Towards an Understanding of EASE and its Properties

Stratis Ioannidis and Peter Marbach
University of Toronto
Department of Computer Science
10 King's College Road, Toronto, ON, Canada
stratis,marbach@cs.toronto.edu

Abstract

We propose a model under which several inherent properties of the Exponential Age SEarch routing protocol can be derived. By making simplifications on this model, we are able to address the issue of the optimality of a parameter of the protocol and to improve the existing upper bound on its performance.

1 Introduction

Routing in ad hoc networks is a challenging task for several reasons. Due to the absence of a hierarchy and a central authority, routing must be performed in a distributed manner; that is, nodes should decide where to forward an incoming packet based only on information maintained locally and information contained in the packet to be routed. Furthermore, the fact that nodes are mobile leads to frequent and abrupt changes in the topology of the network. This makes maintaining accurate information on current topology quite costly.

The Exponential Age SEarch protocol (EASE), proposed by Grossglauser and Vetterli [4], is a novel routing mechanism that uses constrained flooding and does not incur communication overhead for route maintenance. In this protocol, each node maintains in a table the position of each destination at some time the past. This serves as an *estimation* of where the destination currently is. The elapsed time period since the moment that the destination was in this position is also stored and serves as a *measure of accuracy* for this estimation. A source can route a packet using this position information; if the destination is not reached thus, EASE resorts to flooding. More specifically, a node whose estimate of the position is twice as good as the current one (*i.e.* the elapsed time is half) is sought out, and its estimate of the position of the destination can be used for a new forwarding attempt.

Another interesting property of EASE is how frugal the update mechanism it employs is: nodes update their entries every time a destination is within their transmission radius, *i.e.* when two nodes *encounter* each other while moving. This means that no maintenance overhead is incurred and that network mobility plays an important role in the dissemination of information.

There are several questions one would like to answer regarding EASE. First of all, as noted, Grossglauser *et al.* specify that a node containing information that is twice as accurate is sought out when flooding. It is not clear whether some other flooding condition would be better. For example, one may flood looking for *any* node that has better information than the current one, or impose a more stringent condition, like flooding for a node that has thrice as accurate information about the destination than the current one. Another issue is computing the total overhead incurred by the protocol. Since there is no maintenance overhead, this amounts to the cost incurred during flooding.

An analysis that addresses these issues is quite hard. Grossglauser *et al.* propose a model under which, using a compelling argument, they upper-bound the flooding cost. However, they obtain the upper bound by making certain simplifications in their analysis. Furthermore, they do not investigate the optimality of the flooding condition.

In this paper, we propose a new mathematical model for analyzing EASE. This model can be seen as a transfer of the model by Grossglauser *et al.* from the discrete to the continuous realm. Although certain properties of EASE can be derived directly from this model, answering the questions we posed above is still hard. Nevertheless, by making similar simplifications to the ones employed by Grossglauser *et al.* we are able to show several results. First of all, we show that, in our simplified model, the optimal choice while flooding is to look for *any* node that met the destination more recently, instead of only the ones that have estimates that are twice as accurate as the current estimate. We then show that a tighter bound than the one proved by Grossglauser *et al.* can be obtained in our simplified model.

2 The EASE Protocol

In this section we give a more precise definition of EASE. The algorithm can be found in Figure 1. Each node i maintains a routing table RT_i and an elapsed time table T_i. Let $X_i(t)$ denote the position of node i at time t. For any node j, RT_{ij} contains the coordinates of j at the time of the last encounter between i and j. If at time 0 the elapsed time since the last encounter between i and j is t, then the routing table of i will be $RT_{ij} = X_j(-t)$. The elapsed time table entry will be none other than $T_{ij} = t$.

Routing happens in the above setting as follows. Suppose that a source s wishes to send a packet to a destination d. Initially, the source node s can send it to the position $X_d(-T_{sd})$, using any of the location-based forwarding strategies that exist in literature. If the destination is not reached with this forwarding, either because of a forwarding failure or because it is not there anymore, the protocol resorts to flooding. While flooding, the protocol looks for any node n such that $T_{nd} < \frac{T_{sd}}{2}$. After locating such a node n, the protocol resumes forwarding using the position information $X_d(-T_{nd})$. This process is repeated and the protocol alternates between forwarding and flooding until the destination is reached. An illustration of this behavior can be seen in Figure 2.

Route a packet p from s to d
{
    $i = s$; forward p towards d using $X_d(-T_{id})$;
    let a be the node reached with this forwarding;
    while $(a \neq d)$ {
        flood the network from a until a node n
            is reached such that $T_{nd} < \frac{T_{id}}{2}$;
        $i = n$;
        forward p towards d using $X_d(-T_{nd})$;
        let a be the node reached with this forwarding;
    }
}

Figure 1. The EASE algorithm

Grossglauser *et al.* use the term *anchor points* for the points a in which forwarding fails and flooding is initiated, and the term *messenger* nodes for the nodes n that satisfy $T_{nd} \leq \frac{T_{sd}}{2}$. We assume that flooding initiated at a is restricted to the nodes within a distance from a less than the one between a and n. This is plausible if for example flooding is implemented with an increasing TTL mechanism, *i.e.* if consecutive floods with increasing number of hops occur until a (messenger) node replies to the anchor point.

The choice of $T_{nd} < \frac{T_{sd}}{2}$ as the condition for which a node is considered a messenger node is not an apparent one. One may in fact impose any flooding condition of the form $T_{nd} < \gamma T_{sd}$, where $0 < \gamma \leq 1$. Intuitively, small values of γ should make it harder to locate a messenger node, thus increasing the per-step flooding area. On the other hand, a small γ would mean that the protocol makes larger estimation improvements per step and hence fewer flooding steps in total are needed to reach the destination. The optimal value of γ is therefore not obvious. In what follows, we will assume that EASE does not have a specified value γ, and computing its optimal value will be one of our goals.

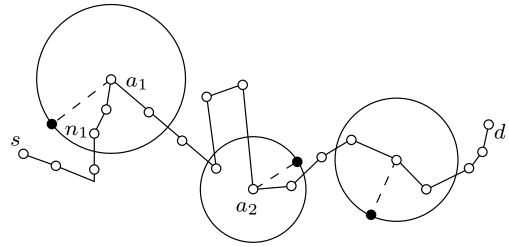

Figure 2. An illustration of the execution of the EASE algorithm.

3 Related Work

Technically, EASE is a *position-based* routing protocol (such as GPSR [6]), since the information used to route a packet to a destination is its position on the plane. Typically position-based protocols consist of a *forwarding strategy*, which specifies how a packet is routed given the destination's position, and a *location service*, which describes how the destination's position can be obtained. EASE combines these two by introducing inaccuracy: each node has its own approximate location service, namely its routing table, whose information is gradually refined while routing.

