

Dynamic Sample Rate Adaptation for Long-Term IoT Sensing Applications

Ulf Kulau, Johannes van Balen, Sebastian Schildt, Felix Büsching and Lars Wolf
Institute of Operating Systems and Computer Networks
TU Braunschweig
Email: [kulau|vanbalen|buesching|schildt|wolf]@ibr.cs.tu-bs.de

Abstract—In long-term sensing applications data patterns can vary significantly over time. Often a multitude of sensors are used to measure different types of environmental conditions. Considering such variations it is hard to select a predefined sample rate that guarantees both, high data quality and energy efficiency.

Hence, this paper presents a dynamic sample rate adaptation that strikes a balance offering optimal energy efficiency while maintaining high data quality. Based on the general concept of Bollinger Bands, a metric is derived that solely depends on the trend of the measured data itself. A real world measurement in the area of smart farming is used to show the effectiveness of this approach.

I. INTRODUCTION

The integration of Wireless Sensor Networks (WSNs) to the internet is more and more common and therefore a vital part of the Internet of Things (IoT). The standard application for most WSNs is to sense data in the field, process it and cooperatively forward it to a sink node. For this purpose WSNs often need to work autonomously in rough environmental conditions, away from any infrastructure for a prolonged time.

Under these conditions, the limited capacities of e.g. batteries lead to the fact that energy is one of the scarcest resources in WSNs [1], [2]. Especially when nodes are difficult to access or the maintainability is limited, the extension of the lifetime of a node is one of the major challenges when designing an application.

One example for an application area that faces these challenges is IoT smart farming. Here, nodes are deployed in rural areas and require a lifetime of several months [3], [4]. In this work we will consider an application tackling the distributed measurement of the crop water stress index [5] of potato plants. As the term 'stress' in connection with plants might be unfamiliar, the following paragraph will provide some background information:

Plants are stressed, when growing conditions are not optimal such as the absence of water or inferior soil. With regard to the climatic change this effect becomes a serious issue as agricultural areas begin to silt and the soil water retention decreases. Thus, with detailed information about the condition of plants, the usage of sprinkling and fertilizers can be optimized. In cooperation with a potato research station¹ we

deployed nodes with a dedicated sensor-set able to measure plant stress influencing factors on a potato field. The various parameters that have to be sensed are depicted in Figure 1.

It can be seen that the rating of the plant's condition is based on several parameters. Besides soil moisture, soil temperature, air temperature and humidity, an important parameter is the surface temperature of the plants which can be measured non-invasively by a simple infra-red temperature sensor. Whenever a plant is 'stressed' it will immediately stop evaporation and the plant heats up which can be revealed by this sensor. Of course the goal is to detect stressful environmental conditions before this happens. However, the actual weighting of the crop water stress index is beyond the scope of this paper. This work discusses a mechanism how a long lifetime of the network can be achieved while ensuring a maximum data quality despite the fact that multiple sensors observe different data patterns.

A simplified model of the power dissipation of a node is given in Equation 1.

$$P_i = \frac{1}{T} (T_{i_{active}} P_{i_{active}} + T_{i_{sleep}} P_{i_{sleep}}) \\ \Rightarrow P_{node} = \sum_i P_i \quad (1)$$

The node consists of i units (processing unit, transceiver unit, sensors, ...) with two power states each: An active state and a sleep state with a much smaller power consumption

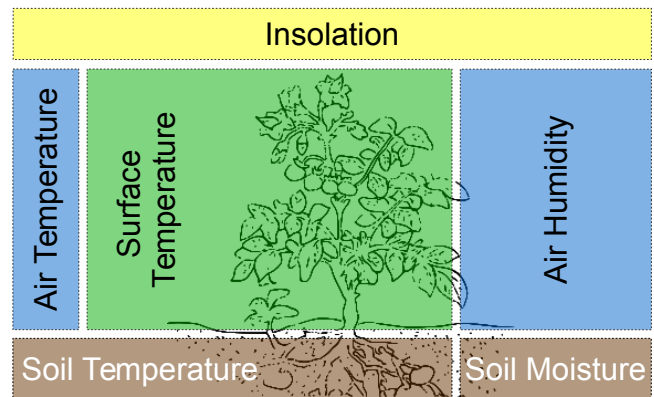


Figure 1: Required parameters to measure the crop water stress index of a potato plant.

