Long-range Low-power Wireless Networks and Sampling Strategies in Electricity Metering

Mauricio C. Tomé, *IEEE Student Member*, Pedro H. J. Nardelli, and Hirley Alves, *IEEE Member*

Abstract—**This paper studies a specific low-power wireless technology capable of reaching a long range, namely LoRa. Such a technology can be used by different applications in cities involving many transmitting devices while requiring loose communication constrains. We focus on electricity grids, where LoRa end-devices are smart-meters that send the average power demanded by their respective households during a given period. The successfully decoded data by the LoRa gateway are used by an aggregator to reconstruct the daily households' profiles. We show how the interference from concurrent transmissions from both LoRa and non-LoRa devices negatively affect the communication outage probability and the link effective bitrate. Besides, we use actual electricity consumption data to compare time-based and event-based sampling strategies, showing the advantages of the latter. We then employ this analysis to assess the gateway range that achieves an average outage probability that leads to a signal reconstruction with a given requirement. We also discuss that, although the proposed analysis focuses on electricity metering, it can be easily extended to any other smart city application with similar requirements, like water metering or traffic monitoring.**

Index Terms—**event-based sampling, stochastic geometry, low-power wide-area wireless networks**

I. INTRODUCTION

The Internet of Things (IoT) denotes the widespread deployment of communication networks between machines without direct human intervention [1]. One important IoT application is in the monitoring of different aspects within cities to help their complex and distributed management [2]. Household management of electricity consumption is a good example case, where IoT allows for a better understanding of how the electricity is consumed, potentially indicating ways to improve its usage efficiency in both individual and aggregate levels [3].

Beyond this, IoT allows for regulatory actions into the physical system based on information [4]. For instance, the utility – informed by IoT devices – may directly control few households' heating devices – via IoT devices – aiming at monetary savings [5]. Appliances connected to the same network may directly coordinate their reactions in respect to

the grid frequency to help balancing supply and demand (e.g. fridges postponing or anticipating their cycles without creating more instabilities [6]).

The term IoT, however, is also very broad, covering extreme application cases: from massive deployment with loose requirements (e.g. air quality measurements) to very specific high-reliability low-latency applications (e.g. robot arms in fully autonomous industrial plants) [7]. In this paper, we target an application with relatively loose requirements related to the communication system, namely electricity metering in households. Specifically, the scenario under investigation consists in a typical household that sends its average demanded power during a given period to an aggregator node (which can be the utility or a micro-grid trader, depending on the distribution arrangement in question) [8], [9]. At the end of each day, the aggregator wants to reconstruct the power demand curve from the successfully decoded samples transmitted by the household. The aggregator may use this information to, for example, plan for the next day operations (e.g. [10]) or profiling consumers (e.g [11]); although the use actually defines the required quality of service provided by the communication network (e.g. [12]), we keep it here unspecified and only look at the relation between the performance of the communication link and the quality of the signal reconstruction.

As to be explained throughout this paper, our contribution to the topic is the following.

- We extend the Long Range (LoRa) technology study case from [13] by (a) including in their proposed stochastic geometry model a density of external interferers, and (b) using a random spreading factor allocation, which is fairer in relation to our specific application.
- The theoretical model developed in [8] is modified to incorporate realistic LoRa setups from [13] (which is supported by actual deployments as discussed in [7].
- We reproduce the sampling strategies from [9] and [14] by testing their proposed approach in a different dataset [15] and providing further comparisons between the time-based and event-based schemes. As in [9], we show that, for most households evaluated, the latter option consistently provides lower reconstruction errors when the number of samples are similar.
- We investigate the trade-off involved between the system variables pointing out what should be the most suitable gateway range to achieve a communication outage probability that can sustain a signal reconstruction by the aggregator within a given quality level.

The rest of this paper is divided as follows. Sec. II provides a short overview of the relevant literature about LoRa and stochastic geometry applied in wireless networks, mainly discussing our main advancements. Sec. III contains the system model based on stochastic geometry and metrics to evaluate the performance of the proposed LoRa deployment, as well as the respective numerical results. We present the sampling strategies and their assessment in Sec. IV. Sec. V discuss about the implications of these results focusing on deployment aspects in smart cities, while Sec. VI concludes this paper, also listing possible research directions.

II. RELATED WORK

A. Long Range Technology – LoRa

Low power wireless technologies covering wide areas are becoming trendy nowadays, as indicated by the survey [7]. The reason for that is the potential to reach the massive number of IoT devices at low cost with reasonably efficient performance. It is worth saying, however, that the so-called "low-power wide-range" technologies are well suited to the massive machine-type Communications (MTC) in contrast to the critical MTC, whose applications require (very) low latency (1-10 ms) and (ultra) high reliability (99.999%) [16]. The main advantages of Low Power Wide Area (LPWA) technologies are, beyond the long range and the low power consumption themselves, their low cost and scalability while guaranteeing some (not so strict) quality of service. For some specific applications as presented in [12], LPWA may become a way to alleviate the data traffic in traditional cellular networks to avoid some problems presented in [17].

