# Maximizing the Functional Lifetime of Sensor Networks

Arvind Giridhar
Coordinated Science Laboratory
University of Illinois, Urbana.
Email: giridhar@uiuc.edu

P.R. Kumar
Coordinated Science Laboratory
University of Illinois, Urbana.
Email: prkumar@uiuc.com

***Abstract*— The *functional lifetime* of a sensor network is defined as the maximum number of times a certain data collection function or task can be carried out without any node running out of energy. The specific task considered in this paper is that of communicating a specified quantity of information from each sensor to a collector node. The problem of finding the communication scheme which maximizes functional lifetime can be formulated as a linear program, under "fluid-like" assumptions on information bits. This paper focuses on *analytically* solving the linear program for some simple regular network topologies.**

**The two topologies considered are a regular linear array, and a regular two-dimensional network. In the linear case, an upper bound on functional lifetime is derived, as a function of the initial energies and quantities of data held by the sensors. Under some assumptions on the relative amounts of the energies and data, this upper bound is shown to be achievable, and the exact form of the optimal communication strategy is derived. For the regular planar network, upper and lower bounds on functional lifetime, differing only by a constant factor, are obtained.**

**Finally, it is shown that the simple collection scheme of transmitting only to nearest neighbors, yields a nearly optimal lifetime in a scaling sense.**

## I. Introduction

Consider a network of sensors, each with a certain quantity of data to be communicated to a designated collector node. The sensor nodes are low power devices, while the collector node might be an information processing center, or a more energy-rich device functioning as a gateway to a higher bandwidth wireless network (such as in [1]). Information packets can be relayed - there is no compression or processing allowed - through an arbitrary sequence of nodes before reaching the collector. There is an energy cost associated with each transmission, which consists of a cost to transmit, that is a function of the transmitter-receiver distance, as well as a cost to receive.

The above is a somewhat idealized model of a sensor network, with a very simple definition of a data collection task. In this paper, we examine the notion of *functional lifetime*, which is defined as follows: Given a quantity of data and an energy budget (for transmitting and receiving) at each node, the functional lifetime is the maximum number of times the task of delivering all the data to the collector node can be repeated before some node runs out of energy. Alternatively, it is the maximum common scale factor by which all the quantities of data at all the sensors can be scaled, while ensuring that it can be delivered to the collector node without some node running out of energy.

In general, a sensor network may be simultaneously performing a variety of tasks, including sensing, processing and communication. In such a context a natural question to ask is the "cost" to longevity associated with carrying out a specific task of interest, and how such a cost is to be minimized. The per task functional lifetime defined above addresses such a question for communication tasks; its inverse represents the maximum fraction of any node's energy, i.e., or "fraction of lifetime" of the network, that is consumed in performing the task. The important point to note is that minimizing the *total* energy consumed may not be optimal for network lifetime, due to the distributed nature of the communication burden as well as energy supplies.

Slight variants of this problem, with essentially the same modeling assumptions and notion of lifetime (i.e., time until the first node failure) have been considered in a number of papers in the literature, in the context of energy aware or maximum lifetime routing. Chang and Tassiulas [2] consider the problem of maximizing lifetime given a certain set of source-destination information rates that must be supported. Similar approaches, leading to linear and/or integer programming formulations which are basically the same as in this paper, have been employed in [3–5]. Most of these papers have focused on finding distributed algorithms to find the optimal routing strategy to maximize lifetime, without specifically analyzing what this solution is.

In this work, we consider a restricted version of the problem considered in the above papers. Two regular spatial topologies are studied, a linear array and a planar circularly symmetric network. Each of these consist of many sensors and one collector. Under these restrictions, we are able to provide sharp analytical bounds, and in some cases exact solutions, to the functional lifetime problem. These analytical results cannot directly be translated to practical distributed algorithms, since constructing the optimal schemes requires non-local information. The advantage of an analytical approach however lies in the structural insight that it provides; while the linear programs could equally well be solved numerically, the numbers may not provide such insight. Specifically, this paper attempts to shed light on natural questions of interest such as:

- How does lifetime depend on the relative quantities and distribution among sensors of data that is to be collected?
- What is the structure of routing strategies that are optimal with respect to lifetime?
- How does lifetime scale with the size of the network, and/or quantity of data to be collected?
- How well do simple routing strategies perform in comparison to the optimal strategy?

We take the approach of addressing these questions for simple network topologies, which nevertheless are fairly good representatives of multi-hop networks. Indeed, our results are indicative of what can be expected in the broader class of "more or less" regular networks as well. Our main contributions can be summarized as follows:

1) We provide an upper bound on the functional lifetime for regular linear networks, which is valid for arbitrary energy profiles and data distribution.
2) Under some restrictions on the energy profile and data distribution, this upper bound is shown to be achievable, and the exact form of the optimal solution is given. This optimal solution causes all nodes to die simultaneously.
3) Similar results are obtained for a class of regular planar

networks.

4) The optimal solutions are compared to a simple suboptimal routing strategy, which consists of each node forwarding all data to its nearest neighbor in the direction of the collector. The results indicate that the simple strategy is nearly optimal in a scaling sense.

All the above results depend on a specific property of the energy cost function, which is derived in Lemmas 3 and 4 in the appendix.

Analytical upper bounds on lifetime are also derived in [6]. These bounds are not very sharp for the "many-to-one" information flows that we consider here, since they deal with total energy consumption rather than per-node energy consumption.

One minor point needs to be noted. While most of the afore-mentioned papers have considered lifetime as being in actual time units, as a function of sets of desired information rates, we have defined functional lifetime as being the number of times a task can be repeated. At one level, this is merely an issue of semantics; one definition can be mapped to the other by mapping a time unit to a "round," which is the time taken for one repetition of the data collection task. However, in the former approach one also needs to address the feasibility of a particular communication flow, which is specified in bits per time unit, given the bandwidth and interference constraints of a wireless network.

In our formulation, however, there is no limitation on the time taken to perform the required communication. For instance, we could allow for the extremely inefficient strategy of only one transmitter in the entire network being active at a time. Such a strategy might have an extremely large associated delay, but the energy consumption would be as specified by our modeling assumptions. Thus, our formulation clearly separates the issue of lifetime optimization from scheduling, delay, and achievable throughputs. In practice, however, sensor network routing algorithms must jointly consider all these. Characterizing the tradeoffs between lifetime, throughput and delay remains an open problem.

The outline of the rest of the paper is as follows. Section II describes the model and linear programming formulation. Sections III and IV deal with the linear and planar networks, respectively. Section V obtains scaling laws for the optimal lifetime, followed by a discussion of conclusions and future work.

