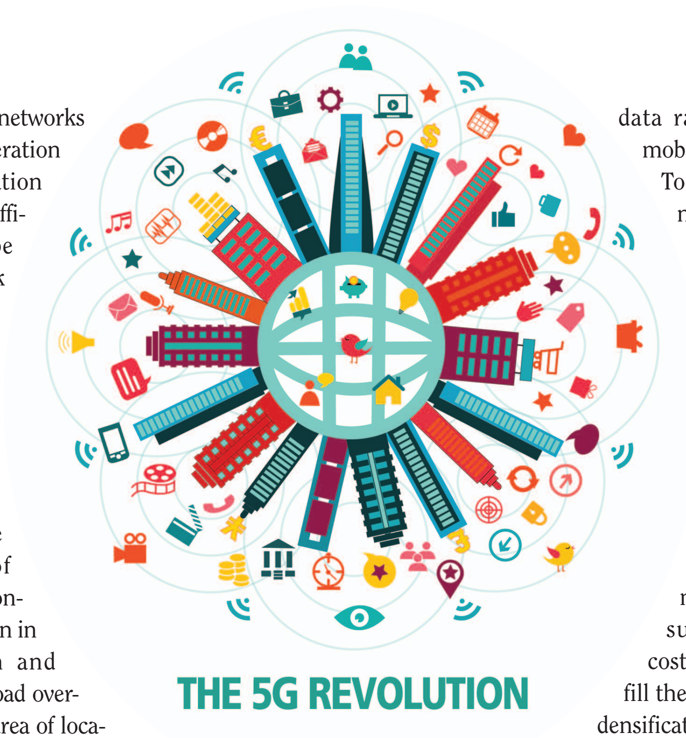
Rocco Di Taranto, Srikar Muppirisetty, Ronald Raulefs, Dirk T.M. Slock,
Tommy Svensson, and Henk Wymeersch

Location-Aware Communications for 5G Networks

How location information can improve scalability, latency, and robustness of 5G

Fifth-generation (5G) networks will be the first generation to benefit from location information that is sufficiently precise to be leveraged in wireless network design and optimization. We argue that location information can aid in addressing several of the key challenges in 5G, complementary to existing and planned technological developments. These challenges include an increase in traffic and number of devices, robustness for mission-critical services, and a reduction in total energy consumption and latency. This article gives a broad overview of the growing research area of location-aware communications across different layers of the protocol stack. We highlight several promising trends, tradeoffs, and pitfalls.



INTRODUCTION AND CHALLENGES

Fifth-generation will be characterized by a wide variety of use cases, as well as orders-of-magnitude increases in mobile data volume per area, number of connected devices, and typical user data rate, all compared to current mobile communication systems [1]. To cope with these demands, a number of challenges must be addressed before 5G can be successfully deployed. These include the demand for extremely high data rates and much lower latencies, potentially down to 1 ms end-to-end for certain applications [2]. Moreover, scalability and reduction of signaling overhead must be accounted for, as well as minimization of (total) energy consumption to enable affordable cost for network operation. To fulfill these requirements in 5G, network densification is key, calling for a variety of coordination and cooperation techniques between various kinds of network elements in an ultradense heterogeneous network. Moreover, by implementing sharing and coexistence approaches, along with new multi-GHz frequency bands, spectrum efficiency can be improved. An overview of a number of disruptive technologies for 5G is provided in [1].

It is our vision that context information in general and location information in particular can complement both traditional and disruptive technologies in addressing several of the challenges in 5G networks. While location information was

Digital Object Identifier 10.1109/MSP.2014.2332611
Date of publication: 15 October 2014

 1053-5888/14©2014IEEE

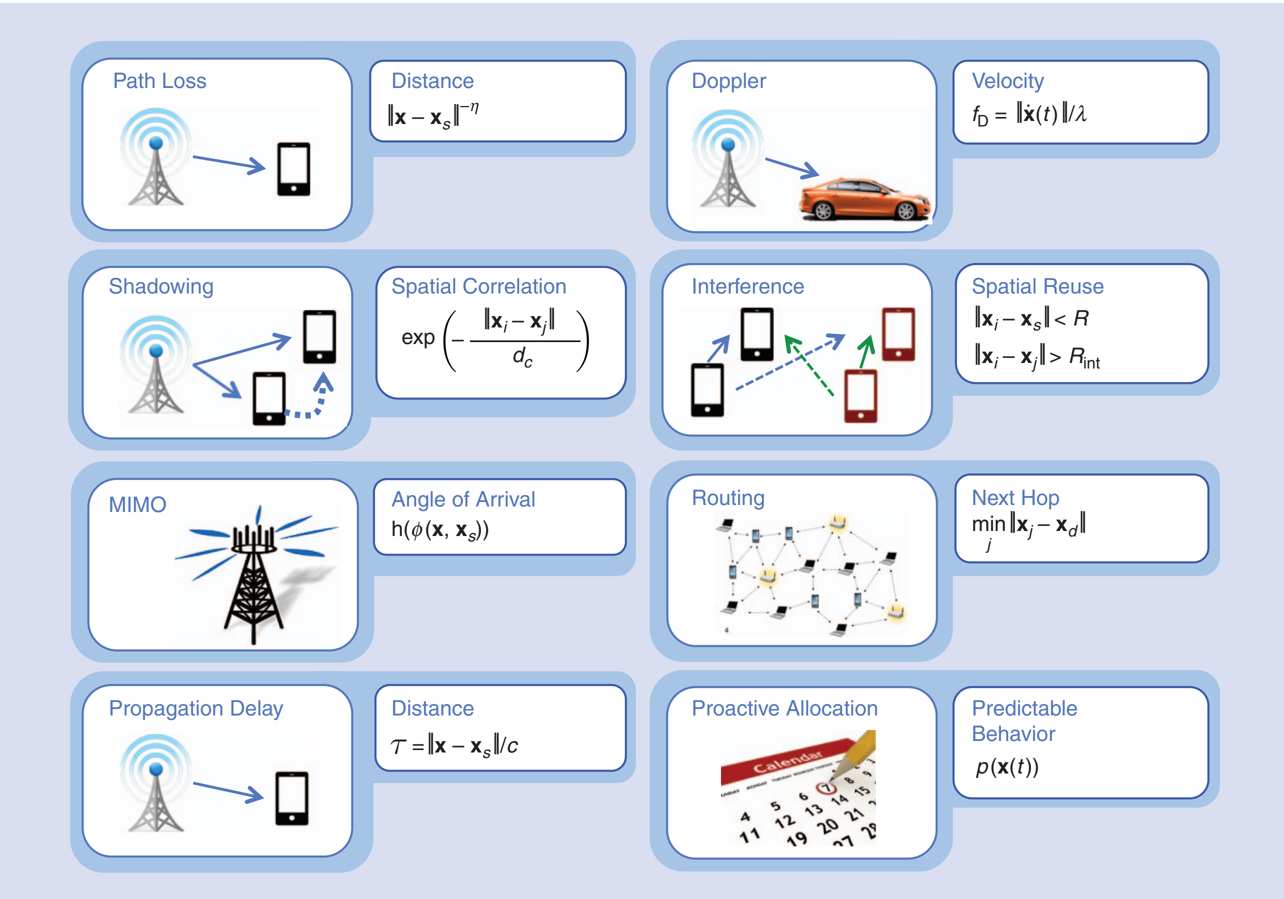
available in previous generations of cellular mobile radio systems, e.g., cell-identifier (ID) positioning in second generation (2G), timing-based positioning using communication-relevant synchronization signals in third generation (3G), and additionally dedicated positioning reference signals in fourth generation (4G), accuracy ranged from hundreds to tens of meters, rendering position information insufficiently precise for some communications operations. In 5G, for the first time, a majority of devices can benefit from positioning technologies that achieve a location accuracy on the order of 1 m.

