Bio-Inspired Data Acquisition in Sensor Networks

Daniel de O. Cunha
Advisor: Otto Carlos M. B. Duarte
Grupo de Teleinformática e Automação - PEE/COPPE - DEL/POLI
Universidade Federal do Rio de Janeiro - Rio de Janeiro, Brazil
doc@gta.ufrj.br, otto@gta.ufrj.br

I. OVERVIEW

Field estimation is an important application of wireless sensor networks. This type of application deploys sensor nodes in a specific region to remotely sense space-temporally variable processes. The spatial and temporal frequencies of sampling directly impact the quality of the estimation. Therefore, there is a tradeoff between the frequency and the number of samples transmitted, which is related to the energy consumption of the network. The most common solution to reduce the number of samples transmitted to the sink is to identify areas where different nodes present similar readings and reduce the sampling spatial frequency by deactivating some of these nodes [1], [2]. Nevertheless, this approach is not efficient in regions of the field with sharp spatial variations, or borders. There is, however, another sampling dimension to regard: the temporal dimension. While many works focus the deactivation of nodes, little effort has been made to determine when active nodes must collect and transmit samples.

We propose to allow nodes to identify patterns in the behavior of the sensed processes and report only uncommon measures. This environment-aware behavior is similar to the response of living beings to the surrounding events.People and animals are continually receiving *stimuli*; however, it is impossible to handle consciously all these *stimuli*. The organisms develop the notion of periphery and center of attention [3]. While the periphery is handled in a sub-conscious manner, the center of attention is the event consciously treated. Generally, an event migrates from the periphery to the center of attention when it differs much from the periphery as a whole. The proposed bio-inspired scheme exploits specific features of the monitored processes in order to reduce the number of transmitted samples. Moreover, the scheme is fully distributed as each node identifies its own periphery (Section II). Thus, each node sends to the sink only the samples that differ from the usually sensed by the node. This scheme can save resources for nodes located in border regions as in smooth varying regions. This temporal technique is orthogonal to space-frequency-reduction techniques and both techniques can be used together to improve the system performance.

II. PROPOSED SCHEME AND PRELIMINARY RESULTS

The proposed scheme is based on the construction of a "periphery of attention" by each sensor node. It is necessary that the monitored process present a expected time behavior. Indeed, many physical processes present regular behaviors, sometimes with well-defined cycles. Temperature, for instance, generally has a similar behavior for consecutive days. Relating to the idea of periphery, a temperature curve that is similar to the average of previous days could remain in the periphery. Otherwise, a curve differing above a certain threshold must be explicitly taken into account. The proposed scheme aims at conserving resources not sending information when the expected behavior occurs. On the other hand, the accuracy of the estimation is preserved sending information for singular situations. Therefore, the sink assumes the behavior is like the one expected if no refining sample is received. Furthermore, refining samples are included in the reconstruction of the process, keeping low the estimation error.

The periphery construction and refining sample generation can take different forms. The first step is to determine the periodicity of the regular behavior. One option is to collect samples for a relatively long time and calculate the autocorrelation of the vector of samples. Based on the autocorrelation, the period of the process can be obtained. Once the period of the regular behavior is identified, the node can begin defining its periphery, or expected behavior. We focus the daily periodicity of the temperature. Hence, the sensor nodes must identify a daily-expected behavior.

It is worth to note that each sensor node builds and analyzes its periphery on its own time basis, which dispenses synchronization between sensor nodes. The timing information for the reconstruction at the sink is relative to the time basis of the specific sensor node. Moreover, the sensor periodically sends expected behavior updates but at a frequency low enough to achieve some gain. Fig. 1 shows the daily procedure, where DB_j is the vector with the expected behavior during day j, D_i the vector with the measurements of day i, $last\_update$ is the vector with the last expected behavior sent to the sink, and $X(k)$ is the k-th element of vector X. The factor α is the weight of the last period obtained in the expected behavior.

```
DB i = α D i + (1−α) DB i−1
If update time
    last_update = DB i
    Send last_update
Else
    For all k samples in D i do
        If |D i(k) − last_update(k)| > |last_update(k)| * configured_error
            Send D i(k)
```

Fig. 1. Daily procedure for each sensor node.

The efficacy of the proposed scheme is evaluated by simulation considering the reduction in the total number of samples sent to the sink. Three main variables are considered: the update frequency, the α parameter, and the maximum error without generating refining samples. In all simulations, the update frequency is one expected behavior vector sent for each $Update$ days. The tolerated sample error is equal to the parameter $maximum\_error$ times the expected behavior of the specific hour. The data is generated based on observation of real-site-collected data. The temperature in consecutive days presents some relation in its average and in the difference between the lower and the higher temperature of the day. It is worth noting that the goal of the generated data model is not to create real temperature curves, but to create curves that exhibit the same main behavior of temperature curves. For a better understanding, the relation between the temperature behavior in consecutive days is modeled by the parameter σ. Greater relations between the temperatures of consecutive days are achieved for smaller values of σ. The proposed scheme is applied to this data set and the fraction of the total samples that must be actually sent to the sink is obtained. The smaller this fraction is, the better the proposed scheme efficacy. All the results shown have a confidence interval smaller than 3% of the average value for a confidence level of 99%.