On the other hand, EASE can also be seen as a hybrid protocol that combines features of both *proactive* or *table-driven* (*e.g.* DSDV [8]) and *reactive* or *demand-driven* protocols (*e.g.* AODV [9], DSR [5]). As in proactive protocols, nodes maintain routing tables; however, the information in them is inaccurate, and no maintenance overhead is thus required. Moreover, flooding is used to dynamically improve the currently available routing information, as in reactive protocols.

The original paper by Grossglauser *et al.* [4] contains an analytical study of EASE, as described in the introduction. Sarafijanovic-Djukic *et al.* [11] study EASE under the random waypoint mobility model. They introduce a modification of EASE that improves the estimate of the destination's position, given that the nodes move according to the random waypoint model, and demonstrate the relative improvement on protocol performance through simulations. FRESH [2] is a simplification of EASE. Its main difference is that nodes maintain only the accuracy tables T_i; the messenger's current position serves as an estimate for the destination's position. The authors use simulation results to argue that, as the protocol improves its accuracy, it also progresses in space and gets closer to the destination. GREP [3] uses next-hop instead of position information and piggy-backing instead of last encounters as an update mechanism. To our knowledge, there is no performance analysis on GREP in existing literature. A proof of its loop freedom can be found in [3].

4 Model

In [4], nodes performed a two-dimensional random walk on an $M \times M$ grid. In our model, nodes perform two-dimensional Brownian motions on the entire plane. A Brownian motion can be seen as the limit of a random walk making infinitely many steps of infinitesimal length.

The network in our model consists of an infinite, countable number of nodes. As in section 2, $X_i(t) \in \mathbb{R}^2$ will denote the position of i at time t. Furthermore, $\boldsymbol{\pi}(t) = \{X_1(t), X_2(t), \dots\}$ will denote the set of all nodes in the network. By definition, $\boldsymbol{\pi}(0)$ is a Poisson field of density ρ spanning over the entire plane $\mathbb{R}^2$, *i.e.* nodes are uniformly distributed in it with density ρ. Each node i moves independently according to a two-dimensional Brownian motion of variance proportional to v. An encounter between nodes i and j occurs at time t if their distance is less than a transmission radius r_0, *i.e.* if $|X_i(t) - X_j(t)| \leq r_0$.

We assume that nodes have been moving in the interval $(-\infty, 0]$ and at time $t = 0$ a node initiates a route discovery to the destination d. Note that since nodes have been moving for an infinite amount of time every node has encountered d with probability 1. As in [4], route discoveries are considered to last for a negligible period of time, so that nodes can be considered static while they take place. In addition, we assume that the node density ρ is large enough so that the network is connected and that a packet can always be forwarded successfully from one anchor point to another.

The parameters of the network are ρ and v. The length unit is assumed scaled in such a way that the transmission radius is $r_0 = 1/\sqrt{\pi}$, *i.e.* the disk defining the neighborhood of a node has unit area. The density ρ of the network has an intuitive influence on protocol performance: the denser the network, the more nodes a destination encounters while moving and the less a source would need to flood in order to find messenger nodes. The Brownian motion variance parameter v is directly linked to how fast nodes move, which also has an intuitive effect on the behavior of the protocol. It seems that the faster a destination moves, the harder it should be to locate; however, a fast destination should also encounter more nodes, generating thus more messengers. We will revisit these issues in Section 6.2.

To answer the questions posed in the introduction, one needs to introduce a cost as a performance metric under which protocol behavior can be evaluated. The cost that we adopt is the total area flooded by the protocol. In defining this, we use the following notation. We denote with $G(T)$ the expected area around an anchor point flooded in one step, given that the anchor point is the position of the destination at time T ago. We will call $G(T)$ the *expected one-step flooding area*. Furthermore, we denote with $p(t, T)$ the density of the accuracy t achieved though flooding conditioned on the fact that the anchor point is the position of the destination at time T ago. In other words, $t = T_{nd}$ is the elapsed time since the last encounter of the destination d with the messenger node n that was located through flooding, and $p(t, T)$ is its density (conditioned on the accuracy T at the anchor point). We will call $p(t, T)$ the *density of the improved accuracy*. We model the cost with a function Q that satisfies the following equation.

$$Q(T) = G(T) + \int_{\alpha}^{T} Q(t)p(t,T)dt \quad T \geq \alpha \tag{1a}$$

$$Q(T) = 0 \quad 0 \leq T < \alpha \tag{1b}$$

The second equation states that the cost is zero if the elapsed time since the last encounter with the destination is less than some parameter α. This models the fact that for a small enough accuracy, the protocol may not have to resort to flooding in order to locate the destination. The first equation is an integral equation of a form known as a linear Volterra equation of the second kind. It is motivated by the fact that the total flooding area can be expressed as the area flooded at the first flooding step, plus the area flooded in all other steps.

Volterra equations arise in a variety of physical problems and have been studied by branches of mathematics such as functional and numerical analysis. From a probabilistic perspective, eq. (1) can be seen as the expected reward of a Markov chain with an uncountable number of states; $p(t, T)$ is then the conditional transition probability density and $G(T)$ the expected reward per transition.

Grossglauser *et al.* use a similar cost metric, where they also assume that the expected cost is the sum of the expected costs per flooding step, a concept that leads to the above Volterra equation in our model. They obtain their cost by also computing the total area flooded and multiplying it with the density of the grid. Multiplying our cost with the density ρ of the network leads to a cost equivalent to theirs.

5 Properties of Messenger Nodes

The aforementioned functions $G(T)$ and $p(t, T)$ depend directly on how many messenger nodes exist, what their accuracies are and how they are positioned around an anchor point. In this section, we will investigate these inherent properties of EASE under our model.

5.1 The number of messenger nodes

Suppose that the destination d follows a deterministic, fixed path given by a continuous function $f(t) : [0, +\infty) \to \mathbb{R}^2$, *i.e.* $X_d = f(t), t \geq 0$. The number of nodes it meets in the interval $[0, t]$ is

$$N_f(t) = \#\{i : \exists 0 \leq \tau \leq t \text{ s.t. } |X_i(\tau) - f(\tau)| \leq r_0\} \quad (2)$$

where $\#A$ denotes the number of elements in set A. Then the following result holds.

Lemma 1. *Let $N_f(t)$ be the number of nodes the destination meets in the interval $[0, t]$ given that it follows the fixed deterministic path $f(t)$, as defined in (2). Then*

$$\mathbf{P}\{N_f(t) = k\} = \frac{(\mu_f(t))^k}{k!} e^{-\mu_f(t)} \qquad t \geq 0, k \geq 0, \quad (3)$$

where $\mu_f(t) = \mathrm{E}[N_f(t)]$ is an increasing function and $\mu_f(0) = \rho$. Furthermore, $N_f(t)$ satisfies the independent increment property.