FIFTH-GENERATION NETWORKS WILL BE THE FIRST GENERATION TO BENEFIT FROM LOCATION INFORMATION THAT IS SUFFICIENTLY PRECISE TO BE LEVERAGED IN WIRELESS NETWORK DESIGN AND OPTIMIZATION.

In this article, we argue why and how such precise location awareness can be harnessed in 5G networks. We first present technologies providing seamless and ubiquitous location awareness for 5G devices, identify associated signal processing challenges, and describe at a high level how location information can be utilized across the protocol stack. We then zoom in on each layer of the protocol stack and provide an overview of recent and relevant research on location-aware communications. We conclude the article by identifying a number of issues and research challenges that must be addressed before 5G technologies can successfully utilize location information and achieve the predicted performance gains.

LOCATION AWARENESS IN 5G NETWORKS

A majority of 5G devices will be able to rely on ubiquitous location awareness, supported through several technological developments: a multitude of global navigation satellite systems (GNSS) are being rolled out, complementing the current global positioning system (GPS). Combined with ground support systems and multiband operation, these systems aim to offer



[FIG1] Communication systems are tied to location information in many ways, including through distances, delays, velocities, angles, and predictable user behavior. The notations are as follows (starting from the top left downward): $\mathbf{x}$ is the user location, $\mathbf{x}_s$ is the base station or sender location, and η is the path loss exponent; $\mathbf{x}_i$ and $\mathbf{x}_j$ are the two-user location and d_c is a correlation distance; $\phi(.)$ is an angle of arrival between a user and a base station and $\mathbf{h}$ is a multiple-input, multiple-output (MIMO) channel; c is the speed of light and τ a propagation delay; f_D is a Doppler shift, $\dot{\mathbf{x}}(t)$ is the user velocity, and λ is the carrier wavelength; R is a communicate range and R_{int} is an interference range; $\mathbf{x}_d$ is a destination; and $p(\mathbf{x}(t))$ is a distribution of a user position at a future time t.

location accuracies around 1 m in open sky [3]. In scenarios where GNSS is weak or unavailable (in urban canyons or indoors), other local radio-based technologies such as ultrawideband (UWB), Bluetooth, ZigBee, and radio frequency identification (RFID), will complement current Wi-Fi-based positioning. Together, they will also result in submeter accuracy.

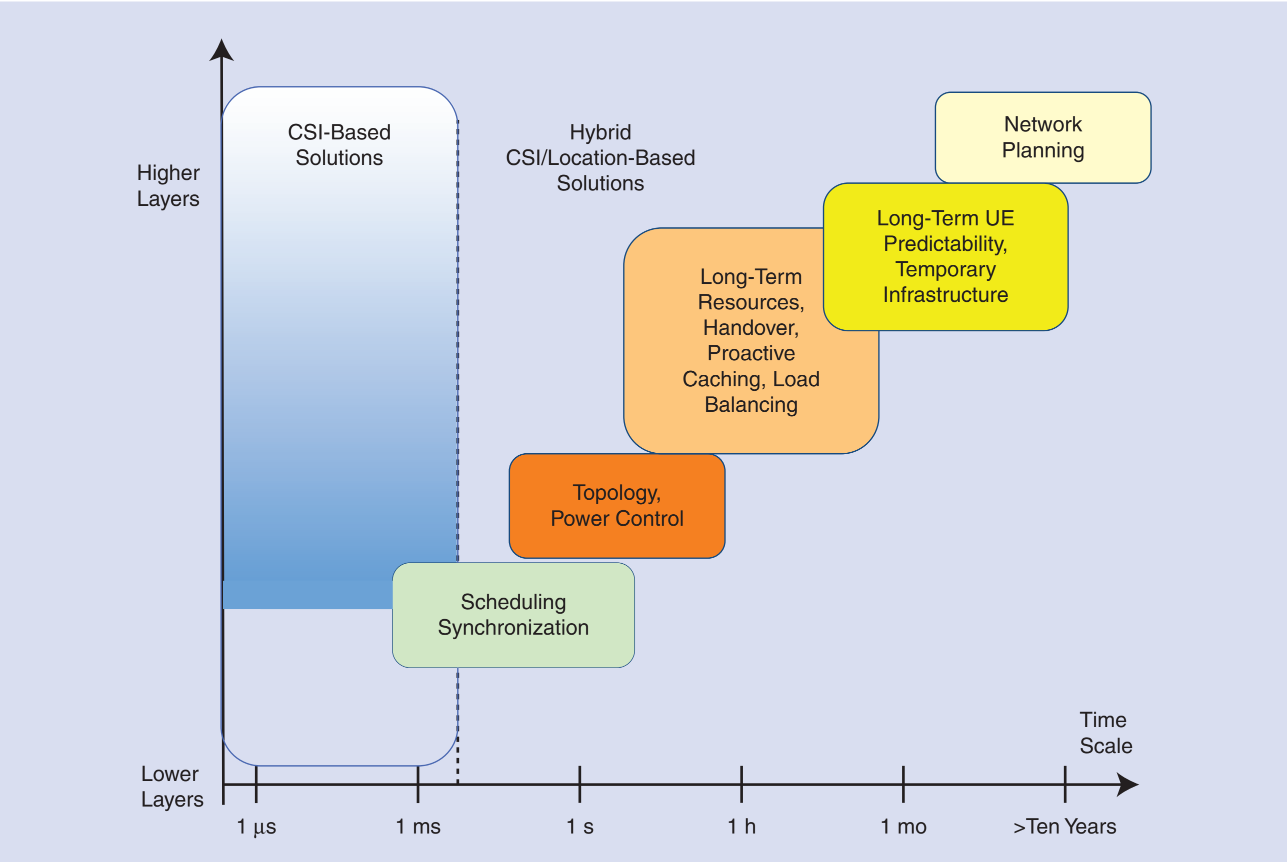
RECENT STUDIES HAVE REVEALED THAT LOCATION INFORMATION CAN BE HARNESSED NOT ONLY BY COGNITIVE NETWORKS, BUT ALSO CELLULAR AND AD HOC CONFIGURATIONS.

Accurate location information can be utilized by 5G networks across all layers of the communication protocol stack [4]. This is due to a number of reasons (see Figure 1), which will be detailed in later sections. First of all, signal-to-noise ratio (SNR) reduces with distance due to path loss, so that location knowledge and thus distance knowledge can serve as an indication of received power and interference level. Thus, if shadowing is neglected, the optimal multihop path between a source-destination pair in a dense network is the one that is shortest in terms of distance. Second, while path loss is the dominant effect in wireless communications, shadowing creates significant localized power differences due to signal propagation through objects. Since shadowing often exhibits decorrelation distances larger than the positioning uncertainty, local channel information can be extrapolated to nearby terminals. Third, most 5G user terminals will largely be predictable in their mobility patterns, since they will be either associated with people or fixed or controllable entities. Finally, at the highest layers, location information is often crucial, not only for location-based services, but also for a variety of tasks in cyberphysical systems, such as robotics and intelligent transportation systems.

Location awareness can be harnessed in a variety of ways to address several of the challenges in 5G networks. In particular, location-aware resource allocation techniques can reduce overheads and delays due to their ability to predict channel quality beyond traditional time scales. In Figure 2, we provide a top-level view of how location information may be utilized, inspired by research activities in 3G and 4G communication networks, while details will be provided in the subsequent sections.

LOCATION AWARENESS ACROSS THE PROTOCOL STACK

Location awareness has received intense interest from the research community, in particular with respect to cognitive radio [5], where location databases are being used to exploit TV



[FIG2] At very short time scales, resource allocation (especially in the lower layers) must rely on instantaneous channel-state information (CSI). At longer time scales, position information can be harnessed to complement CSI.

white spaces. However, recent studies have revealed that location information can be harnessed not only by cognitive networks, but also cellular and ad hoc configurations [6].

