

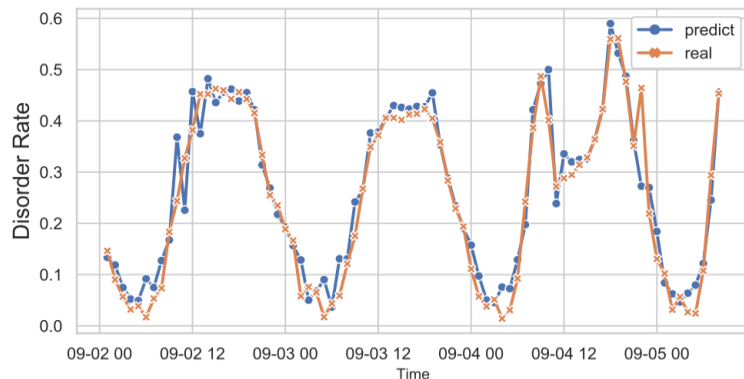
Adaptive watermark generation
mechanism based on time series
prediction for stream processing

Yang SONG, YunChun LI, Hailong YANG, Jun XU, Zerong
LUAN, Wei LI

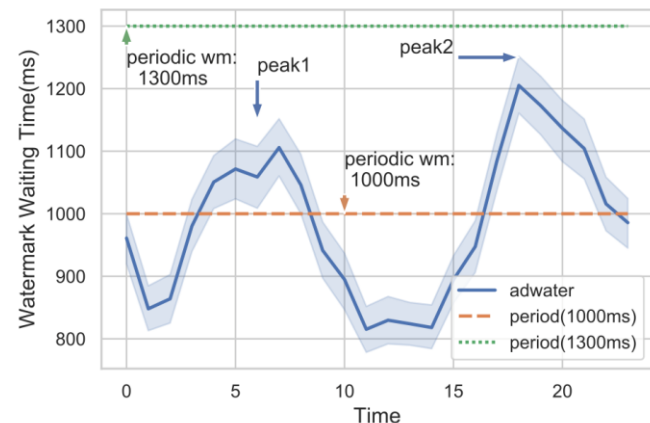
Frontiers of Computer Science, DOI: [10.1007/s11704-020-0206-7](https://doi.org/10.1007/s11704-020-0206-7)

Problems & Ideas

- Problems of the traditional stream-processing watermark distribution strategies
 - Those strategies fail to consider the characteristics of the stream data and cannot adjust to data changes dynamically, which could unnecessarily increase the system response time.
- Ideas: Use predictive models to adjust watermark distribution
 - Use LSTM to predict the degree of data disorder.
 - Use the prediction results of the prediction model to adjust the watermark distribution adaptively.

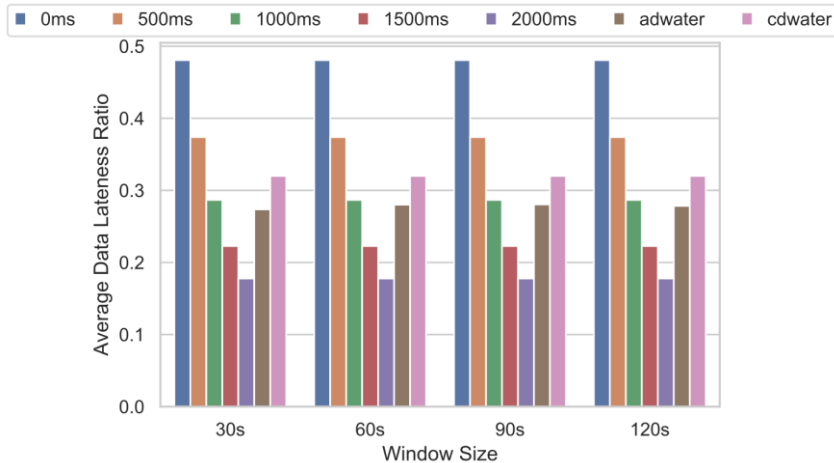


LSTM prediction results

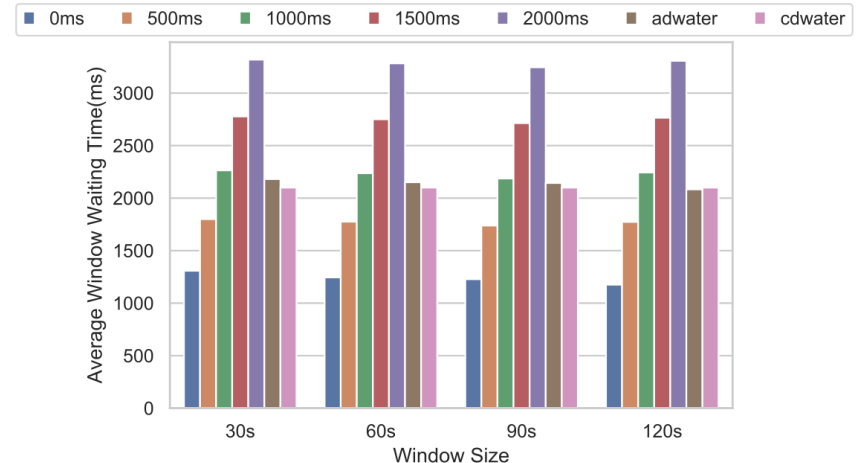


The change of watermark distribution with the adaptive watermark

Main Contributions



Comparison of the lateness ratio using different watermark distribution strategies



Comparison of the average window waiting time using different watermark distribution strategies

- The lateness ratio represents calculation error. The window waiting time represents responding speed.
- The adaptive watermark strategy(adwater) can keep the overall lateness ratio of the data near the set threshold and can ensure a lower window waiting time.