Collective Communication for Dense Sensing Environments

Predrag Jakimovski, Florian Becker, Stephan Sigg, Hedda R. Schmidtke, and Michael Beigl
TecO, Pervasive Computing Systems, Faculty of Informatics
Karlsruhe Institute of Technology
76131 Karlsruhe, Germany
Email: Predrag.Jakimovski@kit.edu, becker@teco.edu, sigg@teco.edu, schmidtke@teco.edu, Michael.Beigl@kit.edu

***Abstract*—Intelligent Environments are currently implemented with standard WSN technologies using conventional connection-based communications. However, connection-based communications may impede progress towards IE scenarios involving high mobility or massive amounts of sensor nodes. We present first results on a novel communications technology for intelligent environments: *collective transmission*. To make the discussion more concrete we focus on a simple, yet practically relevant and soon realizable application example: item level tagging using printed organic electronics.**

Item level tagging is a key enabling technology for next generation business process support. So-called organic printed tags can be used to label large numbers of items in a most cost-efficient way, thus creating sensing environments with massive amounts of communicating nodes. However, this cost advantage comes at the price of lower computational power and less reliable communications. In reading information from such massive amounts of items, traditional connection-based read-out is unfeasible.

We present a novel approach for collective transmission, which implements robust, collective, approximate read-out of large numbers of simple tags. Our approach uses mechanisms for calculation by simultaneous transmission. We detail the collective transmission approach, discuss its implementation in the organic printed label scenario, and show first results of experiments conducted with our smart label test bed. We conclude with an outlook on the potential of collective transmission, and argue that collective transmission is a fundamental building block for realizing distributed intelligence.

I. Introduction

The emergence of intelligence has its origin in communication between different participants and sensing the physical world permanently. 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 standard WSN technologies and conventional connection-based communications. However, connection-based communications may impede progress towards IE scenarios involving high mobility or massive amounts of sensor nodes. The goal of this paper is to present first results on a novel communications technology for intelligent environments, which we call *collective transmission*. The idea of collective transmission is to establish communication not between single senders and single receivers but between collectives.

To make the discussion more concrete, we focus 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 so-called printed organic electronics.

Scenario: The vision of item level tagging of commercial products and goods comes closer to its realization with organic printed electronics. The goal is creating organic printed smart labels, which are 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 will 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 [6]. However, printed electronics cannot compete in terms of performance, reliability, and size.

Applications for organic printed smart labels are e.g. in cost sensitive retail: super markets have on average a shrinkage of 2.77% per year [14]. 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 [16] by spoilage. The principal reasons for spoilage are expired products or interrupted cool chains within supply chains from the manufacturer to the retail stores.

A key scenario for the first organic printed electronics can therefore be temperature monitoring in logistics and supply chain management. First binary organic temperature sensors have been developed[1]. We assume a scenario of a pallet or shelf containing approximately 1000 items[2] to be checked for the maximal temperatures that have been measured. In case of cool chains, for instance, a market could be interested in compliance checks as to whether perishable goods were exposed to higher temperatures during transport. If a pallet shows compliance violation, it could be rejected or, if temperatures were not too high, goods could still be sold at a discount, depending on the amount of violation. For shelves, periodical compliance checks would allow a retailer to detect failures of

[1] PolyIC: http://www.polyic.com
[2] Locostix: http://mstonline.de/mikrosystemtechnik/mst-smart-label/Clust\ermeetingrfid/locostix

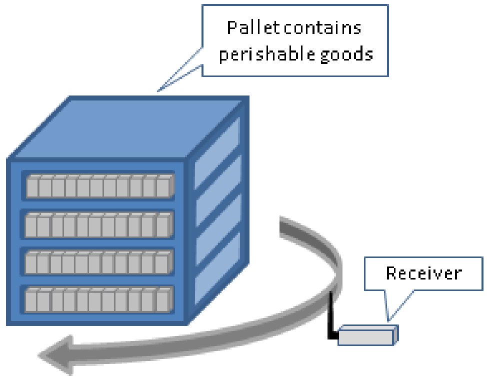

Fig. 1. A typical application scenario for item level tagging in supply chain management is shown. Pallets are investigated by screening the perishable goods.

the cooling system early and to determine whether goods were actually damaged and to what amount.

