Journal of Ambient Intelligence and Smart Environments 4 (2012) 123–134
DOI 10.3233/AIS-2012-0139
IOS Press

Collective communication for dense sensing environments

Predrag Jakimovski [a,*], Hedda R. Schmidtke [b], Stephan Sigg [c], Leonardo Weiss Ferreira Chaves [d] and Michael Beigl [a]

[a] *Pervasive Computing Systems – TecO, Karlsruhe Institute of Technology (KIT), Vincenz-Priessnitz-Str. 1, 76131 Karlsruhe, Germany*
E-mail: Predrag.Jakimovski@kit.edu
[b] *Carnegie Mellon University in Rwanda, Telecom House, Boulevard de l'Umuganda, Kacyiru-Kigali, Rwanda*
E-mail: schmidtke@cmu.edu
[c] *National Institute of Informatics (NII), 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan*
E-mail: sigg@nii.ac.jp
[d] *SAP Research CEC Karlsruhe, Vincenz-Priessnitz-Str. 1, 76131 Karlsruhe, Germany*
E-mail: leonardo.weiss.f.chaves@sap.com

Abstract. Intelligent Environments are currently mostly implemented with WSN technologies using conventional connection-based communications. However, connection-based communications may impede progress towards intelligent environments involving massive amounts of sensor nodes. The goal of this paper is to chart a field of more suitable technologies for communication in intelligent environments, which we call *collective transmission methods*. The idea of collective transmission is to establish communication not between single senders and single receivers but between collectives of senders and receivers, by making use of constructive interference of simultaneously sent signals.

We detail how the collective transmission approach can be realized for a concrete application scenario: item-level tagging using printed organic electronics. We describe an algorithm that can be realized on very simple tags. With a testbed implementation, we show that this algorithm can realize robust, collective, approximate read-out of 21 simultaneously sent signals.

Keywords: Smart labels, wireless sensor networks, item-level tagging, collective information transmission

1. Introduction

The emergence of intelligence has its origin in perception of the physical world and communication between different participants. To realize the vision of intelligent environments, massive amounts of sensor data need to be processed in a spatially distributed way. Communication in intelligent environments is currently mostly implemented using WSN technologies with conventional connection-based communications. However, connection-based communications may be unsuitable for IE scenarios involving massive amounts of relatively simple sensor nodes. The goal of this paper is to point at a spectrum of more suitable technologies for communication in intelligent environments, which we call *collective transmission methods*. The idea of collective transmission is to establish communication not between single senders and single receivers but between collectives of senders and receivers. We focus our discussion on a simple, yet practically relevant and soon realizable application example from the domain of next generation business process management technologies: item level tagging using extremely low-cost tags implemented with printed organic electronics.

The main part of the paper is structured as follows. After an introduction to the application scenario (Sec-

[*]Corresponding author.

tion 2) and a discussion of related works (Section 3), we introduce the general approach and its wider applicability in Section 4. In Section 5, we explore the practical realization of the scenario with our approach. An experimental evaluation with our printed electronics test bed (a simulation using silicon-based hardware) is presented and discussed in Section 6.

2. Business systems for smart supply chain management

Store houses, factories, and retail stores are very likely to be among the first intelligent environments. Many business processes can be automated and optimized using item-level tagging [4,26]. When individual items can be uniquely identified, functions, such as registration of goods received, quality control and in-store processes can be implemented more efficiently:

– *Goods received*: When a pallet with goods arrives, the receiving company usually checks if the right amount and the right goods were delivered. Using item-level tagging, a system can identify which products are packed in a pallet, and compare the quantities and the product identifiers with those in the advanced shipping notification.
– *Quality control*: After checking that a pallet contains the right amount of goods, one checks if the delivered goods fulfill a certain set of quality standards [25]. The data from item-level sensors can be queried to check if for instance a certain temperature threshold is exceeded.
– *In-store processes*: Using item-level tagging, retailers can automatically check whether there are enough goods in the shelves [23]. Also, retailers can check if goods are arranged in a correct way in the shelves, by checking compliance to predefined layout plans, so called planograms [24].

Item-level tagging does not only allow automatization, it also allows more efficient implementation of processes. For instance, the temperature within a pallet with goods may be distributed unevenly during transport and may display different dynamics [9]. Because of that, tagging the complete pallet with only a single smart label captures an incomplete view of the transportation process. As a result, a group of packages in an area that has been over-heated may remain undiscovered if it is far enough from the smart label monitoring the temperature. With item-level tagging this could be avoided, as packages would be monitored individually.

Even though such processes can be efficiently automated and optimized using item-level tagging, item-level data is mostly only needed on a technical level to implement the processes themselves. Many of the processes do not require item-level data but some function computable from item-level data. For instance, for quality control it is in most cases sufficient to know if a pallet is OK or not. If it is, it can be further processed. However, if it does not pass the test, goods might be checked individually before the whole pallet would be sent back.

This scenario makes a strong case for sensing environments using collective communication. Collective communication not only makes the communication process simpler, it also facilitates data processing. This is exemplified using the quality control process:

– *Simpler process*: Using item-level tagging without collective communication involves repeatedly scanning the pallet from different positions, for example with an RFID reader. This is mainly because physical interferences [5] prevent multiple smart labels from being read with one readout. Using collective communication, only one readout is needed, thus making the process simpler and faster to execute.
– *Simpler data processing*: When using item-level tagging without collective communication, the reader device will send the data it gathers to another system where data processing takes place, such as an Enterprise Resource Planning (ERP) system or an Inventory Management System. This system will eliminate duplicate readings and check if the data fulfills predefined standards, for instance, if the temperature of each of the up to 1000 items is below a certain threshold. After that, the system will decide whether the pallet needs further inspection. With collective communication, the aggregated information on the communication channel can be evaluated directly, on the reader itself, to decide if the pallet is OK. Only one reading instead of up to 1000 readings has to be processed.

