Towards Measurement-Based Self-Management of Performance Properties in Wireless Sensor Networks

Adam Bachorek, Bagavathiannan Palanisamy, and Jens B. Schmitt
disco | Distributed Computer Systems Lab
University of Kaiserslautern, Germany
{bachorek,palanisamy,jschmitt}@cs.uni-kl.de

ABSTRACT

Ensuring the aspired outcome quality of network-based applications has ever implied an appropriate prognosis of the performance behavior the underlying communication structures will exhibit prior to potential optimization steps. However, the reliable forecast of correlative metrics is still a challenging task, especially in terms of wireless network systems that reject centralized or manual administration. In the context of our prospective self-management concept for autonomous analysis and control of volatile performance characteristics in *Wireless Sensor Networks (WSNs)*, we relate a common network property known as packet transfer delay to various adjustable parameters peculiar to such radio network technologies. Our hands-on experiments reveal the reproducible influence of packet size, backoff period, and number of active neighbor nodes on the medium access procedure and involved performance indicators. By means of a closed formulation of permutable weighted drivers, we investigate the average-case predictability of inter-node end-to-end delays for arbitrary configurations of given network parameters. We validate our prediction method against basic multi-hop networking scenarios while verifying its practicability on a typical resource-scarce WSN platform. Leveraging measurement-driven inspection and conditional modeling of network attributes based on regression analysis, yield field test results substantiate the high precision of our approach to the estimation of performance-related WSN properties as the basis for application-aware performance optimization subject to projected complements.

Categories and Subject Descriptors

B.4.4 [**Input/Output and Data Communications**]: Performance Analysis and Design Aids—*Verification*

Keywords

Delay; Measurements; Performance; Prediction; Regression Analysis; Self-Management; Wireless Sensor Networks

PE-WASUN'13, November 3–8, 2013, Barcelona, Spain.

http://dx.doi.org/10.1145/2507248.2507271.

1. INTRODUCTION

New-age advances in the development of micro-electro-mechanical systems and wireless communication technologies have facilitated inexpensive network-enabled sensor devices featuring miniaturized shapes and energy-aware operation modes. Ever since, WSNs have evolved into a seminal platform for tackling low-maintenance application domains where efficient capturing, collection, and exchange of physical phenomena data plays a constitutive role, a.o., environment monitoring, medical diagnostics and home automation [1]. However, future considerations of large-scale WSNs envision administration-free deployments of even smaller, poly-functional, and heterogeneous devices that leverage autonomous interaction patterns while serving as ambient information sources for unattended actuation as the eventual step towards pervasive intelligence. In this context, autonomic computing and networking have witnessed great attention from academia as well as the industry in regard to the self-management of distributed network systems of ever-growing size and complexity [2, 3]. The implicated need for a self-governed handling of adjustable technical parameters and implicitly controllable functional properties as encountered in actual network deployments necessitates the reliable and precise estimation of performance metrics without human supervision. This constitutes a critical feature of appliances in inaccessible and/or highly dynamic environments, e.g., traffic control in over-crowded urban areas. Yet, it is obviously challenging to bring these stringent requirements in line with present-day equipment that essentially requires manual intervention to fulfill any task at all due to lacking automation mechanisms [4]. Moreover, even though existing approaches are technically capable of determining such network properties under certain assumptions, they mostly abstract away from primal aspects that inherently have an essential impact on temporal network performance dynamics. Also, they often require auxiliary resources to be deployed along with the actual components to cope with the implied issues of increased complexity. Therefore, we strive for a self-contained performance management concept based on nodal abilities supporting collateral measurements and lightweight prediction methods, the distributed coordination of which noninvasively integrates into any performance-constrained WSN application [5]. In this spirit, Section II introduces the basics of our methodology followed by the analysis of our evaluation approach to identify network attribute correlations in Section III. Our initially proposed prediction technique is discussed and its quality is validated in Section IV before conclusions are given in Section V.

2. PERFORMANCE SELF-MANAGEMENT

The need for application-aware and efficient management of available and controllable resources in question suggests to introduce a comprehensive self-ability concept we recognize as ***Intrinsic/Extrinsic Performance Self-Management (I/EPSM)***. Whereas during a transitional period it is required to provide involved network entities with extrinsic feedback and functional support, the ultimate fusion of intrinsically applied methods in a collaborative manner shall render such networks operable on a fully autonomic basis. In this context, we identify a set of sub-routines to be implemented on I/EPSM-enabled nodes such as measurement, modeling, assessment, and adaptation. These tasks are meant to iterate in a modular loop as defined in the **I/EPSM Task Cycle** and embody two fundamental principles that saturate two recurring phases as detailed below.

