieved. It is responsible for the parsing and encoding the values of the indexed attributes to a constant number of bits, so as to be mapped to a value of a dimension. The attributes to be indexed may belong to the following categories: strings, categorical and dates.

Innovation

The innovation of the SFC component is that enables indexing of multiple attributes by mapping them down to single values. Moreover, the proposed method aim t preserve locality by placing relative metadata files to close nodes, as fas as this is possible. The processing of queries in traditional DHTs overlays is efficient only if the ID of a the searched item is known. The SFC based mapping used for indexing enables the indexing of multiple attributes and the determination of Id ranges corresponding to the queries of users and thus allowing the processing of range queries in the Kademlia based DHT.



Business Impact

The SFC component targets to provide a solution for indexing multiple attributes contained in annotations of multimedia content concurrently. The provided mapping is consistent and it is essential, so as range queries to be supported by DHT systems.

Interoperability

The SFC component can be readily used in any solution that requires mapping of multiple values to a single index and vice versa. Moreover, close indices in the single dimensions correspond to close values in the multidimensional space. The provided mapping can be used for supporting range queries in other DHT structures as well, such as Chord, Pastry. Also, its functionality can also be extended to other applications, such as image processing.



Partners Involved



Contact

Institute of Communication And Computer Systems

<http://www.iccs.gr/>

Dimitrios Tsoumakos



symbian

DEUTSCHE WELLE



School of Informatics
CITY UNIVERSITY LONDON



**Banca Popolare
di Sondrio**