¹<http://www.vsd-dethlingen.de>

$P_{sleep} \ll P_{active}$. Furthermore, during a time T a component i can switch between these two states. While $T_{i_{active}}$ is the total time that a component spends in active mode, $T_{i_{sleep}}$ is the total time in sleep state ($T = T_{i_{active}} + T_{i_{sleep}} \forall i$). All in all P_{node} is the sum of the components power consumption. In sum, to minimize the overall power dissipation we can reduce both, the particular power consumption $P_{i_{active}}$ and the time that a component spends in active state $T_{i_{active}}$. The reduction of $P_{i_{active}}$ of the sensors is limited but an adequate adaptation of the sample rate ($\downarrow T_{i_{active}}$) would help to reduce the total energy consumption significantly. Thus, in the course of this paper we present an approach of a dynamic adaptation of the sample rate that is solely based on the characteristics of the sensed data itself by using the general concept of Bollinger Bands. It will be shown that a self-adjustment of the sample rate is possible while preserving high data quality. Moreover, this approach is sufficiently lightweight to be applied on wireless sensor nodes with limited computational capabilities.

II. RELATED WORK

This paper focuses on the sensing unit of a sensor node. However, as energy efficiency is an important research topic in the area of WSNs, there is a plethora of existing work which aims to increase the energy efficiency of this unit.

Relating to Equation 1, [6] evaluated the impact of the voltage level on the active power dissipation $P_{i_{active}}$ of sensors and memory devices. Normally each peripheral requires a different minimum voltage level but only one fixed node-wide voltage level is applied. Lowering the voltage level to the actual minimum of a peripheral decreases the power dissipation significantly which has also been examined in [7]. These mechanisms can be combined with the approach presented in this paper to increase the energy efficiency even further.

To reduce the active usage of sensors, [8] introduced a scheme where only a random subset of distributed sensors is used to sense the environment. On average the total number of samples decreases which leads to fewer data transfers within the network and therefore less energy consumption. In the smart farming use-case, the randomized request of data is not suitable. For cost reasons you do not over-provision the amount of deployed nodes. Therefore, in a smart farming application you can not afford to lose spatial resolution by randomly sampling nodes, as it is expected that all deployed nodes experience different environmental conditions regarding the monitored properties (c.f. Figure 1).

Another possibility is to control the sample rate by detecting different but distinct activities [9]. With regard to various applications it is hard to define such distinct events.

Using Bollinger Bands to filter data has been proposed in [10]. The goal was to implement a deadband sampling algorithm to optimize the polling of a client server application. The basic idea is similar to the approach presented in this paper. However, while this work also takes up the basic idea of Bollinger Bands, the computation of the bands is modified slightly to be more suitable for the characteristic of sensors and the requirements of WSNs.

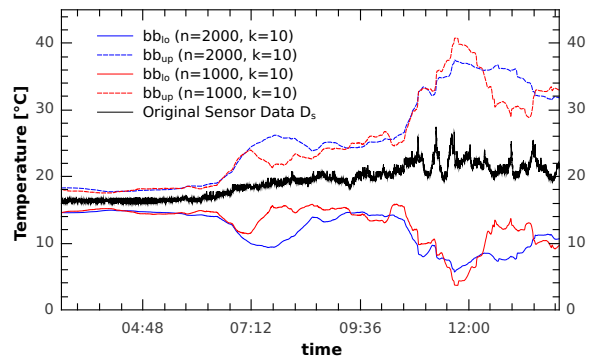


Figure 2: Upper- and lower-band of an exemplary data-set of a temperature sensor using different n .

III. BOLLINGER BANDS

Originally Bollinger Bands are an analysis tool from the financial mathematics. In the 1980s John Bollinger introduced them as a tool to analyze the trend of stock prices. Based on the normal distribution it is assumed that it is more likely that the current stock prices are close to the mean of previous stock prices [11]. However, the concept can also be transferred to the series of sampled data of a sensor node.

