Machine Learning Applied to Failure Detection

Jingyi Zhao, jz3676 Computer Science and Mathematics
Wenbin Qi, wq372 Computer Science

September 2022

Context

The operation of distributed systems highly depends on reliable and continuous communication between nodes. The failure detection phase thus plays an important role in system unavailability [1]. However, authors of [2] proved that one cannot determine the failure of a node with 100% certainty. Therefore, Chandra and Toueg introduced the notion of *unreliable* failure detectors (FD) and proposed that a network of mutual observations via *unreliable* FDs can provide assessment regarding whether the system is still alive or crashed [3].

Chen et al. designed a failure detector algorithm to predict the next arrival time of the signal message based on some proposed quality of service metrics [4]. However, a constant safety margin is added to the expected arrival time in Chen's model. Since this safety margin is fixed, Chen's FD model becomes problematic for highly unstable traces. To address this problem, Bertier et al. proposed an adaptive FD which has a dynamic safety margin [5]. While most FD algorithms rely on statistical methods, Li explored the viability of ML-based FD by proposing an LSTM model [6]. Li's model has great performance but results in a high computational cost. Since we would expect FD to work in high frequency scenarios, heavy computation time will cause problems if it exceeds the actual time interval of arrivals.

Objectives

Our goal of this project is to propose a robust failure detector algorithm that deals with real-world high frequency time series. To achieve this, we mainly need to address two tasks: 1) to design a model which reduces computation time while maintains satisfying performance; 2) to design a model which not only performs well on the existing **PlanetLab** dataset but also adapts to real-world applications. Below are several research directions we plan to work on:

- Find and apply existing machine learning models that deals with high frequency time series.

- Simplify the previous LSTM model (e.g., simplify model structures and change activation functions) to reduce computational cost.
- Modify the training and prediction process, say, to generate the next ten arrivals in every training epoch in order to amortize the computation time.
- Test the computation time of our models against Li's model while maintaining good *probability of availability* and *detection time*, at least not too far off compared to Li's, Chen's, and Bertier et al.'s models.
- Apply the models to real-world situations and evaluate the performance accordingly.

Supervision

– Olivier Marin (ogm2@nyu.edu)

References

[1] J. B. Leners, H. Wu, W.-L. Hung, M. K. Aguilera, and M. Walfish, "Detecting failures in distributed systems with the falcon spy network," in *Proceedings of the Twenty-Third ACM Symposium on Operating Systems Principles*, ser. SOSP '11. New York, NY, USA: Association for Computing Machinery, 2011, p. 279–294. [Online]. Available: https://doi.org/10.1145/2043556.2043583

[2] M. J. Fischer, N. A. Lynch, and M. S. Paterson, "Impossibility of distributed consensus with one faulty process," *J. ACM*, vol. 32, no. 2, p. 374–382, apr 1985. [Online]. Available: https://doi.org/10.1145/3149.214121

[3] T. D. Chandra and S. Toueg, "Unreliable failure detectors for reliable distributed systems," *J. ACM*, vol. 43, no. 2, p. 225–267, mar 1996. [Online]. Available: https://doi.org/10.1145/226643.226647

[4] W. Chen, S. Toueg, and M. Aguilera, "On the quality of service of failure detectors," *IEEE Transactions on Computers*, vol. 51, no. 1, pp. 13–32, 2002.

[5] M. Bertier, O. Marin, and P. Sens, "Implementation and performance evaluation of an adaptable failure detector," in *Proceedings International Conference on Dependable Systems and Networks*, 2002, pp. 354–363.

[6] X. Li and O. Marin, "Towards implementing ML-based failure detectors," in *18th European Dependable Computing Conference, Student Forum*, Zaragoza, Spain, 2022.