The above lemma indicates that $\{N_f(t), t \geq 0\}$ is a non-homogeneous Poisson process in which bulk arrivals can happen at time $t = 0$. The fact that $\mu_f(0) = \rho$ is not surprising, since the expected number of nodes in the unit-area disk around the destination at time 0 is indeed ρ, the density of the Poisson field. The case where $f(t) = f(0)$ for all $t \geq 0$ is the special case in which point d does not move but remains fixed. We will distinguish this case by using the notation $N_0(t)$ and $\mu_0(t)$ for the number of points crossing a fixed disk and its expectation respectively. Révész [10] studied the behavior of $N_0(t)$ in the case where the points forming the Poisson field $\boldsymbol{\pi}$ moved according to standard Brownian motions. Eq. (3) is a generalization of a formula he gave for the fixed disk to any continuous function f and to Brownian motions of variance proportional to v. The proof of the generalized case is quite similar to the one for a fixed disk, for which we refer the reader to Theorem 1.3 of Révész [10].

Although the expected values μ_f cannot be described with basic functions, Révész cites the following result from Spitzer [12] regarding the asymptotic behavior of μ_0:

$$\mu_0(t) = \rho \left(\frac{2\pi v t}{\log vt} + (c_1 + o(1)) \frac{vt}{\log^2 vt} \right) \qquad (4)$$

where c_1 is a constant, independent of v,t and ρ.

In our model, destination d moves according to a Brownian motion $B(t)$ with variance proportional to v, *i.e.* $X_d(t) = B(t), t \geq 0$. We can again define $N_B(t)$ as in (2). The expected number of nodes it meets will then be the expected number of nodes a fixed destination meets in double the time, as the following lemma suggests.

Lemma 2. *Let $\mu_B(t) = \mathrm{E}[N_B(t)]$ be the expected number of nodes met by a destination that moves according to a Brownian motion in the interval $[0, t]$, and $\mu_0 = \mathrm{E}[N_0(t)]$ the expected number of nodes that a fixed destination meets in the same interval. Then*

$$\mu_B(t) = \mu_0(2t). \qquad (5)$$

Proof. Let $\delta_i(t)$ be one if point i has crossed the unit disk around the destination up to time t and zero otherwise. Then $N_B(t) = \sum_i \delta_i(t)$, thus

$$\begin{aligned}
\mathrm{E}[N_B(t)] &= \mathrm{E}\left[\sum_i \delta_i(t)\right] = \sum_i \mathrm{E}[\delta_i(t)] \\
&= \sum_i \mathbf{P}\{\exists 0 \leq \tau \leq t : |X_i(\tau) - X(\tau)| \leq r_0\} \\
&= \sum_i \mathbf{P}\{\exists 0 \leq \tau \leq t : |\hat{X}_i(\tau)| \leq r_0\}
\end{aligned}$$

For each i, the probability $\mathbf{P}\{\exists 0 \leq \tau \leq t : |X_i(\tau) - X(\tau)| \leq r_0\}$ is the probability that i will cross the moving disk around the destination up to time t. This is equal to the probability that $\hat{X}_i = X_i - X_d$ will cross a unit-area disk fixed at the origin. Since Brownian motions X_i, X_d have variances proportional to v, $\hat{X}_i$ are Brownian motions with variances proportional to $2v$, which are not pairwise independent. On the other hand, the sum of these probabilities, *i.e.* of each $\hat{X}_i$ crossing a unit-area disk fixed at the origin, is the equal to the corresponding sum if $\hat{X}_i$ were independent Brownian motions. Therefore $\mathrm{E}[N_B(t)]$ is equal to the expected number of distinct particles crossing a unit-area disk fixed at the origin while moving independently with variances proportional to $2v$. □

Equation (3) implies that $N_B(t)$ *is not* Poisson distributed. The reason is that $N_B(t)$ can be expressed as an average of variables $N_f(t)$ over all possible paths f the destination may follow, and the average of Poisson random variables is, in general, not Poisson. Moreover, $N_B(t)$ cannot be characterized by assuming that the destination is fixed and all other nodes move according to independent Brownian motions with variance $2v$; the distance between the destination and any point in the Poisson field is indeed such a Brownian motion, but these motions are not independent, as mentioned in the above proof. These subtleties complicate our analysis considerably.

5.2 The accuracy of estimates of messenger nodes

As an arrival process, $\{N_B(t), t \geq 0\}$ describes the number of messenger nodes that met the destination in the interval $[0, t]$. As such, it models the epochs of the first entries of messenger nodes into the unit-area disk around

$B(t)$ within the aforementioned interval. However, in our protocol, the metric of accuracy of an estimate at a messenger node is the elapsed time since the last encounter between that node and the destination. The epoch of the last encounter between the destination and a node within an interval $[0, t]$ may be quite different than the first entry epoch within that same interval (*e.g.* it is bound to occur later than the first entry epoch).

To make this distinction more precise, we define the *first entry epoch after time* t of a node i into the unit-area disk around the destination d as

$$t^+(i,t) = \min\{\tau : \tau \geq t, |X_i(\tau) - X_d(\tau)| \leq r_0\}. \quad (6)$$

Similarly, we define the *last exit epoch prior time* t of a node i out of the unit-area disk around the destination d as

$$t^-(i,t) = \max\{\tau : \tau \leq t, |X_i(\tau) - X_d(\tau)| \leq r_0\}. \quad (7)$$

One can show that t^+ and t^- are well defined random variables. Using (6), we can define an arrival process whose epochs are the first entry epochs after some time T. This is none other than

$$N_B^+(T,t) = \#\{i : t^+(i,T) \leq T + t\}, \quad t \geq 0, \quad (8)$$

i.e. $N_B^+(T,t)$ is the number of nodes whose first entry after T in the disk around the destination happened before $T+t$. It is also easy to see that the above process and $N_B(t)$ are related as follows:

$$N_B(t) = N_B^+(0,t) = \#\{i : t^+(i,0) \leq t\}, \quad t \geq 0. \quad (9)$$

We can similarly define an arrival process whose epochs are the last exit epochs prior to some time T:

$$N_B^-(T,t) = \#\{i : t^-(i,T) \geq T - t\}, \quad t \geq 0, \quad (10)$$

i.e. $N_B^-(T,t)$ is the number of nodes whose last exit prior to T out of the disk around the destination happened after $T - t$. Note that, for a fixed T, $\{N_B^-(T,t); t \geq 0\}$ can be seen as a counting process that indicates the number of messenger nodes at time T with accuracy in the interval $[0, t]$. Furthermore, the epoch of the i-th arrival of this counting process is the elapsed time since the last encounter of a node i with the destination d, *i.e.* it is the entry T_{id} regarding the destination at node i's table T_i. In that sense, the epochs of $N_B^-(T,t)$ are quantities of immediate interest for EASE. The following lemma shows that the above two processes are governed by the same probability distribution.