At a time t the Bollinger Bands are calculated by using a fixed number of n previous data-points D_s of a sensor s . The mid-band $bb_{mid_s}(t)$ is calculated as a simple moving average:

$$bb_{mid_s}(t) = \frac{1}{n} \cdot \sum_{i=0}^{n-1} D_s(t-i) \quad (2)$$

The upper- and lower-bands ($bb_{up_s}(t)$, $bb_{lo_s}(t)$) are calculated by using the standard deviation $\sigma_n(t)$ of the n previous data-points which is added and subtracted from the mid-band respectively.

$$\begin{aligned} bb_{up_s}(t) &= bb_{mid_s}(t) + k \cdot \sigma_n(t) \\ bb_{lo_s}(t) &= bb_{mid_s}(t) - k \cdot \sigma_n(t) \end{aligned} \quad (3)$$

The factor k can be used to control the width between the upper- and the lower-band while the standard deviation is given by:

$$\sigma_n(t) = \sqrt{\frac{1}{n} \cdot \sum_{i=0}^{n-1} (D_s(t-i) - bb_{mid_s}(t))^2} \quad (4)$$

By way of example we consider a usual data-trace of a temperature sensor. Figure 2 shows the upper $bb_{up_s}(t)$ and lower $bb_{lo_s}(t)$ bands for values of $n = 1000$ and $n = 2000$. To illustrate the characteristics of the Bollinger Bands applied to real sensor data, a relatively large number of historical data n is used in Figure 2. It can be seen, that growing dynamics of the data itself lead to wider Bollinger Bands. As intuitively increasing variation in a sensed condition demands for a higher sample rate, this effect can be exploited for dynamic adaptation of the sample rate as we will show in the following section.

IV. DYNAMIC SAMPLE RATE ADAPTATION

The general idea is to use the width of the upper- and lower-band as a metric to control the sample rate of a sensor. The goal is to have a fully self-adjusting sampling rate control mechanism that is effective even for a priori unknown sample data. Especially for long-term sensing in various IoT applications this approach offers several advantages: The patterns and dynamics of environmental data changes over time, e.g., due to weather conditions or seasons.

To adapt the sample rate we define a maximum waiting period t_{max} , which limits the minimum sample rate of a sensor. When determining the next point in time ($t_{wait}(t)$) to sample new data, we need to consider how much fluctuation the data series exhibited during the last measurements. We define a dynamic estimation function $dyn(t)$ that quantifies the expected dynamics in the data series based on previously sampled data.

Equation 5 shows the actual calculation of the waiting time which is estimated after every sampling.

$$t_{wait}(t) = \frac{t_{max}}{1 + dyn(t)^\varphi} \quad (5)$$

The exponent φ is used to weight the value of the dynamic estimation $dyn(t)$. The effectiveness of this estimation depends mainly on the chosen dynamic estimation function which will be introduced in the next sections.

A. Dynamic estimation using Bollinger Bands

With regard to Section III the Bollinger Bands seem to be a good basis for an adequate dynamic estimation. With Equation 3 it is possible to calculate the upper- and lower Bollinger Band. The width between these bands is given by:

$$\begin{aligned} \Delta_{bb}(t) &= |bb_{ups}(t) - bb_{lows}(t)| \\ &= 2k \cdot \sigma_n(t) \end{aligned} \quad (6)$$

With $b = 2k$ to adjust the distance between the considered Bollinger Bands the dynamic estimation formula can be written as:

$$dyn_{bb}(t) = b \cdot \sigma_n(t) \quad (7)$$

B. Dynamic estimation using vertical distances

As shown in Figure 2 the standard deviation of a moving mean offers a feasible metric for the dynamic estimation. However, the standard deviation is based on the arithmetic mean which is not always an ideal reference value to calculate the dynamics of sensor data.

In general sensor data D_s can not be described by an echelon form (staircase function) but rather a series of linear functions between the particular data points (cf. Figure 3).

Being based on the arithmetic mean of sample points, the dynamic estimation function $dyn_{bb}(t)$ increases when the sampled data series shows a rising or falling trend. Those trends do not necessitate a higher sampling rate, as constantly

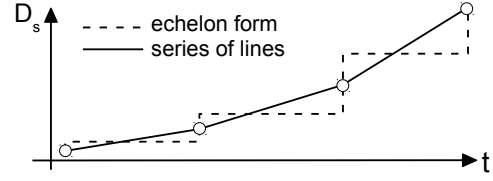


Figure 3: Sensor data are represented by lines between the particular data points.

rising or falling data series contains no fluctuations that might be missed by a lower sampling rate. In many cases the line between two data points is sufficient to approximate the intermediate data points. Therefore, $dyn_{bb}(t)$ would lead to a higher sample rate while fewer sampling would be adequate to represent the sensed data.

