Improving Timeliness of Information through Replicating Requests

Caleb Skinner
Candidate for BSc
Department of Computer Science
The University of Texas at Austin
Austin, TX 78712, USA
calebskinner@utexas.edu

Yu Sang
Department of Computer and Information Sciences
Temple University
Philadelphia, PA 19122, USA
yu.sang@temple.edu

Bo Ji
Department of Computer and Information Sciences
Temple University
Philadelphia, PA 19122, USA
boji@temple.edu

***Abstract*— The Internet of Things has created an influx of data-generating devices. Applications that use these devices require the generated data to be transmitted in a timely manner. In this paper, we consider a Pull model, where remote clients will send requests to retrieve the information generated by multiple sources from the same random process. We assume that the update processes and response times of different sources are different. The timeliness, measured by age of information, depends not only on the freshness of information at the sources but also on the response times. Getting information from a fresh source with a large response time may lead to a larger age of information compared to getting information from a slightly less fresh source with a much smaller response time. Hence, an important objective is to optimize the timeliness of information obtained by the clients, accounting for the tradeoff between the freshness at the sources and the response times. We propose a replication scheme that improves age of information by replicating requests to multiple sources and waiting for a certain number of responses. We show through simulation that depending on the system parameters, the number of responses to wait for could be different for optimizing the information timeliness.**

I. Introduction

The Internet of Things is rapidly increasing the number of data-generating devices. Applications that use these devices require that generated data be transmitted in a timely manner. Timeliness, also known as freshness, can be estimated using a metric known as age of information, or age. Age is defined as the time elapsed since the generation of data. If data is generated regularly, age rises until the next update as seen in Fig. 1.

Much of the existing data freshness research focuses on models in which data can be sent from a source to a client at any point in time [1] - [6]. In such models, data must be queued and the behavior of the queue affects the timeliness of data transmission. We analyze a Pull model in which data must be requested before being sent. It is assumed that both the time at which updates occur and response times are unique to each source. Timeliness is affected by both of these factors. A source with fresh data and a long response time may have a larger age than a source with less fresh data and a shorter response time.

We propose the use of a replication scheme in order to reduce the average age of information. In this scheme, data requests are duplicated to all available sources and a certain number of responses are waited for. Increasing the number of responses waited for also increases the probability of receiving fresher data. This is accompanied by an increase in the amount of time spent waiting for data. We attempt to show through simulation that the optimal number of responses to wait for varies under different system parameters.

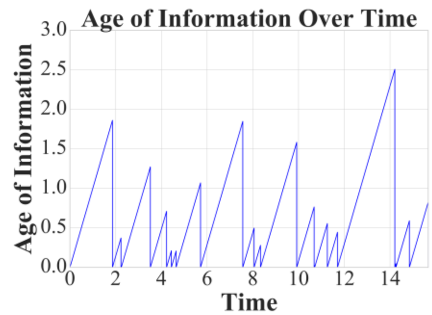

Fig. 1. Behavior of age of information over time

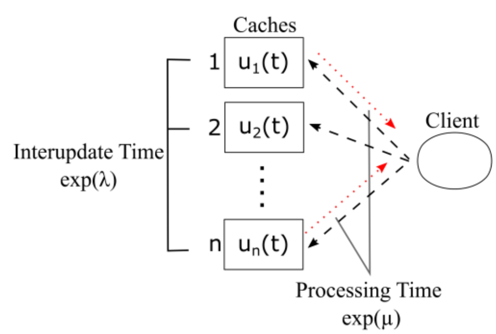

Fig. 2. System model

II. System Model

A. Configuration

This work assumes a model, as seen in Fig. 2, consisting of one client connected to n caches. The client duplicates data requests to each of the caches and waits for d responses. Data transmission between the client and caches is instantaneous.

