Towards Chip-on-Chip Neuroscience: Fast Mining of Neuronal Spike Streams Using Graphics Hardware

Y. Cao, D. Patnaik, S. Ponce, **W. Feng**, and N. Ramakrishnan

VirginiaTech
Invent the Future

SyNeRG
synergy.cs.vt.edu

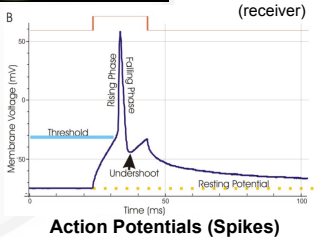
Motivation

GRAND CHALLENGES FOR ENGINEERING

NATIONAL ACADEMY OF ENGINEERING

- Reverse-Engineer the Brain

Question:
How are the neurons connected?



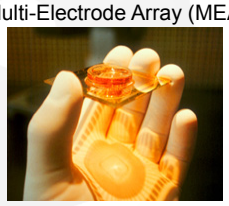
Action Potentials (Spikes)

VirginiaTech
Invent the Future

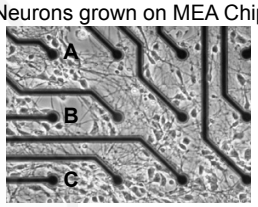
SyNeRG
synergy.cs.vt.edu

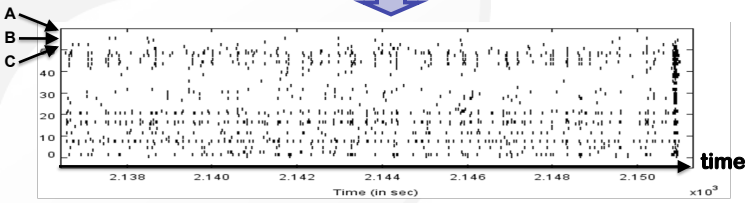
Motivation

Multi-Electrode Array (MEA)



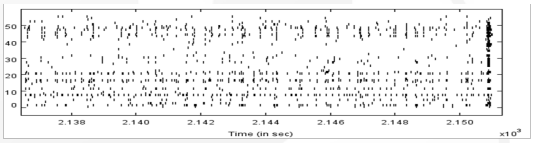
Neurons grown on MEA Chip



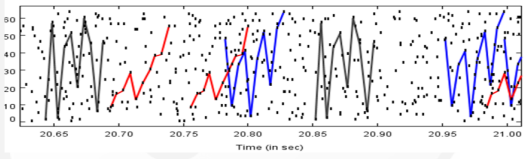


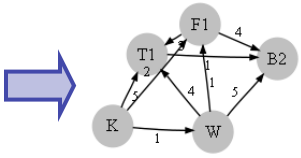
Spike Train Stream

Motivation



Find Repeating Patterns



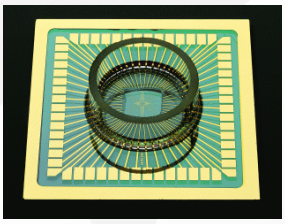


Infer Network Connectivity

Contributions

- Fast data mining of spike train stream on a graphics processing unit (GPU)

MEA Chip → **GPU Chip**



Multi-Electrode Array (MEA)

NVIDIA GTX 280 Graphics Card

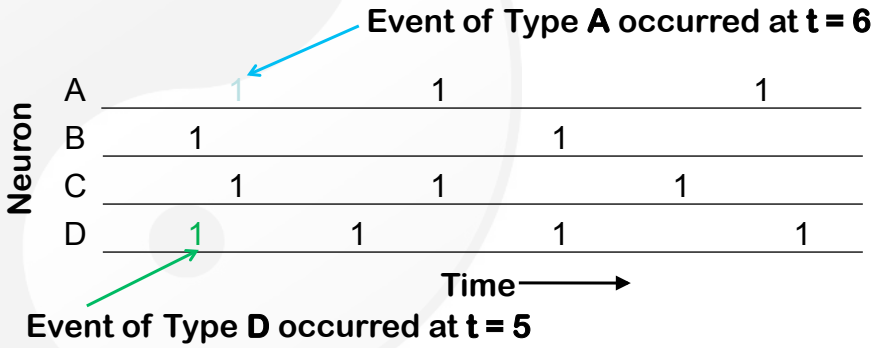
Contributions

- Fast data mining of spike train stream on a graphics processing unit (GPU)
- Two key algorithmic strategies to address scalability problem on GPU
  - A hybrid mining approach
  - A two-pass elimination approach