Among other options discussed in [7], Long Range (LoRa) – a proprietary technology proposed by Semtech and promoted by the LoRa Alliance – provides bidirectional communication based on chirp spread spectrum modulation that spread the narrow-band signals over a wider bandwidth. LoRa uses a star-of-stars topology consisting of end-devices (in our case, smart-meters), gateways and a central network server. The end-devices directly send their message via a wireless channel with six different spreading factors (SF7 to SF12) using unslotted Aloha as its medium access control; in Europe the spectrum used by LoRa is in the 863-870 MHz ISM Band range with channels of bandwidth of 125 kHz. The communication from the gateway to the server occurs through non-LPWA networks (e.g. cellular or Ethernet). End-devices are divided into three classes (A, B and C) that are most related to their downlink capabilities. Aspects relevant to this article, such as data-rate and successful detection at the uplink, will be given in the next section, while more general details about LoRa can be found at [7], [18] and references therein.

B. Stochastic geometry for wireless networks

Communication engineering communities have quite established analytical frameworks to account for channel and/or traffic uncertainties, which are the basis of almost all (if not all) wireless communication systems [19]. The uncertainties related to the relative positions between devices are, however, still under development; positions have been traditionally modeled as regular grid-like topologies (e.g. hexagonal or square

cellular networks) or toy-models with few communication devices (e.g. two-hop systems). As a growing research field, stochastic geometry and spatial point process theories applied in wireless networks (e.g. [20]) are able to capture many trade-offs involved in the system design and deployment explicitly including the uncertainties related to the devices' positions. Their main advantage is that the aggregate interference becomes analytically treatable, so neither time-consuming Monte Carlo simulations nor (over-)simplifying assumptions about the aggregate interference (e.g. completely neglecting it) are needed in order to evaluate the system performance.

Our contribution here is mainly based on three papers that employ such a model. In [13], the authors studied whether LoRa can scale considering Spreading Factor (SF) allocations based on distance. Their results indicate a exponential dependence of the number of end-devices and the outage probability, which is the complement of the probability that a given message is correctly received by the gateway. They show that success probability exponentially decays with the average number of devices, evincing the negative effects of interference in LoRa even with its mitigation techniques. But, the most interesting aspect of this paper is the stochastic geometry treatment of interference. Assuming that end-devices follow a Poisson point process and making use of order statistics, the authors found a closed-form expression for the probability that an outage event being caused by another concurrent end-device using the same SF. This outage event is defined when the received power of a given desired signal is not four times (i.e. 6 dB) stronger than any other concurrent transmission in the same SF. In any case, the effect of the aggregate interference from all users using the same SF is not considered, but only the effect of the dominant (stronger) interferer.

In the other two articles [8], [9], the authors propose a general optimization of the link throughput based on Shannon capacity, where smart meters and aggregators (acting as a base-station, a gateway) are unlicensed users of the uplink channel of the cellular network in a spectrum sharing setting. Their proposed scenario only considers the effects of the interference from the licensed mobile devices (modeled as a Poisson point process) in the aggregator reception, showing that the maximum achievable throughput actually happens with a less strict outage constraint. This higher outage constraint, however, is shown to have a small effect in the average power demand signal reconstruction by the aggregator. We find, nevertheless, that these results are still very analytical, only relying on an abstract conception of cognitive radio and spectrum sharing without specifying any possible technology.

C. Sampling strategy

Another interesting result from [9] is the comparison between time- and event-based sampling, indicating the first may lead to redundant samples for most of the day, while not good enough during other times. The idea of event-based sampling for sensor networks has been discussed for some time (e.g. [21], but only very recently adopted in electricity metering scenarios (e.g. [14], [22]). However, only in [9] the joint performance between sampling and the link outage probability is analyzed (but with the previously discussed limitations).

D. Relation to this contribution

This paper is mainly built upon [8], [9], [13]. In relation to [13], we extend the stochastic geometry modeling by considering outage events due to not only the dominant interferer, but also the aggregate interference in a given SF. Besides, instead of outage caused by Gaussian noise, we consider the aggregate interference of non-LoRa devices by explicitly including a density of interferers that transmit in the same channel. We then have another variable affecting the outage probability when the non-LoRa interference is treated as noise [23].

This approach follows [8], but specifies it to an actual LoRa deployment. In other words, we move from Shannon capacity in bit/s/Hz and their abstract conceptual model (even including highly directional antennas that leads to negligible unlicensed-to-licensed and unlicensed-to-unlicensed user interference). We analyze here a LoRa scenario following the system setting from [13], which is based on the technology specification and several deployment trials (refer to [7] and references therein). Besides, the time-based and event-based schemes employed in [9] are implemented and tested for another electricity consumption database [15], which has different granularity and household composition as well as geographical location.

