Energy-Efficiency of Concast Communication in Wireless Sensor Networks

Joachim Wilke
Institute of Telematics
Karlsruhe Institute of Technology, Germany
Email: joachim.wilke@kit.edu

Christian Haas
Institute of Telematics
Karlsruhe Institute of Technology, Germany
Email: christian.haas@kit.edu

***Abstract*—In monitoring scenarios, Wireless Sensor Networks commonly transmit measurements from a large number of sensor nodes to a central data sink. This communication pattern is known as concast. Different approaches have been proposed to improve the energy-efficiency of concast and thus the lifetime of the WSN. However, energy-efficiency evaluations that are close to reality are missing. This paper systematically analyzes the influence of aggregation strategies, tree topologies, and different MAC protocols on the energy-efficiency of concast communication. We implement a sample concast application and analyze it using the AVRORA+ simulator to gain realistic evaluation results. Our results disproof some common assumptions. We show that aggregation improves energy-efficiency only in a few cases and can even degrade it. Instead, MAC protocol and parametrization have a higher impact on energy-efficiency.**

I. INTRODUCTION

Wireless Sensor Networks (WSNs) are proposed for a large variety of monitoring applications. In these scenarios, a large number of nodes periodically take measurements of environmental data and transmit them to a central data sink. This communication scheme is commonly known as *concast* communication. Such networks have to operate for months to years with a limited energy budget. Further, communication range of sensor nodes requires multi-hop communication. For concast, this often results in a tree topology with the data sink as root node. Here, nodes near the data sink have to forward larger amounts of measurements than other nodes and might run out of energy, first.

Different approaches have been proposed to improve the energy-efficiency of concast. First, different *tree topologies* are suggested, because it is argued that the topology is important for energy-efficiency. Second, *aggregation* is applied. Nodes pre-process measurements before forwarding, e.g., by data or packet aggregation. It is assumed that aggregation dramatically improves energy-efficiency. However, previous evaluation results are either only theoretical estimations or ignore possible cross-layer effects.

Our contribution is to provide *realistic* evaluation results that disproof some common beliefs. We further show, what parameters actually influence energy-efficiency and are thus important to design energy-efficient concast applications.

II. RELATED WORK

Most popular approaches that improve energy-efficiency of concast communication identify two important factors: tree topology and aggregation strategy. Tree topologies can be constructed in various ways. Simple flooding can be used [1] to disseminate a query and use the resulting tree for concast. Various greedy approaches [2], [3] claim to be more energy-efficient than the flooding-based approach. Additionally, aggregation is suggested to increase energy-efficiency. Most approaches fit into two categories:

Loss-less, or *packet* aggregation: The measurements of incoming data packets are stored. At some point of time, all stored measurements are forwarded in one large packet. Different algorithms to determine when to forward data have been proposed [4]. This is an energy–latency trade-off. The longer data packets are stored, the more data packets can be aggregated, but the higher latency is. Packet aggregation mainly reduces the number of transmissions.

Lossy, or *data* aggregation: The measurements of incoming data packets are aggregated using a pre-defined aggregation function. A large variety of functions is possible, heavily depending on the application scenario [5]. Most publications suggest simple functions like average or minimum/maximum to keep computation overhead low [6]. In dense networks, dropping redundant data packets is also possible [7].

In real-world deployments, an energy-efficient MAC protocol should be used to match network lifetime requirements. Various duty-cycling MAC protocols have been proposed in the last years. Energy is saved by switching the radio to sleep-mode as often and as long as possible. Duty-cycling MAC protocols can be divided into asynchronous and synchronous protocols. The simplest synchronous approach is to switch on all node's radios at the same time, communicate and put all radios back to sleep, repeating that in periodic cycles. This is a simple TDMA-based MAC which assumes synchronized nodes. Sensor-MAC [8] was proposed based on the idea of TDMA, but includes its own node synchronization mechanisms. On the other hand, asynchronous protocols do not require synchronization, e.g., TinyOS's Low Power Listening [9]. The drawback of asynchronous protocols is an increased overhead for each transmission.