We present a novel robust read-out mechanism that allows estimation of the distribution of values sent simultaneously from a large number of very simple tags, as in the pallet read-out scenario. Our approach combines a method for calculation by simultaneous transmission with statistical mechanisms for increasing robustness and reliability.

Structure of the article: The main part of the paper is structured as follows. After a discussion of related works (Sect. II), we introduce the general approach and its wider applicability in Sect. III. In Sect. IV, we explore the practical realization of the pallet scenario with our approach. An experimental evaluation with our printed electronics test bed (a simulation using silicon-based hardware) is presented and discussed in Sect. V.

II. Related works

Traditional communication protocols, such as Time-division-multiple-access (TDMA) [18] are commonly used in wireless sensor networks, which do strictly avoid the interference during the data transmission from a source node to a designated sink. The communication and data processing are generally separated from each other. The idea of superimposing the data from each source node to the sink by simultaneous transmission is relatively new [12], [13]. In particular, the interference generated by signals is then exploited to improve the robustness or strength of the signal. The concurrent transmission of data in wireless sensor networks promises to gain more performance in terms of energy efficiency, throughput and latency [15], [17]. The closest related work published recently on the issue to exploit the superposition of the signals come from Goldenbaum and Stańczak [7]–[9]. They employ the multiple access channel (MAC) as a calculator for desired functions, i.e. calculating the arithmetic or geometric mean of the measurements in one step on the channel during simultaneous transmission of the observed data by the sensor nodes. In many application scenarios only the arithmetic or geometric mean of the sensed data is of a practical interest, where the node identity is irrelevant. To realize the computation over MAC, they 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 their approach needs a coarse block/sequence synchronization to initiate a constructive superposition over the MAC. However, an actual implementation under realistic conditions and tests for robustness to noise under such conditions have not yet been performed.

In contrast, our instrumentation features a real hardware set-up and with our encoding scheme we are able to extract a specific data value transmitted from the superimposed received signal and not only the aggregated information. Despite these theoretic calculations, an instrumentation in a realistic setting will likely face considerable additional challenges not considered in these calculations. In particular, for a precise calculation, a very accurate synchronization between nodes is required. Otherwise, the superimposed signals may again impair the ability to decode the encoded, superimposed information from the channel. In particular, the ambient noise figure might change, transmission power, frequency and phase offset of nodes might be different so that the synchronization of the superimposed signal might be harder to read at the receiver side. For instance, the authors in [1], [2] discuss issues related to the estimation of neighboring nodes in practical settings.

A first robust system using constructive interference and statistical properties in a communication scheme has been introduced by Albert Krohn [3]–[5], [11] 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 is estimated in a highly mobile and ad hoc wireless network.

III. General approach

The general problem we study in this paper is how to obtain information from a region, in our scenario the pallet, as a whole. We want to request from the pallet which proportion of tags measured which values. In principle, this could be done by querying each tag using any of the well-established multiple access methods. 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. Therefore, we need special *collective, approximate* versions of the traditional multi access techniques of time division, frequency division, or code division (TDMA, FDMA, CDMA). The SDJS approach of Krohn [3] 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 in which proportion which value was sent. While time slots and frequencies can encode ranges of values very 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 send 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 $s = s_1 + s_2 + \ldots + s_n$, 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 any distance metric on bit sequences $v, w \in \{0,1\}^n$, such as the Hamming distance:

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

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 numberT_n of bits:

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

If the codes c_i were chosen so as to be orthogonal ($d_H(v,w) = 0$), or at least sufficiently different from each other, this entails in particular that we can obtain v_i from s by applying $v'_i = s \oplus c_i$. The result v'_i is so similar to v_i that v_i can then be regenerated from v'_i, e.g., using error correction codes. Codes c_i can be generated so as to be orthogonal, however, when long bit sequences are generated at random, statistical theory predicts that the probability to obtain two sequences of low similarity is the 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 $\oplus$ and also the *circular bitwise shift* are distancing operations, which make their result different from both its operands. CDMA uses the $\oplus c$ encoding to guarantee that the values to be transmitted are not mixed by the 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 v_0 shared by all tags and the receiver. We obtain sufficiently different codes v_i for numerical values i by circularly shifting v_0 by the amount of i bit, since shifting is a distancing operation. In this way, a bit vector $v_0 \in \{0,1\}^n$ can be used to encode n values.

The received signal $s = s_1 + s_2 + \ldots + s_n$ is then simply a sum of encoded numbers v_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 $s = v_7 + v_8 + v_{12}$. The receiver can now check the similarity between s and any value v_i by simply testing $s \sim v_i$. We call this a *binary query*.

In many cases, an estimation of how many tags sent which of the values can be useful. In our scenario, for instance, for checking the amount of damaged products on a pallet. We call this a *proportion query*. One way to do this is least squares estimation.

Due to noise and other problems of collaborative transmission, the model of simultaneous transmission as addition is a highly idealized model. In reality, our algorithms deliver rough approximations. However, we can use statistical methods, such as χ^2 to estimate how reliable our results are and to discard measurements that are too irregular.

The complete resulting algorithm then operates as follows:

1) Tags come initialized with t set to the minimum temperature 0, and transmit code v set to v_0
2) Tags measure their environment continuously over a longer duration: if the measured value $m > t$, then
   a) it sets $t := m$.
   b) it shifts the code v accordingly, that is: set $v := v_t$.[3]
3) Reader sends `start` signal to tags.
4) Tags send their respective v.
5) Reader receives overlayed signal s:
   a) **Binary Query:**
      i) Set $S := \emptyset$.
      ii) For each possible value v: if $v \sim s$ then $S := S \cup \{v\}$.
   b) **Proportion Query:** For each value $v \in S$: use Least Squares Estimation (LSE) to compute proportion of contribution of v:
      i) Generate linear equation system for the found values $v_i \in S$.
      ii) Estimate parameters a_i so that error is minimal.
      iii) Set $M := \{(a_i, v_i) | s = \sum_{v_i \in S} a_i * v_i\}$.
   c) **Output:** return M.

We discuss each step in detail in the next section.

IV. Collective transmission

We can now discuss the details of our implementation of the algorithm. The architecture of our instrumental set-up consists of n wireless sensor nodes (the tags) and a sink node (the reader) processing the received signal (see figure 2). The

[3] In an actual printed electronics implementation these two steps could be combined, e.g., with a destructive, physical temperature sensor that moves a start/end pointer forward through memory.

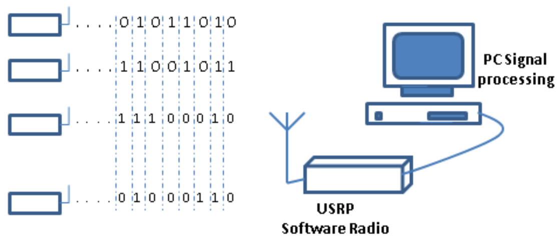

Fig. 2. Principle of the collective information transmission. Each sensor node reached by an external trigger signal is transmitting its binary sequence at the same time. Based on different number of '1' in each time slot different maximal amplitudes are generated. On the receiver side the superimposed binary sequence is captured.