In this section, we aim to group a number of representative works in this developing area, based on the layer of the protocol stack to which they pertain. Since many of the works below are inherently cross-layer, sometimes we had to make choices among two layers. We start by a description of how channel quality metrics can be predicted through a suitable database and inference engine.

THE CHANNEL DATABASE

To predict the channel quality in locations where no previous channel quality measurement was available, a flexible predictive engine is needed. As different radio propagation environments have different statistical model parameters, this engine should be able to learn and adapt. Regression techniques from machine learning can be used for this purpose. Among these techniques, we focus on Gaussian processes (GPs) [7]. GPs have been used to predict location-dependent channel qualities in [8] and [9] in the following manner: users send a channel quality metric (CQM) to the database, along with the time and location at which it was acquired. After a training stage, the GP can provide an estimate of the CQM along with the uncertainty for any other receiver location. Hence, the output of the GP can be considered as a prior distribution on the channel quality. The construction and utilization of such a GP database is shown in Figure 3. The CQM can take on a variety of forms (see also Figure 1), including received power, root mean square delay spread, interference levels, or angular spread and rank profile for multiantenna systems [6], [4]. For the sake of simplicity, we will consider received power and disregard any temporal correlation of the CQM.

To model the received power CQM, we recall that a radio signal is affected mainly by three major components of the wireless propagation channel: distance dependent path-loss, shadowing due to obstacles in the propagation medium, and small-scale fading due to multipath effects. Small scale-fading decorrelates over very short distances for target operational frequencies. Hence, even with highly accurate position information, predictions of small-scale fading in new locations are not possible. This implies that we can only provide coarse channel information, which in many cases must be complemented with instantaneous small-scale information (see Figure 2). We let $P_{\mathrm{RX}}(\mathbf{x}_s, \mathbf{x}_i)$ be the power at a receiver node (located at $\mathbf{x}_i \in \mathbb{R}^2$), averaged over the small-scale fading in either time or frequency, from a source node (located at $\mathbf{x}_s \in \mathbb{R}^2$), which can be expressed in a dB scale as

$$P_{\mathrm{RX}}(\mathbf{x}_s, \mathbf{x}_i) = L_0 - 10\eta\log_{10}(\|\mathbf{x}_s - \mathbf{x}_i\|) + \Psi(\mathbf{x}_s, \mathbf{x}_i), \qquad (1)$$

where η is the path-loss exponent, $\Psi(\mathbf{x}_s, \mathbf{x}_i)$ is the location-dependent shadow fading between the source and the receiver (expressed in dB), and L_0 is a constant that captures antenna and other propagation gains. Although L_0 is assumed to be common to all users, additional user-specific biases, such as different antenna types or transmit powers can be calculated by the user and sent back to the base station. A common choice for shadow fading is to assume a log-normal distribution, i.e., $\Psi(\mathbf{x}_s, \mathbf{x}_i) \sim \mathcal{N}(0, \sigma_\Psi^2)$, where σ_Ψ^2 is the shadowing variance. While the location dependence on path loss is clear from (1), the shadowing also has well-established spatial correlation models, such as [10] for cellular networks, wherein the spatial autocovariance function of shadowing is given by

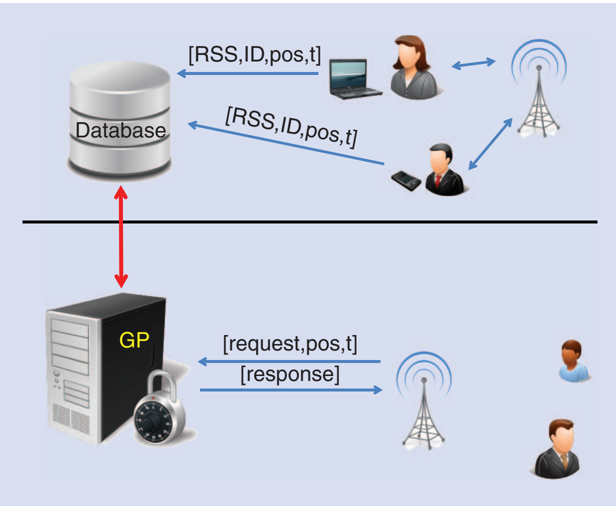
$$1C(\mathbf{x}_i, \mathbf{x}_j) = \mathbb{E}\{\Psi(\mathbf{x}_s, \mathbf{x}_i)\Psi(\mathbf{x}_s, \mathbf{x}_j)\} = \sigma_\psi^2 \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|}{d_c}\right), \quad (2)$$

where d_c denotes the correlation distance.

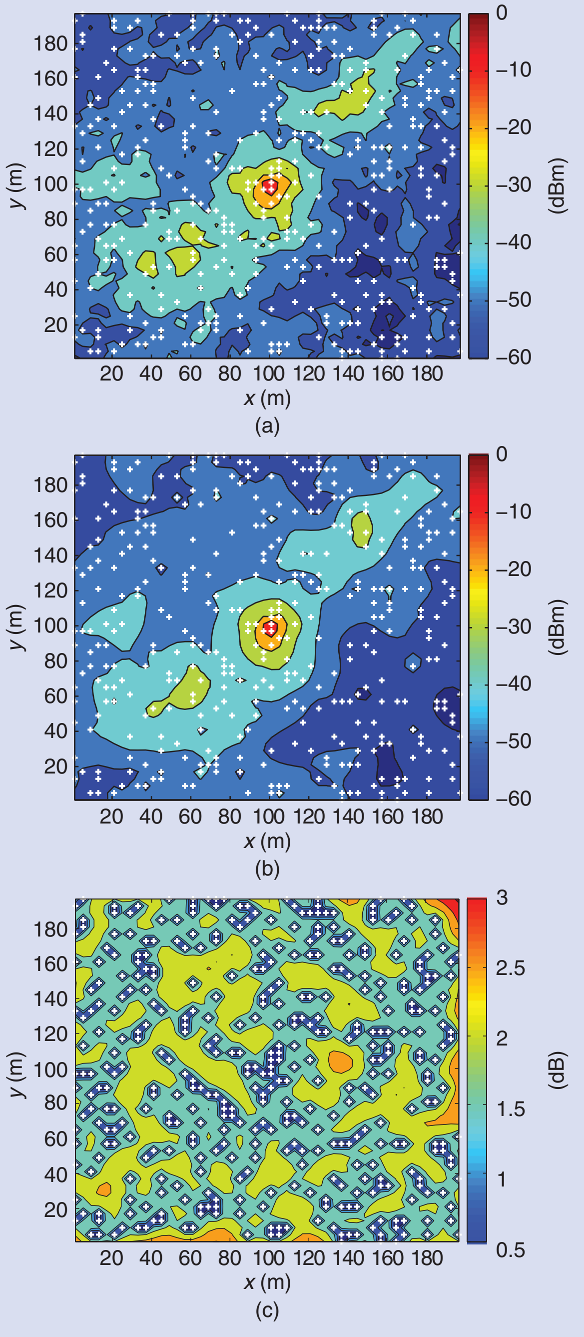
In the case of a common transmitter (e.g., a base station with location $\mathbf{x}_s$), the GP framework operates as follows. The power $P_{\mathrm{RX}}(\mathbf{x}_s, \mathbf{x}_i)$ is considered to be a GP as a function of $\mathbf{x}_i$, with mean function $\mu(\mathbf{x}_i)$ and covariance function $C(\mathbf{x}_i, \mathbf{x}_j)$. If we choose the mean function to be $\mu(\mathbf{x}_i) = L_0 - \eta 10\log_{10}(||\mathbf{x}_s - \mathbf{x}_i||)$, then the covariance function is exactly as defined in (2). To train the GP, let $y_i = P_{\mathrm{RX}}(\mathbf{x}_s, \mathbf{x}_i) + n_i$ be the noisy (scalar) observation of the received power at node i, where n_i is a zero mean additive white Gaussian noise random variable with variance σ_n^2. We introduce $\mathbf{X} = [\mathbf{x}_1^T, \mathbf{x}_2^T, \ldots, \mathbf{x}_N^T]^T$, $\mathbf{y} = [y_1, y_2, \ldots, y_N]^T$, and $\mathcal{D} = \{\mathbf{X}, \mathbf{y}\}$. The joint distribution of the N training observations now exhibits a Gaussian distribution [7]

$$\mathbf{y} \,|\, \mathbf{X}; \theta \sim \mathcal{N}(\mu(\mathbf{X}), \mathbf{K}), \qquad (3)$$

where $\mu(\mathbf{X}) = [\mu(\mathbf{x}_1), \mu(\mathbf{x}_2), \ldots, \mu(\mathbf{x}_N)]^T$ is the mean vector and $\mathbf{K}$ is the covariance matrix with entries $[\mathbf{K}]_{ij} = C(\mathbf{x}_i, \mathbf{x}_j) + \sigma_n^2\,\delta_{ij}$, where $\delta_{ij} = 1$ for $i = j$ and zero otherwise. The Gaussian distribution (3) depends on a number of parameters $\theta = [\sigma_n^2, d_c, L_0, \eta, \sigma_\Psi^2]$, which can be learned using the training