Measurements in real-world deployments provides highest possible accuarcy regarding energy-efficiency evaluating. Because of scalability, experiments using simulators are often preferred. A large number of simulators exists. MiXiM [10] is an extension to the OMNeT++ framework for wireless mobile networks. However, OMNeT++ cannot simulate native WSN applications, protocols have to be reimplemented for analysis. PowerTOSSIM [11] is designed for TinyOS applications but needs code to be instrumented and recompiled. AVRORA [12] is one of the few tools that provides more realistic data on energy consumption in WSNs. It emulates state changes at

978-1-4577-1720-8/13/$31.00 ©2013 IEEE

cycle level, combined with a detailed energy model of all node components. Unmodified application code can be used as AVRORA emulates the sensor node's hardware.

Most publications suggesting tree topology construction algorithms or aggregation strategies use theoretical models and simulations to prove the energy-efficiency of their approach. However, protocols are often evaluated in isolation, ignoring possible cross-layer effects, e.g., by a duty-cycling MAC protocol. A duty-cycling MAC is a requirement in a real network deployment as it prolongs network lifetime significantly. Furthermore, many simulators used in previous works are not suitable for realistic energy-efficiency evaluations. It is therefore of interest, how the suggested approaches perform in a more realistic evaluation.

III. IMPLEMENTATION AND EVALUATION SCENARIO

We implement a sample concast application in TinyOS 2.1.1 to analyze, which parameters influence energy consumption. The application is kept rather generic to fit into most concast scenarios. Time is segmented in rounds of length t_P. In every round, all nodes sample one measurement (6 byte) and transmit it towards the data sink.

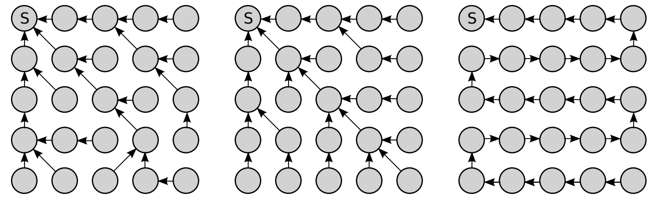
We consider two different kinds of node deployments. 25 nodes are placed in a grid deployment on an area of 35x35m (GRID) in intervals of 8.5m. The radio range is 15 meter, e.g., a node has between 3 and 8 neighbors. For comparison, random deployments, where nodes form a random but connected graph, are also considered (RANDOM).

Three different ways to setup the tree topology are implemented (see Figure 1): Flooding based (FLOOD), shortest path depending on hop count (SHORTEST), and a linear structure (LINEAR). The FLOOD topology is a simple approach taken from related work. It is randomly created at application startup. The SHORTEST topology is fixed and is a best-case regarding hop count to sink. Contrary, the LINEAR topology represents a worst-case. For now, it is sufficient, to analyze these simple, static topologies. If aggregation does not improve energy-efficiency in these idealized scenarios, it will not improve in more complex, dynamic setups.

This setup is sufficient for our goals as it includes a network sink, nodes that aggregate incoming measurements from a different number of sources and nodes that only contribute measurements but do not aggregate itself. At this point, evaluating a larger network would not bring additional insight.

A. Aggregation Strategies

We analyze the following basic aggregation strategies, derived from the related work discussion: No Aggregation (NA), immediately forwards incoming measurements to the data sink without modification. Packet Aggregation (PA) temporally stores incoming measurements. Once each round, at the time the node samples its own measurement, all stored measurements are forwarded as independent measurements but combined into a single packet. The number of transmissions is reduced but *not* the data volume. In contrast to PA, Data Aggregation (DA) does not forward the measurements as independent values, but aggregates them using an aggregation function. We use the minimum aggregation function, e.g., the size of an aggregated measurements is identically to the size of a single measurement. Data aggregation *reduces* communication volume *and* the number of transmissions.



(a) FLOOD (example) (b) SHORTEST (c) LINEAR

Fig. 1. Three ways to setup the tree topology (S = data sink).

Aggregation strategies may increase latency. However, latency evaluation is out of scope of this work. We store measurements rather long, i.e., up to t_P, to get the most possible effect from aggregation to evaluate its influence on energy-efficiency.