III. LoRa deployment

A. System model and performance metrics

Fig.1 depicts an illustrative network deployment. We analyze here the LoRa deployment from [13] with some differences to be explained in the following. We assume that smart-meters need to transmit to a single gateway a 25 byte message in a bandwidth of 125 kHz containing the cumulative energy consumption in a given period of time as well as the period. The meters' locations are randomly distributed in the plane following a Poisson point process Φ_{SF} with density λ_{SF} devices/km^2 [20]. When a given smart-meter wants to send a message to the gateway, the LoRa network server assigns to it a spreading factor (SF), which is randomly and independently allocated from SF7 to SF12. So a smart-meter has a probability of $p_{\mathrm{SF}} = 1/6$ to be allocated to a specific SF, which is independent across the points from λ_{SF}.

The smart-meter traffic is related to the sampling strategy, which will be discussed later in Section IV. Regardless of the strategy adopted, the actual transmission is randomized to decrease the number of concurrent transmissions. A duty cycle limitation of 1% must be also assumed so that each smart meter has a limited number of wireless transmissions per day. The actual density of smart-meter concurrently transmitting is then (much) smaller than λ_{SF}. We assume the density of active smart-meters as $p_0\lambda_{\mathrm{SF}}$, where p_0 is the probability that any smart-meter from Φ_{SF} is active while a given reference link is also transmitting. Using the point process theory nomenclature, the mapping from λ_{SF} to a process in which some of the original points were erased is known as *thinning*.

As the transmissions use LoRa, an outage at the reference link occurs when the other active smart-meters using the same SF lead to a signal-to-interference ratio (SIR) at the gateway in respect to the reference link below a given threshold β_{SF}. As we assume a random SF allocation with probability p_{SF},
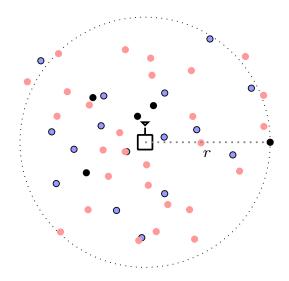

Figure 1. Illustrative figure of the network deployment where LoRa end-devices (smart meters in this paper) are the black and blue circles, while the LoRa gateway is the square node at the center. The black circle linked by a dashed line to the gateway is the reference link, located r km away. The other black circles represents LoRa devices using the same spreading factor (SF) as the reference link, and the blue ones are LoRa devices using other SFs. LoRa devices cause interference to each other only if they are transmitting at the same time with the same SF (represented here by the black circles). Note that LoRa devices may also suffer interference from non-LoRa devices (depicted by red circles). Since these devices use different radio access technology, but in the same frequencies, they are treated as noise by the LoRa gateway.

Table I
LoRa setting adapted from [13]

Spreading factor x	Bit-rate R_{SFx} in kb/s	Minimum SIR β_{SFx}
7	5.47	0.25
8	3.13	0.125
9	1.76	0.06
10	0.98	0.03
11	0.54	0.017
12	0.29	0.01

another thinning happens, resulting in a density $\lambda = p_{\mathrm{SF}}p_0\lambda_{\mathrm{SF}}$ of active smart-meters using the same SF.

We consider that the wireless channel is composed of two components: a distance dependent path-loss with exponent $\alpha > 2$ and a gain due to multipath. We assume the basic path-loss equation, where the received power is proportional to $r^{-\alpha}$ with r being the distance between the smart-meter and the gateway, while the multipath is modeled as independent and identically distributed channel gains related to a Rayleigh fading distribution with unity mean. Then, we can compute the outage probability $P_{\mathrm{out:1}}$ from smart-meter as [8], [20]:

$$P_{\mathrm{out:1}} = 1 - \exp\left(-k\lambda r^2(\beta_{\mathrm{SF}})^{2/\alpha}\right), \qquad (1)$$

where $k = \pi\Gamma(1+1/\alpha)\Gamma(1-1/\alpha)$ with $\Gamma(\cdot)$ being the Gamma function.

In addition to the outages caused by the concurrent LoRa (smart-meter) transmissions, we consider here outages from other (non-LoRa) users that use the same channel. Different from [13] and following [8], we assume the noise is negligible compared to the interference. If the non-LoRa devices' positions are modeled as a Poisson point process Φ_I with density λ_I devices/km^2 and the aggregate interference is treated as noise by the gateway [23], the outage probability $P_{\mathrm{out:2,x}}$ in a link using a spreading factor x with x = 7, ..., 12, with a respective SIR threshold β_{SFx}, is given by:

$$P_{\mathrm{out:2,x}} = 1 - \exp\left(-k\lambda_I r^2 (\beta_{\mathrm{SFx}})^{2/\alpha}\right), \qquad (2)$$

where β_{SFx} is different for each SF (refer to Table I).

Remark: Both (1) and (2) assume that all users transmit with the same power. Different transmit power, or even some kind of channel inversion, may also be incorporated into the proposed formulation, as discussed in, for example, [8] and [20]. Although these differences would affect the overall link performance, it would not change its qualitative behavior in relation to the density of interferers and to the smart-meter-gateway distance, which are our focus in this paper.