[FIG3] Users upload their location (pos) and time-tagged (t) channel quality metrics [e.g., the received signal strength, (RSS)], possibly along with their user ID, to a channel database. The information can be extrapolated to future users, requesting a channel quality metric in other locations for the same base station, using techniques such as GPs.



[FIG4] Radio channel prediction in decibel scale, with hyperparameters $\theta = [\sigma_n^2 = 0.01, d_c = 70\ m, L_0 = 10\ \text{dB}, \eta = 3, \sigma_\Psi = 9\ \text{dB}]$, $N = 400$ measurements (+ signs). The channel prediction is performed at a resolution of 4 m. (a) shows the true channel field, (b) the mean [obtained from (4)] of the predicted channel field; and (c) the standard deviation [obtained from the square root of (5)] of the predicted channel field.

database $\mathcal{D}$ by minimizing negative log-likelihood $-\log(p(\mathbf{y}\,|\,\mathbf{X};\theta))$ with respect to θ. This completes the training process. The predictive distribution of the noise-free signal power $P_{\text{RX}}(\mathbf{x}_s, \mathbf{x}_*)$ at a new node location $\mathbf{x}_*$, given the training database $\mathcal{D}$, is a Gaussian distribution with mean $\overline{P}_{\text{RX}}(\mathbf{x}_s, \mathbf{x}_*)$ and variance $\Sigma_{\text{RX}}(\mathbf{x}_s, \mathbf{x}_*)$, given by [7]

$$\overline{P}_{\text{RX}}(\mathbf{x}_s, \mathbf{x}_*) = \mu(\mathbf{x}_*) + \mathbf{k}_*^T \mathbf{K}^{-1}(\mathbf{y} - \mu(\mathbf{X})) \tag{4}$$

$$\Sigma_{\text{RX}}(\mathbf{x}_s, \mathbf{x}_*) = C(\mathbf{x}_*, \mathbf{x}_*) - \mathbf{k}_*^T \mathbf{K}^{-1}\mathbf{k}_*, \tag{5}$$

in which $\mathbf{k}_*$ is the $N \times 1$ vector of cross-covariances $C(\mathbf{x}_*, \mathbf{x}_i)$ between $\mathbf{x}_*$. and the training inputs $\mathbf{x}_i$.

Figure 4 demonstrates an example of radio channel prediction using a GP. A base station is placed in the center and a two-dimensional radio propagation field is simulated through a computer model according to (3) with sampling points on a square grid of 200 m × 200 m and a resolution of 4 m. Based on measurements at marked locations, the mean and standard deviation of the prediction are obtained for any location. Observe the increased uncertainty in Figure 4(c) in regions where few measurements are available.

In the case where links rely on different transmitters, the model above can still be applied [8], though more advanced models exist. For instance, in the case of ad hoc networks, [9] proposes a model where shadow fading is due to an underlying spatial loss field.

GPs can thus provide a statistical description of the CQM in any location and any time. This description can be used in resource allocation at different layers, e.g., to reduce delays and/or overheads. In the following, we present specific examples that are useful mainly in one layer. We will start with the physical layer.

THE PHYSICAL LAYER

In the lowest layer of the protocol stack, location information can be harnessed to reduce interference and signaling overhead, to avoid penalties due to feedback delays, or to synchronize coordinated communication schemes.

The best known application is spatial spectrum sensing for cognitive radio [11], where a GP allows the estimation of power emitted from primary users at any location through collaboration among secondary users. The resulting power density maps enable the secondary users to choose the frequency bands that are not crowded and to adapt their transmit power to minimize the interference to the primary users. These techniques can be adapted in 5G to perform interference coordination. For instance, significant potential for the exploitation of location information in multiantenna techniques arises in spatial cognitive radio paradigms (underlay, overlay, interweave) [6]. Such location-aided techniques could be compatible with some very recent developments in massive MIMO, where the exploitation of slow fading subspaces in the multiantenna propagation has been advocated.

The GP database also provides useful information in any application that relies on a priori channel information, such as slow adaptive modulation and coding or channel estimation. This is investigated in [12], where location-aware adaptive mobile communication uses both channel and spatial movement coherence in combination with location prediction and a fingerprint database. When at time t a user reports future predicted locations $\mathbf{x}(t), \mathbf{x}(t+1), \ldots, \mathbf{x}(t+T)$ to the database, the corresponding received powers can be determined $\overline{P}_{\mathrm{RX}}(t)$, $\overline{P}_{\mathrm{RX}}(t+1), \ldots, \overline{P}_{\mathrm{RX}}(t+T)$. For each time, the predicted capacity is then

$$C(t) = W \log_2\left(1 + \frac{\overline{P}_{\mathrm{RX}}(t)}{N_0 W}\right), \tag{6}$$

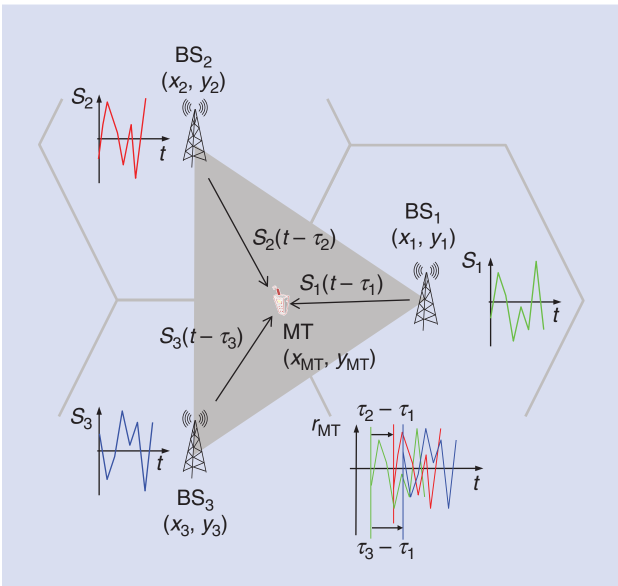
where W is the signaling bandwidth and N_0 is the noise power spectral density. The communication rate is then adapted to not exceed the predicted capacity. It is demonstrated that location-aware adaptive systems achieve large capacity gains compared to state-of-the-art adaptive modulation schemes for medium to large feedback delays. Such delays are especially important in 5G application with fast-moving devices, such as transportation systems, which are also the topic of [13], where the short channel coherence time precludes adaptation based on the fast fading channel. Instead, link adaptation based on path loss is considered, which in turn depends on the locations of the vehicles. Expressions for large-scale coherence time and velocity are derived, and it is found that feeding back location information can substantially reduce feedback overhead without compromising data rate. Location-based feedback latency reduction is also discussed in [6], e.g., for fast relay selection.

