Gathering radio messages in the path

Jean-Claude Bermond, Ralf Klasing, Nelson Morales, Stéphane Pérennes, Patricio Reyes

Gathering radio messages in the path*

Jean-Claude Bermond† Ralf Klasing‡ Nelson Morales§‖ Stéphane Pérennes†
Patricio Reyes¶‖

November 2, 2012

Abstract

In this paper, we address the problem of gathering information in one node (sink) of a radio network where interference constraints are present: when a node transmits, it produces interference in an area bigger than the area in which its message can actually be received. The network is modeled by a graph; a node is able to transmit one unit of information to the set of vertices at distance at most d_T in the graph, but when doing so it generates interferences that do not allow nodes at distance up to d_I ($d_I \geq d_T$) to listen to other transmissions. We are interested in finding a gathering protocol, that is an ordered sequence of rounds (each round consists of non-interfering simultaneous transmissions) such that $w(u)$ messages are transmitted from any node u to a fixed node called the sink. Our aim is to find a gathering protocol with the minimum number of rounds (called *gathering time*).

In this article, we focus on the specific case where the network is a path with the sink at an end vertex of the path and where the traffic is unitary ($w(u) = 1$ for all u); indeed this simple case appears to be already very difficult. We first give a new lower bound and a protocol with a gathering time that differ only by a constant independent of the length of the path. Then we present a method to construct incremental protocols. An incremental protocol for the path on $n + 1$ vertices is obtained from a protocol for n vertices by adding new rounds and new calls to some rounds but without changing the calls of the original rounds. We show that some of these incremental protocols are optimal for many values of d_T and d_I (in particular when d_T is prime). We conjecture that this incremental construction always gives optimal protocols. Finally, we derive an approximation algorithm when the sink is placed in an arbitrary vertex in the path.

Keywords: Radio networks, wireless networks, interference, paths, gathering, approximation algorithms.

*This work was partially supported by the ANR projects AGAPE, ALADDIN, GRATEL, IDEA, and by the INRIA project CEPAGE.

†Mascotte Project, INRIA–I3S(CNRS/UNSA), Sophia Antipolis, France.
E-mails: jean-claude.bermond@inria.fr, stephane.perennes@inria.fr

‡CNRS - LaBRI - Université de Bordeaux, 351 cours de la Libération, 33405 Talence cedex, France.
E-mail: klasing@labri.fr

§Delphos Mine Planning Lab, Advanced Mining Technology Center, University of Chile.
E-mail: nmorales@amtc.cl

¶Department of Statistics, Universidad Carlos III de Madrid, 28903 Getafe Madrid, Spain.
E-mail: preyes@est-econ.uc3m.es

‖Research of these authors started while they were in MASCOTTE Project - INRIA(I3S, CNRS/UNSA) and were supported by CONICYT(Chile)/INRIA.

1 Introduction

1.1 Motivation and modeling aspects

In radio networks, a set of radio devices communicate by using radio transmissions which, depending on the technology used, are subject to different interference constraints (see for instance [Bia00, Gal04, Müh02] for 802.11). This means that only certain transmissions can be performed simultaneously, therefore the devices have to act in a cooperative manner in order to achieve an effective flow of information in the network. In this context, we study a problem proposed by France Telecom, about "how to provide Internet to villages" where there is no high speed wired access (see [BBS05]). The houses of the village are equipped with radio devices and they want to access the *rest of the world* via Internet. For that purpose, they have to send (and receive) information via a gateway called *sink*, connected to the internet (for example via a satellite antenna).

Unlike wired networks, when a node u transmits a message, it does not use a resource as simple as some capacity on a link; instead it produces a signal that may prevent other transmissions to occur. The set of possible concurrent transmissions follows from a complex n-ary interference relation which properly models the idea that the noise intensity must be small enough compared to the signal intensity. In order to obtain tractable models, a widely used simplification consists in associating to each node a *transmission area* in which it can transmit a message and an *interference area* in which it produces a strong noise (see [SW06]). Then, the communication from a node u to a node v is possible if v is in the transmission area of u, and no third node transmitting has v in its interference area. Note that, by doing so, we replace the n-ary relation with a binary relation: two (possible) transmissions can be performed concurrently when they do not interfere with each other.

Like in several other articles, we choose to model the network by an undirected graph $G = (V, E)$, where V is the set of devices in the network, and to use as distance the distance between nodes in the graph. The transmission area (resp. interference area) is then modeled by the ball of radius d_T (resp. $d_I \geq d_T$). Time is considered discrete, i.e., divided into time-steps of fixed length. During a time-step a node might transmit exactly one message to another node at distance at most d_T. Transmissions (called *calls*) can be done simultaneously if they do not interfere, that is if the sender of one call is at distance at least $d_I + 1$ from the receiver of any other call. Note that some authors use the Euclidean distance; but in the case we consider in this paper (i.e. path with nodes equally spaced), these distances are equivalent.

Under this model, the problem raised by France Telecom consists of finding a gathering protocol, that is an ordered sequence of rounds (each round consists of non-interfering simultaneous transmissions) such that $w(u)$ messages are transmitted from node u to a fixed node called the sink. Our aim is to find the minimum number of rounds called gathering time or makespan.

Note that we may as well study the converse problem called personalized broadcast for which we need to send personalized information from the central node to each node. Like in many other communication models, we can simply reverse the order of the communication steps and the direction of the calls to get that gathering and personalized broadcast are formally equivalent; indeed, if two calls are compatible, their reverse calls are also compatible, as we consider undirected graphs with symmetric distances. Due to this equivalence, all the results (algorithms, complexity, bounds) that we give are also valid for personalized broadcast. Here, we focus on gathering issues.

1.2 Related work

Basic communication problems for the dissemination of information (like gathering, broadcasting, gossiping) have been widely studied in classical interconnection networks (see the book [HKP$^+$05]).

The broadcasting and gossiping problems in radio networks with $d_T = d_I = 1$ are studied in [CW91, EK04, GP02] and [CGR02, GM03, CGL02, BGP98, BGRV98], respectively. Note that broadcasting is different from our problem of personalized broadcasting, because in a broadcast the same information has to be transmitted to all the other nodes and therefore flooding techniques can be used.

The problem of data aggregation has been widely studied in particular in sensor networks. But data aggregation usually involves the fusion of data from multiple sensors at intermediate nodes and transmission of the aggregated data to the sink (see for example [KDN03, RV06]) and so it differs from the gathering problem. Furthermore, in many papers authors are interested in minimizing the battery energy more than the makespan.

With respect to the gathering problem (see [BKK$^+$10] for a survey) different cases have been studied. In [BGK$^+$06b] a protocol for general graphs with an approximation factor of at most 4 is presented. An extension of this problem where messages can be released over time is studied in [BKMS08] and a 4-approximation is presented. Using the same interference model, a relaxed approach has been studied in [KMP08] where the problem is studied in terms of collecting the flow demands. The unitary case (where each node has one unit of information to transmit) has been considered under different topologies. The unitary case in the path with $d_T = 1$ and arbitrary d_I is studied in [BCY09]. For the two-dimensional square grid, optimal solutions are provided in [BP12]. For trees, in the case $d_T = d_I = 1$, an optimal solution is presented in [BY10].

In some papers the authors add the constraint that no buffering is allowed in intermediate nodes; this constraint comes from the application in sensor networks. In sensor networks (see [Gar07] for a survey), a model close to ours is considered in [FFM04]. Here, they consider mainly uni-directional antennas and the so-called *primary node interference model* where a node cannot receive and transmit during the same time slot. In [FFM04] they give optimal gathering protocols for paths and trees. The results have been extended to general graphs in the unitary case in [GR06, GR09] where a polynomial algorithm is given. The problem is solved for larger d_I for trees in [BGR10, BGP$^+$11]. Some articles consider symmetric interference models, due to the fact that in the protocol 802.11 when a message is transmitted, acknowledgments need to be transmitted as well. Indeed, in some papers this model is called the *802.11 interference model* (see [Wan09]).

Preliminary versions of some of the results of this paper appeared in [BGK$^+$06a].

1.3 Our results

In this article, we focus on the specific case where the network is a path with the sink at an end-vertex of the path and where the traffic is unitary ($w(u) = 1$ for all u). The hypotheses of the model and the restriction to a path are strong requirements. However, let us note that this simple case appears to be already very difficult and we were unable to find for general d_T a closed formula or at least a polynomial algorithm when the sink is an end-vertex. From a theoretical point of view, it is interesting to know for which topologies of networks a communication problem can be solved polynomially. Most of the communication or routing problems are usually "easy" on paths, trees and other simple topologies; but here it is not the case. Note also that paths are not far from real situations such as remote places where houses (farms) are situated along a unique road. Finally, the values of the gathering time we obtain give lower bounds for the corresponding real-life values. Put differently, given a value for the completion time (or equivalently the bandwidth given to each user), our results give an upper bound on the number of possible

users in the network.

In Section 2, we introduce the notation and precise the problem to solve. In Section 3, we present the classical lower bound and an algorithm which turns out to be optimal when $d_I = (p+1)d_T - 1$ for some positive integer p. In Section 4, we give a new lower bound and we show that the preceding algorithm gives a 1^+-approximation (the number of rounds differs from the new lower bound by a constant independent of the length of the path). Then in Section 5, we give a procedure such that, given a gathering protocol for the path of length n, it produces a solution for the path of length $n+1$. We call this procedure *incremental* as it does not modify the solution for n, but it only adds extra calls and rounds to gather the additional message (sent by the extra vertex). We show that some of these incremental protocols are optimal for many values of d_T and d_I. In particular, if $d_I = pd_T + q$ with $q < d_T$, then we obtain, for n large enough, optimal protocols when d_I and $q+1$ are relatively prime. This gives optimal solutions for $d_T = 2, 3, 5$. We conjecture that an optimal solution can always be obtained for general d_T by the incremental procedure. If the conjecture is true, it would give us an optimal protocol that can be computed in polynomial time in the length of the path. Finally, in Section 6, we extend the protocol and lower bounds of Section 4 to arbitrary positions of the sink obtaining a 1^+-approximation.

2 Preliminaries

In this section, we introduce the model and main notation, and we formally state the problem to solve.

2.1 The model: definitions and notation

In the whole paper, we are given a graph $G = (V, E)$ with n vertices and with a distinguished vertex $t \in V$ called the *sink*, and two integers d_I, d_T, such that $d_I \geq d_T > 0$, where d_I is the *interference distance* and d_T is the *transmission distance*. The distance between two vertices u and v is the length of the shortest path from u to v and is denoted by $d(u, v)$.

A *call* is a couple (s, r) with $s, r \in V$, $0 < d(s, r) \leq d_T$, and where s is the *sender* and r the *receiver*. We denote the call (s, r) as $s \to r$. Call $s \to r$ *interferes* with call $s' \to r'$ if $d(s, r') \leq d_I$ or $d(s', r) \leq d_I$. We say that the two calls $s \to r$ and $s' \to r'$ are *compatible* if they do not interfere, that is both $d(s, r') > d_I$ and $d(s', r) > d_I$. During one unit of time only one (unitary) message can be transmitted during a call.

A *round* is a set of compatible calls. If R is a round and $s \to r \in R$ is a call, we say that $s \to r$ is performed during round R, and this corresponds to the sender s transmitting a message to receiver r if there is one message available. If only one call is performed in a round, we say that the round is *singleton.* In the example of Figure 1 where $d_I = 2$ round 1 consists of the two calls $1 \to 0$ and $5 \to 4$; they are compatible as the distance $d(1, 4) = 3 > d_I$. Note that $1 \to 0$ and $4 \to 3$ are not compatible as $d(1, 3) = 2 = d_I$. In the example round 6 is a singleton consisting of the unique call $2 \to 1$.

In the *gathering problem*, every node $u \in V$ has $w(u)$ unitary pieces of information (called shortly *messages*) which have to reach the sink t, where $w(u)$ is a non-negative integer. A *gathering protocol* is an ordered sequence of rounds that allows to gather the information of the nodes in the sink.