The vision of item-level tagging comes closer to its realization with organic printed electronics. Organic printed smart labels will be capable of recording sensor data such as temperature, humidity or light exposure. Organic smart label technology promises ultra low-cost massive deployment in industry, food, pharmaceutics, healthcare and consumer markets, as tags can simply be printed on packages.

Production of organic electronic circuits can be faster, cheaper and simpler than RFID, as industrial standard printers can be used instead of dust-free fabrication facilities needed for silicon-based electronics, allowing massive deployment [19]. However, printed electronics cannot compete in terms of performance, reliability, and size with RFID.

Applications for organic printed smart labels are, for instance, in cost sensitive retail: super markets have on average a shrinkage of 2.77% per year [17]. This is a significant amount as the average profit margin is only 1.10%. The percentage of perishable goods amounts to 30%, causing more than 56% of the entire shrinkage [2] by spoilage. The principal reasons for spoilage are expired products or interrupted cold chains within supply chains from the manufacturer to the retail stores. A key scenario for the first organic printed electronics is therefore temperature monitoring in logistics and supply chain management, and first destructive binary organic temperature sensors have been developed[1].

In the following, we assume a scenario of a pallet arriving at a storage facility. The pallet contains a large number of items[2], which are checked for the maximal temperatures measured during transport for quality control. We are interested in two specific tests:

- Binary query: have any items been exposed to a certain temperature?
- Proportion query: how many of the items have been exposed to a given temperature?

In case of cold chains, for instance, a pallet could be rejected if a certain temperature threshold has been exceeded during transport. For some good and if temperatures were not too high, goods could still be sold at a discount, reducing the financial damage. Moreover, these checks would allow the transport company to detect failures of the cooling system in a truck early and avoid successive damaging of goods. Figure 1 shows our application scenario for readout of item level tagging in supply chain management. Pallets are investigated by a screening device, which can process the compound signal received from the simultaneously sending tags attached to goods in packages piled up on the pallet.

[1]http://www.polyic.com

[2]http://mstonline.de/mikrosystemtechnik/mst-smart-label/Clustermeetingrfid/locostix

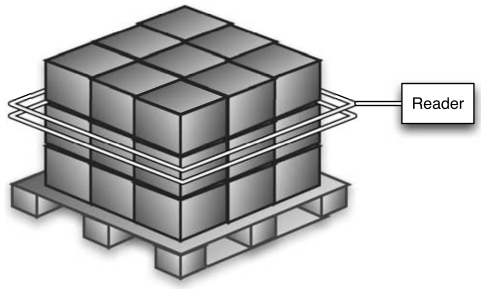

Fig. 1. A typical application scenario for item level tagging in supply chain management: pallets are investigated by a screening device.

3. Related works

Communications in wireless sensor networks between a number of source nodes and a collecting sink node are mostly realized using connection-based communication schemes. Conventional channel access communication schemes, such as time division multiple-access (TDMA) [22], are designed to avoid collisions of data packages on the channel. This creates a range of problems for the scenario of communication between a large number of tags and the collecting reader node, as tags have to be handled by the reader individually. To achieve this, each source node needs to have a unique identity so that communication with it can be separated from communication with other nodes. In the example of TDMA, these individual connections simply take more time. For other protocols it means a larger bandwidth (FDMA) or decreasing throughput (CDMA). In any case, the traditional protocols do not scale easily to large amounts of sensor nodes.

The idea to allow and employ simultaneous data transmission of source nodes to a designated sink in a wireless sensor network is relatively new [15,16]. In these works, constructive interference between superimposing signals of sensor nodes sending the same signal is exploited to improve the robustness or strength of the signal at the receiver. Such concurrent transmission of data in wireless sensor networks promises to gain more performance in terms of energy efficiency, throughput and latency [18,20].

A practical realization of a robust system for counting sensor nodes using constructive interference and statistical properties in a communication scheme has been introduced by Krohn [11–14] with Synchronous Distributed Jam Signaling (SDJS). Based on the fraction of time slots occupied by jam signals among a fixed number of available slots the number of transmitting devices can be estimated in a highly mobile and ad-hoc wireless network. Our method extends on this, as we aim to estimate the amounts of classes with high accuracy.

The work most closely related to our method [6–8], exploits the superposition of the signals on the channel as a calculator for specific functions, as simultaneous transmission realizes an addition function: Goldenbaum and Stańczak show how this can be used for calculating the arithmetic or geometric mean of measurements in one step during simultaneous transmission of the observed data by the sensor nodes.

To realize computation on the channel, nodes perform at first preprocessing on the measurements before transmitting the sensed data simultaneously. To create a constructive superposition, they encode the sensed data with predetermined random phase sequences. On the receiver side a post-processing is performed to recover the result of the calculated function. The authors claim that a coarse block/sequence synchronization is sufficient to initiate a constructive superposition on the channel. Both the reader system of Goldenbaum and Stańczak and the counter implemented by Krohn yield a single numerical value. In the above scenario of the proportion query however, we need to transmit a multiset of values. Our reader system processes the superimposed codes to estimate the proportion of sensor nodes that sent a certain value.

Theoretical evaluations, as proposed e.g. in [6–8], yield promising results, however an instrumentation in a realistic setting might face considerable additional challenges, especially regarding accuracy [1,3] and robustness against

- interactions between neighboring and distant nodes,
- lack of accuracy in synchronization between nodes,
- ambient noise,
- changes in ambient noise,
- changes in transmission power, frequency and phase offset of nodes.

Questions of robustness are particularly important for low-cost nodes, such as printed organic electronic smart labels. We therefore implemented a testbed for collaborative transmission methods and tested our algorithm under realistic conditions. All tests were performed in the student lab at TecO under regular working conditions. Our implementation for a set of 21 nodes thus allows a realistic evaluation of the practical feasibility of the method in the scenario.

