Nils Richerzhagen, Roland Kluge, Björn Richerzhagen, Patrick Lieser, Boris Koldehofe, Ioannis Stavrakakis and Ralf Steinmetz: **Better Together: Collaborative Monitoring for Location-based Services**, In Proc. 19th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), IEEE, 2018, ISBN 978-1-5386-4725-7

Better Together: Collaborative Monitoring for Location-based Services

Nils Richerzhagen*, Roland Kluge‡, Björn Richerzhagen*, Patrick Lieser*,
Boris Koldehofe*, Ioannis Stavrakakis† and Ralf Steinmetz*
* Multimedia Communications Lab (KOM), Technische Universität Darmstadt, Germany
{nils.richerzhagen|bjoern.richerzhagen|patrick.lieser|boris.koldehofe|ralf.steinmetz}@kom.tu-darmstadt.de
‡ Real-Time Systems Lab (ES), Technische Universität Darmstadt, Germany
roland.kluge@es.tu-darmstadt.de
† National and Kapodistrian University of Athens, Greece – IMDEA Networks & UC3M, Spain
ioannis@di.uoa.gr

***Abstract*—Mobile applications increasingly rely on frequent and accurate position updates—e.g., with GPS- or Wi-Fi–assisted localization techniques—to provide for functionality to their users. The service quality and acceptance of the application depend strongly on the localization accuracy and the introduced costs, in form of the resource consumption, of the used localization technique. Current mechanisms for location retrieval, however, are limited to non-mobile scenarios or still introduce high costs while obtaining the location. In this work, we propose a collaborative location retrieval service for location-based services in mobile scenarios that combines the location information of a subset of users with the connectivity information between users to enable accurate and cost-efficient location estimations. We evaluate a prototype of our solution to study the impact of service compositions in changing environments and to assess the potential of our proposed service compared to the current state-of-the-art used within location-based services. Our results reveal that, depending on the localization technique, the costs can be reduced significantly while the achieved sensing accuracy and fairness among users improves strongly at the same time.**

I. Introduction

Wireless networking communications and technologies witness an evolution in both their growth and their applications. Especially in future Internet scenarios, such as the Internet of Things (IoT) and mobile social networks, wireless communications offer great potential. Many of the applications used in these scenarios are implicitly bound to the users and their personal surroundings. A very prominent class of applications are location-based services, which comprise, e.g., augmented reality games, such as Google *Ingress* and *Pokémon Go*. This class of applications requires frequent and accurate location updates of the users to provide functionality. Unfortunately, accurate location retrieval can be difficult with localization technologies such as the global position system (GPS) or Wi-Fi / cellular triangulation, offering diverse accuracies. Furthermore, erroneous sensors readings may further decrease the accuracy. With frequent individual location measurements, the acceptance of the location-based services shrinks as every accurate location measurement is a battery-intensive task [1]. Especially on resource-constrained devices, high battery usage is unintended.

978-1-5386-4725-7/18/$31.00 ©2018 IEEE

Current state-of-the-art mechanisms for location retrieval in mobile wireless scenarios rely on fingerprinting, geometric estimations [2], or local sharing of positions between devices that have access to accurate positions and those who have not [3], [4]. These mechanisms have one thing in common: they pursue a given utility function—mostly providing accurate position estimations—while not considering the introduced cost or the fairness of the resources consumed on the mobile devices. Additionally, many of the approaches rely on measurements of every device having access to localization technologies, irrespective of the accuracy or cost of the specific localization technology.

In this work, we target the currently missing joint consideration of (i) position accuracy, (ii) cost-awareness, and (iii) fair usage of resources within location retrieval in mobile wireless scenarios. To this end, we propose a collaborative monitoring approach for location-based services on the example of location estimation. Our proposed service combines the selection of relevant anchor nodes, which measure their location, with connectivity information between neighboring nodes to provide for accurate and cost-efficient position estimations of all nodes in the network. By design, based on an in-depth analysis of current location estimation approaches, the proposed collaborative monitoring service is able to pursue different utility functions such as low energy usage, fair resource consumption, and high accuracy. We rely on the Simonstrator platform [5], [6] to assess the performance characteristics of the proposed collaborative location retrieval service and to compare it with the current de facto standard (i.e., GPS- and Wi-Fi–assisted location retrieval). In our evaluation, we analyze (i) the impact of different compositions of our proposed service, such as the used anchor node selection strategy, and the influence of changing environmental conditions on the position accuracy, cost, and fairness. Additionally, (ii) we assess the performance–cost–trade-off using our proposed service in comparison with the de facto standard used for location-based services today. Our results reveal that, depending on the location-sensing technology at hand, the introduced cost can be reduced significantly while the achieved sensing accuracy and fairness can be improved considerably at the same time.

