



NEW YORK UNIVERSITY

Final Project Presentation

Department of Electrical and Computer Engineering
Tandon School of Engineering, New York University

Present By: Haoze He
NYUID: hh2537



Agenda

01 / Executive Summary

02 / Approach

03 / Main Results

04 / Observation & Conclusion

05 / GitHub Link

Executive Summary

Asynchronized Stochastic Gradient Descent

Advantages

High accuracy

Disadvantages

Waste time and lead to slow converge

Synchronous Distributed Stochastic Gradient Descent

Advantages


Time saving

Disadvantages

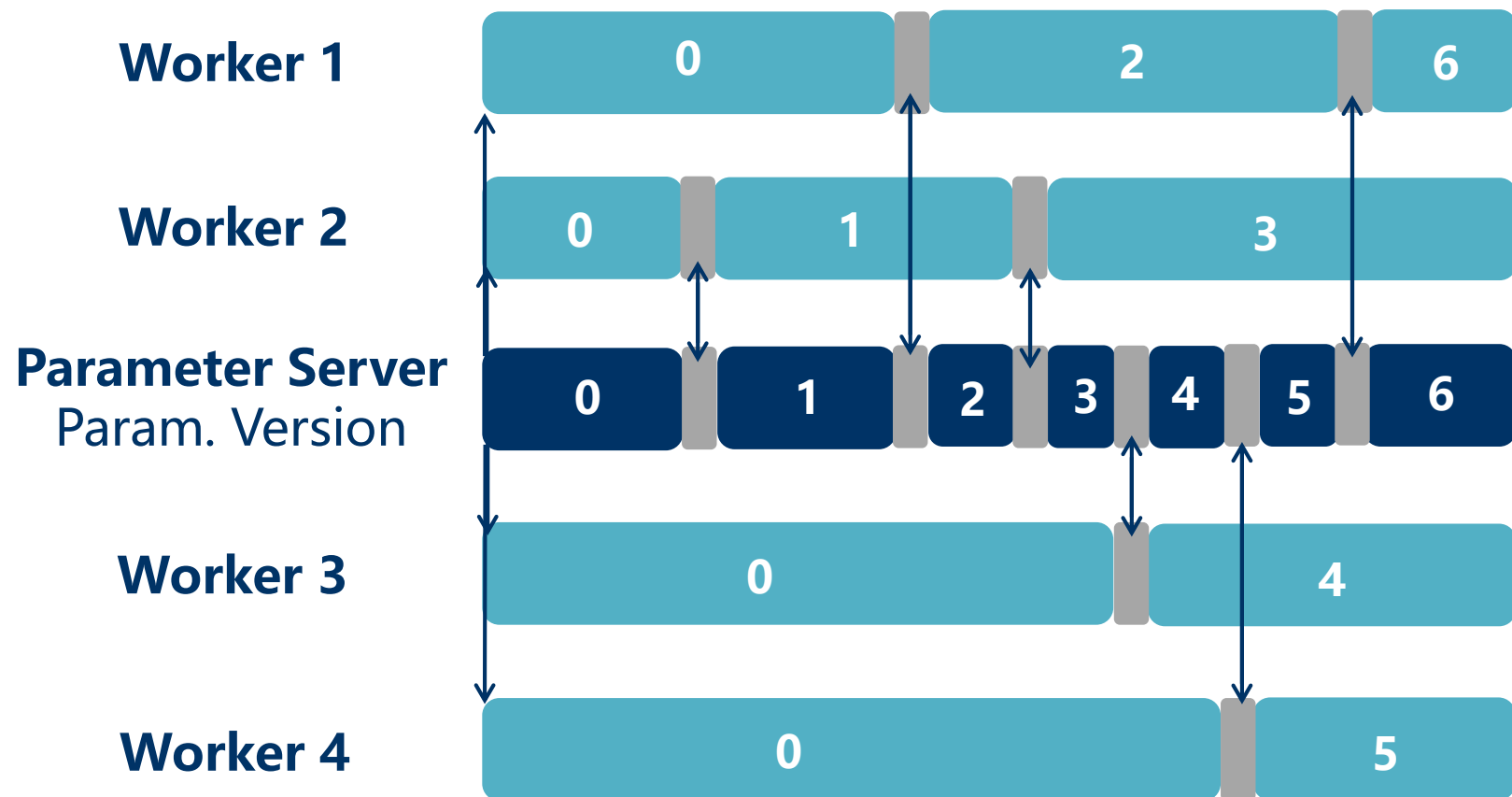
Lower accuracy


Approach A – Synchronize SGD



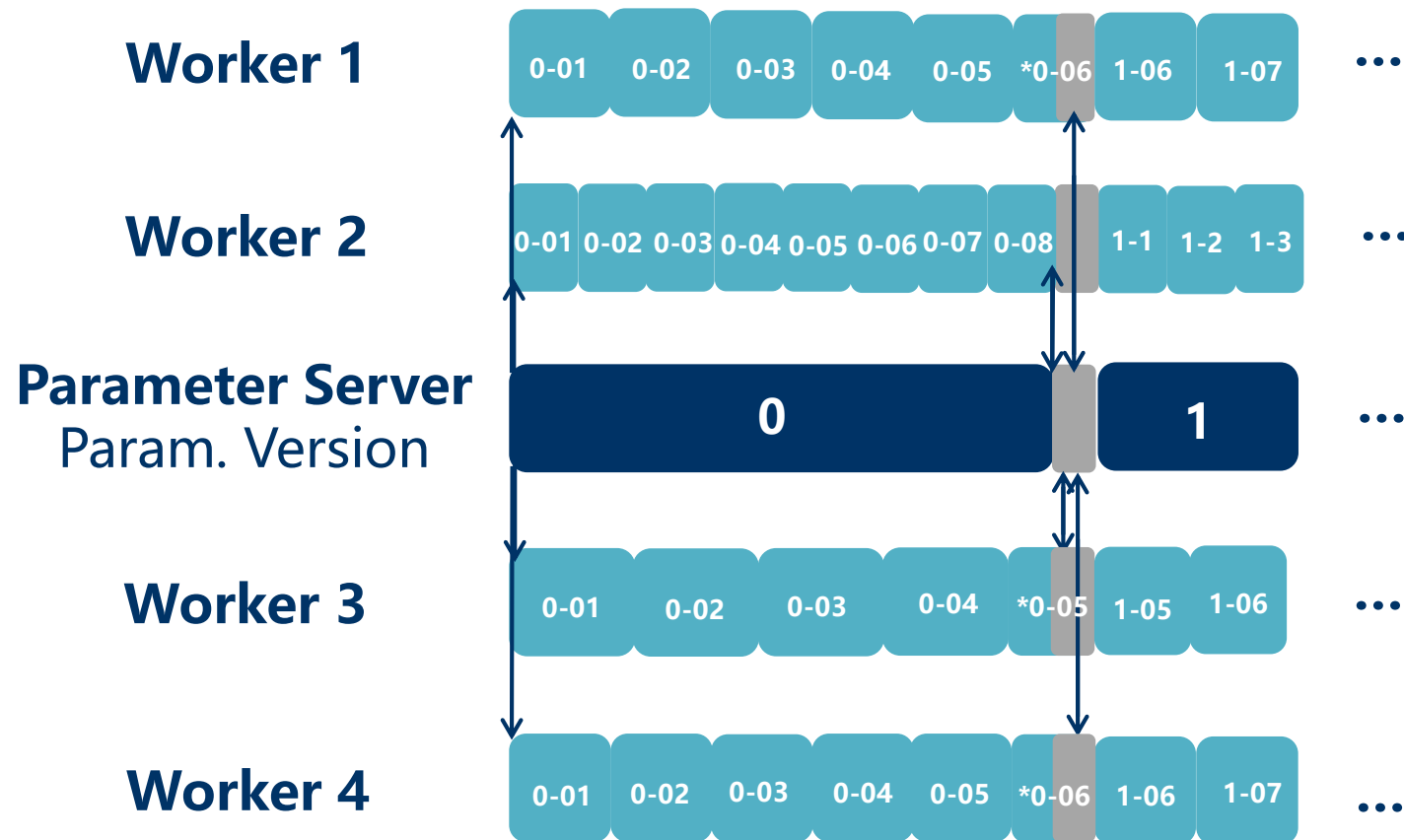
- Batch size = 256
- Cons: Waste time
-  : Communication time