4. General approach

The general problem we study in this paper is how to obtain information from a set of simultaneously sending nodes, in our scenario the tags attached to goods on the pallet. We aim to request from the pallet which proportion of tags measured which values. In principle, this could be done by querying tags individually using any of the well-established protocols. Implementing protocols that assign a distinct channel to each sender however is not feasible in our scenario, since the senders need to be simple and we assume a large number of senders.

For these reasons, we suggest to use novel collective, approximate versions of the traditional multiple access techniques of time division, frequency division, or code division (TDMA, FDMA, CDMA). The SDJS approach of Krohn [12] for counting the number of senders, for instance, can be viewed as a collective, approximate version of TDMA: all tags send a single burst signal in a random time slot of a given base interval and the reader then statistically analyses from the number of filled time slots, how many tags there might have been. Similar time-slot techniques could be used by the reader to ask the pallet, whether a certain value was measured and even how many tags have measured a certain value.

In a similar way as SDJS but using code division instead of time division, our goal was to develop an algorithm that can statistically analyze the superimposed signals from all tags on the pallet and estimate which proportion of senders sent which value. While time slots and frequencies can encode ranges of values particularly well, our code-based method can be generalized to encode any type of value.

CDMA is based on bit sequences c that are shared between a sender S and a receiver R. A bit sequence v is sent from S as $s = c \oplus v$, where $\oplus$ is the bitwise *exclusive or*. The receiver extracts v from s by computing $v = s \oplus c$. The double application of $\oplus c$ cancels out c and v is regained. Simultaneous connections between a number of senders S_i and corresponding receivers R_i can then be achieved: simultaneous transmission yields the superimposed signal as the sum $s = s_1 + s_2 + \cdots + s_n$ of signals s_i sent, since the amplitudes of synchronized signals of the same frequency are approximately added to each other when the bit sequences s_i are sent.

The resulting signal s is *similar* to each of the original signals s_i, where similarity can be based on various distance metrics on bit sequences $v, w \in \{0,1\}^n$, e. g. on the Hamming distance:

$$d_H(v, w) = \sum_{i=1}^{n} |v_i - w_i|. \qquad (1)$$

The similarity can then be defined by choosing a threshold T_n suitable for the length of the vectors n. Two bit sequences $v, w \in \{0,1\}^n$ are called *similar* if they differ only in a small number T_n of bits:

$$v \sim w \overset{\text{def}}{\Leftrightarrow} d_H(v, w) \leqslant T_n.$$

A number of pairs of senders and receivers can thus communicate via codes c_i. If the codes c_i are chosen so as to be orthogonal, or at least sufficiently different from each other, this entails that we can obtain v_i from s by applying $v_i' = s \oplus c_i$. The result v_i' is similar to v_i so that v_i can then be regenerated from v_i', using error correcting codes. Codes c_i can be generated so as to be orthogonal, however, sufficiently long random bit sequences, are also suitable: statistical theory suggests that the probability to obtain two random bit sequences of low similarity is higher, the longer the sequences are.

The key properties employed in this encoding are the notions of similarity and difference and of similarity preserving operations and distancing operations: addition is an operation that preserves similarity, whereas $\oplus$ but also the *circular bitwise shift* are distancing operations, which make their result different from both its operands. CDMA uses the $\oplus c_i$ encoding to ensure that the signals s_i sent are sufficiently different, and thus not mixed during simultaneous transmission.

In our scenario, we only need to ensure that different values transmitted can be retrieved from the superimposed signal. Moreover, the individual tags are much too simple and their number n is too large, as to allow for any complex protocol or encoding mechanism to be implemented. We therefore directly encode numerical values using a single random bit vector z_0 shared by all tags and the receiver. We obtain sufficiently different codes z_i for numerical values i by circularly shifting z_0 by the amount of i bits, since shifting is a distancing operation. In this way, a single bit vector $z_0 \in \{0,1\}^n$ can be used to encode n values.

The received signal $s = s_1 + s_2 + \cdots + s_n$ is then simply a sum of encoded numbers z_i, directly encoding the multi-set of measured values. If three tags, for instance, send the values $\{7, 8, 12\}$ the received signal would be $s = z_7 + z_8 + z_{12}$. The receiver can now check the similarity between s and any value z_i by simply testing $s \sim z_i$.

Using this method, we can already resolve the *binary query* outlined in Section 2, to check whether some goods have been exposed to a temperature higher than a given threshold. In many cases, however, an estimation of how many tags sent which of the values is needed (*proportion query*). One way to do this is least squares estimation (LSE), as we will show in more detail in the next section. The algorithm for the reader and tags can then be realized:

1. Tags come initialized with a register t set to the minimum temperature 0, and transmit code z set to z_0.
2. Each tag measures its environment continuously over a longer duration: if the measured value is $m > t$, then
   (a) it sets $t := m$.
   (b) it shifts the code z accordingly, that is: sets $z := z_t$.
3. Reader sends `start` signal to tags.
4. Tags send their respective z.
5. Reader receives overlayed signal s:
   (a) *Binary Query*:
      i. Set $S := \emptyset$.
      ii. For each possible value z_i:
      if $z_i \sim s$ then $S := S \cup \{z_i\}$.
   (b) *Proportion Query*: For each value $z \in S$: use Least Squares Estimation (LSE) to compute proportion of contribution of z:
      i. Generate linear equation system for the found values $z_i \in S$.
      ii. Estimate parameters a_i so that error is minimal.
      iii. Set $M := \{(a_i, z_i) | s = \sum_{z_i \in S} a_i * z_i\}$.
   (c) *Output*: return M.

In an actual printed electronics implementation, the register t and the variable z of steps 1 and 2 would be combined. It would be possible, for instance, to implement the two steps with a destructive, physical temperature sensor that shifts a start/stop pointer forward along the fixed random vector z_0 in response to higher measured temperatures (Fig. 2). When the readout signal is sent the tag can then respond correctly by sending from start point to end point.