The outage probability that the reference link experiences when their transmission occur using a spreading factor x can be computed as the complement of a successful transmission probability: $1 - (1 - P_{\mathrm{out:1}})(1 - P_{\mathrm{out:2,x}})$. Using (1) and (2),

$$P_{\mathrm{out,x}} = 1 - \exp\left(-kr^2\left(\lambda(\beta_{\mathrm{SF}})^{2/\alpha} + \lambda_I(\beta_{\mathrm{SFx}})^{2/\alpha}\right)\right). \quad (3)$$

As the SF allocation is random and independent, the average outage probability P_{out} can be computed as follows:

$$P_{\mathrm{out}} = p_{\mathrm{SF}} \sum_{\mathrm{x}} P_{\mathrm{out,x}}, \qquad (4)$$

remembering that, for LoRa, x = 7, .., 12 and $p_{\mathrm{SF}} = 1/6$ when the allocation is uniform across the SFs.

Besides, each SFx has a different bit-rate R_{SFx}, as shown in Table I, This, however, does not include outages and it might be interesting to evaluate the effective bit-rate, here defined as $(1 - P_{\mathrm{out,x}})R_{\mathrm{SFx,eff}}$. Similar to (4), we can evaluate the average effective bit-rate $R_{\mathrm{SF,eff}}$ as:

$$R_{\mathrm{SF,eff}} = p_{\mathrm{SF}} \sum_{\mathrm{x}} (1 - P_{\mathrm{out,x}})R_{\mathrm{SFx,eff}}. \qquad (5)$$

B. Numerical results

We present here the numerical results assuming: a path-loss exponent $\alpha = 4$ (urban environment), a SIR threshold related to LoRa end-devices using the same SF $\beta_{\mathrm{SF}} = 4$ (6dB) and a thinning probability $p_0 = 0.025$ related to the traffic and duty cycle constraint; the numerical setting related to each different SF is given in Table I. When not explicitly mentioned, we assume the distance between the reference smart-meter to the gateway to be $r = 1.5$ km, while the density of active non-LoRa devices to be $\lambda_I = 0.05$ devices/km^2. These two values were arbitrarily chosen since they do not qualitatively change the analysis [20]; the way they affect the link performance will be presented when discussing Fig. 6 and 7.
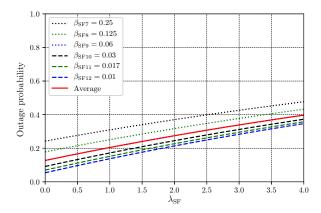

Figure 2. Outage probability as a function of the density λ_{SF} of users served by the gateway for the different spreading factors (SFs) and $r = 1.5$km. The red curve is its average considering $p_{\mathrm{SF}} = 1/6$.
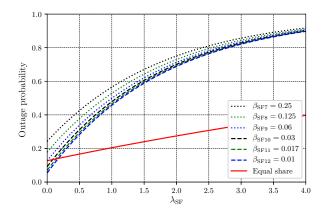

Figure 3. Outage probability as a function of the density λ_{SF} of smart-meters served by the gateway for the different SFs and $r = 1.5$km. The red curve is its average considering $p_{\mathrm{SF}} = 1/6$ while the other curves consider that all users are allocated only to one SF.

Fig. 2 shows how the density λ_{SF} of smart-meters affects the outage probability $P_{\mathrm{out,x}}$ for each possible SFx and its respective average P_{out}. Although the relation between these variable shown by (3) is exponential, the range of interest does not present a steep behavior, regardless of the SF. This effect comes with the thinning processes related to the active concurrent transmissions at the same SF. Therefore, the effects of the smart-meters' interference is not so dramatic.

Fig. 3 reinforces this idea by showing the link outage probability when *all* smart-meters are allocated in the same SF. This case, compared with the average using an equal share (i.e. $p_{\mathrm{SF}} = 1/6$, shows the steep exponential behavior so that the outage probability grows much faster with λ_{SF}. Besides, from both figures, we confirm that the lower the spreading factor, the higher the outage probability for the same setting. This fact is expected since higher SFs imply lower SIR thresholds for successful reception, obtained at expense of the link bit-rate.

Fig. 4 shows the effective bit-rate for the different SFs and its average from (5). As in Fig. 2, one can see a smooth decrease (not steep as expected in exponential relations) in respect to λ_{SF}, regardless of the SF. The link can nevertheless transmit, in average, with a bit-rate between 1 and 2 kb/s.
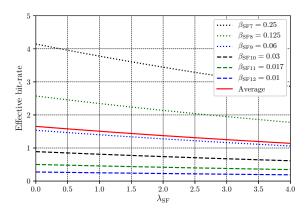
Figure 4. Effective bit-rate as a function of the density λ_{SF} of smart-meters served by the gateway for the different SFs and $r = 1.5$km. The red curve is the average outage probability considering that the end-users are randomly assigned to a specific SF with probability $1/6$.
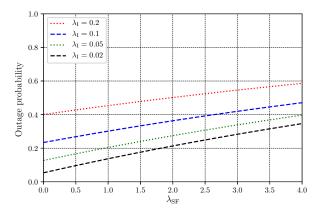

Figure 6. Average outage probability considering a random allocation of SF with probability $1/6$ as a function of the density λ_{I} of non-LoRa devices for different densities λ_{SF} of users served by the gateway and $r = 1.5$km.

