Optimal Update Times for Stale Information Metrics Including the Age of Information

Chris Ferguson and Leonard Kleinrock, *Life Fellow, IEEE*

***Abstract*—In this paper we examine the general problem of determining when to update information that can go out-of-date. Not updating frequently enough results in poor decision making based on stale information. Updating too often results in excessive update costs. We study the tradeoff between having stale information and the cost of updating that information. We use a general model, some versions of which match an idealized version of the Age of Information (AoI) model. We first present the assumptions, and a novel methodology for solving problems of this sort. Then we solve the case where the update cost is fixed and the time-value of the information is well understood. Our results provide simple and powerful insights regarding optimal update times. We further look at cases where there are delays associated with sending a request for an update and receiving the update, cases where the update source may be stale, cases where the information cannot be used during the update process, and cases where update costs can change randomly.**

***Index Terms*—Information decay, age of information, optimal update times, out-of-date information, stale information, update frequency.**

I. Introduction

MANY forms of information are constantly changing or going out-of-date over time. Examples include weather reports, stock quotes, web pages, data files, medical test results, traffic information and much more. Getting this information to where it needs to be and keeping it as up-to-date as practical is a challenging task. There has been considerable research for particular models regarding updating out-of-date information in which sources send time-stamped status updates to interested recipients. These applications desire status updates at the recipients to be as timely as possible. In this paper we model the declining value of information as it decays and solve for the optimum time to update information that has gone stale. The optimal update time depends on the value of the information to the user, as well as the cost of obtaining an update. This general problem and our solution of determining when to optimally update information as it goes out-of-date was reported by us earlier in [1]. Subsequent to our model, a related end-to-end metric of staleness, called Age of Information (AoI), has been reported extensively in the literature [2], [3]. Similar to us, AoI addresses optimal update issues including threshold-based policies. These have been reported in [4], [5], [6] for energy harvesting, for example. We analyze some simple versions of that AoI metric as well as more general metrics. As mentioned above, there are many applications that face this problem. Such applications include: updating automobile traffic information [7], when to repeat a measurement of atmospheric data for weather forecasting [8]; determining the optimal time to update routing tables in data networks [9]; when to update medical testing (such as cancer screenings, mammograms, infectious disease testing such as COVID-19 tests, etc.) [10]; updating stock prices [11]; updating a transaction database [12]; updating data regarding a security or surveillance system [13]; updating cached pages [14]; refreshing website data [15]; updating the relative position of drones in a swarm [16]; updating measures of reputation [17]; status update policies under an energy harvesting setting [4], [5], [6]; and many more. These applications range over a wide variety of systems.

We consider a number of different models. These models are relatively simple and lead to uncomplicated closed-form expressions that result in intuitive understanding of the underlying behavior. In Section II we state our basic assumptions. Then in Section III we consider the simple case where the update cost is fixed and we know the value of the information as a function of how stale it is. In Section IV, we expand our model to include the AoI metric and show when best to initiate an update. In Section V we look at the case where there are delays associated with requesting an update and receiving it and therefore the received information starts somewhat out-of-date and also where the information source itself may be out-of-date. In Section VI we assume there is no gain during the update process. In Section VII we introduce variable update costs. In Section VII-A we consider cases where users become intermittently disconnected and are unable to receive updates, and in Section VII-B where users intermittently move between states where the cost is low and high. Finally, Section VIII provides a short conclusion.

Manuscript received 11 March 2023; revised 14 August 2023 and 27 November 2023; accepted 15 December 2023. Date of publication 22 December 2023; date of current version 28 December 2023. This work was supported in part by the Advanced Research Projects Agency through Advanced Networking and Distributed Systems under Contract MDA-972-91-J-1011, through Networking Control Algorithms for Mobile Wireless Data Communications under Contract DAAH-04-04-G-0364, and through Computer Aided Design of High Performance Wireless Networked Systems under Contract J-FBI-93-112. *(Corresponding author: Leonard Kleinrock.)*

Chris Ferguson is with the Computer Science Department, University of California at Los Angeles, Los Angeles, CA 90095 USA, and also with the IT Department, RP1, Las Vegas, NV 89158 USA.

Leonard Kleinrock is with the Computer Science Department, University of California at Los Angeles, Los Angeles, CA 90095 USA (e-mail: lk@cs.ucla.edu).

Digital Object Identifier 10.1109/JSAIT.2023.3344760

II. Assumptions

The primary assumptions we use are:

1) There is a finite cost $C > 0$ of updating a given piece of information.

2) There is a value per unit time associated with having this piece of information, and this value changes over time depending on the amount of time, t, since it was last updated. This value we denote by $f(t)$. The units of cost and value are assumed to be the same, e.g., dollars.

3) $f(t)$ is a non-negative monotonically non-increasing function.[1]

4) For Section III, we further assume[2] that $\lim_{t\to\infty} f(t) = 0$.

5) Updates occur immediately upon being requested.[3]

In the first assumption above, C is the actual cost of obtaining an update. As examples: for COVID testing, it is the dollar price of the test; for updating routing tables, it is the cost in dollars for using resources to process the path information; for updating a remote file cache, it is the bandwidth cost in dollars plus any cost the server may charge for the update. In some applications, the cost C is not easy or obvious to determine, such as updating weather data. For the AoI models, a specific cost is not usually addressed, but rather, the focus is typically on the timeliness of status updates, which may degrade due to congestion if updates are generated too often.

In the second assumption above, it can be much harder to evaluate $f(t)$, namely, the instantaneous benefit rate resulting from having the information that is t units old. As examples: for COVID testing, it is the rate of value obtained for limiting the spread of the infection and improved treatment by knowing the test results; for updating routing tables, it is the instantaneous dollar value of improved response time; for updating a remote file cache, it is the instantaneous dollar value to the network for having access to the cache.

Assumptions 2 and 3 are very important. Assumption 2 states that there is a value for holding on to a piece of information over time which depends on how out-of-date the information is. If the information was last updated at time 0, then the accumulated value received for having this information from t_1 to t_2, (where $0 \le t_1 \le t_2$) is $\int_{x=t_1}^{t_2} f(x)dx$. Assumption 3 states that the instantaneous value of the information, $f(t)$, cannot increase over time as the information goes further and further out-of-date. Thus, an older piece of information can never be more valuable than a newer piece. Furthermore, each update subsumes all past updates in that no past updates have any value once a new update is received, i.e., it is a Markov process. We rely on this part of Assumption 3 in order for us to prove our main Theorem 1. The idea of information giving value over time is an innovative way of looking at information; we point out that a special case of this is the AoI metric (which can take on negative values as information goes out-of-date) to which we extend our analysis in Section IV.

Example 1: As stated above, in practice it may be difficult to come up with the function $f(t)$. Here we give a real-world example where $f(t)$ can be easily determined. Consider the case of a cache server that can store local copies of files received from a host. The files on the host are subject to random modifications at a Poisson rate. When the cache server sends a request to update its local copy of a file, it obtains the latest version of that file from the host. Consider a particular file which costs C to update. Assuming that the value per unit time of the cached file is equal to the probability that no modifications have been made to the source file since the last update, then $f(t) = e^{-\rho t}$, where ρ is the Poisson rate of modifications for this file. This function is graphed in Figure 1 along with the accumulated value gained since the last update for $\rho = 1$. The decreasing curve is the instantaneous value of the information, and the rising curve is the accumulated value gained since the last update, which is the integral of $f(x)$ from 0 to t.

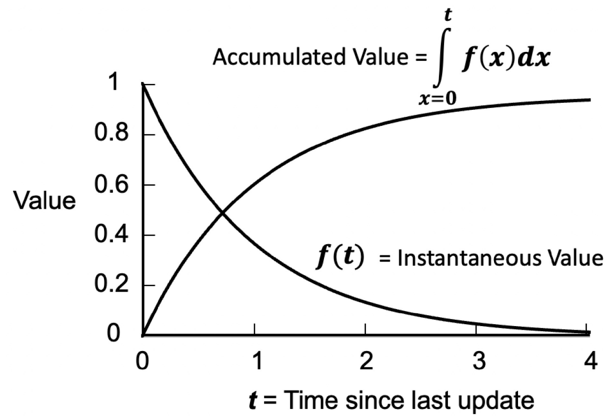


Fig. 1. Instantaneous Value Gained and Accumulated Value Gained Over Time Since the Last Update.

III. Problem Statement and Solution

Given that we know C and $f(t)$, when and how often should a piece of information be updated? Let t be the amount of time since the last update was received. We seek to find the optimum time to update (defined as t^*) that maximizes the average value gained per unit time between updates. Once an update is received, this repeating process starts all over again with $f(t)$ starting at its orginal (now updated) value, and as the new information goes out-of-date, one must once again decide when to update. Our main result from [1] is in the following Theorem:

Theorem 1: To maximize the average value gained per unit time, a piece of information should be updated as soon as t satisfies

$$\frac{-C + \int_{x=0}^{t} f(x)dx}{t} \ge f(t). \tag{1}$$

(this value being t^*) where t is the time since the last update. However, if C is so large such that Eq. (1) is never satisfied, then one should never update.[4]

Proof: The cost incurred is the cost of one update, C. The total value gained by time t is the integral of the instantaneous value of the information from the time it is received to

[1] In Section IV we allow $f(t)$ to be positive and/or negative.

[2] In Section IV we relax this condition.

[3] In Section V we relax this assumption.

