Adaptive Caching With Follow the Perturbed Leader Replacement Policy

Mikkel Mäkelä
Supervisors: Georgios Iosifidis, Tareq Si Salem

EEMCS, Delft University of Technology, The Netherlands
m.makela@student.tudelft.nl

Background

- Internet traffic is growing exponentially [1].
- Caching is mechanism that places frequently requested content in faster to access memory, reducing network load and latencies.
- Traditionally a low-level problem with hard computational constraints on algorithms.
- Web caching has made machine learning algorithms possible due to bigger I/O overhead.
- Caching can be modelled as an online learning problem.
- Online Convex Optimization has produced many promising replacement policies, including Follow the Perturbed Leader (FTPL) [2].
- FTPL performs poorly when popularities of requested objects change over time.

Research Goal

- Develop a replacement policy that can adapt to different request patterns, but maintain a mathematical upper bound.
- Strike a balance between adversarial resistance and practicality.

Contributions

- Introduce a constant discount rate to FTPL, intentionally omitting the upper bound.
- Use adaptive FTPL with various discount rates (including the one with an upper bound), as experts in an Incrementally Adaptive Weighted Majority (IAWM) algorithm [3].
- IAWM makes our algorithm adaptible by learning the optimal rate, but also comes with a perfromance guarantee which carries over to our algorithm.

Method

- Python3 simulator that runs traces and benchmarks algorithms.
- Uses both realistic and adversarial traces.

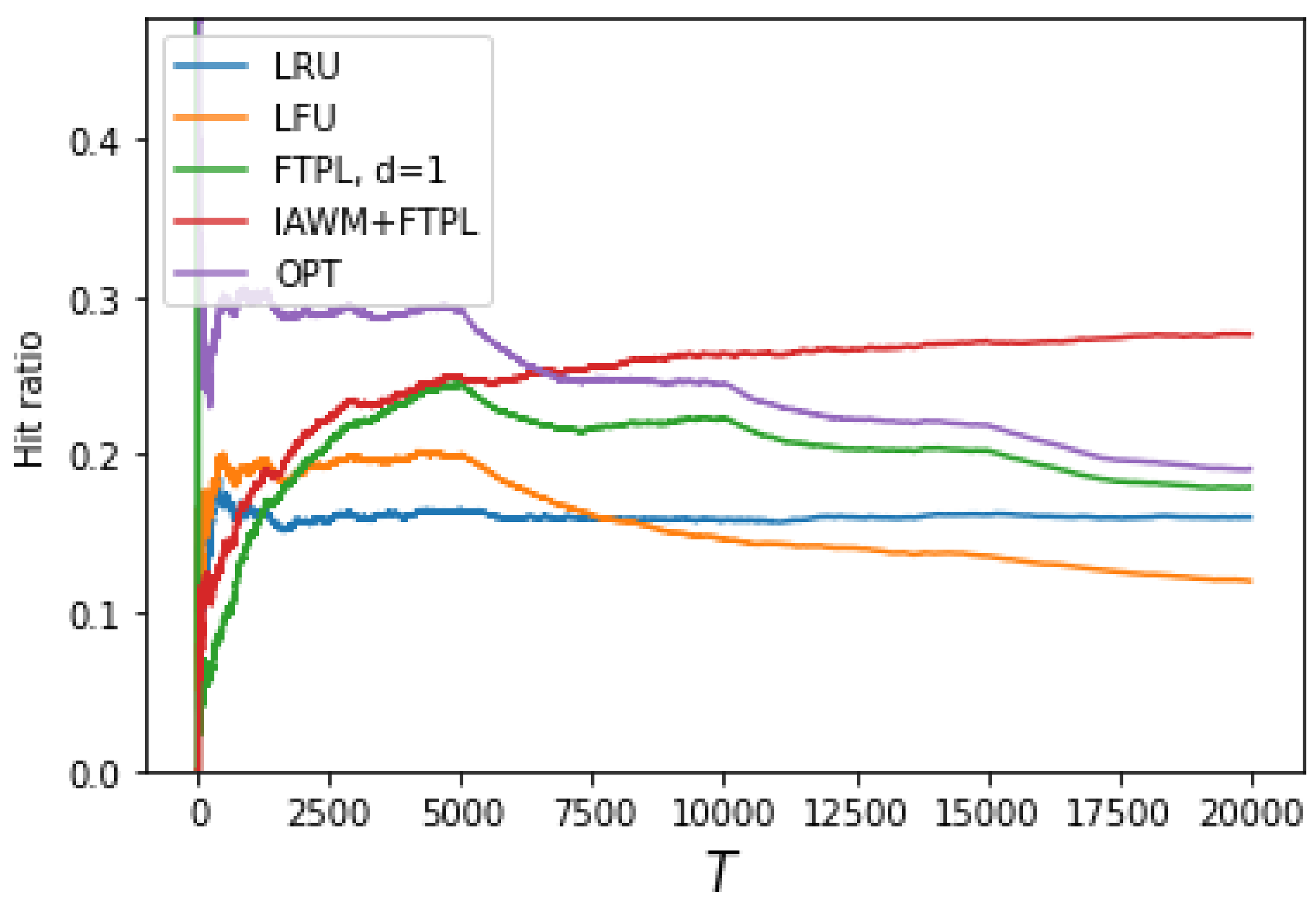

Figure 1: Hit ratios for each time slot for a synthetic changing popularity trace.