Approach B – Asynchronized SGD



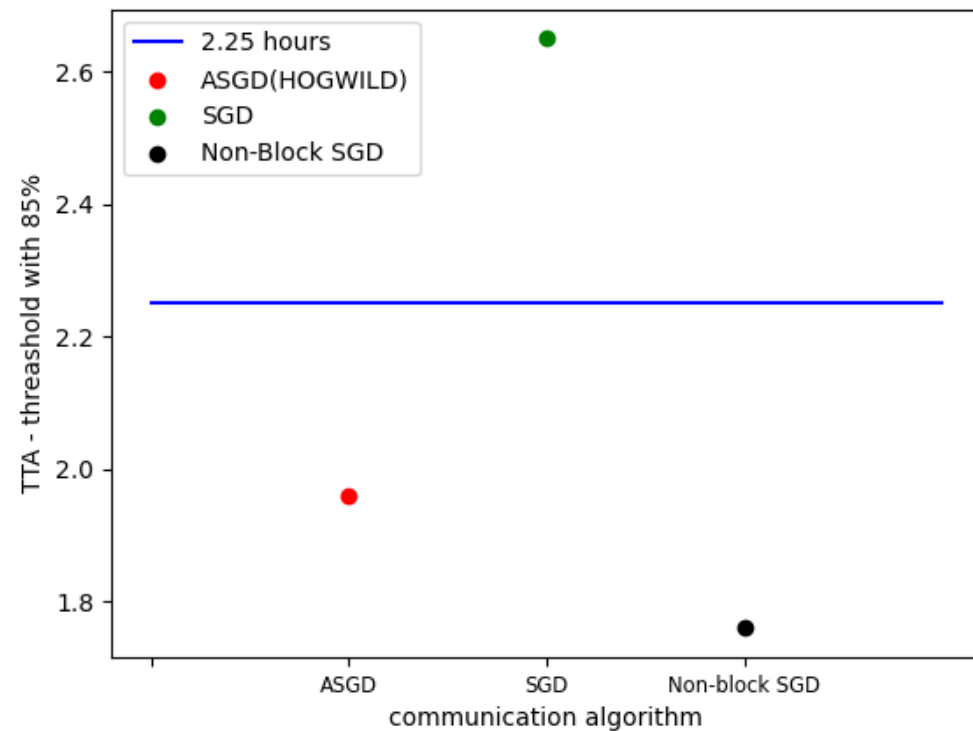
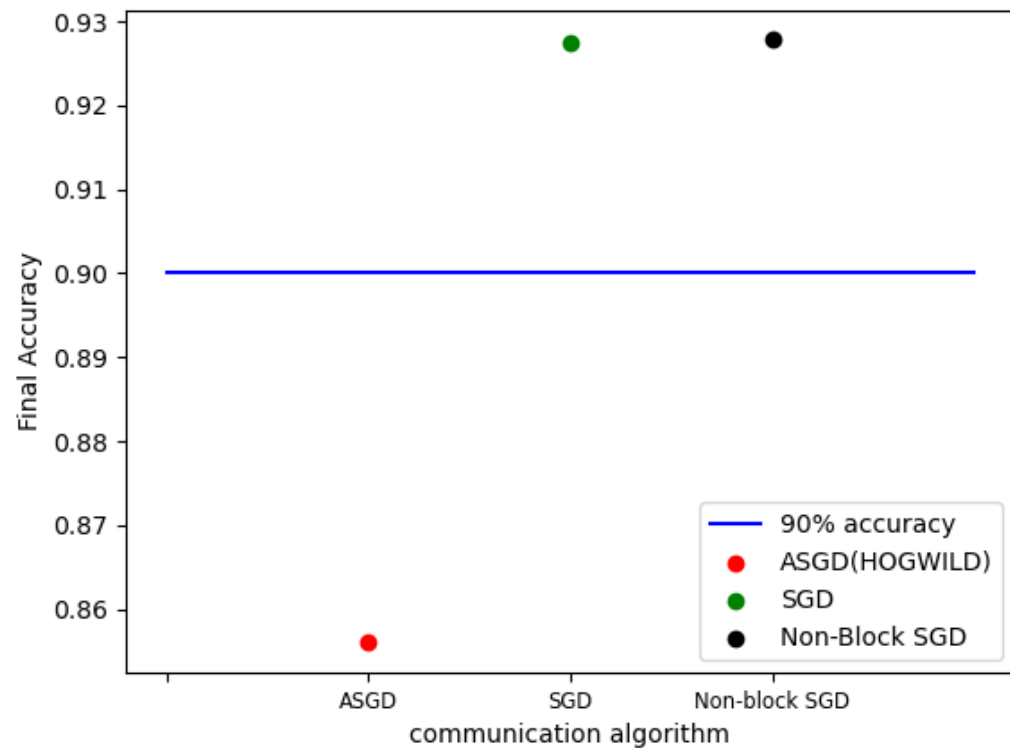
- Batch size = 256
- Cons: Lower Accuracy
-  : Communication time

Approach C – Non-Block SGD



- Batch size = 256
- Mini-batch size = 32
-  : Unfinished part
-  : Communication time

Main Results



Observations & Conclusion

- Non-blocking SGD has similar accuracy with SGD and much higher than ASGD
- Non-blocking SGD has similar TTA with ASGD and much quicker than SGD
- Non-blocking SGD has even better TTA than ASGD when threshold is high enough(85%)



GitHub Link

GitHub Link: <https://github.com/HectorHHZ/HPML>





NEW YORK UNIVERSITY

THANK YOU!