## II. Model, Assumptions and Formulation of Linear Program

The general problem setting we consider is very similar to those considered in [2–5]. Suppose there are sensors $1, 2, \ldots, n$ located on a plane, along with a collector node labeled $0$. Sensor $i$ located at $(x_i, y_i)$, has $b_i$ bits to send to the collector node, and has initial energy level $E_i$. The energy consumed by node $i$ in sending $m$ units of information to node $j$, which is a distance $d(i,j)$ away, is $mf(d(i,j))$, while the energy consumed by $j$ (for $j \neq 0$; we do not count the energy consumed by the collector node) in receiving $m$ units of information, is $f_R m$, for a given constant $f_R \geq 0$. We seek the communication scheme that maximizes the number of times the information collection process, i.e., communicating $b_i$ bits from each sensor to the collector node, can be repeated before one of the sensors dies, i.e., has no remaining energy.

We make the simplifying assumption that information is infinitely divisible and incompressible. As a consequence of this "fluid-like" assumption, flow conservation is preserved.

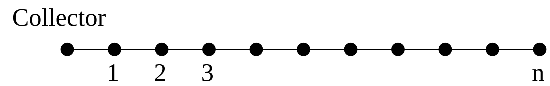


Fig. 1. A regular linear sensor network.

Consider the following linear program:

$$\text{Min } z \tag{1}$$

subject to:

$$\frac{1}{E_i}\left(\sum_{0\leq j\leq n} \lambda_{ij} f(d(i,j)) + \sum_{1\leq j\leq n} \lambda_{ji} f_R\right) \leq z, \quad 1 \leq i \leq n \tag{2}$$

$$\sum_{0\leq j\leq n} \lambda_{ij} - \sum_{1\leq j\leq n} \lambda_{ji} = b_i, \quad 1 \leq i \leq n \tag{3}$$

$$\lambda_{ij} \geq 0,\ 1 \leq i,j \leq n.$$

Here $\lambda_{ij}$ denotes the amount of information transmitted from node $i$ to node $j$, and $z$ is essentially the maximum of the *normalized costs*, or energy cost divided by initial energy, incurred by each node. The set of $\lambda_{ij}$'s satisfying the constraints specifies a communication *flow* from the set of sensors to the collector node.

This formulation is in a different form than the linear program described in [2], but the two can be mapped to each other via simple change of variables. The following lemma establishes the connection between the functional lifetime problem and the linear program (1).

*Lemma 1:* The solution to the optimization problem (1) yields the strategy that maximizes the number of times the information collection process described above can be repeated before some node dies. The functional lifetime is given by $\frac{1}{z^*}$, where $z^*$ is the optimal value of the objective function.

**Proof.** Assuming the fluidity of information, any feasible solution to (1) can be translated to a communication scheme with appropriate scheduling, since the flow conservation constraints (3) are satisfied. If $\lambda^*$ is the solution to the optimization problem, with optimal value $z^*$, one can replace all the $b_i$'s by $\frac{b_i}{z^*}$ and the $\lambda^*_{ij}$'s by $\frac{\lambda^*_{ij}}{z^*}$ . Then, from the constraints (2), we have that for each $i$,

$$\frac{1}{z^*}\left(\sum_{0\leq j\leq n} \lambda_{ij} f(d(i,j)) + \sum_{1\leq j\leq n} \lambda_{ji} f_R\right) \leq E_i, \tag{4}$$

meaning that no sensor runs out of energy. Thus, the optimal communication scheme can repeat the operation $\frac{1}{z^*}$ times without any sensor running out of energy.

On the other hand, any communication scheme to repeat the operation $K$ times, i.e., transmit $Kb_i$ bits from each node to the collector, must involve sending bits between nodes in such a way so as to satisfy the flow conservation constraints (3). Furthermore, the total energy consumed by each node $i$ must be no more than $E_i$, so a constraint of the form (4) must be satisfied as well. Thus, the constraints (3) and (2) are necessary as well as sufficient, and $\frac{1}{z^*}$ is the maximum number of times the operation can be repeated. □

The linear program given above can be numerically solved for any number of nodes, any spatial placement, and energy function. The rest of the paper, however, will be devoted to explicitly solving this problem in some restricted cases.

## III. Functional Lifetime of Regular Linear networks

Consider the uniform linear configuration (Figure 1), where $n$ sensors spaced evenly at a distance $d$ from each other on a line, with a single collector node located at the leftmost point. The collector node is located at the origin, and sensor $i$ at the point $(id, 0)$.

The energy cost function $f(\cdot)$ is given by

$$f(x) = E_t x^{\alpha} e^{\gamma x}, \tag{5}$$

where $\alpha \geq 2$ is the path loss exponent, $\gamma \geq 0$ the absorption coefficient, and $E_t$ is some positive constant. We make the following assumption on the constant $f_R$, which represents the energy to receive a unit of information:

**Assumption**: $f_R \leq E_t d^{\alpha}(1 - (1/2)^{\alpha-1})$.

This is a fairly reasonable condition; for $\alpha = 3$, it states that the power to receive must be no more than $\frac{3}{4}$ times the power to transmit to the nearest neighbor. As stated, it is a sufficient condition for Lemma 4 to hold; we expect that a somewhat weaker condition would also suffice. Nonetheless, the structure of the energy cost function, together with this assumption, are critical to all the results in this paper. The only specific property we require of the energy cost function is stated in Lemma 4, and this property is essential for any of our results to hold. Essentially, what it describes is the relative costs of short and long hops; the disadvantage of a long hop is that due to the sharp degradation in signal power with distance, it is more inefficient in terms of transmit power than multiple short hops (see [7]). However, multiple short hops involve more receptions and thus increase the power spent in receiving, and in addition increase the energy consumption of the relaying nodes. The optimal strategy described in Subsection II B optimally balances these two effects. However, if power to receive is too high (i.e., the assumption on $f_R$ is violated), then relaying becomes too expensive, and such a strategy is no longer optimal.

The form of the cost function $f(\cdot)$ given by (5), and the assumption on $f_R$ will be implicitly assumed for the rest of the paper.

Denote $f_i := f(id)$ for brevity. The linear program (1) has the following form:

$$\begin{aligned}
&\text{Min } z \qquad\qquad (6)\\
&\text{subject to:}\\
&E_1 z - \left( \sum_{0\leq j\leq n} \lambda_{1j} f_{|j-1|} + \sum_{1\leq j\leq n} \lambda_{j1} f_R \right) \geq 0\\
&E_2 z - \left( \sum_{0\leq j\leq n} \lambda_{2j} f_{|j-2|} + \sum_{1\leq j\leq n} \lambda_{j2} f_R \right) \geq 0\\
&\ldots\\
&E_n z - \left( \sum_{0\leq j\leq n} \lambda_{nj} f_{|j-n|} + \sum_{1\leq j\leq n} \lambda_{jn} f_R \right) \geq 0\\
&\lambda_{10} + \lambda_{12} + \ldots + \lambda_{1n} - \lambda_{21} - \lambda_{31} - \ldots - \lambda_{n1} = b_1\\
&\lambda_{20} + \lambda_{21} + \ldots + \lambda_{2n} - \lambda_{12} - \lambda_{32} - \ldots - \lambda_{n2} = b_2\\
&\ldots\\
&\lambda_{n0} + \lambda_{n1} + \ldots + \lambda_{nn-1} - \lambda_{1n} - \ldots - \lambda_{n-1n} = b_n\\
&\lambda_{ij} \geq 0, \text{ for each } 1 \leq i, j \leq n.
\end{aligned}$$

### A. Upper Bound on Lifetime

We now give the first main result, which provides an upper bound to functional lifetime via linear programming duality.