[4] We examine when the condition is satisfied below in Theorem 2.

time t, namely $\int_{x=0}^{t} f(x)dx$. The objective function we seek to maximize is the average gain per unit time which equals the total information value gained minus the update cost, divided by the time, or in this case[5]

$$\frac{-C + \int_{x=0}^{t} f(x)dx}{t}. \tag{2}$$

To find t^*, the value of t that maximizes this average gain per unit time, we take the derivative[6] of this expression and set it equal to zero, namely

$$\frac{C}{t^2} - \frac{\int_{x=0}^{t} f(x)dx}{t^2} + \frac{f(t)}{t} = 0. \tag{3}$$

Multiplying by t and rearranging terms yields the following equation defining the optimal value of t, namely, t^*

$$\frac{-C + \int_{x=0}^{t^*} f(x)dx}{t^*} = f\left(t^*\right). \tag{4}$$

We now show that the objective function in Eq. (2) has a unique maximum value by proving that its slope is positive when $t < t^*$ and is negative for $t > t^*$. We introduce the following definitions:

$$A(t) \triangleq \int_{x=0}^{t} f(x)dx. \tag{5}$$

$A(t)$ is the accumulated value gained by time t.

$$D(t) \triangleq A(t) - tf(t). \tag{6}$$

Examining $dD(t)/dt$ we note that

$$\begin{aligned} dD(t)/dt &= d\left[A(t) - tf(t)\right]/dt, \\ &= dA(t)/dt - f(t) - tdf(t)/dt, \\ &= f(t) - f(t) - tdf(t)/dt. \end{aligned}$$

But by Assumption 3 in Section II, $df(t)/dt \leq 0$ which shows that

$$dD(t)/dt \geq 0. \tag{7}$$

We denote the objective function given in Eq. (2) by $Q(t)$. Note that the left-hand-side of Eq. (3) is the derivative of our objective function $Q(t)$. Multiplying this derivative by t^2 (clearly this does not change the sign of the derivative) we get $C - \int_{x=0}^{t} f(x)dx + tf(t)$ and from Eqs. (5) and (6) this is merely $C - D(t)$. Since C is a constant and $D(0) = 0$ and from Eq. (7) $D(t)$ cannot decrease with t, we have

$$dQ(t)/dt = \begin{cases} > 0 \text{ if } C > D(t) \\ = 0 \text{ if } C = D(t) \\ < 0 \text{ if } C < D(t). \end{cases} \tag{8}$$

This shows that $Q(t)$ has a unique maximum if $\lim_{t\to\infty} D(t) \geq C$. Further, from Eq. (4) we know that the maximum occurs at $t = t^*$ for which $D(t^*) = C$, that is,

$$dQ(t)/dt = \begin{cases} > 0 \text{ if } t < t^* \\ = 0 \text{ if } t = t^* \\ < 0 \text{ if } t > t^*. \end{cases} \tag{9}$$

Hence, the objective function has a unique maximum. ■

[5]A discrete version of Eq. (2) can be found in [18].

[6]If $f(t)$ is discontinuous at $t = x$, then we consider $f(x)$ to hold all values between $f(x^-)$ and $f(x^+)$.

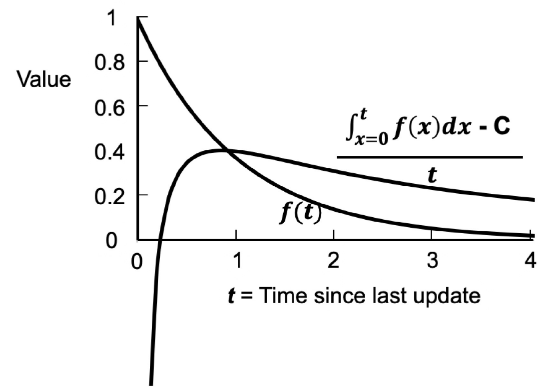


Fig. 2. Instantaneous Value Gained and Average Value Gained Per Unit Time Since the Last Update Including the Update Cost.

We repeat that if Eq. (4) is never satisfied, then no updates should ever take place (and the exact conditions are addressed in Theorem 2). Also note that since $f(t)$ may be constant over an interval ($t_0 \leq t \leq t_1$), if t^* occurs in this interval, then all values of t in this interval are optimal. In such a case, we choose to use the earliest, t_0, as t^*.

It is intriguing to note that $t^* f(t^*)$ is the net value gained per cycle and the net value gained per unit time over each cycle is $f(t^*)$ which we denote by V. Updates should be performed to maximize the average gain per unit time which equals the total information value gained minus the update cost, divided by the time, namely, our objective function, $[-C + \int_{x=0}^{t} f(x)dx]/t$. See Figure 2. We have just proven in Eq. (4) that this function is at its maximum when its value equals $f(t)$.

Figure 2 shows a plot of the sample function $f(t) = e^{-t}$ and also the average value gained per unit time since the last update, assuming $C = 0.2$, that is, $[-0.2 + \int_{x=0}^{t} f(x)dx]/t = [0.8 - e^{-t}]/t$. This average value function starts at $-\infty$ since $C > 0$ and we are dividing by time which starts at zero. According to our theorem, this function hits its maximum when it equals $f(t)$. This can be explained by understanding the two functions. At the point where they cross, the average value gained per unit time since the last update (including the update cost) equals the current instantaneous value of the most recent update, $f(t)$, i.e., the value currently being gained per unit time. Since $f(t)$ is non-increasing, at no time t beyond the crossing does the information have a greater instantaneous value $f(t)$ than it does at the crossing. Thus, since $[-C + \int_{x=0}^{t} f(x)dx]/t = f(t)$ at the crossing, this objective function must decrease continuously past the point where it meets $f(t)$. Furthermore at any point before the two lines meet, the instantaneous value gained by having the information is greater than the average value gained per unit time. Thus as long as, $f(t) > [-C + \int_{x=0}^{t} f(x)dx]/t$, waiting a little longer without updating increases the average value gained per unit time over the period of one update. Furthermore this shows that there is only a single local maximum value as shown in the proof of Theorem 1. This is true for any monotonically non-increasing $f(t)$ and positive C.

Of interest is the following Corollary which shows that there is a class of functions including f(t) all of which produce the same t*.

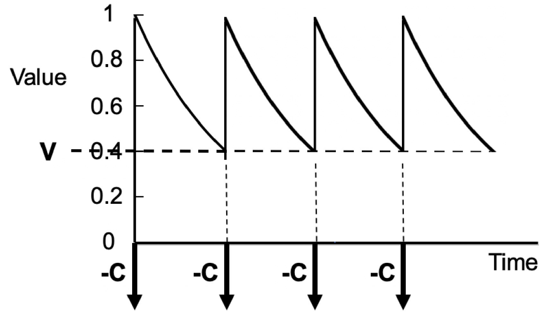


Fig. 3. The Recurring Update Process.

Corollary 1: For a given C, if t^* is the optimal update time for $f(t)$, then t^* is also the optimal update time for all non-increasing $g(t)$ where $g(t^*) = f(t^*)$ and $\int_{x=0}^{t^*} f(x)dx = \int_{x=0}^{t^*} g(x)dx$.

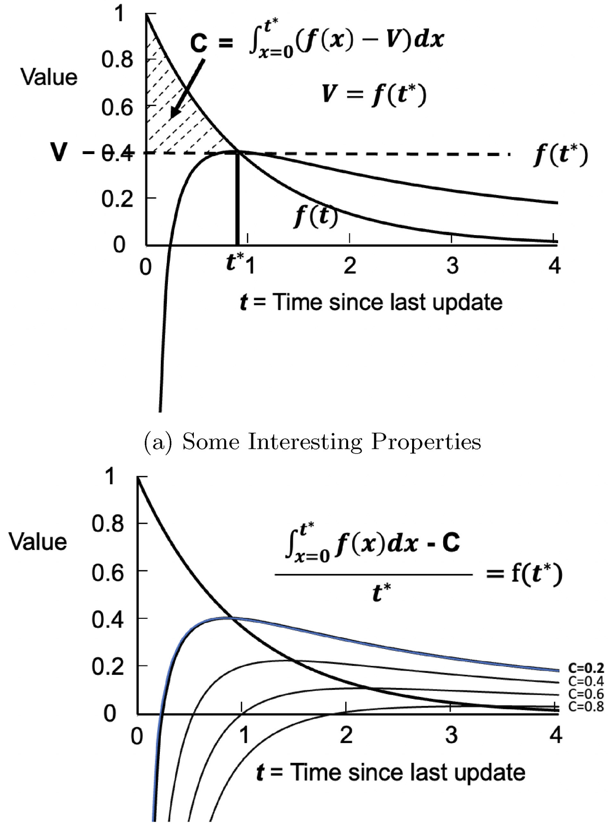
Proof: From Eq. (4), we see that if $\int_{x=0}^{t^*} f(x)dx = \int_{x=0}^{t^*} g(x)dx$ and $f(t^*) = g(t^*)$, then Eq. (4) clearly holds for such $g(t)$. ∎

We can also see from this Corollary that the future form of $f(t)$ beyond t^* is irrelevant. Nothing that happens after t^* can ever affect the value of the optimal update time, t^*. It is interesting and useful to note that one does not even have to know the function f(t) ahead of time. One only needs a way to measure it on the fly, and as soon as the current rate of productivity goes below the average rate since the last update, the next update should be performed. Furthermore, the shape of $f(t)$ prior to t^* is irrelevant as long as $\int_{x=0}^{t^*} f(x)dx = C + t^* f(t^*)$.

Recall that the process of updating information is recurrent. After the update is obtained, the process starts over. When this new copy becomes far enough out-of-date, i.e., after another t^* time units, it is replaced by a newer (and, according to our model, an instantaneous) update. Figure 3 is a graph of the value per unit time of the current copy on the positive Y axis. On the negative Y axis, we see the cost paid every t^* time units for each new copy.