The first analysis concerns the impact of the uncertainty of the sensed process over the fraction of samples sent. Fig. 2 shows the fraction of the samples sent when σ varies for a tolerated error of 5%.

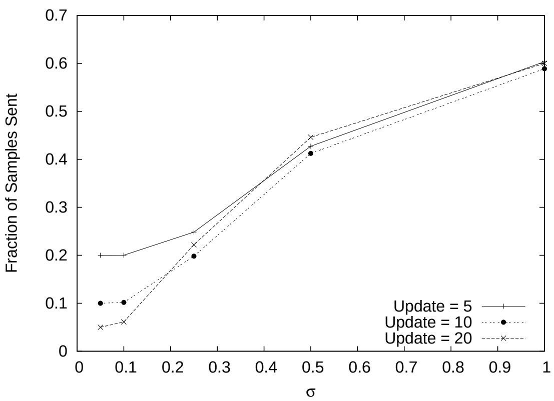

Fig. 2. Samples sent as a function of σ with $\alpha = 0.25$.

Preliminary results show that small values of the tolerated error severely reduces the influence of the α and $Update$ parameters. As the error restriction is relaxed, these parameters show more significant influence in the results [4]. Analyzing Fig. 2 it is possible to observe that for processes with smaller variation (low σ) the fraction of samples sent tends to be $\frac{1}{Update}$. It is worth noting that the smaller update frequency used, $Update = 20$, sends less samples when the process does not vary much, but has a worse performance when the variation of the process increases (higher σ). Under the conditions showed in Fig. 2, the proposed scheme reduces around 40% the number of samples sent for σ = 1. Moreover, for higher values of σ, the intermediate $Update$ value, 10, presents better results. This suggests that this intermediate value of $Update$ achieves a better tradeoff between the fixed part of the total samples sent ($\frac{1}{Update}$) and the samples sent due to the difference from the expected behavior.

Fig. 3 shows the results as a function of the tolerated error for $Update$ = 10 and for an intermediate value for α. As

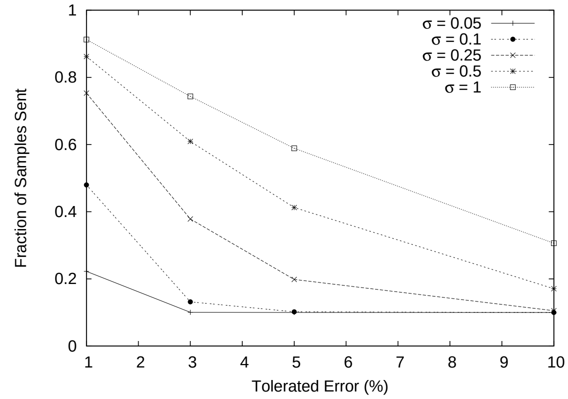

Fig. 3. Samples sent as a function of the tolerated error for $\alpha = 0.25$.

the tolerated error decreases, the variations of σ have worse impact on the scheme performance. For high tolerated errors, small variations of σ are negligible. It is worth to note the limit $\frac{1}{Update}$ to which the fraction of samples tends when the tolerated error grows or σ is reduced (more regular processes). Moreover, even for less regular processes, σ = 1, it is possible to reduce between 10 and 20% the number of samples sent for tolerated errors as low as 1 and 3%.

III. Discussion and Future Directions

The proposed scheme exploits specific features of the sensed physical process, identifying a regular behavior of the process, which will define the periphery of attention of the sensor node. Peripheral events, or events similar to the periphery, are not reported to the sink. Even the nodes placed at borders regions, which do not take advantage of the techniques based on spatial density reduction, benefit of the proposed scheme. This is an important feature that can increase the lifetime of the sensor network. Moreover, the proposed scheme can be used together with these conventional techniques to obtain greater gains. Preliminary results show that for very regular processes the scheme can reduce upto 90% the total amount of samples sent in the network. For less regular processes the proposed scheme can still reduce between 10 and 20% the amount of samples sent with small reconstruction errors. As future works we intend to analyze the effects of packet losses in the network on the information reconstruction error and to evaluate the scheme with real-site data.

Acknowledgements

This work has been supported by CNPq, CAPES, FAPERJ, RNP, FINEP, and FUNTTEL.

References

[1] R. Willett, A. Martin, and R. Nowak, "Backcasting: adaptive sampling for sensor networks," in *Information Processing In Sensor Networks - IPSN'04*, apr 2004, pp. 124 – 133.

[2] M. Rahimi, R. Pon, W. J. Kaiser, G. S. Sukhatme, D. Estrin, and M. Sirivastava, "Adaptive sampling for environmental robotics," in *IEEE International Conference on Robotics & Automation*, apr 2004, pp. 3537–3544.

[3] M. Weiser and J. S. Brown, "The coming age of calm technolgy," in *Beyond calculation: the next fifty years*. Copernicus, 1997, pp. 75 – 85.

[4] D. de O. Cunha, R. P. Laufer, I. M. Moraes, M. D. D. Bicudo, P. B. Velloso, and O. C. M. B. Duarte, "Bio-inspired field estimation with wireless sensor networks," Grupo de Teleinformática e Automação, http://www.gta.ufrj.br/~doc/TechBio.pdf, Tech. Rep., Jan. 2005.