*Theorem 1:* Let $z^*$ be the optimal value of the linear program (6). $z^*$ is lower bounded as follows:

$$z^* \geq \frac{\sum_{i=1}^n b_i \frac{\prod_{j=i+1}^n 1 - \frac{f_1}{f_j}}{\prod_{j=i}^n 1 + \frac{f_R}{f_j}}}{\sum_{i=1}^n \frac{E_i}{f_i} \frac{\prod_{j=i+1}^n 1 - \frac{f_1}{f_j}}{\prod_{j=i}^n 1 + \frac{f_R}{f_j}}}. \tag{7}$$

**Proof.** We find a feasible solution to the dual program of (6) with objective function equal to the RHS of (7). The result then follows by weak duality.

Associating the first $n$ constraint equations in (6) with variables $y_1, y_2, \ldots, y_n$, and the last $n$ constraint equations with $w_1, w_2, \ldots, w_n$, the dual linear program has the following form:

$$\begin{aligned}
&\text{Max } b_1 y_1 + b_2 y_2 + \ldots + b_n y_n \qquad (8)\\
&\text{subject to:}\\
&E_1 w_1 + E_2 w_2 + \ldots + E_n w_n \leq 1\\
&-f_1 w_1 + y_1 \leq 0\\
&-f_1 w_1 - f_R w_2 + y_1 - y_2 \leq 0\\
&-f_2 w_1 - f_R w_3 + y_1 - y_3 \leq 0\\
&\ldots\\
&-f_{n-1} w_1 - f_R w_n + y_1 - y_n \leq 0\\
&-f_2 w_2 + y_2 \leq 0\\
&-f_1 w_2 - f_R w_1 + y_2 - y_1 \leq 0\\
&-f_1 w_2 - f_R w_3 + y_2 - y_3 \leq 0\\
&\ldots\\
&-f_{n-2} w_2 - f_R w_n + y_2 - y_n \leq 0\\
&\ldots\\
&-f_n w_n + y_n \leq 0\\
&-f_{n-1} w_n - f_R w_1 + y_n - y_1 \leq 0\\
&\ldots\\
&-f_1 w_n - f_R w_{n-1} + y_n - y_{n-1} \leq 0\\
&w_1, w_2, \ldots, w_n \geq 0.
\end{aligned}$$

Consider the following choices of $y_i$'s and $w_i$'s: For each $1 \leq i \leq n$, set

$$y_i = \frac{\frac{\prod_{j=i+1}^n 1 - \frac{f_1}{f_j}}{\prod_{j=i}^n 1 + \frac{f_R}{f_j}}}{\sum_{i=1}^n \frac{E_i}{f_i} \frac{\prod_{j=i+1}^n 1 - \frac{f_1}{f_j}}{\prod_{j=i}^n 1 + \frac{f_R}{f_j}}} \text{ and } w_i = \frac{y_i}{f_i}. \tag{9}$$

Then, $\sum_i b_i y_i = z^*$ in equation (7), so the result is proved if this set of choices actually constitutes a feasible solution to the dual. Clearly, the $w_i$'s are all non-negative. Now,

$$\sum_i E_i w_i = \sum_i \frac{\frac{E_i}{f_i} \frac{\prod_{j=i+1}^n 1 - \frac{f_1}{f_j}}{\prod_{j=i}^n 1 + \frac{f_R}{f_j}}}{\sum_{i=1}^n \frac{E_i}{f_i} \frac{\prod_{j=i+1}^n 1 - \frac{f_1}{f_j}}{\prod_{j=i}^n 1 + \frac{f_R}{f_j}}} = 1.$$

Also, $-f_i w_i + y_i = 0$ for each $1 \leq i \leq n$, by definition of $w_i$. Next, it must be proved that for each $1 \leq j \leq n$, $1 \leq k \leq n$,

$$-f_{|k-j|} w_j - f_R w_k + y_j - y_k \leq 0.$$

For $k > j$, we have

$$\begin{aligned} -f_{k-j}w_j + y_j &= \left(1 - \frac{f_{k-j}}{f_j}\right) \frac{\prod_{i=j+1}^{k} 1 - \frac{f_1}{f_i}}{\prod_{i=j}^{k-1} 1 + \frac{f_R}{f_i}} y_k \\ &< \left(1 + \frac{f_R}{f_k}\right) y_k. \end{aligned}$$

Thus, the only remaining case is in which $k < j$, for which we need to prove that

$$\begin{aligned} \left(1 - \frac{f_{j-k}}{fj}\right) y_j &\leq \left(1 + \frac{f_R}{f_k}\right) f_k \\ &= \left(1 + \frac{f_R}{f_k}\right) \left( \frac{\prod_{i=k+1}^{j} 1 - \frac{f_1}{f_i}}{\prod_{i=k}^{j-1} 1 + \frac{f_R}{f_i}} \right) y_j. \end{aligned}$$

In other words,

$$1 - \frac{f_{j-k}}{f_j} \leq \frac{\prod_{i=k+1}^{j} 1 - \frac{f_1}{f_i}}{\prod_{i=k+1}^{j-1} 1 + \frac{f_R}{f_i}}.$$

This is proved in Lemmas 3 and 4 in the Appendix. □

The crucial step in this proof is the validity of Lemma 3, which guarantees the feasibility of the conjectured dual solution. The validity of this Lemma depends on the exact structure of the cost function $f(\cdot)$, as will be evident in its proof.

### *B. Attaining the Dual Upper Bound*

It turns out that the RHS of (7) is the objective function value corresponding to a particular assignment of variables for the primal problem, which is however not always feasible. But when it is, it is automatically optimal due to strong duality. We now describe how to obtain this assignment of variables.