As shown in Figure 4(a), this function has some interesting properties. The average value, V, gained over time under optimal updating, is equal, by Eq. (4), to $f(t^*)$, where t^* is the optimal update time. Rewriting Eq. (4) we have $C = \int_{x=0}^{t^*} f(x)dx - f(t^*)t^*$. But since $f(t^*) = V$, we have that $C = \int_{x=0}^{t^*} (f(x) - V)dx$. Figure 4(b) shows different curves for the different values of C. Note that t^* always occurs at the peak of $[\int_{x=0}^{t} f(x)dx - C]/t$. The higher the update cost, the less frequently updates should be performed, and the less value that can be extracted from a piece of information. Figure 5(a) shows the relationship between C and t^* for our sample function $f(t) = e^{-t}$. As expected the larger the cost of obtaining an update, the longer one waits between updates. Also if the update cost is greater than the total value possibly gained by a copy over time, in this case, if $C > \int_{x=0}^{\infty} e^{-t}dt = 1$, then $t^* = \infty$ and of course no updates are made. Figure 5(b) shows the relationship between C, and the value gained per unit time using the optimal update times.

As we promised in Theorem 1, we now address the issue as to when Eq. (4) has a solution.



(a) Some Interesting Properties

(b) Curves for Different Costs C.

Fig. 4. Some Interesting Properties and Behavior.

We define

$$A \triangleq \lim_{t \to \infty} A(t) \tag{10}$$

In Figure 6 we show $D(t)$, $A(t)$ and A graphically. Note that $A(t)$ is the total value gained by having a new update from time 0 to time t. Further, A is the maximum gain by holding an update forever.

Using Eq. (6), we can rewrite Eq. (4) as

$$D(t^*) = A(t^*) - t^* f(t^*) = C. \tag{11}$$

We now inquire under what conditions Eq. (11) has a solution for t^*.

Theorem 2: We break this theorem into two parts:

Part a: If $A < \infty$, then t^* exists iff $A \geq C$.

Part b: If $A = \infty$, then t^* always exists.

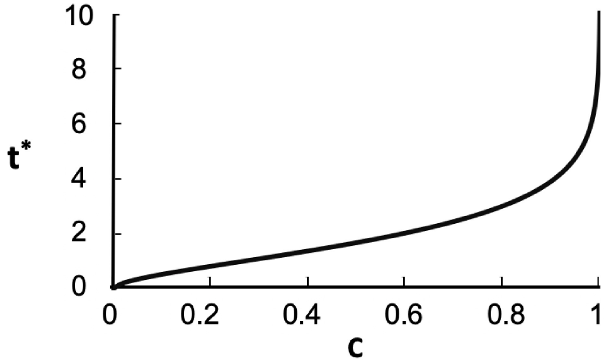
This Theorem is proven in the Appendix.

Let us consider two examples related to Theorem 2.

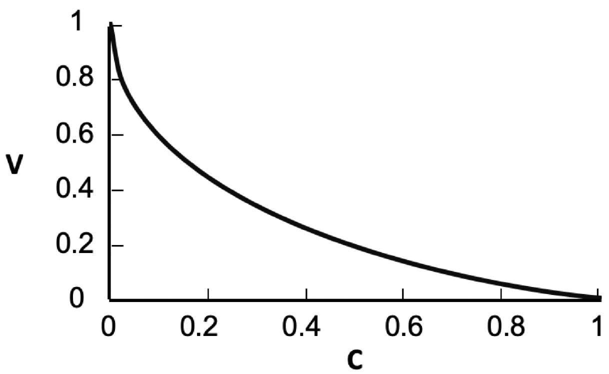
Example 2: We begin with a simple example illustrating *Part a* where $f(t) = ae^{-\rho t}$. For this function, we see that $A = a/\rho$. For cases where $a/\rho < \infty$ then Theorem 2 *Part a* applies and so there exists a finite t^* iff $A \geq C$.

In Figure 7 we show a particular case of Example 2 where we have selected the same parameter values as from Figure 2, namely, $a = 1$, $\rho = 1$ and $C = 0.2$. In this case, $A = 1$ and since $C = 0.2$, we have $A \geq C$ and $t^* = 0.825$.

Example 3: As another example, let us illustrate *Part b* by considering $f(t) = a/\sqrt{t}$. For this function, we see that



(a) Optimal Updates t^* as a Function of C.



(b) Maximum Expected Value Gained Per Unit Time as a Function of C.

Fig. 5. Variation with C for $f(t) = e^{-t}$.

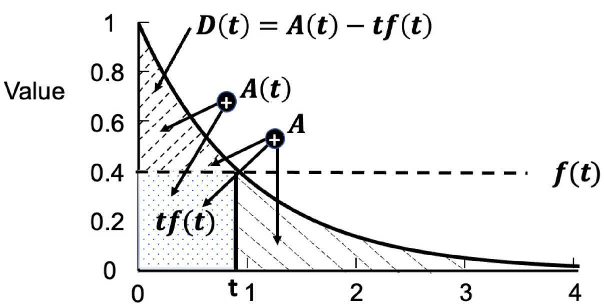


Fig. 6. Graphical representations of $D(t), A(t), A$.

$A(t) = 2a\sqrt{t}$. We now have $\lim_{t\to\infty} A(t) = \infty$, and so from Theorem 2 *Part b*, then t^* always exists. In this case, it is easy to find the exact value of t^* in terms of the two system parameters, namely, a and C as follows. From Eq. (11), we know that t^* must satisfy

$$A\left(t^*\right) - t^* f\left(t^*\right) = C,$$

substituting for $A(t)$ and $f(t)$ we get

$$2a\sqrt{t^*} - t^* a/\sqrt{t^*} = C,$$

and this gives

$$t^* = (C/a)^2,$$

and also

$$A\left(t^*\right) = 2C.$$

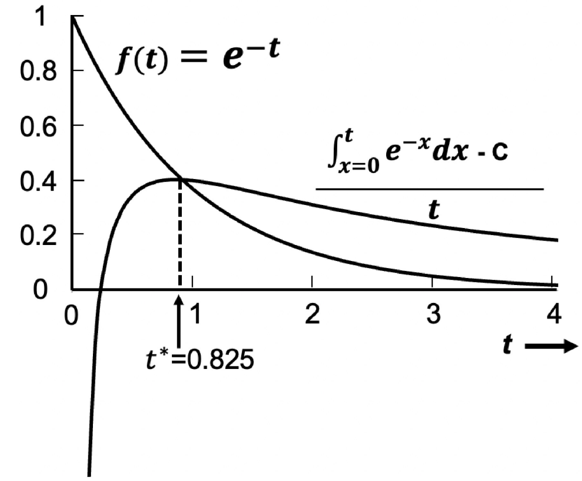


Fig. 7. Example 2 illustrating *Part a* of Theorem 2 for $f(t) = e^{-t}$ with $a = 1$, $\rho = 1$, $A = a/\rho = 1$, $C = 0.2$ and $t^* = 0.825$.

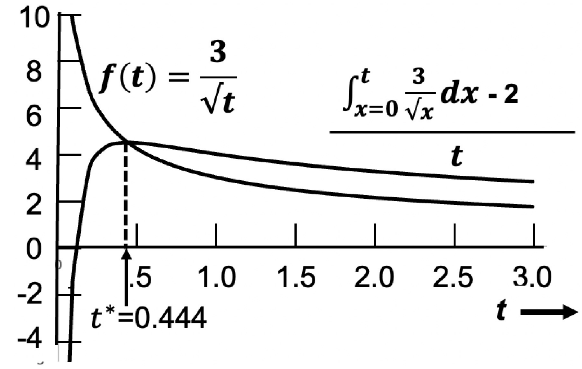


Fig. 8. Example 3 illustrating *Part b* of Theorem 2 for $f(t) = a/\sqrt{t}$ with $a = 3$, $C = 2$ and $t^* = (C/a)^2 = (2/3)^2 = 0.444$.

Moreover,

$$f\left(t^*\right) = a^2/C,$$

and so

$$t^* f\left(t^*\right) = C.$$

Recall that, under optimal updating, the net value gained per unit time is $f(t^*)$. Also $t^* f(t^*)$ is the net value that the system accumulates over each period which, in this example, is half of $A(t^*)$, the other half of which is also C, the amount that has been paid for the update. That is, as said earlier, it is always true that $A(t^*) - C = t^* f(t^*)$.

In Figure 8 we show a particular case of Example 3 where we have $a = 3$ and $C = 2$. In this case we have $t^* = (C/a)^2 = 0.444$.

In the following Theorem, we show an invariance of t^* to a vertical shift of $f(t)$ by a constant B.

Theorem 3: The value of t^* is invariant to any "shifting" of $f(t)$ by a (positive or negative) constant, say B.

Proof: We know that t^* is the value of t that satisfies Eq. (4). Rewriting that equation, we seek the value of t that satisfies

$$\int_{x=0}^{t} f(x)dx - C = tf(t). \tag{12}$$

Let us now consider finding t^* for the shifted function $f(t)+B$ which must satisfy the equation above, namely,

$$\int_{x=0}^{t} [f(x)+B]dx - C = t[f(t)+B],$$

or,

$$\int_{x=0}^{t} f(x)dx + Bt - C = tf(t) + Bt.$$

Note specifically that the terms involving B cancel out yielding

$$\int_{x=0}^{t} f(x)dx - C = tf(t).$$

But this is exactly the same as Eq. (12) which proves that t^* is the same for $f(t)$ and $f(t)+B$. ∎

Since we allow $f(t)$ to be monotonically non-increasing, we must consider the case where $f(t)$ remains "flat" at the constant value $f(t_0)$ in some region such as $0 \le t_0 \le t \le t_1$. Consider the following Theorem.