B. MAC Protocols

Based on related work, four MAC protocols can be enabled in our sample application to study their influence on energy-efficiency: First, TinyOS Low Power Listening (LPL) is the default duty-cycling MAC protocol provided by TinyOS. Nodes wake up periodically for clear channel assessments (CCA) and transmission of data packets. The time a node sleeps between two CCAs is t_{LPL} in seconds. Higher values of t_{LPL} implicate a lower duty cycle. Second, Sensor-MAC (SMAC) synchronizes wake-up schedules of neighboring nodes. SMAC can be configured using a duty cycle parameter t_{DC} in percent. The lower the duty cycle, the longer the sleep phase. For comparison, the third protocol is a simple TDMA protocol that requires pre-synchronized nodes. Again, t_{DC} is used to configure the duty-cycle. Forth, IEEE 802.15.4, the default MAC for most node platforms which provides no duty cycling.

C. Expectations

We would like to state some a priori expectations based on experiences from related work. First, one would expect that DA improves energy-efficiency significantly compared to PA (less data volume) and NA (less transmissions). Second, one would assume that synchronized MAC protocols (e.g., SMAC or TDMA) are more energy-efficient in a periodic concast scenario compared to asynchronous protocols. Third, the SHORTEST topology is expected to be the best, the LINEAR topology to be the worst case regarding energy-efficiency.

IV. EVALUATION

To get results close to reality, we use AVRORA+, which is an improved version of the AVRORA [12] sensor network simulator that emulates sensor nodes using native application code. It has been cross-checked with reality using a testbed [13]. Its deviation from reality regarding energy consumption in a generic concast scenario is below 5% [14].

We are interested in comparing the energy-efficiency of the different aggregation strategies and topologies in periodic concast scenarios. We evaluate each combination of parameters using AVRORA+, the MICAz platform and 20 different simulation seeds. The most interesting results have been included into this Section.

Each simulation run is divided into rounds. The round R0 of length $t_I = 3t_P$ is used to construct the tree topology and

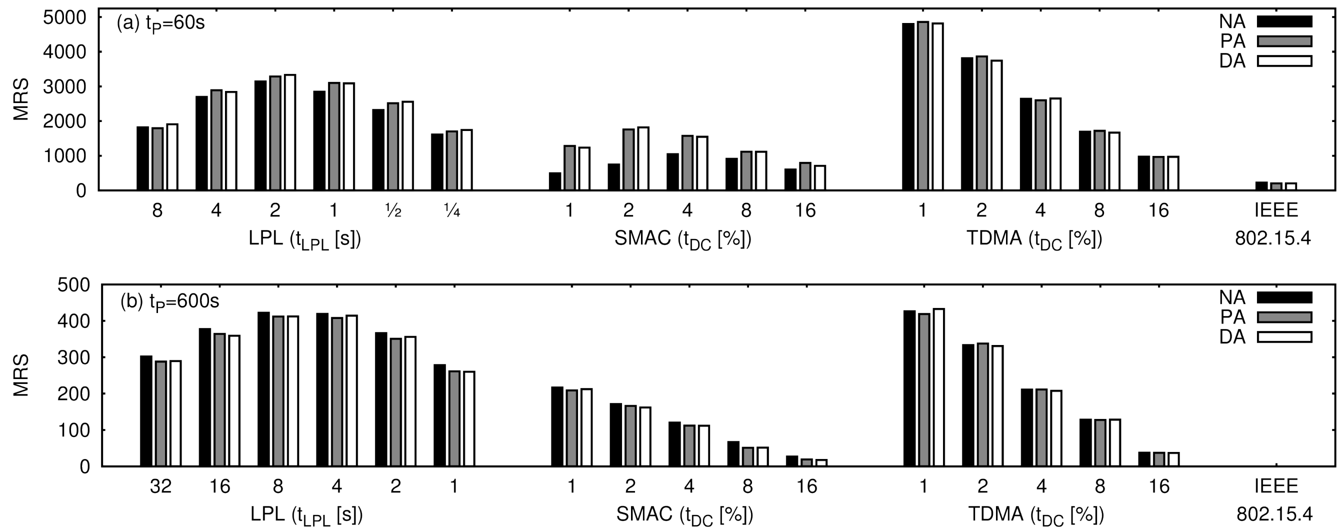


Fig. 2. MRS until depletion of energy resources using FLOOD topology and GRID deployment. The figures compare LPL, SMAC, TDMA and IEEE 802.15.4.

initialize the MAC protocol. Also, R0 provides time to start nodes randomly, to avoid timing-based effects. All following rounds $Rx, x \geq 1$ are of length t_P and are used to sample one measurement per node. Each node samples its measurement at a randomly selected but fixed time within the interval t_P. We present results using $t_P = 60s$ to represent a high-traffic scenario, and tp= $600s$ to represent a low-traffic scenario.