Based on this insight an optimized dynamic estimation function can be derived. Instead of using the arithmetic mean, the vertical distances between a line from the first and the last data point and the linear connection of the previous n data points is considered. Figure 4 illustrates this procedure.

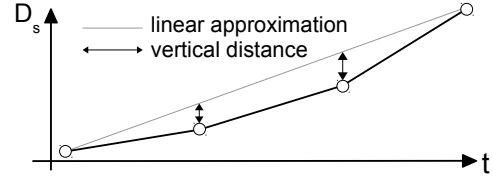


Figure 4: Vertical distance between the sampled data and the linear approximation of the n previous data points.

Here the intention is, the more linear the characteristic of a data series has been in the past, the lower the sample rate can be. The dynamic estimation function based on vertical distances $dyn_{vd}(t)$ is given by the following Equation 8:

$$dyn_{vd}(t) = \frac{k}{n} \cdot \sum_{i=0}^{n-1} |f(t-i) - D_s(t-i)| \quad (8)$$

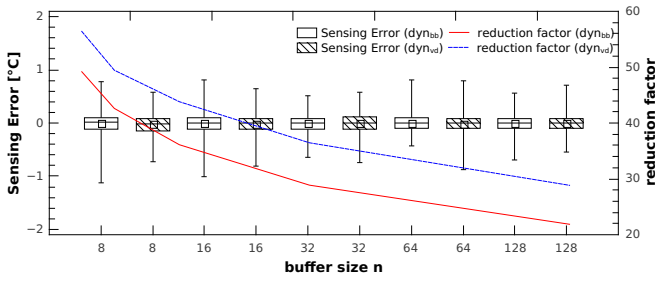
The linear function $f(t-i)$ is therefore given by:

$$\begin{aligned} f(t-i) &= (t-i) \cdot m + D_s(t-(n-1)) \\ \text{with } m &= \frac{D_s(t) - D_s(t-(n-1))}{n-1} \end{aligned} \quad (9)$$

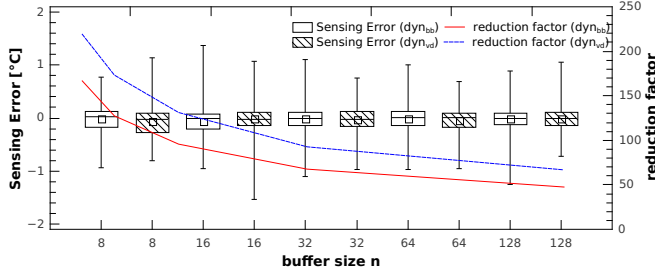
By inserting $dyn_{vd}(t)$ in Equation 5 the next waiting time can be estimated similar to the Bollinger approach.

V. ANALYSIS AND COMPARISON OF THE WAITING TIME ESTIMATION

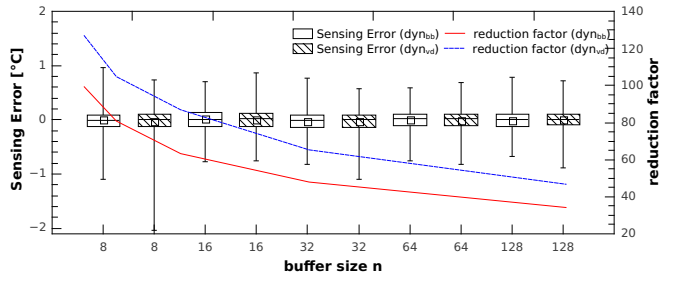
To evaluate the efficiency of the dynamic estimation functions $dyn_{bb}(t)$ and $dyn_{vd}(t)$ as well as the impact of different buffer sizes n and maximum waiting times t_{max} , we use an exemplary dataset of a one day temperature measurement. The exponent φ is set to a fixed value of 2. A temperature sensor (located outdoors) with a resolution of 0.02°C and a sample



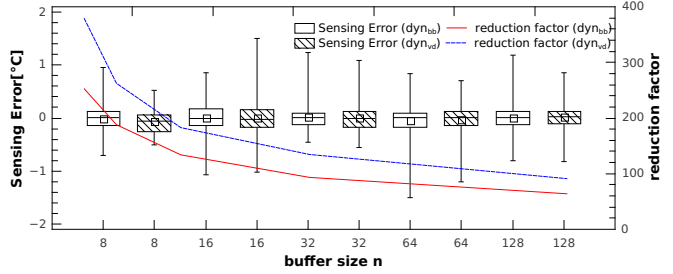
(a) Maximum waiting time $t_{max} = 300$ s.



