Streaming Queries over
Streaming Data

- Discussion -

OUTLINE

- Contributions
- Limitations
- Analysis
- Presentation
Ideas in PSoup

- Streaming Data (subject of research – especially sensor data)
- Streaming Queries (NiagaraCQ, CACQ, …)
- Queries on old and new data using materialization (Wave indices)
- Intermittent Connectivity (delivery on demand – is it a big contribution?)

Limitations

- PSoup builds on CACQ but lacks some of its features (e.g. Disjunctive predicates)
- Aggregation queries (data structures are only shared if we have the same SELECT-PROJECT-JOIN clause for different queries)
- MAIN MEMORY Implementation: Maximum window size for data is limited.
Analysis

• Most of the analysis is to validate implementation choices.
• Most comparison of results involve variations of PSoup only.
• Is data generated and used for analysis representative of streaming data?
• If data streams can be archived on disk, how would it change the analysis?

Presentation

• The authors do not sufficiently motivate the general problem of streaming data and continuous queries
• Very little details on previous related work
• Too many implementation details
• Some figures are difficult to read (see Figure 11) + names (e.g. pSoup-C_UnShrd)
Conclusion

• The authors have the credit to provide an efficient and flexible system
• I look forward to future work. E.g. “Temporal Browser”

• What are your critiques?

Thank you.