A COMPARATIVE STUDY OF PROTOCOLS FOR EFFICIENT
DATA PROPAGATION IN SMART DUST NETWORKS

IOANNIS CHATZIGIANNAKIS
Computer Technology Institute, P.O. Box 1122, 26110 Patras, Greece
Dept. of Computer Engineering and Informatics, University of Patras, 26500 Patras, Greece

and

TASSOS DIMITRIOU
Athens Information Technology
P.O. Box 68, 19002 Peania, Athens, Greece

and

MARIOS MAVRONICOLAS
Dept. of Computer Science, University of Cyprus
CY-1678 Nicosia, Greece

and

SOTIRIS NIKOLETSEAS
Computer Technology Institute, P.O. Box 1122, 26110 Patras, Greece
Dept. of Computer Engineering and Informatics, University of Patras, 26500 Patras, Greece

and

PAUL SPIRAKIS
Computer Technology Institute, P.O. Box 1122, 26110 Patras, Greece
Dept. of Computer Engineering and Informatics, University of Patras, 26500 Patras, Greece

Received [received date]
Revised [revised date]
Communicated by [Name of Editor]

* A preliminary version of this paper has appeared at the 9th International Conference on Parallel
† This work has been partially supported by the 1ST Programme of the European Union under
ABSTRACT

Smart Dust is comprised of a vast number of ultra-small fully autonomous computing and communication devices, with very restricted energy and computing capabilities, that co-operate to accomplish a large sensing task. Smart Dust can be very useful in practice i.e. in the local detection of a remote crucial event and the propagation of data reporting its realization to a control center. In this work, we have implemented and experimentally evaluated four protocols (PFR, LTP and two variations of LTP which we here introduce) for local detection and propagation in smart dust networks, under new, more general and realistic modelling assumptions. We comparatively study, by using extensive experiments, their behavior highlighting their relative advantages and disadvantages. All protocols are very successful. In the setting we considered here, PFR seems to be faster while the LTP based protocols are more energy efficient.

1. Introduction, our Results and Related Work

Recent dramatic developments in micro-electro-mechanical (MEMS) systems, wireless communications and digital electronics have already led to the development of small in size, low-power, low-cost sensor devices. Such extremely small devices integrate sensing, data processing and communication capabilities. Examining each such device individually might appear to have small utility, however the effective distributed co-ordination of large numbers of such devices may lead to the efficient accomplishment of large sensing tasks.

Large numbers of small in size, low-power, low-cost sensor devices can be deployed in areas of interest (such as inaccessible terrains or disaster places) and use self-organization and collaborative methods to form a sensor network. Their wide range of applications is based on the possible use of various sensor types (i.e. thermal, visual, seismic, acoustic, radar, magnetic, etc.) in order to monitor a wide variety of conditions (e.g. temperature, object presence and movement, humidity, pressure, noise levels etc.). Thus, sensor networks can be used for continuous sensing, event detection, location sensing as well as micro-sensing. Hence, sensor networks have important applications, including (a) military (like forces and equipment monitoring, battlefield surveillance, targeting, nuclear, biological and chemical attack detection), (b) environmental applications (such as fire detection, flood detection, precision agriculture), (c) health applications (like telemonitoring of human physiological data) and (d) home applications (e.g. smart environments and home automation). For an excellent survey of wireless sensor networks see [2] and also [3,4].

Note however that the efficient and robust realization of such large, highly-dynamic, complex, non-conventional networking environments is a challenging algorithmic and technological task. Features including the huge number of sensor devices involved, the severe power, computational and memory limitations, their dense deployment and frequent failures, pose new design and implementation aspects which are essentially different not only with respect to distributed computing and systems approaches but also to ad-hoc networking techniques.

1.1. Contribution

We focus on an important problem under a particular model of sensor networks
that we present. More specifically, continuing the research of our team (see [5], [6]), we study the problem of local detection and propagation, i.e. the local sensing of a crucial event and the energy and time efficient propagation of data reporting its realization to a (fixed or mobile) control center. This center may be some human authorities responsible of taking action upon the realization of the crucial event. We use the term “sink” for this control center. We note that the protocols we present here can also be used for the more general problem of data propagation in sensor networks (see [7]).

As opposed to [8] (where a 2-dimensional lattice deployment of particles is used) we extend the network model to the general case of particle deployment according to a random, uniform distribution. We study here the more realistic case when the control center is not a line in the plane (i.e. as in [5]) but a single point.