Theorem 4: Given that $f(t)$ is "flat" in a region, i.e., $f(t) = f(t_0)$ in the region $t_0 \le t \le t_1$, and if, in addition, we have that $t^* > t_0$, then t^* cannot occur in the flat interval $t_0 \le t \le t_1$.

Proof: Recall from Eq. (11) that the non-decreasing function $D(t) = A(t) - tf(t)$ must rise to the value C at which point, $t = t^*$. However, in the region $t_0 \le t \le t_1$ we have

$$A(t) = A(t_0) + f(t_0)(t - t_0),$$

and

$$tf(t) = t_0 f(t_0) + f(t_0)(t - t_0).$$

So

$$A(t) - tf(t) = A(t_0) + f(t_0)(t - t_0) - t_0 f(t_0) - f(t_0)(t - t_0),$$

or

$$A(t) - tf(t) = A(t_0) - t_0 f(t_0). \quad (13)$$

which is constant in the interval $t_0 \le t \le t_1$.

But since we know that $t^* > t_0$, then $A(t) - tf(t)$ has not yet reached C by time t_0. Moreover, from Eq. (13) we know that $D(t)$ is constant at the value $A(t_0) - t_0 f(t_0)$ in the region $t_0 \le t \le t_1$, and so it cannot possibly reach C in that flat interval. Hence, $t^* > t_1$. ∎

Note that Theorem 3 extends Assumption 4 to allow $\lim_{t\to\infty} f(t)$ to be any finite value. Specifically, if $\lim_{t\to\infty} f(t) = 0$ as in Assumption 4, then adding B to $f(t)$ gives us the new function $g(t) = f(t)+B$ and for this function, $\lim_{t\to\infty} g(t) = B$. Note further that Theorem 2 *Part a* requires $A \ge C$ in order for t^* to exist, yet when we consider $g(t)$, then A depends upon B. However, Theorem 3 shows that $g(t)$ and $f(t)$ have the same t^*. Thus the A we must compare to C (i.e., is $A \ge C$?) is the A for $f(t)$ (i.e., $B = 0$).

IV. Application to the Age of Information Metric

Note that the analysis in Section III was based on Assumption 3 which stated that $f(t) \ge 0$. Moreover, from Assumption 4 we assumed that $\lim_{t\to\infty} f(t) = 0$. In fact, we were never really constrained by $f(t) \ge 0$ and, further, we noted from Theorem 3 that we could also relax Assumption 4. Hence, in this Section IV, we relax both of those Assumptions.

Once we allow $f(t) < 0$ as well, we may extend our results to include the common Age of Information (AoI) metric. The AoI literature assumes a penalty for holding stale information. In many AoI studies that penalty grows linearly[7] as that information goes further out-of-date. As a simple way to model such AoI studies, we can set $f(t) = -t$. In this Section IV, we consider combining the case of positive values for information (as in Section III) mixed with negative values for information (as with AoI).

References [2], [3] are excellent introductions and summaries of AoI. Perhaps the simplest model of AoI consists of updates delivered by an update source to a destination. If an update is generated at time t_a, then at time $t \ge t_a$, the Age of Information (AoI) metric associated with that update at time t is $t - t_a$. The larger is the value of the AoI metric, the greater is the penalty for not updating. The simplest model for AoI we can generate in this paper (assuming $t_a = 0$) is

$$f(t) = -t. \quad (14)$$

The general literature of AoI considers far more complex AoI scenarios[8] as described in [2] and [19]. We choose to consider our simple model[9] (and some extensions below) to enable us to use the "simple" and intuitive optimization results from Section III and some more details as developed in this Section IV.

We now introduce and analyze some interesting AoI examples.

Example 4: We begin with the linear model presented in Eq. (14), namely $f(t) = -t$. We note that this is perhaps the simplest pure AoI metric where the negative value of $f(t)$ just gets linearly worse as t increases since AoI customarily penalizes the information value of $f(t)$ for its increasing staleness. For this case, we see that $A(t) = -(t^2)/2$. From Eq. (11) we see that in order to find t^*, we must solve the equation

$$A(t^*) - t^* f(t^*) = C. \quad (15)$$

Substituting for $A(t)$ and $f(t)$ we get

$$-(t^*)^2/2 + (t^*)^2 = C,$$

and this gives

$$t^* = \sqrt{2C},$$

and also

$$A(t^*) = -C.$$

[7]Non-linear cases are also studied in the AoI literature as in [4], [6], [20], [21], [22].

[8]For example, if updates are sent by the source repeatedly, then they may form a queue at the destination and the order in which they are used by the destination will determine the value of the information for that AoI scenario [23], [24]. In addition, the updates may arrive at the destination out of order due to random network delays [25]. Incorrect information may arrive from the source as investigated in [26]; this might occur if, for example, the communication channel is noisy.

[9]Note that our model contains no queues.

Moreover,

$$f\left(t^*\right) = -\sqrt{2C},$$

and so

$$t^* f\left(t^*\right) = -2C.$$

The net value gained per unit time over a cycle is $f(t^*)$. We note again that $t^* f(t^*)$ is the net value that the system accumulates over each period which, in this example, is $-2C$; of course it is pure negative since in this model, we are penalized for staleness (i.e., an amount $A(t^*) = -C$) and once we "pay" the additional amount of $-C$ for the update, we have a net "loss" of $-2C$ which is $t^* f(t^*)$ and is also $A - C$. Note further that $A(t) - tf(t) = t^2/2$ and so, from Eq. (1) we see that $A(t) - tf(t)$ will always exceed the cost C for some t and so t^* always exists.

In Figure 9(a) we show a specific case for Example 4 and in Figure 9(b) we show a specific case for the following Example 5.

Example 5: As another example, let us slightly generalize the linear model in Example 4 and consider the "shifted" AoI function

$$f(t) = a - bt.$$

For this function, we see that $A(t) = at - bt^2/2$. As earlier we must solve Eq. (15). Substituting for $A(t)$ and $f(t)$ we get

$$at^* - b\left(t^*\right)^2/2 - t^*\left(a - bt^*\right) = C,$$

and this gives

$$t^* = \sqrt{2C/b},$$

and also

$$A\left(t^*\right) = a\sqrt{2C/b} - C.$$

Moreover,

$$f\left(t^*\right) = a - \sqrt{2Cb},$$

and so

$$t^* f\left(t^*\right) = a\sqrt{2C/b} - 2C.$$

Again we note that since $A(t) - tf(t) = bt^2/2$, then there is always some value of t for which this function equals C and so t^* always exists (and equals $\sqrt{2C/b}$).

Example 5 shows us a combined case where we allow $f(t)$ to be positive (as in Section III) when $f(t) \geq 0$ in the range $0 \leq t \leq a/b$ as well as negative (i.e., stale information for the AoI metric) when $f(t) \leq 0$ in the range $a/b \leq t$.

We now consider a class of non-linear combined cases.

Example 6: We consider

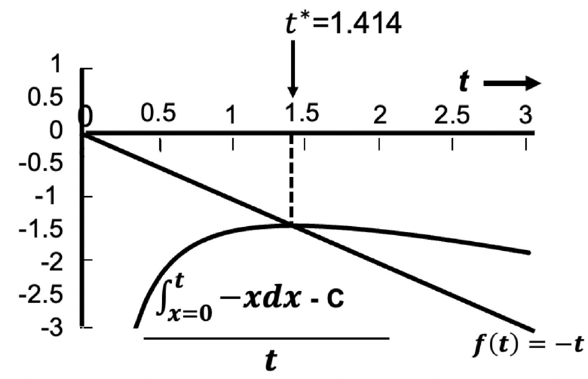
$$f(t) = a - bt^k.$$

For this function, we see that $A(t) = at - bt^{k+1}/(k+1)$. Once again, we substitute for $A(t)$ and $f(t)$ in Eq. (15) to get

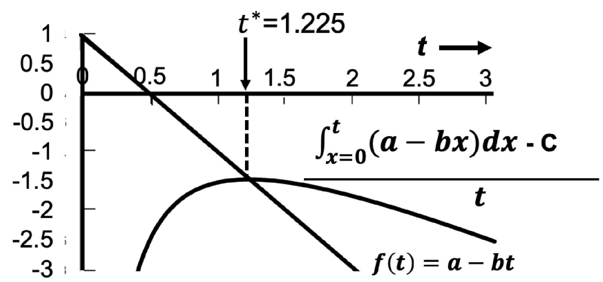
$$at^* - b\left(t^*\right)^{k+1}/(k+1) - t^*(a - b\left(t^*\right)^k = C,$$

and this gives

$$t^* = \sqrt[k+1]{(k+1)C/kb},$$



(a) Example 4 for $f(t) = -t$, $C = 1$ and $t^* = \sqrt{2C} = 1.414$.



(b) Example 5 for $f(t) = a - bt$, $a = 1$, $b = 2$, $C = 1.5$ and $t^* = \sqrt{2C/b} = 1.225$.

Fig. 9. Specific cases for Examples 4 and 5.

and also

$$A\left(t^*\right) = a\sqrt[k+1]{(k+1)C/kb} - C/k.$$

Moreover,

$$f\left(t^*\right) = a - b\left[(k+1)C/kb\right]^{k/(k+1)},$$

and so

$$t^* f\left(t^*\right) = a\sqrt[k+1]{(k+1)C/kb} - C(k+1)/k.$$

Again we note that since $A(t) - tf(t) = bt^{k+1}/(k+1)$, then there is always some value of t for which this function equals C and so t^* always exists.