Lemma 3. *For any fixed T, $N_B^-(T,t)$ is identically distributed as $N_B(t) = N_B^+(0,t)$.*

Sketch of proof. Assume that at time T all points start moving "backwards" following the trajectories that they followed in the interval $(-\infty, T]$. At time T, as at any other time, the point process $\boldsymbol{\pi}(T)$ is a Poisson field. As points move backwards they are performing independent Brownian motions by the time reversibility property of the Brownian motion. Finally, the last exit epochs of the original process become the first entry epochs in the backward one, since the last time a point exited the disk becomes the first time it enters it now. Thus, in this backward model counting the last exit epochs of the forward model is probabilistically equivalent to counting the first entry epochs of points forming a Poisson field moving according to independent Brownian motions. This however is described by the process $\{N_B^+(0,t), t \geq 0\}$. □

5.3 The spatial distribution of messenger nodes

While computing $G(T)$ and $p(t,T)$, it is important to know where messenger nodes are positioned. In fact, we need their joint spatial distribution around the anchor point, since the messenger node that will be located through flooding is the node closest to the anchor point. According to our model, if the current anchor point at time 0 is the position of the destination at time T ago, a messenger node is a node that met the destination in the interval $(-T, 0]$. Under our model, for large enough T, the displacement of a messenger node from the anchor point at time 0 can by seen as the sum of two normal random variables: the destination's displacement from the anchor point until the time of an encounter with the messenger and the messenger node's displacement from then on. From this observation one can derive that a messenger node will be normally distributed around the anchor point with variance vT, *i.e.* the distribution of its coordinates will be

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma^2} e^{-[x^2+y^2]/2\sigma^2}, \quad (11)$$

where $\sigma^2 = vT$. However, these distributions are not independent; although the latter of the two displacements they are comprised of are independent, the first ones are not. Note however that they are independent given the trajectory of the destination.

The joint spatial distribution of messenger nodes is thus not easy to obtain, due to the aforementioned interdependence of their positions. This, along with the fact that the distribution of $N_B(t)$ is not known, makes the analysis of EASE under our model difficult.

6 An Analysis Under a Simplified Model

As noted in the previous section, it is difficult to determine the spatial distribution of messenger nodes and the distribution of $N_B(t)$; indeed, it is not clear whether these quantities can be obtained analytically. For our analysis, we introduce two additional assumptions. First, we approximate $N_B(t)$ (and $N_B^-(T,t)$) with the process $\hat{N}_B(t)$ where

$$\mathbf{P}\{\hat{N}_B(t)=k\} = \frac{(\mu_0(2t))^k}{k!}e^{-\mu_0(2t)}, t \geq 0, k \geq 0, \quad (12)$$

with μ_0 described by equation (4). Note that $\hat{N}_B(t)$ has the same mean as $N_B(t)$, but is now given by a non-homogeneous Poisson process with bulk arrivals at time zero. Second, we ignore the effect of the correlated part of the positions of messenger nodes, assuming thus that they are independent. We note that the nature of the above simplifying assumptions is similar to the ones made in [4].

For this simplified model, we will denote with $\hat{G}(T)$ and $\hat{p}(t,T)$ the expected one-step flooding area and the density of the improved accuracy under the model respectively. Similarly to (1), the function $\hat{Q}(T)$ that describes the cost under this model will be a function that satisfies the equations

$$\hat{Q}(T) = \hat{G}(T) + \int_{\alpha}^{\infty} \hat{Q}(t)\hat{p}(t,T)dt \quad T \geq \alpha \qquad (13)$$

and $\hat{Q}(T)=0$ for $0 \leq T < \alpha$.

6.1 Results

Under simplified model, the positions of messenger nodes and the destination are identically, normally and independently distributed around the current anchor point with variance vT on each axis. An immediate implication of this assumption is that messenger nodes are equally likely to be located through flooding. One can use the above facts to address the problem of the optimality of the flooding condition, as described in section 2.

Proposition 1. *The value $\gamma = 1$ is optimal for the EASE protocol, with respect to function $\hat{Q}$.*

This proposition is proved in Appendix A. Using this result and the assumption that $\{N_B(t), t \geq 0\}$ can be approximated by the Poisson process $\{\hat{N}_B(t), t \geq 0\}$ in the simplified model one can derive the expected one-step flooding area $\hat{G}(T)$ and the density of the improved accuracy $\hat{p}(t,T)$. The resulting functions are given in the following two lemmas whose proofs can be found in Appendices B and C.

Lemma 4. *Under the simplified model, for $\gamma = 1$, the expected one-step flooding area is:*

$$\hat{G}(T) = \frac{2\pi vT}{\mu_0(2T)}(1-e^{-\mu_0(2T)}). \qquad (14)$$

Lemma 5. *Under the simplified model, for $\gamma = 1$, the density of the improved accuracy is*

$$\hat{p}(t,T) = \frac{\mu'(t)}{\mu(T)-\mu(0)}(1-\chi(T)) \qquad (15)$$

where $\mu(t) = \mu_0(2t)$ and

$$\chi(T) = \frac{\mu(0)}{\mu(T)} + \frac{\mu(T)-\mu(0)}{\mu(T)} \cdot \frac{1-e^{-\mu(T)}}{\mu(T)}$$

for $\alpha \leq t \leq T$, $\alpha > 0$.

Using the above two lemmas, one can solve the Volterra equation in 13 and obtain the following upper bound on protocol performance which is proved in Appendix D.

Proposition 2. *Function $\hat{Q}(T)$ is upper-bounded by*

$$\frac{1}{\rho}\left(\frac{1}{4}\log^2 T + (c(v)+o(1))\log T\right) \qquad (16)$$

where $c(v) = \frac{1}{2}\log(2v) - \frac{c_1}{4\pi}$ and c_1 is as given in Eq. (4).

6.2 Discussion

Let us briefly comment on the above results. Proposition 1 states that the optimal value for γ is 1, *i.e.* the protocol should search for *any* node that has better information than the current. As noted in Section 2, this result is not obvious. Proposition 2 states that using EASE (with $\gamma = 1$) to discover a route to a destination node that was last met at T time units ago introduces a flooding cost in the order of $\log^2 T$. We note that the expected distance (shortest path) between a source and a destination node which met T time units ago is of order $\sqrt{T}$. This means that the flooding cost of EASE (with $\gamma = 1$) is negligible compared to the overhead required for routing the packet. In [4], the flooding cost for EASE (with $\gamma = 1/2$) was shown to be at most of the order $\sqrt{T}$ and hence comparable with the overhead required for routing the packet. Proposition 2 is therefore an improvement on the bound in [4].