A. Rating Energy-Efficiency

AVRORA+ provides detailed statistics for energy consumption in the network. However, energy-efficiency is a metric that cannot be directly measured. We use the ability of AVRORA+ to assign all emulated sensor nodes a fixed amount of energy. If a node's total energy consumption exceeds that amount, AVRORA+ removes the node from the simulation.

In general, to improve energy-efficiency, the amount of energy spent for a certain value has to be reduced. This in turn means, if the value can be increased using the same amount of energy, energy-efficiency is also improved. We thus require metrics that express this value.

In concast scenarios, the number of distinct *measurements that are received at the data sink* (MRS) is a metric that expresses such a value. Referring to the definition of energy-efficiency, if the same amount of energy is provided at start-up and a parametrization achieves a higher MRS, this parametrization is more energy-efficient. For NA and PA, we just count distinct measurements that the data sink receives. For DA, we count the number of measurements the aggregated value is based on. In our evaluation, all nodes start with the same amount of energy (50 Joule).

B. Comparison of MAC Protocols and Aggregation Strategies

In Figure 2, each plot shows the MRS for different parametrizations and all three aggregation strategies, FLOOD topology, and GRID deployment. Figure 2a and Figure 2b show results for $t_P = 60s$ and $t_P = 600s$ respectively. For each MAC protocol, the duty cycle is decreased from left to right.

We first compare the results using different MAC protocols, ignoring the different aggregation strategies: It can clearly be seen that TDMA performs better than SMAC with respect to MRS. This is because TDMA can assume all nodes following the same wake-up schedule. With SMAC, in a multi-hop scenario, many nodes follow more than one schedule causing additional wake time. With TDMA, the whole network wakes up synchronously. Every node just follows that single schedule. This has a direct effect on the MRS metric: Even using the best possible MAC parametrization, MRS is reduced by about factor 1.5-2 using SMAC compared to TDMA. However, due to possible clock drift, at some point of time in a longer running network there would be no communication possible any more. If there is no external mechanism that re-synchronizes the nodes, TDMA would be no realistic option.

The result for IEEE 802.15.4 clearly show that a duty-cycling MAC layer is obligatory in almost any sensor network application. Results in both scenarios are very poor. In fact, using $t_P = 600s$, most nodes run out of energy before reaching concast round R1 and no measurements reach the data sink.

As results show, an interesting alternative to SMAC and TDMA are asynchronous MACs like LPL. LPL performs better than SMAC for all parametrizations. With $t_P = 600s$ it even performs better than TDMA. Another important fact is that results largely vary within each MAC protocol depending on its parametrization.

Looking at the previous expectations, LPL performs surprisingly better than SMAC. TDMA can provide an interesting alternative if node synchronization is provided by other means and needed by the application anyway.

Regarding the different aggregation strategies, the impact highly depends on which MAC protocol is used. For TDMA, there is almost no significant influence neither for the $t_P = 60s$ nor for $t_P = 600s$. The reason is that with TDMA, the node is active for a constant time (defined by t_{DC}). The duty cycle has major influence on the energy consumption and thus on MRS. If the node's radio is in transmit, receive or idle state during its active phase has only a minor influence, as the power draw of these three states is very similar.

