WatDiv: How to Tune-Up Your RDF Data Management System

Güneş Aluç    Olaf Hartig    M. Tamer Ö兹su    Khuzaima Daudjee

This presentation is sponsored in part by the Linked Data Benchmark Council (LDBC).
Questions

• Which of the existing SPARQL benchmarks, if any, should I use to diagnose (and fix) potential problems with the physical design of my system?

• How can I use the Waterloo SPARQL Diversity Test Suite (WatDiv) where existing benchmarks fall short?
Contributions

Waterloo SPARQL Diversity Test Suite
(WatDiv)

http://db.uwaterloo.ca/watdiv/
Contributions

Waterloo SPARQL Diversity Test Suite (WatDiv)

http://db.uwaterloo.ca/watdiv/

Measures to Evaluate Diversity in SPARQL Workloads

- Structural
- Data-driven
Contributions

Waterloo SPARQL Diversity Test Suite (WatDiv)

http://db.uwaterloo.ca/watdiv/

Measures to Evaluate Diversity in SPARQL Workloads
- Structural
- Data-driven

Analysis of WatDiv and Popular SPARQL Benchmarks
Contributions

Waterloo SPARQL Diversity Test Suite (WatDiv)

Measures to Evaluate Diversity in SPARQL Workloads
- Structural
- Data-driven

Analysis of WatDiv and Popular SPARQL Benchmarks

Debugging with WatDiv

http://db.uwaterloo.ca/watdiv/
Structural Features
Structural Features
Structural Features
Structural Features
[ Triple Pattern Count ]

Diagram:

- Nodes labeled with numbers: 1, 2, 3, 4, 5, 6, 7
- Connections labeled with numbers: 1, 2, 3, 4, 5, 6, 7
- Pattern count:
  - Triple Pattern Count: 1
  - Pattern 2
  - Pattern 3
  - Pattern 4
  - Pattern 5
  - Pattern 6
  - Pattern 7
Structural Features
[ Join Vertex Count ]
Structural Features
[ Join Vertex Degree ]
Structural Features
[ Join Vertex Degree ]

Diagram showing vertex degrees and connections.
Structural Features
[ Join Vertex Degree ]

Diagram showing a network with labeled vertices and edges.
Structural Features

<table>
<thead>
<tr>
<th>Join Vertex Count</th>
<th>Mean Join Vertex Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>2.0</td>
</tr>
</tbody>
</table>
Structural Features

<table>
<thead>
<tr>
<th>Join Vertex Count</th>
<th>Mean Join Vertex Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>2.0</td>
</tr>
<tr>
<td>5</td>
<td>2.6</td>
</tr>
</tbody>
</table>
Structural Features

Join Vertex Count | Mean Join Vertex Degree
--- | ---
8 | 2.0
5 | 2.6
3 | ~3.7
Structural Features

<table>
<thead>
<tr>
<th>Join Vertex Count</th>
<th>Mean Join Vertex Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>2.0</td>
</tr>
<tr>
<td>5</td>
<td>2.6</td>
</tr>
<tr>
<td>3</td>
<td>~3.7</td>
</tr>
<tr>
<td>2</td>
<td>5.0</td>
</tr>
</tbody>
</table>
Structural Features

<table>
<thead>
<tr>
<th>Join Vertex Count</th>
<th>Mean Join Vertex Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>2.0</td>
</tr>
<tr>
<td>5</td>
<td>2.6</td>
</tr>
<tr>
<td>3</td>
<td>~3.7</td>
</tr>
<tr>
<td>2</td>
<td>5.0</td>
</tr>
<tr>
<td>1</td>
<td>9.0</td>
</tr>
</tbody>
</table>
Structural Features
[ Join Vertex Type ]

SS$^+$ Type
Structural Features

[ Join Vertex Type ]

OO$^+$ Type
Structural Features
[ Join Vertex Type ]

SO⁺ Type
How Diverse are SPARQL Benchmarks?