We will often specify protocols by giving simply the sequence of rounds, without specifying which message is sent (indeed, when gathering, which message is sent from a vertex in a given round is irrelevant if the vertex has several messages to transmit). Also, observe that, when gathering, it is not useful to have multiples copies of a message in different vertices: it suffices to keep the copy that reaches the sink first. This allows us to consider

simply calls of the type $s \to r$, meaning that the sender can select a unique receiver between the potential ones.

2.2 The Minimum Time Gathering Problem

Let us now precise the problem to solve. We call it the Minimum Time Gathering problem. The input of the problem is given by a tuple (G, w, t, d_I, d_T) with

1. A base graph $G = (V, E)$.
2. A *sink* $t \in V$.
3. A weight function $w : V \to \mathbb{N}$, $w(u)$ being the number of messages to gather from vertex u into the sink t.
4. A transmission distance a positive integer $d_T \geq 1$.
5. An interference distance a positive integer $d_I \geq d_T$.

Definition 1 (Gathering protocol) *A* gathering protocol *(or simply* protocol*) is an ordered sequence of rounds such that, once all the rounds of the protocol are executed, exactly $w(u)$ messages have been gathered from each vertex $u \in V$ into the sink t.*

The goal of the Minimum Time Gathering problem is to find a protocol that requires a minimum number of rounds, called the *gathering number*.

Definition 2 (Gathering number) *Given an instance (G, w, t, d_I, d_T) of the gathering problem, the minimum number of rounds for any gathering protocol for the instance will be called the* gathering number *and will be denoted as $g_{d_I,d_T}(G, w, t)$.*

For example, Figure 1 shows an optimal gathering protocol using 18 rounds for a path with 7 vertices (each having one piece of information), with $d_T = 1$, $d_I = 2$ and sink $t = 0$.

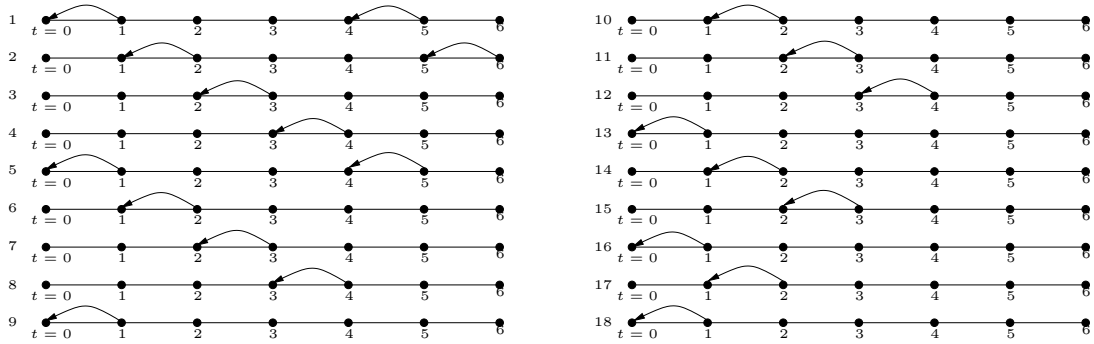

Figure 1: A gathering protocol in the path P_7 when $d_T = 1, d_I = 2$ and every vertex has one message to send to the sink $t = 0$.

In this article, we restrict ourselves to the case where $G = P_n$, the path with n vertices. Formally, P_n is the graph with vertex set $\{0, 1, \ldots, n-1\}$ and edges between vertices i and j if and only if $|i - j| = 1$. We consider protocols only for Unitary Minimum Time Gathering which is the case where $w(u) = 1$ for all $u \neq t$. However, lower bounds are given for general values of w.

In the rest of the paper, we suppose d_T and d_I, $d_I \geq d_T$, are given. Let $\mathbf{d_I = pd_T + q}$ with p and q integers, $p \geq 1$ and $0 \leq q < d_T$.

Table 1 summarizes important notation used in this paper.

V	The set of devices in the path.
$G = (V, E)$	The *base graph*.
P_n	The path with n vertices $\{0, 1, 2, \ldots, n-1\}$.
$t \in V$	The *sink*.
$d(u, v)$	Distance between vertices $u \in V$ and $v \in V$.
$d_T, d_T > 0$	The *transmission distance*. d_T positive integer.
$d_I, d_I \geq d_T$	The *interference distance*. d_I positive integer.
p, q	$d_I = pd_T + q$. p and q integers, $p > 0$ and $0 \leq q < d_T$.
D	$D = d_I + d_T + 1$.
$w : V \to \mathbb{N}$	Number of messages that must be gathered from a node.
w unitary	If $w(u) = 1\,(\forall u \neq t)$.
$m(u)$	Any message originated in vertex $u \in V$.
$s \to r$	A call from vertex $s \in V$ to vertex $r \in V$, $s \neq r, d(s, r) \leq d_T$.
R	A round, i.e., a set of compatible calls.
(G, w, t, d_I, d_T)	An instance of the gathering problem. Shortened to (G, w, t) when clear. Shortened to (G, t) if w is unitary.
$g_{d_I,d_T}(P_n, w, t)$	The length of the shortest gathering protocol for the instance (P_n, w, t). Shortened to $g_{d_I,d_T}(P_n, w)$ if $t = 0$. Simply $g_{d_I,d_T}(P_n, t)$ if w is unitary.
$g_{d_I,d_T}(P_n)$	$g_{d_I,d_T}(P_n, w, t)$ with w unitary and $t = 0$.
$A(P_n)$	A gathering protocol for the path P_n with w unitary and $t = 0$.
$f : X \to X$	$x \mapsto f(x) = [(x + q) \mod d_T] + 1$, with $X = \{1, 2, \ldots, d_T\}$.

Table 1: Summary of the general notation.

3 Lower Bounds and Simple Protocols for the sink as an end-vertex

In the rest of the paper (except the last section), we suppose the sink is the end-vertex of the path, i.e. $t = 0$. We will use the simplified notation $g_{d_I,d_T}(P_n, w)$ for $g_{d_I,d_T}(P_n, w, t = 0)$ and $g_{d_I,d_T}(P_n)$ when w is unitary. We will also denote by $A(P_n)$ instead of $A_{d_I,d_T}(P_n, w, t = 0)$ a gathering protocol that gathers one message from each vertex $i \neq 0$ into the sink $t = 0$. We will also use intensively the notation $\mathbf{D = d_I + d_T + 1}$.

3.1 A first lower bound

In [BGK+06b] the authors give a general lower bound which is presented in the following proposition for the path P_n with the sink at vertex 0 and general weights w.

Proposition 1 ([BGK+06b]) *We have* $g_{d_I,d_T}(P_n, w) \geq \mathrm{LB}_0(P_n, w)$, *where*

$$\mathrm{LB}_0(P_n, w) = \sum_{i \leq d_I+1} w(i) \left\lceil \frac{i}{d_T} \right\rceil + \left\lceil \frac{d_I + 2}{d_T} \right\rceil \sum_{i > d_I+1} w(i)\,.$$

Note that the bound can be easily derived, in the case of the path, by noting that there is at most one call (s, r) per round with $r \leq d_I + 1$.

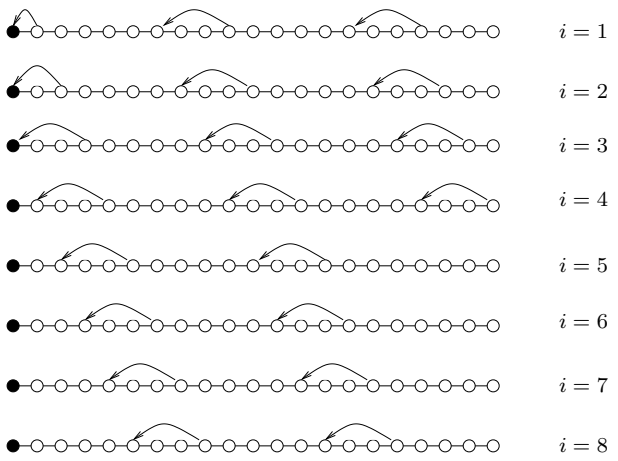

Figure 2: The rounds $\{i + kD \to \max[0, i + kD - d_T] : k \geq 0, i + kD \leq n - 1\}$ for $1 \leq i \leq D$ in P_{21} when $t = 0, d_I = 4, d_T = 3$ and hence $D = 8$.

3.2 Optimal Protocols for P_n, $n \leq (p+1)d_T + 1$

Using a greedy protocol we can obtain the value of $g_{d_I,d_T}(P_n)$ for small n, more precisely for $n \leq (p+1)d_T + 1$.

Proposition 2 *Let $d_I = pd_T + q$, $0 \leq q < d_T$. For $n \leq (p+1)d_T + 1$, $g_{d_I,d_T}(P_n, w) = \sum_{i \leq n-1} w(i) \left\lceil \frac{i}{d_T} \right\rceil$.*

Proof: From Proposition 1, by noting that $\lceil \frac{d_I+2}{d_T} \rceil \geq p + 1$ and that for $d_I + 2 \leq i \leq n - 1 = (p+1)d_T$ we have $\lceil \frac{i}{d_T} \rceil = p + 1$, we obtain $\mathrm{LB}_0(P_n, w) \geq \sum_{i \leq n-1} w(i) \lceil \frac{i}{d_T} \rceil$.

Now the bound is attained by considering the greedy protocol consisting of singleton rounds of length d_T if possible. More precisely, for a message located at a vertex $i = \alpha d_T + \beta$ with $1 \leq \beta \leq d_T$, the protocol performs $\alpha + 1 = \lceil \frac{i}{d_T} \rceil$ rounds which are $i - jd_T \to i - (j+1)d_T$ for $0 \leq j \leq \alpha - 1$ and $(i - \alpha d_T = \beta) \to 0$. ■

3.3 A simple gathering protocol

In the rest of this section, we consider only unitary weights. The algorithm we describe is very similar to the general algorithm of [BGK+06b] which gives a $\frac{3}{2}$-approximation in the particular case of P_n. But as we consider only the unitary case ($w(u) = 1, \forall u \neq 0$), our algorithm is simpler. Furthermore, it will be sufficient to solve completely the case $q = d_T - 1$ ($d_I = pd_T + d_T - 1$), and that implies a general 1^+-approximation. This case can also be viewed as an extension of the algorithm given in [BCY09] for $d_T = 1$ (recall $D = d_T + d_I + 1$).

The algorithm consists of 2 phases: a loop that reduces the instance into an instance of P_k, where $k \leq D$ and a simple greedy gathering for P_k.

Within the loop, we use the set of calls $\{i + kD \to \max[0, i + kD - d_T] : k \geq 0, i + kD \leq n - 1\}$ defined for $i = 1, \ldots, D$ (see Figure 2 for an example on P_{21}). These sets of calls form rounds; indeed, the calls of each set are compatible, because the distance between two consecutive transmitters is D.

Input: n, d_I, d_T
1 **while** $n-1 \geq D$ **do**
2 | **for** $i = 1$ *to* D **do**
| | Apply the round $\{i + kD \to \max[0, i + kD - d_T] : k \geq 0, i + kD \leq n-1\}$
| **end**
| $n \leftarrow n - d_T$
end
3 Gather each message independently, using calls of maximum length.

Algorithm 1: Algorithm A_1 solves gathering in P_n for $t = 0$

Proposition 3 *For the path P_n, Algorithm A_1 gathers in $|A_1(P_n)|$ rounds, where*

$$|A_1(P_n)| = \begin{cases} |A_1(P_{n-d_T})| + D & \text{if } n-1 \geq D, \\ \sum_{i=1}^{n-1} \left\lceil \frac{i}{d_T} \right\rceil & \text{if } n-1 < D \end{cases}$$

Proof: Clearly, the result holds if $n-1 < D$, thus we focus on the case $n-1 \geq D$. For $n-1 \geq D$, we have that each iteration of the inner `for loop` (Step 2) requires D rounds and transforms the instance $(P_n, 0)$ into the instance $(P_{n-d_T}, 0)$, hence the claim. ■

3.4 Case $q = d_T - 1$ $(d_I = pd_T + d_T - 1)$

In the case of $q = d_T - 1$, we can give exact values as we will see that $\mathrm{LB}_0(P_n)$ and $|A_1(P_n)|$ are equal. This case also generalizes the results given in [BCY09] for $d_T = 1$ as $q < d_T$ implies $q = 0 = d_T - 1$.

