Errata for 'Removal Policies in Network Caches for World-Wide Web Documents'

Stephen Williams, Marc Abrams, Charles R. Standridge, Ghaleb Abdulla, and Edward A. Fox

Abstract

Due to an error in the simulation used to conduct removal policy experiments, one of the primary conclusions from our SIGCOMM'96 paper[1] is in error. We present here the corrections, and some new results. The conclusion "sorting by SIZE is the best policy to maximize hit rate," is correct. However, the other primary conclusion "sorting by SIZE is the best policy to maximize weighted hit rate," is *incorrect*. In fact, sorting by SIZE was found to be the *worst* policy. A brief description of changes in each of the four experiments conducted follows. The text of the revised paper can be found at `http://www.cs.vt.edu/~chitra/docs/96sigcomm/`.

1 Simulation Error

The simulation used in the original paper contained a bug that failed to correctly map identical URLs from different trace files to the same document identifier. The modified simulation was verified with several hand checked test cases, and with another simulation written at the University of Saskatchewan.

2 Experiments

Experiment 1: All of the basic results and conclusions are correct. However, the shape of the graphs presented is different. Also, some of the suppositions, based on the curve of the infinite cache HR and WHR have changed. In particular, the Classroom (C) workload exhibits a desirable effect. Both the HR and WHR increased significantly in the final days of the semester, which may indicate that students were reviewing previously viewed pages in preparation for a final exam.

Experiment 2: The results from this experiment, where the cache size was fixed at 10% of the size needed to hold all unique documents, changed substantially.

Results from tests on the primary sort key for hit rate (HR) show the best policy across all workloads to be SIZE, as was concluded in the original paper. The order of ranking of other the policies changed, however. NREF (corresponding to LFU) ranks second in four out of five workloads. ETIME (time of first access) is the worst in all workloads.

Results from tests on the primary sort key for weighted hit rate (WHR) show that SIZE is the worst policy in all workloads, contrary to our previous results. Instead, NREF ranks highest in two of the five workloads. In workloads C, G, and U there is no significant difference in performance of the non-size based sort keys.

Results from the tests on secondary sort key show that none of the policies makes a significant difference in HR or WHR when the primary sort key is $\lfloor \log_2(\text{SIZE}) \rfloor$. For WHR the secondary key never improves the average performance more than 1% of that when a random selection is made. For HR the performance gain is less than 0.5%.

Experiment 3: No changes in the general conclusions were made.

Experiment 4: There are no changes in the general conclusions. However, the original graphs showed WHR with respect to ONLY the form of traffic being cached. The revised graphs show the WHR with respect to the total traffic.

[1] S. Williams, M. Abrams, C.R. Standridge, G. Abdulla, and E.A. Fox, "Removal Policies in Network Caches for World-Wide Web Documents," Proc. of ACM SIGCOMM '96 Conference, August 1996, Stanford, CA, USA, pp.293-305

3 Acknowledgments

We thank Martin Arlitt of the University of Saskatchewan for his help in verifying the results from our new simulation. We also thank Ilhwan Kim of Seoul National University, Korea for bringing one of our original conclusions into question.