THE UNIVERSITY OF MELBOURNE

A Sensor Network Data Structure for Centralized Post-Processing of Events and Trends

Joost Kuckartz & Philip Collier
Department of Infrastructure Engineering
The University of Melbourne
Victoria, Australia

Australian Government
Australian Research Council

UNSW THE UNIVERSITY OF NEW SOUTH WALES

Leica Geosystems

C.R. KENNEDY SURVEY SOLUTIONS

VEKTA

NPO SODIS

Introduction

- SHM as sensor networks
- Data management issues
- Data structures
- The TB-tree
- Conclusion

SHM as sensor networks

- Multitude of sensors
- Centralized location for post-processing
- Decentralized research does/can not address SHM need for long-term and multi-sensor analysis

Data management issues

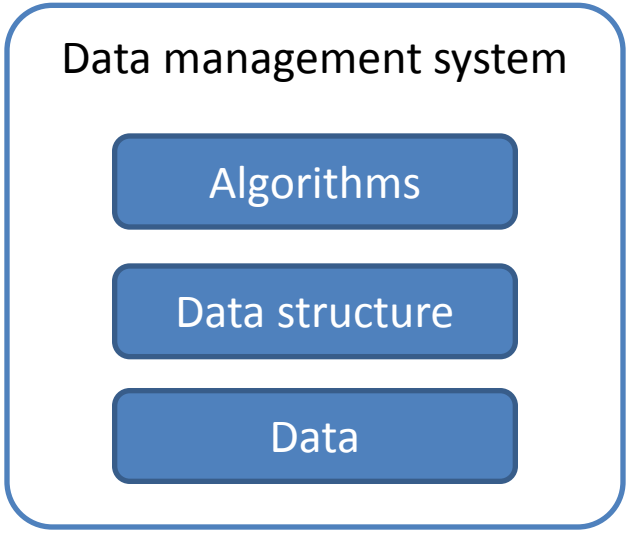
1. Different storage systems
2. Data at different frequencies
3. Different data representations
4. Post-processing
   a) Multiple sensors
   b) Different lengths of time
   c) New data created
   d) Efficient data access needed
5. Data storage limitation

Data management issues

	Storage system	Indexing scheme	Storage limitation mitigation
Burj Khalifa	Files & database	Statistical info of files in database	Delete data not associated with events
Svinesund Bridge	Files	File with 10-minute min, max, avg, stdev.	Unknown
Zhanjiang Bay Bridge	Database	Default database indexing	Move to archive, delete data after certain period

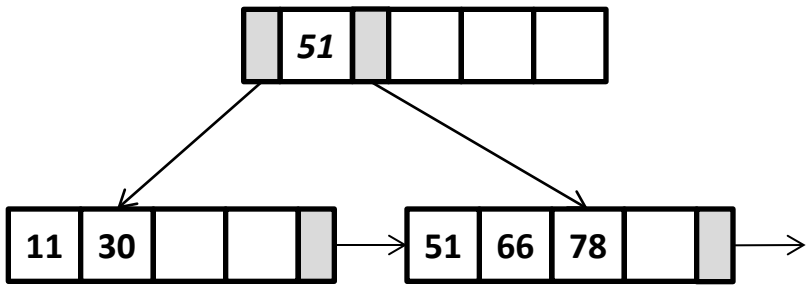
Data structures

- Definition: “An organization of information, usually in memory, for better algorithm efficiency”