There are other interesting conclusions one can deduce from Proposition 2. First, the asymptotic behavior of the upper bound is independent of v. To see this, note that although v appears in (16), it does not affect its dominant term (*i.e.* the $\log^2(T)$ term); hence, for large values of T the influence of v on the protocol cost vanishes. This suggests that the protocol scales well as far as mobility is concerned. As noted in section 4, the intuition behind this is that, although high mobility makes a destination harder to locate, it also contributes to the creation of more messenger nodes and thus also to the restriction of the area flooded.

Moreover, Proposition 2 also suggests a nice behavior in terms of ρ. The quantity $\rho\hat{Q}(T)$ should indicate the number of nodes that forwarded flooding traffic. By Proposition 2, the upper bound of $\rho\hat{Q}(T)$ is independent of ρ. This result suggests that the number of nodes involved in the EASE flooding phases does not depend on the density of the network. A non-rigorous argument for this behavior would be that although in a dense network flooding the same area is more costly than in a sparse network, as noted in section 4, in a dense network more messenger nodes will be generated, thus leading to a decrease in the flooding area and the number of nodes flooded.

7 Conclusions and Future Work

We have proposed a model under which several inherent properties of the EASE protocol can be computed. Furthermore, we have obtained several interesting results regarding the performance of EASE under a simplified model. These results may be seen as incentives to investigate EASE even further and provide a more thorough analysis. Such an analysis may in fact be possible; there are indications that the simplifications made in this paper could be amended from our model and lead to an exact description of EASE. Our current work focuses on this direction.

Acknowledgements. We are indebted to Bálint Virág for his invaluable input on the work presented in this paper.

References

[1] H. Brunner and P. J. van der Houwen. *The Numerical Solution of Volterra Equations*. CWI monograph. Elsevier Science Pub. Co., 1986.

[2] H. Dubois-Ferriére, M. Grossglauser, and M. Vetterli. Age matters: Efficient route discovery in mobile ad hoc networks using encounter ages. In *Proceeding of the ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHOC)*, 2003.

[3] H. Dubois-Ferriére, M. Grossglauser, and M. Vetterli. Space-time routing in ad hoc networks. In *Ad Hoc Now 03*, Montréal, Canada, October 2003.

[4] M. Grossglauser and M. Vetterli. Locating nodes with EASE: Mobility diffusion of last encounters in ad hoc networks. In *IEEE Infocom2003*, San Francisco, 2003.

[5] D. B. Johnson and D. A. Maltz. Dynamic source routing in ad hoc wireless networks. In *Mobile Computing*, volume 353. Kluwer Academic Publishers, 1996.

[6] B. Karp and H. Kung. Greedy perimeter stateless routing for wireless networks. In *Sixth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, 2000.

[7] I. N. Kovalenko, N. Y. Kuznetsov, and V. M. Shurenkov. *Models of Random Processes: A Handbook for Mathematicians and Engineers*. CRC Press, 1996.

[8] C. E. Perkins and P. Bhagwat. Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers. In *ACM SIGCOMM'94 Conference on Communications Architectures, Protocols and Applications*, pages 234–244, October 1994.

[9] C. E. Perkins and E. M. Royer. Ad-hoc on-demand distance vector routing. In *MILCOM '97 panel on Ad Hoc Networks*, 1997.

[10] P. Révész. *Random Walks of Infinitely Many Particles*. World Scientific Publishing, 1994.

[11] N. Sarafijanovic-Djukic and M. Grossglauser. Last encounter routing under random waypoint mobility. In *NETWORKING 2004*, Athens, Greece, May 2004.

[12] F. Spitzer. Electrostatic capacity, heat flow and brownian motion. *Z. Wahrscheinlichkeitstheorie*, 3:110–121, 1964.

A A proof of Proposition 1

To prove Proposition 1, we will make use of a series of preliminary lemmas. The following describes the expected area we need to flood if the number of messenger nodes is given. In particular, the number of messenger nodes is assumed to be $n-1$, and thus there are n possible nodes in total that may be located through flooding (including the destination).

Lemma 6. *Let $(X_i, Y_i) \in \mathbb{R}^2$, $1 \le i \le n$, be n independent, omnidirectional, two-dimensional, normal random variables with zero mean and variance $\sigma^2 = vT$ on each axis. Let $A_i = \pi \cdot (X_i^2 + Y_i^2)$, $1 \le i \le n$, and $A = \min A_i$. Then*

$$\mathrm{E}\left[A\right] = 2\pi \frac{vT}{n}. \tag{17}$$

Proof. The density distribution of a two-dimensional, omnidirectional normal random variable (X, Y) is described by (11). Since $dxdy = rdrd\theta$, in polar coordinates we get $f_{R,\Theta}(r,\theta) = \frac{r}{2\pi\sigma^2} e^{-r^2/2\sigma^2}$. The marginal density of R is thus $f_R(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2}$. The probability that the area A_i is greater than πr^2 is therefore

$$\mathbf{P}\{A_i > \pi r^2\} = \int_r^\infty \frac{\rho}{\sigma^2} e^{-\frac{\rho^2}{2\sigma^2}} d\rho = e^{-\frac{r^2}{2\sigma^2}}$$

where $\sigma^2 = vT$. Thus, the probability that the area A that is greater than πr^2 is

$$\mathbf{P}\{A > \pi r^2\} = \prod_i \mathbf{P}\{A_i > \pi r^2\} = e^{-\frac{nr^2}{2\sigma^2}}$$

since A_i are independent. Therefore, the probability that the area A is greater than a value a is $\mathbf{P}\{A > a\} = e^{-\frac{na}{2\pi\sigma^2}}$ and its expected value is

$$\mathrm{E}\left[A\right] = \int_0^\infty \mathbf{P}\{A > a\} da = 2\pi \frac{\sigma^2}{n}$$

where $\sigma^2 = vT$. □

Assume now that we wish to route a packet to the destination at time 0 and that we are given the accuracies of messenger nodes, *i.e.* if S_i are the epochs of the arrival process formed by the accuracies of the messenger nodes (see

Appendix C for a more elaborate description of this process), we have that $S_i = t_i$, $i \in \{1, 2, \ldots\}$ are given and $S_0 = t_0 = 0$ is the accuracy of the destination itself. According to our simplified model (and our original model as well), many nodes may have zero accuracies, but all other nodes have distinct accuracies with probability one. Hence, we can order nodes so that $t_i > t_j$ if $i > j > s$, for some value $s \geq 0$, and $t_i = 0$ for all $i \leq s$.