Nils Richerzhagen, Roland Kluge, Björn Richerzhagen, Patrick Lieser, Boris Koldehofe, Ioannis Stavrakakis and Ralf Steinmetz: **Better Together: Collaborative Monitoring for Location-based Services**, In Proc. 19th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), IEEE, 2018, ISBN 978-1-5386-4725-7

The remainder of this paper is structured as follows. In Section II, characteristics of the scenario are detailed. The related work on location estimation in mobile networks is classified and discussed in Section III. Details of the design of the proposed collaborative monitoring service for location estimation, which is the core contribution of this work, are given in Section IV. The proposed service relies, other than its related work, on the selection of relevant anchor nodes, which execute the location measurement in combination with connectivity information of the mobile nodes to estimate the locations of all nodes. In Section V, we evaluate a prototype of our proposed service and present the results in an in-depth simulation-based evaluation. Finally, Section VI concludes.

II. Scenario

In the scenario considered in this work, mobile devices are able to connect to the Internet and to other mobile devices based on ad hoc wireless communication technologies such as Wi-Fi ad hoc or Bluetooth LE. This enables a multitude of mobile social applications such as the well-known augmented reality games Google *Ingress* and *Pokémon Go*. For all of these location-based services, determining the user's location continuously is essential.

As mobile social applications exhibit strong bounds to the user and her smart device, the prevailing environmental conditions are determined by the users' movement and the current location characteristics. The movement of people is influenced by their attraction to specific places, social ties, and interaction patterns with others. Obviously, obstacles such as roads and buildings need to be considered as well. Thus, people are restricted to walk on streets and pathways in the urban area, while being able to directly communicate with others in the range of the currently used communication interface. The smart devices, carried by the users, are equipped with off-the-shelve communication interfaces for cellular communication, Wi-Fi, Bluetooth and Bluetooth LE.

We assume that the proposed location retrieval service is running on all devices carried by the users and a cloud-based entity. Additionally, a location-based service runs on the mobile devices to generate location requests in certain intervals as workload for the proposed location retrieval service. This location-based service is used to assess the accuracy and achieved utility function of the location estimation service under changing environmental conditions.

III. Related Work

Location retrieval approaches such as LOCALE or GMAN [7], [8] or the work by Kampis [9] focus on location estimation in sparsely populated networks based on readings of positions of neighbors over time used for prediction of the own position. However, these approaches are relying on all nodes having access to localization techniques to perform the measurement to provide for accurate results, which is not cost-efficient. Furthermore, nodes are assumed to be non-mobile, which is unrealistic in today's networks and not sensible in the area of mobile social applications. Chan et al. [10] propose an approach for indoor localization of mobile nodes. However, the assumed maximum distance of 0.5m between nodes in the formed clusters is not feasible in most scenarios. Additionally, assuming the presence of static anchor nodes used to re-calibrate positioning errors needs pre-configuration of the environment, which further limits the usability of the approach presented in [10]. A collaborative approach for location estimation relying on so-called multidimensional scaling and maximum-likelihood estimation is proposed in [3]. Their approach focuses on accurate location retrieval for non-mobile environments with densely interconnected nodes in a grid structure. In a static network of sensors, a grid structure may be plausible; however, in the presence of mobile social networks, this assumption does not hold true. The authors of [4] use GPS-equipped mobile anchor nodes in a wireless sensor network to provide for location information by broadcasting the location information to non-mobile nodes. Hu et al. [2] use the Monte Carlo Localization method for location retrieval in mobile sensor networks. Still, introducing additional overhead for local communication among nodes limits the applicability of the approach. Availability of neighborhood information, as used in our approach, in location-based services is likely.

Similar to the approaches presented before, the authors of [3] also rely on reference points that know their exact position without any additional error. None of the approaches focus on the per-node or overall cost of the location retrieval process, which is crucial concerning the resource-constrained devices in today's and future networking scenarios. Doherty et al. [11] suggest to use connectivity information in the location retrieval process similar to the proposed approach in this work. But, the authors assume the following unrealistic requirements (i) error-free sensor readings, (ii) non-mobile nodes, and (iii) a-priori knowledge of the network characteristics. The authors limit their work on a rectangular scenario with different densities while needed anchor nodes are placed, beneficial for their approach in the corners of the network, based on the a-priori knowledge of the network characteristics.

The influence of location error on the application performance is discussed using the example of a geo-based content sharing approach in [12]. The results indicate that often used movement models such as the random waypoint [13] or the Gaussian movement model [14] exhibit stronger dependency to location errors because users do not interconnect as frequently as in movement models that incorporate realistic social ties between users. Obviously, the authors of [12] indicate that the impact of location error is depending on the application as well as on the location retrieval approach. Current state-of-the-art location retrieval approaches are mainly focusing on the accuracy of the location estimation they provide. Cost and fairness of the resource consumption are considered rarely. However, especially with respect to today's resource-constrained smart devices and different application requirements, a joint consideration of (i) position accuracy, (ii) cost-awareness, and (iii) fair usage of resources within location retrieval is essential to allow for wide acceptance and applicability.