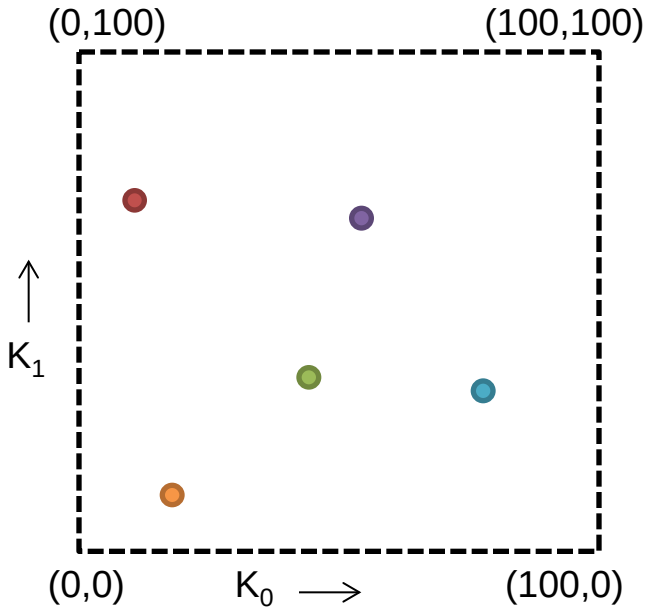
Data structures

- Many different structures exist
- Databases: optimized for data search, insert, delete
- Common: B-tree derivations
  - B+-tree:



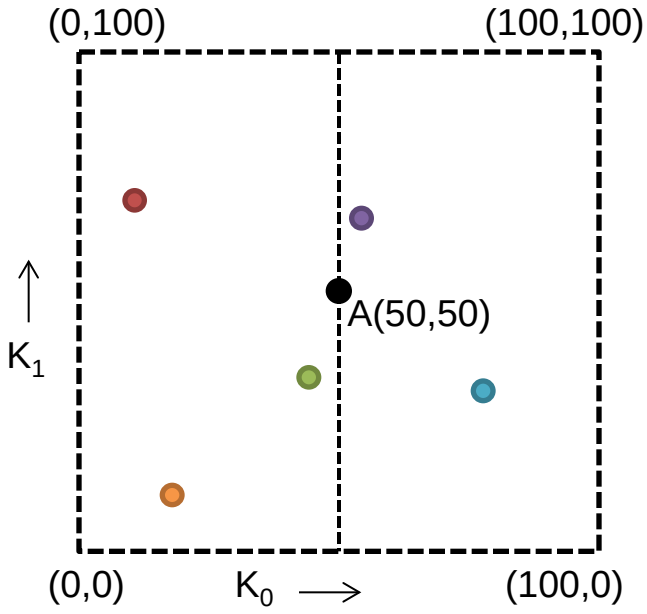
Data structures

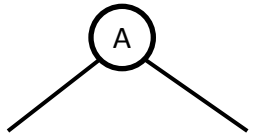
- Two-dimensional structure (spatial)



Data structures

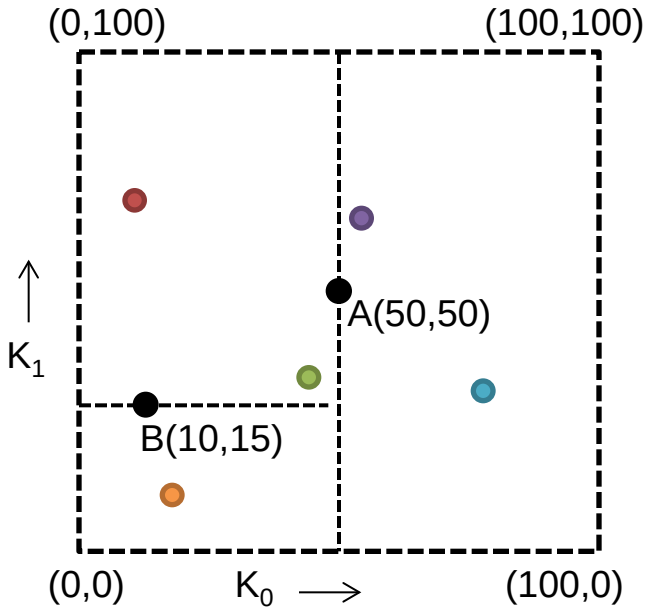
- Two-dimensional structure (spatial)