By Lemma 6, the expected one-step flooding area from a point where the accuracy is t_n is $\hat{G}(t_n) = \frac{2\pi\sigma^2(t_n)}{n}$, where $\sigma^2(t) = vt$. Since all messenger nodes and the destination are equally likely to be located through flooding, the density of the improved accuracy from a point with accuracy t_n (given that the accuracies of all messenger nodes are known), where $n > s$, is $\hat{p}(t_i, t_n) = \frac{1}{n}$, for $0 \leq i \leq n-1$. Hence, equation (13) becomes

$$\hat{Q}(t_n) = \frac{2\pi\sigma^2(t_n)}{n} + \sum_{i=0}^{n-1} \frac{\hat{Q}(t_i)}{n}, \quad t_n \geq \alpha \tag{18}$$

and $\hat{Q}(t_n) = 0$ otherwise. The following lemma is a consequence of the monotonicity of $\sigma^2(t)$ and we omit its proof for reasons of brevity.

Lemma 7. *Function $\hat{Q}$, as defined in* (18)*, is an increasing function,* i.e. *for every $t_i > t_j$ we have $\hat{Q}(t_i) \geq \hat{Q}(t_j)$.*

We will now show that, by adding a new messenger node, function $\hat{Q}$ can only be improved.

Lemma 8. *Let p be a new messenger node with accuracy t_p such that $t_{n-1} < t_p < t_n$ for some $n \geq 1$. Let $\hat{Q}'(t)$ be the cost function if we consider this messenger node when flooding and $\hat{Q}(t)$ the one if we do not consider it. Then, $\hat{Q}'(t_i) \leq \hat{Q}(t_i)$ for all $i \in \mathbb{N}$.*

Proof. If $t_p < \alpha$ this is trivially true. Assume thus that $t_p \geq \alpha$. Notice that

$$\hat{Q}'(t_i) = \hat{Q}(t_i), \quad \text{for all } 0 \leq i \leq n-1 \tag{19}$$

since, by the definition of the functions $\hat{Q}$, $\hat{Q}'$, they depend only on messenger nodes with accuracy better than t_i and are thus not influenced by the existence of messenger node p. At the anchor point with accuracy t_n, $\hat{Q}(t_n)$ is described by (18). On the other hand, $\hat{Q}'(t_n)$ is

$$\begin{aligned}\hat{Q}'(t_n) &= \frac{2\pi\sigma^2(t_n)}{n+1} + \sum_{i=0}^{n-1} \frac{\hat{Q}'(t_i)}{n+1} + \frac{\hat{Q}'(t_p)}{n+1} \\ &\stackrel{(19)}{=} \frac{2\pi\sigma^2(t_n)}{n+1} + \sum_{i=0}^{n-1} \frac{\hat{Q}(t_i)}{n+1} + \frac{\hat{Q}'(t_p)}{n+1}\end{aligned} \tag{20}$$

where

$$\begin{aligned}\hat{Q}'(t_p) &= \frac{2\pi\sigma^2(t_p)}{n} + \sum_{i=0}^{n-1} \frac{\hat{Q}'(t_i)}{n} \\ &\stackrel{(19)}{=} \frac{2\pi\sigma^2(t_p)}{n} + \sum_{i=0}^{n-1} \frac{\hat{Q}(t_i)}{n}\end{aligned} \tag{21}$$

Therefore, $\hat{Q}'(t_n) \leq \hat{Q}(t_n)$ becomes

$$\frac{2\pi\sigma^2(t_n)}{n+1} + \sum_{i=0}^{n-1} \frac{\hat{Q}(t_i)}{n+1} + \frac{\hat{Q}'(t_p)}{n+1} \leq \frac{2\pi\sigma^2(t_n)}{n} + \sum_{i=0}^{n-1} \frac{\hat{Q}(t_i)}{n}$$

which is equivalent to

$$\left(2\pi\sigma^2(t_n) + \sum_{i=0}^{n-1} \hat{Q}(t_i)\right)\left(\frac{1}{n} - \frac{1}{n+1}\right) \geq \frac{\hat{Q}'(t_p)}{n+1}$$

or

$$\left(\frac{2\pi\sigma^2(t_n)}{n} + \sum_{i=0}^{n-1} \frac{\hat{Q}(t_i)}{n}\right)\frac{1}{n+1} \geq \frac{\hat{Q}'(t_p)}{n+1} \tag{22}$$

and by eq. (21) can be written as

$$\frac{2\pi\sigma^2(t_n)}{n} + \sum_{i=0}^{n-1} \frac{\hat{Q}(t_i)}{n} \geq \frac{2\pi\sigma^2(t_p)}{n} + \sum_{i=0}^{n-1} \frac{\hat{Q}(t_i)}{n}$$

which is true iff $\sigma^2(t_n) \geq \sigma^2(t_p)$. This holds by the monotonicity of $\sigma^2(t) = vt$. Note that inequality (22) implies that

$$\hat{Q}(t_n) \geq \hat{Q}'(t_p) \tag{23}$$

We prove the remaining cases by induction. We know that $\hat{Q}'(t_n) \leq \hat{Q}(t_n)$. Assume that $\hat{Q}'(t_i) \leq \hat{Q}(t_i)$ for all $n \leq i \leq k-1$ for some $k > n$. We will prove that $\hat{Q}'(t_k) \leq \hat{Q}(t_k)$. This is equivalent to

$$\frac{2\pi\sigma^2(t_k)}{k+1} + \sum_{i=n}^{k-1} \frac{\hat{Q}'(t_i)}{k+1} + \frac{\hat{Q}'(t_p)}{k+1} + \sum_{i=0}^{n-1} \frac{\hat{Q}'(t_i)}{k+1} \leq \frac{2\pi\sigma^2(t_k)}{k} + \sum_{i=0}^{k-1} \frac{\hat{Q}(t_i)}{k}.$$

Since $\hat{Q}'(t_i) \leq \hat{Q}(t_i)$ for all $n \leq i \leq k-1$ by the induction hypothesis and $\hat{Q}'(t_i) = \hat{Q}(t_i)$ for all $0 \leq i \leq n-1$, it suffices that

$$\frac{2\pi\sigma^2(t_k)}{k+1} + \sum_{i=n}^{k-1} \frac{\hat{Q}(t_i)}{k+1} + \frac{\hat{Q}'(t_p)}{k+1} + \sum_{i=0}^{n-1} \frac{\hat{Q}(t_i)}{k+1} \leq \frac{2\pi\sigma^2(t_k)}{k} + \sum_{i=0}^{k-1} \frac{\hat{Q}(t_i)}{k}$$

or, equivalently, $\hat{Q}(t_k) \geq \hat{Q}'(t_p)$ which is true since $\hat{Q}(t_k) \geq \hat{Q}(t_n)$ by the monotonicity of $\hat{Q}$ and $\hat{Q}(t_n) \geq \hat{Q}'(t_p)$ by (23). □

Lemma 8 shows that by considering an existing messenger node, the LER protocol can only improve its performance with respect to function $\hat{Q}$ under our simplified model. This immediately implies Proposition 1, since any halting condition other than the maximal will discard some messenger nodes.