Significant opportunities for location-aware communications concern resource allocation aspects, especially for multiuser (MU) and MIMO systems. In such systems, recent information theory progress shows that an optimized handling may lead to significant system capacity increases, though only in the presence of very precise channel-state information at the transmitter (CSIT). In the single-cell case (or at the cell center), one can consider location-aware downlink MU-MIMO. Multiple antennas at the user side do not allow a base station with M antennas serving M users to send more streams in a cell, but a user can use its N antennas to suppress the effect of $N-1$ multipath components. Hence, if the overall propagation scenario involves a line-of-sight (LoS) path and up to $N-1$ multipath components, the user can use receive beamforming (BF) to transform its channel into a pure LoS channel, allowing the base station to perform zero-forcing (ZF) transmission with only location information [14]. In the multicell case, which in information-theoretic terms corresponds to the interference channel (IC) and in practice to the macrocellular environment or to HetNets (coexistence of macro and femto/small cells), there are opportunities for location-aided MIMO interference channels [14]. In particular the feasibility of joint transmitter/receiver (Tx/Rx) ZF BF is of interest in the case of reduced rank MIMO channels (with LoS being the extreme case of rank one). Whereas in the full rank MIMO case, the joint Tx/Rx design is complicated by overall coupling between all Tx and all Rx, i.e., a requirement of overall system CSI at all base stations, some simplifications may occur in the reduced rank case. In particular, for the LoS case (the easiest location-aided scenario, higher rank cases requiring databases), the Tx/Rx design gets decoupled, leading to only local (e.g., location-based) CSI requirements [14].

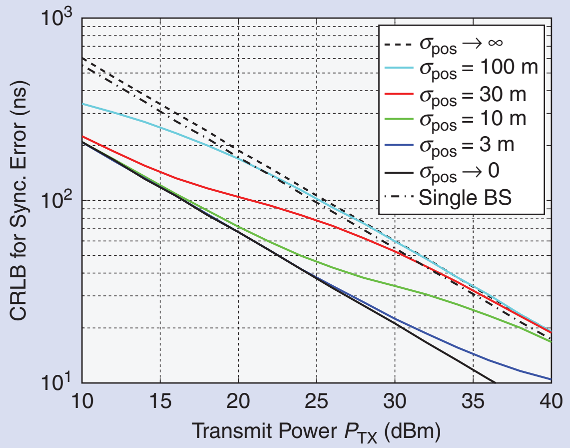
Locations can also be utilized in a different manner, by converting them not to a CQM, but to other physical quantities, such as Doppler shifts (proportional to the user's relative velocity), arrival angles (used in [4] for location-based spatial division multiple access), or timing delays (which are related to the distance between transmitter and receiver). This latter idea is taken up in [6] and applied for coordinated multipoint (CoMP) transmission, illustrated in Figure 5, showing a mobile node receiving synchronization signals from three base stations, deployed with a frequency reuse of 1. CoMP transforms interference experienced by the mobile users to signal power, especially at the cell edge, by coordinating the signals of all involved base stations.

CoMP relies on accurately synchronized signals, a process that can be aided through a priori location information, which determines the potential window to exploit the synchronization signals from different base stations. Figure 6 shows the potential gains in terms of required transmit power at the base station to achieve a certain synchronization performance for different values of the location uncertainty of the mobile node. The communication system benefits if the synchronization requirements are at least in the range of the location accuracy (1-ns timing uncertainty corresponds to 30-cm position uncertainty). For example, comparing a system with multiple base stations for a desired synchronization accuracy of 20 ns, 40 dBm is required when no position information is available, while less than 32 dBm is required when positioning accuracy is around 3 m.

As the aforementioned works indicate, location information provides valuable side-information about the physical layer. It can be harnessed to reduce delays and feedback overhead, and even to improve performance. Determining when to utilize location-based CQM and when to rely on instantaneous CSI is an important topic in the optimization of 5G communications. Next, we move up to the medium access control (MAC) layer, where even richer opportunities for the use of location information arise than



[FIG5] By knowing the location of the mobile node (x_{MT}, y_{MT}), the propagation delays τ_i of the signal components $s_i(t)$ coming from different base stations [with known locations (x_i, y_i)] can be related to each other to improve the CoMP transmission.



[FIG6] The Cramér–Rao lower bound (CRLB) for the synchronization error (standard deviation) versus the base station transmit power P_{TX} for different positioning uncertainties (σ_{pos}) of the mobile node. The Third-Generation Partnership Project (3GPP) long-term evolution secondary synchronization signals were used with IDs 142, 411, and 472. (Figure based on [6].)

at the physical layer, especially without the need to estimate channel gains based on position and/or distance information.

THE MAC LAYER

With more devices communicating with each other, scalability, efficiency, and latency are important challenges in designing efficient protocols for MAC. In this section, we provide an overview of some of the existing works on the use of location information at the MAC layer to address these design challenges. In particular, multicasting, scheduling, and selection protocols are considered. Again, we can make a distinction between approaches that tie locations to channel and approaches where locations are exploited in a different way.

In the first group, we find works such as [6] and [15]–[17]. The basic premise is that a link between transmitter with position $\mathbf{x}_s$ and receiver with position $\mathbf{x}_i$ can be scheduled with the same resource as an interfering transmitter with position $\mathbf{x}_j$, provided that

$$\frac{P_{RX}(\mathbf{x}_s, \mathbf{x}_i)}{N_0 W + P_{RX}(\mathbf{x}_j, \mathbf{x}_i)} > \gamma, \tag{7}$$

where γ is a signal-to-interference-plus-noise ratio (SINR) threshold. SINR expressions such as (7) can easily be combined with a CQM database. In [6], a location-aided round-robin scheduling algorithm for fractional frequency reuse is proposed, where allowing temporary sharing of resources between cell-center and cell-edge users is shown to achieve higher total throughput with less and less frequent feedback than the conventional method. In the same paper, location-based long-term power setting in heterogeneous cochannel deployments of macro and femto base stations is investigated. In [15], location-based multicasting is considered, assuming a disk model, and is shown to both reduce the number of contention phases and increase the reliability of packet delivery, especially in dense networks. Time division with spatial reuse is considered in [17], which investigates location-aware joint scheduling and power control for IEEE 802.15.3, leading to lower latencies and higher throughput compared to a traditional round-robin type scheduling mechanism. Location information is also beneficial in reducing the overhead associated with node selection mechanisms (e.g., users, relays), by allowing base stations to make decisions based solely on the users' positions [6]. Finally, location information is a crucial ingredient in predicting interference levels in small/macrocell coexistence, in multicell scenarios, and in all cognitive radio primary/secondary systems. For example, [6] and [14] demonstrate the use of location information to allow to significantly improve intercell interference coordination techniques. Location-based modeling of attenuation and slow fading components will bring about progress in the design of multicellular systems, complementing the recent significant progress that has focused almost exclusively on the fast fading component (e.g., interference alignment). For underlay cognitive radio systems, location-based prediction of interference caused to primary users may be a real enabling approach. These works indicate that significant gains in terms of throughput and latency can be reaped from location-aware MAC in 5G networks, provided appropriate channel models are used.

In the second group, we find approaches that utilize location information in a different way [16], [18], [19]. All turn out to relate to vehicular networks. In [16], a family of highly efficient location-based MAC protocols is proposed, whereby vehicles broadcast information to other vehicles only when they pass through predetermined transmission areas. When the traffic flow rate increases, the proposed location-based protocols have a smaller message delivery time compared with conventional random access schemes. A similar idea is proposed in [18], where a decentralized location-based channel access protocol for intervehicle communication is studied. Channels are allocated based on vehicles' instantaneous geographic location, and unique channels are associated to geographic cells. Using a prestored cell-to-channel mapping, vehicles know when to transmit on which channel, alleviating the need for a centralized coordinator for channel allocation. This leads to efficient bandwidth use and avoids hidden node problems, since neighboring cells do not use the same channel. In addition, communication delay is bounded and fairness among the vehicles is maintained as each vehicle gets a channel regularly to transmit. Finally, [19] introduces the concept of geocasting, whereby multicast regions are formed based on the geographical location of the nodes and packets are sent to all the nodes in the group. Specialized location-based multicasting schemes are proposed to decrease the delivery overhead of packets when compared to multicast flooding mechanisms.