Consider the following $2n$ equations:

$$\begin{aligned} & \textit{Feasible flow constraints} && (10) \\ \lambda_{10}^* f_1 + \lambda_{21}^* f_R &= z^* E_1 && (11) \\ \lambda_{20}^* f_2 + \lambda_{21}^* f_1 + \lambda_{32}^* f_R &= z^* E_2 && (12) \\ &\cdots \\ \lambda_{k0}^* f_k + \lambda_{kk-1}^* f_1 + \lambda_{k+1k}^* f_R &= z^* E_k && (13) \\ &\cdots \\ \lambda_{n0}^* f_n + \lambda_{nn-1}^* f_1 &= z^* E_n, \\ \lambda_{10}^* - b_1 &= \lambda_{21}^* && (14) \\ \lambda_{21}^* + \lambda_{20}^* - b_2 &= \lambda_{32}^* && (15) \\ &\cdots \\ \lambda_{k-1k-2}^* + \lambda_{k-10}^* - b_{k-1} &= \lambda_{kk-1}^* && (16) \\ &\cdots \\ \lambda_{n-1n-2}^* + \lambda_{n-10}^* - b_{n-1} &= \lambda_{nn-1}^* \\ \lambda_{nn-1}^* + \lambda_{n0}^* - b_n &= 0, \end{aligned}$$

with unknowns $\lambda_{ii-1}^*$ and $\lambda_{i0}^*$ for each $1 \leq i \leq n$, and $z^*$. If these equations are solvable, and if the $\lambda_{ii-1}^*$'s and $\lambda_{i0}^*$'s are all nonnegative, then the solution to these equations clearly yields a feasible solution to the original linear program (6) (after setting all the other flows to zero). The communication flow defined by this solution consists of each node sending a part of its information to the collector node, and the rest to its nearest neighbor in the direction of the collector. The exact quantities are chosen in such a way as to equalize the normalized costs incurred by each node (if possible, otherwise the solution is not valid). This corresponds to all nodes depleting their energy supply at the same time.

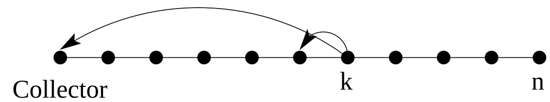


Fig. 2. The optimal communication "flow" for node $k$.

Suppose that the equations are indeed solvable and that the appropriate solutions are all nonnegative. We can then explicitly calculate the objective function $z^*$. Solving (11) and (14), we get

$$\lambda_{10}^* = \frac{\frac{E_1}{f_1} z^* + \frac{b_1 f_R}{f_1}}{1 + \frac{f_R}{f_1}}, \qquad (17)$$

$$\lambda_{21}^* = \frac{\frac{E_1}{f_1} z^* - b_1}{1 + \frac{f_R}{f_1}}. \qquad (18)$$

Substituting these in (12) and (15) and solving for $\lambda_{32}^*$, we have

$$\lambda_{32}^*(1 + \frac{f_R}{f_2}) = z^* \left( \frac{E_2}{f_2} + \frac{\frac{E_1}{f_1}(1 - \frac{f_1}{f_2})}{1 + \frac{f_R}{f_1}} \right) - \frac{b_1(1 - \frac{f_1}{f_2})}{1 + \frac{f_R}{f_1}} - b_2.$$

Repeating this calculation, we obtain the following expression for $\lambda_{kk-1}^*$,

$$\lambda_{kk-1}^* = z^* \sum_{i=1}^{k-1} \frac{E_i}{f_i} \frac{\prod_{j=i+1}^{k-1} 1 - \frac{f_1}{f_j}}{\prod_{j=i}^{k-1} 1 + \frac{f_R}{f_j}} - \sum_{i=1}^{k-1} b_i \frac{\prod_{j=i+1}^{k-1} 1 - \frac{f_1}{f_j}}{\prod_{j=i}^{k-1} 1 + \frac{f_R}{f_j}}. \qquad (19)$$

By setting $\lambda_{n+1n}^* = 0$ (since there is no node $n+1$), we then obtain the following expression for $z^*$:

$$z^* = \frac{\sum_{i=1}^{n} b_i \frac{\prod_{j=i+1}^{n} 1 - \frac{f_1}{f_j}}{\prod_{j=i}^{n} 1 + \frac{f_R}{f_j}}}{\sum_{i=1}^{n} \frac{E_i}{f_i} \frac{\prod_{j=i+1}^{n} 1 - \frac{f_1}{f_j}}{\prod_{j=i}^{n} 1 + \frac{f_R}{f_j}}}, \qquad (20)$$

which is the same as the upper bound derived in Theorem 1. The above calculations show that the equations (10) do indeed admit solutions. However, these solutions correspond to a valid flow only if the $\lambda_{k0}^*$'s and $\lambda_{kk-1}^*$'s are *non-negative*. Duality thus yields the following result.

*Theorem 2:* If the set of equations (10) have non-negative solutions, the corresponding communication flow, in which each node $i$ communicates only to the two nodes, node $i-1$ and the collector node $0$, is optimal with respect to the primal linear program (6). The normalized cost incurred by each node under this flow is the same, and consequently all nodes in the network die at the same time. The optimal functional lifetime is given by the inverse of the expression in (20).

Figure 2 shows the form of the optimal strategy. The proof of Theorem 2, depending as it does on the construction of the dual solution and the solubility of the set of equations (10), seemingly gives no intuitive reason why such a particular combination of long and short hops is optimal. In fact, it is possible to construct a direct proof that such a strategy is optimal, by employing variational arguments to show that any other strategy is suboptimal. Such a proof is much more lengthy than the one provided here. It turns out that the inequality proved in Lemma 4 is one of the key steps in this alternate proof.

Theorem 2 provides sufficient conditions which are purely in terms of the $b_i$'s and $E_i$'s. However, the form of the conditions is somewhat involved. The following lemma gives simpler sufficient conditions.

*Lemma 2:* If $\frac{b_i f_i}{E_i}$ and $E_i$ are non-decreasing in $i$, then the set of equations (10) have non-negative solutions.

**Proof.** We first prove that $\lambda^*_{k+1k} \geq 0$ for all $k$. For convenience, denote $g(i,j) := \frac{\prod_{l=i+1}^{j} 1-\frac{f_1}{f_l}}{\prod_{l=i}^{j} 1+\frac{f_R}{f_l}}$. Substituting (19), this is the same as proving that

$$z^* \geq \frac{\sum_{i=1}^{k} b_i g(i,k)}{\sum_{i=1}^{k} \frac{E_i}{f_i} g(i,k)}.$$

Cross multiplying and canceling the common terms, we need to prove that

$$\left(\sum_{i=k+1}^{n} b_i g(i,n)\right)\left(\sum_{i=1}^{k} \frac{E_i}{f_i} g(i,k)\right) \geq \left(\sum_{i=k+1}^{n} \frac{E_i}{f_i} g(i,n)\right)\left(\sum_{i=1}^{k} b_i g(i,k)\right).$$

Now, for any $k+1 \leq l \leq n$, and $1 \leq m \leq k$, by assumption we have $b_l \frac{E_m}{f_m} \geq b_m \frac{E_l}{f_l}$. This implies that the $lm^{th}$ cross term in the LHS is greater than or equal to the $lm^{th}$ cross term in the RHS, for each $l$ and $m$. This proves that $\lambda^*_{k+1k} \geq 0$ for all $k \geq 0$.