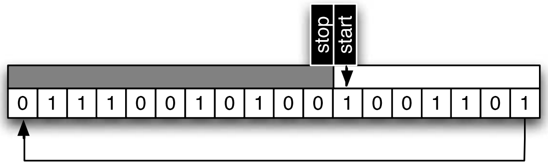

Fig. 2. A model of a simple tag using a destructive, physical temperature sensor (gray and white bar): the start/stop pointer is shifted forward along the fixed random vector z_0 as increasing temperatures permanently alter the material of the sensor (grey). When a readout signal is received the correspondingly shifted vector would be transmitted.

5. Collective transmission

We now discuss the details of our implementation of the algorithm. The architecture of our instrumental setup consists of n wireless sensor nodes (the tags) and a sink node (the reader) processing the received signal (see Fig. 3). The data transmission of the sensor nodes is triggered through an external signal (step 3) as in the case of RFID tagging. After initiating the transmission process each node in the sensor field is transmitting its measured sensory value simultaneously. The bit vector encoding a measured value v to be sent is transmitted in step 4 by a node sending out a sinusoidal signal in a time slot if in the sequence of bits a '1' occurs, otherwise it keeps silent.

In Fig. 3 a possible scenario is depicted. When two or more nodes are simultaneously transmitting a sinusoidal signal, the signal components interfere on the channel and are received in a superposition by a receiver. Consequently, the amplitude of the superimposed electromagnetic waves is either intensified or becomes less intense.

In Fig. 4 an example of a superposition between three sine waves is shown. The amplitude strength depends on the number of participating nodes, their individual transmission power, the dominance of the line of sight components to the scattered multi-path signal components and the distance between receiver and sensor nodes. During the transmission of the bit sequences from the n sensor nodes, the maximum can therefore vary in each time slot making measurement of the strength of the signal difficult. An example for a received raw signal is depicted in Fig. 5.

By detecting the maximal amplitude in each time slot a vector of maximal amplitudes is created on the receiver side (Fig. 6, step 5), which is then used to extract the sensory information of the collective information transmission.

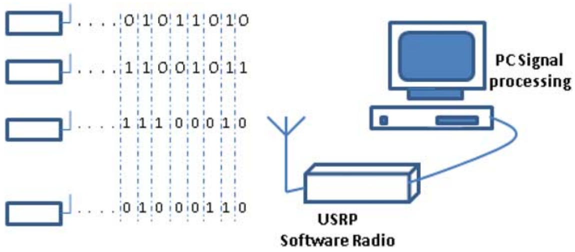

Fig. 3. Principle of collective transmission. Each sensor node reached by an external trigger signal is transmitting its binary sequence at the same time. Based on the different number of '1's in each time slot different maximal amplitudes are generated. On the receiver side, the superimposed binary sequence is captured.

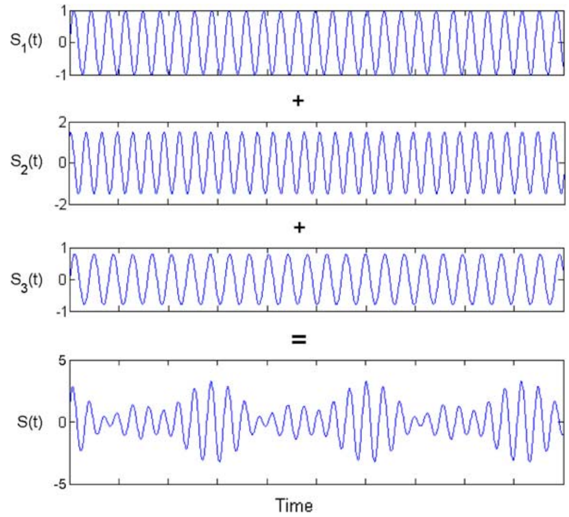

Fig. 4. The superposition principle: $s(t)$ is a superimposed signal generated by three sine signals $s_1(t), s_2(t)$ and $s_3(t)$. The sine signals are chosen slightly different from each other in frequency, phase and amplitude strength, i.e. $f_{s_1} = 16$ Hz, $f_{s_2} = 18$ Hz and $f_{s_3} = 20$ Hz. Thus, when two or more waves traverse the same space, the amplitude at each point is the sum of the amplitudes of the individual waves.

For encoding values, we chose a 100-bit-long random vector z_0 in such a way that $z_i \nsim z_j$ for $i \neq j$. The vector thus allows robust encoding of 100 values by shifting. Moreover, the relatively long random sequence makes it possible to benefit from statistical methods for robust retrieval of vectors from the superimposed signal. By statistical properties, a noisy version of a random vector may differ in more than a third, and it is still recognizable [10].

The main steps of the algorithm, the binary query (step 5a) and the proportion query (step 5b), have distinct applications. The binary query is a simple and highly reliable method to find out whether a value has

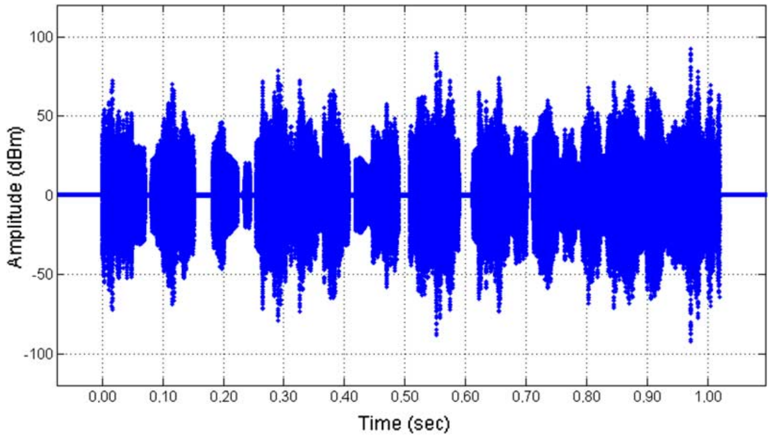

Fig. 5. Raw data of a superimposed signal caused by 21 transducers transmitting different binary sequences simultaneously. The signal length is set to 100 time slots.