- **Inspection Phase:** Measurement of parameter values and property-related statistics as indicated by application requirements (*Control Principle*) and modeling, i.e., identification and instantiation of the best-fitting performance model (*Analysis Principle*).
- **Tuning Phase:** Assessment of current performance quality against application requirements (*Analysis Principle*) and adaptation of parameter value constellations to adhere to imposed network property demands at corresponding entities (*Control Principle*).

Based on resource availability and mode of operation, the flexible task cycle procedure shall include time-based, event-triggered, and entity-controlled execution which also depends on the desired performance sustainment quality. Due to the higher computational/communicative burden imposed on nodes involved in task accomplishment, a selective load-balancing mechanisms shall further be applied for energy saving purposes. The choice of the model type and prediction technique shall also be subject to setup conditions.

For the sake of referable terminology, we further distinguish two variable network metric types and the viewpoint on them. Whereas **network parameters** include explicitly or implicitly adjustable settings when acting as drivers for network property control, **network properties**, denote attributes merely obtainable by means of measurements and/or computations based on statistical or analytical inference. **Centrical parameters** refer to intra-node variables that can be determined and tuned without any interaction with other entities (e.g., power amplifier level, packet size) while **peripheral parameters** encompass inter-node system variables usually accessible and modifiable only by collaboration (e.g., number of active nodes in cluster, number of neighbor connections). Correspondingly, **centrical properties** pertain to measurable or calculable attributes that relate to the status of a single network node (e.g., medium access delay, traffic arrival rate) in contrast to **peripheral properties** which are concerned with attributes quantifying network-wide conditions (e.g., end-to-end delay, network backlog). In some cases, network parameters might also need to be measured or deduced from captured phenomena (e.g., node position, distance to neighbors). However, as long as they remain uncontrollable, they are not considered drivers for network properties but at most predictor variables for network property determination based on a predefined abstract intuition of real circumstances, generally referred to as **performance model**. For more detailed explanations see [5].

3. SELF-MEASUREMENT EXPERIMENTS

As part of this work, we focus on a representative selection of peripheral and centrical parameters to investigate their relation to the ***End-to-End Delay (E2ED)*** as a fundamental and descriptive network property. As per convention, the E2ED in packet-switched communication networks encompasses the duration between the start of the sending process at the source node and the end of reception at the ultimate destination node [6]. In between, several delay components known as processing, queuing, transmission, and propagation delay can be identified dividing the E2ED into per-hop time segments. Without loss of generality, we cumulatively redefine that division into *Packet Processing Delay (PPD)*, *Medium Access Delay (MAD)*, and *Packet Transmission Delay (PTD)* all of which summing up to the *Packet Sending Delay (PSD)* for practical examination reasons. This allows to consider the otherwise peripheral notion of E2ED also from a centrical view of a single network entity.

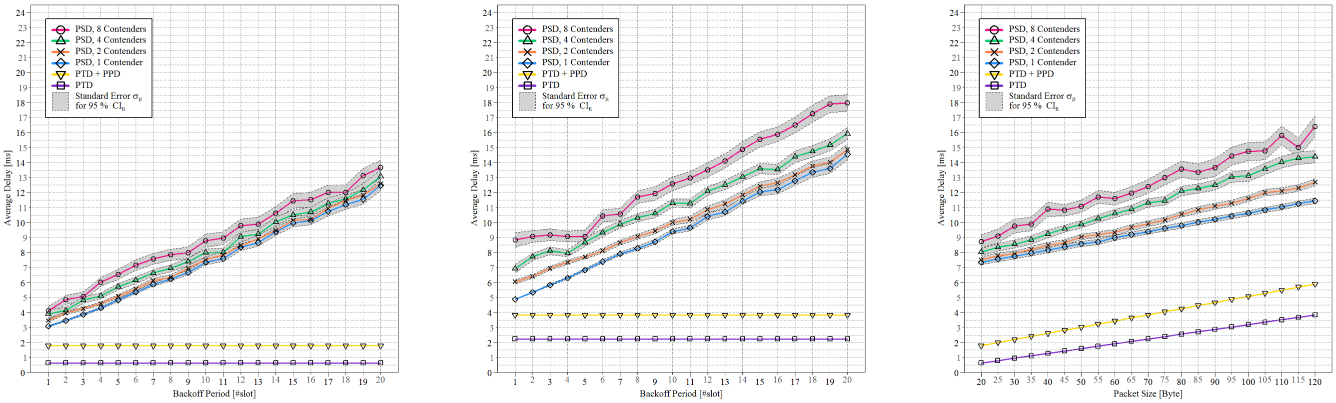
General Node Configuration		Experiment Run Configuration	
Parameter	**Setting**	**Parameter**	**Setting**
B_{min}	10 (320 μs)	B_P	[1,20] by 1
PAL	31 (0 dBm)	P_S [Byte]	[20,120] by 5
Channel	26 (2.48 GHz)	N_C	{1,2,4,8}
ACKs	0 (disabled)	Samples [packets]	1000
Duty Cycle	1 (always on)	Event rate [1/s]	10