We observe that in the MAC layer there is a more varied use of location information than in the physical layer, especially without direct need of the channel database. In all cases, improvements in terms of latency, overhead, or throughput were reported. The emphasis on dense mobile networks, and the limited need for centralized infrastructure make these techniques promising for 5G networks. We now move up to the network and transport layers, where geographic routing plays an important role.

WITH MORE DEVICES COMMUNICATING WITH EACH OTHER, SCALABILITY, EFFICIENCY, AND LATENCY ARE IMPORTANT CHALLENGES IN DESIGNING EFFICIENT PROTOCOLS FOR MAC.

NETWORK AND TRANSPORT LAYERS

At the network and transport layers, location information has been shown to improve scalability and to reduce overhead and latency. A full-fledged location-based network architecture is proposed in [5] for cognitive wireless networks, dealing with dynamic spectrum management, network planning and expansion, and in handover. In particular, a location-aided handover mechanism significantly reduces the number of handovers compared with signal strength-based methods [20], which are subject to delay and hysteresis effects.

Location-aided techniques, especially using mobility information to forecast future channel capacities for the mobile, become particularly powerful when vertical temporary handovers are considered to systems with larger channel capacity to offload data. Such large capacity systems may exhibit short windows of opportunity due to their limited coverage.

Most other works at the network layer have focused on the routing problem. A well-known technique in this area is geographic routing (georouting), which takes advantage of geographic information of nodes (actual geographic coordinates or virtual relative coordinates) to move data packets to gradually approach and eventually reach their intended destination. In its most basic form, given a destination d, a node i with neighbors $\mathcal{N}_i$ will choose to forward data to a neighbor closest to the destination:

$$j^* = \arg\min_{j \in \mathcal{N}_i} \| \mathbf{x}_j - \mathbf{x}_d \|. \tag{8}$$

Recently, georouting has gained considerable attention, as it promises a scalable, efficient, and low-latency solution for information delivery in wireless ad hoc networks. For a comprehensive survey of the existing literature on georouting, investigating how location information can benefit routing, we refer to [21].

Georouting is mainly limited due to two factors: it is sensitive to localization errors and it does not exploit CQM, favoring latency (measured in this context in terms of progress toward the destination) over throughput. The first issue is investigated in [22], where it is shown that georouting quickly degrades as location information becomes imprecise. More robust routing mechanisms are proposed, combining progress toward the destination with an error measure in the locations. The second issue is treated in [23] and [24]. In [23], where positions are mapped to a CQM, a centralized routing algorithm aims to maximize end-to-end flow. The mismatch between the estimated and true channels is mitigated using a distributed algorithm, whereby nodes locally adjust their rate, but not the routes. While [23] no longer directly optimizes progress toward the destination, [24] considers both throughput and latency in a fully distributed manner. In [24], the network consists of power-constrained nodes that transmit over wireless links with adaptive transmission rates. Packets randomly enter the system at each node and wait in output queues to be transmitted through the network to their destinations. The data flows from source to destination according to the enhanced dynamic routing and power control (EDRPC) algorithm, which is proven to stabilize the network with a bounded average delay. In EDRPC, each of the N nodes in the network maintains N queues, $Q_i^{(d)}$ denoting the queue at node i with stored information destined to node d (note that $Q_d^{(d)} = 0$ for all destinations). Each link, say (i, j), locally decides the destination to serve, such as

$$d_{ij}^* = \arg\max_{d \in \{1,\ldots,N\}} (\tilde{Q}_i^{(d)} - \tilde{Q}_j^{(d)}), \tag{9}$$

where $\tilde{Q}_i^{(d)} = Q_i^{(d)} + V_i^{(d)}$, in which $V_i^{(d)} \geq 0$ is a design parameter. When $V_i^{(d)} = 0,\ \forall i$, the destination with the largest backlog will be served over link (i, j). Setting the values $V_i^{(d)} = f(\| \mathbf{x}_i - \mathbf{x}_d \|)$, where $f(\cdot)$ is a monotonically increasing function will incentivize data to flow toward the geographic position of the destination (i.e., given equal backlogs, the destination will be chosen that maximizes $f(\| \mathbf{x}_i - \mathbf{x}_d \|) - f(\| \mathbf{x}_j - \mathbf{x}_d \|)$, favoring small $\| \mathbf{x}_j - \mathbf{x}_d \|$). Following the choice of d_{ij}^*, EDRPC performs a (centralized) power allocation for each link, leading to an allowable rate per link. Finally, each node i will serve destination d_{ij}^* over link (i, j) with an amount of data at the allowable rate and thus reduces its queue length $Q_i^{(d_i^*)}$.

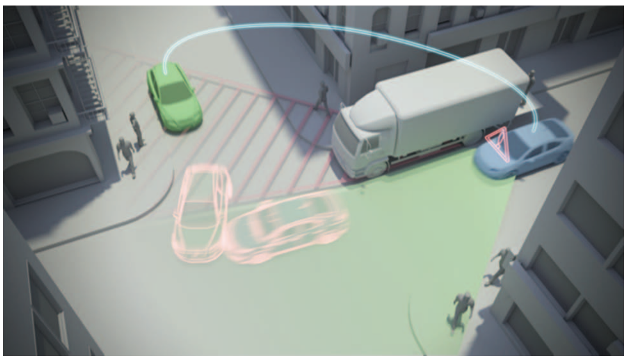
The focus in [22]–[24] is on relatively static networks, where there are no drastic topology changes. In certain applications, such as vehicular networks, this assumption is no longer valid, as is treated in [25] and [26]. In [25], the use of mobility prediction to anticipate topology changes and perform rerouting prior to route breaks is considered. The mobility prediction mechanism is applied to some of the most popular representatives of the wireless ad hoc routing family, mainly an on-demand unicast routing protocol, a distance vector routing protocol, and a multicast routing protocol. Routes that are the most stable (i.e., routes that do not become invalid due to node movements) and stay connected longest are chosen by utilizing the mobility prediction. The mobility characteristics of the mobile nodes are taken into account in [26], and a velocity-aided routing algorithm is proposed, which determines its packet forwarding scheme based on the relative velocity between the intended forwarding node and the destination node. The routing performance can further be improved by the proposed predictive mobility and location-aware routing algorithm, which incorporates the predictive moving behaviors of nodes in protocol design. The region for packet forwarding is determined by predicting the future trajectory of the destination node.

We clearly see that at the network and transport layer, harnessing location information appropriately can aid in reducing overhead and latency, while offering scalable solutions, even for highly mobile networks. In such networks, location information also plays an important role in the higher layers, as we will detail next.

HIGHER LAYERS

At the higher layers, location information will naturally be critical to provide navigation and location-based services. While we do not aim to provide a complete overview of such services, we briefly detail several applications of importance in the context of 5G networks.

First, we have classical context awareness, which finds natural applications in location-aware information delivery [27] (e.g., location-aware advertising) and multimedia streaming [28]. For the latter application, [28] tackles the problem of guaranteeing continuous streaming of multimedia services while minimizing the overhead involved, by capturing correlated mobility patterns, predicting future network planning events. A second class of applications is in the context of intelligent transportation systems. Several car manufacturers and research centers are investigating the development of intervehicle communication protocols. In this context, [29] focuses on the problem of providing location-aware services (e.g., traffic-related, time-sensitive information) to moving vehicles by taking advantage of short-range, intervehicle wireless communication and vehicular ad hoc networks. Location information is also critical for autonomous vehicles to coordinate and plan the vehicle's actions with respect to the environment and current traffic conditions (see Figure 7). Highly related are the tactile Internet [2] and other mobile cyber-physical systems, such as groups of unmanned aerial vehicles or robots [8], where localization and communication are closely intertwined.