Proposition 4 *If $d_I = pd_T + d_T - 1$ $(q = d_T - 1)$, and $n > (p+1)d_T + 1$, then $g_{d_I,d_T}(P_n) = \mathrm{LB}_0(P_n) = \sum_{i \leq d_I+1} \left\lceil \frac{i}{d_T} \right\rceil + (p+2)(n - d_I - 2)$.*

Proof: Let $n > (p+1)d_T + 1$, and let k be such that $n = D - \gamma + kd_T$, where $0 \leq \gamma < d_T$, then by Algorithm A_1

$$|A_1(P_n)| = kD + \sum_{i=1}^{D-\gamma-1} \left\lceil \frac{i}{d_T} \right\rceil = kD + \sum_{i=1}^{d_I+1} \left\lceil \frac{i}{d_T} \right\rceil + \sum_{i=d_I+2}^{D-\gamma-1} \left\lceil \frac{i}{d_T} \right\rceil$$

But, for $d_I + 2 \leq i \leq D - \gamma - 1$, $\left\lceil \frac{i}{d_T} \right\rceil = p + 2$ and so $|A_1(P_n)| = k(p+2)d_T + (p+2)(D - \gamma - d_I - 2) + \sum_{i \leq d_I+1} \left\lceil \frac{i}{d_T} \right\rceil = (p+2)(n - d_I - 2) + \sum_{i \leq d_I+1} \left\lceil \frac{i}{d_T} \right\rceil$.

This matches $\mathrm{LB}_0(P_n)$ because when $q = d_T - 1 \Rightarrow d_I = (p+1)d_T - 1$ and hence

$$\mathrm{LB}_0(P_n) = \sum_{i \leq d_I+1} \left\lceil \frac{i}{d_T} \right\rceil + (n - 1 - d_I - 1) \left\lceil \frac{d_I + 2}{d_T} \right\rceil = (p+2)(n - d_I - 2) + \sum_{i \leq d_I+1} \left\lceil \frac{i}{d_T} \right\rceil.$$

■

For the other cases $q \neq d_T - 1$, we obtain that for $n > (p+1)d_T + 1$, $\mathrm{LB}_0(P_n)$ and $|A_1(P_n)|$ are different. Indeed when n increases by 1, $\mathrm{LB}_0(P_n)$ increases by $p+1$ (as $q < d_T - 1$, then $\left\lceil \frac{d_I+2}{d_T} \right\rceil = p + 1$) and so when n increases by d_T, $\mathrm{LB}_0(P_n)$ increases by $(p+1)d_T$, but $|A_1(P_n)|$ increases by $D > (p+1)d_T$.

4 A new lower bound and a 1^+-approximation

In this section we give another lower bound which increases by D when n increases by d_T and hence, we will deduce a 1^+-approximation result.

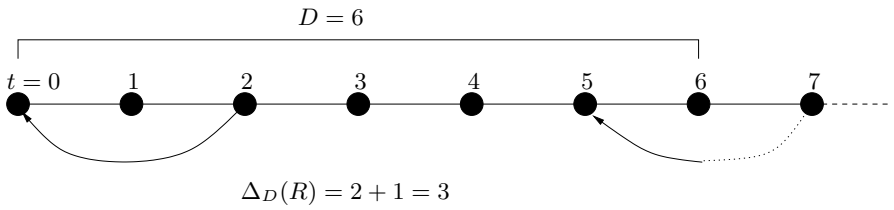

Figure 3: An example of the distance contribution in a specific neighborhood of the sink when $d_T = d_I = 2$. Round R contains the calls $2 \to 0$ and $7 \to 5$.

4.1 A new lower bound

Let us define the distance contribution $\Delta_D(R)$ of a round R in the interval $[0, D]$ as the distance that the messages transmitted during round R advance towards the sink $t = 0$ inside the interval $[0, D]$ (see Figure 3 for an example). More precisely

$$\Delta_D(R) = \sum_{s \to r \in R} \max[0, \min[d(s,0) - d(r,0), D - d(r,0)]$$

Note that if r is not in $[0, D-1]$, then $D - d(r,0) \leq 0$ and hence, the call contributes 0 in $\Delta_D(R)$. If a call is backwards $s < r$, $d(s,0) - d(r,0) < 0$ and then such a call also contributes 0 in $\Delta_D(R)$.

If $\mathcal{R} = (R_j)_{j \in J}$ is a sequence of rounds, we define its contribution as the sum of the contribution of its rounds $\Delta_D(\mathcal{R}) = \sum_{j \in J} \Delta_D(R_j)$.

These definitions are useful to prove the following lower bound. We give it for general w although we will use it only for the unitary case.

Proposition 5 *$g_{d_I,d_T}(P_n, w) \geq \mathrm{LB}_1(P_n, w)$, where*

$$\mathrm{LB}_1(P_n, w) = \frac{1}{d_T} \left(\sum_{i=1}^{D-1} iw(i) + D \sum_{i \geq D} w(i) \right)$$

In particular for the unitary case

$$g_{d_I,d_T}(P_n) \geq \mathrm{LB}_1(P_n) = \frac{D(n-D)}{d_T} + \frac{D(D-1)}{2d_T}$$

Proof: Let $A = (R_j)_{j=1}^{|A|}$ be a gathering protocol. We observe that, even when two receptions can be performed inside the interval $[0, D]$ during the same round (because the distance between a vertex receiving a message and a vertex transmitting another is at least $d_I + 1$), then

$$(\forall j = 1, \ldots, |A|), \quad \Delta_D(R_j) \leq d_T \tag{1}$$

We also observe that

- if $i \geq D$, a message from node i has to travel at least a distance D inside $[0, D]$ to reach the sink and there are $\sum_{i \geq D} w(i)$ such messages; and
- if $i < D$, a message from node i needs to travel a distance i inside $[0, D]$ to reach the sink and in the beginning there are $w(i)$ messages at vertex i, thus overall these messages need to travel a distance $iw(i)$ towards the sink.

Adding these values for $i = 1, \ldots, n-1$, it follows that

$$\Delta_D(A) \geq \sum_{i=1}^{D-1} iw(i) + D \sum_{i \geq D} w(i) \tag{2}$$

but from the definition of distance contribution and (1)

$$\Delta_D(A) = \sum_{j=1}^{|A|} \Delta_D(R_j) \leq d_T|A|. \tag{3}$$

Using (2) and (3), we have that for any gathering protocol $d_T|A| \geq \sum_{i=1}^{D-1} iw(i) + D\sum_{i \geq D} w(i)$, which corresponds to the first claim.

Now, for the second claim, we distinguish two cases.

If $n \geq D$, then $\sum_{i=1}^{D-1} i = \frac{(D-1)D}{2}$ and $\sum_{i=D}^{n-1} 1 = n - 1 - (D-1) = n - D$, hence

$$g_{d_I,d_T}(P_n) \geq \frac{D}{d_T}(n-D) + \frac{D(D-1)}{2d_T}.$$

If $n < D$, $\sum_{i=1}^{n-1} i = \sum_{i=1}^{D-1} i - \sum_{i=n}^{D-1} i \geq \frac{D(D-1)}{2} + D(n-D)$ as $i \leq D-1$.

■

4.2 A 1^+-approximation

Recall that an algorithm A calculates a 1^+-approximation for the Unitary Minimum Gathering Time problem if there exists a constant $C = C(d_I, d_T)$ independent of n such that $|A(P_n)| \leq g_{d_I,d_T}(P_n) + C$. That means that the gap between the number of rounds of algorithm A and the optimum value is an additive constant which does not increase with the size of the path.

Theorem 1 *Algorithm A_1 gives a 1^+-approximation for $g_{d_I,d_T}(P_n)$.*

Proof: If $n \leq (p+1)d_T + 1$, we have by Proposition 2 an optimal algorithm. So, let $n \geq (p+1)d_T + 1$, and let k be such that $n = D - \gamma + kd_T$, where $0 \leq \gamma < d_T$, then by Algorithm A_1

$$\begin{aligned} |A_1(P_n)| &= kD + \sum_{i=1}^{D-\gamma-1} \left\lceil \frac{i}{d_T} \right\rceil = \frac{(n-D+\gamma)}{d_T} D + \sum_{i=1}^{D-\gamma-1} \left\lceil \frac{i}{d_T} \right\rceil \\ &= \frac{D(n-D)}{d_T} + C_1(d_I, d_T) \end{aligned}$$

By Proposition 5

$$\mathrm{LB}_1(P_n) = \frac{D(n-D)}{d_T} + \frac{D(D-1)}{2d_T} = \frac{D(n-D)}{d_T} + C_2(d_I, d_T)$$

and so $\frac{|A_1(P_n)|}{\mathrm{LB}_1(P_n)} \to 1$ as $n \to \infty$. Said more precisely,

$$|A_1(P_n)| - g_{d_I,d_T}(P_n) \leq C(d_I, d_T) = C_1(d_I, d_T) - C_2(d_I, d_T).$$

■

5 Incremental Protocols

In what follows, it will be convenient to define $X = \{1, 2, \ldots, d_T\}$, the set of possible transmission lengths and consider the translation function $f : X \to X, x \mapsto f(x) = [(x + q) \mod d_T] + 1$.

In this section, we are interested in constructing protocols incrementally from P_n to P_{n+1} by adding new calls (without changing the former calls) and perhaps extra rounds. More formally, protocol $A^+ = (R_j^+)_{j \leq |A^+|}$ for the path P_{n+1} is an increment of $A = (R_j)_{j \leq |A|}$ for the path P_n, if $R_j \subset R_j^+$, for all $1 \leq j \leq |A|$. We will show how to construct a specific increment of a gathering protocol A for P_n, using a singleton round $\{d \to 0\}$ of A. We will call it A^+ or $\mathrm{Inc}(A, d)$ if we want to precise the call $d \to 0$ used.

5.1 Construction of the Incremental Protocol

Before going into the construction (called *Construction of Inc*), let us give a simple example to show how the construction works. Let $d_T = 2$, $d_I = 2$ ($p = 1, q = 0$), so $D = 5$. For $n = 5 = (p + 1)d_T + 1$ the greedy protocol of Proposition 2 consists of 6 singleton rounds $R_1 = \{4 \to 2\}$, $R_2 = \{3 \to 1\}$, $R_3 = \{2 \to 0\}$, $R_4 = \{1 \to 0\}$, $R_5 = \{2 \to 0\}$, $R_6 = \{1 \to 0\}$. Using $d = 1$ we obtain the increment $A^+ = \mathrm{Inc}(A, 1)$ for $n = 6$ by keeping rounds R_1 to R_5, replacing R_6 by $R_6^+ = \{1 \to 0, 5 \to 4\}$ and adding $R_7^+ = \{4 \to 2\}$, $R_8^+ = \{2 \to 0\}$. Here, $f(1) = 2$ and the number of rounds of $\mathrm{Inc}(A, 1)$ is $8 = 6 + (p + 1)$. Starting from A^+, as $n = 6 \geq D + 1$, we can increment it using again $d = 1$ (round R_4) obtaining a protocol $A^{++} = \mathrm{Inc}(A^+, 1)$ for $n = 7$ using the same rounds of R^+ except R_4^+ replaced by $R_4^{++} = \{1 \to 0, 6 \to 4\}$ and two new rounds $R_9^{++} = \{4 \to 2\}$, $R_{10}^{++} = \{2 \to 0\}$. Note that A^{++} is optimum as $\mathrm{LB}_1(P_7) = \frac{5 \cdot 2}{2} + \frac{5 \cdot 4}{4} = 10$. Now, in A^{++} there are 4 rounds $\{s \to 0\}$ but all of them of the form $\{2 \to 0\}$, and so an increment of A^{++} will have 3 more rounds giving a protocol for $n = 8$ with 13 rounds ($\mathrm{LB}_1(P_8) = 12.5$) but with a new singleton round $\{1 \to 0\}$ which can be used to obtain an increment for P_9 with $15 = \mathrm{LB}_1(P_9)$ rounds. Figure 4 shows the resulting protocol for P_9.

