

Project Proposal Handout

CS554: Data-Intensive Computing

Assigned: 02-21-2023

Due: 11:59PM 03-01-2023

Overview

A major component of your grade will be based on a project. The topic of the project will be chosen by the student(s), and will require possibly the implementation of a real/simulated system, a written report, and oral presentations. You have been given a project idea based on your interests and background; you must expand it into a 2 to 3-page proposal (>1000 words), according to the outline below.

Your proposal should include the following information:

- Title
- people involved (in case it is a group project)
- abstract
- background information
- problem statement
- related work
- proposed solution
 - clearly state the nature of the project (e.g. implementation of a real system, simulation, theoretical, empirical performance evaluation, survey, etc)
 - be specific about what techniques you plan to use, what existing software and systems you will use, etc
- evaluation: be specific with what metrics you plan to examine
- timeline with weekly goals
- deliverables
- conclusion
 - what do you think you will learn (keep it short)
 - be specific with what how you will evaluate if your project is a success
- references (use Google Scholar at <https://scholar.google.com> to find at least 5 formal references; web pages such as Wikipedia do not count as formal references)

I know this is a lot of material to fit in 2 to 3 pages, but keep text to a minimum, where you can keep it short and to the point. Also, you should use single spaced 10pt font, with 1" margins. If you cannot stay in 3 pages, feel free to make it longer to cover all the things mentioned above. Also, you should utilize my office hours to meet with me in person and talk about your project. Please submit your proposal write-up through BB by the deadline stated at the top of this assignment.

Project Specific Details

The projects I mentioned in class are:

- **DATASET:** Fine-grained Crypto Dataset Publishing ([Binance.us](https://binance.us) 100+ coins since 2017)
- **PREDICT:** Time-series Prediction (deep learning vs [statistics](#), multivariate data input)
- **HASH:** Cryptographic Hashing Algorithms (SHA256, [BLAKE3](#), etc)
- **POWER:** Decentralized Power-Efficient Consensus Algorithms (Proof of Space vs. Proof of Work vs. Proof of Stake)

- **SHARDING**: Achieving 1K~1M transactions per second through block sharding (among other techniques)
- **CHIA**: Performance profiling of Chia Blockchain
- **SIMULATE**: Explore block propagation techniques at large scales through simulations

Relevant code repositories are:

- **Bamboo**: <https://github.com/pandanite-crypto/pandanite>
- Litecoin (LTC): <https://github.com/litecoin-project/litecoin>
- Chia blockchain: <https://github.com/Chia-Network/chia-blockchain>
- Other useful repositories:
 - Bitcoin (BTC): <https://github.com/bitcoin/bitcoin>
 - Ethereum (ETH): <https://github.com/ethereum/go-ethereum>
 - Blake3: <https://github.com/BLAKE3-team/BLAKE3>
 - SHA256: <https://www.movable-type.co.uk/scripts/sha256.html>

Testbeds you should be thinking about when you write your proposals are:

- Mystic
 - <http://mystic.cs.iit.edu>
- Chameleon
 - <https://www.chameleoncloud.org>