[FIG7] The use of location information in intelligent transportation systems. After self-positioning, the vehicles become aware of each other through wireless communication and are able to avoid an accident.

Finally, location information also has implications in the context of security and privacy. For example, [30] studies the management of encryption keys in large-scale clustered sensor networks. In particular, a novel distributed key management scheme is proposed that reduces the potential of collusion among compromised sensor nodes by factoring the geographic location of nodes in key assignment. In [6], location information is utilized to detect wormhole attacks, which disrupt the network topology, as perceived by the benign nodes.

FIFTH-GENERATION MOBILE AND WIRELESS COMMUNICATION SYSTEMS WILL REQUIRE A MIX OF NEW SYSTEM CONCEPTS TO BOOST SPECTRAL EFFICIENCY, ENERGY EFFICIENCY, AND THE NETWORK DESIGN.

While we have focused on existing applications, we can expect novel, unforeseen location-based services in 5G networks, following us at all times, anticipating our needs, and providing us with information when and where we need it. With this comes a number of risks related to security and privacy, which should be addressed explicitly.

RESEARCH CHALLENGES AND CONCLUSIONS

Fifth-generation mobile and wireless communication systems will require a mix of new system concepts to boost spectral efficiency, energy efficiency, and the network design. There are many open issues to be addressed before these systems will be able to enter the market. In the following, we focus our attention on challenges related to the use of location information in 5G networks.

- *Achieving location awareness*: Throughout this article, we have assumed accurate location information is available. However, to realize the predicted position accuracies, significant signal processing challenges must be addressed so that seamless and ubiquitous localization can be made possible. The challenges include 1) handover, fusion, and integration of different positioning technologies; 2) coping with errors due to harsh propagation environments and interference; and 3) decentralization and reduction of complexity. In addition, 5G technologies themselves may have tight interactions with positioning. For example, millimeter wave systems may require accurate user tracking through BF; novel waveforms such as those used in filter bank multicarrier have relaxed synchronization demands, and may therefore reduce time-based positioning accuracy.
- *Ad hoc networking*: In ad hoc and certain machine-to-machine (M2M) networks, availability of a CQM database is questionable. In addition, accessing the database would require a preexisting communication infrastructure. Hence, distributed databases (or database-free methods) may be required in such networks, to capitalize on location awareness. The construction, maintenance, and exploitation of these databases will rely on distributed signal processing and deserves further study. Location knowledge can also be leveraged to find low-latency control and data paths in ad hoc networks, enabling wireless control systems. The appropriate storage, utilization, and combination of location-based with pilot-based CQM is an open issue.

■ *Signaling overhead*: While the ratio of signaling overhead with respect to data payload is generally increasing, this is particularly apparent in M2M and Internet of Things signaling, as the typically used protocols are inefficient for such traffic. Even in combination with location awareness, overhead will be a major bottleneck, and dedicated representation and compression mechanisms as well as localized protocols need to be designed. The choice of CQM also plays an important role as more precise information can yield better gains, but at costs in terms of complexity, robustness, and overhead.

■ *Spatial channel modeling*: The wide variety of use cases requires a flexible and robust inference engine. GPs, as presented earlier, are a promising candidate, but they are faced with challenges in terms of storage and computational complexity. Sparsifying techniques to build and maintain the database, decentralized processing, as well as structured approaches in the prediction are among the main signal processing challenges. In addition, various sources of uncertainty must be accounted for explicitly in the GP framework (e.g., in terms of the position), as well as inherent nonstationarities in the channel statistics. Yet another challenge is to keep the database of the different CQMs updated and synchronized. The updates may be delivered by different 5G radio devices and could drive the synergy between the different radio types, such as M2M or mobile radio devices. Compared to today's drive tests, the autonomous refinement of network resources would allow to increase the coherence time of the database content.

■ *PHY/MAC/NET layers*: Location information can be exploited in a number of ways, both through databases and channel modeling, as well as more directly at the PHY/MAC/NET layers. An important challenge is to identify the right tradeoff between relying on location-based information and on pilot-based CQM information. A second challenge involves the amount of centralized versus decentralized processing. An open question on the network level is how to best utilize location information for identifying when network-assisted device-to-device communication is beneficial and aiding neighbor discovery. Finally, the issue of energy-efficiency deserves further study. For example, location information could be used to decide when to power down certain small cell base stations.

■ *Higher layers*: In 5G networks using location information, there are great possibilities for resource allocation (power, bandwidth, rate) based on prediction of user behaviors/trajectories, predicted load levels at various network nodes, channel statistics, and interference levels (previously stored in databases, for example). Some initial research has been done, but there is a large space for designing completely new algorithms and solutions. In addition, sharing location information raises important privacy and security issues. Secure and private computing in a location-aware context are promising, but they pose technical challenges in which signal processing can aid in masking and hiding information and in developing attack-resistant algorithms and protocols.

LOCATION AWARENESS BEARS GREAT PROMISE TO THE 5G REVOLUTION, PROVIDED WE CAN UNDERSTAND THE RIGHT TRADEOFFS FOR EACH OF THE POSSIBLE USE CASES.

In summary, location awareness bears great promise to the 5G revolution, provided we can understand the right tradeoffs for each of the possible use cases. In this article, we have given an overview of how location awareness can be leveraged across the different layers of the (traditional) protocol stack, and we highlighted a number of important technical challenges.

ACKNOWLEDGMENTS

This work is supported, in part, by the European Research Council under grant number 258418 (COOPNET), the European Union under ICT-248894 FP7 project WHERE2, and by the Swedish Research Council under grant numbers 2011-6864 and 2009-4555.

AUTHORS

Rocco Di Taranto (taranto@chalmers.se) received the Laurea in Ingegneria delle Telecomunicazioni from Politecnico di Torino in 2005 and the Ph.D. degree in wireless communications from Aalborg University in 2010. From March to June 2009 he was a visiting researcher at the New Jersey Institute of Technology, United States. From February 2011 to May 2012 he was a postdoctoral fellow at the University of Waterloo, Ontario, Canada. From June 2012 to May 2014, he was a postdoctoral fellow at Chalmers University of Technology, Sweden. His research focuses on resource allocation in wireless and fiber-optic communications.

Srikar Muppirisetty (srikar.muppirisetty@chalmers.se) received his B.Tech degree in electronics and communication engineering in 2005 from R.V.R. and J.C. College of Engineering, India. He received his M.Sc. degree in communication engineering in 2009 and is currently pursuing a Ph.D. degree at Chalmers University of Technology. He has over four years of industrial experience in physical layer algorithm development. From 2005 to 2007, he worked as a member of technical staff at Digibee Microsystems, Bangalore. During 2009–2012, he worked as a senior system software engineer and as a technical leader on a simulation team at ST-Ericsson R&D Center, Bangalore. His current research focuses on developing algorithms for resource allocation using position information in wireless communications.

Ronald Raulefs (r.raulefs@dlr.de) is senior researcher and project manager at the Institute of Communications and Navigation of the German Aerospace Center (DLR) in Wessling, Germany. He received the Dr.-Ing. degree from the University of Erlangen-Nuremberg, Germany, in 2008. He was a visiting researcher at the City University of Hong Kong in September 2004. He initiated and coordinated the European Union Seventh Framework Programme for Research projects WHERE and WHERE2. He holds numerous patents in the area of mobile communications

and radio-based localization. His current research interests include various aspects of mobile radio communications and localization, such as dynamic resource allocation for cooperative localization.