Looking at the results using SMAC, aggregation can only improve results in the $t_P = 60s$ scenario. This is because

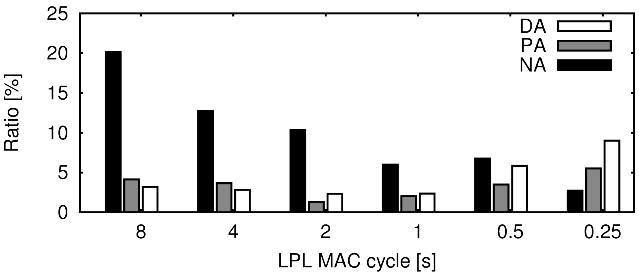


Fig. 3. Relative difference of MRS using LINEAR topology compared to FLOOD topology (LPL MAC, $t_P = 60s$).

aggregation reduces the number of independent transmissions. SMAC can only transmit one data packet per cycle. Using a low t_{DC} and NA, SMAC cannot forward incoming measurements fast enough. For $t_P = 600s$ this is no issue, so PA and DA do not differ. The length of data packets has no significant influence on the radio's energy consumption. More relevant is the time, the radio chip is awake (either being in idle, transmit or receive state).

For LPL, a lower duty cycle is not automatically increasing MRS. The reason is LPL's internal trade-off between the number of transmissions and t_{LPL}: Each transmission requires a preamble which is up to t_{LPL} long. The more transmissions, the better it is to shorten t_{LPL}. On the other hand, this means more CCA checks. As a consequence, there is an optimal value for t_{LPL}, depending on the number of transmissions per time interval. For the evaluated set of duty cycles this optimum is $t_{LPL} \approx 2s$ (at $t_P = 60s$) and $t_{LPL} \approx 4s$ (at $t_P = 600s$). As PA and DA reduce the number of transmissions and each transmission requires an expensive preamble, energy-efficiency improves. On the other hand, package loss can decrease MRS. Loosing an aggregated package has higher impact. Also, the influence of communication on the total energy consumption of the node decreases with increasing t_P. This is why aggregation slightly improves energy-efficiency for $t_P = 60s$, but slightly decreases energy-efficiency for $t_P = 600$.

Summarized, aggregation does not improve MRS if a synchronous MAC protocol with $t_P = 600s$ is used. In case of LPL, using PA can improve energy-efficiency, but only if t_P is rather short. If time synchronization is available or required anyway by the application, simple TDMA without any aggregation is suggested. In other cases where long term monitoring, e.g., using tp$\geq 600s$, is intended, LPL would be the best option. Although DA reduces data volume, it does not further improve MRS in any scenario, compared to the simpler PA. The results therefore definitely contradict expectations.

C. Influence of Node Deployment

To analyze influence of node placement, all experiments were also run using a random deployment. No significant tendency can be observed for any of the three MAC protocols. Results vary by $\pm 2\%$ for LPL and $\pm - 1\%$ for SMAC and TDMA, i.e., the deployment has no significant influence to MRS with regard to different MAC protocols.

D. Influence of Tree Topology Structure

Results using SMAC and TDMA do not significantly vary between the topologies, results using LPL are, however, quite interesting. We therefore show the ratio of MRS using LINEAR compared to the MRS using FLOOD tree topology structure in Figure 3. A positive ratio means a higher MRS using FLOOD. It can be seen that the FLOOD tree topology is always more energy-efficient than the LINEAR one. However, the differences vary depending on the aggregation strategy. The negative impact is larger for NA than for PA and DA and increases with lower duty cycles. This means in turn that aggregation has a higher impact using the LINEAR tree topology compared to the FLOOD tree topology.

Data packets have to pass much more hops in this topology. Using aggregation, only one transmission is required per node in each round. This is why aggregation has a larger impact using the LINEAR topology than a FLOOD topology. Using SMAC and TDMA no significant influence occurs. As before, the reason is that with SMAC and TDMA the node's wake-time only depends on the duty-cycle and not on the number and size of data packets.

E. Other Metrics for rating Energy-Efficiency

One could argue that the MRS metric could be misleading, as it shows the cumulative number of received measurements over the total network lifetime only. It does not reveal the temporal distribution of measurements. In this Section, we analyze this distribution.

First, we look at the temporal distribution by analyzing the MRS separated by *concast round*. Figure 4 shows selected results for the two parametrizations that were identified to benefit from aggregation. For clarity the results are shown in blocks of 5 concast rounds, i.e., the first bar shows how many measurements were received at the data sink from rounds R1 through R5. The maximum per bar is $5 \cdot 24 = 120$ measurements.