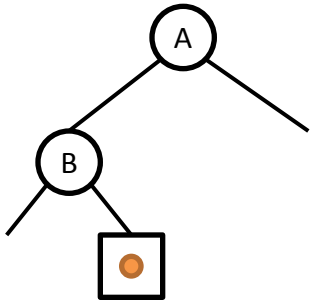




Data structures

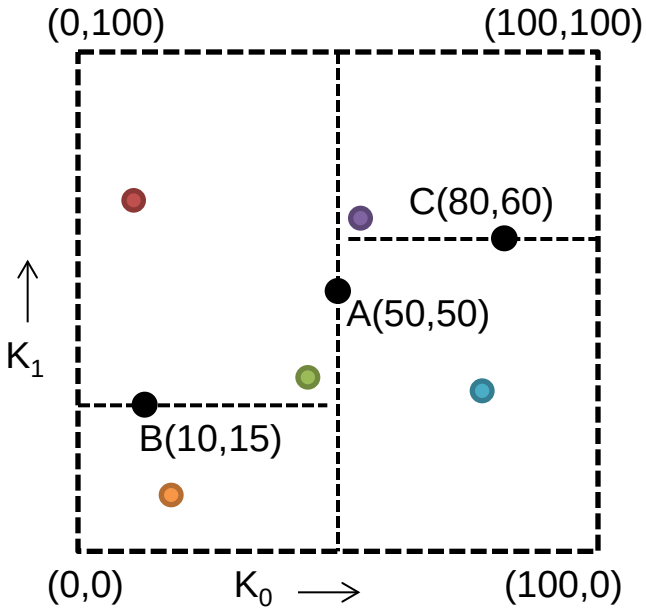
- Two-dimensional structure (spatial)

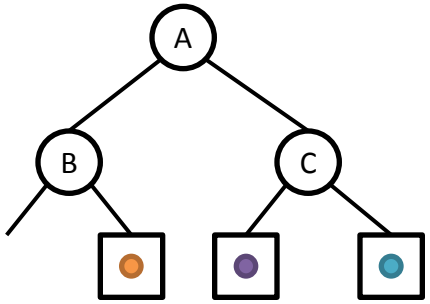




Data structures

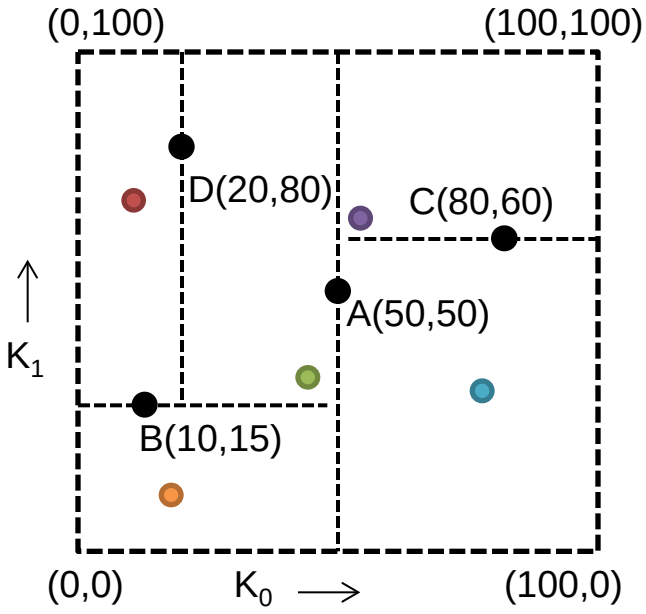
- Two-dimensional structure (spatial)

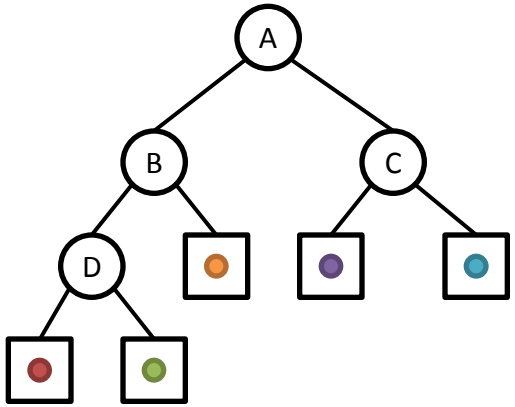




Data structures

- Two-dimensional structure (spatial)