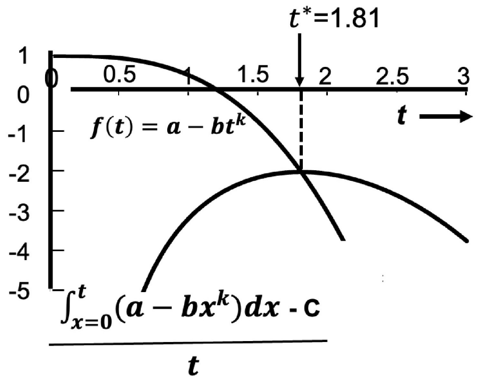
It is interesting to note that t^* is independent of the "shifting" parameter in Examples 5 and 6 above. In fact, Theorem 3 (where we used the general "shifting parameter" B) has proven this for all functions $f(t)$.

In Figure 10(a) we show a specific case for Example 6 and in Figure 10(b) we show a specific case for the following Example 7.

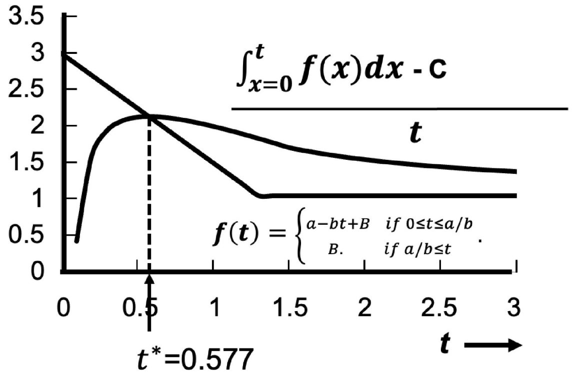
Example 7: We now consider a slightly generalized version of the linearized function shown in Example 5. Specifically we now consider

$$f(t) = \begin{cases} a - bt + B & \text{if } 0 \leq t \leq a/b \\ B & \text{if } a/b \leq t. \end{cases}$$

Note that we have added two features: (1) we have explicitly introduced the "shifting" parameter B where B is any finite value (positive or negative), and (2) we have frozen $f(t) = B$



(a) Example 6 for $f(t) = a - bt^k$, $a = 1$, $b = 0.5$, $k = 3$, $C = 4.$ and $t^* = 1.81$



(b) Example 7 for $f(t) = a - bt + B$, $a = 2$, $b = 1.5$, $B = 1$, $C = .25$, $A = a^2/2b = 1.33$ and $t^* = \sqrt{2C/b} = 0.577$.

Fig. 10. Specific cases for Examples 6 and 7.

when $a/b \le t$. For this function, we see that $A(t) = at - bt^2/2 + Bt$ in the range $0 \le t \le a/b$.

As earlier we must solve Eq. (15) to find t^*. However, we recall from Theorem 4, that if $t_0 < t^*$, then t^* cannot occur in the "flat" interval $t_0 \le t \le t_1$. In this example, the "flat" region is $(a/b) \le t$. Note further that this example is an example of Theorem 3 and recall that the "shifting" parameter B can be set to 0 and still get the same t^*; once we do that, we have an $f(t)$ that fits the conditions of Theorem 2 *Part a* and for t^* to exist, A must be such that $A \ge C$ where A is simply the area of the "triangle" for $B = 0$ in this example, i.e., $A = a^2/2b$.

Thus we seek to find t^* in the range $0 \le t \le a/b$. Substituting for $A(t)$ and $f(t)$ we get

$$at^* - b\left(t^*\right)^2/2 + B\left(t^*\right) - t^*\left(a - bt^* + B\right) = C,$$

and this gives

$$t^* = \sqrt{2C/b} \qquad \text{for } 0 \le t^* \le a/b.$$

If we check the condition on a, b and C for t^* to exist, we get

$$t^* = \sqrt{2C/b} \le a/b,$$

or

$$C \le a^2/2b.$$

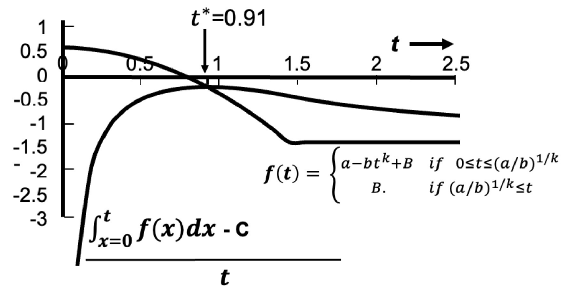


Fig. 11. Example 8 for $f(t) = a - bt^k + B$ for $0 \le t \le (a/b)^{1/k}$ and $f(t) = B$ for $(a/b)^{1/k} \le t$, $a = 2$, $b = 1$, $k = 2$, $B = -1.5$, $C = 0.5$, $A = 1.886$ and $t^* = 0.91$.

But we know that $A = a^2/2b$ and so the condition for t^* to exist is $A \ge C$. If $A < C$ then t^* does not exist, hence

$$t^* = \begin{cases} \sqrt{2C/b} & \text{if } C \le a^2/2b \\ \text{Does not exist} & \text{if } C > a^2/2b. \end{cases} \tag{16}$$

Example 8: We now generalize the case shown in Example 6 in the same way Example 7 extended Example 5. That is we introduce the "shifting" parameter B and we freeze $f(t) = B$ when $(a/b)^{1/k} \le t$.

$$f(t) = \begin{cases} a - bt^k + B & \text{if } 0 \le t \le (a/b)^{1/k} \\ B & \text{if } (a/b)^{1/k} \le t. \end{cases}$$

For this function, we see that $A(t) = at - bt^{(k+1)}/(k+1) + Bt$ in the range $0 \le t \le (a/b)^{(1/k)}$. As earlier we must solve Eq. (15) in this range. Substituting for $A(t)$ and $f(t)$ we get for $0 \le t^* \le (a/b)^{(1/k)}$

$$at^* - b\left(t^*\right)^{(k+1)}/(k+1) + Bt^* - t^*\left(a - b\left(t^*\right)^k + B\right) = C,$$

and if $t^* \le (a/b)^{(1/k)}$ this gives (where the argument for $(a/b)^{1/k} < t^*$ is the same as in Example 7)

$$t^* = \begin{cases} \sqrt[k+1]{(k+1)C/kb} & \text{if } 0 \le t^* \le (a/b)^{1/k} \\ \text{Does not exist} & \text{if } (a/b)^{1/k} < t^*. \end{cases}$$

As in Example 7, if $t^* > t_0$, then t^* cannot occur in the "flat" interval $t_0 \le t \le t_1$. In this example, the "flat" region is $(a/b)^{1/k} \le t$. Thus we conclude that if t^* is to occur at all, it must occur in the interval $0 \le t \le (a/b)^{(1/k)}$. Following the discussion in Example 7, the "shifting" parameter B can then be set to 0 in order to find A and then we see that the condition for t^* to exist is, as usual, $A \ge C$ where now the appropriate value of A is $A = ((ak/(k+1))(a/b)^{1/k}$. In Figure 11 we show a specific case for this Example 8.

It is interesting to see that for Examples 4, 5 and 7, the value of the optimum update time is the same, that is, $t^* = \sqrt{2C/b}$ and only Example 7 requires the condition $A \ge C$ (since the limit A is unbounded for the other two cases). Further, we note that both Examples 6 and 8 have the same value for t^* and only Example 8 requires the condition $A \ge C$. All of these statements regarding t^* hold regardless of the value of B. Finally note in all these examples, that t^* occurs at the maximum value of $(A(t) - C)/t$.

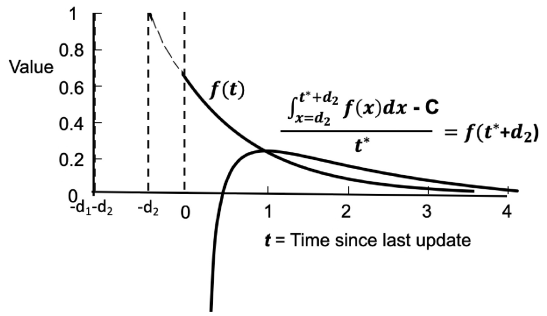


Fig. 12. Average Value Gained Per Unit Time With Transmission Delays.

V. Transmission Delays and Stale Updates

Here we assume that there is a transmission delay, d_1, for an update request to reach the information source, and a delay, d_2, for the information to be returned. When it returns, it is already d_2 time units old. Let t be the time between receiving updates. The problem becomes: Given d_1, d_2, and C, choose t to maximize the average value gained per unit time.

Figure 12 shows the cycle of an update with delays for a user. At time $-(d_1 + d_2)$, the update is requested. At time $-d_2$, the request reaches its destination, and the response is sent to the user. At time 0, the data reaches the user. Note that at time 0, the data is already d_2 time units old. Let $t - d_1 - d_2$ be the time the next update request is sent. At time t, the next update arrives at the user.

Using the same approach as above, we identify the objective function for our optimization to be the total net value gained between the receipt of two updates, namely,

$$\frac{-C + \int_{x=d_2}^{t+d_2} f(x)dx}{t}. \tag{17}$$

Taking the derivative of this objective function with respect to t shows this function is maximized when it equals $f(t+d_2)$ which is the value the old information will have when the update arrives. The formula for defining the optimal t is thus given by

$$\frac{-C + \int_{x=d_2}^{t^*+d_2} f(x)dx}{t^*} = f\big(t^* + d_2\big). \tag{18}$$

Note that in this case, the actual request is made at time $t^* - d_1 - d_2$. Thus it is important to know $f(t)$ ahead of time since it can sometimes be optimal to request an update before the previous update request has been satisfied.