Figure 5. Average outage probability with a random allocation of SF with probability $1/6$ as a function of the density λ_{SF} smart-meters served by the gateway for different densities λ_{I} and $r = 1.5$km.

Figure 7. Outage probability vs. the distance r for $\lambda_{\text{SF}} = 0.5$ and $\lambda_{\text{I}} = 0.05$. The red curve is the average outage probability considering the smart-meters are assigned to a specific SF with probability $1/6$.

Figs. 5 and 6 present how the average outage probability changes with the density λ_{I} of non-LoRa devices. For the range studied here, we can see that, in the worst case scenario with $\lambda_{\text{SF}} = 4$ and $\lambda_{\text{I}} = 0.2$ devices/km^2, the average outage is 60% (a relatively high value). Similar to the preset noise level in [13], the interference from non-LoRa devices imposes a floor level in the outage probability, but dependent on another parameter, namely λ_{I}. In specific terms, the higher λ_{I}, the higher the outage floor, as indicated by (2). Consequently, if the density of non-LoRa devices in a given region is high enough, LoRa deployments will experience a very poor performance, specially for long range links.

To evaluate the effect of the distance on the link performance, Fig. 7 shows the average outage probability when the distance between the reference smart-meter and the gateway changes. As stated in (3), the relation is exponential and depends on r^2, so growing r implies in a steep increase of outage events. This plot is important for the network deployment when considering the sampling strategies since it helps to determine the worst case scenario so the most suitable position of the gateway can be chosen so as to achieve a minimum quality related to the signal reconstruction.
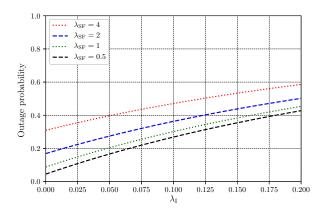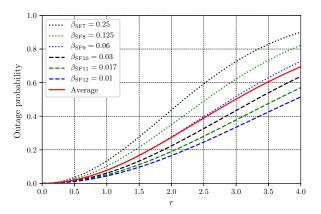
For example, if the gateway is planned to have a range of $r = 4$ km, the worst case average outage probability is about 70%, which probably lead to a poor signal reconstruction (to be assessed in Section IV). This plot may also indicate that the SF allocation strategy used in [13] based on distance ranges may outperform the random strategy. However, from Fig. 2, this is not so obvious and requires further studies, since there is a clear trade-off involved between sharing the spectrum and the link distance. This will be further discussed in Sec. V.

IV. Sampling strategies

We followed here the results presented in [9], but considering a different database [15], which is comprised of around 400 houses with measurement lengths ranging from a few days to a whole year with 10 minutes granularity. The sampling strategies chosen are (i) time-based with a sampling frequency of 30 minutes, and (ii) event-based as described in [14]. While the time-based implementation is straightforward, the other depends on some simple processing to identify a prescribed event. Following [9], [14], the event is defined by two situations: (a) a certain amount of energy consumption E_{lim} is reached; or (b) a sudden change in the power demand denoted by P_{lim} is detected.

Algorithm 1 Event-based setting

$P_{\mathrm{lim}} \leftarrow$ Set Power threshold
$P_{\mathrm{step}} \leftarrow$ Set Min Power Increase
$E_{\mathrm{lim}} \leftarrow$ Set Energy threshold
$OK \leftarrow$ False
while OK is False **do**
    $measEvent \leftarrow EventMeasuring(P_{\mathrm{lim}}, E_{\mathrm{lim}})$
    if $len(measEvent) \geqslant len(measTime)$ **then**
        $P_{\mathrm{lim}} \leftarrow P_{\mathrm{lim}} + P_{\mathrm{step}}$ ▷ Increase threshold
        if $sum(measEvent('Power')) == 0$ **then**:
                     ▷ Power threshold set too high
            $E_{\mathrm{lim}} \leftarrow 2 * E_{\mathrm{lim}}$ ▷ Increase energy limit
            $P_{\mathrm{lim}} \leftarrow P_{\mathrm{step}}$ ▷ Reset threshold
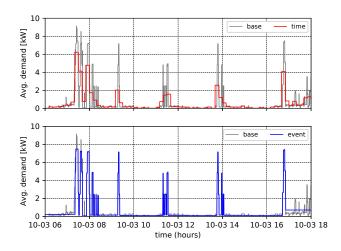   **else**
        $OK \leftarrow$ True

Figure 8. Comparison between the actual measurements and sampled signals for a single house from 6:00 to 18:00 on October 3, 2016 (x-axis is presented as mm-dd hh). Top: Time-based; Bottom: Event-based.