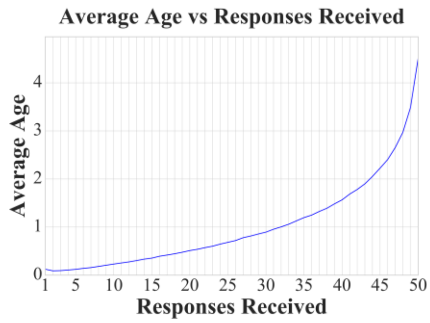

Fig. 3. Simulation results for λ = 10 and μ = 1

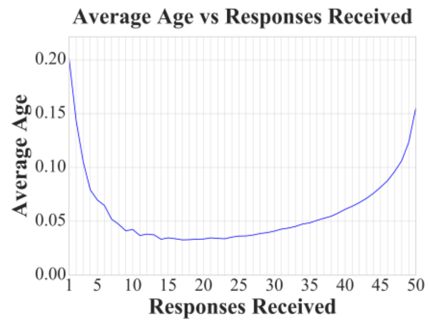

Fig. 4. Simulation results for λ = 3 and μ = 30

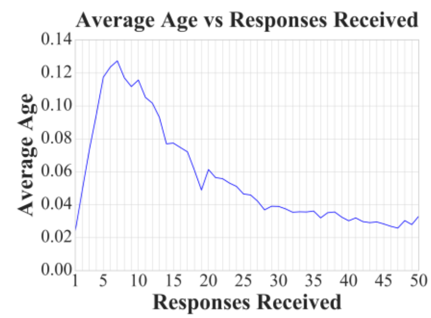

Fig. 5. Simulation results for λ = 1 and μ = 400

B. Update Policy

Each caches possesses a timestamp, *u(t)*, of the time, *t*, at which its last update occurred. Like [5], updates occur according to a Poisson distribution of λ. Random variables drawn from an exponential distribution with a mean of 1/λ are used to represent the time between each update.

C. Response Time

Upon reception of a data request, caches processes their current data before responding. Similar to [7] and [5], processing times are determined by random variables drawn from an exponential distribution with a mean of 1/μ.

D. Calculating Data Fitness

Upon reception of *d* responses, the client computes the age, Δ*t*, using

$$\Delta t = t - u(t). \tag{1}$$

III. Results

We hypothesized that the following behaviors would occur as *d* increased. Age would increase when response times were faster than processing times. Age would decrease and then increase when updates and responses took approximately the same amount of time. Age would decrease when response times were faster than update times. A simulation, using the described model, was used to test these hypotheses. Parameter *n* was set to 50 and the results of 1000 data request were averaged. The simulation produced the following results as *d* increased.

A. Increasing

In cases where update times were faster than response times, such as in Fig. 3, the age increased. In these cases, the waiting time increase caused by increasing d was more than the difference between timestamps and caused age to increase.

B. Decreases then Increases

As expected, the age decreased and then increased when updates and responses took approximately the same amount of time. We found that this behavior also occurred when response times were less than update times, such as in Fig. 4. In both case increasing d reduces age, but eventually the time it takes for another response to arrive is greater than the difference between possible timestamps. This caused age to increase.

C. Increases then Decreases

When response times were significantly faster than update times, such as in Fig. 5, age increased before decreasing. Few or none of the caches updated during the simulation at the lower values of *d*. As a result, the wait time increased without the possibility of obtaining fresher data. This caused age to increase. Larger values of *d* increased wait time enough for caches to update during the duration of the simulation. Within this range, increasing *d* results in more caches updating and generally a lower average age.

IV. Conclusions

Our results suggest the effects of increasing the number of responses waited for are dependent on the relationship between average update times and average response times. When response times are either approximately equal to or significantly faster than update times waiting for more than one response generally reduces the average age. When update times are faster than response times it is better to wait for one response or a few responses in some cases.

References

[1] S. Kaul, R. Yates and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE INFOCOM*, 2012.

[2] A. Bedewy, Y. Sun, and N. Shroff, "Optimizing Data Freshness, Throughput, and Delay in Multi-Server Information-Update Systems," in *Proc. IEEE ISIT*, 2016.

[3] Y. Sun, E. Uysal-Biyikoglu, R. Yates, C. Koksal, and N. Shroff, "Update or Wait: How to Keep Your Data Fresh," in *Proc. IEEE INFOCOM* , April 2016.

[4] R. Yates and S. Kaul, "Real-time status updating: Multiple sources," in *Proc. IEEE ISIT*, 2012.

[5] M. Costa, M. Codreanu and A. Ephremides, " Age of Information with Packet Management," in *Proc. IEEE ISIT*, 2014.

[6] Q. He, D. Yuan and A. Ephremides, "Optimizing freshness of information: On minimum age link scheduling in wireless systems," in Proc. *WiOpt*, 2016.

[7] C. Kam, S. Kompella and A. Ephremides, "Age of information under random updates," in Proc. IEEE ISIT, 2013.