

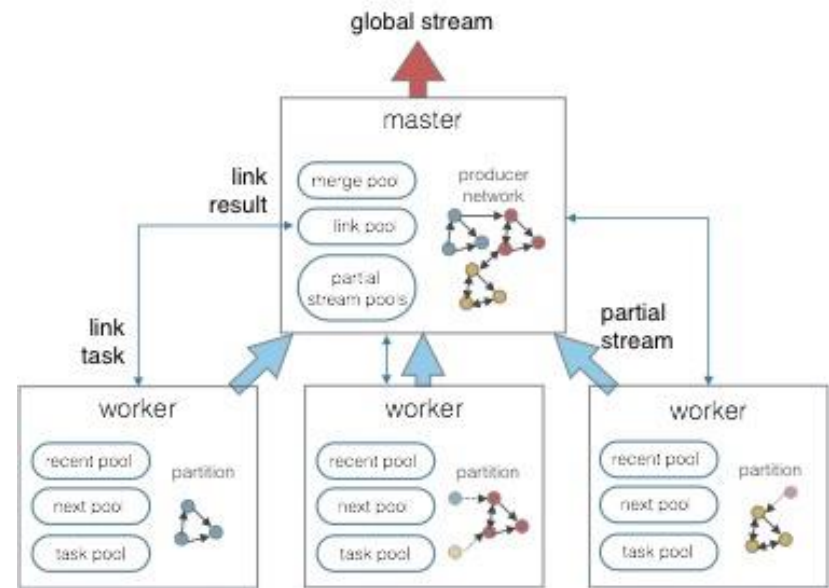
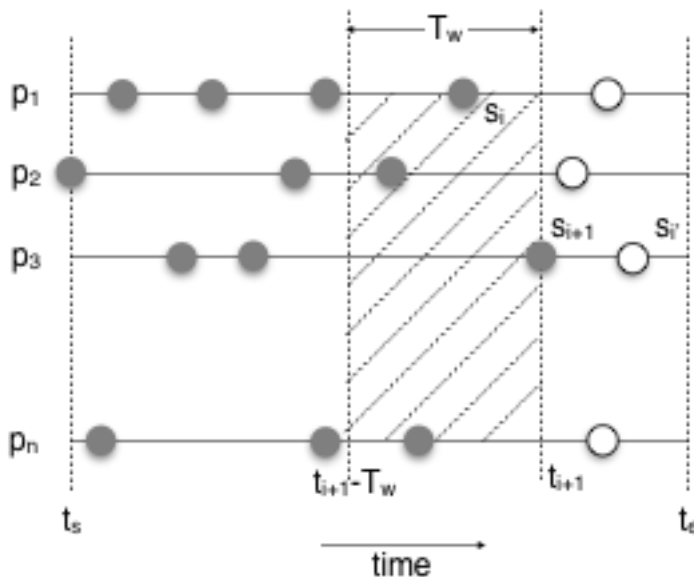
Parallel data generator for efficiently generating “realistic” social streams

Chengcheng YU, Fan XIA, Weining QIAN, Aoying ZHOU

Frontiers of Computer Science, DOI: [10.1007/s11704-018-8022-z](https://doi.org/10.1007/s11704-018-8022-z)

Problems & Ideas

- Problems: A social stream generator for *flexibly* and *efficiently* generating “realistic” synthetic social stream data.
 - The generated data should capture “realistic” social stream properties.
 - The social stream generator system with high generation throughput.
- Ideas: A parallel social stream generator with master-worker architecture.



Social stream generated framework of workers

The architecture of the parallel social stream generator

Main Contributions

- A flexible framework, which can be tuned to capture “realistic” social stream properties, including fundamental patterns of human dynamic and network.
- A parallel social stream generator using the master-worker architecture is implemented to achieve high generation throughput.
- Intensive experiments show that our data generator can generate “realistic” data compared to other data generators *flexibly* and *efficiently*.