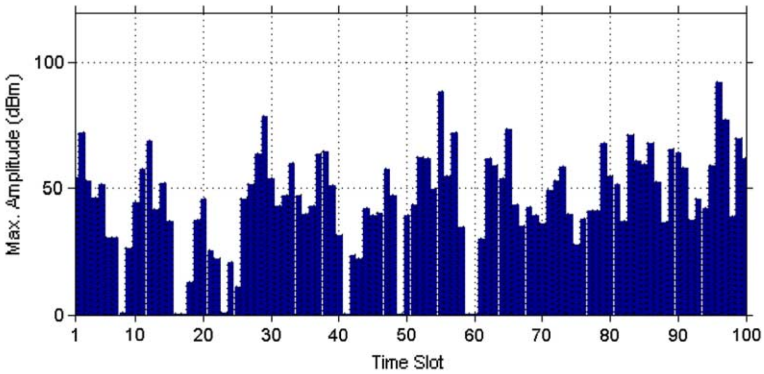

Fig. 6. Quantification of the superimposed signal shown in Fig. 5. In each time slot the maximal amplitude is detected and visualized by a single bar.

been sent at all. The proportion query uses this information to additionally compute which percentage of senders have sent a certain value.

5.1. Binary query

The advantage of collective information transmission is that we can get sensory information at once in an environmental monitoring application. Often one is not interested in single sensory values, but rather in estimating the state of a sensor field, by detecting whether or not a certain property is present. The Hamming distance has the property of being suitable to identify vectors contained in a received superimposed signal. For calculating the Hamming distance between two vectors $v = (v_1, v_2, \dots, v_n), w = (w_1, w_2, \dots, w_n) \in [0,1]^n \subset \mathbb{R}^n$ equation (1) also applies. However, if two vectors are not in the interval $[0,1] \subset \mathbb{R}$ they need to be normalized. For measuring the difference between a measured input vector $v \in \mathbb{R}^n$ and an expected vector $w \in \mathbb{R}^n$, we normalize to the maximal amplitudes $A_v = \max_i v_i$ and $A_w = \max_i w_i$, yielding the generalized definition:

$$d_H(v, w) = \sum_{i=1}^{n} \left| \frac{v_i}{A_v} - \frac{w_i}{A_w} \right|.$$

The similarity can then be defined by

$$v \sim w \overset{\text{def}}{\Leftrightarrow} d_H(v, w)/n < T_n,$$

where T_n is a threshold suitable for the length of the vectors n.

In practice the usage of the Hamming distance has its limits [10], the Hamming metric is applicable only if small sets of different vectors are added. The more vectors are used to encode entities, the lower the probability of identification. With large sets of different signals, synchronization issues become more critical and noise increases. However, the method still scales with larger numbers of nodes that transmit a small set of values, as in collective transmission. While noise also increases in the case of a large number of nodes, especially due to synchronization issues, values still were detected reliably in our testbed.

5.2. *Proportion query*

Using binary query alone the following applications can already be realized

- detection of an abnormality, for instance, the pallet containing perished goods,
- detection of the presence of classes A, B, C indicate temperature intervals, such as $A = [0 \ldots 8]^\circ$, $B = [10 \ldots 25]^\circ$, $C = [26 \ldots 100]^\circ$.

However, the capabilities of the system can be extended considerably when we can estimate the proportions of the classes A, B, C, e.g. computing the percentages of senders in the classes as $A = 30\%, B = 60\%, C = 10\%$.

To realize the proportion query, a mathematical formalization of the superposition principle combined with the statistical mechanism is required. Thus, the first step is to collect the possible vectors s_i that can be sent in a matrix A. Therefore, let

$$A = (s_1 s_2 \cdots s_M) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \cdots & 0 \end{pmatrix}$$

be an $N \times M$ matrix that contains the M vectors $s_1, s_2, \ldots, s_M$ of length N.

The modeling of superposition is based on a linear system, which is additive and homogeneous. Hence, the physical model is given by the linear system

$$z(a_1 s_1 + a_2 s_2 + a_3 s_3 + \cdots + a_M s_M) = y\,, \quad (2)$$

where the parameters $a_i \in \mathbb{R}$ indicate the number of sensor nodes that have sent out the binary sequence $s_i \in \{0,1\}^N$. The variable $y \in \mathbb{R}^N$ contains the recorded N maximal amplitudes of the N time slots, cf. Fig. 6 the received collective information transmission. To adapt the model to the reality numerically, a trade-off is required, which is expressed by $z \in \mathbb{R}$.

In the implementation, the first step is to solve the linear system in (2) without considering z, i.e.

$$\hat{a_1} s_1 + \hat{a_2} s_2 + \hat{a_3} s_3 + \cdots + \hat{a_M} s_M = y\,. \quad (3)$$

In the second step the solution $\hat{a} = (\hat{a_1}, \ldots, \hat{a_M})^T$ of the linear system and the number of the participating sensor nodes is used to calculate the trade-off z. The number of the sensor nodes is usually not known. We therefore operate with percentages of senders, and assume the number of sensor nodes n to be 100 in the following. If the number of senders is known, n can be set accordingly.

$$z = \frac{n}{\sum_{i=1}^{M} \hat{a_i}}.$$

Finally, the solution of the parameters $a = (a_1, \ldots, a_M)$ can be estimated as

$$a_i = \hat{a_i} z \quad \text{for } i = 1, \ldots, M\,.$$

The component a_i of the solution vector a then gives the estimated percentage of sensor nodes transmitting the bit sequence v_i.