(c) Maximum waiting time $t_{max} = 1500$ s.



(b) Maximum waiting time $t_{max} = 750$ s.



(d) Maximum waiting time $t_{max} = 3000$ s.

Figure 5: Evaluation of the dynamic estimation functions $dyn_{bb}(t)$ and $dyn_{vd}(t)$ ($\varphi = 2$) by using temperature measurement of an exemplary day (reference data: sample rate ≈ 0.3 Hz $\hat{=}$ 26500 samples, resolution 0.02 °C). The reference data are compared against the results of the dynamic adaptation of the sample rate by using the online-estimation of the next waiting time.

rate of 0.3 Hz was used to generate a fine-grained temperature profile. Figure 6 shows such a temperature profile.

By using this dataset the dynamic adaptation of the sample rate can be emulated. As already depicted in Section IV after each sample the next waiting time is estimated using either $dyn_{bb}(t)$ or $dyn_{vd}(t)$.

We evaluated the error of measurement compared to the fine-grained raw data and the reduction in energy usage. For different parameters of buffer size n and maximum waiting time t_{max} we compared the results against the full reference dataset.

Figure 5 shows the results. In general it can be seen, that the sample rate can be reduced significantly while incurring only a relatively small sensing error. Moreover, with regard to the error of measurement, difference between the dynamic estimation functions $dyn_{bb}(t)$ and $dyn_{vd}(t)$ is small in all scenarios. However, the dynamic estimation function based on the vertical distances $dyn_{vd}(t)$ shows a significant higher

reduction of sampling rate and thus energy usage.

While the variation of buffer size n has no significant impact on the error of measurement, the reduction factor decreases with a growing buffer size. This effect is due to the fact that low frequency changes fade out slowly as the buffer contains many previous sampled data. In sum it is recommended to use a relatively small buffer size n as this will increase energy savings while at the same allow the dynamic estimation to react faster to changes.

Another expected observation is that the maximum waiting time t_{max} , which determines the lowest allowed sample rate, influences both, the error of measurement and the reduction factor. It can be seen that the number of measurement errors growth slightly with a higher t_{max} . Especially the minimum and maximum deviations to the reference data increases. However, the reduction factor strongly depends on t_{max} . While $t_{max} = 300$ s leads to a reduction factor of up to 56, a waiting time of $t_{max} = 3000$ s yields a reduction factor of 378.

VI. EXPERIMENT

As mentioned in the introduction (cf. Section I), this work was motivated by enabling long-term IoT sensing applications. In particular the measurement of the crop water stress index was introduced as a use case demanding the use of multiple sensors.

To enable the measurement of the required data illustrated in Figure 1 a simple node based on an 8-Bit micro controller unit (MCU) was developed and integrated into our WSN testbed PotatoNet [4].

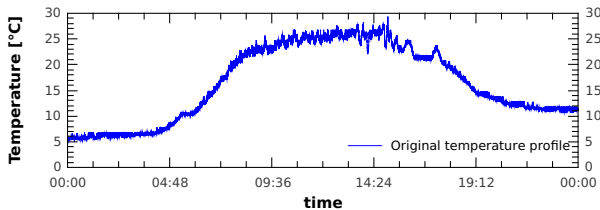


Figure 6: Exemplary temperature profile of one day measurement.