[ Triple Pattern Count ]
How Diverse are SPARQL Benchmarks?
[ Triple Pattern Count ]
How Diverse are SPARQL Benchmarks?

[ Join Vertex Count ]

- WatDiv
- DBSB
- BSBM
- SP2Bench
- LUBM

* * * * * (WatDiv)
* * * * * (DBSB)
* * * * * (BSBM)
* * * (SP2Bench)
* (LUBM)

0 5 10
How Diverse are SPARQL Benchmarks?
[ Join Vertex Count ]
How Diverse are SPARQL Benchmarks?

[ Join Vertex Degree – mean ]

- WatDiv
- DBSB
- BSBM
- SP2Bench
- LUBM
How Diverse are SPARQL Benchmarks?

[ Join Vertex Degree – mean ]
How Diverse are SPARQL Benchmarks?
[ Join Vertex Type – % Queries w/in Workload ]

<table>
<thead>
<tr>
<th></th>
<th>SS⁺</th>
<th>OO⁺</th>
<th>SO⁺</th>
</tr>
</thead>
<tbody>
<tr>
<td>LUBM</td>
<td>78.6 %</td>
<td>0.0 %</td>
<td>42.9 %</td>
</tr>
<tr>
<td>SP²Bench</td>
<td>81.0 %</td>
<td>33.3 %</td>
<td>57.1 %</td>
</tr>
<tr>
<td>BSBM</td>
<td>84.8 %</td>
<td>5.6 %</td>
<td>52.8 %</td>
</tr>
<tr>
<td>DBSB</td>
<td>41.1 %</td>
<td>4.4 %</td>
<td>5.4 %</td>
</tr>
<tr>
<td>WatDiv</td>
<td>61.3 %</td>
<td>18.0 %</td>
<td>61.3 %</td>
</tr>
</tbody>
</table>
How Diverse are SPARQL Benchmarks?
[ Join Vertex Type – % Queries w/in Workload ]

<table>
<thead>
<tr>
<th></th>
<th>SS⁺</th>
<th>OO⁺</th>
<th>SO⁺</th>
</tr>
</thead>
<tbody>
<tr>
<td>LUBM</td>
<td>78.6 %</td>
<td>0.0 %</td>
<td>42.9 %</td>
</tr>
<tr>
<td>SP²Bench</td>
<td>81.0 %</td>
<td>33.3 %</td>
<td>57.1 %</td>
</tr>
<tr>
<td>BSBM</td>
<td>84.8 %</td>
<td>5.6 %</td>
<td>52.8 %</td>
</tr>
<tr>
<td>DBSB</td>
<td>41.1 %</td>
<td>4.4 %</td>
<td>5.4 %</td>
</tr>
<tr>
<td>WatDiv</td>
<td>61.3 %</td>
<td>18.0 %</td>
<td>61.3 %</td>
</tr>
</tbody>
</table>
Data-Driven Features

• Why are data-driven query features important?
  – Why are structural features not sufficient?
  – Why is analysis based purely on the data not sufficient?
Data-Driven Features

[ Result Cardinality ]

Diagram:

Graph with nodes labeled a, b, c, d, e, f, g, and edges connecting them.

Table:

<table>
<thead>
<tr>
<th>?a</th>
<th>...</th>
<th>?g</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The table is incomplete, with only 1 and 2 entries provided.
Data-Driven Features
[ Filtered Triple Pattern (f-TP) Selectivity ]

Diagram:

<table>
<thead>
<tr>
<th>P</th>
<th>S</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;9..</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;A&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;A&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;A&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;A&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;B..</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data-Driven Features
[ Filtered Triple Pattern (f-TP) Selectivity ]

\[ f\text{-TP Selectivity} = \frac{k}{n} \]
Data-Driven Features
[ f-TP Selectivity, BGP-Restricted ]

```
<table>
<thead>
<tr>
<th>P</th>
<th>S</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;9..</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;A&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;A&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;A&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;A&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;B..</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```
Data-Driven Features
[ f-TP Selectivity, BGP-Restricted ]

BGP-Restricted
f-TP Selectivity = |blue| / |orange|

<table>
<thead>
<tr>
<th>P</th>
<th>S</th>
<th>O</th>
<th>?a</th>
<th>...</th>
<th>?g</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>9..</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td>compatible</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td>compatible</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B..</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data-Driven Features
[ f-TP Selectivity, Join-Restricted ]
Data-Driven Features
[ f-TP Selectivity, Join-Restricted ]

Join-Restricted
f-TP Selectivity = |blue| / |orange|

<table>
<thead>
<tr>
<th>P</th>
<th>S</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;9..</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;A&gt;</td>
<td>compatible</td>
<td></td>
</tr>
<tr>
<td>&lt;A&gt;</td>
<td>compatible</td>
<td></td>
</tr>
<tr>
<td>&lt;A&gt;</td>
<td>compatible</td>
<td></td>
</tr>
<tr>
<td>&lt;A&gt;</td>
<td>compatible</td>
<td></td>
</tr>
<tr>
<td>&lt;B..</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
How Diverse are SPARQL Benchmarks?

[ Result Cardinality ]

WatDiv
DBSB
BSBM
SP2Bench
LUBM
How Diverse are SPARQL Benchmarks?
[ Result Cardinality ]
How Diverse are SPARQL Benchmarks?
[ f-TP Selectivity – mean ]

- WatDiv
- DBSB
- BSBM
- SP2Bench
- LUBM
How Diverse are SPARQL Benchmarks?

[f-TP Selectivity – mean]
How Diverse are SPARQL Benchmarks?

[ f-TP Selectivity – \textit{stdev} ]
How Diverse are SPARQL Benchmarks?

[f-TP Selectivity – stdev]
How Diverse are SPARQL Benchmarks?

[ BGP-Restricted f-TP Selectivity – mean ]
How Diverse are SPARQL Benchmarks?

[ BGP-Restricted f-TP Selectivity – mean ]

![Diagram showing selectivity distribution for different benchmarks]
How Diverse are SPARQL Benchmarks?

[BGP-Restricted f-TP Selectivity – stdev]
How Diverse are SPARQL Benchmarks?

[BGP-Restricted f-TP Selectivity – stdev]
How Diverse are SPARQL Benchmarks?

[ Join-Restricted f-TP Selectivity – mean ]
How Diverse are SPARQL Benchmarks?

[ Join-Restricted f-TP Selectivity – mean ]
How Diverse are SPARQL Benchmarks?

- WatDiv
- DBSB
- BSBM
- SP$^2$Bench
- LUBM
- Other SPARQL benchmarks
- Production workload
- ...

---

25
WatDiv Tools

- Data Generator
  - Customizable data description model
- Query Template Generator
- Query Instantiator
WatDiv Dataset

Entities generated according to the *default* data description model

<table>
<thead>
<tr>
<th>Entity</th>
<th>No. of Instances (per scale factor)</th>
</tr>
</thead>
<tbody>
<tr>
<td>wsdbm:Purchase</td>
<td>1500</td>
</tr>
<tr>