Nils Richerzhagen, Roland Kluge, Björn Richerzhagen, Patrick Lieser, Boris Koldehofe, Ioannis Stavrakakis and Ralf Steinmetz: **Better Together: Collaborative Monitoring for Location-based Services**, In Proc. 19th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), IEEE, 2018, ISBN 978-1-5386-4725-7

IV. CoMon: Collaborative Location Retrieval

The goal of this work is to enable the joint consideration of (i) position accuracy, (ii) cost-awareness, and (iii) fair usage of resources within location retrieval. To this end, we propose the collaborative location retrieval approach for location-based services CoMon. The service allows for different compositions of its components to pursue various utility functions for location retrieval. Furthermore, instead of relying on every node's measurements CoMon is able to select a subset of anchor nodes for cost-intensive location measurement, resulting in a significant reduction of the induced cost while still providing accurate position estimations. In the following, the design requirements of CoMon are outlined. Afterwards, the main components of CoMon are explained in detail. This includes the anchor point selection and clustering strategies used in our approach. Additionally, we elaborate on how different utility functions can be mapped with CoMon.

Design of the Collaborative Location Retrieval Service

The collaborative location retrieval service needs to allow for accurate location estimation of all nodes in the network while permitting different utility functions based on the requirements of the location-based services. CoMon combines the dynamic selection of anchor nodes—a subset of the nodes in the network that measure their location—and connectivity information of the nodes to allow for cost-efficient but still accurate location estimation. Accordingly, CoMon contains the following components (see Figure 1). The *anchor selection* and *clustering components* are responsible for determining anchor nodes and represent the most influential components with respect to the utility function of our approach. In the *layout component*, the location measurements of the selected anchor nodes are combined with the connectivity information obtained by the nodes to estimate each node's location. As the receiving range of the communication mean used to obtain the connectivity graph is not known a-priori the *range learner component* is used to learn the maximum communication range over time. Different service compositions—also allowing for multiple utility functions—are organized within the *composition component*.

a) Anchor Node Selection and Clustering: Selection and clustering is also highly relevant in the research area of offloading applications [15], [16]. However, most of the approaches used in this area need recent position measurements of all nodes. Thus, beside applicable selection strategies proposed for offloading, such as energy-aware approaches [17], [18], we propose selection strategies optimized for the use case of location estimation in this work. The proposed strategies are all based on information gained from the connectivity graph or available within CoMon to reduce further costs from otherwise needed additional measurements. Those strategies are used within the *selection component*, as shown in Figure 1.

Related to the centrality in social networks [16], [19], selecting anchor nodes that are well connected to other nodes and central according to [19] in the connectivity graph seems beneficial. It would cause anchors to have many adjacent nodes, resulting in potentially lower offsets for the estimated positions. However, as initial evaluation results revealed, such a selection leads to a very dense population of anchor nodes around the detected central point of the connectivity graph leading to strong offset variations on the edges of the connectivity graph. As a result, we use the *negative interference selection* (NegInt) strategy. This strategy selects the best-connected node as anchor, but then removes all nodes with a direct connection to the selected anchor node from the list of potential anchors. This allows for distances of at least two hops between anchors, which still provides for a sufficient amount of anchors in dense populated areas but results in a better density-aware coverage of the network.

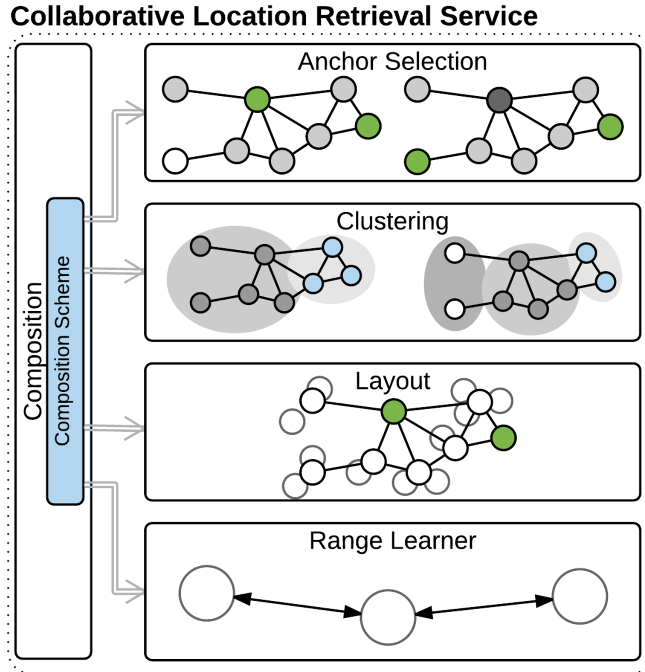


Figure 1: Illustration of the components of CoMon.

Contrary to the negative interference selection, the *cluster corner selection* (ClusCor) strategy selects anchor nodes from the outside of a cluster resulting in other non-anchor nodes being surrounded by anchor nodes. If non-anchor nodes are very likely within a given area spanned by the selected anchors, the estimation of the individual locations can benefit from the selected boundaries. However, without location information, the selection of anchors becomes more demanding. The authors of [20] found a correlation between the number of connected neighbors with the distance to the center of a node in a cluster. Accordingly, we start from the best-connected node(s) in the connectivity graph and exclude these node(s) and their neighbors. The resulting ring of nodes around the excluded nodes forms the initial set of potential anchor nodes. The process of excluding further nodes if they have more neighbors can be continued from there on. We stop this process after three iterations, which is sufficient for most location-aware clusters prevalent in mobile networks [21].