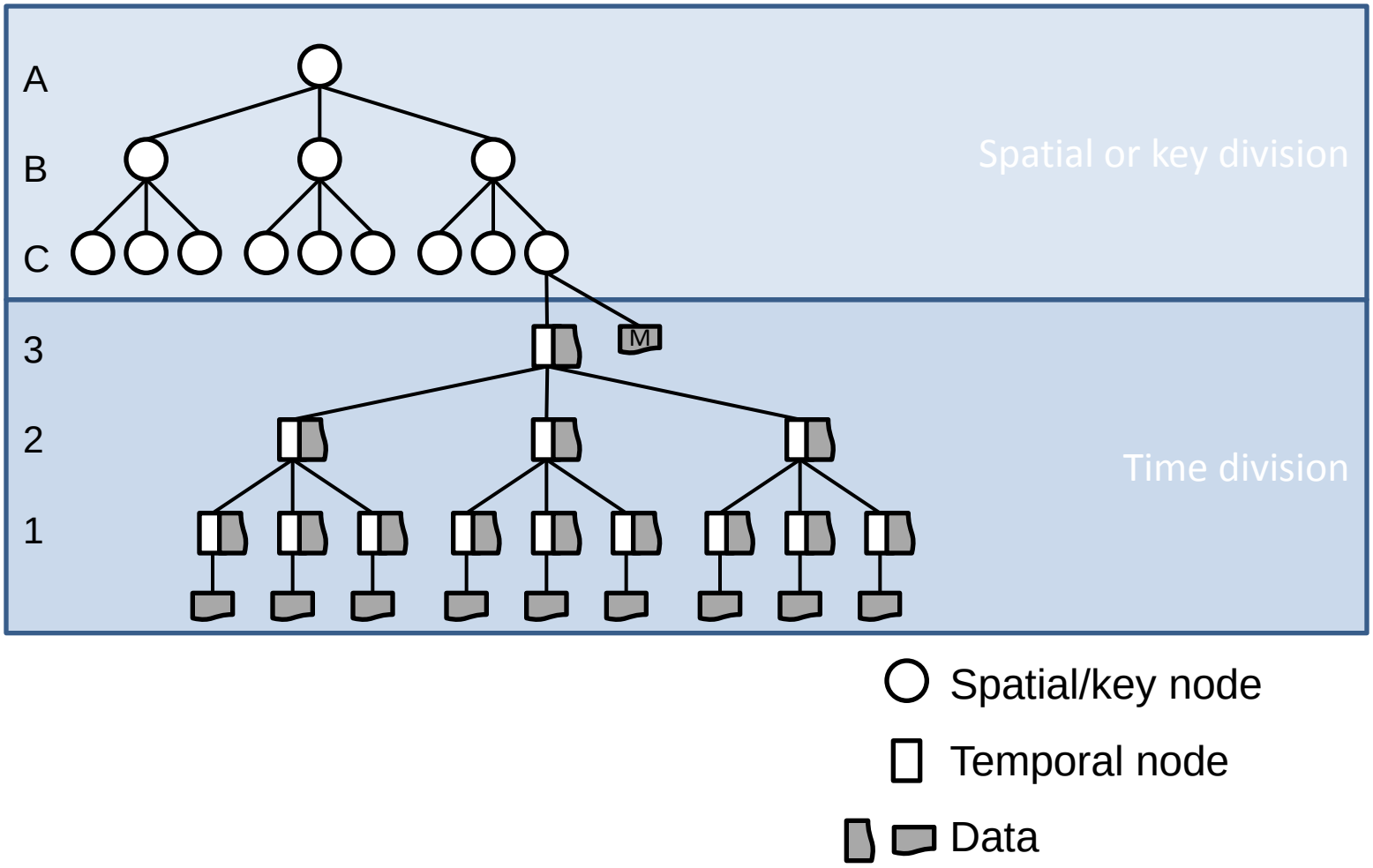




The TB-tree

- Issues with existing data structures:
  - Assume spatial movement over time: sensor locations are quasi-static and time is continuous
  - Data search requires traversal of all data
  - No long-term evaluation or statistical information present
- TB-tree (temporally blurred tree) solves this
  - Plus allows addressing of the data management issues mentioned before