B A proof of Lemma 4

Lemma 6, proved in Appendix A, describes the expected area flooded for a given number of messenger nodes. The following lemma gives the expected flooding area for $\gamma = 1$ and for a Poisson distributed number of messenger nodes, as dictated by the simplified model.

Lemma 9. *Let $(X_i, Y_i) \in \mathbb{R}^2$, $1 \leq i \leq N+1$, be $N+1$ independent omnidirectional, two-dimensional, normal random variables with zero mean and variance $\sigma^2 = vT$ on each axis, where N is a Poisson distributed random variable with mean μ. Let $A_i = \pi(X_i^2 + Y_i^2)$, $1 \leq i \leq N+1$ and $A = \min A_i$. Then*

$$\mathrm{E}\,[A] = \frac{2\pi v T}{\mu}(1 - e^{-\mu}). \tag{24}$$

Proof. Given that $N = n$, by Lemma 6 the conditional expected minimum area is $\mathrm{E}\,[A \mid N = n] = \frac{2\pi\sigma^2}{n+1}$. Therefore, the expected area is

$$\begin{aligned}\mathrm{E}\,[A] &= \sum_{n=0}^{\infty} \frac{2\pi\sigma^2}{n+1} \frac{e^{-\mu}\,(\mu)^n}{n!} = \frac{2\pi\sigma^2 e^{-\mu}}{\mu} \sum_{n=0}^{\infty} \frac{(\mu)^{n+1}}{n+1!} \\ &= \frac{2\pi\sigma^2}{\mu}(1 - e^{-\mu})\end{aligned}$$

where $\sigma^2 = vT$. □

Lemma 4 thus follows from Lemma 9 and equation (12).

C A proof of Lemma 5

Let $\{N(t); t \geq 0\}$, be the counting process of the number of messenger nodes with accuracy in the interval $[0, t]$ at time 0, and let $\mathrm{E}\,[N(t)] = \mu(t)$ for $t \geq 0$. We know that this process is none other than $\{\hat{N}_B^-(0,t), t \geq 0\}$, and that it is thus identically distributed to $\{\hat{N}_B(t), t \geq 0\}$ by Lemma 3. Moreover, $\mu(t) = \mathrm{E}\,[N(t)] = \hat{\mu}_B(t) = \mu_0(2t)$. W.l.o.g. we assume that the node that corresponds to the epoch S_i of this arrival process is node i.

By (12), $\{N(t); t \geq 0\}$ is a non-homogeneous Poisson process with bulk arrivals at time zero. We will denote the arrival epochs of this process with S_i, $i \geq 1$. Furthermore, $\tilde{N}(t) = N(t) - N(0)$ is a non-homogeneous Poisson process independent of $N(0)$ with $\mathrm{E}\,[\tilde{N}(t)] = \mu(t) - \mu(0)$. The arrival epochs of this process will be denoted with $\tilde{S}_i$, $i \geq 1$.

We assume that we start flooding at time 0 from the position of the destination at T time ago, *i.e.* at $X_d(-T)$. All existing messenger nodes and the destination are equally likely to be found, therefore, if $N(T) = n$, the probability that the i-th messenger node (or the destination) will be found is $\frac{1}{n+1}$. Let X be the time estimate that we will get, *i.e.* the time S_i since node i had accurate information about the destination, if we located messenger node i, and zero if we located the destination. We can formally describe X as

$$\mathbf{P}\{X = S_i \mid N(T) = n\} = \frac{1}{n+1} \quad 0 \leq i \leq n$$

where S_i, $i \geq 1$, is the epoch of the i-th arrival of the process $\{N(t), t \geq 0\}$ and $S_0 \equiv 0$, corresponding to the event of finding the destination itself.

We will use the following lemma

Lemma 10. *Let $\{N(t), t \geq 0\}$ be a non-homogeneous Poisson process with $\mathrm{E}\,[N(t)] = \mu(t)$ (possibly with $\mu(0) \neq 0$). Given that $N(T) = n$ for some $T > 0$ and $n > 0$, let Y be a random variable defined as*

$$\mathbf{P}\{Y = S_i \mid N(t) = n\} = \frac{1}{n}$$

where S_i is the epoch of the i-th arrival, $1 \leq i \leq n$. Then

$$\mathbf{P}\{Y \leq t \mid N(T) = n\} = \frac{\mu(t)}{\mu(T)}, \qquad 0 \leq t \leq T.$$

This lemma is merely a consequence of the properties of the Poisson process and we omit its proof. A variation of it can be found in Kovalenko *et al.* [7, p. 76].

The following lemma describes the probability that the node located through flooding is the destination or is a messenger node aware of the destination's current position.

Lemma 11.

$$\mathbf{P}\{X = 0\} = \frac{\mu(0)}{\mu(T)} + \frac{\mu(T) - \mu(0)}{\mu(T)} \cdot \frac{1 - e^{-\mu(T)}}{\mu(T)}.$$

Proof. Given that $N(T) = n$, we have that the probability that $X = 0$ is equal to the probability that the destination was located or that a messenger node that has accuracy zero was located. The probability that the destination was located is $\frac{1}{n+1}$, since all nodes and the destination are equally likely to be found.

The probability that a messenger node is chosen is $\frac{n}{n+1}$. We wish to compute the probability that, given that a messenger node was chosen, it was one with accuracy zero.