In some cases we may choose an information source with a lower cost, but whose information may be out-of-date to begin with (such as a cache file). For this case, we assume no transmission delays. Let d be how far out-of-date the information source itself is. The model for transmission delays in this Section and this model for stale updates are essentially the same in that the information is already somewhat out-of-date by the time it reaches the user. Thus, the delay model and its equations herein can easily be applied to stale updates, and so we find that the optimal update time t^*, is the t that satisfies $\frac{-C + \int_{x=d}^{t^*+d} f(x)dx}{t^*} = f(t^* + d)$. The models for stale updates and for transmission delays are particularly useful in networks that have multiple sources of information and/or multiple paths to retrieve information each with its own independent cost and delay. The formulae we have derived can be used to determine the average value over time gained from each path to each source under optimal update times. The one with the highest value over time should be used.

VI. No Gain During Updates

Often, no value can be gained during the update process. We model this by introducing a new parameter L, the amount of time necessary to perform an update during which no value can be gained.

For example, in some ad hoc mobile radio networks, the information that needs updating are power gains, code assignments, connectivity, routing tables, and other information regarding the status of the network. The process of updating this network information is referred to as a global control phase. The accuracy of this information is crucial to the efficiency of the network. Since this algorithm is working in a mobile environment, this information will change over time. As it changes, the performance of the system will degrade until another global control phase is performed. However, the network may be unusable during global control phases (as in [1]). Obviously, if the radios cannot communicate, it doesn't matter how up-to-date the network information is, and no real value is gained by the network during the global control phases.

We now present a method for determining the optimal update times for problems of this type. As usual, let the monotonically decreasing function $f(t)$ be the instantaneous value of the system, given that the update process finished t seconds ago, and let C be the cost of the update (C can be 0). Let the amount of time necessary to perform an update be the constant $L > 0$.

What is the optimal amount of time, t^*, between completing the last update and starting the next one for this model? We now establish the following Theorem:

Theorem 5: A piece of information should be updated whenever $\frac{-C + \int_{x=0}^{t} f(x)dx}{L+t} \geq f(t)$ where t is the time since the last update.

Proof: Once again the optimal update times are chosen to maximize the average net gain per unit time. The total time between updates is now $(L+t)$, the amount of time to perform the update, L, plus the amount of time the system is allowed to run between updates, t. The net value gained between updates is $-C$ plus the integral of the instantaneous value of the information from the time it is received to time t, namely $-C + \int_{x=0}^{t} f(x)dx$. Thus, our objective function, namely, the average gain per unit time, is as follows:

$$\frac{-C + \int_{x=0}^{t} f(x)dx}{L+t}. \tag{19}$$

To find the maximum point, we take the derivative and set it to 0, namely

$$f(t)(L+t) + C - \int_{x=0}^{t} f(x)dx = 0. \tag{20}$$

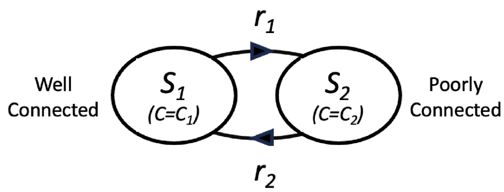


Fig. 13. The Mobility Model.

Simplifying yields

$$\frac{-C+\int_{x=0}^{t} f(x)dx}{L+t}=f(t). \tag{21}$$

■

This formula is the same as the one derived for the optimal scheduling of load balancing on a SIMD machine in [27] where $C=0$. The explanation of this formula is similar to that of the original formula in Section III. The update should be performed at the point where the instantaneous amount gained per unit time, $f(t)$, is equal to the average gained during the current update cycle from the time when the update began, which is $\frac{-C+\int_{x=0}^{t} f(x)dx}{L+t}$. Since $f(t)$ is monotonically non-increasing, either waiting longer or updating sooner will result in a lower expected gain per unit time over the course of the update cycle.

It turns out that these results have many far-reaching applications outside the field of information management. For example, suppose we own a factory and are able to judge the productivity of this factory as a function of the last time it was retooled, $f(t)$. We also know how much it will cost to retool, C, and the length of time for which operations will have to be shut down, L. The formula above will give the optimal times to retool a factory.

VII. Variable Update Costs

Let us now consider the case of dynamically changing update costs. This may be caused by a user moving around a network, a network that is unreliable, or using wired and wireless access alternately. Again we revert to the assumption that there are no transmission delays. We represent this changing cost by using a Markov process with two states, S_1 and S_2 which respectively represent well-connected and poorly-connected. The cost to update information from S_1 is C_1 and the cost from S_2 is C_2 ($C_1 < C_2$). We further assume that the user moves from state S_1 to S_2 at rate r_1, and moves from S_2 to S_1 at rate r_2. That is, the user movement among states is a Markov Process and spends independently chosen exponentially distributed amounts of time with means $1/r_1$ and $1/r_2$ in S_1 and S_2 respectively (see Figure 13).

A. Disconnected Users

We first look at the case where the user is totally disconnected in S_2. That is to say, no information can be accessed from S_2, or equivalently, the update cost C_2 is prohibitively high. Let t_1 be the amount of time since the last update beyond which the user will request a new update if in the connected state. We choose t_1 to maximize the total average value gained per unit time.

To find the optimal t_1, we use the same idea as earlier. We look at the expected value gained over the period from one update to the next, divide by the expected length of the period, and choose t_1 to maximize this value. This is a Markov renewal reward process [28] where we seek to find the expected value gained per unit time. This leads us to Equation (22) below which is a direct application of the Renewal Reward Theorem [28] to this problem.

No matter which state the user is in at time t_1, the value gained so far is $-C_1+\int_{x=0}^{t_1} f(x)dx$. The time taken is t_1. Since the user starts in S_1 at time 0, the probability of being in state S_1 at time t_1 is $\frac{r_2}{(r_1+r_2)}+\frac{r_1}{(r_1+r_2)}e^{-(r_1+r_2)t_1}$ [29]. The probability of being in S_2 at time t_1 is $\frac{r_1}{(r_1+r_2)}(1-e^{-(r_1+r_2)t_1})$.

If the system is in S_1 at time t_1, an update is immediately requested. The information arrives immediately and the cycle starts over again. However, if the user is in S_2 at time t_1, no update can be made immediately. But, as soon the user moves back to S_1, the update will be requested. This must be optimal for any non-increasing $f(t)$. The expected amount of time the user stays in S_2, if there at time t_1, is $\int_{x=0}^{\infty} e^{-r_2 x}dx$ which equals $1/r_2$. Thus the expected time between updates is $t_1+\frac{r_1}{(r_1+r_2)}(1-e^{-(r_1+r_2)t_1})(1/r_2)$.

The expected net value gained between updates equals minus the information cost plus the value gained up to time t_1, namely, $-C_1+\int_{x=0}^{t_1} f(x)dx$, plus the probability that the user is in S_2 at time t_1, times the expected value gained while the user stays in S_2 before returning to S_1. If the user is in S_1 time t_1, an update will be obtained immediately; otherwise, it will be obtained as soon as the user returns to S_1. If the user is in S_2 at time t_1, the expected total value gained before returning to S_1 and updating is $\int_{x=t_1}^{\infty} f(x)e^{-r_2(x-t_1)}dx = e^{r_2 t_1}\int_{x=t_1}^{\infty} f(x)e^{-r_2 x}dx$. We define this last integral as $I \triangleq \int_{x=t_1}^{\infty} f(x)e^{-r_2 x}dx$. Furthermore, for convenience we define the quantity J as $J \triangleq \frac{r_1}{(r_1+r_2)}(1-e^{-(r_1+r_2)t_1})$.

The complete formula for the expected value gained per unit time in this model is thus:

$$\frac{-C_1+\int_{x=0}^{t_1} f(x)dx + IJe^{r_2 t_1}}{t_1+J/r_2}. \tag{22}$$

In order to maximize this function we take its derivative and set it to 0, noting that if $(h/g)'=0$ then $\frac{g'}{g}=\frac{h'}{h}$ and $g'h=gh'$. We thus get

$$\left[f(t_1)(1-J)+\frac{r_1 I}{r_1+r_2}\left(r_2 e^{r_2 t_1}+r_1 e^{-r_1 t_1}\right)\right]\left(t_1+\frac{J}{r_2}\right) = \left[-C_1+\int_{x=0}^{t_1} f(x)dx+IJe^{r_2 t_1}\right]\left[1+\left(\frac{r_1}{r_2}\right)e^{-(r_1+r_2)t_1}\right]. \tag{23}$$

Thus an update should be performed when time $t=t_1$ satisfies this equation. Note that the more frequently one becomes disconnected and the longer one stays disconnected, the more frequently one should update when in the connected state.

B. Poorly-Connected Users

In this section we consider users who move at exponential rates r_1 and r_2, as shown in Figure 13, between a well-connected state, S_1, and, in this case, a poorly-connected state, S_2. We model well-connected and poorly-connected users by having different costs for obtaining updates from S_1 and S_2 which are C_1 and C_2 respectively, where $C_1 < C_2$. Assume that we make an update at $t = t_1$ when in S_1 and make an update at $t = t_1 + t_2$ when we are in S_2. As soon as we move from S_2 to S_1, we make an update immediately if $t \geq t_1$.

If it is optimal to perform an update from S_1 at the time since the last update, t_1, it is also optimal to perform an update from S_1 at any time greater than t_1. Thus the optimal strategy has the following form:

Never perform an update from either state until time t_1. If in state S_1 at time t_1, immediately perform an update. If in state S_2 at time t_1, then perform an update as soon as you transition from S_2 to S_1 for any time up to t_1+t_2. If no transitions to S_1 have occurred by time $t_1 + t_2$, then perform an update from S_2 at time $t_1 + t_2$.

