Source code topics identifying

Juraj Vincúr supervised by Ing. Ivan Polášek, PhD.

STU FIIT

`our://approach/software_visualization`

Spanning tree construction

ST

- Algorithm inspired by fire ant colony's ability to build **living raft**.
- Use of concurrency leads to optimal trade-off between equally sized and natural clusters.
- Spanning tree based algorithms can handle irregular boundaries.
- Does not require full covariance matrix.

VSM

Pre-processing

- Incremental model based on global weights proposed.
- First adaptation of Random Indexing in source code domain.

Cluster analysis

- Implementation of new clustering algorithm with linear complexity which requires only spanning tree close to minimum.
- 3 layer hierarchy with parametric level of abstraction.
- Precise technique for medoid approximation (cluster-point) proposed.
- Cluster-point based labeling and search space reduction.
- Aspect mining technique based on packages and identified clusters proposed.

JSON

Visualization

- Lightweight, no installation required.
- Tested on 10k + classes and over 800 clusters.
- Especially suitable for integration with version control systems like Git.
- Adaptation of intuitive and space efficient bubble layout.

- Clusters
- Cluster-points
- Classes

Publications — Problems — **EXAMPLE** — Results
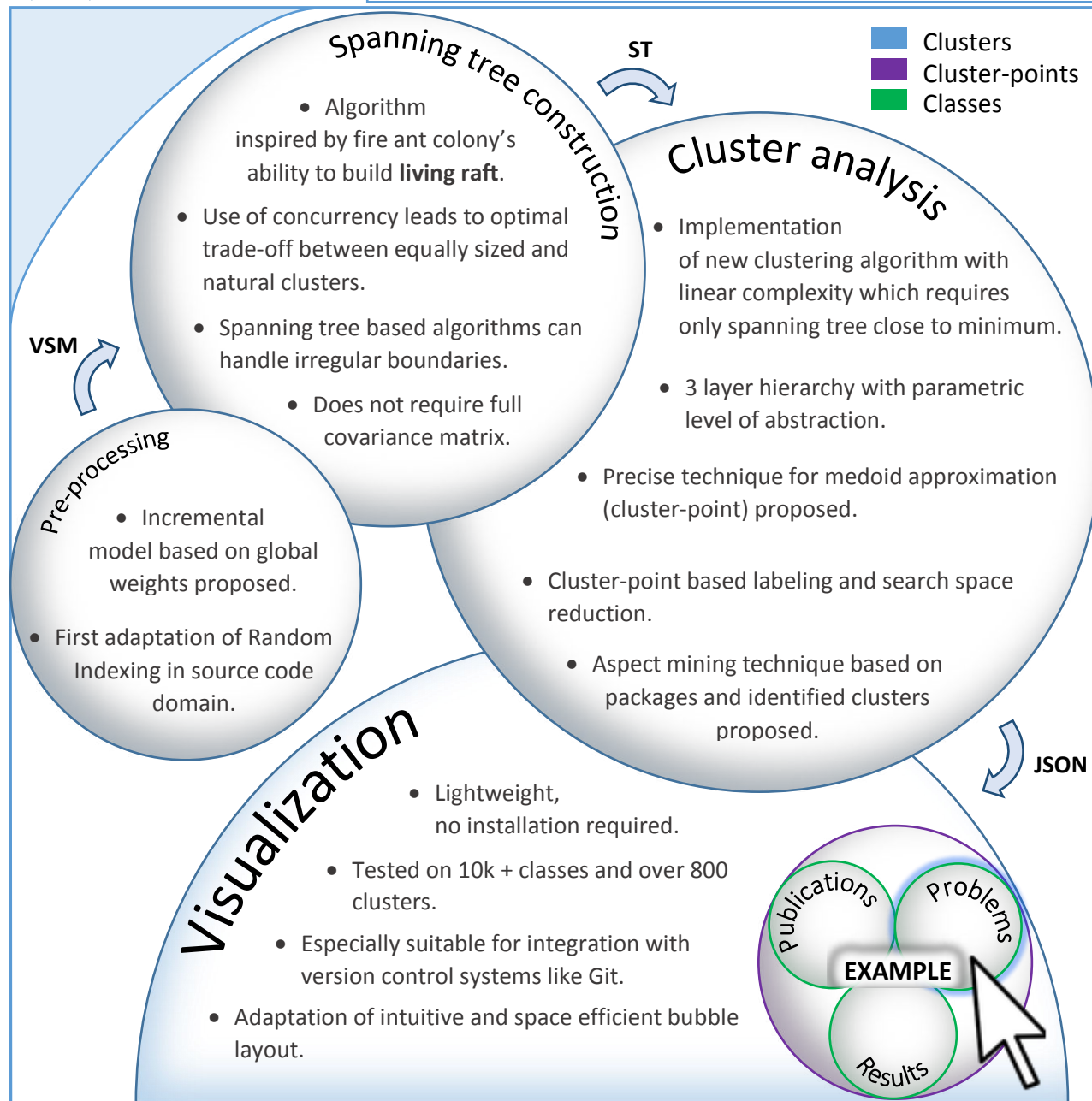
```
class Problems {

    • Understanding of large and complex software system
      requires huge effort.

    • Significant amount of developer knowledge is hidden
      in the code naming.

}

class Results {
```

- Strong correlation between covariance matrices obtained from reduced and non-reduced space: 0.66
- Proposed incremental approach for model construction yield significant performance improvement.

- Medoid approximation average error: 0.02 – 0.05
- Number of identified clusters scales linearly with number of extracted classes.
- Based on relevance judgments natural and relatively equally sized clusters were identified.
- Proposed aspect mining technique combines advantages of identifier and fan-in analysis.

```
}

class Publications {
```

- Paper "Software visualization using clustering inspired by ant colony" accepted by IFIP AITP 2015.

```
}
```