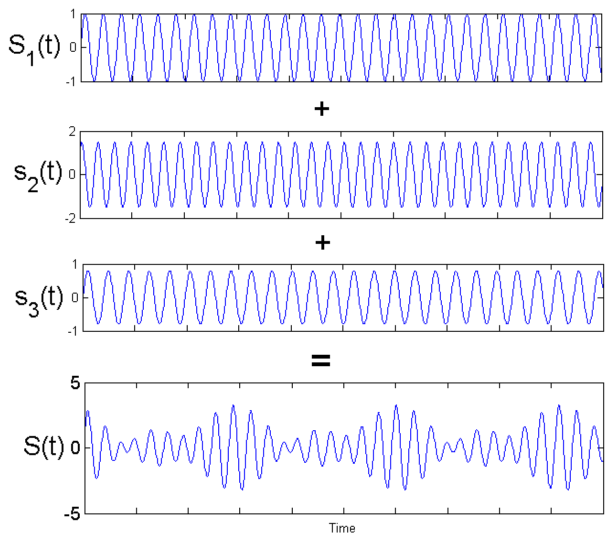

Fig. 3. Illustrates the superposition principle: $s(t)$ is a superimposed signal generated by three sine signals $s_1(t), s_2(t)$ and $s_3(t)$. The sine signals are chosen slightly different from each other in frequency, phase and amplitude strength, i. e. $f_{s_1} = 16$ Hz, $f_{s_2} = 18$ Hz and $f_{s_3} = 20$ 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.

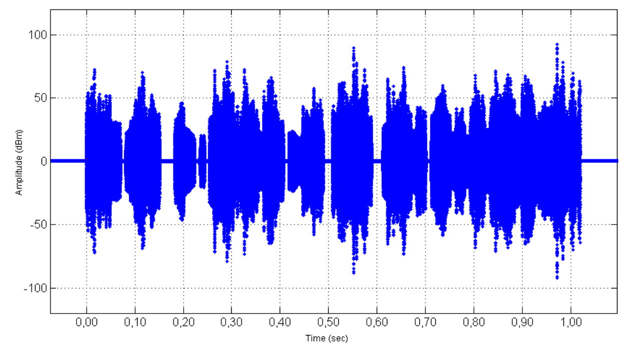

Fig. 4. Raw data of a superimposed signal caused by 21 transducers transmitting different binary sequences simultaneously. The signal length is set by 100 time slots.

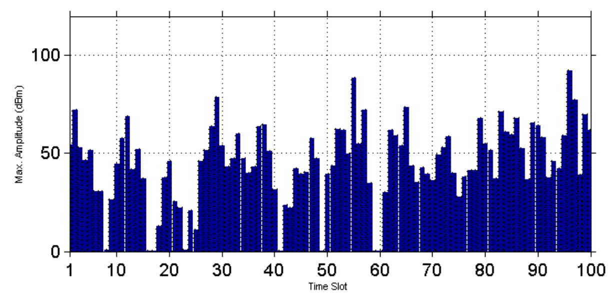

Fig. 5. Illustrates the quantification of the superimposed signal shown in figure 4. In each time slot the maximal amplitude is detected and visualized by a single bar.

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 figure 2 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 superimposition by a receiver. Consequently, the amplitude of the superimposed electromagnetic waves is either intensified or becomes less intense.

In figure 3 an example of a superimposition 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. Therefore, during the transmission of the bit sequences from the n sensor nodes the maximum can vary in each time slot making measurement of the strength of the signal difficult. An example for a received raw signal is depicted in figure 4.

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

For encoding values, we chose a 100-bit-long random vector v_0 in such a way that $v_i \neg \sim v_j$ 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 been sent at all. The proportion query uses this information to additionally compute what percentage of senders have sent a certain value.

A. 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. The generalized Hamming distance d_H between two vectors $v = (v_1, v_2, \dots, v_n), w = (w_1, w_2, \dots, w_n) \in [0,1]^n \subset \mathbb{R}^n$ can be defined as:

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

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 most general definition:

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

The similarity can then be defined by

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

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 for small sets of vectors concerning the addition. The more vectors are used to encode entities the worse the identification. The cause for worse recognition lies in the overlapping and magnitude of the vectors while transmitting them on the MAC. When all time slots are occupied by sent signals of the sensor nodes the Hamming distance becomes inefficient. In other words, the Hamming distance fails to extract the vectors correctly, which are contained in the received superimposed signal.