Dirk T.M. Slock (slock@eurecom.fr) received an engineering degree from the University of Ghent, Belgium, in 1982, and the M.S. degree in electrical engineering (1986), M.S. degree in statistics (1989), and Ph.D. degree in electrical engineering (1989), all from Stanford University (Fulbright grant). He has been a professor at EURECOM since 1991. He invented semiblind channel estimation, the chip equalizer-correlator receiver (3G), MIMO-CDD now part of LTE, and the Single Antenna Interference Cancellation (SAIC) (GSM standard). He has (co) authored 420 papers. He received a Best Paper Award from the IEEE Signal Processing Society and from EURASIP in 1992, and two IEEE Globecom'98, one IEEE SIU'04 and one IEEE SPAWC'05 Best Student Paper Award.

Tommy Svensson (tommy.svensson@chalmers.se) received his Ph.D. degree in information theory from Chalmers University of Technology in 2003. He is an associate professor of communication systems at the same university, leading the research on air interface and wireless backhaul networking technologies for wireless systems. He actively contributes to European projects, currently within the EU FP7 METIS project targeting solutions for fifth generation. His main research interests are in design and analysis of physical layer algorithms, multiple access, resource allocation, cooperative systems, and moving relays/cells/networks. He is chair of the IEEE Sweden joint Vehicular Technology/Communications/Information Theory Societies Chapter and the coordinator of the Communication Engineering Master's Program at Chalmers. He is a Senior Member of the IEEE.

Henk Wymeersch (henkw@chalmers.se) is an associate professor of communication systems in the Department of Signals and Systems, Chalmers University of Technology, Sweden. He received the Ph.D. degree in 2005 from Ghent University. For his Ph.D. dissertation, he was awarded the 2006 Alcatel Bell Scientific Award. Between 2005 and 2009, he was a postdoctoral researcher at the Massachusetts Institute of Technology. His research interests include Bayesian inference, cooperative networks, and optical communications. He is the author of *Iterative Receiver Design* (Cambridge University Press, 2007), and is an associate editor of *IEEE Transactions on Wireless Communications* and *IEEE Transactions on Emerging Telecommunications Technologies*.

REFERENCES

[1] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[2] G. Fettweis, "The tactile internet: Applications and challenges," *IEEE Veh. Technol. Mag.*, vol. 9, no. 1, pp. 64–70, Mar. 2014.

[3] M. de Reuver, D. Skournetou, and E.-S. Lohan, "Impact of Galileo commercial service on location-based service providers: Business model analysis and policy implications," *J. Location Based Serv.*, vol. 7, no. 2, pp. 67–78, 2013.

[4] D. Slock, "Location aided wireless communications," in *Proc. Int. Symp. Communications Control and Signal Processing*, 2012, pp. 1–6.

[5] H. Celebi and H. Arslan, "Utilization of location information in cognitive wireless networks," *IEEE Wireless Commun.*, vol. 14, no. 4, pp. 6–13, 2007.

[6] A. Dammann, G. Agapiou, J. Bastos, L. Brunel, M. García, J. Guillet, Y. Ma, J. Ma, J. J. Nielsen, L. Ping, R. Raulefs, J. Rodriguez, D. Slock, D. Yang, and N. Yi, "WHERE2 location aided communications," in *Proc. European Wireless Conf.*, Apr. 2013, pp. 1–8.

[7] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press, 2006.

[8] J. Fink, "Communication for teams of networked robots," Ph.D. thesis, Elect. Syst. Eng., Univ. Pennsylvania, Philadelphia, PA, Aug. 2011.

[9] P. Agrawal and N. Patwari, "Correlated link shadow fading in multi-hop wireless networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4024–4036, 2009.

[10] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electron. Lett.*, vol. 27, no. 23, pp. 2145–2146, 1991.

[11] I. Nevat, G. W. Peters, and I. B. Collings, "Location-aware cooperative spectrum sensing via Gaussian processes," in *Proc. Australian Communications Theory Workshop*, 2012, pp. 19–24.

[12] S. Sand, R. Tanbourgi, C. Mensing, and R. Raulefs, "Position aware adaptive communication systems," in *Proc. Asilomar Conf. Signals, Systems and Computers*, 2009, pp. 73–77.

[13] R. C. Daniels and R. W. Heath, "Link adaptation with position/motion information in vehicle-to-vehicle networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 505–509, 2012.

[14] A. Dammann et al., A. Dammann, J. Bastos, L. Brunel, M. García, J. Guillet, J. Ma, J. J. Nielsen, L. Ping, R. Raulefs, J. Rodriguez, D. Slock, D. Yang, and S. Zazo, "Location aided wireless communications," *IEEE Commun. Mag.,* to be published.

[15] M. T. Sun, L. Huang, S. Wang, A. Arora, and T. H. Lai, "Reliable MAC layer multicast in IEEE 802.11 wireless networks," *Wireless Commun. Mobile Comput.*, vol. 3, no. 4, pp. 439–453, 2003.

[16] N. Wen and R. Berry, "Information propagation for location-based MAC protocols in vehicular networks," in *Proc. Annu. Conf. Information Sciences and Systems*, 2006, pp. 1242–1247.

[17] S. B. Kodeswaran and A. Joshi, "Using location information for scheduling in 802.15.3 MAC," in *Proc. Int. Conf. Broadband Networks*, 2005, pp. 718–725.

[18] S. Katragadda, C. N. S. G. Murthy, R. Rao, S. M. Kumar, and R. Sachin, "A decentralized location-based channel access protocol for inter-vehicle communication," in *IEEE Vehicular Technology Conf.*, 2003, vol. 3, pp. 1831–1835.

[19] Y. B. Ko and N. H. Vaidya, "Geocasting in mobile ad hoc networks: Location-based multicast algorithms," in *Proc. IEEE Workshop on Mobile Computing Systems and Applications*, 1999, pp. 101–110.

[20] J. J. Nielsen, "Location based network optimizations for mobile wireless networks—A study of the impact of mobility and inaccurate information," Ph.D. dissertation, Dept. Electronic Systems, Aalborg Univ., Mar 2011.

[21] F. Cadger, K. Curran, J. Santos, and S. Moffett, "A survey of geographical routing in wireless ad-hoc networks," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 2, pp. 621–653, 2013.

[22] A. M. Popescu, N. Salman, and A. H. Kemp, "Geographic routing resilient to location errors," *IEEE Wireless Commun. Lett.*, vol. 2, no. 2, pp . 203–206, 2013.

[23] R. Di Taranto and H. Wymeersch, "Simultaneous routing and power allocation using location information," in *Proc. Asilomar Conf. Signals, Systems and Computers*, 2013, pp. 1700–1704.

[24] M. J. Neely, E. Modiano, and C. E. Rohrs, "Dynamic power allocation and routing for time-varying wireless networks," *IEEE J. Select. Areas Commun.*, vol. 23, no. 1, pp. 89–103, 2005.

[25] W. Su, S. J. Lee, and M. Gerla, "Mobility prediction and routing in ad hoc wireless networks," *Int. J. Network Manage.*, vol. 11, no. 1, pp. 3–30, 2001.

[26] K. T. Feng, C. H. Hsu, and T. E. Lu, "Velocity-assisted predictive mobility and location-aware routing protocols for mobile ad-hoc networks," *IEEE Trans. Veh. Technol.*, vol. 57, no. 1, pp. 448–464, 2008.

[27] N. Marmasse and C. Schmandt, "Location-aware information delivery with ComMotion," in *Proc. Handheld and Ubiquitous Computing*, 2000, pp. 157–171.

[28] B. Li and K. H. Wang, "NonStop: Continuous multimedia streaming in wireless ad hoc networks with node mobility," *IEEE J. Select. Areas Commun.*, vol. 21, no. 10, pp. 1627–1641, 2003.

[29] M. D. Dikaiakos, A. Florides, T. Nadeem, and L. Iftode, "Location-aware services over vehicular ad-hoc networks using car-to-car communication," *IEEE J. Select. Areas Commun.*, vol. 25, no. 8, pp. 1590–1602, 2007.

[30] M. F. Younis, K. Ghumman, and M. Ektiweissy, "Location-aware combinatorial key management scheme for clustered sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 17, no. 8, pp. 865–882, 2006.

[SP]