Background

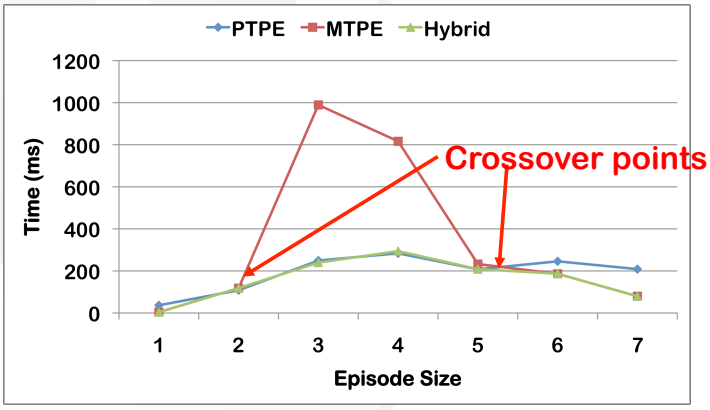
- Event stream data: sequence of neurons firing

$$\langle (E_1,t_1),(E_2,t_2),\ldots,(E_n,t_n) \rangle$$



Results

- Performance of the hybrid approach

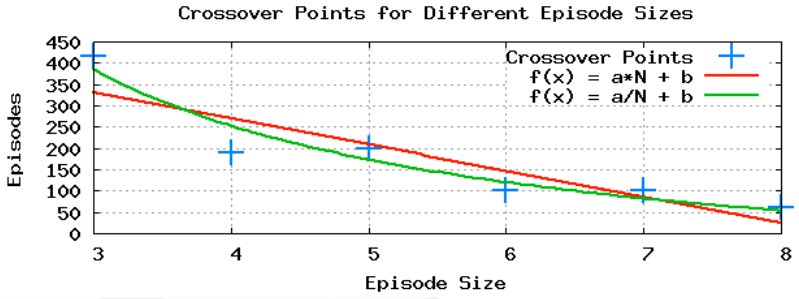


Episode Number:	26	650	4075	1288	228	63	3

Sym26 dataset, Support = 100

Results

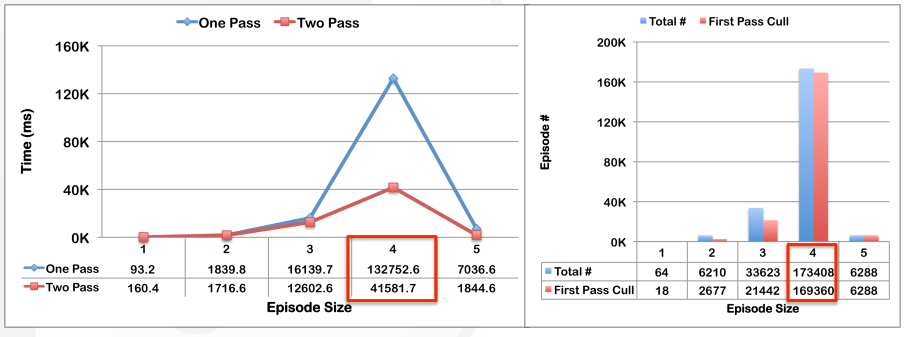
- Crossover point estimation



- $f(size) = \frac{a}{size} + b$ is a better fit.
- A least square fit is performed.

Results

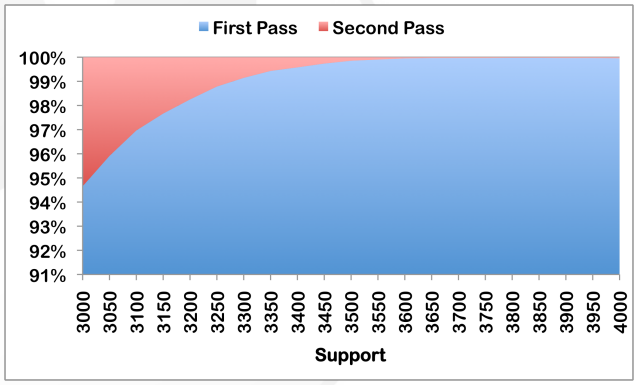
- Performance of the two-pass approach



2-1-35 **dataset, Support =** **3150**

Results

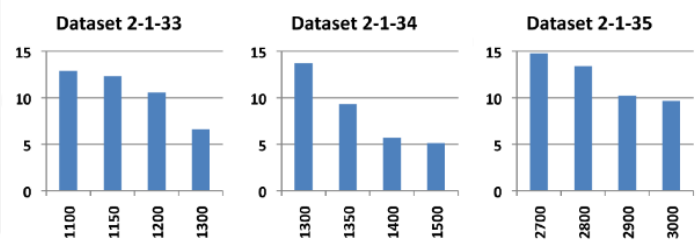
- Percentage of episodes eliminated by each pass



2-1-35 dataset, episode size = 4

Results

- GPU vs CPU



- GPU is always faster than CPU
  - 5x - 15x speedup
  - Fair comparison
    - Two-pass algorithm used
    - Maximum threading for both

Conclusion

- A fast temporal data mining framework on GPUs
  - Commoditized system
  - Massive parallel execution architecture
- Two programming strategies
  - A hybrid approach
    - Increase level of parallelism (data segmentation + map-reduce)
  - Two-pass elimination approach
    - Decrease algorithm complexity (Task decomposition)