Figure 1: General and experimental configuration settings

Considering major network parameters common to wireless network technologies, we opt for the ***Number of Contenders (N_C)*** as a peripheral and ***Packet Size (P_S)*** and ***Backoff Period (B_P)*** as two centrical parameters considered the objects of investigation due to their anticipated influence on E2ED. In this regard, as part of the control principle in the I/EPSM inspection-phase, nodes participating in network activity are supposed to capture parameter values that are considered important regarding their impact on network properties. To this end, self-measurable performance-related data is of particular interest that can be obtained during regular network operation, i.e., collected or derived when sending, receiving or processing data in the context of orderly application execution (e.g., medium access delay, traffic input) without auxiliary network communication.

The WSN node platform used throughout the experiments is the popular MICAz mote [7] running TinyOS 2.1.2. Network nodes are randomly arranged in a star-shaped topology with a radius of $\approx 1\,m$ allowing for line-of-sight communication (see Fig. 1 (left) for further configuration settings). One mote being attached to a gateway board is in charge of test run coordination and relaying of experiment data. All other motes are endued with experiment run-specific execution settings and act as parts of a sample homogeneous WSN providing the Linux-based evaluation back-end system with feedback on gathered results. Throughout experimental runs, the selected network parameters are systematically permuted over the specified value range (see Fig. 1 (right)).

Regarding all influencing factors separately, Fig. 2 includes a subset of all experiment results for representative parameter constellations. Fig. 2 (a) reveals the developing of mean delays against increasing B_P also in view of 4 con-



(a) Packet Size $P_S = 20$ Byte (b) Packet Size $P_S = 70$ Byte (c) Backoff Period $B_P = 10$

Figure 2: Influence of backoff period, packet size, and number of contenders on packet transmission/processing/sending delay

tender configurations while P_S is minimized to 20 Byte. The PSD increases gradually with all settings of B_P. Although it also shows a virtual dependency on N_C, the confidence intervals for 0.05 significance level overlap especially for higher B_P values, signifying their statistical indistinguishability. This is due to the increasing standard deviation of the samples lying between 776 μs and 5729 μs for no neighbors and between 1324 μs and 6044 μs for 7 neighbors, respectively, as far as backoff periods between 1 and 20 slots are concerned. This also emphasizes the high probabilistic impact, albeit ascertainable as almost linear, of the B_P parameter on overall delay. A more convincing statistical relevance of the N_C becomes evident not before the P_S is increased as exemplified for $P_S = 70$ Byte in Fig. 2 (b), where, additionally, a slightly wider spreading of the confidence intervals and also a definite elevation of the PSD can be observed. The latter trend, which obviously suggest a sensitivity of the PSD to the packet length, is captured in Fig. 2 (c). In terms of the default B_P setting of 10 slots and incrementation of the P_S by 5 Byte over the available range of values, the average PSD increases apparently in a linear manner. Again, a severe statistical relevance of the number of contenders for the mean PSD, at least for the first two proximate instances of the traversed value range, cannot be assumed before a packet size of 50 Byte is reached. Throughout all experimental runs, we observe the devolution of the PPD as well. It turns out that the PPD is exclusively dependent on P_S as expected. However, in order to shed light on its portion of the overall time readings in isolation, we conducted disjoint measurement runs freezing the random number to $r = 100$ for $B_P = 10$ within the backoff process, i.e., observing a constant MAD of 3520 μs for an initial backoff value of $z = 31$ as applied on first transmission attempt in case of only one sender. In this way, we numerically divided the PSD into its individual components, i.e., PPD, MAD, PTD, the relationship of which becomes evident in Fig. 2 (a)-(c). Whereas the PTD can be assumed to be consistently dependent on the transceiver transmission speed, the measurements of the PPD yield a median standard deviation of about 11 μs proving its constancy in view of any parameter constellation. In summary, we can act on the assumption of a linearly proportional interconnection between P_S and not only the PTD but also the PPD. More details on the measurement campaigns including methodology, data analysis and further experiment results, e.g., parameter influence on the packet loss rate can be found in [5].