Under these more general and realistic (in terms of motivation by applications) modelling assumptions, we implemented and experimentally evaluated four information propagation protocols: (a) The Probabilistic Forwarding Protocol (PFR) that avoids flooding by favoring in a probabilistic way certain “close to optimal” transmissions, (b) the Local Target Protocol (LTP), where data is propagated by each time selecting the best (with respect to progress towards the sink) particle for passing information and (c) we propose two variations of LTP according to different next particle selection criteria. We note that we had to carefully design the protocols to work under the new network models.

The extensive simulations we have performed show that all protocols are very successful. In the particular setting we considered here, PFR seems to achieve high success rates in terms of time and hops efficiency, while the LTP based protocols manage to reduce the energy spent in the process by activating less particles.

1.2. Selected Related Work

A family of negotiation-based information dissemination protocols suitable for wireless sensor networks is presented in [9]. A data dissemination paradigm using interests for named data, called directed diffusion, is presented in [7]. In [11] a clustering-based protocol is given that utilizes randomized rotation of local cluster heads to evenly distribute the energy load among the sensors in the network. In [12] a new energy efficient routing protocol is introduced where nodes transmit data only when sudden and drastic changes are sensed (as opposed to [11]). A modified version of the PFR protocol presented here has been proposed and comparatively evaluated with [11,12] in [13].

We note, as opposed to the work presented in this paper, the above research focuses on energy consumption without examining the time efficiency of their protocols. Furthermore, note that our protocols are quite general in the sense that (a) we do not assume global network topology information, (b) we do not assume geolocation information (such as GPS information) and (c) we use very few control message exchanges, thus having low communication overhead.

2. The Model

Sensor networks are comprised of a vast number of ultra-small homogenous sen-
sors. Each particle is a fully-autonomous computing and communication device, characterized mainly by its available power supply (battery) and the energy cost of computation and transmission of data. Such particles (in our model here) cannot move.

Each particle is equipped with a set of monitors (sensors) for light, pressure, humidity, temperature etc. Each particle has a broadcast (digital radio) beacon mode which can be also a directed transmission of angle $\alpha$ around a certain line (possibly using some special kind of antenna, see Fig. 2).

We adopt here (as a starting point) a two-dimensional (plane) framework: A smart dust cloud (a set of particles) is spread in an area (see Fig. 1). Note that a two-dimensional setting is also used in [7, 9, 10, 11, 12].

![Fig. 1. A Smart Dust Cloud](image1)

![Fig. 2. Example of the Search Phase](image2)

**Definition 1** Let $n$ be the number of smart dust particles and let $d$ (usually measured in numbers of particles/m$^2$) be the density of particles in the area. Let $R$ be the maximum (beacon/laser) transmission range of each grain particle.

There is a single point in the network area, which we call the sink $S$, and represents a control center where data should be propagated to. Note that, although in the basic case we assume the sink to be static, in a variation it may be allowed to move around its initial base position, to possibly get data that failed to reach it but made it close enough to it.

Furthermore, we assume that there is a set-up phase of the smart dust network, during which the smart cloud is dropped in the terrain of interest, when using special control messages (which are very short, cheap and transmitted only once) each smart dust particle is provided with the direction of $S$. By assuming that each smart-dust particle has individually a sense of direction, and using these control messages, each particle is aware of the general location of $S$.

We feel that our model, although simple, depicts accurately enough the technological specifications of real smart dust systems. Similar models are being used by other researchers in order to study sensor networks. In contrast however to [7, 14], our model is weaker in the sense that no geolocation abilities are assumed (e.g. a GPS device) for the smart dust particles leading to more generic and thus stronger results.

3. The Problem
Assume that a single particle, \( p \), senses the realization of a local crucial event \( \mathcal{E} \). Then the propagation problem \( \mathcal{P} \) is the following:

“How can particle \( p \), via cooperation with the rest of the grain particles, efficiently propagate information \( \text{info}(\mathcal{E}) \) reporting realization of event \( \mathcal{E} \) to the sink \( S \)?”

Because of the dense deployment of particles close to each other, multi-hop communication consumes less power than a traditional single hop propagation. Also, multi-hop communication can effectively overcome some of the signal propagation effects in long-distance wireless transmissions. Furthermore, short-range hop-by-hop transmissions may help to smoothly adjust propagation around obstacles. Finally, the low energy transmission in hop-by-hop may enhance security, protecting from undesired discovery of the data propagation operation. To minimize the energy consumption in the sensor network we wish to minimize the number of hops (directed transmissions) performed in the data propagation process. Note that the number of hops also characterizes propagation time, assuming an appropriate MAC protocol [15].

Furthermore, an interesting aspect is how close to the sink data is propagated (in the case where data does not exactly reach the sink). Note that proximity to the sink might be useful in the case where the sink is mobile or it performs a “limited flooding” to get the information from the final position it arrived.

