

TIME^oDIFF: a Visual Approach to Compare Period Data

V. Del Fatto, A. Dignös, J. Gamper

Faculty of Computer Science, Free University of Bozen-Bolzano

July 13th 2018

Outline

Introduction

Motivation

Running Example

Approach

The $\text{TIME}^\circ\text{DIFF}$ Approach

Real World Scenario

Conclusion and Future Work

Motivation

Nowadays, applications in different domains require to collect and analyze information about the past (stored as period data).

Example: Flight Routes

Motivation

Nowadays, applications in different domains require to collect and analyze information about the past (stored as period data).

Example: Contracts in a Company

Problem

- ▶ Database systems are capable of supporting the storage and management of period data.
- ▶ However, the importance of an appropriate **visualization** and comparison of period data has **not been sufficiently studied**.

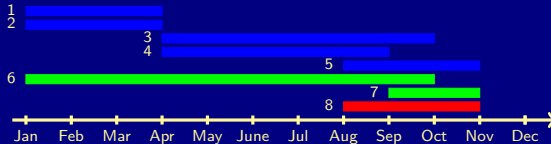
Contributions

- ▶ Novel visualization approach Time^odiff: suitable for visually comparing data with time periods.
- ▶ Time^odiff visualizes **amount** of data, **changes** in the data, and **amount of changes** in the data **at the same time** in a concise way.
- ▶ Enables decision makers to easily summarize, analyze and compare period data.

Running Example

A company database that stores employees working in departments and its graphical representation.

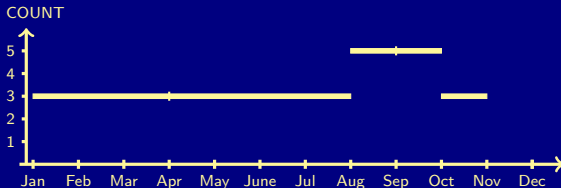
SSN	Dept	Start	End
1	D1	01.01.2018	01.04.2018
2	D1	01.01.2018	01.04.2018
3	D1	01.04.2018	01.10.2018
4	D1	01.04.2018	01.09.2018
5	D1	01.08.2018	01.11.2018
6	D2	01.01.2018	01.10.2018
7	D2	01.09.2018	01.11.2018
8	D3	01.08.2018	01.11.2018



Running Example

Temporal aggregation: counts the number of tuples in the data at each time point and merges identical results over consecutive time points into time periods.

Count	Start	End
3	01.01.2018	01.08.2018
5	01.08.2018	01.10.2018
3	01.10.2018	01.11.2018

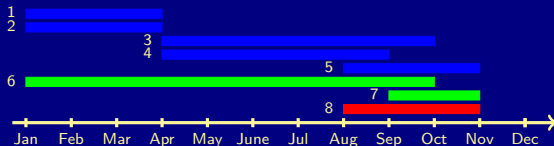
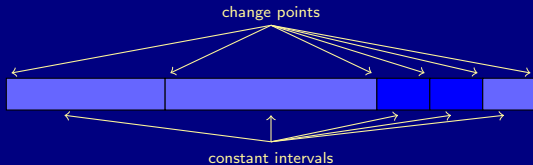


Features

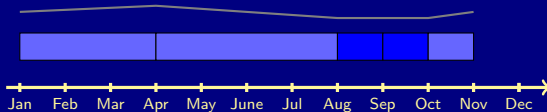
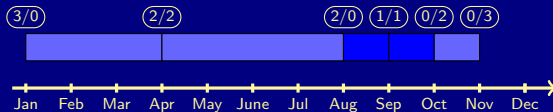
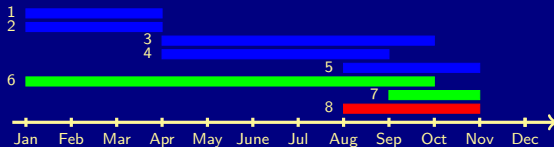
We identified the requirements for a visualization approach suitable for visualizing and comparing period data:

- ▶ Visualization of: historical amount of data, changes in the data, and the amount of changes in the data
- ▶ Scalability for large datasets
- ▶ Support for different groupings/datasets in a single chart

TIME^oDIFF Approach



TIME^oDIFF Approach



TIME°DIFF Approach

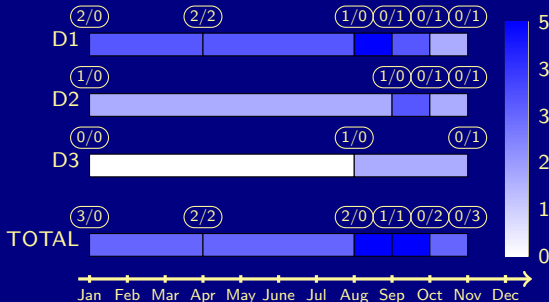
SSN	Dept	Start	End
1	D1	01.01.2018	01.04.2018
2	D1	01.01.2018	01.04.2018
3	D1	01.04.2018	01.10.2018
4	D1	01.04.2018	01.09.2018
5	D1	01.08.2018	01.11.2018
6	D2	01.01.2018	01.10.2018
7	D2	01.09.2018	01.11.2018
8	D3	01.08.2018	01.11.2018



Count	Dept	Start	End
2	D1	01.01.2018	01.04.2018
2	D1	01.04.2018	01.08.2018
3	D1	01.08.2018	01.09.2018
2	D1	01.09.2018	01.10.2018
1	D1	01.10.2018	01.11.2018
1	D2	01.01.2018	01.09.2018
2	D2	01.09.2018	01.10.2018
1	D2	01.10.2018	01.11.2018
1	D3	01.08.2018	01.11.2018