Nils Richerzhagen, Roland Kluge, Björn Richerzhagen, Patrick Lieser, Boris Koldehofe, Ioannis Stavrakakis and Ralf Steinmetz: **Better Together: Collaborative Monitoring for Location-based Services**, In Proc. 19th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), IEEE, 2018, ISBN 978-1-5386-4725-7

Both approaches presented above do not consider the cost introduced by the selected anchors as they are tuned toward coverage. To this end, we rely on the cost-aware *minimum cost selection* (MinC) and the *fair-cost selection* (FairC) strategies. The minimum cost selection reduces (i) the number cost-intensive location measurements, and (ii) the number of state changes on the nodes, i.e., how often cost-intensive sensors such as GPS are turned on and off. While this strategy may reduce the number of state changes to allow for less expensive continuous measurements on the nodes for a longer period of time, the energy consumption is not distributed among nodes in a fair manner. A fair share of the overall energy consumption is important to not only allow for longer device lifetimes but also for better acceptance of the approach. We target this limitation with the fair-cost selection strategy. Here, fairness is most important while still reducing the overall energy consumption. Thus, few anchor nodes are selected but, in contrast to the minimum cost approach, the nodes that are selected as anchors are changed more frequently over time to provide for a fairer utilization of resources on the nodes.

Clustering within CoMon is used to split the complete connectivity graph into smaller sub-graphs. This may lead to more accurate position estimations because subsequent computations (anchor selection and layout) are performed on more regional graph structures. For clustering, we use (i) a variant of the well-known density-based DBScan [22] clustering algorithm called *connection-based clustering* (CBC), (ii) the partition-based k-Means++ [23], and (iii) a grid-density–approach, which builds upon [24]. The original DBScan algorithm may classify some nodes as *noise*, which is contradictory to the requirement of CoMon to provide for location estimations of all nodes in the network. Thus, nodes that are not part of a cluster resulting from DBScan are individually assigned to their own one-node clusters enabling the reliable selection of anchors per-cluster for all nodes. CBC starts with the best-connected nodes unlike the random node sampling that DBScan uses. It then iterates over each node, in descending order of the number of connections each node has. For each node, the direct one-hop neighborhood is searched for cluster affiliation. If no affiliation is found, the node is assumed to be part of its own cluster, guaranteeing position estimation for all nodes. The idea behind CBC is that well-connected nodes are likely close together in each cluster. By starting with the most connected node, and then iterating outwards, we gather all nodes belonging to that cluster, but not the nodes that are closer to another equally well-connected node.

b) Layout of the Connectivity Graph: The selected anchor nodes, representing a small subset of all nodes, measure their position using a localization method such as GPS or Wi-Fi triangulation. This information is combined with the one-hop connectivity information of all nodes, resulting in a connectivity graph per-cluster or for the whole network. Still the resulting connectivity graphs might be partitioned. To estimate the positions of non-anchor nodes utilizing the connectivity graph we rely on a spring force layout algorithm based on [25], [26] as well as on the layout algorithm proposed by Kamada and Kawai in [27]. However, int the layout the distance between locations of nodes without a connection but with line-of-sight are further apart than the maximum communication range of the used communication mean must be considered. Thus, connected nodes must not be further apart from each other than the maximum communication range.

Within the layout(s), anchor nodes are fixed in their position, while non-anchor nodes are displaced after calculating so-called *forces* for each node. The calculation of forces and the following displacement is referred to as *step* [25]. After a given amount of steps, the layout algorithm terminates with the estimated positions for all non-anchor nodes. In the used spring force layout, we iterate once over each node pair, without repetition. If the node pair has a connection in the connectivity graph, only attractive forces are computed. If there is no connection between the pair, repulsive forces are only computed if the two nodes positions in the layout are closer together than the currently assumed communication range (as provided by the range learner component). With this needed customization of [25], [26], the computational overhead is reduced drastically, as pairwise relations are taken into account only once. The resulting complexity for n nodes per step is given by Equation (1).

$$\mathcal{O}\binom{n}{2} = \sum_{m=1}^{n-1} (m) = \frac{(n-1)^2 + (n-1)}{2} \tag{1}$$

Node distances do not necessarily behave like actual springs. The likelihood for connected nodes being closer together than the assumed maximum distance is larger than it is for the case that they are much further apart. Accordingly, the utilized spring forces are non-uniform. Instead of using a single spring force multiplier ς as in [26], we rely on three spring forces to improve the approximation of actual node distances and the location estimation accuracy. A ς of 0.1 is used for connected nodes that are closer together than the average distance in the connectivity graph. For connected nodes whose estimated distance is larger than the maximum communication range a ς of 0.8 is used so that those nodes move together. A ς of 0.3 is used for connected nodes further apart than the average distance in the connectivity graph, but closer than the maximum communication range. The same ς of 0.3 is used for unconnected nodes that are closer together than the maximum communication range. In this case the desired distance is at least the maximum communication range obtained from the range learner component.

$$\mathcal{F}(d_{A\leftrightarrow B}) = \frac{d_{\text{desired}} - d_{A\leftrightarrow B}}{d_{A\leftrightarrow B}} \tag{2}$$

The spring force multiplier value ς for two nodes A and B is multiplied with the *distance dependent force* $\mathcal{F}(d_{A\leftrightarrow B})$ seen in Equation (2). To prevent high degree nodes from *slingshot* to another position we split the resulting forces between nodes A and B reverse proportionally to their respective number of connections. Thereby, a slingshot of a high-degree node in step i and a resulting drag of all connected nodes in the next step $i+1$ is prevented.