Here, the initial parameters for the event-based approach were set as $E_{\mathrm{lim}} = 2$ kWh, $P_{\mathrm{lim}} = 1$ kW, with increments of $P_{\mathrm{step}} = 0.5$ kW. These values were arbitrarily chosen to lead to a smaller or equal number of samples compared to the time-based approach (48 samples per day). In the households that the event-based approach leads to more samples than the time-based, we implemented a simple procedure (presented in Alg. 1) to modify the parameters in order to achieve a similar number of samples between the two cases. Such procedure is neither optimal nor exhaustive; however, it builds a smaller set of measurements that is suitable for the present study. On average, a reduction of about 17% in the total number of measures was observed, but depending on the consumption patterns of the houses, the reduction ranges from about 70% (14 measures per day on average) to 0 (no reduction).

Fig. 8 shows an example of the two strategies compared to the original measurements. The results illustrate a single household between 6:00 and 18:00 on Oct. 3, 2006. Due to the thresholds chosen for plotting the event-based strategy, some smaller peaks were skipped (the power threshold is related to power variation, not to the absolute values). Remember also that the amount of points generated with these thresholds is designed to be similar to the time-based strategy's amount.
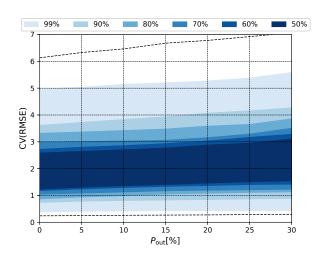

Figure 9. CV(RMSE) vs. the outage probability for the time-based strategy. Colored regions indicate the percentage of houses which belong to the range. Dashed lines indicate the extreme values.
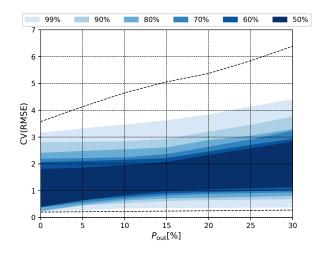

Figure 10. CV(RMSE) vs. the outage probability for the event-based strategy. Colored regions indicate the percentage of houses which belong to the range. Dashed lines indicate the extreme values.

We ran the same procedure illustrated in Fig. 8 for 350 houses for a full week to assess how outage events from the LoRa system would affect the curve reconstruction by the aggregator, following [8], [9]. The outage probabilities were varied from 0 to 30% so the aggregator observes less points, which are lost with such a probability (note that this value comes from the LoRa specification, as discussed in Sec. III). At the end of each day, the aggregator makes a linear interpolation between subsequent points to reconstruct the average power curve. In the case that a sample is lost, the reconstruction error is computed as the relation between the interpolated point and the original (from the database). The error analysis is based on 100 simulations for each household. The quality metric was the root-mean-square error (RMSE) error between the measurements, normalized by the consumption of each one of the houses, herein called CV(RMSE). While the measurements cannot be used to directly compare two houses with reasonably different consumptions, it provides a fair tool to compare the different sampling strategies as well as compare the quality degradation due to different values of outage.
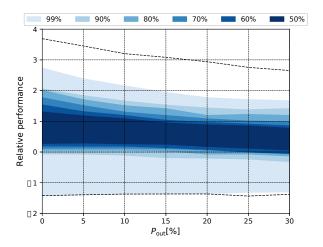
Figure 11. Relative performance between time-based and event-based strategies. Results above zero imply the event-based strategy is better (i.e. smaller error), while below zero the time-based strategy is better.

Figs. 9 and 10 show the results for the time-based and event-based strategies. The color bands are the percentage of houses that fall into that range, whereas the dotted lines indicate the extreme values. One can see that the time-based strategy is less sensitive to outage events, that is, its CV(RMSE) value has small changes as the outage probability grows, but its performance is overall poor when compared to the event-based strategy. This indicates the time-based sampling is more redundant in many periods (e.g. during the night, when the electricity demand tends to be minimal) and missing points can be successfully reconstructed. Conversely, the event-based samples tend to contain more information so missing points have a bigger impact in the signal reconstruction.

Fig. 11 compares the two strategies by subtracting the errors from both measurements to evaluate their relative performance. The positive values indicate that the event-based strategy outperform the time-based. It is possible to see that the event-based strategy provides better reconstruction in almost 90% of the cases, regardless of the amount of error involved. On the other hand, as previously discussed, this advantage tends to decrease if the outages become too frequent. It is important saying that this effect is a byproduct of the design of the event-based strategy where the number of samples cannot grow indefinitely, but is limited to about 48 per day (the number of time-based samples). With such a limited number of samples, the information of each measurement grows.

V. Discussions

The results presented in the previous two sections showed that: (i) the outage probability is affected by the interference and the smart-meter-gateway distance and (ii) the event-based strategy usually outperforms the time-based in terms of reconstruction error. Now, let us consider the following situation: a communication engineer needs to deploy LoRa gateways in a city so that the reconstruction error is within a given quality limit. From the proposed scenario, the engineer can only decide about the gateway range and the sampling strategy; the other parameters like density of smart-meters and non-LoRa devices, as well as the power demand, are given.