TIME^oDIFF Approach

Count	Dept	Start	End
2	D1	01.01.2018	01.04.2018
2	D1	01.04.2018	01.08.2018
3	D1	01.08.2018	01.09.2018
2	D1	01.09.2018	01.10.2018
1	D1	01.10.2018	01.11.2018
1	D2	01.01.2018	01.09.2018
2	D2	01.09.2018	01.10.2018
1	D2	01.10.2018	01.11.2018
1	D3	01.08.2018	01.11.2018

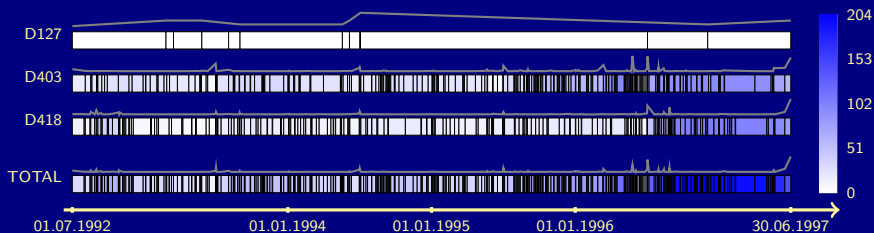


Real World Scenario

- ▶ We applied our Time^odiff approach on a real world dataset.
- ▶ The University Information systems (UIS) database¹ from the University of Arizona that records 83,857 job assignments of employees in departments over a period of 17 years at a granularity of days.
- ▶ For our purpose we compared three departments with names D127, D403, and D418, over a period of 5 years from 01.07.1992 till 30.06.1997.

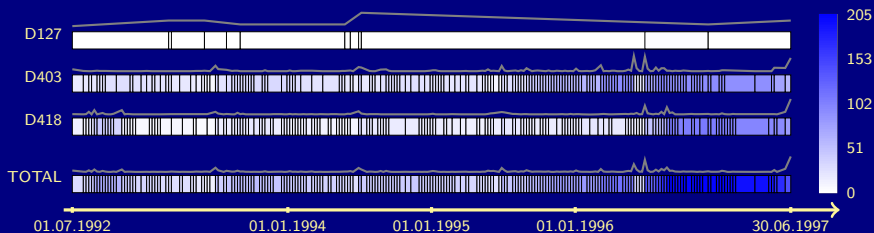
¹J. A. G. Gendrano, R. Shah, R. T. Snodgrass, and J. Yang: University information system (UIS) dataset. TimeCenter CD- 1, 1998.

Real World Scenario - granularity of days



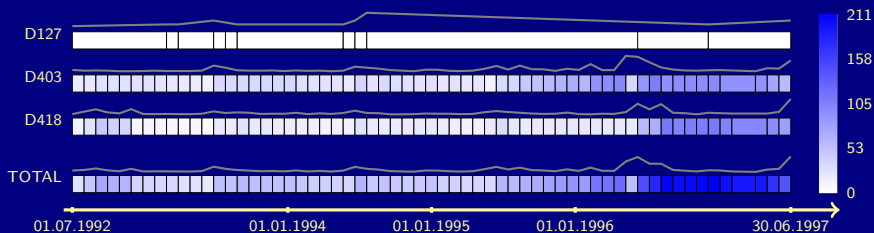
- ▶ number of employees increased over the years
- ▶ D127 has very few employees compared to the others
- ▶ D403 grew before D418 in terms of number of employees

Real World Scenario - granularity of weeks



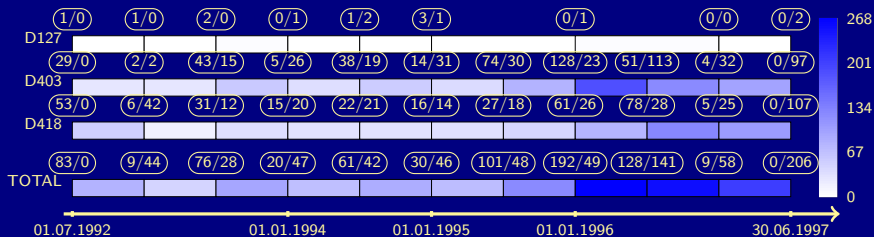
- ▶ number of employees increased over the years
- ▶ D127 has very few employees compared to the others
- ▶ D403 grew before D418 in terms of number of employees
- ▶ dominating change points were in mid-summer

Real World Scenario - granularity of months



- ▶ number of employees increased over the years
- ▶ D127 has very few employees compared to the others
- ▶ D403 grew before D418 in terms of number of employees
- ▶ dominating change points were in mid-summer

Real World Scenario - granularity of six months



- ▶ In the first half of the year 1994 (particularly for D418) the number of employees is approximately the same as in the second half of the year 1994. However, many changes occurred.

Conclusion

Time^odiff: a visual approach based on timebars for comparing information containing period data.

It contributes in helping decision makers to easily summarize and analyze information containing period data, by exploiting the following characteristics:

- ▶ visualization of the historical amount of data,
- ▶ visualization of changes in the data,
- ▶ visualization of the amount of changes in the data,
- ▶ scalability for large datasets, and
- ▶ support for different groupings/datasets in a single chart.

Future Work

- ▶ to implement a system capable of generating Time^odiff visualizations from input data
- ▶ to develop an interactive interface based on Time^odiff that allows filtering and zooming the data.
- ▶ to study in deep the scalability for very large datasets

Thanks for your attention!



vincenzo.delfatto@unibz.it