4. PREDICTING NETWORK PROPERTIES

For the purpose of network property estimation, determining an appropriate performance model based on a lightweight technique in addition to its concrete instantiation with available measurement results constitutes the key enabling procedure as per analysis principle of the I/EPSM inspection-phase. In view of our experiment results, (multiple) linear regression qualifies as most suitable for finding the model instance of choice that relates PSD, considered as predicted response y_{PSD}, to the nonstochastic network parameters x from set $\Phi = \{B_P, P_S, N_C\}$, that act as predicting explanatory variables according to Eq.(1) with $m \in \{1, 2, 3\}$.

$$y_{PSD} = r_0 + \left(\sum_{i=1}^{m} r_i \cdot x_i\right) + \epsilon \ , x_i \in \Phi \setminus \bigcup_{k=1}^{i} \{x_{k-1}\} \quad (1)$$

where $x_0 = \varnothing$, r_i are regression coefficients, and ϵ denotes the residual error term. To identify the best fitting model variant Ψ_j for any parameter constellation, we include all previously measured value observations. First, we derive a simple linear model to relate PSD to any individual parameter from Φ and subsequently consider it in view of all parameter combinations by virtue of a multiple linear model each. In order to find the mean predicted response with as little variability as possible, we apply the least-squares criterion procedure minimizing the sum of squared errors in view of the observed mean response for all observations. Expectedly, every network parameter shows a distinct influence on the PSD to a certain extent as can be told by the coefficient of determination Ω [8]. Model variant Ψ_1 based on just B_P appears to have highest single explanatory rate of about 62.53 % on PSD in contrast to N_C that seems to play a minuscule role with only 9 % followed by P_S with 26.03 %. However, the permuted combination of all three parameters turns out to increase Ω in a cumulative manner so that Ψ_4 with $PSD \sim B_P + P_S$ exhibits 88.56 % whereas Ψ_5 with $PSD \sim B_P + N_C$ relies on an explanatory value of 71.53 %. Accordingly, Ψ_7 is assumed to most precisely explain the PSD on the average by 97.56 %, i.e., it can predict its mean value more reliably than all the other models do. For model testing purposes, we examine the *confidence intervals (CIs)* for the regression coefficients of each model to reveal the statistical significance of the regression with respect to the response variation. Since all corresponding CIs are non-zero for even a 0.05 significance level in terms of any single predictor variable tested against PSD as in isolation so

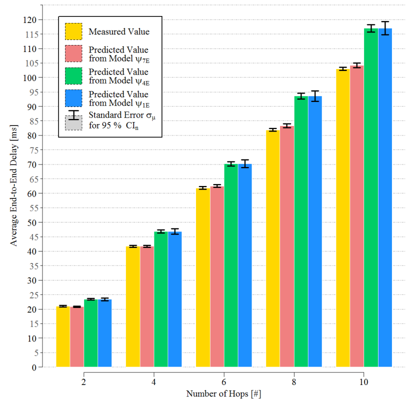


Figure 3: Comparison of model prediction quality for E2ED in multi-hop scenarios for $\mathrm{B}_P = 10$, $\mathrm{P}_S = 70$ Byte, and $\mathrm{N}_C = 1$

in combination, they are considered statistically significant. Also, all parameter combinations pass the zero correlation test excluding multicollinearity effects of the predictor variables that might otherwise reduce the statistical accuracy of the applied regression method. At last, a clear additive parameter cohesion throughout all model alternatives is evident indicating their pertinence for PSD prediction.

In terms of verifying the viability of the analysis principle in the I/EPSM inspection phase, all mentioned procedures are implemented on MICAz motes. Results confirm that existing resource limitations do not hamper the precision of the proposed modeling and prediction technique. A moving average function implemented to consider only the last captured value into a continuously updated mean PSD for all setting constellations from Φ, allowed for storing just a subset of values of $(\|\Phi\| + 2) \cdot 4\, Byte$ for each compound observation with μs precision so as to conduct the sequence of calculations needed for regression model instantiation [8].