For example, a distribution company requires that the quality of the reconstructed signals by the aggregator has, in the worst case, a CV(RMSE) of 4 for the 90% of the households under its coverage. Using the event-based scheme, this performance can be achieved with an outage probability up to 30%, as determined by Fig. 10. If the density of smart-meters is 0.5 and of non-LoRa devices is 0.05, then we can use the red curve from Fig. 7 to find the gateway shall cover a range of 2 km, at most.

This simple case illustrates how a LoRa communication system together with a "smarter" sampling strategy can be used to deploy electricity metering in cities. Other important feature introduced by the event-based approach is an inherent peak detection in relation to variations in power demand, which may be useful to identify unusual or critical situations. It is worth saying that, although the results are based on realistic numbers and on actual electricity demand data, they are used here to test a concept that must be further analyzed and optimized. The same methodology may be directly applied to, for instance, water and gas metering since the consumption patterns are similar [24]. Another interesting use case might be related to commuting and traffic, whose daily patterns are somehow similar to electricity demand (e.g. [24]); this, however, would require a more thorough study about the event definition and the gateway points. In general, despite the particularities of each specific application, we understand that LoRa combined with an event-based sampling strategy provides a scalable solution for massive machine-type communications and IoT deployments needed in the future smart cities.

VI. Conclusions

This paper studied a LoRa wireless network deployment for electricity metering, where this technology combined with a event-based metering strategy led to a fairly good quality signal reconstruction. We presented here new results, which provide advances in the topic as follows.

We extended the LoRa analysis of [13] by using a stochastic geometry approach that includes LoRa and non-LoRa devices as interferers. We employed a randomized (distance-independent) spreading factor allocation without favoring devices closer to the gateway, providing a fairer allocation for the specific metering application. Note that other applications where fairness is not an issue would best suit with the usual distance-dependent allocation (e.g. temperature monitoring in cities). Therefore, we see our results as complementary to and consistent with [13].

Our study also moved beyond [8], [9] by considering a LoRa technology in contrast to the previously developed cognitive radio approach based on abstract Shannon limits of interference-limited networks and perfect directional antennas. While [8], [9], focused on optimizing the communication system performance under such idealized conditions, we assume LoRa specifications to serve as guidance for actual deployments. Nevertheless, although the present model and our objectives are quite different from [8], [9], the system performance (evaluated in terms of outage) is still limited by the same factor, namely co-channel interference.

Besides, we reinforced the strength of the event-based metering introduced by Simonov et al. [14], [22] (also used in [9]) by showing that a better signal reconstruction can be consistently achieved in comparison with the time-based one using a different dataset (i.e. different demand profiles and data granularity). To reach this result, we proposed a simple algorithm that defines the events' thresholds from the consumption data (which are determined by each different household) so the number of samples generated by the event- and time-based are approximately the same. This algorithm is general and can be used to determine the thresholds based on historical data, and can be easily adapted to provide "real-time" adjustments.

All in all, we argue that the proposed approach may be used in planning actual deployments by (*i*) defining event-based metering using thresholds set based on historical data, and (*ii*) defining LoRa gateways' locations since their range can be directly related to the signal reconstruction. Although the results are quite abstract at this stage, we plan to develop this framework to real case-studies (ranging from dense cities to remote rural zones), as well as assess its feasibility to other smart city applications, as water metering and heating.