Figure 4a shows the results for LPL using most energy-efficient value for t_{LPL}. The difference between NA and both aggregation strategies is clearly visible. Aggregation increases the number of rounds the network delivers measurements from a large part of the network by $\approx 20\%$. Figure 4b shows results

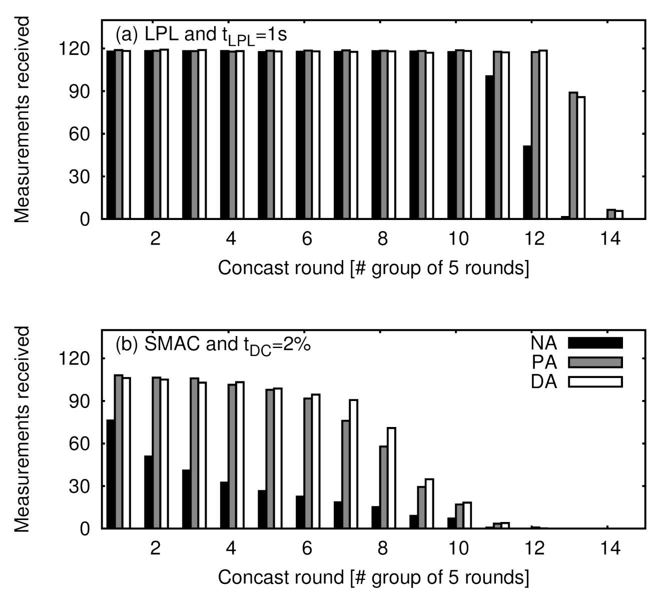


Fig. 4. Temporal distribution of measurements received at the data sink for $t_P = 60s$ using a GRID tree topology.

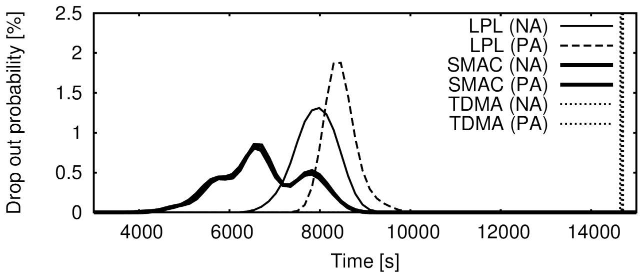


Fig. 5. Probability to run out of energy over time. LPL ($t_{LPL} = 1s$), SMAC ($t_{DC} = 2\%$), and TDMA ($t_{DC} = 1\%$) compared using GRID, $t_P = 60s$.

for SMAC using the most energy-efficient value for t_{DC}. Here, it is clearly shown that aggregation does not prolong network lifetime, but continuously increases the ratio of measurements that arrive at the data sink in each round. This confirms previous argumentation that SMAC suffers from a throughput issue in this parametrization.

Additionally, we analyzed the *node lifetime*. Figure 5 shows the probability a node runs out of energy at a certain interval of time. The Figure approximates over intervals of length 15s and uses the same parametrizations as for Figure 4. Additionally, TDMA is shown using $t_{DC} = 1\%$. Note, as NA and PA differ only marginal with SMAC and TDMA, they are plotted using the same line-style.

The results match the previous discussion. First, the node lifetime for SMAC is shorter than for LPL, which corresponds to the lower MRS shown before. TDMA outperforms all other MAC protocols. Second, for SMAC there is no normal distribution as for LPL. Instead, there are several peaks where nodes drop out with a higher probability. This confirms previous argumentation that in multi-hop network some nodes follow more than one wake-up schedule causing additional overhead. The aggregation strategy however does not influence node lifetime at all, which additionally proves that the differences in MRS are a throughput issue. Contrary to SMAC, using TDMA all nodes have nearly the same lifetime. Third, for LPL the difference between NA and PA is clearly visible. Node lifetime's increase using PA and the distribution is narrower than for NA meaning differences in lifetime between nodes in the same network are less. Both are reasons why MRS is higher for PA in this scenario.

Summarized, MRS turns out to be a suitable, realistic metric to rate the energy-efficiency of concast-based applications. Other metrics like node lifetime can give additional insights but do not confute previous results using MRS as metric.