4. The Probabilistic Forwarding Protocol (PFR)

The basic idea of the protocol lies in probabilistically favoring transmissions towards the sink within a thin zone of particles around the line connecting the particle sensing the event \( \mathcal{E} \) and the sink (see Fig. 3). Note that transmission along this line is optimal w.r.t. energy and time. However it is not always possible to achieve this optimality, basically because, even if initially this direct line was appropriately occupied by sensors, certain sensors on this line might become inactive, either permanently (because their energy has been exhausted) or temporarily (because they might enter a sleeping mode to save energy).

![Figure 3: Thin zone of particles](image)

![Figure 4: Angle \( \phi \)](image)
The protocol evolves in two phases:

**Phase 1: The “Front” Creation Phase.** Because of the probabilistic decision on whether to forward messages or not, initially we build (by using flooding) a sufficiently large “front” of particles, in order to guarantee the survivability of the propagation process. During this phase, each particle having received the data, deterministically forwards them towards the sink. In particular, and for a sufficiently large number of steps, each particle broadcasts the information to all its neighbors, using a directed (towards the sink) angle transmission.

**Phase 2: The Probabilistic Forwarding Phase.** During this phase, each particle possessing the information under propagation, probabilistically favors its transmission within a thin zone of sensors lying close to the (optimal) line between the particle that sensed \( \mathcal{E} \) and \( \mathcal{S} \). In other words, data is propagated with a suitably chosen probability \( p \), while it is not propagated with probability \( 1 - p \), based on a random choice.

This probability is calculated as follows: Let \( \phi \) the angle defined by the line connecting \( \mathcal{E} \) and the sensor performing the random choice, and the line defined by the position of this particle and \( \mathcal{S} \) (see Fig. 4).

To limit the propagation zone we choose the forwarding probability \( P_{\text{forward}} \) to be

\[
    P_{\text{forward}} = \frac{\phi}{\pi}
\]

Remark that indeed a bigger angle \( \phi \) suggests a sensor position closer to the direct line between \( \mathcal{E} \) and \( \mathcal{S} \). Clearly, when \( \phi = \pi \), then the sensor lies on this line. Also note that calculations of \( \phi \) needs only local information. Figure 4 displays this graphically.

Thus, we get that \( \phi_1 > \phi_2 \) implies that for the corresponding particles \( p_1, p_2, p_1 \) is closer to the \( \mathcal{E}-\mathcal{S} \) line than \( p_2 \), thus

\[
    P_{\text{forward}}(p_1) = \frac{\phi_1}{\pi} > \frac{\phi_2}{\pi} = P_{\text{forward}}(p_2)
\]

**Angle \( \phi \) calculation.** Under appropriate (and realistic) modelling assumptions for sensor particles, angle \( \phi \) can be locally calculated running a simple subprotocol (see [8] for detailed discussion). Such modelling assumptions include a) the ability of sensor particles to estimate the direction of a received transmission (e.g. by direction-sensing antennae), b) to estimate the distance from a nearby particle that did the transmission (e.g. via signal attenuation estimation techniques), c) to have a common a common co-ordinates system and d) to know the direction towards the sink (this is possibly done during a set-up phase). Note that we do not need GPS information or global network structure knowledge.

5. The Local Target Protocol (LTP) and its Variations

We now present the “Local Target” protocol and two variations that use different next particle selection criteria. In this protocol, each particle \( p' \) that has received \( \text{info(}\mathcal{E}\text{)} \) from \( p \) (via, possibly, other particles) does the following:
**Phase 1: The Search Phase.** It uses a periodic low energy broadcast of a beacon in order to discover a particle nearer to \( S \) than itself. Among the particles returned, \( p' \) selects a unique particle \( p'' \) that is “best” with respect to progress towards the sink. We here consider two criteria for measuring the quality of this progress and also a randomized version to get a good average mix:

(a) **Euclidean Distance.** In this case, the particle \( p''_E \) that among all particles found achieves the bigger progress on the \( p'/S \) line, should be selected. We call this variation of our protocol LTP\(_E\). This is considered as the “basis” LTP.

(b) **Angle Optimality.** In this case, the particle \( p''_A \) such that the angle \( \hat{p''_A}p'/S \) is minimum, should be selected. We call this variation of our protocol LTP\(_A\).

(c) **Randomization.** Towards a good average case performance of the protocol, we use randomization to avoid bad behavior due to the worst case input distributions for each selection (i.e., particles forming big angles with the optimal line in LTP\(_E\) and particles resulting to small Euclidean progress in LTP\(_A\)). Thus, we find \( p''_E, p''_A \) and randomly select one of them, with probability \( \frac{1}{2} \). We call this protocol LTP\(_R\).