Construction of Inc(A,d): Let A be a gathering protocol for P_n containing a singleton round $\{d \to 0\}$. Let $(n \geq D + 1)$ or $((p + 1)d_T + 1 \leq n \leq D$ with $d \leq n - 2 - d_I)$. Let $m(n)$ denote the message of n, the last vertex of P_{n+1}. The idea of the construction is that, given the gathering protocol A for the instance $(P_n, 0)$ (i.e. A gathers messages from vertices $i = 1, \ldots, n - 1$ into the sink), we will show that there exist rounds in A such that $m(n)$ (recall that n is the last vertex in P_{n+1}) can be transmitted near to the sink by extending these rounds of A with some additional calls. Once message $m(n)$ is close to the sink, we will add x additional singleton rounds to complete gathering for P_{n+1}.

Let R_{j_0} be the round $R_{j_0} = \{d \to 0\}$, which exists by hypothesis.

For an integer k such that $k \geq 1$ and $d + kd_T \leq n - 1$, define j_k in such a way that the last round in A with a transmitter $s, d + (k - 1)d_T + 1 \leq s \leq d + kd_T$ is R_{j_k}. Notice that we have that $j_{k+1} < j_k$ and if s transmits during round R_{j_k} then $s \leq d + kd_T$. Let also k_d be the largest k such that $d + d_I + 1 + kd_T \leq n - 1$. (See Figure 5 for an example of the construction.). Note that k_d exists ($k_d \geq 0$); indeed by hypothesis either $n \geq D + 1$ or $d \leq n - 2 - d_I$.

For $j = 1, \ldots, |A|, j \neq j_k, k = 0, \ldots, k_d$ we set $R_j^+ = R_j$.

For $k = 0, \ldots, k_d - 1$ we set $R_{j_k}^+ = R_{j_k} \cup \{d + d_I + 1 + (k+1)d_T \to d + d_I + 1 + kd_T\}$. The call added exists as, by maximality of k_d, for $k \leq k_d - 1$, then $d + d_I + 1 + (k+1)d_T \leq n - 1$. Furthermore, round $R_{j_k}^+$ is valid ; indeed, any transmitter s in R_{j_k} is such that $s \leq d + kd_T$, hence the distance from the receiver of the new call to the largest transmitter in R_{j_k} is $d(s, d + d_I + 1 + kd_T) \geq d + d_I + 1 + kd_T - d - kd_T = d_I + 1$.

For $k = k_d$ we observe that the distance from vertex n to $d + d_I + 1 + k_d d_T$ is at most d_T, hence we can set $R_{j_{k_d}}^+ = R_{j_{k_d}} \cup \{n \to d + d_I + 1 + k_d d_T\}$.

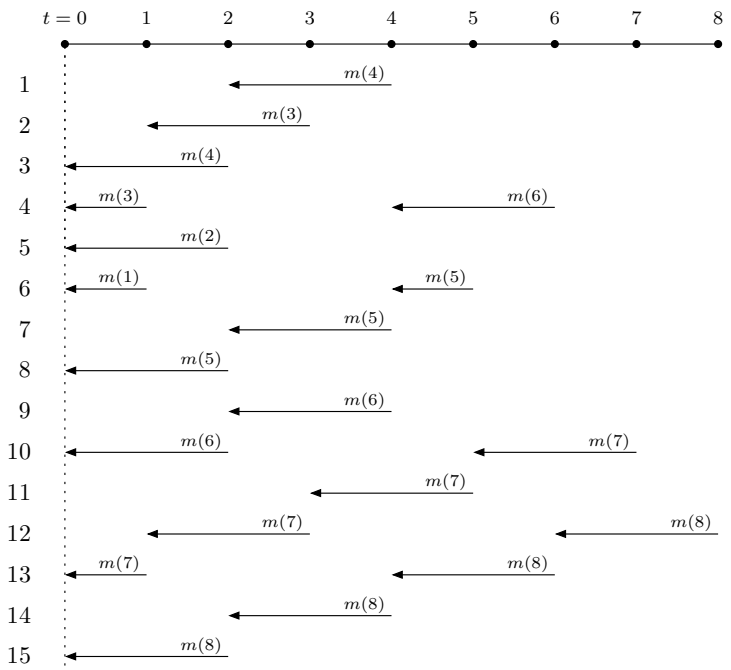

Figure 4: An example of the step for constructing incremental protocols, starting from P_5 to P_9. In this example $d_T = d_I = 2$. Above each call depicted as an arrow, $m(u)$ denotes that message transmitted comes from node u.

The protocol Inc we have devised consists of $|A|$ rounds, it gathers the same messages as A, and transmits message $m(n)$ from vertex n up to vertex $v_0 = d + d_I + 1$. Note that there always exists a call ending in v_0. Indeed, either $n \geq D + 1$ and as $d \leq d_T$, $v_0 = d + d_I + 1 \leq d_T + d_I + 1 = D < n$; or $(p + 1)d_T + 1 \leq n \leq D$, but in this case we choose d such that $v_0 = d + d_I + 1 \leq n - 2 - d_I + d_I + 1 < n$.

Now, we can add extra singleton rounds to transmit the message from vertex v_0 up to the sink $t = 0$. We do so using only calls of length d_T (except, maybe, the last one).

The following result characterizes $\mathrm{Inc}(A, d)$.

Lemma 1 *Let A be a gathering protocol for P_n containing a singleton round $\{d \to 0\}$. Let $(n \geq D + 1)$ or $((p + 1)d_T + 1 \leq n \leq D$ with $d \leq n - 2 - d_I)$. There exists an incremental protocol for P_{n+1} denoted by $Inc(A, d)$ satisfying the following:*

(i)

$$|Inc(A, d)| = |A| + \begin{cases} p + 1 & d \leq d_T - q - 1, \\ p + 2 & d > d_T - q - 1 \end{cases}$$

(ii) The family of singleton rounds of type $\{s \to 0\}$ in $Inc(A, d)$ is obtained from the family of singleton rounds of type $\{s \to 0\}$ in A, by deleting the round $\{d \to 0\}$ and perhaps another round $\{d' \to 0\}$ (which are no more singletons) and then by adding the round $\{f(d) \to 0\}$, where f is the translation function $f(d) = [(d + q) \mod d_T] + 1$.

Proof: Let x denote the number of calls needed to transmit $m(n)$ from v_0 (v_0 as in *Construction of Inc*) to the sink. We have $x = \left\lceil \frac{d+d_I+1}{d_T} \right\rceil$; hence $x = p + \left\lceil \frac{d+q+1}{d_T} \right\rceil$ and therefore $x = p + 1$ if $d + q + 1 \leq d_T$ ($\iff d \leq d_T - q - 1$) or $x = p + 2$ if $d + q + 1 > d_T$ (notice that $d + q + 1 \leq 2d_T$ as $q < d_T$ and $d \leq d_T$). We also obtain that the very

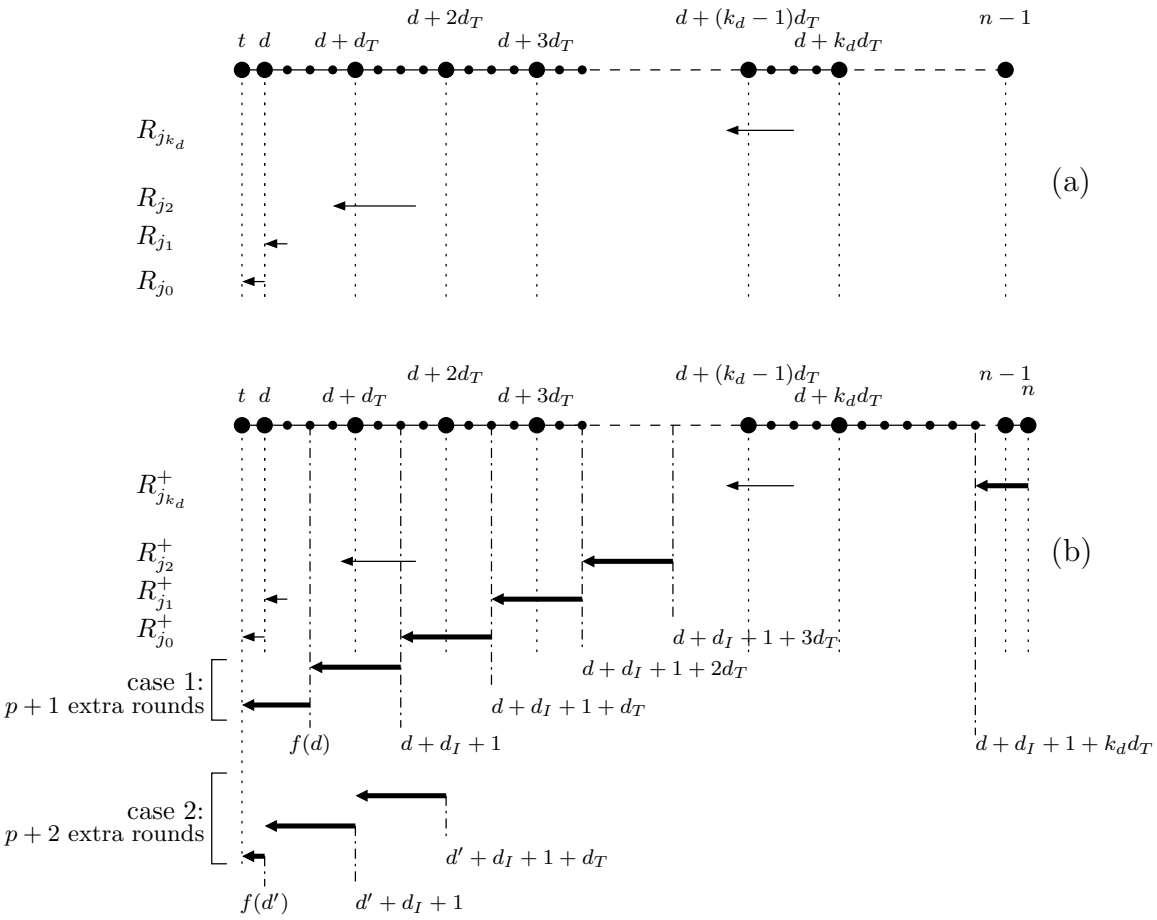

Figure 5: An example of the step for constructing incremental protocols. $x = p + 1$ if $d \leq d_T - q - 1$, or $x = p + 2$ otherwise.

last call performed in this way, which is the only call transmitting $m(n)$ ending in 0, is $d+q+1 \to 0$ if $d \leq d_T - q - 1$ or $d+q+1-d_T$ if $d > d_T - q - 1$, that is $f(d) \to 0$. Note that for $k = 1$, R_{j_1} contains a unique call with sender s such that $d+1 \leq s \leq d+d_T$. It might happen that $s \leq d_T$ and the call of R_{j_1} is of type $s \to 0$. So, we might loose a second singleton round of type $s \to 0$. All the other singleton rounds remain singleton. ■

5.2 Upper bound for incremental protocol

We can repeat the incremental *Construction of Inc* from Subsection 5.1. Let us start at some value n_0 with a protocol $A(P_{n_0})$ containing the family S_0 of singleton rounds of the form $\{s \to 0\}$. Let $\{d_0 \to 0\} \in S_0$, then applying Lemma 1 we obtain a protocol $A(P_{n_0+1}) = \mathrm{Inc}(A(P_{n_0}), d_0)$. Let S_1 be the set of singleton rounds of $A(P_{n_0+1})$. We repeat the procedure as follows: at step j we start from the protocol $A(P_{n_0+j})$, we choose a d_j such that $\{d_j \to 0\} \in S_j$ and apply Lemma 1 to obtain a protocol $A(P_{n_0+j+1}) = \mathrm{Inc}(A(P_{n_0+j}), d_j)$ and we let S_{j+1} be the set of singleton rounds of $A(P_{n_0+j+1})$.