Nils Richerzhagen, Roland Kluge, Björn Richerzhagen, Patrick Lieser, Boris Koldehofe, Ioannis Stavrakakis and Ralf Steinmetz: **Better Together: Collaborative Monitoring for Location-based Services**, In Proc. 19th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), IEEE, 2018, ISBN 978-1-5386-4725-7

c) Learning the Communication Range: Communication ranges used in the layout algorithm may change over time, as nodes venture into areas with different environmental interference on the used communication mean. Considering the heterogeneity of devices, they may also change from device to device. Learning the possible communication ranges at runtime not only improves the system's adaptability to environmental changes but also makes it more resistant to inaccurate initial assumptions. The distances are computed from information in the connectivity graph and accompanying locations of anchor nodes. However, with the intended limited number of position measurements, respectively anchor nodes, we also incorporate the complete but biased values generated by the layout algorithm for the learning process.

With no prior information on how the distance values behave or change, as the kind and amplitude of change is also unknown, we use two kinds of learners in addition to a *No Learning* baseline. These are based on filter methods used for example in signal processing. In our approach, they serve as low pass filters, minimizing the effects of jitter values extracted from the input.

The *exponential smoothing algorithm* takes the smoothed communication estimate of the former measurement interval i_{t-1} into account for the current interval i_t. The estimated distance e_t for interval i_t is computed as given by Equation (3).

$$e_t = \alpha \cdot d + (1 - \alpha) \cdot e_{t-1} \tag{3}$$

Here, d is the current computed distance and α is a weighting factor for the ratio of the current value over the value from the previous interval. The *autoregressive filter algorithm* is similar to the exponential smoothing algorithms, but it takes N previous outputs into account. The formula of the filter is given by Equation (4).

$$e_t = \alpha \cdot d + (1 - \alpha) \cdot \frac{1}{N} \cdot \sum_{n=1}^{N+1} e_{t-n} \tag{4}$$

d) Multiple Utility Functions with Composition: To allow for different utility functions within the collaborative location retrieval service, the *composition component* is used. It determines the order in which the main components, i.e., anchor selection, clustering, and layout, are used and combined. The outcome of the collaborative location retrieval service changes if, e.g., the anchor selection is done on a per-cluster basis instead of selecting anchors first and then building clusters around the anchor nodes. The used layout can be determined for all nodes (including potentially unconnected graphs) or on a per-cluster basis to reduce the computational load. With different combinations and used settings, a variation of utility functions can be pursued. However, a main influencing factor is the anchor selection, which determines the nodes that perform the measurements.

The *select-and-layout* (SL) composition starts with the anchor node selection followed by a layout of the connectivity graph with the selected anchors. Clustering is not used in this composition variant. The *cluster-select-layout* (CSL) composition generated clusters in the beginning, followed by the selection of anchor nodes in each cluster. For clustering the presented approaches DBScan [22], k-Means++ [23], and grid density [24] are used together with the proposed connection-based clustering approach. After the clusters and anchor nodes are computed, the CSL composition executes the layout in each cluster for the given amount of steps. This is followed by a final layout on the merged connectivity graph to reduce the potential of unlike placement decisions in clusters such as placing several nodes within a cluster at the same location.

The *select-layout-cluster* (SLC) composition is similar to the SL composition. However, a clustering is done after the selection and the initial layout, which is followed by an additional layout per-cluster. This is done to refine the position estimates after the initial selection and layout in the smaller clusters. In doing so, negative side effects caused in the layout by nodes far away from a cluster are prevented.

V. Evaluation

The evaluation of the proposed collaborative location estimation service addresses two main aspects. First, different compositions of the location estimation service are compared using variations in parameter and environmental settings (Section V-B). Here, the impact of parameter settings and environmental conditions is evaluated. Additionally, compositions of CoMon are compared with respect to different utility functions. Second, we assess the potential of the collaborative location estimation in comparison to the de facto standard, i.e., GPS- or Wi-Fi–assisted localization, in Section V-C. In the following section, we begin with explaining the setup as well as the metrics used in the evaluation.