The TB-tree

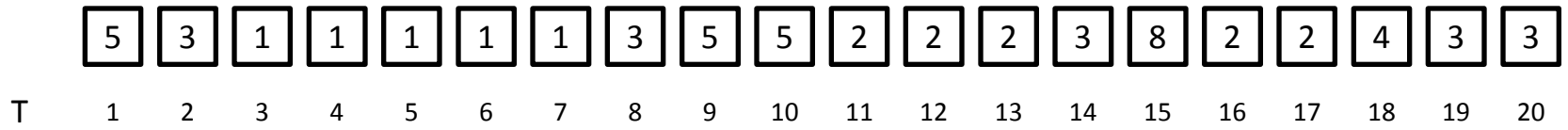

A
B
C
3
2
1
M
Spatial or key division
Time division
Spatial/key node
Temporal node
Data

The TB-tree

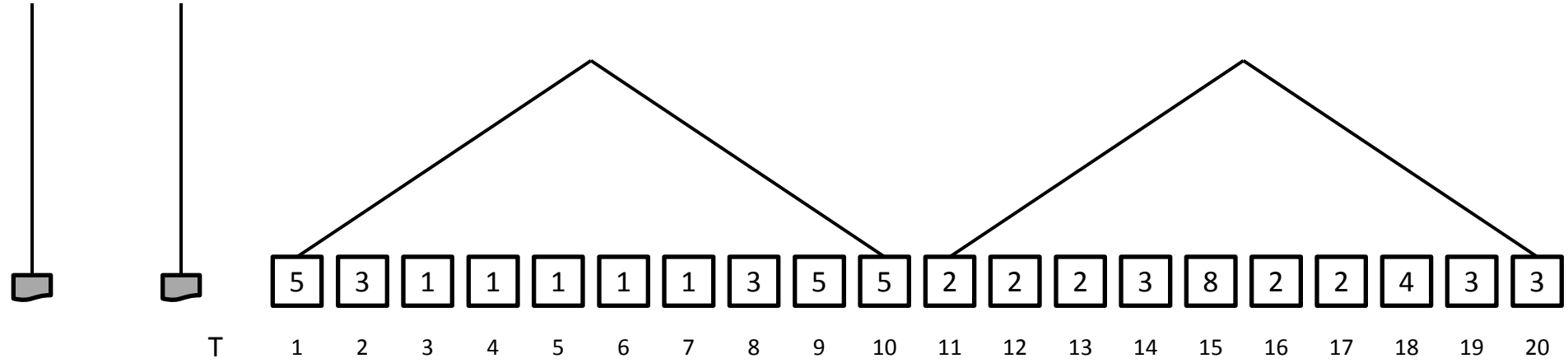
- Split between ‘sensors’ and ‘data’ (spatial/key node and temporal nodes)
- High-frequency data in lowest level temporal nodes, higher levels are averages over time

The TB-tree

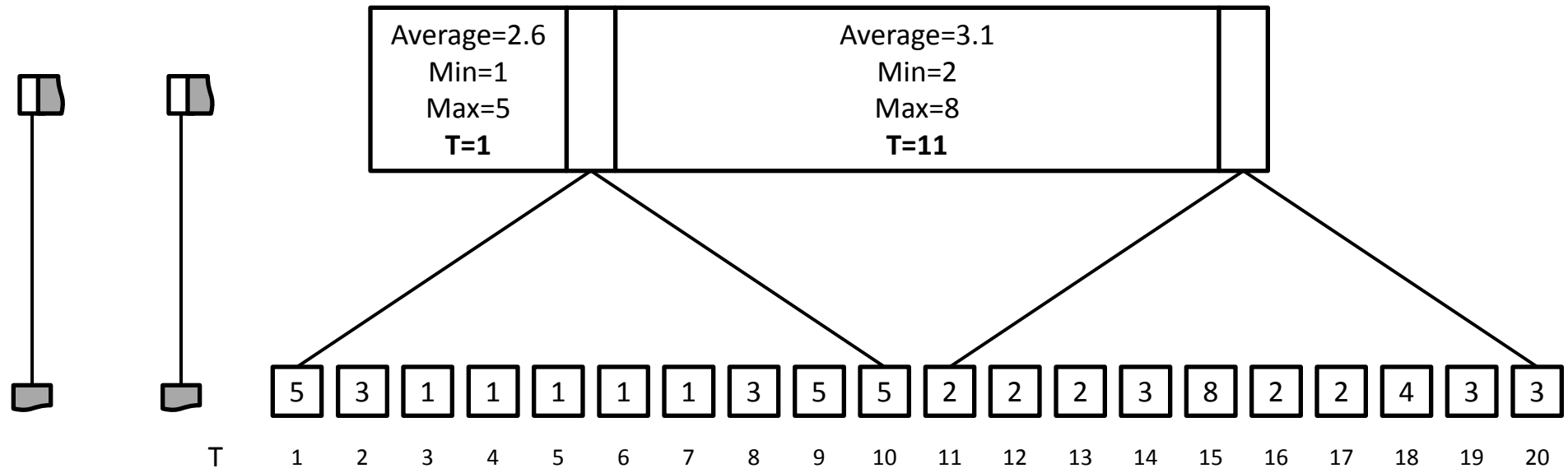
- Blurring (averaging) of data values
- Example data set and process:



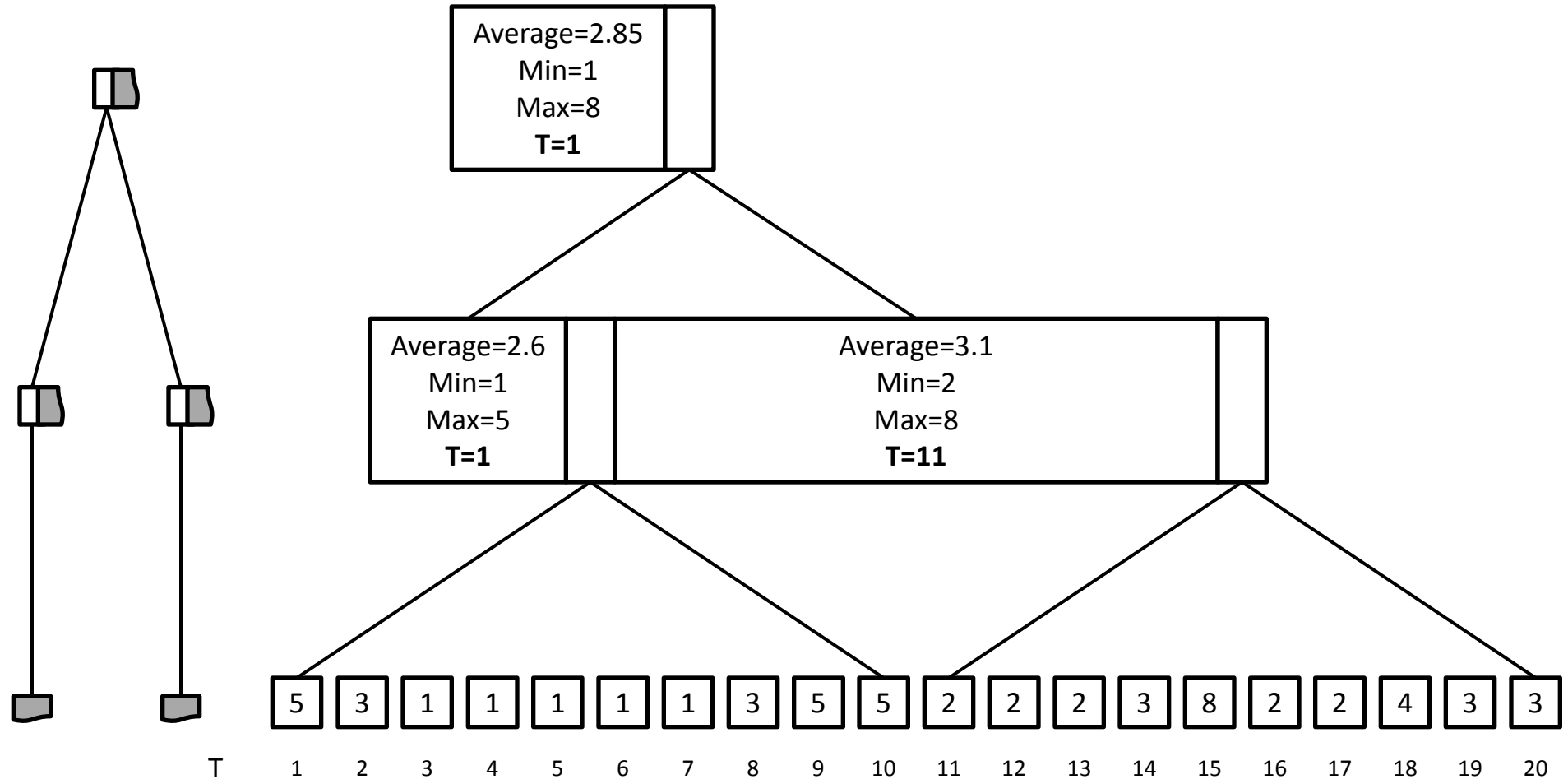
The TB-tree


5 3 1 1 1 1 1 3 5 5 2 2 2 3 8 2 2 4 3 3
T 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

The TB-tree


Average=2.6
Min=1
Max=5
T=1
Average=3.1
Min=2
Max=8
T=11
5 3 1 1 1 1 1 3 5 5 2 2 2 3 8 2 2 4 3 3
T
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

The TB-tree


Average=2.85
Min=1
Max=8
T=1
Average=2.6
Min=1
Max=5
T=1
Average=3.1
Min=2
Max=8
T=11
5 3 1 1 1 1 1 3 5 5 2 2 2 3 8 2 2 4 3 3
T
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

The TB-tree

- Metadata file incorporated for each sensor limits overhead

Data	Value
Metadata	Sensor id, type, accuracy, etc.
Approximate location	Coordinates
Tree identifier	ID value
Measurement frequency	TB-tree level 1 TB-tree level 2 TB-tree level *n*
Calculation results 1	Link to tree identifier
Calculation results 2	Link to tree identifier
Calculation results *n*	Link to tree identifier

The TB-tree

Data management issue	How it is addressed
Different storage systems	Linkage to files and/or database through the indexing scheme
Data at different frequencies	Each sensor has its own branch with its own frequency
Different data representations	Each sensor has its own branch with its own representation
Data storage limitation mitigation	Non-event and/or low blur-level data can be removed, archiving can be performed

The TB-tree

Data management issue	How it is addressed
Post-processing uses multiple sensors	Selection of multiple spatial/key node equals selection of multiple sensors