Let us define a sequence $(d_0, \ldots, d_j, \ldots, d_{n-n_0-1})$ as *admissible* if $d_j \in S_j$, where S_j is the family of values d of singleton rounds $\{d \to 0\}$ for the protocol at step j. By construction, $S_{j+1} \subset S_j \setminus \{d_j\} \cup \{f(d_j)\}$. Then, for any admissible sequence $(d_0, \ldots, d_{n-n_0-1})$ we obtain, by using the preceding construction at each step j, a protocol $A(P_n)$ which satisfies $|A(P_n)| = |A(P_{n_0})| + (n-n_0)(p+1) + \delta$, where δ is the number of d_j such that $d_j > d_T - q - 1$. We will call such values *bad values*. Otherwise, the values d_j for which $d_j \leq d_T - q - 1$ are denoted *good values*.

The aim of this subsection is to show that for some choices of d we obtain an increment of $|A(P)|$ equal to the increment in the lower bound LB_1 (Proposition 6). This will imply that, if for some value N_0 we have an optimal protocol reaching exactly the lower bound LB_1, we obtain an optimal protocol for any $n \geq N_0$ (Theorem 2). We show how to find N_0 for $q = 0$ in Subsection 5.3 and more generally when $q+1$ and d_T are relatively prime in Subsection 5.4.

Examples (see the examples of Subsection 5.3 and Subsection 5.4) show that it is not easy to determine what is the right choice for d at a given step. One possible way consists of choosing the smallest possible available d. This works in many cases (and we can prove it is optimal if we do not have to choose at some step a call $s \to 0$ for a round R_{j_1} as defined in *Construction of Inc* in Section 5.1). Another choice consists of taking in a clearer way sequences of d, all strictly less than d_T to ensure small increments. However, we can overcome this difficulty by always choosing a value of d for which we are sure it belongs to S_j (S_j as defined above). It suffices to take the sequence $(d, f(d), f^2(d), \ldots, f^{h-1}(d))$ starting with a d in S_0 as by Lemma 1 we are sure that $f(d) \in S_1$ and more generally $f^j(d) \in S_j$. Using this sequence of values of d, we will obtain an upper bound which has the same behavior as $\mathrm{LB}_1(P_n)$. The following lemma indicates that we have an interest to choose the smallest d.

Lemma 2 *For any integer h, if $d < d'$ the sequence $(d, f(d), \ldots, f^{h-1}(d))$ contains at least as many* good values *(values such that $f^i(d) \leq d_T - (q+1)$) as $(d', f(d'), \ldots, f^{h-1}(d'))$.*

Proof: We will prove that if we have α *good values* in the sequence $(d, f(d), \ldots, f^{h-1}(d))$ with $d \geq 2$, we have at least α good values in the sequence $(d-1, f(d-1), \ldots, f^{h-1}(d-1))$. Suppose $f^i(d)$ is a good value. We consider two cases:

- Case $f^i(d) \geq 2$, which is the case for $i = 0$. Then $f^i(d-1) = f^i(d) - 1$. As $f^i(d) \leq d_T - q - 1$, then $1 \leq f^i(d-1) < d_T - q - 1$ and so $f^i(d-1)$ is also a good value.

- Case $f^i(d) = 1$, with $i \geq 1$. Then, $f^{i-1}(d) = d_T - q$ and so $f^{i-1}(d)$ is a *bad value*. But $f^{i-1}(d-1) = d_T - q - 1$ and $f^{i-1}(d-1)$ is a good value. So altogether we have at least as many *good values* for $d-1$ as for d (perhaps one more).

■

Lemma 3 *The sequence* $(d_T, f(d_T), \ldots, f^{h-1}(d_T))$ *contains at most* δ_h bad values*, where* $\delta_h = \left\lceil \frac{h(q+1)}{d_T} \right\rceil$.

Proof: We proceed by induction on h. Let $\delta_h = \left\lceil \frac{h(q+1)}{d_T} \right\rceil$. The lemma is true for $h = 1$ as $\delta_h = 1$ and the sequence has one value d_T (bad in that case). Note that $f^i(d_T) = [i(q+1) - 1] \mod d_T + 1$. Suppose the fact is true for $h-1$. We distinguish two cases:

- If $(h-1)(q+1) \leq (\delta_h - 1)d_T$ and so $\delta_{h-1} = \delta_h - 1$ and by induction hypothesis we have at most $\delta_h - 1$ *bad values* and so at most δ_h *bad values* for h.
- If $(h-1)(q+1) > (\delta_h - 1)d_T$. Then, $f^{h-1}(d_T) = [(h-1)(q+1) - 1] \mod d_T + 1 = (h-1)(q+1) - (\delta_h - 1)d_T$. By definition of δ_h, $h(q+1) \leq \delta_h d_T$; therefore $f^{h-1}(d_T) \leq \delta_h d_T - (q+1) - (\delta_h - 1)d_T = d_T - (q+1)$. Therefore $f^{h-1}(d_T)$ is a good value and so the number of bad values is $\delta_{h-1} \leq \delta_h$.

■

Proposition 6 *For any pair of integers* n, h

$$g_{d_I,d_T}(P_{n+h}) \leq g_{d_I,d_T}(P_n) + \left\lceil \frac{hD}{d_T} \right\rceil.$$

Proof: We have to show that there exists a sequence of h increments such that the number of *bad values* δ (those values $d > d_T - q - 1$) satisfies $(p+1)h + \delta \leq \left\lceil \frac{hD}{d_T} \right\rceil$. As $D = d_I + d_T + 1 = (p+1)d_T + q + 1$ it suffices to show that $\delta \leq \left\lceil \frac{h(q+1)}{d_T} \right\rceil$. As in any protocol the last round is necessarily a singleton round, there always exists a sequence $(d, f(d), \ldots, f^{h-1}(d))$. For this sequence the number of bad values is by Lemma 2 at most that of the sequence $(d_T, f(d_T), \ldots, f^{h-1}(d_T))$, which itself is according to Lemma 3 at most δ_h. ■

Note that, taking $h = d_T$ in Proposition 6, we find again Proposition 3.

Theorem 2 *If there exists an integer* N_0 *such that* $\mathrm{LB}_1(P_{N_0}) = |A^*(P_{N_0})|$ *where* $A^*(P_{N_0})$ *is a gathering protocol for the path* P_{N_0}*, then there exists an optimal incremental protocol for the path* (P_{N_0+h}) *for any* $h \geq 0$ *with value* $|A^*(P_{N_0+h})| = |A^*(P_{N_0})| + \left\lceil \frac{hD}{d_T} \right\rceil$.

Proof: Because $|A^*(P_{N_0})| = \mathrm{LB}_1(P_{N_0})$, we have that $A^*(P_{N_0})$ is optimum. Now, notice that

$$\mathrm{LB}_1(P_{N_0+h}) = \mathrm{LB}_1(P_{N_0}) + \frac{hD}{d_T},$$

and that from Proposition 6 we have $g_{d_I,d_T}(P_{N_0+h}) \leq g_{d_I,d_T}(P_{N_0}) + \left\lceil \frac{hD}{d_T} \right\rceil = |A^*(P_{N_0})| + \left\lceil \frac{hD}{d_T} \right\rceil$. But $|A^*(P_{N_0})| = \mathrm{LB}_1(P_{N_0})$, and we can write

$$|A^*(P_{N_0})| + \frac{hD}{d_T} \leq g_{d_I,d_T}(P_{N_0+h}, 0) \leq |A^*(P_{N_0})| + \left\lceil \frac{hD}{d_T} \right\rceil,$$

from which the result follows. ■

Remark 1 *Note that the proof works because* $|A^*(P_{N_0})| = \mathrm{LB}_1(P_{N_0})$*. It implies that* $\mathrm{LB}_1(P_{N_0})$ *is an integer (see the example of Section 5.4).*

5.3 Optimal protocol for the case $q = 0$

We will show how to choose d, when doing increments, in order to obtain optimal solutions for all values of n, when $q = 0$. In that case, $d_I = pd_T$ and we suppose $q \neq d_T - 1$ that is $d_T \neq 1$ as we deal with this case in Proposition 4 and the result is already known ([BCY09]).

Theorem 3 *Let $q = 0$ ($d_I = pd_T$) and $n_0 = D = (p+1)d_T + 1$. There exists an optimal protocol $A^*(P_n)$ for any n such that*

$$|A^*(P_n)| = \begin{cases} \mathrm{LB}_0(P_n) & \text{if } n \leq N_0, \\ \lceil \mathrm{LB}_1(P_n) \rceil & \text{if } n \geq N_0 \end{cases}$$

where

- $N_0 = D + \frac{D-1}{2}(d_T - 1) = \frac{(p+1)d_T(d_T+1)}{2} + 1$,
- $\mathrm{LB}_0(P_{n_0}) = \frac{(p+1)(p+2)}{2} d_T$,
- $\mathrm{LB}_0(P_n) = \mathrm{LB}_0(P_{n_0}) + (p+1)(n - n_0)$, *for* $n \geq n_0$,
- $\mathrm{LB}_0(P_{N_0}) = \mathrm{LB}_1(P_{N_0}) = \frac{D(D-1)}{2}$,
- $\mathrm{LB}_1(P_n) = \frac{D(n-D)}{d_T} + \frac{D(D-1)}{2d_T} = \mathrm{LB}_1(P_{N_0}) + \frac{D}{d_T}(n - N_0)$, *for* $n \geq N_0$.

Proof: Note that as $q = 0$, for any $d < d_T$, $f(d) = d + 1$ and the number of rounds in $\mathrm{Inc}(A, d)$ is $|A| + p + 1$.

The idea is to proceed by phases, each consisting of sequences of the form $d, d + 1, \ldots, d_T - 1$, taking as d the smallest value available and $d \neq d_T$. The second idea is to use as call R_{j_1} a call $\{d + d_T \to d\}$ (we will see that there are just enough such calls).

Let us now describe precisely the process. Let $n_0 = D = (p+1)d_T + 1$.

For $n \leq n_0$ the answer is given by Proposition 2 and we have an optimal protocol with $|A(P_n)| = \sum_{i \leq n-1} \left\lceil \frac{i}{d_T} \right\rceil = \mathrm{LB}_0(P_n)$.

For $n = n_0$, $A(P_{n_0}) = \mathrm{LB}_0(P_{n_0}) = \frac{(p+1)(p+2)}{2} d_T$. Furthermore, $A(P_{n_0})$ contains $p + 1$ singleton rounds $\{d \to 0\}$ and also p singleton rounds $\{d_T + d \to d\}$, for each $1 \leq d \leq d_T - 1$. In the first phase, we do increments using the sequence $1, 2, \ldots, d_T - 1$. For each d, we put the call $n_0 + d - 1 \to pd_T + d + 1$ in the same round as $\{d \to 0\}$; that is possible as $(n_0 + d - 1) - (pd_T + d + 1) = d_T - 1 \leq d_T$ and $(pd_T + d + 1) - d = d_I + 1 > d_I$. Then we add the $p+1$ rounds $\{(p-j)d_T + d + 1 \to (p-j-1)d_T + d + 1\}$ for $j = 0, \ldots, p-1$ and the round $\{f(d) = d + 1 \to 0\}$. During this phase, for each d, $1 \leq d \leq d_T - 1$, we have deleted one round $\{d \to 0\}$, but created one round $\{d + 1 \to 0\}$ and also a round $\{d_T + d + 1 \to d + 1\}$. The protocol for $n_0' = n_0 + d_T - 1$ contains now p rounds $\{1 \to 0\}$ and $\{d_T + 1 \to 1\}$, and for $2 \leq d \leq d_T - 1$, $p + 1$ rounds $\{d \to 0\}$ and $\{d_T + d \to d\}$ (and $p + 2$ rounds $\{d_T \to 0\}$ and $p + 1$ rounds $\{2d_T \to d_T\}$). Furthermore for $n_0 \leq n \leq n_0'$ we have $|A(P_n)| = |A(P_{n-1})| + p + 1$ and so $A(P_n) = \mathrm{LB}_0(P_n)$.