What is left to prove is that $\lambda^*_{k0} \geq 0$ for all $k$. From (17), it is clear that $\lambda^*_{10} > 0$. Solving for $\lambda^*_{k0}$ in (13) and (16), we obtain that for $k \geq 2$,

$$\lambda^*_{k0}(1 + \frac{f_R}{f_k}) = \frac{1}{f_k}(E_k z^* - (f_R + f_1)\lambda^*_{kk-1} + f_R b_k).$$

Thus, we need to prove that for $2 \leq k \leq n$,

$$E_k z^* - (f_R + f_1)\lambda^*_{kk-1} + f_R b_k \geq 0.$$

We prove this using induction on $k$. For the base case $k = 2$, we have

$$\begin{aligned} & E_2 z^* - (f_R + f_1)\lambda^*_{21} + f_R b_2 \\ & \geq E_1 z^* - (f_R + f_1)\left(\frac{\frac{E_1}{f_1} z^* - b_1}{1 + \frac{f_R}{f_1}}\right) + f_R b_2 \quad (21) \\ & = b_1 f_1 + b_2 f_R \geq 0, \end{aligned}$$

where (21) follows by substituting (18) and the assumption that $E_2 \geq E_1$. Suppose now that the induction hypothesis is true for $k-1$. Substituting for $\lambda^*_{k-1k-2}$ in (19), we get that

$$\lambda^*_{kk-1} = \left(\frac{1 - \frac{f_1}{f_{k-1}}}{1 + \frac{f_R}{f_{k-1}}}\right)\lambda^*_{k-1k-2} + \frac{\frac{E_{k-1}}{f_{k-1}} z^* - b_{k-1}}{1 + \frac{f_R}{f_{k-1}}}.$$

Then,

$$\begin{aligned} & E_k z^* - (f_R + f_1)\lambda^*_{kk-1} + f_R b_k \\ \geq \quad & E_{k-1} z^* - (f_R + f_1)\left(\frac{1 - \frac{f_1}{f_{k-1}}}{1 + \frac{f_R}{f_{k-1}}}\right)\lambda^*_{k-1k-2} \\ & -(f_R + f_1)\left(\frac{\frac{E_{k-1}}{f_{k-1}} z^* - b_{k-1}}{1 + \frac{f_R}{f_{k-1}}}\right) + f_R b_k \quad (22) \\ = \quad & \left(\frac{E_{k-1} + E_{k-1}\frac{f_R}{f_{k-1}} - \frac{f_R}{f_{k-1}} E_{k-1} - \frac{f_1}{f_{k-1}} E_{k-1}}{1 + \frac{f_R}{f_{k-1}}}\right) z^* \\ & +(f_R + f_1)\left(\frac{1 - \frac{f_1}{f_{k-1}}}{1 + \frac{f_R}{f_{k-1}}}\right)\lambda^*_{k-1k-2} + \frac{(f_R + f_1) b_{k-1}}{1 + \frac{f_R}{f_{k-1}}} \\ & + f_R b_k \\ = \quad & \left(\frac{1 - \frac{f_1}{f_{k-1}}}{1 + \frac{f_R}{f_{k-1}}}\right)(E_{k-1} z^* - (f_R + f_1)\lambda^*_{k-1k-2} + f_R b_{k-1}) \\ & + b_{k-1}\frac{f_{k-1} + f_R - f_R + \frac{f_R f_1}{f_{k-1}}}{1 + \frac{f_R}{f_{k-1}}} + b_k f_R \\ \geq \quad & f_R b_k + f_1 b_{k-1} \quad (23) \\ \geq \quad & 0, \end{aligned}$$

where (22) follows from the assumption that $E_k \geq E_{k-1}$, and (23) from the induction hypothesis. □

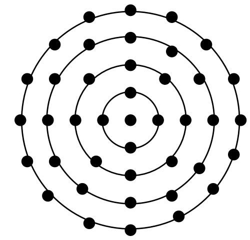

Fig. 3. The regular planar network

## IV. Functional Lifetime of a Regular Planar Network

Consider the planar network shown in Figure 3. The center node is the collector, and there are $N$ concentric circles, each containing nodes along its circumference. The $i^{th}$ ring, of radius $iR$, contains $Mi$ nodes, equally spaced along the circumference. There are thus a total of $\frac{N(N+1)}{2}M$ nodes. This particular configuration for the network is chosen simply as a convenient example of a regular planar network, due to its circular symmetry. Any other regular arrangement of nodes would admit similar results, though the analysis might be more cumbersome.

The direct approach of the last section is considerably harder in the planar case. However, we can use the following simple idea to obtain an upper bound. Suppose that all intra-ring communication, i.e., communications between nodes belonging to the same ring, have zero cost. Further, let the cost of a unit of transmission from any node in ring $i$ to any node in ring $j \neq i$ be $f((i-j)R)$, i.e. as if the distance between the two nodes were equal to the distance between the two rings (when in fact the former is larger than or equal to the latter). Since all link or hop communication costs are thereby either reduced or remain the same, it is clear that such a modified cost network would have functional lifetime greater than or equal to the original planar network. Therefore, an upper bound on the lifetime of this modified cost network is also an upper bound on the lifetime of the original network.

The treatment of this modified cost network is simplified by the fact that it is equivalent to the following linear network: Replace the $i^{th}$ ring with a super-node $s_i$, located at a distance $iR$ from the collector node. The initial energy of this super-node is $\sum_j E_{ij}$, and the number of bits initially held by it is equal to $\sum_{ij} b_{ij}$, where $E_{ij}$ and $b_{ij}$ are the amount of energy and number of bits respectively of the $j^{th}$ node in ring $i$ of the original planar network. The set of super-nodes thus forms a linear array with inter-node distance $R$. The upper bound of Theorem 1 can therefore be directly applied (with $f_i$ now denoting $f(iR)$) to obtain the following upper bound on the

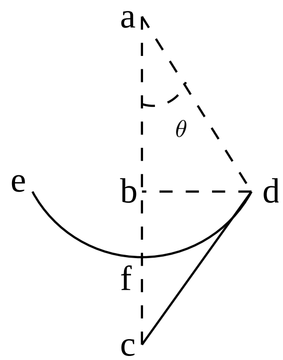


Fig. 4. Bounding the constant K.

functional lifetime $\frac{1}{z_p^*}$ of the regular planar network:

$$\frac{1}{z_p^*} \le \frac{\sum_{i=1}^{N}\left(\frac{\sum_{j=1}^{Mi} E_{ij}}{f_i}\right)\left(\frac{\prod_{j=i+1}^{n} 1-\frac{f_1}{f_j}}{\prod_{j=i}^{n} 1+\frac{f_R}{f_j}}\right)}{\sum_{i=1}^{N}\left(\sum_{j=1}^{Mi} b_{ij}\right)\left(\frac{\prod_{j=i+1}^{n} 1-\frac{f_1}{f_j}}{\prod_{j=i}^{n} 1+\frac{f_R}{f_j}}\right)}. \quad (24)$$

This upper bound is valid for any set of $E_{ij}$'s and $b_{ij}$'s.