A. Evaluation Setup, Scenario Model and Metrics

We evaluate the prototype of CoMon in the Simonstrator framework [5]. It allows us to use a social movement model that builds on OpenStreetMap map data for realistic node mobility [6]. We rely on the Wi-Fi 802.11g model from the ns-3 network simulator for local ad hoc communication between nodes [28]. As the connectivity characteristics are heavily influenced by human mobility, we compare the performance achieved with our social movement model against a Gaussian [14] and a random waypoint (RWP) [13] mobility model. We simulate mobility within the urban city center of Darmstadt with a area of $1500\,\mathrm{m} \times 1500\,\mathrm{m}$. Considering different node densities allows us to assess the usability of the presented approach in sparsely and densely populated scenarios. To assess the performance of CoMon and to allow for comparison with the state-of-the-art, we assume that the position information is requested in intervals of different length. Table I summarized the simulation setup, with default values being underlined. We simulated $30\,\mathrm{min}$ of operation and recorded measurements after a warm-up period of $10\,\mathrm{min}$.

To assess the performance of our proposed collaborative location retrieval service, we consider the following metrics: (i) the achieved *position offset* (distance between the true and the estimated position of a node) directly after the computations

Nils Richerzhagen, Roland Kluge, Björn Richerzhagen, Patrick Lieser, Boris Koldehofe, Ioannis Stavrakakis and Ralf Steinmetz: **Better Together: Collaborative Monitoring for Location-based Services**, In Proc. 19th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), IEEE, 2018, ISBN 978-1-5386-4725-7

of COMON, (ii) the *offset over time* sampled every $5\,\mathrm{s}$, which changes with longer measurement intervals and the movement of nodes, (iii) the *measurement state ratios* of the sensors on the nodes as they allow us to assess the cost introduced with the location retrieval service, and (iv) *Jain's fairness index (*FI*)* [29] as given in Equation (5) to assess the fair distribution of the cost-intensive measurement tasks. Both offset metrics are used to assess the accuracy of the approach.

$$\mathrm{FI}(X) = \frac{[\sum_{i=1}^{n} x_i]^2}{n * \sum_{i=1}^{n} {x_i}^2} \tag{5}$$

It ranges between 0 and 1, while an index of 0.1 implies that the system is fair to $10\,\%$ of the users. An index of 1 implies fair resource sharing for all users. However, the index does not imply to which extent resources are consumed. For example, with the cost-intensive measurement task, the index might result in a $100\,\%$ fair share, but the total number of cost-intensive measurement tasks performed by each user may highlight potential weaknesses of a system. Thus, when considering quantitative fairness metrics, we must also take into account the total value of the shared resource.

We rely on the measurement state of the sensor as abstract cost metric instead of using a specific cost model because the measurement state is more universal than a single cost model and provides the same insights to the costs. The used measurement states describe the life cycle phase of localization sensors (e.g., GPS), which are *activation*, *continuous measurement*, *deactivation* and *no measurement*. In the following we will use the *measurement state* as a combination of the *activation* and *continuous* state because these states introduce costs at the nodes. We observe the ratios of the measurement states because the overall number of states does only change with the measurement interval. Box plots show the distribution of the results. A solid line inside the box represents the median, while the lower (upper) quartile are represented by the boxes lower (upper) end of the box. Whiskers show the lower and upper data point within 1.5 of the interquartile range. As box plots show the results of a single simulation run, a marked dot with error bars is plotted to the left side of the boxes indicating the confidence intervals over 30 repetitions with different random seeds.

B. Impact of Parameter and Environmental Settings

The de facto standard in location-based services is to rely on measurements of all nodes—often using the accurate but cost-intensive GPS localization technique. We assume different sensing errors, depending on the localization techniques (e.g.,, GPS or Wi-Fi triangulation). In the following, we assume a perfect location sensing technique with zero error if not otherwise stated; the impact of the sensing error is assessed specifically in Section V-C.

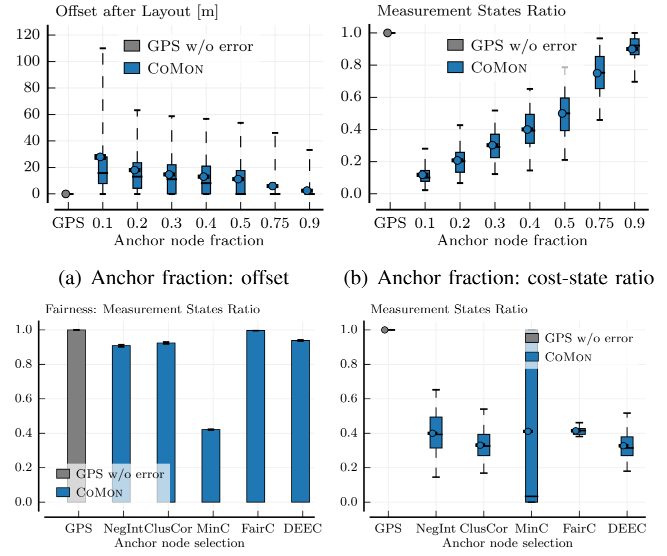
The number of selected anchor nodes influences the resulting estimation accuracy. With fewer selected anchors, the collaborative location retrieval service has to estimate the positions of a large share of the nodes in the layout. Here, the *select-and-layout* composition strategy with the negative interference selection (see Table I) is used, and the fraction of anchor nodes is varied between $10\,\%$ and $90\,\%$ for COMON whereas the baseline approach uses $100\,\%$ of the nodes as anchors. The position estimation achieved with COMON using only $40\,\%$ of the nodes as anchors is less accurate than the perfect zero-error localization (see Figure 2(a)). However, the localization error ranges between $0\,\mathrm{m}$ and $20\,\mathrm{m}$ for three quarters of the nodes, which is sufficiently accurate for many applications. Still, the localization error may go up to $55\,\mathrm{m}$ for nodes that are not well connected to neighbors, e.g., over a chain of nodes with only two adjacent nodes each. With more anchor nodes, the achieved offset improves as more fix points can be used in the layout. But, compared to the zero-error localization, where all nodes use their error-free position sensor, COMON reduces the cost significantly. Figure 2(b) shows the ratio of measurement states for the different anchor node fractions. With $40\,\%$ of the nodes being anchors, COMON reduces the ratio of nodes performing active measurements in average by a factor larger than two. Obviously, the introduced measurement costs show a linear dependency to the anchor node fraction. Still, the box ranges in Figure 2(b) indicate an unfair distribution of the measurement task among the nodes. Active measurement states on nodes make up between $18\,\%$ and $70\,\%$ of all states on the nodes (includign *deactivation* and *no measurement*).