We now execute p times the phase consisting in the sequence $1, 2, \ldots, d_T - 1$, obtaining an optimal protocol as $|A(P_n)| = |A(P_{n-1})| + p + 1$, for $n_0' \leq n \leq n_1 = n_0' + p(d_T - 1)$. On the one hand, we have altogether used p calls $\{d \to 0\}$ for $1 \leq d \leq d_T - 1$ and in R_{j_1} p calls $\{d_T + d \to d\}$ for $1 \leq d \leq d_T - 1$. On the other hand, we have created p calls $\{d \to 0\}$ and $\{d_T + d \to d\}$ for $2 \leq d \leq d_T - 1$. Summarizing, in $A(P_{n_1})$ we have no more calls $\{1 \to 0\}$ and $\{d_T + 1 \to 1\}$, but still $(p + 1)$ calls $\{d \to 0\}$ and $\{d_T + d \to d\}$ for $2 \leq d \leq d_T - 1$.

Then, if $d_T \geq 3$, we execute $p + 1$ times the sequence $2, 3, \ldots, d_T - 1$; we use all the calls $\{2 \to 0\}$ and $\{d_T + 2 \to 2\}$. We also use, but create the same number $p + 1$ of calls $\{d \to 0\}$ and $\{d_T + d \to d\}$ for $3 \leq d \leq d_T - 1$. So for $n_2 = n_1 + (p + 1)(d_T - 2)$ we have available $(p + 1)$ rounds $\{d \to 0\}$ and $\{d_T + d \to d\}$ for $3 \leq d \leq d_T - 1$ and $A(P_{n_2}) = \mathrm{LB}_0(P_{n_2})$.

We repeat the process by taking successively $(p+1)$ sequences $d, d+1, \ldots, d_T-1$ for $d = 3, \ldots, d_T-1$ obtaining always optimal protocols satisfying $A(P_n) = \mathrm{LB}_0(P_n)$ for $n_{d-1} \leq n \leq n_d$, where $n_d = n_{d-1} + (p+1)(d_T - d)$ for $d \geq 1$. At the end, we arrive at a value $n_{d_T-1} = N_0$, where there are no more calls $\{d \to 0\}$ with $d < d_T$. The value of N_0 is

$$\begin{aligned} N_0 &= n_0 + (p+1)\left((d_T-1) + (d_T-2) + \ldots + 1\right) \\ &= D + \frac{D-1}{2}(d_T - 1) \\ &= \frac{(p+1)d_T(d_T+1)}{2} + 1 \end{aligned}$$

Furthermore, for N_0 we have

$$\begin{aligned} |A(P_{N_0})| = \mathrm{LB}_0(P_{N_0}) &= |A(P_{n_0})| + (p+1)^2\frac{d_T(d_T-1)}{2} \\ &= \frac{(p+1)d_T}{2}\Big((p+2) + (p+1)(d_T-1)\Big) \\ &= \frac{(D-1)D}{2} \end{aligned}$$

But, we also have

$$\begin{aligned} \mathrm{LB}_1(P_{N_0}) &= \frac{D(N_0-D)}{d_T} + \frac{D(D-1)}{2d_T} \\ &= \frac{D}{2d_T}\Big((D-1)(d_T-1) + (D-1)\Big) \\ &= \frac{D(D-1)}{2} \end{aligned}$$

So we have $|A(P_{N_0})| = \mathrm{LB}_1(P_{N_0})$ and by Theorem 2 there exists an (incremental) optimal protocol for any $n \geq N_0$ with $|A(P_n)| = \lceil \mathrm{LB}_1(P_n) \rceil$. ■

Remark 2 *Note that we could have chosen to do increments using always the smallest values for d. So we increment first with the $p+1$ calls $1 \to 0$ then with all the calls $2 \to 0$ and so on. But doing so we do not create enough $d_T + d \to d$ and for $d \geq 4$ we are obliged to use calls of the type $a \to 0$ for R_{j_1} and for $d \geq 6$ we are obliged to use calls of the type $a \to 0$ with $a < d_T$ and we do not have enough d's to reach the optimal N_0.*

Our method works because by chance for $n_0 \leq n \leq n_1$ we did not use calls $d_T + d \to d$ but we were able to create one new call $d_T + d \to d$ for $d \geq 2$, obtaining the right number $p+1$ of such calls to be used as R_{j_1}.

So it is not true that choosing the smallest d available gives an optimal solution (in particular it does not work for $q = 0$ and $d_T \geq 8$).

5.4 Optimal case when $q+1$ and d_T are relatively prime

According to Theorem 2, if we are lucky enough to find a value N_0 for which $|A(P_{N_0})| = \mathrm{LB}_1(P_{N_0})$ (which in particular means $\mathrm{LB}_1(P_{N_0})$ is an integer), we can conclude that for $n \geq N_0$ the increments A^{+h} (i.e. incrementing A h times) are all optimum. That is what happens for $q = 0$, where we found such an $N_0 = D + \frac{D-1}{2}(d_T - 1)$. In that case, we started from the greedy protocol for $n = n_0 = (p+1)d_T + 1$. But in general, it is not the right protocol to start with; on the contrary, we might have to start from some non-optimal protocol, but with more singleton rounds in S_0. To see what happens, let us consider a case where $q \neq 0$ and $q \neq d_T - 1$ for example $d_T = 3, d_I = 4$ ($p = 1, q = 1$) and $D = 8$. For $n_0 = 7$, we have the greedy optimal protocol with $|A| = 9$ rounds, containing two

rounds $\{1 \to 0\}$, two $\{2 \to 0\}$ and two $\{3 \to 0\}$. Using twice $d = 1$, we obtain a protocol for $n = 9$ with $9 + 2 + 2 = 13$ rounds containing as singleton rounds twice $\{2 \to 0\}$ and 4 times $\{3 \to 0\}$ as $f(1) = 3$. It is optimal as $\mathrm{LB}_0(9) = 13$. Then we increment with $d = 2$, obtaining as $f(2) = 1$ one round $\{1 \to 0\}$. We can in turn increment with $d = 1$ obtaining a protocol for $n = 11$ with $13 + 3 + 2 = 18$ rounds. Here $\mathrm{LB}_0(11) = 17$, but $\mathrm{LB}_1(11) = 17 + 2/3$ and so it is optimal. We do again an increment with $d = 2$, followed by an increment with $d = 1$ obtaining for $n = 13$ a protocol with $18+5 = 23$ rounds. Here $\mathrm{LB}_0(13) = 21$, but $\mathrm{LB}_1(13) = 22 + 2/3$ and so it is optimal. But then, we are obliged to use $d = 3$ which gives $f(d) = 2$ and then $d = 2$. So for $n = 15$ we have a protocol with $23 + 3 + 3 = 29$ rounds but $\mathrm{LB}_1(15) = 28$ and $\mathrm{LB}_0(15) = 25$. Table 2 gives the values of the number of rounds using the best increment and the corresponding values of LB_0 and LB_1. Note that sometimes $|A(P_n)| = \lceil \mathrm{LB}_1(P_n) \rceil$ but never $|A(P_n)| = \mathrm{LB}_1(P_n)$.

n	7	8	9	10	11	12	13	14	15	16	17	18
$\lvert A(P_n)\rvert$	9	11	13	16	18	21	23	26	29	31	34	37
$\lvert A^*(P_n)\rvert$		12	14	16	18	21	23	26	28	31	34	36
$\mathrm{LB}_1(P_n)$	7	$9\ \frac{2}{3}$	12	$14\ \frac{2}{3}$	$17\ \frac{1}{3}$	20	$22\ \frac{2}{3}$	$25\ \frac{1}{3}$	28	$30\ \frac{2}{3}$	$33\ \frac{1}{3}$	36
$\mathrm{LB}_0(P_n)$	9	11	13	15	17	19	21	23	25	27	29	31
Optimum	*	*	*		*		*	*	*	*	*	*

Table 2: Number of rounds using the best increment

In fact, we will see in the next proposition, that if we start with the non-optimal greedy protocol for $n_1 = D$ containing only singleton rounds, we obtain an optimal protocol for $n \geq N_0$. In the example above, $n_1 = D = 8$ and the greedy protocol has 12 rounds (not optimal), with singleton rounds three $\{1 \to 0\}$, two $\{2 \to 0\}$ and two $\{3 \to 0\}$. Incrementing three times with $d = 1$, we obtain, for $n = 11$, a protocol with $|A^*| = 12 + 3 \cdot 2 = 18$ rounds. Using twice the rounds $\{2 \to 0\}$ and then the rounds $\{1 \to 0\}$ created, we obtain, for $n = 15$, a protocol with $|A^*| = 18 + 2 \cdot 5 = 28$ rounds, which meets exactly $\mathrm{LB}_1(P_n)$. The values of $A^*(P_N)$ are by Theorem 2 optimum for $n \geq 15$ (in fact, they are also optimal for $n = 13, 14$).

The next theorem is a generalization of Theorem 3.

Theorem 4 *Let $d_I = pd_T+q$ and $D = (p+1)d_T+q+1$. Then, if $q+1$ and d_T are relatively prime, for $n \geq N_0$, where $N_0 = D + \frac{D-1}{2}(d_T - 1)$, there exists an optimal protocol $A(P_n)$ with $|A(P_n)| = \lceil \mathrm{LB}_1(P_n) \rceil = \mathrm{LB}_1(P_{N_0}) + \left\lceil \frac{D}{d_T}(n - N_0) \right\rceil$, where $\mathrm{LB}_1(P_{N_0}) = \frac{D(D-1)}{2}$.*

Proof: The proof is illustrated in the example after with $d_T = 5, d_I = 12, (p = 2, q = 2, D = 18)$. Let us start with the trivial non-optimal solution B_1 for $n_1 = D = (p + 1)d_T + q + 1$ consisting of the singleton rounds for $i < n_1$, $i = \alpha d_T + \beta$, $1 \leq \beta \leq d_T$, $\{i \to i - d_T\}, \{i - d_T \to i - 2d_T\}, \ldots, \{i - (\alpha - 1)d_T \to i - \alpha d_T\}, \{i - \alpha d_T = \beta \to 0\}$.

B_1 contains $p + 2$ rounds $\{d \to 0\}$ for $1 \leq d \leq q$ and $p + 1$ rounds $\{d \to 0\}$ for $q + 1 \leq d \leq d_T$. It also contains $p + 1$ rounds $d + d_T \to d$ for $1 \leq d \leq q$ and p rounds $d + d_T \to d$ for $q + 1 \leq d \leq d_T$.

Like for the case $q = 0$, we first do increments by using the sequence $1, 2, \ldots, d_T - 1$. For $n'_0 = D + d_T - 1$ we have used one call $\{d \to 0\}$ for each d, $1 \leq d \leq d_T - 1$ and created one call $\{f(d) \to 0\}$, that is one call $\{d \to 0\}$ for each $d \neq q + 1$. We have also created calls $\{f(d) + d_T \to f(d)\}$, that is one call $\{d + d_T \to d\}$ for each $d \neq q + 1$. In summary, we have: - for $1 \leq d \leq q$: $p+2$ rounds $\{d \to 0\}$ and $p+2$ rounds $\{d+d_T \to d\}$; - for $d = q + 1$: p rounds $\{q + 1 \to 0\}$ and p rounds $\{q + 1 + d_T \to q + 1\}$; - and for $q + 2 \leq d \leq d_T - 1$: $p + 1$ rounds $\{d \to 0\}$ and $p + 1$ rounds $\{d + d_T \to d\}$.