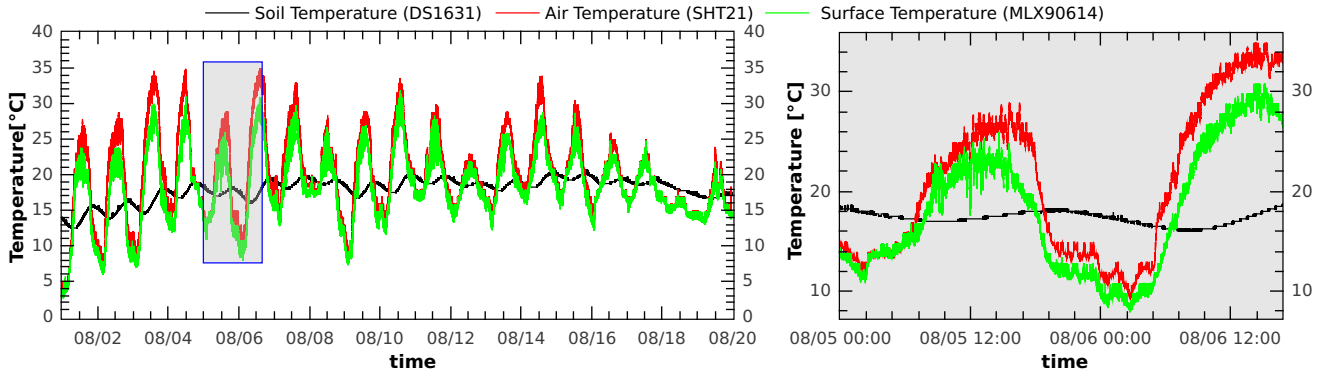


Figure 7: Exemplary data pattern of the soil temperature, air temperature and surface temperature of one measuring station.

The following sensors have been used to measure the parameters:

- *Melexis MLX90614* [12]: Inter-Integrated Circuit (I²C) Infra-red temperature sensor to allow a non-invasive measurement of the surface temperature.
- *Sensiron SHT21* [13]: I²C temperature and air humidity sensor.
- *Maxim DS1631* [14]: I²C temperature sensor to measure the soil temperature.
- *Pyranometer*: Analogue sensor based on a photo-diode PDB-C139.
- *Soil moisture*: Analogue sensor based on an oscillating circuit with a capacitive probe.

The sensors are connected to a ATmega328P low-power 8-bit MCU [15] and assembled to a combined measuring station as shown in Figure 8. All digital sensors are connected to the I²C master of the MCU while the analog-digital converter (ADC) of the MCU is used to integrate the analogue sensors. The MCU implements a slave I²C which is used to connect the measuring station to nodes of the WSN testbed PotatoNet [4].

In total four measuring stations were deployed on a potato field for a time span of about 38 days. The sample rate of each sensor was set to ≈ 0.3 Hz, as the gathered data

Table I: Sample time and current consumption of the considered temperature sensors.

Sensor	Parameter	Sample time [ms]	Current [mA]	
			Datasheet	In practice
SHT21	Temperature	≈ 63.0	0.2 - 0.33	0.235
DS1631	Temperature	≈ 585.1	≤ 1.0	0.640
MLX90614	Temperature	≈ 150.0	1.3 - 2.5	1.354

should provide the reference for the evaluation of the presented approach. Figure 7 shows the soil temperature, air temperature and surface temperature of one measuring station. For the sake of clarity we limit the evaluation to those 3 data series, which already represent largely varying characteristics.

To evaluate the impact on the energy efficiency we firstly measured the current consumption of each sensor with the aid of an online-oscilloscope [16]. Table I summarizes the sensors and gives the time they need to sample a single value and the current consumption during this process (with a fixed voltage supply for all components of 3.3 V).

Similar to the analysis in Section V, the results of the dynamic sample rate are compared against the reference data to derive the sensing errors and reduction factors. However, in this case only the dynamic estimation based on the vertical distances $dyn_{vd}(t)$ was used, as the analysis revealed its advantages. Instead of a single day, the entire temperature profile depicted in Figure 7 has been used. The maximum waiting time t_{wait} was set to 500 s and the exponent $\varphi = 2$.