Thus, the impact of different selection strategies on the introduced cost and, very importantly, the achieved fairness of the cost-intensive measurement task must be assessed. To this end, the anchor selection strategies presented in Section IV are used: negative interference (NegInt), cluster corner (ClusCor), minimum cost (MinC), fair-cost (FairC) and the energy-aware selection strategy DEEC [17]. Jain's fairness index [29], shown in Figure 2(c), reveals that the fair-cost selection strategy is able to provide for equal distribution of the measurement tasks. As explained before, while the error-free baseline approach, where all nodes measure their own position provides high fairness, too, we have to consider the extent of measurement tasks compared to other tasks. Figure 2(d) shows that, using

Table I: Scenario and simulation setup.

Max. Wi-Fi Range	$78\,\mathrm{m}$
Mov. Speed	$1.5\,\frac{\mathrm{m}}{\mathrm{s}}$–$2.5\,\frac{\mathrm{m}}{\mathrm{s}}$
Movement Model	social [6], gauss [14], RWP [13]
Density $[\frac{\mathrm{nodes}}{km^2}]$	22–222; 88.8
Baseline Systems	GPS ($0\,\mathrm{m}$), GPS (0–$15\,\mathrm{m}$), Wi-Fi (0–$130\,\mathrm{m}$)
Layout Steps	30
Measurement Interv.	$10\,\mathrm{s}$, $20\,\mathrm{s}$, $30\,\mathrm{s}$, $60\,\mathrm{s}$, $120\,\mathrm{s}$
Composition	select-and-layout (SL), cluster-select-layout (CSL), select-layout-cluster (SLC)
Anchor Selection	negative interference, cluster corner, minimum cost, fair-cost, DEEC [17]
Anchor Fraction	0.1, 0.2, 0.3, 0.4, 0.5, 0.75, 0.9
Clustering	connection-based (CBC), DBScan [22], k-Means++ [23], grid density [24]
Layout	spring force [26], Kamada Kawai [27]
Range Learner	none, exp. smooth, autoregressive
Sensing Error	no error ($0\,\mathrm{m}$), GPS (0–$15\,\mathrm{m}$), Wi-Fi (0–$130\,\mathrm{m}$)

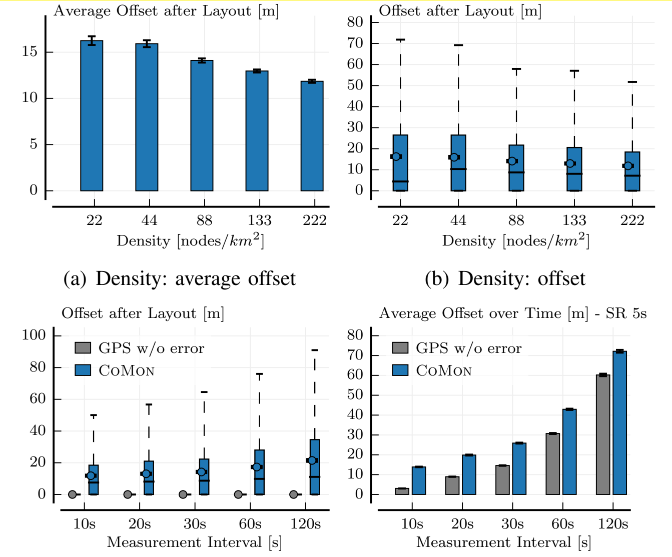
Nils Richerzhagen, Roland Kluge, Björn Richerzhagen, Patrick Lieser, Boris Koldehofe, Ioannis Stavrakakis and Ralf Steinmetz: **Better Together: Collaborative Monitoring for Location-based Services**, In Proc. 19th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), IEEE, 2018, ISBN 978-1-5386-4725-7



(a) Anchor fraction: offset (b) Anchor fraction: cost-state ratio

(c) Anchor selection: fairness (d) Anchor selection: cost-state ratio

Figure 2: Impact of the fraction of anchor nodes and the anchor node selection strategy.



