Tuning the Epidemical Algorithm in Wireless Sensor Networks

Kostis Gerakos[1], Christos Anagnostopoulos[2], and Stathes Hadjiefthymiades[1]

[1] Pervasive Computing Research Group, Department of Informatics and Telecommunications, University of Athens, Athens, Greece
[2] Department of Informatics, Ionian University, Corfu, Greece
kostis@dtps.unipi.gr,{bleu,shadj}@di.uoa.gr

Abstract. We discuss the networking dimension of the Integrated Platform for Autonomic Computing (IPAC). IPAC supports the development and running of fully distributed applications that rely on infrastructureless (ad-hoc) network with multi-hop transmission capabilities. Such environment is typically used for the realization of collaborative context awareness where nodes with sensors "generate" and report context while other nodes receive and "consume" such information (i.e., feed local applications with it). Due to its highly dynamic character this application environment, an efficient solution for the dissemination of information within the network involves the adoption of epidemical algorithms. With the use of such algorithms, a certain node spreads information probabilistically to its neighborhood. Evidently this is a rational approach since the neighborhood changes frequently and nodes are not necessarily in need of the generated contextual stream. IPAC mainly targets embedded devices such as OS-powered sensor motes, smartphones and PDAs. The platform relies on the OSGi framework (a popular middleware for embedded devices) for component deployment, management and execution. We discuss implementation issues focusing on the broad spectrum of IPAC services that were developed in order to facilitate applications. We elaborate on the networking stack that implements epidemical dissemination. We also discuss how such infrastructure has been used to realize applications related to crisis management and environmental protection. We present an adaptive flavor of the epidemical dissemination which expedites delivery by tuning the forwarding probability whenever an alarming situation is detected.

Keywords: wireless sensor networks, dissemination, epidemical algorithm, wildfire, epidemic model, forwarding probability.

1 Introduction

During the last few years advantage of WSN (wireless sensor networks) gave affordable and smart solutions to "real life" problems and situations. IPAC aims at providing all the communication functionality for fully distributed applications including medical monitoring and emergency response applications, monitoring remote or inhospitable habitats, disaster relief networks, crisis management applications, early fire detection in forests, and environmental monitoring.

H. Jahankhani et al. (Eds.): ICGS3/e-Democracy 2011, LNICST 99, pp. 31–37, 2012.

Following nature's example a good way of disseminating information over WSN are the so called epidemical algorithms and gossip protocols solving the underlying problems that comes with the lack of a secure direct path for the data to be delivered. In this paper we discuss the efficiency in information dissemination of IPAC platform in a emergency case of a wildfire. In such cases multiple factors can be used for understanding the state of fire and predict the spreading so the data to be delivered safely. We use a temperature vs time model behavior as an example to explain the spreading algorithms of IPAC project.

2 Wildfire Behavior

Before further explaining the dissemination of information we must describe the factors that are taken into consideration during the spread of a wildfire into a forest monitored from IPAC sensors. In such cases sensors measuring temperature are preferred from others more sophisticated, like for example heat flux measuring sensors, due of being inexpensive and more reliable.

2.1 General Behavior

According to [1] the temperature course of a wildfire may be divided into three periods:

1. The growth period.
2. The fully developed period
3. The decay period.

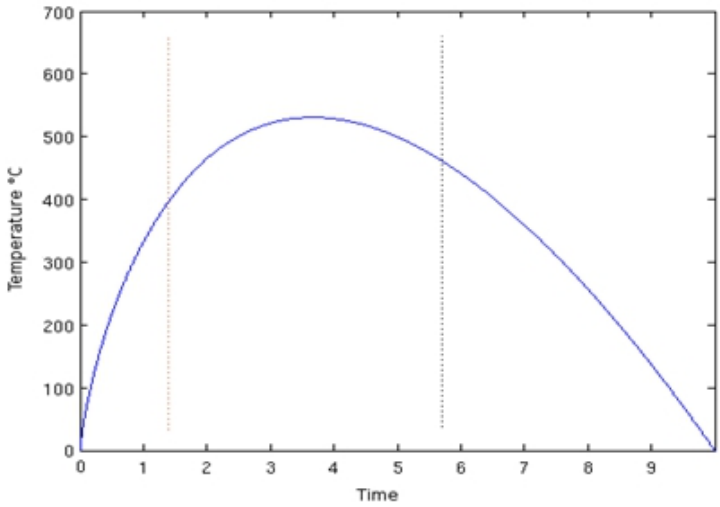

Fig. 1. Temperature produced by a wildfire versus time measured from a constant point. Can be used for emulation of the results that a sensor will receive while measuring absolute values of temperature.