<td>wsdbm:User</td>
<td>1000</td>
</tr>
<tr>
<td>wsdbm:Offer</td>
<td>900</td>
</tr>
<tr>
<td>wsdbm:Product</td>
<td>250</td>
</tr>
<tr>
<td>wsdbm:Website</td>
<td>50</td>
</tr>
<tr>
<td>wsdbm:Retailer</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Entity</th>
<th>No. of Instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>wsdbm:Topic</td>
<td>250</td>
</tr>
<tr>
<td>wsdbm:City</td>
<td>240</td>
</tr>
<tr>
<td>wsdbm:SubGenre</td>
<td>145</td>
</tr>
<tr>
<td>wsdbm:Language</td>
<td>25</td>
</tr>
<tr>
<td>wsdbm:Country</td>
<td>25</td>
</tr>
<tr>
<td>wsdbm:Genre</td>
<td>21</td>
</tr>
<tr>
<td>wsdbm:ProductCategory</td>
<td>15</td>
</tr>
<tr>
<td>wsdbm:AgeGroup</td>
<td>9</td>
</tr>
<tr>
<td>wsdbm:Role</td>
<td>3</td>
</tr>
<tr>
<td>wsdbm:Gender</td>
<td>2</td>
</tr>
</tbody>
</table>

*The entities above do not scale.*
WatDiv Dataset

Characteristics of the dataset at scale-factor=1 (default model)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Triples</td>
<td>105257</td>
</tr>
<tr>
<td>Distinct subjects</td>
<td>5597</td>
</tr>
<tr>
<td>Distinct predicates</td>
<td>85</td>
</tr>
<tr>
<td>Distinct objects</td>
<td>13258</td>
</tr>
<tr>
<td>URIs</td>
<td>5947</td>
</tr>
<tr>
<td>Literals</td>
<td>14286</td>
</tr>
<tr>
<td>Distinct literals</td>
<td>8018</td>
</tr>
</tbody>
</table>

*Numbers are approximate and may vary slightly in each dataset generation*
WatDiv Dataset

[ What Sets It Apart ]

(1) Heavily Relies on Optional Attributes
WatDiv Dataset
[ What Sets It Apart ]

(1) Heavily Relies on Optional Attributes
WatDiv Dataset
[ What Sets It Apart ]

(1) Heavily Relies on Optional Attributes
WatDiv Dataset
[ What Sets It Apart ]

(1) Heavily Relies on Optional Attributes
WatDiv Dataset
[ What Sets It Apart ]

(2) Parts of the database are well-structured
(2) Parts of the database are well-structured while remaining parts are less well-structured
WatDiv Dataset
[ What Sets It Apart ]

// Attributes for wsdbm:ProductCategory4 (i.e., NewsArticle)

... 

<pgroup> 0.8 @wsdbm:ProductCategory4
    #predicate  sorg:publisher    string
</pgroup>

<pgroup> 0.7 @wsdbm:ProductCategory4
    #predicate  sorg:datePublished date
</pgroup>

<pgroup> 0.2 @wsdbm:ProductCategory4
    #predicate  sorg:printPage    integer 1 999
    #predicate  sorg:printSection integer 1 9
</pgroup>
WatDiv Dataset
[ What Sets It Apart ]

- // Associations for wsdbm:ProductCategory2 (i.e., Movie)

- #association
  wsdbm:Product sorg:actor wsdbm:User
  2 25[normal] 0.8 UNIFORM
  @wsdbm:ProductCategory2 @wsdbm:Role2

- #association
  wsdbm:Product sorg:director wsdbm:User
  2 1 0.8 ZIPFIAN
  @wsdbm:ProductCategory2 @wsdbm:Role2

- #association
  wsdbm:Product sorg:trailer wsdbm:Website
  2 3[uniform] 0.1 UNIFORM
  @wsdbm:ProductCategory2 @null
How Robust are Systems across WatDiv Workloads?

WatDiv 100M triples, queries w/ single join vertex, result cardinality ≤ 2000
How Robust are Systems across WatDiv Workloads?

**WatDiv 10M triples**

- **linear** = \{ \text{mean join vertex degree} \leq 3.0, \text{join vertex count} \geq 3 \}
- **star/snowflake** = \{ \text{mean join vertex degree} \geq 5.0, \text{join vertex count} \leq 2 \}
Conclusions

- Which of the existing SPARQL benchmarks, if any, should I use to diagnose (and fix) potential problems with the physical design of my system?
  - *Analyze your production workload and find the best-matching benchmarks*

- How can I use the Waterloo SPARQL Diversity Test Suite (WatDiv) where existing benchmarks fall short?
  - *Drill down into different classes of queries until you hit problematic spots*
Questions

Waterloo SPARQL Diversity Test Suite
(WatDiv)

http://db.uwaterloo.ca/watdiv/