Note that this strategy makes no use of the state in which the last update occurred and only uses knowledge of the current state. This is because the movement between states is a Markov Process, thus knowing the current state renders the knowledge of historic movement between states and, in particular, knowing the state from which the last update occurred, irrelevant for predicting the future.

We wish to develop a formula that gives the expected gain per unit time for given values of t_1 and t_2. We first compute the steady state probabilities, P_1 and P_2, of performing an update from S_1 and S_2 respectively for different values of t_1 and t_2.

Each update occurs from either S_1 or S_2. If the last update occurred in S_1, let $p_{1,1}$ and $p_{1,2}$ be the probabilities that the next update occurs from S_1 and S_2, respectively (where $p_{1,1}+ p_{1,2} = 1$). Similarly, let $p_{2,1}$ and $p_{2,2}$ be the probabilities that the next update occurs from S_1 and S_2, respectively, given that the last update occurred from S_2 (where $p_{2,1}+ p_{2,2} = 1$).

Now, to compute $p_{1,1}, p_{1,2}, p_{2,1}$ and $p_{2,2}$, assume the last update occurred from state S_1. The probability, $p_{1,2}$, that the next update occurs in state S_2 is the probability that the user is in state S_2 at time t_1, times the probability that no state transitions occur between time t_1 and $t_1 + t_2$, that is

$$p_{1,2} = \frac{r_1}{r_1 + r_2}\left(1 - e^{-(r_1+r_2)t_1}\right)e^{-r_2 t_2}. \tag{24}$$

Similarly, $p_{2,2}$ equals the probability that the user is in state S_2 at time t_1 given that the last update occurred from S_2, times the probability that the user does not become well-connected between time t_1 and $t_1 + t_2$, that is

$$p_{2,2} = \left[\frac{r_1}{r_1 + r_2} + \frac{r_2}{r_1 + r_2}e^{-(r_1+r_2)t_1}\right]e^{-r_2 t_2}. \tag{25}$$

Recall that $p_{1,1} = 1 - p_{1,2}$ and that $p_{2,1} = 1 - p_{2,2}$.

Now we compute P_1 and $P_2 = 1 - P_1$, the two steady-state probabilities for the last update having occurred in S_1 and S_2, respectively. We have

$$P_2 = P_1 p_{1,2} + P_2 p_{2,2}.$$

Recalling that $P_1 = 1 - P_2$ yields

$$P_2 = \frac{p_{1,2}}{p_{1,2} + 1 - p_{2,2}}.$$

Some algebraic substitutions yield

$$P_2 = \frac{1 - e^{-(r_1+r_2)t_1}}{1 - e^{-(r_1+r_2)t_1} + \frac{r_1 + r_2}{r_1}e^{r_2 t_2} - 1 - \frac{r_2}{r_1}e^{-(r_1+r_2)t_1}}. \tag{26}$$

Now that we have the steady-state probabilities, we are in a position to create an equation for the value gained per unit time under the strategy defined by t_1 and t_2. To derive this equation, we assume that the last update was from S_1 with probability P_1 and from S_2 with probability P_2. Given the definition of the steady state, we know that the next update will be from S_1 with probability P_1 and from S_2 with probability P_2. Because of this, the expected net value gained per unit time over a single update cycle is the overall expected value gained per unit time for the strategy. It is convenient to first define R to be the steady-state probability that we are in S_2 at time t_1, which yields

$$\begin{aligned} R = {} & P_1 \frac{r_1}{r_1 + r_2}\left(1 - e^{-(r_1+r_2)t_1}\right) + \\ & P_2\left[\frac{r_1}{r_1 + r_2} + \frac{r_2}{r_1 + r_2}e^{-(r_1+r_2)t_1}\right]. \end{aligned} \tag{27}$$

The expected time between updates is thus: $t_1 + R\int_{x=0}^{t_2} e^{-r_2 x}dx$. The expected cost of an update is $P_1 C_1 + P_2 C_2$. The expected value gained from the information between updates is $\int_{x=0}^{t_1} f(x)dx + R\int_{x=0}^{t_2} e^{-r_2 x} f(x + t_1)dx$. Finally, the average value gained per unit time (using the Renewal Reward Theorem [28]) is thus

$$\frac{-P_1 C_1 - P_2 C_2 + \int_{x=0}^{t_1} f(x)dx + R\int_{x=0}^{t_2} e^{-r_2 x} f(x + t_1)dx}{t_1 + R\int_{x=0}^{t_2} e^{-r_2 x}dx}. \tag{28}$$

The optimal values of t_1 and t_2 are found by maximizing this equation with respect to t_1 and t_2. Note that the more frequently one becomes poorly-connected and the longer one stays poorly-connected, the more frequently one should update when in the well-connected state. Furthermore, the more frequently one moves from the poorly-connected state, the less frequently one should update in the poorly-connected state.

VIII. Conclusion

We have presented methods for finding the optimal update times for out-of-date information. To solve this problem we introduced a new way of looking at the value of information as a function, $f(t)$, of how far out-of-date information is. These methods are very general, and work for any arbitrary non-increasing $f(t)$, and have intuitive and graphical simplicity. Using this we have found optimal update times for a number of simple cases, including some that use the Age of Information metric. We further solved the model in which there is a delay between an information request being made and that information arriving. Next we used the same ideas to present a methodology for determining optimal update times when no value can be gained from information during the update process. We also presented a more general model in which there is an explicit cost of an update and in which useful work must cease while the update is performed. We extended this to a case where the user becomes intermittently disconnected, and a case where the user moves from well-connected to poorly-connected. All these results make no assumptions about the shape of $f(t)$ except that it must be monotonically non-increasing.

Appendix

Proof of Theorem 2

Proof (Part a): Rewriting Eq. (11) we see that we obtain t^* if the following equation is solvable:

$$D(t^*) = A(t^*) - t^* f(t^*) = C. \tag{A.1}$$

From Eq. (6) we have

$$\begin{aligned} D(t) &= A(t) - tf(t), \\ &= \int_{x=0}^{t} f(x)dx - tf(t), \\ &= \int_{x=0}^{t} [f(x) - f(t) + f(t)]dx - tf(t), \\ &= \int_{x=0}^{t} [f(x) - f(t)]dx + tf(t) - tf(t), \\ &= \int_{x=0}^{t} [f(x) - f(t)]dx. \end{aligned}$$

But by Assumption 3 in Section II, the integrand $f(x) - f(t) \geq 0$ for $x < t$ which shows that

$$D(t) \geq 0. \tag{A.2}$$

Taking the limit of $D(t)$ as $t \to \infty$, we have

$$\begin{aligned} \lim_{t\to\infty} D(t) &= \lim_{t\to\infty} A(t) - \lim_{t\to\infty} tf(t) \\ &= A - \lim_{t\to\infty} tf(t). \end{aligned} \tag{A.3}$$

We now set out to prove for this Part a that if $A < \infty$, then, $\lim_{t\to\infty} tf(t) = 0$.

Let $g(t) = tA(t)$. Consider $\lim_{t\to\infty} g(t)/t$ which naively can be considered to be indeterminant. So let's use L'Hopital's rule to give

$$\lim_{t\to\infty} g(t)/t = \lim_{t\to\infty} \frac{(d/dt)g(t)}{(d/dt)t}. \tag{A.4}$$

In general, L'Hopital's rule requires the $\lim_{t\to\infty}(d/dt)g(t)$ to exist. From its definition, we have

$$\begin{aligned} \lim_{t\to\infty} (d/dt)g(t) &= \lim_{t\to\infty} (d/dt)tA(t), \\ &= \lim_{t\to\infty} tdA(t)/dt + \lim_{t\to\infty} A(t), \\ &= \lim_{t\to\infty} tf(t) + A. \end{aligned}$$

However, from Figure 6 we clearly see that $tf(t) < A(t)$ and so we see that $\lim_{t\to\infty} tf(t) < \lim_{t\to\infty} A(t) = A < \infty$. Hence we have shown that $\lim_{t\to\infty}(d/dt)g(t) < 2A < \infty$ and so $\lim_{t\to\infty}(d/dt)g(t)$ exists. The left-hand side of Eq. (A.4) gives us

$$\lim_{t\to\infty} g(t)/t = \lim_{t\to\infty} A(t) = A, \tag{A.5}$$

and the right-hand side of Eq.(A.4) gives us

$$\lim_{t\to\infty} \frac{(d/dt)g(t)}{(d/dt)t} = \lim_{t\to\infty} \frac{tf(t) + A}{(d/dt)t}, \tag{A.6}$$

Equating Eq. (A.5) and Eq. (A.6) gives us

$$A = \lim_{t\to\infty} tf(t) + A,$$

and hence we have established that if $A < \infty$, then we have

$$\lim_{t\to\infty} tf(t) = 0. \tag{A.7}$$

Applying Eq. (A.7) to Eq. (A.3) we see that

$$\lim_{t\to\infty} D(t) = A. \tag{A.8}$$

Clearly $D(0) = 0$ and from Eqs. (A.2) and (7) $D(t)$ cannot decrease as t increases and its limit is A as given in Eq. (A.8). Now, if $A \geq C$, then $D(t)$ must reach the value C on its way to its limit A and the time t when this occurs is, by Eq. (A.1), equal to t^*. If $A < C$, then clearly $A(t)$ will never reach C so t^* will not exist. Thus *Part a* is proven. ■

Part b: We seek to show that Eq. (A.1) is always satisfied if $\lim_{t\to\infty} A(t) = \infty$. From Eq. (6) and Eq. (A.2) we see that

$$A(t) \geq tf(t). \tag{A.9}$$

We now examine the slope of both sides of Eq. (A.9). By definition, the differential of the left-hand side of the inequality is clearly $dA(t)/dt = f(t)$. Differentiating the right-hand side of this inequality we get

$$d(tf(t))/dt = df(t)/dt + f(t).$$

Again, by Assumption 3 in Section II, we know that $df(t)/dt \leq 0$ and so

$$dA(t)/dt = f(t) = d(tf(t))/dt - df(t)/dt \geq d(tf(t))/dt.$$

This establishes that

$$dA(t)/dt \geq d(tf(t)/dt. \tag{A.10}$$

From its definition in Eq. (6) we see that $D(t)$ is the difference between $A(t)$ and $tf(t)$. In Eq. (A.10) we have shown that the slope for $A(t)$ is greater than the slope for $tf(t)$ (except in those intervals where $f(t)$ is constant), so these two curves must diverge as t increases. Now we have assumed that $\lim_{t\to\infty} A(t) = \infty$ and so $A(t)$ grows without limit. That means that $D(t)$ will grow without limit (unless f(t)=constant for all t) and so $D(t)$ will eventually reach the value for any finite C at some value of $t = t^*$ which will then satisfy Eq. (A.1), and so *Part b* is proven. ■

REFERENCES

[1] C. Ferguson, "The virtual network algorithm," Ph.D. dissertation, Dept. Comput. Sci., Univ. California, Los Angeles, CA, USA, 1999.