Figure 9 shows the results of this experiment where the total energy consumption is based on the values of Table I. On the left Figure 9a we can see that for each sensor a significant reduction is achieved. Moreover the reduction strongly depends on the characteristic of the sampled data itself. The reduction factor of the DS1631 (soil temperature) is almost two times higher than the reduction factors of the MLX90614 (surface temperature) and SHT21 (air temperature). This is due to the fact that the sensors measure different temperature profiles. The thermal capacity of the ground leads to the effect that the temperature curve of the soil is dampened compared to air and surface temperature. Thus, higher frequency fluctuations in air or surface temperature do not increase the dynamics of

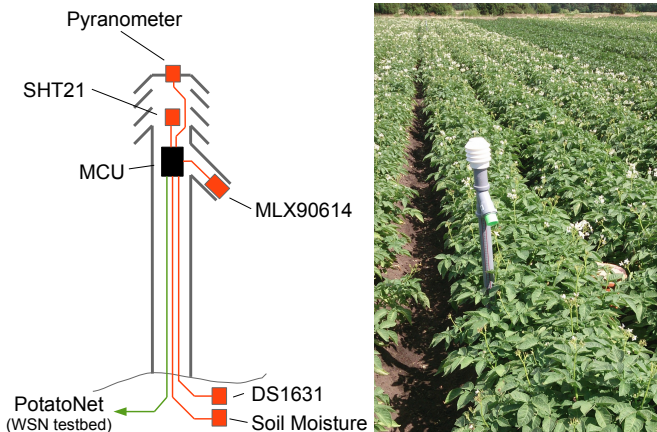


Figure 8: Illustration and picture of one measuring station to evaluate the crop water stress index of potato plants.

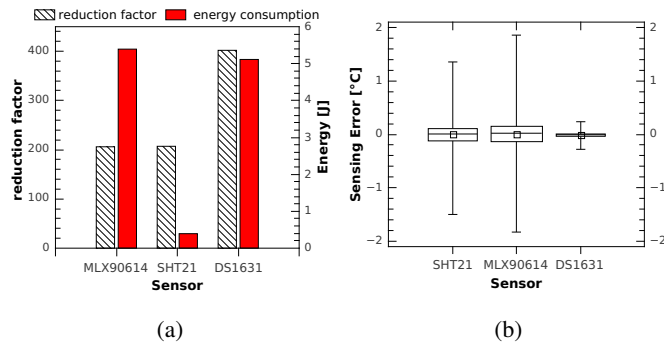


Figure 9: Experimental results of the reduction error, the total energy consumption and the overall sensing error.

the data series experienced by the soil temperature sensor. The presented approach is able to handle such individual variations and reaches a balanced state between high data-quality and energy efficiency for each sensor (cf. Figure 9b).

A. Practical test on low-power MCUs

Due to the energy constraints, the computational performance of wireless sensor nodes is limited. So far in this analysis we performed the sample rate adaptation offline in a simulation using the complete high resolution dataset collected in the field.

To show both, the actual online adaptation and the low computational overhead, we implemented the dynamic adaptation with the dynamic estimation $dyn_{vd}(t)$ on an 8-bit ATmega328P MCU [15] and a single temperature sensor (exemplary for a typical sensor node). A second MCU was placed close to the node with the variable sample rate, but sampled the temperature with a fixed rate of 1 Hz (reference node). Both nodes were located outdoors and sampled temperatures. In sum, the node using the dynamic adaptation of the sample rate performed only 0.104% of sensor readouts compared to the reference node. Due to the computational overhead and the non-negligible power dissipation during the sleep state, the consumed energy was 0.319% of the reference node. However, the sensing error for 91.3% of the dynamically sampled data was below $\pm 0.5^\circ\text{C}$.

VII. CONCLUSION

Motivated by a real-world sensing application in the area of smart farming, this work presented a dynamical adaptation of the sample rate solely based on the sensed data itself.

The concept of Bollinger Bands, which are usually used in financial mathematics, offers a metric to control the waiting time till the next sample. The metric was improved with regard to the characteristic of data gathered by a WSN. As a result the energy efficiency of sampling can be increased with only negligible impact on the data quality.

The mechanism was evaluated against the background of a real world sensing application trying to measure the crop water stress index of potato plants. The results show that the sample rate can be reduced significantly while keeping the sensing error small.

The presented approach is applicable to sensed data series exhibiting a wide range of individual characteristics as the sample rate is adapted accordingly.

An implementation of the presented approach on an low-power 8-bit MCU illustrates the applicability of the approach to common WSN and IoT applications.

