University of Zurich[UZH]

Department of Informatics

University of Zürich
Department of Informatics
Binzmühlestr. 14
CH-8050 Zürich
Phone. +41 44 635 43 11
Fax +41 44 635 68 09
www.ifi.uzh.ch/dbtg

UZH, Dept. of Informatics, Binzmühlestr. 14, CH-8050 Zürich

Oliver Leumann

Prof. Dr. Michael Böhlen
Professor
Phone +41 44 635 43 33
Fax +41 44 635 68 09
boehlen@ifi.uzh.ch

Zürich, February 16, 2015

MSc Project

Topic: Structuring and Processing Temporal Probabilistic Data in Migros

Big companies, such as Migros, monitor the shopping habits of their clients through loyalty or credit cards. Every time a client reaches the cashier and presents such a card, tuples of the form (clientID, productID, brandID, locationID, timestamp) are stored, fully recording the contents of their shopping cart. In order to derive insights from such data, we want to restructure them in a temporal probabilistic form where each tuple represents an event associated with a time interval and a probability value.

This MSc project is performed in co-operation with Migros and involves working with anonymous data provided by Migros. The responsible person from Migros is Peter Niederberger, BI expert, Migros IT-Services, Heinrichstrasse 216, 8031 Zürich.

Tasks

1. Study the aggregation of groups of point-timestamped tuples into interval-timestamped tuples [1].
2. Determine proper aggregation functions such that, apart from the interval timestamp, there is a probability value associated with each tuple that quantifies the frequency of the event [2].
3. Implementation and evaluation of the temporal aggregation and probability computation process.
4. Collection of a set of reference queries to be performed over the constructed temporal probabilistic database.
5. Analysis of the results of each query at a tuple-level with the help of the probabilistic lineage tree (PLT) and the probabilistic bubble chart (PBC).
6. Evaluation of the execution times and scalability of the reference queries.
7. Technical Report (20 pages).

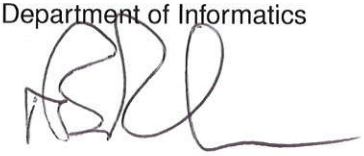
References

[1] Juozas Gordevicius, Johann Gamper, and Michael H. Böhlen. Parsimonious temporal aggregation. In *EDBT*, 2009.

[2] Sheldon M. Ross. *A First course in probability*. Prentice Hall, 6. ed edition, 2002.

Supervisor: Prof. Dr. Michael Böhlen
Start date: 16-02-2015
End date: 16-07-2015

University of Zürich
Department of Informatics



Prof. Dr. Michael Böhlen