Assuming uncorrelated measurements and equal Gaussian error σ^2, the parameters in $a = (a_1, \ldots, a_M)$ can be estimated by using linear least squares estimation. Thus, the preliminary solution is given by evaluating

$$\hat{a} = (A^T A)^{-1} \cdot (A^T y)\,.$$

Afterwards, the output vector $\hat{a}$ is used to get the final estimation of a by applying $a = z\hat{a}$, where $a, \hat{a} \in \mathbb{R}^N$ and $z \in \mathbb{R}$, as described above.

6. Evaluation

For testing our approach, we explicitly chose a monitoring scenario in which the sensors are observing some environmental parameter and a designated receiver is reading out the measurements from all sensor nodes located within the range simultaneously. The monitoring of perishable goods in a cold chain is such an example. The purpose of the following experiments is to illustrate the performance and robustness of our approach in our organic electronics test bed.

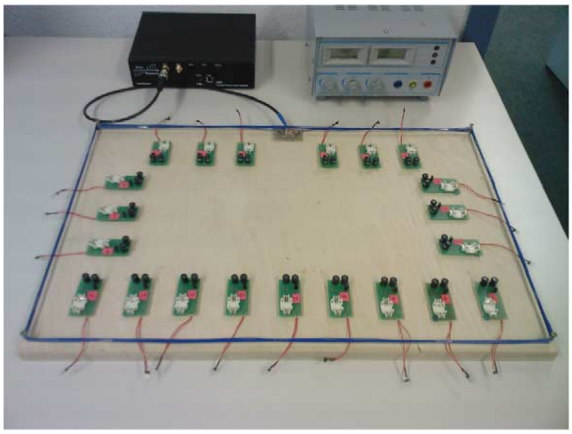

Fig. 7. The research platform consists of 21 transducers and one receiver connected to a PC. During experiments, transducers are controlled by sensor nodes (not shown).

6.1. Pre-study

According to studies on the constraints of printed organic electronics [19,21], organic electronics will behave and develop very differently from traditional electronics. Thus, for testing purposes, we created thirty transducers on PCB with off-the-shelf components, which conform to the constraints of organic electronics and in this manner mimic their behavior. The operating transmission frequency is set to 135 kHz, because tests have shown that an analog oscillator of the transducer is generating a stable sinusoidal signal in this low frequency domain when using a small number of electronic components. Additionally, it has been considered that first working printed circuits will be operating in the lower frequency domain. In Fig. 7 the entire experimental platform is shown, which consists of the mentioned transducers, a loop antenna operating in the low frequency domain and one receiver[3] connected to a PC, where the computation and visualization of a received transmission is performed.

To see how the collective transmission of the transducers behaves under real environmental conditions, we carried out about 360 experiments as follows: By changing the parameters such as the number of transducers sending out simultaneously and the distance between antenna and transducers, we obtained a detailed behavior of the superimposition on the channel. For instance, Fig. 8 depicts four histograms gained by experiments with 5, 10, 15, and 30 transducers, respectively. Transducers transmit a sinusoidal signal simultaneously. Here, the positive amplitudes of the superimposed signals are considered. Their histograms resemble a Rician distribution as a probability density function (PDF). The distribution fit for the four experiments is shown in Fig. 9. The distribution of amplitudes indicates that an amplitude with a high magnitude is a rare event, that is, the exact coincidence of amplitudes from several signals concurring at the same time point is not probable given the hardware. Our testbed thus reflects lack of synchronization in real systems.

[3]Ettus Research: http://www.ettus.com/products

6.2. Testbed experiment

Experiment design To give proper evaluation results to the approach proposed in Section 5, we set transducers to transmit a certain bit sequence corresponding to a certain value. In this way, the evaluation was performed under realistic and controlled conditions. The values in the experiment were programmed into the transducers, but the collective information transmission (steps 4 and 5) took place as in the case of a real environmental monitoring scenario. In this way, we arranged several different setups in which the position and vector sent by transducers were varied.

To create the required vectors for encoding temperature values, we first generated a 100-bit long equally distributed pseudo-random bit vector. Then the randomly drawn binary vector was circularly bit-wise shifted to encode further values, as described in Section 4. With respect to circuitry design for organic printed smart labels the circular bit-wise shifting operator was used for testing, as it would be cost-efficient to print tags with only one vector, but synchronization problems could be critical in this case, leading to false values being read.

The choice for using a specific 100-bit long vector was based on preliminary study. The relatively small bit-length of the z_0 vector was chosen carefully. The use of higher dimensional binary vectors did not improve or make the performance worse, whereas with the use of lower dimensionality the performance suffers. Moreover, it is possible to encode 100 numerical values using the shifting method with this vector. In our experiments the transducers were positioned as it is shown in Fig. 7 reflecting the intended usage scenario (Fig. 1).

Through the circular bit-wise shifting operator 100 different temperature classes can be represented, however with maximally seven possible different values being used in our instrumental set-up, we limited our

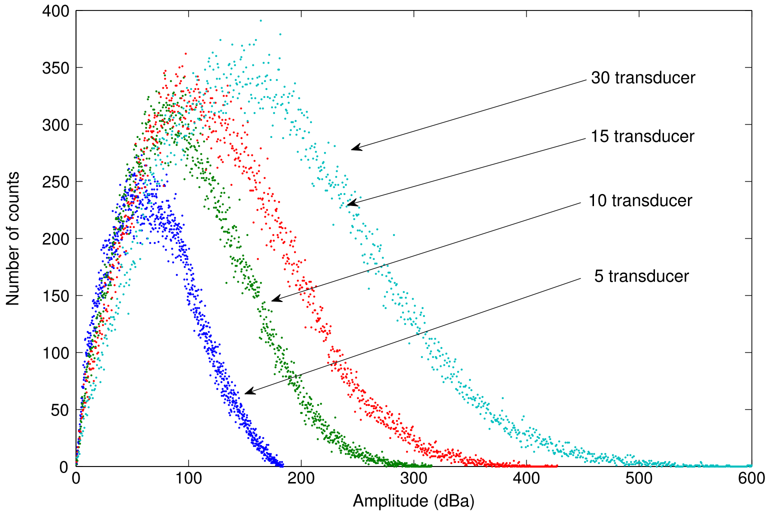

Fig. 8. The four histograms illustrate the distributions of amplitudes which are generated by 5, 10, 15 and 30 transducers sending a sinusoidal signal simultaneously.