A more stable test is the Pearson correlation coefficient, which is defined by

$$\rho_{s,y} = corr(s,y) = \frac{E[(s-\mu_s)(y-\mu_y)]}{\sigma_s \sigma_y} .$$

Here, s denotes a vector or a linear combination of vectors that is to be screened in the captured superimposed signal y. μ_s and μ_y are the means and, σ_s, σ_y are the variances of s and y.

The Pearson correlation coefficient has the property of being sensitive only to a linear relationship between two variables, which is the case here: s and y show a high degree of correlation $\rho_{s,y}$, if the signals s and y are strong related to each other, else the correlation is less indicated.

B. Proportion query

Using binary query the following applications can be realized

- detection of an abnormality, for instance, the pallet containing perished goods shows a compliance violation;
- different classes A, B, C indicate different temperature intervals, such as $A = [0...8]\,°\text{C}$, $B = [10...25]\,°\text{C}C = [26...100]$.

In the following, the capabilities of the system are extended by estimating the proportions of the classes A, B, C, e.g. computing class $A = 30\%, B = 60\%, C = 10\%$.

To realize this capability a mathematical formalization of the superposition principle combined with the statistical mechanism is required. Thus, the first modeling step is to collect the vectors v_i 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 $s_1, s_2, \cdots, s_M$ 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

$$z(a_1 s_1 + a_2 s_2 + a_3 s_3 + \cdots + a_M s_M) = y\,, \qquad (1)$$

where the parameters $\{a_1, a_2, a3, \cdots, a_M\} \in \mathbb{R}$ indicate the number of sensor nodes that have sent out the same sensory value associated with the binary sequence $s_i \in \mathbb{B}^N$. The variable $y \in \mathbb{R}^N$ contains the occurred N maximal amplitudes of the N time slots. Here, y represents 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 (1) without considering z, i.e.

$$\hat{a_1} s_1 + \hat{a_2} s_2 + \hat{a_3} s_3 + \cdots + \hat{a_M} s_M = y\,. \qquad (2)$$

In the second step the solution $\hat{a} = (\hat{a_1}, \hat{a_2}, \cdots, \hat{a_M})^T$ 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 $a = (a_1, a_2, \cdots, a_M)^T$ can be estimated as

$$a_i = \hat{a}_i z \text{ for } i = 1, \cdots, M\,.$$

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 $a = (a_1, a_2, a_3, \cdots, a_M)^T$ can be estimated by using linear least squares estimation (LSE). 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.

V. Evaluation

For testing our approach, we explicitly chose an environmental 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 cool 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 testbed.

A. *Experimental setting*

According to the constraints of the printed organic electronics [6], [19], organic electronics will behave and develop very differently from traditional electronics. Thus, for testing purposes, we created thirty transducers on PCB with components off-the-shelf, which conform fully to the requirements of the 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 at this low frequency domain by 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 figure 6 the entire experimental platform is shown, which consists of the mentioned transducers, a loop antenna operating in the low frequency domain and one receiver[4] connected to an ordinary PC, where the computation and visualization of a received transmission is performed.

B. *Results*

To give proper evaluation results to the proposed approach in Sect. IV, we set every transducer to transmit a certain bit sequence corresponding to a certain sensory value. Hence, the evaluation was performed under realistic and controlled conditions. The sensory values in the experiment were fixed to the transducers, but the collective information transmission (step 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 sensory value of transducers were varied.

To create the required vectors for encoding temperature values, we first generate a 100-bit long vector following the Bernoulli process. Afterwards the randomly drawn binary vector was shifted bit-wise to create additional vectors for further temperature values.