These periods are illustrated in the figure below, where an idealized fire temperature course is shown. During the growth period, heat produced by the burning trees increasing rapidly to high values. After a certain temperature the fully developed period starts and then follows the decay period.

2.2 Wildfire Modeling

According to [5] a relationship between temperature, distance from flame and flame length is the following:

$$I = 300 \cdot L^2 \tag{1}$$

$$Q = 60\left[1 - exp\left(\frac{-1}{300 \cdot D}\right)\right] \tag{2}$$

where I, L, Q and D denote fire intensity (in kW/m), flame length (in meters), radiation intensity (in kW/m2) and the distance from the flame position (in meters) respectively. Assuming a linear T = a Q relationship (with a=10 as inferred from [3] and [4]) we can establish a temperature distance T = f_L(D) relationship for any given L. On the other hand for large flame height, the authors in [6] have presented a model that estimates the net radiant energy transfer to a fire fighter standing at a specified distance D from a fire of a height L. It is observed that the larger the flame height the larger is the distance from the flame that would result to a specific heat flux (and sensed temperature) value. We can use this work to extract a T = g_L(D) relationship for large flame heights, L > 8m. It is interesting that the two different methods, for small [5] and large [6] flame sizes, produce approximately the same temperature estimates at the cutoff point of L=8m.

3 Adaptive Epidemic Model

In this section we introduce the concept of adjusting parameters of the SIS epidemic model for achieving efficient valid information dissemination in an IPAC WSN. Consider a discrete time domain, i.e., t ∈ N. We consider a IPAC WSN consisting of N + 1 nodes. Each node is indexed by an integer i ∈ N. The node 0 is the source, which generates and transmits data values. The node i = 1, . . . ,N is the (consumer) relay node, which receives, stores and forwards. Nodes disseminate data if they are within the communication range of each other. The source node 0 is equipped with a sensor (for our example a temperature sensor). At each discrete time instance t, the node 0 generates a data value x(t) of a contextual parameter (e.g., temperature data) from an attached sensor and a temporal validity value v(t) ∈ R. The parameter is sampled with frequency z. The υ(t) value indicates the maximum time-horizon that a value x is considered valid, that is, υ(t) = t_E − t. The υ(t) value decreases with time. A υ(t) = 0 indicates that the value x(t) turns obsolete at t. The source forwards the pairs (x(0), υ(0)), (x(1), v(1)), . . .) with a time–varying forwarding probability β(t) ∈ (0, 1]. Any

relay node i = 1, . . . ,N, which receives x(t), becomes infected at time t once $\upsilon(t) > 0$ and forwards x(t) to its neighbors with constant forwarding probability $\gamma \in (0, 1]$. Since we have only one source, an infected node, which has recently received (x(t), $\upsilon(t)$), can be re-infected with some received (x(t+1), $\upsilon(t+1)$) at time instance t+1 since $\upsilon(t+1) > \upsilon(t)$, i.e., x(t+1) is more valid data than x(t).

We now discuss the significance of the forwarding probabilities γ for relay nodes and $\beta(t)$ for the source. A high value of γ and $\beta(t)$ leads to (almost) full network coverage (i.e., information diffusion among the nodes) but at the expense of increased energy consumption due to redundant transmissions –receptions of messages. For $\gamma = 1$ we obtain the Flooding scheme. On the other hand, low values of γ and β (t) lead to a global ageing of information throughout the network, i.e., the consumer nodes receive (and process) information of lower quality due to the elongation of the time interval $\gamma - t_G$(reception time-generation time). In addition, the delay of a received piece of information is measured as the time interval between t_G and the reception time at some node in which the information is considered usable. A consumer node relies on the last received piece of data (for further processing by the upper layers) until a new one is received. With a low value of $\beta(t)$ and a low value of γ a newly generated piece of information is received with increased delay. Hence, a random relay node is badly synchronized with the source. A relay node is considered well synchronized with the source if the time interval $(\gamma - t_G)$ is relatively small. On the other hand, synchronization is negatively impacted if the discussed time interval increases. Let us introduce a global error indicator e to clearly indicate the impact of delayed data delivery to consumer nodes. The error indicator captures the timed data value difference between the source and relays. Surely, the e indicator should be kept at low levels and can be adopted as a metric for the assessment of the proposed scheme. Good synchronization leads to a low e value while the opposite holds for poor synchronization.

Evidently, there is a trade-off between data dissemination efficiency and validity. Our idea is to adjust the $\beta(t)$ value on the source prior to injecting information to the WSN according to the data stream(DS) variability experienced there. The DS variability is quantified through the rate of change of the disseminated pieces of information. Intuitively, a DS of high variability has to be disseminated by the source with higher $\beta(t)$ than a DS with low variability. High variability in the DS is manifested through frequent changes in the observed (sampled) quantity.