To provide lower bounds on functional lifetime, however, we will have to impose further restrictions on the energies and information quantities of the sensor nodes. Suppose that it is further true that for each ring $i$,

$$E_{ij} = \bar{E}_i \text{ and } b_{ij} = \bar{b}_i, \text{ for all } 1 \le j \le Mi, \quad (25)$$

and that

$$\frac{\bar{b}_i f_i}{\bar{E}_i} \text{ and } i\bar{E}_i \text{ are non-decreasing in } i. \quad (26)$$

Then, Lemma 2 guarantees the achievability of the following functional lifetime for the linear network of super-nodes:

$$\frac{1}{z_p^*} \ge \frac{\sum_{i=1}^{N}\left(\frac{\bar{E}_i Mi}{f_i}\right)\left(\frac{\prod_{j=i+1}^{n} 1-\frac{f_1}{f_j}}{\prod_{j=i}^{n} 1+\frac{f_R}{f_j}}\right)}{\sum_{i=1}^{N}\left(\bar{b}_i Mi\right)\left(\frac{\prod_{j=i+1}^{n} 1-\frac{f_1}{f_j}}{\prod_{j=i}^{n} 1+\frac{f_R}{f_j}}\right)}. \quad (27)$$

Now we construct a routing scheme for the actual planar network which achieves this solution to within a constant factor. This scheme attempts to mimic the optimal scheme for the linear super-node network, with each ring equally dividing the flow that it transmits to the next inner ring among all the nodes in that inner ring. The constant factor gap arises because the distance between two nodes in successive rings is larger than the inter-ring distance.

Let $\lambda_{ii-1}^*$ and $\lambda_{i0}^*$ be the respective amounts of information communicated from super-node $s_i$ to super-node $s_{i-1}$ and to the collector node respectively, under the optimal flow. Define $A_i := \frac{\lambda_{ii-1}^*}{Mi}$, and $B_i := \frac{\lambda_{i0}^*}{Mi}$.

The nodes in the first ring communicate all their bits to the collector. This scheme will ensure that each node in the $i^{th}$ ring has a total load of $A_i + B_i$ bits, i.e., each node receives an equal amount of incoming flow, and further that each such node will incur the same cost.

Consider the $i^{th}$ ring, for $i \ge 2$. Index each node by the angle of the radial line connecting the node to the origin. These angles are multiples of $\frac{2\pi}{Mi}$. Consider the node located at the angle $\frac{2\pi j}{iM}$. Let $p$ and $q$ respectively be the quotient and remainder obtained upon dividing $j$ by $i$, i.e., $j = pi + q$. This node communicates a fraction $\frac{B_i}{A_i+B_i}$ of its total load to the collector node, and the remaining to ring $(i-1)$. Of these remaining bits, fractions $\frac{q}{i-1}$ and $\frac{i-1-q}{i-1}$ respectively are transmitted to nodes $\frac{(p(i-1)+q-1)2\pi}{M(i-1)}$ and $\frac{(p(i-1)+q)2\pi}{M(i-1)}$ of the $i-$1th ring.

This ensures that for each $i \ge 2$, each node in the $(i-1)$th ring receives a total of $\frac{i}{i-1}B_i$ bits from ring $i$. To see this, suppose that each node in the $i$th ring sends the same number of bits to the $(i-1)^{th}$ ring. In the $(i-1)^{th}$ ring, node $\frac{(p'(i-1)+q')2\pi}{M(i-1)}$ receives a fraction $\frac{q'+1}{i-1}$ from node $\frac{(p'i+q'+1)2\pi}{Mi}$ in ring $i$, and a fraction $\frac{i-1-q'}{i-1}$ from node $\frac{(p'i+q')2\pi}{Mi}$ in ring $i$, for a total of $\frac{i}{i-1}$ times the number of bits sent by each node in ring $i$ to ring $i-1$. Backwards induction from the outermost node, together with the flow conservation constraints, complete the proof.

What is the normalized cost incurred by each node? The distance between each node in the $i^{th}$ ring and the collector node is $iR$. However, the distance between each such node and the node(s) it transmits to in the $(i-1)^{th}$ ring is greater than $R$. Let the largest such distance for any ring be $KR$. In that case, each node in the $i^{th}$ ring incurs a normalized cost that is less than $K$ times the normalized cost incurred by the $i^{th}$ super-node under the optimal scheme derived above (because total load and energy are scaled by the same factor $Mi$).

The factor $K$ can be bounded as follows. Consider Figure 4. Note that the angle $\theta = \frac{2\pi}{Mi}$, $|af| = iR$, and $|fc| = R$. The two nodes in the $i^{th}$ ring nearest to node $j$ must both be within the arc $efd$. Thus, the maximum distance between $j$ and any of these nodes is no more than $|cd|$. We then have,

$$\begin{aligned}
|cd|^2 &= |bc|^2 + |bd|^2 \\
&= (iR(1-\cos\theta) + R)^2 + (iR\sin\theta)^2, \\
&= R^2(1 + 2i(i+1)(1-\cos\theta)), \\
&= R^2(1 + 2i(i+1)2\sin^2(\theta/2)), \\
&\le R^2(1 + 2i(i+1)2(\frac{\pi}{Mi})^2, \\
&\le R^2(1 + 8(\frac{\pi}{M})^2).
\end{aligned}$$

Thus,

$$K \le \frac{f(R\sqrt{1+8(\frac{\pi}{M})^2})}{f(R)}. \quad (28)$$

The results of this section are summarized in the following theorem.

*Theorem 3:* Consider the regular planar network of Figure 3. Its optimal functional lifetime is upper-bounded by the right hand side of (24). If the initial energies and information quantities in the planar network further satisfy conditions (25) and (26), then there exists a routing strategy achieving a functional lifetime of $\frac{1}{K}$ times the right hand side of (27), where $K$ is given by (28).

It should be noted that as $m \to \infty$, the factor $K$ approaches 1, yielding asymptotic optimality. This indicates that the routing strategy described above is nearly optimal in large dense networks.

## V. Scaling Implications

We now evaluate the simple routing strategy consisting of communicating only with the nearest neighbor in the direction of the collector. Because it involves only short hop communication, this scheme is an appealing one for operational reasons, such as minimizing medium access contention. However, from the point of view of lifetime, it is not clear whether it is a good routing strategy, because it places a higher load on nodes close to the sink. Thus, it is worthwhile to compare it with the optimal strategy derived in this paper. Henceforth, it will be referred to as simply the nearest neighbor strategy.

Consider first the linear case. Let $\alpha = 2$, $\gamma = 0$, and $f_R = 0$. These are the smallest values within the range we consider, so the lifetime obtained will be the largest possible. Let $E_i = E$ and $b_i = b$ for all $i$, and let the inter-node distance be $d = 1$. Under the nearest neighbor strategy, the node next to the collector is the the bottleneck, and it will transmit $nb$ bits. Hence, the functional lifetime under this scheme is $\frac{E}{bf_1 n}$.