Results

- Adaptive FTPL significantly outperforms FTPL and the static benchmark when popularities change.
- Performance is similar on fixed popularity traces.
- Adaptive FTPL can effectively tackle various adversarial traces.
- Adaptive FTPL strongly outperforms FTPL in a bipartite network when edges are connected optimally.
- As edge configurations become more adversarial, Adaptive FTPL begins to resemble FTPL.

Future work

- Different machine learning techniques to learn a discount rate.
- Optimally matching clients with caches in networks.
- Expert algorithm with a better bound.

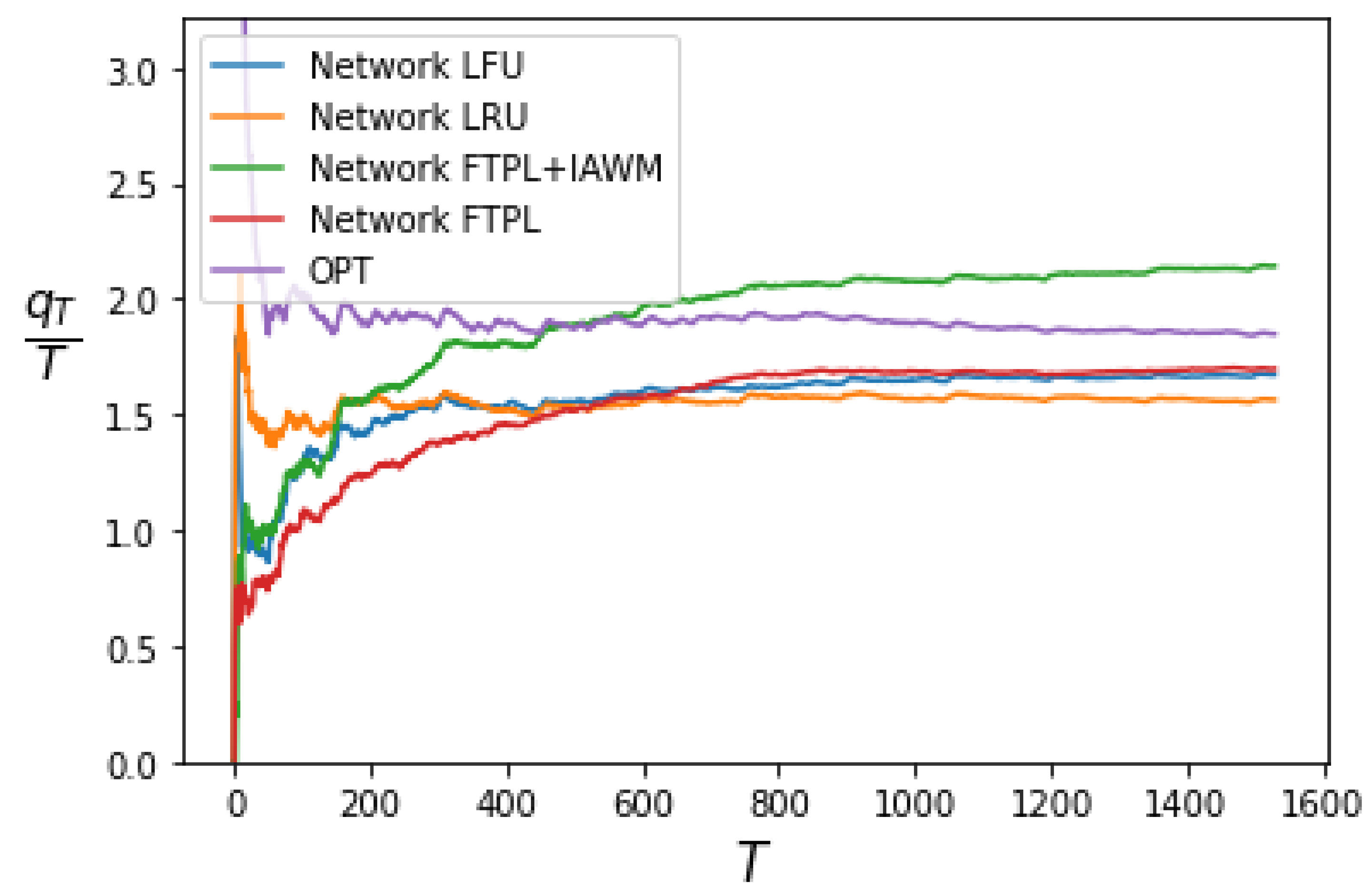

Figure 2: Reward for each time slot in an edge-optimal bipartite network.

References

[1] T. Barnett, S. Jain, U. Andra, and T. Khurana, "Cisco visual networking index (vni) complete forecast update, 2017–2022," *Americas/EMEAR Cisco Knowledge Network (CKN) Presentation*, pp. 1–30, 2018.

[2] R. Bhattacharjee, S. Banerjee, and A. Sinha, "Fundamental limits on the regret of online network-caching," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 4, no. 2, pp. 1–31, 2020.

[3] P. Auer, N. Cesa-Bianchi, and C. Gentile, "Adaptive and self-confident on-line learning algorithms," *Journal of Computer and System Sciences*, vol. 64, no. 1, pp. 48–75, 2002.