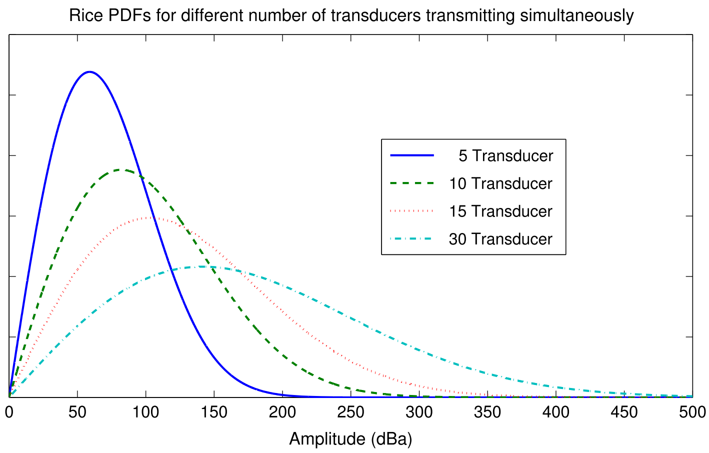

Fig. 9. The four distributions indicate various Rice PDFs obtained by fitting data of 5, 10, 15 and 30 transducers which sent a sinusoidal signal simultaneously.

evaluation to a comparison of only these seven possible classes. We tested all 15 different settings possible. For each setting, ten trials were executed and evaluated. The testing environment was part of the TecO student computing lab, and experiments were conducted during regular usage of the facilities.

Results Table 1 shows our evaluation results. The first column describes the setting by chosen values and transducers. For example a setting of 21 describes the trials where all 21 transducers sent the same value and a setting of 9, 6, 6 means nine transducers sent a value A, six send a value B and the remaining six send a value C. Column two presents the average amount of correctly recognized temperature values using the binary query algorithm exclusively. Column three shows the mean error with respect to the seven possible classes for each set of trials when the proportion query algorithm was applied. When comparing against all 100 possible values (fourth column), the mean error is again considerably lower between 0.01% and 1.16%.

6.3. Discussion

The experiments suggest, that collective information transmission is possible. The simple algorithm already yields results that would be acceptable for a range of

Table 1
Results

Setting	Binary query (accuracy)	Proportion query (mean error per class)	
		7 classes	100 classes
21	88.57%	1.77%	0.0001%
18.3	89.05%	1.70%	0.3963%
15.6	91.90%	1.22%	0.1359%
12.9	97.62%	0.34%	0.0579%
15.3.3	90.00%	1.56%	0.5585%
12.6.3	89.05%	1.63%	0.5199%
9.9.3	89.05%	1.56%	0.3796%
9.6.6	82.38%	2.52%	0.1277%
12.3.3.3	82.86%	2.52%	0.9124%
9.6.3.3	82.38%	2.45%	0.6107%
6.6.6.3	80.95%	2.59%	0.3811%
9.3.3.3.3	80.00%	3.13%	0.6047%
6.6.3.3.3	85.71%	2.11%	0.6991%
6.3.3.3.3.3	80.00%	2.99%	1.1610%
3.3.3.3.3.3.3	79.52%	2.79%	0.9540%

applications, such as estimating whether a pallet has been damaged during transport. For the given number of maximally seven values, the simple communication scheme seemingly is robust enough. The different classes of sensory information sent in a collective information transmission are reliably detected, and the number of the senders, which have sent the same sensory information are estimated with an averaged error of 2.06% in comparison with seven classes.

Problems of the current testbed are its still low number of senders when compared to the pallet scenario of 1000 tags. Moreover, conditions in the testbed are presumably much better than in a pallet. Different packaging materials and larger distances could increase synchronization problems.

7. Conclusions

We outlined a novel approach for communication between large numbers of senders a single receiver. This method of *collective readout* was shown to be a robust, collective, approximate communication method for reading out massive amounts of sensor nodes by combining communication with computation on the channel. We described and tested an implementation to realize collective read-out that can be realized in an efficient manner on very simple tags. The experiments suggest the general feasibility of this mechanism in the economically meaningful scenario of item-level tagging for next-generation business process support.

However, our results have further reaching consequences. While computation on the channel has been advocated previously on theoretical grounds, its practical use for intelligent environments was so far questionable, as experimental results regarding robustness to noise and inaccurate synchronization under realistic conditions were missing. The implementation of the proposed collective transmission method, however, has shown that statistical methods can be employed to improve tolerance to noise and phase shifts. Future work should include expanding the testbed with more sending nodes and a more realistic pallet design.

The robustness of collective transmission comes from the use of random vector encodings of numerical values. In our example application, collective transmission makes it possible to communicate simultaneously with the complete pallet. Collective transmission does not aim to communicate with individual senders but with the collective. The transmitted signal, the sum of all transmissions, is an approximate representation of a multi-set of values. Future works will further elaborate such construction of representations through collective transmission. A disadvantage of the simple example scenario is its centralized architecture: intelligent environments with massive amounts of sensor nodes should not rely on a central processing unit, and instead employ the spatial distribution of nodes. Approaches on distributed representations and computations, such as Vector Symbolic Architectures [10], can further guide this work. Collective transmission and read-out could be fundamental building blocks for realizing distributed intelligence.

Acknowledgment

This research was partially supported by the project Polytos, in the BMBF-funded excellence cluster "Printed Organic Circuits and Chips"[4], and also by the Deutsche Forschungsgemeinschaft (DFG) in the project "SenseCast: Kontextprognose zur Optimierung von Netzwerkparametern in drahtlosen Sensornetzen" (BE4319/1). We further acknowledge partial funding by a fellowship within the Postdoc-Programme of the German Academic Exchange Service (DAAD).