Post-processing of different lengths of time	The blurring system allows any timespan to be easily processed
Post-processing creates new data	New data is stored in a new spatial/key node
Post-processing requires efficient data access	Sensors are selected through spatial/key nodes, time through the TB-tree and statistical information assists with the sensor data values

Conclusion

- 8 data management issues for sensor networks identified
- TB-tree proposed with a spatial/key and temporal division, in which the temporal division is blurred
- TB-tree offers better long-term and event post-processing through blurring and statistical information
- Flexibility in TB-tree allows easy integration to a multitude of applications

THE UNIVERSITY OF MELBOURNE

A Sensor Network Data Structure for Centralized Post-Processing of Events and Trends

Joost Kuckartz

j.kuckartz@student.unimelb.edu.au

Australian Government
Australian Research Council

UNSW THE UNIVERSITY OF NEW SOUTH WALES

Leica Geosystems

C.R. KENNEDY SURVEY SOLUTIONS

VEKTA

NPO SODIS

TB-tree: blurring levels create extra data?

- Yes, but it allows easier low-level data removal.
- Also, the TB-tree has less overhead compared to traditional databases:

ID	Sensor id	Timestamp	Value
1	1	2013-10-24 12:12:30.1	2
2	1	2013-10-24 12:12:30.2	3
3	1	2013-10-24 12:12:30.3	4
4	1	2013-10-24 12:12:30.4	3
5	1	2013-10-24 12:12:30.5	2
6	1	2013-10-24 12:12:30.6	2

Metafile	
TB-tree level 0 frequency	10 Hz

TB-tree level 1 timestamp
2013-10-24 12:12:30.1

TB-tree level 0 data file
2, 3, 4, 3, 2, 2