F. Lessons Learned

Recalling the expectations of Section III-C, we gained some surprising results. First, DA did not outperform PA in any case. Second, aggregation improves energy-efficiency only in few cases. Third, results can be generalized independently of node deployment and tree topology.

Recommendations depend on the given MAC protocol and application scenario. If nodes are synchronized, a simple TDMA scheme provides best energy-efficiency. Here, aggregation does not improve energy-efficiency at all. If no node synchronization is provided, LPL should be preferred over SMAC. The impact of aggregation on one scenario with SMAC is implementation specific and cannot be generalized. With LPL, aggregation improves energy-efficiency slightly, especially if t_P is short. Summarized, energy-efficiency improvements of aggregation in concast scenarios is overrated. Choice and parametrization of MAC protocol are far more important to achieve the highest possible energy-efficiency.

V. CONCLUSION

This work provides a more realistic evaluation of the impact of aggregations strategies, tree topology and MAC protocol on energy-efficiency in periodic concast scenarios than given by most previous work. Using AVRORA+, we profit from the scalability of simulations while at the same time gaining *realistic energy consumption data*, *cross-checked* with a testbed.

Some results are quite surprising as they contradict common assumptions. The impact of aggregation strategies and tree topology is less than often assumed. Cross-layer effects caused by MAC layer are by far more important. Energy-efficiency evaluations of communication protocols either using simulation or theoretical models should thus always respect the MAC layer and its influence to energy consumption.

Future work is to include more parameters to improve generalizability, e.g., transmission power, use of acknowledgments or further MAC protocols like RI-MAC or ContikiMAC. Also, non-energy metrics like latency would be of interest.

REFERENCES

[1] C. Intanagonwiwat, R. Govindan, and D. Estrin. Directed diffusion: A scalable and robust communication paradigm for sensor networks. In *MobiCom 2000*, pages 56–67.

[2] B. Krishnamachari, D. Estrin, and S. Wicker. The impact of data aggregation in WSN. In *ICDCS 2002*, pages 575–578.

[3] S. Upadhyayula, V. Annamalai, and S.K.S. Gupta. A low-latency and energy-efficient algorithm for convergecast in WSN. In *GLOBECOM 2003*, pages 3525–3530.

[4] D. Petrovic, R.C. Shah, K. Ramchandran, and J. Rabaey. Data funneling: routing with aggregation and compression for WSN. In *Workshop on Sensor Network Protocols and Applications*, pages 156–162, 2003.

[5] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi. In-network aggregation techniques for WSN: a survey. *IEEE Wireless Communications*, 14(2):70–87, 2007.

[6] S. Madden and M. J. Franklin. TAG: a Tiny AGgregation Service for Ad-Hoc Sensor Networks. *ACM SIGOPS Operating Systems Review*, 36(SI):131–146, 2002.

[7] A. Boulis. Aggregation in sensor networks: an energy-accuracy trade-off. *Ad Hoc Networks*, 1(2-3):317–331, 2003.

[8] W. Ye, J. Heidemann, and D. Estrin. An Energy-Efficient MAC Protocol for Wireless Sensor Networks. In *INFOCOM 2002*, pages 1567–1576.

[9] D. Moss, J. Hui, and K. Klues. Low Power Listening. Technical report, TinyOS Core Working Group, 2007.

[10] A. Köpke, M. Swigulski, K. Wessel, D. Willkomm, T. E. V. Parker, P. T. Kleinhaneveld, O. W. Visser, H. S. Lichte, and S. Valentin. The MiXiM Vision. In *SimuTOOLS*, 2008.

[11] V. Shnayder, M. Hempstead, B. Chen, G. W. Allen, and M. Welsh. Simulating the power consumption of large-scale sensor network applications. In *SenSys 2004*, pages 188–200.

[12] B. L. Titzeri, K. D. Lee, and J. Palsberg. Avrora: scalable sensor network simulation with precise timing. In *IPSN 2005*, page 67.

[13] A. Hergenröder, J. Wilke, and D. Meier. Distributed Energy Measurements in WSN Testbeds with a Sensor Node Management Device (SNMD). In *ARCS 2010*, pages 341–438.

[14] C. Haas, V. Stöhr, and J. Wilke. Realistic Simulation of Energy Consumption in WSN. In *EWSN 2012*, pages 82–97.