[4]http://www.innovationlab.de/en/research/excellence-cluster-forum-organic-electronics/research-projects/printed-organic-circuits-and-chips-polytos/

References

[1] H. Adam, E. Yanmaz, W. Elmenreich and C. Bettstetter, Contention-based neighborhood estimation, in: *Vehicular Technology Conference, IEEE 71st*, 2010, pp. 1–5.

[2] P. Beswick, M. Isotta and S. Winter, A retailer's recipe for fresher food and far less shrink, in: *Oliver Wyman Journal Issue 24*, 2008, pp. 47–53.

[3] C. Budianu and L. Tong, Estimation of the number of operating sensors in a sensor network, in: *Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers*, vol. 2, 2003, pp. 1728–1732.

[4] C. Decker, M. Berchtold, L. Weiss Ferreira Chaves, M. Beigl, D. Röhr, T. Riedel, M. Beuster, T. Herzog and D. Herzig, Cost-benefit model for smart items in the supply chain, in: *Proc. of Internet of Things'08*, 2008.

[5] C. Floerkemeier and M. Lampe, Issues with RFID usage in ubiquitous computing applications, in: *Proc. of Pervasive'04*, 2004.

[6] M. Goldenbaum and S. Stanczak, Computing the geometric mean over multiple-access channels: Error analysis and comparisons, in: *43rd Asilomar Conference on Signals, Systems and Computers*, Monterey, USA, November 2009.

[7] M. Goldenbaum and S. Stanczak, Computing functions via SIMO multiple-access channels: How much channel knowledge is needed? in: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '10)*, Dallas, USA, March 2010, pp. 3394–3397.

[8] M. Goldenbaum, S. Stanczak and M. Kaliszan, On function computation via wireless sensor multiple-access channels, in: *IEEE Wireless Communications and Networking Conference (WCNC '09)*, Budapest, Hungary, April 2009.

[9] R. Jedermann, L. Ruiz-Garcia and W. Lang, Spatial temperature profiling by semi-passive RFID loggers for perishable food transportation, *Computers and Electronics in Agriculture*, 2009.

[10] P. Kanerva, Hyperdimensional computing: An introduction to computing in distributed representation with high-dimensional random vectors, *Cognitive Computation*, 2009, pp. 139–159.

[11] A. Krohn, Überlagerte Funksignale in drahtlosen Sensornetzwerken, PhD thesis, Karl-Friedrich-Gauss-Fakultät für Mathematik und Informatik der Technischen Universität Carolo-Wilhelmina zu Braunschweig, TU Braunschweig, 2007.

[12] A. Krohn, M. Beigl and S. Wendhack, SDJS: Efficient statistics for wireless networks, in: *Proc. of the 12th IEEE International Conference on Network Protocols*, 2004, pp. 262–270.

[13] A. Krohn, M. Hazas and M. Beigl, Removing systematic error in node localisation using scalable data fusion, in: *Fourth European Conference on Wireless Sensor Networks*, 2007, pp. 341–356.

[14] A. Krohn, T. Zimmer, M. Beigl and C. Decker, Collaborative sensing in a retail store using synchronous distributed jam signalling, in: *International Conference on Pervasive Computing*, Springer Verlag, 2005, pp. 237–245.

[15] J. Laneman, G. Wornell and D. Tse, An efficient protocol for realising cooperative diversity in wireless networks, in: *Proc. of the IEEE International Symposium on Information Theory*, 2001, p. 294.

[16] J. Lanemann, D. Tse and G. Wornell, Cooperative diversity in wireless networks: Efficient protocols and outage behaviour, *Transactions on Information Theory*, December 2004, pp. 3062–3080.

[17] L. Miller and J. Allen, National supermarket shrink report, Technical report, National Supermarket Research Group, 2005.

[18] R. Mudumbai, G. Barriac and U. Madhow, On the feasibility of distributed beamforming in wireless networks, *IEEE Transactions on Wireless Communications*, **6** (2007), 1754–1763.

[19] Organic Electronics Association, OE-A Roadmap for organic and printed electronics, *White Paper*, 2008.

[20] S. Sigg, R.M.E. Masri, and M. Beigl, Feedback-based closed-loop carrier synchronization: A sharp asymptotic bound, an asymptotically optimal approach, simulations, and experiments, *IEEE Transactions on Mobile Computing (TMC)*, **10**(11) (Nov. 2011), 1605–1617.

[21] V. Subramanian, J.B. Chang, A. de la Fuente Vornbrock, D.C. Huang, L. Jagannathan, F. Liao, B. Mattis, S. Molesa, D.R. Redinger, D. Soltman, S.K. Volkman and Q. Zhang, Printed electronics for low-cost electronic systems: Technology status and application development, in: *Solid-State Device Research Conference 2008, ESSDERC 2008, 38th European*, 2008, pp. 17–24.

[22] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.

[23] L. Weiss Ferreira Chaves, E. Buchmann and K. Böhm, Tagmark: Reliable estimations of RFID tags for business processes, in: *Proc. of KDD'08*, 2008.

[24] L. Weiss Ferreira Chaves, E. Buchmann and K. Böhm, RPCV: RFID Planogram Compliance Verification, in: *Proc. of ICDE'10*, 2010.

[25] L. Weiss Ferreira Chaves and F. Kerschbaum, Industrial privacy in RFID-based batch recalls, in: *Proc. of the International Workshop on Security and Privacy in Enterprise Computing InSPEC'08*, 2008.

[26] L. Weiss Ferreira Chaves and Z. Nochta, Printed organic smart labels: Breakthrough towards the internet of things? in: D.C. Ranasinghe, Q.Z. Sheng, and S. Zeadally, eds, *Unique Radio Innovation for the 21st Century*, Springer, 2010.