**Phase 2: The Direct Transmission Phase** Then, \( p' \) sends \( \text{info}(\mathcal{E}) \) to \( p'' \) and sends a **success** message to \( p \) (i.e., to the particle that it originally received the information from).

**Phase 3: The Backtrack Phase.** If the search phase fails to discover a particle nearer to \( S \), then \( p' \) sends a **fail** message to \( p \).

In the above procedure, propagation of \( \text{info}(\mathcal{E}) \) is done in two steps: (i) particle \( p' \) locates the next particle \( (p'') \) and transmits the information and (ii) particle \( p' \) waits until the next particle \( (p'') \) succeeds in propagating the message further towards \( S \). This is done to speed up the backtrack phase in case \( p'' \) does not succeed in discovering a particle nearer to \( S \).

6. **Efficiency Measures**

**Definition 2** Let \( h_A \) (for “active”) be the **number of “active” sensor particles** participating in the data propagation and let \( E_{TR} \) be the **total number of data transmissions** during propagation. Let \( T \) be the **total time** for the propagation process to reach its final position and \( H \) the **total number of “hops”** required.

Clearly, by minimizing \( h_A \) we succeed in avoiding flooding and thus we minimize energy consumption. Remark that in LTP we count as active those particles that transmit \( \text{info}(\mathcal{E}) \) at least once.

Note that \( h_A, E_{TR}, T \) and \( H \) are random variables. Furthermore, we define the **success probability** of our algorithm where we call success the eventual data propagation to the sink.

**Definition 3** Let \( P_s \) be the **success probability** of our protocol.

We also focus on the study of the following parameter. Suppose that the data propagation fails to reach the sink. In this case, it is very important to know “how
close” to the sink it managed to get. Propagation reaching close to the sink might be very useful, since the sink (which can be assumed to be mobile) could itself move (possibly by performing a random walk) to the final point of propagation and get the desired data from there. Even assuming a fixed sink, closeness to it is important, since the sink might in this case begin some “limited” flooding to get to where data propagation stopped. Clearly, the closer to the sink we get, the cheaper this flooding becomes.

**Definition 4** Let $F$ be the final position of the data propagation process. Let $D$ be $F$’s (Euclidean) distance from the sink $S$.

Clearly in the case of total success $F$ coincides with the sink and $D = 0$.

7. Experimental Evaluation

We evaluate the performance of the four protocols by a comparative experimental study. The protocols have been implemented as C++ classes using the data types for two-dimensional geometry of LEDA [16] based on the environment developed in [5], [6]. Each class is installed in an environment that generates sensor fields given some parameters (such as the area of the field, the distribution function used to drop the particles), and performs a network simulation for a given number of repetitions, a fixed number of particles and certain protocol parameters.

![Success Probability](image)

Fig. 5. Success Probability $P_{su}$ over particle density $d = [0.01, 0.3]$.

In our experiments, we generate a variety of sensor fields in a 100m by 100m square. In these fields, we drop $n \in [100, 3000]$ particles randomly uniformly distributed on the smart-dust plane, i.e. for densities $0.01 \leq d \leq 0.3$. Each smart dust particle has a fixed radio range of $R = 5m$ and $\alpha = 90^\circ$. The particle $p$ that initially senses the crucial event is always explicitly positioned at $(x, y) = (0, 0)$ and the sink is located at $(x, y) = (100, 100)$. Note that this experimental setup is based on and extends that used in [10,11,12]. We repeated each experiment for more than 5,000 times in order to achieve good average results.

We start by examining the success rate of the four protocols (see Fig. 5), for different particle densities. Initially, when the density is low (i.e. $d \leq 0.06$), the
protocols fail to propagate the data to the sink. However as the density increases, the success rate increases quite fast and for high densities, all four protocols almost always succeed in propagating the data to the sink. Thus, all protocols are very successful. We remark a similar shape of the success rate function in terms of density. This is due to the fact that all protocols use local information to decide how to proceed by basically selecting (all protocols) the next particle with respect to a similar criterion (best progress towards the sink).

In the case when the protocols fail to propagate the data to the sink, we examine “how close” to the sink they managed to get. Figure 6 depicts the distance of the final point of propagation to the position of the sink. Note that this figure should be considered in conjunction with Fig. 5 on the success rate. Indeed, failures to reach the sink are very rare and seem to appear in very extreme network topologies due to bad particle distribution on the network area.
Figure 7 depicts the ratio of active particles over the total number of particles \( r = \frac{A^+}{A^-} \) that make up the sensor network. In this figure we clearly see that PFR, for low densities (i.e. \( d \leq 0.07 \)), indeed activates a small number of particles (i.e. \( r \leq 0.3 \)) while the ratio \( r \) increases as the density of the particles increases. The LTP based protocols seem more efficient and the ratio \( r \) seems to be independent of the total number of particles (because only one particle in each hop becomes active).