Now, we apply successively the sequence $d, f(d), f^2(d), \ldots, f^{h-1}(d)$, where d is the smallest available and h such that $f^h(d) = d_T$ (such an h exists as $q+1$ and d_T are relatively prime). The values $d, f(d), f^2(d), \ldots, f^{h-1}(d)$ are therefore all different from d_T. For a given d, applying such a sequence decreases the set of rounds $\{d \to 0\}$ and $\{d + d_T \to d\}$ by exactly one. It leaves the number of other singleton rounds unchanged (except for $\{d_T \to 0\}$ where it increases by one). Note that the result follows from the fact that (thanks to the first phase), we always have a call $d + d_T \to d$ available for round R_{j_1}. Like for the case $q = 0$, if we did not do the first phase, we would need to use, for $d_T \geq 8$, a call $d \to 0$ in round R_{j_1} and the proof would not work.

We repeat the process until we have no more $d < d_T$ available, which happens for some value N_0. We claim that for this value N_0 the protocol obtained is optimal. Indeed, consider the intersection of a round with the interval $[0, D]$. Either it contains a unique call of length d_T (of the form $\{d_T \to 0\}$ or $\{a \to a - d_T\}$ for $d_T < a \leq D$). Otherwise, the round contains two calls intersecting $[0, D]$: a call $d \to 0$ for some $d < d_T$ which has been completed in the incremental process with a call $d + d_I + d_T + 1 \to d + d_I + 1$ whose contribution in $[0, D]$ is exactly $d_T - d$ and so the total contribution of the round in the interval $[0, D]$ is d_T and so the lower bound is exactly attained.

Computation of N_0:

Note that altogether (including the first phase), we have used $p+2$ (resp. $p+1$) times the values of the sequences $d, f(d), \ldots, f^{h_d-1}(d)$ for $1 \leq d \leq q$ (resp. $q+1 \leq d \leq d_T - 1$). So, to compute N_0, we have to determine the length h_d of a sequence starting at d. By definition h_d is the smallest integer h such that $f^{h_d}(d) = d_T$ that is $d + h(q+1)$ is a multiple of d_T. As $q+1$ is relatively prime to d_T, we have that: $h_d \neq h_{d'}$ for $d \neq d'$; $h_{q+1} = d_T - 1$ and $h_d + h_{q+1-d} = d_T - 1$ for $1 \leq d \leq q$. This last equality follows from the fact that, $d + (q+1)h_d$ and $q + 1 - d + (q+1)h_{q+1-d}$ being multiple of d_T, their sum $(q+1)(h_d + h_{q+1-d} + 1)$ is also a multiple of d_T; and as $q+1$ is relatively prime to d_T, we obtain $h_d + h_{q+1-d} + 1 = d_T$. Now we can compute N_0:

$N_0 = D + (p+2)\sum_{d=1}^{q} h_d + (p+1)\sum_{i=q+1}^{d_T-1} h_d = D + (p+1)\sum_{d \neq d_T} h_d + \sum_{d=1}^{q} h_d$. As $h_d \neq h_{d'}$ for $d \neq d'$, $\sum_{d \neq d_T} h_d = \sum_{i=1}^{d_T-1} i = \frac{d_T(d_T-1)}{2}$. Now, using $h_d + h_{q+1-d} = d_T - 1$ for $1 \leq d \leq q$, we obtain: if q is even $\sum_{d=1}^{q} h_d = \sum_{d=1}^{\frac{q}{2}} (h_d + h_{q+1-d}) = \frac{q}{2}(d_T - 1)$ and if q is odd $\sum_{d=1}^{q} h_d = \sum_{d=1}^{\frac{q-1}{2}} (h_d + h_{q+1-d}) + h_{\frac{q+1}{2}} = \frac{q}{2}(d_T - 1)$.

Therefore, $N_0 = D + (p+1)\frac{d_T(d_T-1)}{2} + \frac{q}{2}(d_T - 1) = D + \frac{D-1}{2}(d_T - 1)$.

Finally, note that we already know that $LB_1(P_{N_0}) = \frac{D(D-1)}{2}$. ■

Example: $d_T = 5, d_I = 12, (p = 2, q = 2, D = 18)$. For P_{18}, B_1 contains 4 rounds $\{1 \to 0\}$, $\{2 \to 0\}$; 3 rounds $\{3 \to 0\}, \{4 \to 0\}, \{5 \to 0\}$ and then $|A(P_{18})| = 38$. We first use the sequence $1, 2, 3, 4$ and we obtain $|A(P_{22})| = 38 + 3 + 3 + 4 + 4 = 52$. $A(P_{22})$ contains 4 rounds $\{1 \to 0\}, \{2 \to 0\}$; 2 rounds $\{3 \to 0\}$; 3 rounds $\{4 \to 0\}$ and 4 rounds $\{5 \to 0\}$. Notice that we have only 2 rounds $\{3 \to 0\}$ because we have not used any round $\{5 \to 0\}$, so we have not created a new round $\{3 \to 0\}$. Then we will use 4 times the sequence $1, 4, 2$ (with increments $3 + 4 + 3 = 10$); 4 times the sequence 2 (increment 3); 2 times the sequence $3, 1, 4, 2$ (increment $4 + 3 + 4 + 3 = 14$) and 3 times the sequence $4, 2$ (increment $4 + 3 = 7$) obtaining the value $N_0 = 22 + 4 \cdot 3 + 4 \cdot 1 + 2 \cdot 4 + 3 \cdot 2 = 52$ and $|A(P_{N_0})| = 52 + 4 \cdot 10 + 4 \cdot 3 + 2 \cdot 14 + 3 \cdot 7 = 153 = \frac{18 \cdot 17}{2}$. Table 3 describes the number of rounds $\{a \to 0\}$ with $1 \leq a \leq 5$ at each incremental step starting from P_{18}. In this table, the value in row d and column P_j denotes the round $\{d \to 0\}$ used to increment P_{j-1} to P_j.

Corollary 1 *If d_T is prime, we have an optimal protocol $A(P_n)$ for $n \geq N_0$, where $N_0 = D + \frac{D-1}{2}(d_T - 1)$.*

	P_{18}	P_{19}	P_{20}	P_{21}	P_{22}
d		1	2	3	4
$\lvert\{1 \to 0\}\rvert$	4	3	3	4	4
$\lvert\{2 \to 0\}\rvert$	4	4	3	3	4
$\lvert\{3 \to 0\}\rvert$	3	3	3	2	2
$\lvert\{4 \to 0\}\rvert$	3	4	4	4	3
$\lvert\{5 \to 0\}\rvert$	3	3	4	4	4

	P_{23}	P_{24}	P_{25}	P_{26}	P_{27}	P_{28}	P_{29}	P_{30}	P_{31}	P_{32}	P_{33}	P_{34}
d	1	4	2	1	4	2	1	4	2	1	4	2
$\lvert\{1 \to 0\}\rvert$	3	3	3	2	2	2	1	1	1	0	0	0
$\lvert\{2 \to 0\}\rvert$	4	5	4	4	5	4	4	5	4	4	5	4
$\lvert\{3 \to 0\}\rvert$	2	2	2	2	2	2	2	2	2	2	2	2
$\lvert\{4 \to 0\}\rvert$	4	3	3	4	3	3	4	3	3	4	3	3
$\lvert\{5 \to 0\}\rvert$	4	4	5	5	5	6	6	6	7	7	7	8

	P_{35}	P_{36}	P_{37}	P_{38}
d	2	2	2	2
$\lvert\{1 \to 0\}\rvert$	0	0	0	0
$\lvert\{2 \to 0\}\rvert$	3	2	1	0
$\lvert\{3 \to 0\}\rvert$	2	2	2	2
$\lvert\{4 \to 0\}\rvert$	3	3	3	3
$\lvert\{5 \to 0\}\rvert$	9	10	11	12

	P_{39}	P_{40}	P_{41}	P_{42}	P_{43}	P_{44}	P_{45}	P_{46}
d	3	1	4	2	3	1	4	2
$\lvert\{1 \to 0\}\rvert$	1	0	0	0	1	0	0	0
$\lvert\{2 \to 0\}\rvert$	0	0	1	0	0	0	1	0
$\lvert\{3 \to 0\}\rvert$	1	1	1	1	0	0	0	0
$\lvert\{4 \to 0\}\rvert$	3	4	3	3	3	4	3	3
$\lvert\{5 \to 0\}\rvert$	12	12	12	13	13	13	13	14

	P_{47}	P_{48}	P_{49}	P_{50}	P_{51}	P_{52}
d	4	2	4	2	4	2
$\lvert\{1 \to 0\}\rvert$	0	0	0	0	0	0
$\lvert\{2 \to 0\}\rvert$	1	0	1	0	1	0
$\lvert\{3 \to 0\}\rvert$	0	0	0	0	0	0
$\lvert\{4 \to 0\}\rvert$	2	2	1	1	0	0
$\lvert\{5 \to 0\}\rvert$	14	15	15	16	16	17

Table 3: Number of rounds of the type $\{a \to 0\}$ at each increment

For $d_T = 4$, $q = 0, 2$ (or for $d_T = 6$, $q = 0, 4$), by applying Theorem 4, we obtain an optimal protocol $A(P_n)$ for $n \geq N_0$; but for $d_T = 4$, when $q = 1$, $q+1 = 2$ is not relatively prime to 4 and then we cannot apply the proof of the theorem. Using increments we can find a protocol with one more round than $\mathrm{LB}_1(P_n)$ (we conjecture that this protocol is optimal which will need an improvement of the lower bound).

We have seen how to obtain in some cases exact or asymptotic results by using incremental protocols. Clearly, not all increments are optimal and the choice of the starting protocol is not evident, as we have seen.

We conjecture that:

Conjecture 1 *For any $n \geq D+1$, there exists an optimal protocol B for P_n obtained by repeated applications of Construction of Inc increments of some protocol A for P_{D+1}.*

If Conjecture 1 is true, we could prove the main conjecture:

Conjecture 2 Unitary Minimum Time Gathering *in the path P_n with $t = 0$ is polynomial in the length n of the path.*

Remark 3 *We have only been able to prove Conjecture 1 for only one increment and obtain for $n \geq (p+1)d_T + 1$ from a protocol A for $n+1$ a protocol A^- for P_n such that there exists an increment B of A^- with a number of rounds at most that of A. In particular if A is optimum, then B is optimum.*

6 Gathering into an arbitrary vertex of the path

So far, we have discussed only the case where $t = 0$. In this section, we remove this constraint and take $0 < t < n - 1$.

In [BCY09], results are given for $d_T = 1$ and an optimal solution is only given for $d_I \in \{1, 2, 3, 4\}$. So, for $d_T > 1$ there is no hope to find general optimal protocols. However, in this section we will give a 1^+-approximation in a similar manner than in Section 4.

Proposition 7

$$g_{d_I,d_T}(P_n, t) \geq \frac{d_I + d_T + 1}{d_T}(\max[t, n-1-t] - D + 1) + \frac{(d_I + d_T + 1)(d_I + d_T)}{2d_T}.$$

Proof: Recall that we assume that $\max[t, n-1-t] = n-1-t$; consider the interval of vertices $I = \{t, t+1, \ldots, t+D-1\}$, with $D = d_I + d_T + 1$. Like in Proposition 5, we have that for $i \in I$, a message originated in i has to travel (within I) a distance $i - t$ to the sink, and that for $j \geq t + D$, a message originated in j has to travel a distance D (within I) in order to reach the sink t. We also have that even if two messages can move inside I, overall these two messages cannot progress more than d_T vertices toward the sink, from where it follows that

$$g_{d_I,d_T}(P_n, t) \geq \frac{D}{d_T}(n - t - D) + \frac{D(D-1)}{2d_T}.$$

■

Notice that this bound assumes perfect synchronization between the calls in the two sides, i.e., that gathering messages from the shortest side does not delay gathering messages from the longest side. In general, transmitting messages in one side produces interference in the other side, thus some extra rounds may be required (see [BCY09]).