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

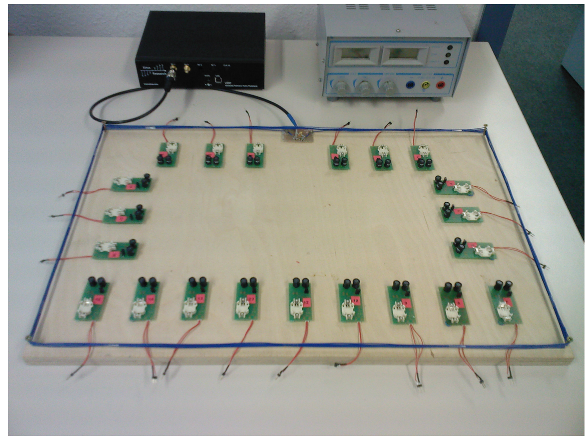

Fig. 6. The shown research platform consists of 21 transducers and one receiver connected to a PC

Due to the small bit-length of the original vector is chosen carefully, having equal number of '1' and '0'. Additionally, the minimal Hamming distance between two possible combinations has been considered.

In our experiments the transducers have been positioned as it is shown in figure 6. Based on seven possible temperatures and our instrumental set-up, we had fifteen different settings to perform. For every setting ten trails have been made and evaluated. The following table shows our evaluation results.

The first column describes the setting by chosen temperatures and transducers. For example a setting of 21 describes the trials where all transducers send the same temperature value and a setting of 9,6,6 means nine transducers send temperature A, six send temperature B and the remaining six send temperature C. Column two states the mentioned number of trials for every setting and column three presents the average amount of correctly recognized temperature values using the binary query algorithm exclusively. Column four shows the average error for each set of trials when the proportion query algorithm was applied, at which the last column reduces this to the average error mean per class.

C. *Discussion*

The experiments have first shown, that collective information transmission is possible, which enables the reading out of a wireless sensor network at once. It has proved that superimposed signals can be used efficiently to transmit data simultaneously by concurrent usage of simple analog sensor nodes. The simple communication scheme is robust as it is shown in Sect. V-B. The different classes of sensory information sent in a collective information transmission are almost all detected by using binary query. Further on, in the proportion query the number of the senders, which have sent the same sensory information are estimated with a small deviation error.

TABLE I
RESULTS

		Binary Query	Proportion Query	
Setting	**Number of Trials**	**Correctly identified Msg.**	**Average Error Sum**	**Average Error Mean p. Class**
21	10	88,57%	12,38%	1,77%
18,3	10	89,05%	11,90%	1,70%
15,6	10	91,90%	8,57%	1,22%
12,9	10	97,62%	2,38%	0,34%
15,3,3	10	90,00%	10,95%	1,56%
12,6,3	10	89,05%	11,43%	1,63%
9,9,3	10	89,05%	10,95%	1,56%
9,6,6	10	82,38%	17,62%	2,52%
12,3,3,3	10	82,86%	17,62%	2,52%
9,6,3,3	10	82,38%	17,14%	2,45%
6,6,6,3	10	80,95%	18,10%	2,59%
9,3,3,3,3	10	80,00%	21,90%	3,13%
6,6,3,3,3	10	85,71%	14,76%	2,11%
6,3,3,3,3,3	10	80,00%	20,95%	2,99%
3,3,3,3,3,3,3	10	79,52%	19,52%	2,79%

VI. CONCLUSION

We presented a novel organic computing approach for *collective transmission*, a robust, collective, approximate communication method for massive amounts of sensor nodes that combines communication with computation on the channel. The aim of this paper was to detail and test an implementation to realize collective read-out. Our experiments prove 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 it lacks robustness to noise and requires exact synchronization of phases. Collective transmission in contrast employs statistical methods and error correction, thus allowing for high tolerance to noise and phase shifts.

