

## **Improving Query Processing on Imprecise Data Streams**

*Eugenia Gabrielova, Senior, Computer Science, Northwestern University*

*Mentor: Magdalena Balazinska, CSE*

*Mentor: Julie Letchner, CSE*

Many applications, such as monitored health care and theft detection, depend on higher-level data inferred from low-level location-dating sensors such as RFID and GPS. This high level data occurs in the form of imprecise, correlated sequences, which are modeled by Markovian streams. Such data is too difficult to manage with traditional databases. Lahar is a system that warehouses and processes queries on such streams, returning a set of query answers annotated with probabilities. Some queries return many partial results, which wastes computing resources. Processing streams and queries in a reversed direction may result in fewer partial matches. In this poster, I present an application developed to reverse Markovian streams and queries. This application also compares the efficiency of processing a query on forward and backward streams. Some properties of queries, such as a rare element at the end of a query, may make backward processing a more efficient choice. The ability to reverse and process Markovian streams backward to process such queries improves the efficiency of the Lahar system.