Now, consider the lifetime corresponding to the optimal strategy. We have

$$\prod_{j=i+1}^{n} 1 - \frac{f_1}{f_j} = \prod_{j=i+1}^{n} \frac{(j-1)(j+1)}{j^2} = \frac{i(n+1)}{(i+1)n}.$$

Substituting in (24), we get

$$\begin{aligned} \frac{1}{z^*} &= \frac{E}{bf_1} \frac{\sum_{i=1}^{n} \frac{1}{i(i+1)}}{\sum_{i=1}^{n} \frac{i}{i+1}} \\ &= \frac{E}{bf_1} \frac{\sum_{i=1}^{n} \left(\frac{1}{i} - \frac{1}{i+1}\right)}{\sum_{i=1}^{n} \left(1 - \frac{1}{i+1}\right)} \\ &\approx \frac{E}{bf_1} \frac{1}{n+1} \frac{1}{1 - \frac{\log n}{n}}. \end{aligned}$$

Thus, not only is the optimal lifetime of the same order as the lifetime under the nearest neighbor strategy, but the ratio of the two converges to 1 as the size of the network goes to $\infty$. Thus, in the linear case at least, the nearest neighbor strategy is nearly optimal.

A similar calculation can be performed for the regular planar network of Figure 3. The nearest neighbor strategy in this case involves each node in the $i^{th}$ ring transmitting to its two nearest neighbors in the $(i-1)^{th}$ ring just as in the scheme described in the last section. The nodes in the first ring would then incur the greatest normalized cost, which would be $\frac{Kf(R)\sum_{i=1}^{N} ib}{E} = \frac{N(N+1)bf(R)}{2E}$. The functional lifetime associated with this scheme is thus upper bounded by $\frac{2E}{KN(N+1)bf(R)}$. On the other hand, the optimal functional lifetime can be bounded by substituting $\gamma = 0$ and $\alpha = 2$ in (24), to get

$$\begin{aligned} \frac{1}{z^*} &= \frac{E}{bf(R)} \frac{\sum_{i=1}^{N} \frac{1}{i+1}}{\sum_{i=1}^{N} \frac{i^2}{i+1}} \\ &= \frac{E}{bf(R)} \frac{\sum_{i=1}^{N} \frac{1}{i+1}}{\sum_{i=1}^{N} i - 1 + \frac{1}{i+1}} \\ &\approx \frac{E}{bf(R)} \frac{2}{N(N+1)} \frac{logN}{\frac{N-1}{N+1} - \frac{\log N}{N(N+1)}} \\ &\approx \frac{2E}{bf(R)N(N+1)} logN. \end{aligned}$$

There is thus a $\log N$ factor improvement in this case. Note that for larger values of $\alpha$, the ratio between the gap in performance between the optimal scheme and the nearest neighbor scheme will further reduce.

Some conclusions follow from these scaling results. First, the lifetime scales down sharply with the size of the network. This is not very surprising in itself, but underscores one of the problems with the "many sensors one collector" network configuration. This seems to be one of the major tradeoffs to be made in sensor networks, since collector nodes, which are essentially gateways to a larger and more powerful communication infrastructure, are more expensive to maintain, whereas sensor nodes are cheap. On the other hand, the number of sensor nodes per collector node is limited by such lifetime and throughput considerations. One of the motivations for performing in-network processing is precisely to limit the communication burden on nodes which are close to the collector.

## VI. Conclusions and Future Work

The main contribution of this paper is to provide sharp bounds, and in some cases exact solutions, to the functional lifetime problem for spatially regular networks. Even here, some gaps remain. The characterization of the optimal solution holds only under certain conditions on the energies and traffic requirements of the network. Nevertheless, we believe that these conclusions hold broadly for large sensor networks.

Another interesting extension involves considering more general cost functions. In practice, power levels belong to a discrete set, and all pairs of nodes may not be connected. Such effects could be modeled by suitably changing the cost function. The key property of the cost function that underlies all our results is the inequality in Lemma 4. It would be worth investigating if a similar result holds for more general functions.

One major limitation of our model is that we treat information as incompressible, whereas *in-network processing* [8] and compression of correlated sources might substantially reduce the relaying burden. Indeed, our results suggest the need for such techniques to increase lifetime.

## Acknowledgments

This material is based upon work partially supported by AFOSR under Contract No. F49620-02-1-0217, NSF under Contract Nos. NSF ANI 02-21357 and CCR-0325716, USARO under Contract Nos. DAAD19-00-1-0466 and DAAD19-01010-465, DARPA/AFOSR under Contract No. F49620-02-1-0325, and DARPA under Contact Nos. N00014-0-1-1-0576.

## References

[1] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *WSNA '02: Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*. ACM Press, 2002, pp. 88–97.

[2] J.-H. Chang and L. Tassiulas, "Maximum lifetime routing in wireless sensor networks," *IEEE Trans. Netw.*, vol. 12, no. 4, pp. 609–619, 2004.

[3] A. Sankar and Z. Liu, "Maximum lifetime routing in wireless ad-hoc networks," in *Proceedings of IEEE Infocom '04*, Hong Kong, 2004, pp. 1089–1097.

[4] K. Kalpakis, K. Dasgupta, and P. Namjoshi, "Efficient algorithms for maximum lifetime data gathering and aggregation in wireless sensor networks," *Comput. Networks*, vol. 42, no. 6, pp. 697–716, 2003.

[5] R. Madan and S. Lall, "Distributed algorithms for maximum lifetime routing in wireless sensor networks," in *Proceedings of IEEE Globecom '04*, Dallas, 2004, pp. 748–753.

[6] M. Bhardwaj, A. Chandrakasan, and T. Garnett, "Upper bounds on the lifetime of sensor networks," in *Proceedings of IEEE ICC 2001*, Helsinki, Finland, 2001, pp. 785–790.

[7] S. Narayanaswamy, V. Kawadia, R. Sreenivas, and P. Kumar, "Power control in ad-hoc networks: Theory, architecture, algorithm and implementation of the COMPOW protocol," in *Proceedings of European Wireless Conference 2002*, Florence, Italy, 2002, pp. 156–162.

[8] A. Giridhar and P. Kumar, "Computing and communicating functions over sensor networks," *IEEE Journal on Selected Areas of Communication*, to appear.

## APPENDIX

*Lemma 3:* For any $1 \leq k < j$,

$$1 - \frac{f_{j-k}}{f_j} \leq \frac{\prod_{i=k+1}^{j} 1 - \frac{f_1}{f_i}}{\prod_{i=k+1}^{j-1} 1 + \frac{f_R}{f_i}},$$

where $f_i := (id)^\alpha e^{\gamma(id)}$, with $\alpha \geq 2$ and $\gamma \geq 0$.