[2] R. Yates, Y. Sun, D. R. Brown, S. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE J. Sel. Areas Commun*. vol. 39, no. 5, pp. 1183–1210, May 2021.

[3] A. Kosta, N. Pappas, and V. Angelakis, *Age of Information: A New Concept, Metric, and Tool*, vol. 12. Delft, The Netherlands: Now Publ. Inc., 2017, pp. 162–259.

[4] X. Wu, J. Yang, and J. Wu, "Optimal status update for age of information minimization with an energy harvesting source," *IEEE Trans. Green Commun. and Netw.*, vol. 2, no. 1, pp. 193–204, Mar. 2018.

[5] B. T. Bacinoglu, E. T. Ceran, and E. Uysal-Biyikoglu, "Age of information under energy replenishment constraints," *Inf. Theory Appl. Workshop*, 2015, pp. 25–31.

[6] B. T. Bacinoglu, Y. Sun, E. Uysal, and V. Mutlu, "Optimal status updating with a finite-battery energy harvesting source," *J. Commun. Netw.*, vol. 21, no. 3, pp. 280–294, Jun. 2019.

[7] S. Chen and L. Du, "Simulation study of the impact of local real-time traffic information provision strategy in connected vehicle systems," *Int. J. Transp. Sci. Technol.*, vol. 6, no. 4, pp. 229–239, Dec. 2017.

[8] D. S. Stamoulis and P. A. Giannopoulos, ""My personal forecast": The digital transformation of the weather forecast communication using a fuzzy logic recommendation system," *Adv. Sci. Res.*, vol. 19, pp. 9–12, Mar. 2022.

[9] Y. C. Hu and D. B. Johnson, "Ensuring cache freshness in on-demand Ad Hoc network routing protocols," in *Proc. 2nd ACM Int. Workshop Princ. Mobile Comput.*, 2002, pp. 25–30.

[10] S. Upadhyay and H. Hu, "A qualitative analysis of the impact of electronic health records (EHR) on healthcare quality and safety: Clinicians' lived experiences," *Sage Publ., Health Serv. Insights*, vol. 15, pp. 1–7, Mar. 2022.

[11] Y. Ren, F. Liao, and Y. Gong "Impact of news on the trend of stock price change: An analysis based on the deep bidirectional LSTM model," *Procedia Comp. Sci.*, vol. 174, pp. 128–140, Oct. 2020.

[12] M. Xiong and K. Ramamritham, "Deriving deadlines and periods for real-time update transactions," in *Proc. 20th IEEE Real-Time Syst. Symp.*, 1999, pp. 32–43.

[13] J. S. Kim, M. G. Kim, and S. B. Pan, "A study on implementation of real-time intelligent video surveillance system based on embedded module," *J. Image Video Process.*, vol. 2021, no. 1, pp. 1–22, Nov. 2021.

[14] H. Yu, L. Breslau, and S. Shenker, "A scalable Web cache consistency architecture," *SIGCOMM Comput. Commun. Rev.*, vol. 29, no. 4, pp. 163–174, Aug. 1999.

[15] J. Cho, and H. Garcia-Molina, "Effective page refresh policies for Web crawlers," *ACM Trans. Database Syst. (TODS)*, vol. 8, no. 4, pp. 390–426, Dec. 2003.

[16] A. Reizenstein, "Position and trajectory control of a quadcopter using PID and LQ controllers," M.S. Thesis, Dept. Electr. Eng., Linköping Univ., Linköping, Sweden, 2017.

[17] F. Hendrikx, K. Bubendorfer, and R. Chard, "Reputation systems: A survey and taxonomy," *J. Parallel Distrib. Comput.*, vol. 75, pp. 184–197, Jan. 2015.

[18] Z. Liu, B. Li, Z. Zheng, Y. T. Hou, and B. Ji, "Towards optimal tradeoff between data freshness and update cost in information-update systems," in *Proc. Int. Conf. Comp. Commun. Netw.*, 2022, pp. 1–10.

[19] Y. Sun, I. Kadota, R. Talak, and E. Modiano, *Age of Information: A New Metric for Information Freshness* (Synthesis Lectures on Learning, Networks, and Algorithms). San Rafael, CA, USA: Morgan & Claypool, 2020, pp. 1–204.

[20] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, "Age and value of information: Non-linear age case," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, 2017, pp. 326–330.

[21] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, "Non-linear age of information in a discrete time queue: Stationary distribution and average performance analysis," in *Proc. ICC IEEE Int. Conf. Commun. (ICC)*, 2020, pp. 1–6.

[22] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Trans. Inf. Theory*, vol. 63, no. 11, pp. 7492–7508, Nov. 2017.

[23] A. Kosta, N. Pappas, and V. Angelakis, *Age of Information: A New Concept, Metric, and Tool*. Delft, The Netherlands: Now Publ. Inc., 2017, pp. 162–259.

[24] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE INFOCOM*, 2012, pp. 2731–2735.

[25] C. Kam, S. Kompella, and A. Ephremides, "Age of information under random updates" in *Proc. IEEE Int. Symp. Inf. Theory*, 2013, pp. 66–70.

[26] A. Maatouk, S. Kriouile, M. Assaad, and A. Ephremides "The age of incorrect information: A new performance metric for status updates," *IEEE/ACM Trans. Netw.*, vol. 28, no. 5, pp. 2215–2228, Oct. 2020.

[27] C. Powley, C. Ferguson, and R. Korf, "Depth-first heuristic search on a SIMD machine," *Artif. Intell.*, vol. 60, no. 2, pp. 199–242, Apr. 1993.

[28] Renewal Reward Processes. "Section 14 part 6 in probability, mathematical statistics, stochastic processes." Accessed: 2023. [Online]. Available: https://www.randomservices.org/random/renewal/Reward.html

[29] L. Kleinrock, *Queueing Systems Volume I: Theory*. Hoboken, NJ, USA: Wiley, 1975.

Chris Ferguson received the B.S. degree from the University of California at Los Angeles (UCLA) in 1986, and the Ph.D. degree in computer science specializing in artificial intelligence from UCLA in 1999, under the tutelage of his coauthor Prof. L. Kleinrock. He has been an avid poker player and a Game Theorist his entire life, and worked with his father UCLA Professor T. Ferguson on optimal strategies for the game of poker. These studies greatly contributed to his success in the world of poker, including his victory in the main event of the World Series of Poker in 2000. He is currently working for RP1, a company he founded with other UCLA alumni. RP1 is developing the technologies to enable a "True" metaverse where the entire population of the world can share and interact in a universe-sized virtual environment. This is considered by many to be the natural progression of the Internet.

Leonard Kleinrock (Life Fellow, IEEE) received the B.E.E. degree from the City College of New York in 1957, and the M.S. and Ph.D. degrees from the Massachusetts Institute of Technology (MIT) in 1959 and 1963, respectively.

He is a Distinguished Professor of Computer Science with the University of California at Los Angeles (UCLA). He is considered a father of the Internet, having developed the mathematical theory of data networks, the technology underpinning the Internet as an MIT graduate student in 1962. His UCLA Host computer became the first node of the Internet in September 1969 from which he directed the transmission of the first Internet message. He has served as a Professor of Computer Science with UCLA since then, and was the Department Chairman from 1991 to 1995. He has received eight honorary degrees, has published over 250 papers, authored six books, and has supervised the research for over 50 Ph.D. students on a wide array of subjects, including packet switching networks, packet radio networks, local area networks, broadband networks, queueing theory, congestion control, gigabit networks, nomadic computing, intelligent software agents, performance evaluation, peer-to-peer networks, and blockchain performance. He received the 2007 National Medal of Science, the highest honor for achievement in science bestowed by the President of the United States. Among his many honors, he is the recipient of the National Medal of Science, the Ericsson Prize, the NAE Draper Prize, the Marconi Prize, the Dan David Prize, the Okawa Prize, the BBVA Frontiers of Knowledge Award, the ORSA Lanchester Prize, the ACM SIGCOMM Award, the IEEE Leonard G. Abraham Prize Paper Award, the IEEE Harry M. Goode Award, and the IEEE Alexander Graham Bell Medal. He is a member of the National Academy of Engineering and the American Academy of Arts and Sciences, an ACM Fellow, an INFORMS Fellow, a CHM Fellow, an IEC Fellow, an Inaugural Member of the Internet Hall of Fame, a Guggenheim Fellow, and an Eminent Member of Eta Kappa Nu.