A high value of γ for the relay nodes safeguards the rapid dissemination of information throughout the network. Distinct values generated by the source (with probability $\beta(t)$) reach the various consumers in the network rapidly. Hence the induced error indicator drops and synchronization improves. A low $\beta(t)$ value increases the inter-arrival time for data readings messages at relay nodes. Received values are not promptly updated and become stale and obsolete. Despite their ageing and expiration, such values can still be exploited by relays, since the DS is relatively static i.e., exhibits low variability. Apart from the discussed $\beta(t)$ tuning, another, closely related parameter (that still qualifies for tuning) is the time-to-live (TTL) of the disseminated data. The maximum temporal validity (TTL) of a sampled piece of information depends on the observed variability of the DS and the relay probability of WSN nodes.

On a high variability DS, the sensed data values have to cover the whole WSN in order all nodes to 'follow' the rate of change of the DS, which is experienced by the source. The proposed adaptive SIS model is entirely data-centric meaning that specific characteristics of the generated DS tune the propagation (dissemination) of information throughout the network.

3.1 The Adaptive Epidemic Model

We use the epidemic model in terms of the adaptive behavior of the source and the forwarding capability of the relay nodes. We assume that the WSN operation starts at t = 0. At that time all the relay nodes are susceptible and the source is infected.

3.2 The Behavior of the Source

Let x_0(t) the temperature that is sensed by the source node 0 at time t. At time t ≥ 0, the source determines the value for the forwarding probability β(t) ∈ (0, 1] by taking into account the rate of change $\frac{\Delta x_0(t)}{\Delta t}$ of the sensed x_0 (t) value(the change of temperature). In addition, the temporal validity value υ(t) at time t, i.e., the TTL of the sensed data value x_0(t), depends also on the variability of the temperature and the relay probability of the relay node γ as described later. We introduce the real functions f and g such that

$$\beta(t) = f\left(\frac{\Delta x_0(t)}{\Delta t}\right), f: \mathbb{R} \longrightarrow (0,1] \tag{3}$$

$$\upsilon(t) = g(\beta(t), \gamma), g: (0,1] \times (0,1] \rightarrow \mathbb{R} \tag{4}$$

The f (·) and g(·,·) functions rely on the nature of the sampled DS x_0(t). The following paragraphs provide details of f (·) functions.

3.3 Adaptive Forwarding Probability

In this section we discuss the characteristics of the f (·) function.

1. The f (·) function produces probability values (forwarding probability β(t))
2. The f (·) function is increased in the interval [0, ∞) of the DS variability percentage change $\left(\frac{|\Delta x_0(t)|}{|x_0 t|}\right)$ The higher the percentage change in DS variability gets, the higher the probability of forwarding x_0 (t) to the neighbors of the source becomes. In such case the relay node I can reconstruct a DS $x_i(t)$ with high variability assuming a small e value. On the other hand, a DS with low rate of change can be disseminated with lower forwarding probability since the x_0 DS remains quite constant. The reduction in β(t) reduces the transmissions for the sake of energy.

3. The f (·) function should be tunable so as to treat the $\frac{|\Delta x_0(t)|}{|x_0 t|}$ values in a non-uniform way. In other words, the f (·) function should be able to assign varying significance to $\frac{|\Delta x_0(t)|}{|x_0 t|}$ ratio values depending on the application and the actual utility of the transmitted data value. In certain cases, a small change in the DS $\left(\frac{|\Delta x_0(t)|}{|x_0 t|}\right)$ can be regarded as noise and suppressed by the network to preserve energy efficiency while a significant change in the DS would be considered as highly important (e.g., an emergency alarm) and affect the f (·) value accordingly.

We adopt the sigmoid s-shape function (structured as shown below) as it allows the ad hoc, discriminative treatment of the DS changes as discussed above. Specifically,

$$f\left(\frac{\Delta x_0(t)}{\Delta t}\right) = \frac{1}{1 + exp\left(-c\left(\frac{|\Delta x_0(t)|}{\Delta t}\frac{1}{|x_0(t)|} - \omega\right)\right)} \tag{5}$$