References

[1] J. Manyika et al., "The Internet of Things: Mapping the value beyond the hype," *McKinsey Global Institute*, no. June, p. 144, 2015.
[2] A. Zanella et al., "Internet of Things for Smart Cities," *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 22–32, 2 2014.
[3] P. Siano, "Demand response and smart grids - A survey," *Renewable and Sustainable Energy Reviews*, vol. 30, pp. 461–478, 2 2014.
[4] F. Kühnlenz and P. H. J. Nardelli, "Dynamics of Complex Systems Built as Coupled Physical, Communication and Decision Layers," *PLOS ONE*, vol. 11, no. 1, p. e0145135, 1 2016.
[5] P. Palensky and D. Dietrich, "Demand side management: Demand response, intelligent energy systems, and smart loads," *IEEE Transactions on Industrial Informatics*, vol. 7, no. 3, pp. 381–388, 8 2011.
[6] J. Evora et al., "Swarm Intelligence for Frequency Management in Smart Grids," *Informatica (Netherlands)*, vol. 26, no. 3, pp. 419–434, 2015.
[7] U. Raza, P. Kulkarni, and M. Sooriyabandara, "Low Power Wide Area Networks: An Overview," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 855–873, 2017.
[8] P. H. Nardelli et al., "Maximizing the link throughput between smart meters and aggregators as secondary users under power and outage constraints," *Ad Hoc Networks*, vol. 41, pp. 57–68, 5 2016.
[9] M. C. Tomé et al., "Joint sampling-communication strategies for smart-meters to aggregator link as secondary users," in *2016 IEEE International Energy Conference (ENERGYCON)*. IEEE, 4 2016.
[10] Z. Ma et al., "The Role of Data Analysis in the Development of Intelligent Energy Networks," *IEEE Network*, vol. 31, no. 5, pp. 88–95, 2017.
[11] R. Li, F. Li, and N. D. Smith, "Load Characterization and Low-Order Approximation for Smart Metering Data in the Spectral Domain," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 3, pp. 976–984, 6 2017.
[12] Z. Dawy et al., "Toward Massive Machine Type Cellular Communications," *IEEE Wireless Communications*, vol. 24, no. 1, pp. 120–128, 2 2017.
[13] O. Georgiou and U. Raza, "Low Power Wide Area Network Analysis: Can LoRa Scale?" *IEEE Wireless Communications Letters*, vol. 6, no. 2, pp. 162–165, 4 2017.
[14] M. Simonov, "Hybrid scheme of electricity metering in smart grid," *IEEE Systems Journal*, vol. 8, no. 2, pp. 422–429, 6 2014.
[15] J. P. Zimmermann, "End-use metering campaign in 400 households In Sweden Assessment of the Potential Electricity Savings," Enertech. Eskilstuna: Swedish Energy Agency, Tech. Rep., 2009. [Online]. Available: http://www.energimyndigheten.se/globalassets/statistik/festis/elmatning-i-bostader/final_report.pdf
[16] P. Popovski, "Ultra-Reliable Communication in 5G Wireless Systems," in *Proceedings of the 1st International Conference on 5G for Ubiquitous Connectivity*. ICST, 2014.
[17] G. C. Madueno et al., "Assessment of LTE Wireless Access for Monitoring of Energy Distribution in the Smart Grid," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 3, pp. 675–688, 3 2016.
[18] F. Adelantado et al., "Understanding the Limits of LoRaWAN," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 34–40, 2017.
[19] M. D. Yacoub, *Foundations of Mobile Radio Engineering*. CRC Press, 1993.
[20] M. Haenggi, *Stochastic geometry for wireless networks*. Cambridge University Press, 2013.
[21] M. Miskowicz, "Send-On-Delta Concept: An Event-Based Data Reporting Strategy," *Sensors*, vol. 6, no. 1, pp. 49–63, 1 2006.
[22] M. Simonov, H. Li, and G. Chicco, "Gathering Process Data in Low-Voltage Systems by Enhanced Event-Driven Metering," *IEEE Systems Journal*, vol. 11, no. 3, pp. 1755–1766, 9 2017.
[23] P. H. J. Nardelli et al., "Throughput analysis of cognitive wireless networks with Poisson distributed nodes based on location information," *Ad Hoc Networks*, vol. 33, pp. 1–15, 10 2015.
[24] G. Mcneill, J. Bright, and S. A. Hale, "Estimating local commuting patterns from geolocated Twitter data," *EPJ Data Sci.*, vol. 6, 2017.

Mauricio de Castro Tomé received the B.S and M.Sc. in electrical engineering from the State University of Campinas, Brazil, in 2011 and 2014, respectively. He is currently pursuing his Ph.D. degree at the University of Oulu, Finland, researching power and energy measurements and load estimation in smart grids.

Pedro H. J. Nardelli received the B.S. and M.Sc. degrees in electrical engineering from the State University of Campinas, Brazil, in 2006 and 2008. In 2013 he received his doctoral degree from University of Oulu, Finland, and State University of Campinas following a dual-degree agreement. Nowadays he is assistant professor in IoT in energy systems (tenure track) at the Laboratory of Control Engineering and Digital Systems, School of Energy Systems, Lappeenranta University of Technology, Finland, as well as adjunct professor (docent) in information processing and communication strategies for energy systems from at Centre for Wireless Communications, University of Oulu. More information: https://sites.google.com/view/nardelli/

Hirley Alves received the B.Sc. and M.Sc. degrees from Federal University of Technology - Paraná (UTFPR), Brazil, in 2010 and 2011, respectively, both in Electrical Engineering. Hirley received dual D.Sc. degree from University of Oulu and UTFPR in 2015. Dr. Alves is Adjunct Professor on Machine-type Wireless Communications at Centre for Wireless Communications (CWC), University of Oulu, Oulu, Finland. Dr Alves has acted as organizer, chair, and serves TPC and tutorial lecturer to several renowned international conferences. He is a co-recipient of the 2017 IEEE International Symposium on Wireless Communications and Systems (ISWCS) Best Paper Award, and a co-recipient of the 2016 Research Award from the Cuban Academy of Sciences. Dr. Alves is actively working on massive connectivity and ultra-reliable low latency communications. His research interests are wireless and cooperative communications, wireless full-duplex communications, PHY-security and ultra-reliable communications mechanisms for future machine type wireless networks.