For validation purposes, we investigate hop-variant network topologies of tandem-like shape while measuring E2ED between the source and destination node for default parameter settings as per Fig. 3. Along with estimating PSD by statistically most accurate model instances selected based on Ω, our actual models are enhanced (Ψ_{jE}) by including PPD, which is assumed to additionally incur at every receiving node r, and by postulating an additive coherence between the E2ED and the number of hops h according to

$$y_{PSD,E} = h \cdot \left(y_{PSD} + \frac{1}{2} PPD_r \right) \qquad (2)$$

Indeed, yield results show an almost perfect forecast of the E2ED for up to 10 hops using enhanced model Ψ_{7E} as indicated by overlapping CIs for 95 % confidence level. Other model candidates, in turn, seem to mismatch the E2ED to a large extent. Interestingly, the prediction based on Ψ_{1E} is equal to that based on Ψ_{4E}, albeit, with a lower certainty, even though P_S is unconsidered in Ψ_{1E} and $\Omega_{\Psi_1} < \Omega_{\Psi_4}$. Nevertheless, a general trend towards overestimation of the E2ED by any model can be observed that seems to sums up with the increase in number of hops. As further node-local measurements reveal, the cause lies in the fact that PPD during the sending process differs from PPD during reception by about the half. These observations point out the need for diversity in consulting statistical measures when comparing prediction quality as well as completeness when collecting or concluding details about the constituent and ambient parts of network entities when adhering to the I/EPSM control principle. Summing up, we verified the investigated modeling technique to be what a WSN node is able to apply for meaningful performance model derivation in virtue of its given capabilities. This is in line with our original aim to keep complexity as rational as possible rendering the methodology what is considered lightweight and eligible for I/EPSM. For more details on the proposed methodology, results and implementation see [5].

5. CONCLUSION AND FUTURE WORK

This work introduced the basics of a measurement-driven approach to self-management of network attributes based on regression analysis for performance property value prediction. In this context, we studied the influence of selected network parameters on a fundamental network property. The results of node-local measurements confirmed their linear relationships under isolated conditions. We derived several regression models based on empirical data enabling property forecast for arbitrary parametrical settings beside reliable model validation even on resource-scarce WSN nodes. The practical verification of yield speculations on performance behavior in basic multi-hop networks revealed the high precision of an augmented performance model in view of end-to-end delay prediction. After all, we have demonstrated viable methods for the aspired measurement and modeling tasks based on principles as part of our comprehensive I/EPSM approach that strives for lightweight procedures with as little demands on available resources as possible by using network inherent information drawn from regular network node operation and communication behavior. In future works, we intend to explore further WSN peculiarities as influence factors, test the quality of more sophisticated prediction and validation methods for other network properties and include simulations to cover more involved topologies and parameter interactions in larger-scale networking scenarios.

6. REFERENCES

[1] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, “A Survey on Wireless Multimedia Sensor Networks,” *Computer Networks*, vol. 51, no. 4, pp. 921–960, 2007.

[2] J. O. Kephart, “Research Challenges of Autonomic Computing,” in *IEEE/ACM International Conference on Software Engineering (ICSE)*, pp. 15–22, May 2005.

[3] N. Samaan and A. Karmouch, “Towards Autonomic Network Management: an Analysis of Current and Future Research Directions,” *IEEE Communications Surveys & Tutorials*, vol. 11, pp. 22–36, August 2009.

[4] Z. Movahedi, M. Ayari, R. Langar, and G. Pujolle, “A Survey of Autonomic Network Architectures and Evaluation Criteria,” *IEEE Communications Surveys and Tutorials*, vol. 14, pp. 464–490, June 2012.

[5] A. Bachorek, B. Palanisamy, and J. B. Schmitt, “Measurement and Prediction of Centrical/Peripheral Network Properties based on Regression Analysis,” Tech. Report arXiv:1308.3855 [cs.NI], August 2013.

[6] L. Peterson and B. Davie, *Computer Networks: A Systems Approach*. M. Kaufmann, 5th ed., 2011.

[7] Crossbow Technology Inc., *”Mote Processor Radio (MPR) Platforms and Mote Interface Boards (MIB), Rev. B”*. Manual, June 2006.

[8] R. K. Jain, *The Art of Computer Systems Performance Analysis*. New York: J. Wiley & Sons, 1st ed., 1991.