**Proof.** We prove this for fixed $k$, by induction on $j$. For $j = k+1$, the LHS and the RHS are the same, so the inequality is true with equality. Suppose it is true for some $j > k$. From the induction hypothesis,

$$\frac{\prod_{i=k+1}^{j+1} 1 - \frac{f_1}{f_i}}{\prod_{i=k+1}^{j} 1 + \frac{f_R}{f_i}} \quad \geq \quad \left(1 - \frac{f_{j-k}}{f_j}\right)\left(\frac{1 - \frac{f_1}{f_{j+1}}}{1 + \frac{f_R}{f_j}}\right).$$

It is thus enough to prove that

$$\left(1 - \frac{f_{j-k}}{f_j}\right)\left(\frac{1 - \frac{f_1}{f_{j+1}}}{1 + \frac{f_R}{f_j}}\right) \geq \left(1 - \frac{f_{j+1-k}}{f_{j+1}}\right).$$

Cross multiplying and rearranging terms, this is equivalent to proving that

$$f_{j-k}f_1 - f_1 f_j + f_j f_{j+1-k} \geq f_{j+1}f_{j-k} + f_R(f_{j+1} - f_{j-k+1}). \quad (29)$$

Let $a := k, b := j-k, c := 1$. Adding and subtracting $f_{j+1-k}f_{j-k}$, and canceling out the common factor $d^{2\alpha}$ (recall that $f_i = f(id)$), this is equivalent to proving that

$$((a+b)^\alpha e^{\gamma(a+b)} - b^\alpha e^{\gamma b})((b+c)^\alpha e^{\gamma(b+c)} - c^\alpha e^{\gamma c}) \geq$$
$$((a+b+c)^\alpha e^{\gamma(a+b+c)} - (b+c)^\alpha e^{\gamma(b+c)})(b^\alpha e^{\gamma b} + f'_R),$$

where $f'_R := \frac{f_R}{E_t d^\alpha}$. This in turn is proved in Lemma 4, which follows. □

*Lemma 4:* Given any real numbers $a \geq 0$, $b \geq c \geq 1$, $\alpha \geq 2$, $\gamma \geq 0$, $K \leq 1 - (1/2)^{\alpha-1}$,

$$((a+b)^\alpha e^{\gamma(a+b)} - b^\alpha e^{\gamma b})((b+c)^\alpha e^{\gamma(b+c)} - (c)^\alpha e^{\gamma(c)}) \geq$$
$$((a+b+c)^\alpha e^{\gamma(a+b+c)} - (b+c)^\alpha e^{\gamma(b+c)})(b^\alpha e^{\gamma b} + K).$$

**Proof.** The common factors $e^{\gamma(b+c)}$ can be cancelled out from both sides. Denote the left hand side and right hand side of the resulting inequality, as functions of the variable $a$, respectively by $LHS(a)$, and $RHS(a)$.

$F(a) := LHS(a) - RHS(a)$ is a differentiable function of $a$, and $F(0) = 0$. It is enough to prove that $\frac{dF(a)}{da} > 0$ for all $a > 0$. We have

$$\frac{dLHS(a)}{da} = (\gamma(a+b) + \alpha)(a+b)^{\alpha-1}e^{\gamma a}((b+c)^\alpha e^{\gamma b} - c^\alpha).$$

Also,

$$\frac{dRHS(a)}{a} = (\gamma(a+b+c) + \alpha)(a+b+c)^{\alpha-1}e^{\gamma a}(b^\alpha e^{\gamma b} + K).$$

We can cancel out the common factor $e^{\gamma a}$. Letting $L(a) := e^{-\gamma a}\frac{dRHS(a)}{a}$, $R(a) := e^{-\gamma a}\frac{dLHS(a)}{a}$, we have

$$\begin{aligned} L(a) \quad = \quad & (\alpha + \gamma(a+b))(a+b)^{\alpha-1}b(b+c)^{\alpha-1}e^{\gamma b} \\ & + \gamma bc(a+b)^{\alpha-1}(b+c)^{\alpha-1}e^{\gamma b} \\ & + (\gamma a + \alpha)c(a+b)^{\alpha-1}(b+c)^{\alpha-1}e^{\gamma b} \\ & - (\alpha + \gamma(a+b))(a+b)^{\alpha-1}c^\alpha). \end{aligned} \quad (30)$$

Let the four summands in (30) be $L_1, L_2, L_3, L_4$. Now,

$$\begin{aligned} R(a) \quad = \quad & (\gamma(a+b) + \alpha)(a+b+c)^{\alpha-1}e^{\gamma a}b^\alpha e^{\gamma b} \\ & + \gamma c(a+b+c)^{\alpha-1}e^{\gamma a}b^\alpha e^{\gamma b} \\ & + K(\alpha + \gamma(a+b+c))(a+b+c)^{\alpha-1}. \end{aligned} \quad (31)$$

Let the three summands in (31) be $R_1, R_2, R_3$. Since $(a+b)(b+c) > (a+b+c)b$, we have $L_1 > R_1$ and $L_2 > R_2$. It remains to prove that $L_3 + L_4 \geq R_3$. Now,

$$\begin{aligned} L_3 \quad & \geq \quad (\alpha + \gamma a)(1 + \gamma b)c(a+b)(b+c)^{\alpha-1} \\ & = \quad (\alpha + \gamma a + \gamma\alpha b + \gamma^2 ab)c(a+b)(b+c)^{\alpha-1} \\ & \geq \quad (\alpha + \gamma(a+b+c))c(a+b)(b+c)^{\alpha-1}, \end{aligned} \quad (32)$$

since $\alpha \geq 2$ and $b \geq c$. Thus,

$$L_3 + L_4 > (\alpha + \gamma(a+b+c))c(a+b)^{\alpha-1}((b+c)^{\alpha-1} - c^{\alpha-1}).$$

Since $c \geq 1$, it remains to prove that

$$(a+b)^{\alpha-1}((b+c)^{\alpha-1} - c^{\alpha-1}) \geq K(a+b+c)^{\alpha-1}. \quad (33)$$

Consider the expression $P(a) := \frac{(a+b)^{\alpha-1}((b+c)^{\alpha-1} - c^{\alpha-1})}{(a+b+c)^{\alpha-1}}$. Differentiating with respect to $a$, we get

$$\begin{aligned} \frac{dP(a)}{da} \quad & = \quad (\alpha - 1)\frac{c(a+b)^{\alpha-2}((b+c)^{\alpha-1} - c^{\alpha-1})}{(a+b+c)^\alpha} \\ & \geq \quad 0. \end{aligned} \quad (34)$$

Therefore, the ratio is minimum at $a = 0$. Thus,

$$\begin{aligned} P(a) \quad & \geq \quad P(0) \\ & = \quad b^{\alpha-1}(1 - \frac{c^{\alpha-1}}{(b+c)^{\alpha-1}}) \\ & \geq \quad 1 - (\frac{1}{2})^{\alpha-1} \geq K. \end{aligned}$$

□

Note: The conditions on $a, b, c$ and $K$ for the above inequality to hold may seem somewhat arbitrary. This is mainly due to the presence of the constant $K$. If $K = 0$, the inequality has a more symmetric structure, and is true for any $a, b, c \geq 0$, and $\alpha > 1$. If, further, the constant $\gamma = 0$, the inequality follows in straightforward fashion from Jensen's inequality, due to the convexity of the function $x^\alpha$ (but not *only* from its convexity; the inequality does not hold for all convex functions). We are not aware if any form of this inequality is already known.