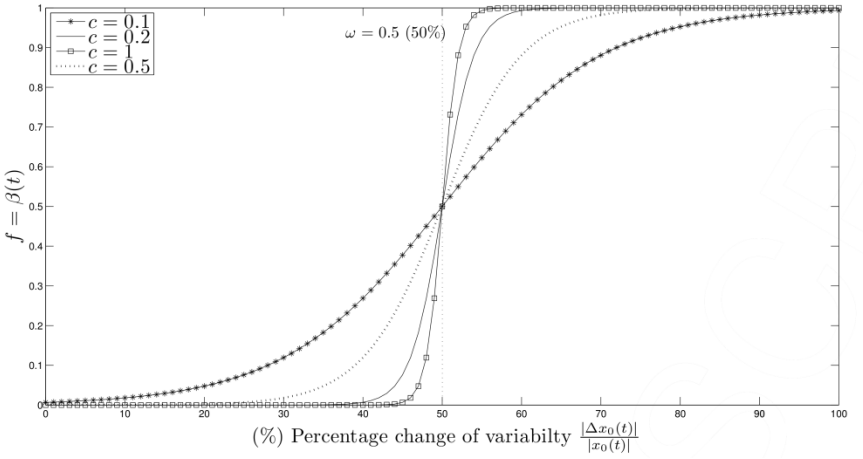

Fig. 2. We observe the different s-shaped graphs for the sigmoid f function versus the percentage change of variability for the different values of constant c. For c = 0.1 the graph approach a linear form.

The c ∈ [0, 1] and ω ∈ (0, ∞) in (5) are the tuning factors of f (·). Without loss of generality, we assume that Δt = 1. The ω parameter is the 'significance threshold', which indicates the percentage change of the variability of the temperature, i.e., $\frac{|\Delta x_0(t)|}{|x_0 t|}$ that yields source forwarding probabilities greater than 0.5. The c parameter is a 'bias factor' that determines the shape of the f (·) function around the ω value. The higher the c the higher the rate of change of f from ω− to ω+ . Evidently, a zero value

of c yields a constant f (·) = 0.5 function which is completely independent of $\frac{|\Delta x_0(t)|}{|x_0 t|}$ (β(t) = 0.5). In this paper, we deal with a fire-detection application[8] . By inspection of the collected data, we noticed percentage value changes of the DS variability lower than 50%. Hence, we adopt as significance threshold ω = 0.5. In the unlikely case $\frac{|\Delta x_0(t)|}{|x_0 t|} > 100\%$ appears in the DS, the β(t) = 1 will be adopted as the ceiling value. In addition, we adopt c = 0.1. Such values are aligned with the characteristics points (3) to (5) of f (·) discussed above.

We applied the fire progress details found in studies like [8] and observed the way the β(t) parameter fluctuates. Our findings are presented in the following table.

Table 1. Results of β(t). Anything above 0.4 alarms the dissemination process.

Condition	β(t)
Regular Sensor Operation(Ambient Temperature)	0.3-0.4
Alarming Conditions(Possible Fire)	0.4+

4 Conclusions

The proposed scheme satisfies two important requirements for sensor networks in critical safety applications (a) it saves energy throughout regular operation, thus, extending the lifetime of the network and improving dependability and (b) is totally reactive to changes of the sensed environmental parameter, thus, guaranteeing the timely issue of alarms and the required follow-up operations.

References

1. Lie, T.T.: Fire temperature time relations. The SFPE Handbook of Fire Protection Engineering (1988)
2. Manolakos, E.S., Manatakis, D.V., Xanthopoulos, G.: Temperature field modeling and simulation of wireless sensor network behavior during a spreading wildfire. In: 16th European Signal Processing Conference, EUSIPCO (2008)
3. Stroup, D., DeLauter, L., Lee, J., Roadarmel, G.: Fire Test of Men's Suits on Racks. In: Proc. Report of Test FR 4013, p. 25 (2001)
4. Pitts, W., Braun, E., Peacock, R., Mitler, H., Johnsson, E., Reneke, P., Blevins, L.: Temperature Uncertainties for Bare-Bead and Aspirated Thermocouple Measurements in Fire Environments. In: Proc. Annual Conference on Fire Research, pp. 15–16 (1999)
5. Kurt, F.: Prometheus Fire Growth Model: Design and Incorporation of Spotting and Breaching of Fire Break Functionality. In: Post-Fire Research Workshop (2005)
6. Butler, B., Cohen, J.: A Theoretical Model Based on Radiative Heating. Int. J. Wildland Fire 8(2), 73–77 (1998)
7. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.: Complex Networks: Structure and Dynamics. Phys. Rep. 424, 175–308 (2006)
8. Zervas, E., Mpimpoudis, A., Anagnostopoulos, C., Sekkas, O., Hadjiefthymiades, S.: Multisensor Data Fusion for Fire Detection. In: Information Fusion, vol. 12(3). Elsevier (2011)