ACKNOWLEDGMENTS

We would like to thank the VSD Dethlingen research station for providing the field and technical support. This research was partially supported by the German Research Council (DFG) under the grant no. BU 3282/2-1.

REFERENCES

- [1] F. Schlachter, "No Moore's Law for batteries," *Proceedings of the National Academy of Sciences*, vol. 110, no. 14, pp. 5273–5273, 2013.
- [2] Holger Karl and Andreas Willig, *Protocols and architectures for wireless sensor networks*. John Wiley & Sons, 2007.
- [3] K. Langendoen, A. Baggio, and O. Visser, "Murphy loves potatoes: experiences from a pilot sensor network deployment in precision agriculture," in *the 20th IEEE International Parallel and Distributed Processing Symposium*, ser. IPDPS 2006, April 2006, p. 8.
- [4] U. Kulau, S. Schildt, S. Rottmann, B. Gernert, and L. Wolf, "Demo: PotatoNet – Robust Outdoor Testbed for WSNs: Experiment Like on Your Desk. Outside." in *the 10th ACM MobiCom Workshop on Challenged Networks*, ser. CHANTS 2015, Paris, France, September 2015. [Online]. Available: <http://doi.acm.org/10.1145/2799371.2799374>
- [5] R. D. Jackson, S. Idso, R. Reginato, and P. Pinter, "Canopy temperature as a crop water stress indicator," *Water resources research*, vol. 17, no. 4, pp. 1133–1138, 1981.
- [6] U. Kulau, F. Büsching, and L. Wolf, "A Node's Life: Increasing WSN Lifetime by Dynamic Voltage Scaling," in *the 9th IEEE International Conference on Distributed Computing in Sensor Systems 2013*, ser. DCoSS 2013, Cambridge, USA, May 2013. [Online]. Available: <http://www.ibr.cs.tu-bs.de/papers/kulau-dcss2013.pdf>
- [7] S. Friedrichs, U. Kulau, and L. Wolf, "Energy-Efficient Voltage Scheduling of Peripheral Components on Wireless Sensor Nodes," in *IEEE International Conference on Communications Workshop on Energy Efficiency in Wireless Networks & Wireless Networks for Energy Efficiency*, ser. E2Nets 2014, Sydney, Australia, June 2014.
- [8] F. Fazel, M. Fazel, and M. Stojanovic, "Random access compressed sensing for energy-efficient underwater sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1660–1670, 2011.
- [9] Z. Yan, V. Subbaraju, D. Chakraborty, A. Misra, and K. Aberer, "Energy-efficient continuous activity recognition on mobile phones: An activity-adaptive approach," in *2012 16th international symposium on wearable computers*. IEEE, 2012, pp. 17–24.
- [10] N. M. Torrisi, "Preliminary evaluation of bollinger deadband filtering using upper and lower bands," in *Industry Applications (INDUSCON), 2014 11th IEEE/IAS International Conference on*. IEEE, 2014, pp. 1–6.
- [11] J. Bollinger, "Using bollinger bands," *Stocks & Commodities*, vol. 10, no. 2, pp. 47–51, 1992.
- [12] Melexis, "Datasheet MLX90614, single and dual zone infrared thermometer," 2015. [Online]. Available: <https://www.melexis.com/-/media/files/documents/datasheets/mlx90614-datasheet-melexis.pdf>
- [13] Sensirion, "Datasheet SHT21 Humidity and Temperature Sensor IC," 2014. [Online]. Available: https://www.sensirion.com/fileadmin/user_upload/customers/sensirion/Dokumente/Humidity_Sensors/Sensirion_Humidity_Sensors_SHT21_Datasheet_V4.pdf
- [14] Maxim Integrated, "Datasheet DS1631/DS1631A/DS1731 High-Precision Digital Thermometer and Thermostat," 2015. [Online]. Available: <http://datasheets.maximintegrated.com/en/ds/DS1631-DS1731.pdf>
- [15] Atmel Corporation, "Atmega328p datasheet," 2015. [Online]. Available: <http://www.atmel.com/devices/atmega328p.aspx>
- [16] R. Hartung, U. Kulau, and L. Wolf, "Distributed Energy Measurement in WSNs for Outdoor Applications," in *IEEE International Conference on Sensing, Communication and Networking Conference*, ser. SECON 2016, London, UK, June 2016.