The robustness of collective transmission comes from the use of random vector encodings of numerical values. In our example application, simultaneous transmission made it possible to communicate with the pallet as a whole. 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 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 can be fundamental building blocks for realizing distributed intelligence.

REFERENCES

[1] H. Adam, E. Yanmaz, W. Elmenreich, and C. Bettstetter. Estimation of the number of operating sensors in a sensor network contention-based neighbourhood estimation. In *Conference Record of the Thirty-Seventh Asilomar Conference on International Symposium on Information Theory*, 2004.

[2] H. Adam, E. Yanmaz, W. Elmenreich, and C. Bettstetter. Contention-based neighbourhood estimation. In *Proceedings of the IEEE International Symposium on Information Theory*, 2010.

[3] Albert Krohn, Michael Beigl, Sabin Wendhack. SDJS: Efficient Statistics for Wireless Networks. In *Proceedings of the 12th IEEE International Conference on Network Protocols*, Berlin, Germany, 2004.

[4] Albert Krohn, Mike Hazas, Michael Beigl. Removing Systematic Error in Node Localisation Using Scalable Data Fusion. In *Fourth European conference on Wireless Sensor Networks*, 2007.

[5] Albert Krohn, Tobias Zimmer, Michael Beigl, Christian Decker. Collaborative Sensing in a Retail Store Using Synchronous Distributed Jam Signalling. In *International Conference on Pervasive Computing*, pages 237–245, Munich, Germany, 2005. Springer Verlag.

[6] Organic Electronics Association. OE-A Roadmap for Organic and Printed Electronics. *White Paper*, 2008.

[7] Mario Goldenbaum and Slawomir Stanczak. Computing the Geometric Mean Over Multiple-Access Channels: Error Analysis and Comparisons. In *Proc. 43rd Asilomar Conference on Signals, Systems and Computers*, Monterey, USA, November 2009.

[8] Mario Goldenbaum and Slawomir Stanczak. Computing Functions via SIMO Multiple-Access Channels: How Much Channel Knowledge Is Needed? In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '10)*, pages 3394–3397, Dallas, USA, March 2010.

[9] Mario Goldenbaum, Slawomir Stanczak, and Michal Kaliszan. On Function Computation via Wireless Sensor Multiple-Access Channels. In *Proc. IEEE Wireless Communications and Networking Conference (WCNC '09)*, Budapest, Hungary, April 2009.

[10] Pentti Kanerva. Hyperdimensional computing: An introduction to computing in distributed representation with high-dimensional random vectors. *Cognitive Computation*, 1(2):139–159, 2009.

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

[12] J. Laneman, G. Wornell, and David Tse. An efficient protocol for realising cooperative diversity in wireless networks. In *Proceedings of the IEEE International Symposium on Information Theory*, page 294, 2001.

[13] J. Lanemann, David Tse, and G. Wornell. Cooperative diversity in wireless networks: Efficient protocols and outage behaviour. *Transactions on inforamtion Theory*, 50(12), December 2004.

[14] Larry Miller and Janet Allen. National Supermarket Shrink Report. Technical report, National Supermarket Research Group, 2005.

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

[16] Paul Beswick, Mathew Isotta, Stefan Winter. A Retailer's Recipe for Fresher Food and Far Less Shrink. In *Oliver Wyman Journal Issue 24*, 2008.

[17] Stephan Sigg, Rayan Merched El Masri, and Michael Beigl. A sharp asymptotic bound for feedback based closed-loop distributed adaptive beamforming in wireless sensor networks. *Transactions on mobile computing*, 2011 (accepted for publication).

[18] David Tse and Pramod Viswanath. *Fundamentals of wireless communication*. Cambridge University Press, 2005.

[19] Vivek Subramanian, Josephine B. Chang, Alejandro de la Fuente Vornbrock, Daniel C. Huang, Lakshmi Jagannathan, Fank Liao, Brain Mattis, Steven Molesa, David R. Redinger, Daniel Soltman, Steven K. Volkman, Qintao 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.