**Remark:** Because of the way PFR attempts to avoid flooding (by using angle \( \phi \) to capture “distance” from optimality) its merits are not sufficiently shown in the setting considered here. We expect PFR to behave significantly better with respect to energy in much larger networks and in cases where the event is sensed in an average place of the network. Also, stronger probabilistic choices (i.e. \( P_{fwd} = \left( \frac{\phi}{\pi} \right)^a \), where \( a > 1 \) a constant) may further limit propagation time.
Furthermore, examining the total number of transmissions performed by the
particles (see Fig. 8), it is evident that because the LTP based protocols activate
a small number of particles, the overall transmissions are kept low. This is a sur-
prising result, since the PFR protocol was originally designed to work without the
need of any control messages so that the energy consumption is low. However, the
comparative study clearly shows that avoiding the use of control messages does not
achieve the expected results. So, even though all four protocols succeed in propa-
gating the data, it is evident that the LTP based protocols are more energy efficient
in the sense that less particles are involved in the process.

We continue with the following two parameters: (a) the “hops” efficiency and
(b) the time efficiency, measured in terms of rounds needed to reach the sink. As
can be seen in Fig. 9, all protocols are very efficient, in the sense that the number
of hops required to get to the sink tends below 40 even for densities \( d \geq 0.17 \).
The value 40 in our setting is close to optimality since in an ideal placement, the
diagonal line is of length \( 100\sqrt{2} \) and since the transmission range \( R = 5 \) the optimal
number of hops (in an ideal case) is roughly 29. In particular PFR achieves this for
very low densities \( d \geq 0.07 \). On the other hand, the LTP based protocols exhibit a
certain pathological behavior for low densities (i.e. \( d \leq 0.12 \)) due to a high number
of executions of the backtrack mechanism in the attempt to find a particle closer to
the sink (see also Fig. 10).

![Graph](image)

Fig. 10. Average Number of Backtracks over particle density \( d = [0.01, 0.3] \).

Finally, in Fig. 10 we compare the three LTP based protocols and the number
of backtracks invoked in the data propagation. It is evident that for very low
particle densities (i.e. \( d \leq 0.12 \)), all three protocols perform a large number
of backtracks in order to find a valid path towards the sink. As the particle density
increases, the number of backtrack reduces fast enough and almost reaches zero.

8. Future Work

We plan to also investigate alternative probabilistic choices for favoring certain data
transmissions for the PFR protocol and consider alternative backtrack mechanisms.
for the LTP protocol.

Such different probabilistic choices may include a) stronger choices i.e. when the forwarding probability becomes smaller. We expect this to further limit the number of particles participating in the data propagation process (thus further reducing energy consumption in the network). However, there is a trade-off with the time needed to get to the sink and fault-tolerance as well, that should be investigated, both by analytic and experimental means. b) energy-related probabilistic choices: to explicitly focus on energy consumption, the forwarding probability at each sensor can be made proportional to the remaining energy at the sensor, thus discouraging transmissions at sensors whose remaining energy is low, in order to avoid early energy depletion of sensors. This is expected to contribute towards a certain balance of the energy load in the network, thus prolonging the network’s lifetime. c) other choices may include forwarding probabilities depending on certain topology-related, local metrics, aiming at satisfying other (besides energy efficiency) important properties (such as obstacle avoidance, fault-tolerance etc.).

Also, we wish to study different network shapes, various distributions used to drop the sensors in the area of interest and the fault-tolerance of the protocols.

We indeed plan to carry out a detailed study (both by theoretical analysis and simulation) on the particular effect of the network characteristics. Important input parameters include the density and the distribution of the sensors in the network. We will further analyze a main finding in this work: that the PFR protocol seems more suitable in low density networks, while LTP is more energy-efficient in the case of high densities.

In particular, we aim at evaluating explicit relations between energy consumption and density in each case, and accordingly locate ranges of best performance for each protocol. We also plan to examine hybrid combinations of the two protocols and switching mechanisms between the two as the network characteristics vary.

Finally, we plan to provide performance comparisons with other protocols mentioned in the related work section, as well as implement and evaluate hybrid approaches that combine the PFR and LTP protocols in a parameterized way.

References

[5] I. Chatzigiannakis, S. Nikoletseas and P. Spirakis, Smart Dust Protocols for Local De-