

CAE Research Project Meeting Minutes

11-11-2022

- **Participants**
 - Dr. Waleed Farag (PI)
 - Dr. Xin-Wen Wu
 - Drew Rado
 - Emily Himes
 - Hector Linarez
 - Meghan Schilpp
 - Maria Balega
 - Sky Semone
 - Zach Keyser
 - Zaryn Good
- Meeting started at 6:00 pm
- Unanimously approved minutes for meeting on November 4th, 2022

Main Points:

- **Progress report**
 - Drew took 5 hours of data collection which resulted in 2300 lines of data. It was preprocessed removing his personal devices
 - Make sure number of data is correct
 - Are IP addresses correct
 - Nate has no updates
 - Sky mentioned that “hacking” the board is not recommended with the current configuration on IUP’s wifi. Worked on getting the new data to run with the XGBoost scripts from part one.
 - Meghan was meeting with Sky and Maria as they figured out XGBoost modifications.
 - Emily was working on OneDrive and reviewing documents for future presentations and reports.
 - Maria got the data to run on XGBoost, submitted paper.
 - Zaryn worked on paper, met with Maria and others for ML advice. Sent completed draft to Dr. Wu.
 - Zach and Hector met with Sky and determined that it was not safe to hack Arduino devices. Started researching synthetic data creation and the types of anomalies from the IoT-23 dataset.

Discussion points

- Primary focus was to get larger more representative dataset.
- Dr Ezekiel mentioned the following suggestions regarding creating the anomalous data and ML training:
 - Using VAE
 - Use a synthetic dataset
 - Train the model with our data
 - If the result it still bad, change parameters
- Dr. Farag and Dr. Ezekiel mentioned to Sky that this new dataset should be used on its own, and not with other datasets like the IoT23 set.
- Continue documentation, writing down procedures for how tasks are accomplished.

- **Tasks:**
 - Zaryn
 - Work with Dr. Wu on paper. Once data is present, run data through DCNNs
 - Sky
 - Meet with Dr. Ezekiel to make sure data collection methods are valid.
 - Process/generate synthetic data for the “anomalous” cases.
 - Hector/Zach
 - Work with Sky to create synthetic data and review data-viz tools in Python.
 - How will anomalous data work?
 - Drew
 - Collect more data and verify the IP addresses are correct. Review the pcap files.
 - Contribute to documentation efforts
 - Sky
 - Think about ways to inject anomalies, make sure data viz and processing tools are ready.
 - Meghan
 - Continue learning about SVM/DCNN and other ML processes with Maria and Zayrn
 - Use these for the new Zeek dataset.
 - Emily
 - Review OneDrive materials
 - Work on documentation
 - Maria/Zaryn
 - Start working with the new data (small dataset for now)
 - Update Meghan on procedures

Adjournment at 6:35

Next meeting will be held on Friday, November 11th at 6:00 pm