Given that a messenger node is chosen, each of them is equally likely to be located with probability $\frac{1}{n}$. Since $\{N(t), t \geq 0\}$ is a non-homogeneous Poisson process with expected value $\mu(t)$, by Lemma 10, choosing one of its epochs with equal probability will yield a zero epoch with probability $\mu(0)/\mu(T)$. Hence we get that

$$\begin{aligned}
\mathbf{P}\{X = 0 \mid N(T) = n\} &= \frac{1}{n+1} + \frac{n}{n+1}\frac{\mu(0)}{\mu(T)} \\
&= \frac{1}{n+1} + \left(1 - \frac{1}{n+1}\right)\frac{\mu(0)}{\mu(T)} \\
&= \frac{\mu(0)}{\mu(T)} + \left(1 - \frac{\mu(0)}{\mu(T)}\right)\frac{1}{n+1}
\end{aligned}$$

Thus $\mathbf{P}\{X = 0\}$ can be computed as the expectation over $N(T)$, where $N(T)$ is Poisson distributed with average $\mu(T)$. This computation gives us the lemma. □

Lemma 12.

$$\mathbf{P}\{X \leq t \mid X \neq 0\} = \frac{\mu(t) - \mu(0)}{\mu(T) - \mu(0)}, \quad 0 < t \leq T. \quad (25)$$

Proof. If $X \neq 0$, the messenger node we located is one of the nodes that are counted in $\tilde{N}(T)$, where $\tilde{N}(t) = N(t) - N(0)$. Furthermore, $\tilde{N}(T) \geq 1$, since at least one such messenger node exists (the one we located). We are therefore contemplating the probability

$$\mathbf{P}\{\tilde{X} \leq t \mid \tilde{N}(T) \geq 1\}$$

where $\tilde{X}$ is the epoch we get if we choose among the epochs in $(0, T]$ with equal probability, *i.e.*

$$\mathbf{P}\{\tilde{X} = \tilde{S}_i \mid \tilde{N}(T) = n\} = \frac{1}{n}, \qquad 1 \leq i \leq n.$$

However, $\{\tilde{N}(t), t \geq 0\}$ is a Poisson process with expectation $\mu(t) - \mu(0)$. Hence, by Lemma 10 we get that

$$\mathbf{P}\{\tilde{X} \leq t \mid \tilde{N}(T) = n\} = \frac{\mu(t) - \mu(0)}{\mu(T) - \mu(0)},$$

for $0 \leq t \leq T$. Since $\mathbf{P}\{\tilde{X} \leq t \mid \tilde{N}(T) \geq 1\}$ is equal to

$$\frac{\sum_{n=1}^{\infty} \mathbf{P}\{\tilde{X} \leq t \mid \tilde{N}(T) = n\}\mathbf{P}\{\tilde{N}(T) = n\}}{\mathbf{P}\{\tilde{N}(T) \geq 1\}}$$

we have

$$\mathbf{P}\{\tilde{X} \leq t \mid \tilde{N}(T) \geq 1\} = \frac{\mu(t) - \mu(0)}{\mu(T) - \mu(0)}$$

which gives us the lemma. □

Let $\chi(T) = \mathbf{P}\{X = 0\}$. From Lemmas 11,12 we get

$$\mathbf{P}\{X \leq t\} = \frac{\mu(t) - \mu(0)}{\mu(T) - \mu(0)}(1 - \chi(T)) + \chi(T)$$

for $0 \leq t \leq T$. Function $\hat{p}(t, T)$ is actually $\frac{\partial}{\partial t}\mathbf{P}\{X \leq t\}$, which exists for all $t \geq \alpha > 0$. This implies Lemma 5.

D A proof of Proposition 2

Lemmas 4 and 5 give us $\hat{G(T)}$ and $\hat{p}(t, T)$ under the simplified model. We can thus insert them in the Volterra equation in (13) and try to obtain a solution for $\hat{Q}(T)$.

The kernel $\hat{p}(t, T)$ of (13) is separable, so the solution to the equation can be obtained by the following formula (see e.g. Brunner and van der Howen [1]):

$$\hat{Q}(T) = \hat{G}(T) + \int_{\alpha}^{T} \hat{p}(t, T)e^{\left(\int_t^T \hat{p}(u,u)du\right)}\hat{G}(t)dt.$$

We have that

$$\begin{aligned}
\int_t^T \hat{p}(u, u)du &= \int_t^T \frac{\mu'(u)}{\mu(u) - \mu(0)}(1 - \chi(u))du \\
&\leq \int_t^T \frac{\mu'(u)}{\mu(u) - \mu(0)}du = \log\frac{\mu(T) - \mu(0)}{\mu(t) - \mu(0)}
\end{aligned}$$

We thus have

$$\begin{aligned}
\hat{Q}(T) &\leq \hat{G}(T) + \int_{\alpha}^{T} \frac{\mu'(t)}{\mu(t) - \mu(0)}\hat{G}(t)dt \\
&\leq \frac{2\pi vT}{\mu(T)} + \int_{\alpha}^{T} \frac{\mu'(t)}{\mu(t) - \mu(0)}\frac{2\pi vt}{\mu(t)}dt \\
&\leq \frac{2\pi vT}{\mu(T) - \mu(0)} + \int_{\alpha}^{T} \frac{\mu'(t)}{(\mu(t) - \mu(0))^2}2\pi vtdt \\
&= \int_{\alpha}^{T} \frac{2\pi v}{\mu(t) - \mu(0)}dt + \frac{2\pi v\alpha}{\mu(\alpha) - \mu(0)}
\end{aligned}$$

However $\mu(t) - \mu(0) = \mu_0(2t) - \mu_0(0)$ where $\mu_0(t)$ is described by (4). Hence we have

$$\begin{aligned}
\hat{Q}(T) \leq\ & \frac{1}{\rho}\int_{\alpha}^{T} \frac{2\pi v}{\frac{2\pi 2vt}{\log 2vt} + (d + o(1))\frac{2vt}{\log^2 2vt}}dt + \\
& \frac{2\pi v\alpha}{\rho\left(\frac{2\pi 2v\alpha}{\log 2v\alpha} + (c_1 + o(1))\frac{2v\alpha}{\log^2 2v\alpha}\right)} \\
=\ & \frac{1}{\rho}\int_{\alpha}^{T} \frac{\log 2vt}{2t}\frac{\log 2vt}{\log 2vt + \frac{c_1}{2\pi} + o(1)}dt + \frac{o(1)\log T}{\rho} \\
=\ & \frac{1}{\rho}\left[\frac{1}{4}\log^2 vT - \int_{\alpha}^{T} \frac{\log 2vt}{2t}\frac{\frac{c_1}{2\pi} + o(1)}{\log 2vt + \frac{c_1}{2\pi} + o(1)}dt \right. \\
& \left. + o(1)\log T\right]
\end{aligned}$$

Note that

$$\lim_{T\to\infty} \frac{\int_{\alpha}^{T} \frac{\log 2vt}{2t}\frac{c_1/2\pi + o(1)}{\log 2vt + (c_1/2\pi + o(1))}dt}{\log T} - c_1/4\pi = 0$$

Hence

$$\hat{Q}(T) \leq \frac{1}{\rho}\Bigg(\frac{1}{4}\log^2 2vT - (c_1/4\pi + o(1))\log T +$$
$$o(1)\log(T)\Bigg)$$
$$= \frac{1}{\rho}\left(\frac{1}{4}\log^2 2vT + (-c_1/4\pi + o(1))\log T\right)$$

which gives us the upper bound in the proposition.