Now we want to extend the 1^+-approximation result of Section 4.2. Since Proposition 7 establishes that the number of rounds required to gather the path that has two sides is, roughly, lower bounded by the number of rounds needed to gather its longest side, the algorithm we introduce works as follows: Given the protocol that gathers the longest side, its rounds are modified by adding calls in the shorter side so messages coming from that side are gathered into the sink at the same time. Moreover, this is done in such a way that when finished, we can guarantee that only the vertices at distance at most $D-1$ from the sink have messages that are still unknown for the sink, and that they have at most one message. Because these are $O(1)$ vertices, each at a distance $O(1)$ from the sink, we deduce that we can gather these messages in constant time.

Theorem 5 *For the* Unitary Minimum Time Gathering *problem where the base graph is a path P_n, the interference distance is d_I, and transmission distance is d_T, there exists a 1^+-approximation.*

Proof: First, let $\ell_1 = t, \ell_2 = n-1-t$ and recall that we suppose $\ell_1 \leq \ell_2$. If $D \geq \ell_2$, the size of the network is bounded so to gather a message at a distance i from t requires at most $c(i) = \lceil i/d_T \rceil \leq \lceil D/d_T \rceil \leq 2 + d_I/d_T$ rounds. Therefore to gather the path in this case can be done in at most $2\sum_{i=1}^{D} c(i) \leq 2(2+d_I/d_T)D$ rounds, which does not depend on n.

To analyze the case $\ell_2 \geq D$, let us rename the vertices of the path in such a way that vertex i becomes $i-t$, so the sink is vertex 0, the left side consists of vertices $-\ell_1, -(\ell_1-1), \ldots, -1$ and the right (and longer) side corresponds to vertices $1, 2, \ldots, \ell_2$.

We define $D_i = \{i + kD \to \max[0, i+kD-d_T] : k \geq 0, i+kD \leq \ell_2\}$. We have that, after applying $D_i, i = 1, \ldots, D$, each message at the right side of the sink either has reached the sink or it has advanced a distance d_T towards the sink. Similarly, we define $C_i = \{i-D-1-kD \to \min[0, i-D-1-kD+d_T] : k \geq 0, i-D-1-kD \geq -\ell_1\}$ and have that after applying $C_i, i = 1, \ldots, D$, each message at the left side of the sink has reached it or it moved a distance d_T towards the sink.

The protocol we use is computed using the following algorithm (recall that $\ell_1 \leq \ell_2$):

Input: $n, \ell_1, \ell_2, d_I, d_T$
1 **while** $\ell_2 - 1 \geq D$ **do**
2 | **for** $i = 1$ *to* D **do**
| | Apply $C_i \cup D_i$.
| **end**
| $\ell_1 \leftarrow \ell_1 - d_T$, $\ell_2 \leftarrow \ell_2 - d_T$.
end
3 Gather each message independently, using a shortest path to the sink.

Algorithm 2: Solves gathering in P_n for arbitrary t.

This algorithm is almost identical to Algorithm A_1, the only difference is that we have replaced "Apply $\{i+kD \to \max[0, i+kD-d_T] : k \geq 0, i+kD \leq n-1\}$" with "Apply $C_i \cup D_i$" (in fact, D_i is precisely the set $\{i + kD \to \max[0, i+kD-d_T] : k \geq 0, i+kD \leq n-1\}$ where we replaced $n-1$ with ℓ_2).

We only need to check that each call in C_i is compatible with each call in D_i. Indeed, for any i the closest calls are $i \to \max[0, i-d_T] \in D_i$ and $i-D-1 \to \min[0, i-D-1+d_T] \in C_i$, and we have $d(i, \min[0, i-D-1+d_T]) \geq D+1-d_T \geq d_I+1$ (as if $i-D-1+d_T > 0$ then $i > D+1-d_T$) and similarly $d(\max[0, i-d_T], i-D-1) \geq D+1-d_T \geq d_I+1$.

Because Step 3 requires $O(1)$ rounds, we focus on Step 1. This step requires at most $\max\left[0, D\left\lceil \frac{\ell_2 - D}{d_T} \right\rceil\right] \leq D\frac{\ell_2}{d_T}$ rounds. Adding up the overall number of rounds performed by the algorithm and using that $\ell_2 = \max[t, n-1-t]$ we obtain

$$g_{d_I,d_T}(P_n,t) \leq \frac{D}{d_T}(\max[t, n-1-t] - D + 1) \ + \ 2(2 + d_I/d_T)D.$$

But the second term does not depend on n, hence by using Proposition 7, we obtain that this is a 1^+-approximation.

■

7 Conclusions

We studied the problem of finding the minimum number of rounds needed to gather information in a path in the unitary case. This problem appears to be a difficult one (much more difficult than we thought when starting the research).

We have obtained a 1^+-approximation for any position of the sink. When the sink is an end-vertex of the path, we have also described an incremental procedure which produces optimal protocols when $q+1$ is relatively prime to d_T, which includes the case where d_T is prime and in particular works for $d_T = 2, 3, 5$.

We conjecture that this procedure always gives optimal protocols. One challenging problem is to show that the minimum time gathering problem can be solved in polynomial time in a general or at least in the unitary case or even with the sink at the end of the path. Extending the result to other topologies like trees will be of interest, although optimal solutions might be difficult to obtain in view of the complicated proofs for the case $d_T = 1$, $d_I = 1$ (see [BY10]).

Acknowledgments. We thank the anonymous referees for their suggestions which helped improve the paper, and J. Galtier (France Telecom R&D) for stimulating discussions.

References

[BBS05] P. Bertin, J.-F. Bresse, and B.L. Sage. Accèss haut débit en zone rurale: une solution "ad hoc". *France Telecom R&D*, 22:16–18, 2005.

[BCY09] J-C. Bermond, R. Corrêa, and M.-L. Yu. Optimal gathering protocols on paths under interference constraints. *Discrete Mathematics*, 309(18):5574–5587, September 2009.

[BGK+06a] J-C. Bermond, J. Galtier, R. Klasing, N. Morales, and S. Pérennes. Gathering in specific radio networks. In *Huitièmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications (AlgoTel'06)*, pages 85–88, Trégastel, France, May 2006.

[BGK+06b] J-C. Bermond, J. Galtier, R. Klasing, N. Morales, and S. Pérennes. Hardness and approximation of gathering in static radio networks. *Parallel Processing Letters*, 16(2):165–183, June 2006.

[BGP98] J-C. Bermond, L. Gargano, and S. Pérennes. Optimal sequential gossiping by short messages. *Discrete Applied Mathematics*, 86:145–155, 1998.

[BGP+11] J-C. Bermond, L. Gargano, S. Pérennes, A.A. Rescigno, and U. Vaccaro. Optimal time data gathering in wireless networks with omni-directional antennas. In *SIROCCO 2011*, volume 6796 of *Lecture Notes in Computer Science*, pages 306–317, Gdansk, Poland, June 2011. Springer-Verlag.

[BGR10] J-C. Bermond, L. Gargano, and A.A. Rescigno. Gathering with minimum completion time in sensor tree networks. *Journal of Interconnection Networks*, 11(1-2):1–33, 2010.

[BGRV98] J-C. Bermond, L. Gargano, A.A. Rescigno, and U. Vaccaro. Fast gossiping by short messages. *SIAM Journal on Computing*, 27(4):917–941, 1998.

[Bia00] G. Bianchi. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal of Selected Areas of Communication*, 18:535–547, 2000.

[BKK+10] V. Bonifaci, R. Klasing, P. Korteweg, L. Stougie, and A. Marchetti-Spaccamela. *Graphs and Algorithms in Communication Networks*, chapter Data Gathering in Wireless Networks, pages 357–377. Springer Monograph Springer-Verlag. A. Koster and X. Munoz, editors, 2010.

[BKMS08] V. Bonifaci, P. Korteweg, A. Marchetti-Spaccamela, and L. Stougie. An approximation algorithm for the wireless gathering problem. *Operations Research Letters*, 36(5):605 – 608, 2008.

[BP12] J-C. Bermond and J. Peters. Optimal gathering in radio grids with interference. *Theoretical Computer Science*, 457:10–26, October 2012.

[BY10] J-C. Bermond and M.-L. Yu. Optimal gathering algorithms in multi-hop radio tree networks with interferences. *Ad Hoc and Sensor Wireless Networks*, 9(1-2):109–127, 2010.

[CGL02] M. Christersson, L. Gasieniec, and A. Lingas. Gossiping with bounded size messages in ad-hoc radio networks. In *Proceedings of 29th International Colloquium on Automata, Languages and Programming (ICALP'02)*, volume 2380 of *Lecture Notes in Computer Science*, pages 377–389. Springer-Verlag, 2002.

[CGR02] M. Chrobak, L. Gasieniec, and W. Rytter. Fast broadcasting and gossiping in radio networks. *Journal of Algorithms*, 43(2):177–189, 2002.

[CW91] I. Chlamtac and O. Weinstein. The wave expansion approach to broadcasting in multihop radio networks. *IEEE Transaction on Communications*, 39(3):426–433, 1991.

[EK04] M.L. Elkin and G. Kortsarz. Logarithmic inapproximability of the radio broadcast problem. *Journal of Algorithms*, 52(1):8–25, 2004.

[FFM04] C. Florens, M. Franceschetti, and R.J. McEliece. Lower bounds on data collection time in sensory networks. *IEEE Selected Areas in Communications*, 22(6):1110–1120, 2004.

[Gal04] J. Galtier. Optimizing the IEEE 802.11b performance using slow congestion window decrease. In *Proc. 16th ITC Specialist Seminar on Performance Evaluation of Wireless and Mobile Systems*, pages 165–176, 2004.

[Gar07] L. Gargano. Time optimal gathering in sensor networks. In Giuseppe Prencipe and Shmuel Zaks, editors, *Structural Information and Communication Complexity, 14th International Colloquium, SIROCCO 2007, Castiglioncello, Italy, June 5-8, 2007, Proceedings*, volume 4474 of *Lecture Notes in Computer Science*, pages 7–10. Springer, 2007.

[GM03] I. Gaber and Y. Mansour. Centralized broadcast in multihop radio networks. *Journal of Algorithms*, 46(1):1–20, 2003.

[GP02] L. Gasieniec and I. Potapov. Gossiping with unit messages in known radio networks. In *Proceedings of the IFIP 17th World Computer Congress - TC1 Stream / 2nd IFIP International Conference on Theoretical Computer Science*, pages 193–205. Kluwer, B.V., 2002.

[GR06] L. Gargano and A. A. Rescigno. Optimally fast data gathering in sensor networks. In Rastislav Kralovic and Pawel Urzyczyn, editors, *MFCS*, volume 4162 of *Lecture Notes in Computer Science*, pages 399–411. Springer, 2006.

[GR09] L. Gargano and A.A. Rescigno. Collision-free path coloring with application to minimum-delay gathering in sensor networks. *Discrete Applied Mathematics*, 157(8):1858–1872, 2009.

[HKP+05] J. Hromkovič, R. Klasing, A. Pelc, P. Ružička, and W. Unger. *Dissemination of Information in Communication Networks: Part I. Broadcasting, Gossiping, Leader Election, and Fault-Tolerance*. Springer Monograph. Springer-Verlag, 2005.

[KDN03] K. Kalpakis, K. Dasgupta, and P. Namjoshi. Efficient algorithms for maximum lifetime data gathering and aggregation in wireless sensor networks. *Computer Networks*, 42(6):697–716, August 2003.

[KMP08] R. Klasing, N. Morales, and S. Pérennes. On the complexity of bandwidth allocation in radio networks. *Theoretical Computer Science*, 406(3):225 – 239, 2008.

[Müh02] P. Mühlethaler. *802.11 et les réseaux sans fil*. Eyrolles, 2002.

[RV06] R. Rajagopalan and P. Varshney. Data-aggregation techniques in sensor networks: a survey. *IEEE Communications Surveys & Tutorials*, 8(4):48–63, 2006.

[SW06] S. Schmid and R. Wattenhofer. Algorithmic models for sensor networks. In *14th International Workshop on Parallel and Distributed Real-Time Systems (WPDRTS), Island of Rhodes, Greece*, April 2006.

[Wan09] P.-J. Wan. Multiflows in multihop wireless networks. In *ACM MobiHoc*, pages 85–94, 2009.