(a) Density: average offset (b) Density: offset

(c) Interval: offset (d) Interval: avg. offset over time

Figure 3: Impact of the node density and the measurement interval.

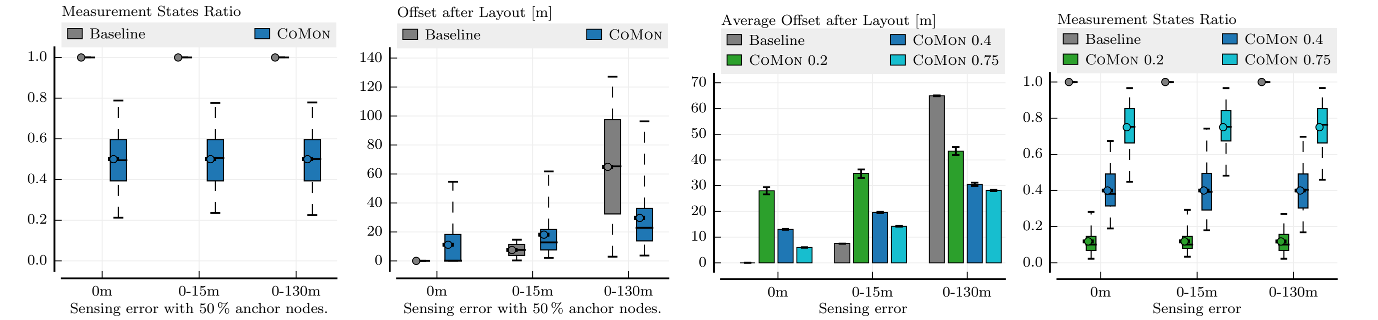
the baseline approach, all nodes need to use their cost-intensive measurements, whereas using the fair-cost selection strategy, only the subset of anchor nodes ($40\,\%$) is required to measure continuously. In total more measurements are required with the fair-cost strategy compared to the cluster corner (ClusCor) or DEEC selection strategies (see Figure 2(d)). However, compared to the minimum-cost selection strategy, where the anchor node fraction of $40\,\%$ is selected once and then the selected anchors are used for further measurement tasks, the equal distribution of measurement tasks among the nodes guarantees the fairness of the system (see Figure 2(c)). Furthermore, by distributing the measurement tasks equally among the nodes, the lifetime of the system can be improved because no node measures its location throughout the entire simulation.

The large box range for the minimum-cost selection strategy (see Figure 2(d)) is based on the characteristic that the strategy uses a set of nodes that is selected once to perform the measurements. Thus, $40\,\%$ of the nodes perform the measurement over the whole simulation while the other $60\,\%$ never measure their own position. Surprisingly, unlike the anchor selection, the used clustering approach does not has a strong influence on the estimation quality or introduced costs in CoMon. The reason is that with the selection strategy at hand a direct effect on the nodes is achieved; with clustering, in contrast, nodes may not be equally clustered, but still be selected as anchors.

As CoMon is using the connectivity information of nodes to estimate the positions of non-anchor nodes, the approach shows a clear dependency on the node density in the network (see Figure 3(a) and 3(b)). Denser environments result in more connections among nodes, which is beneficial for the layout used within CoMon. Still, our approach is able to estimate the positions of all nodes with an average offset of $16\,\mathrm{m}$ utilizing $40\,\%$ as anchor nodes even in sparse environments with populations of 22 nodes per km^2 (Figure 3(a)). Our approach reduces the introduced cost by $50\,\%$ compared to the baseline approach while still delivering sufficiently accurate position information for typical applications. In denser environments, CoMon benefits from more anchor nodes (while the fraction remains the same) and more connections between nodes that are used in the layout (see Figure 3(b)). The selected measurement interval has two main influencing factors. First, the amount of the resulting measurements, and thus the introduced cost, increases significantly with shorter intervals. Second, for mobile nodes, the retrieved location information is only valid for the point in time at which it was obtained. Thus, location measurements and estimations are time-dependent information—significantly influenced by the duration between new measurements and estimation calculations. Figure 3(c) shows the offset after the layout for the different measurement intervals ranging between $10\,\mathrm{s}$ and $120\,\mathrm{s}$. CoMon uses the position estimates from the previous measurement interval i_{t-1}. Hence, the offset increases with an increased measurement interval as the estimates are very likely less accurate with increasing age. The average offset over time (sampled every $5\,\mathrm{s}$) changes accordingly (see Figure 3(d)). With speeds distributed between $1.5\,\frac{\mathrm{m}}{\mathrm{s}}$ and $2.5\,\frac{\mathrm{m}}{\mathrm{s}}$, the offset over time increases up to $60\,\mathrm{m}$ even if all nodes measure their position using GPS without sensor error. We employ a measurement interval of $20\,\mathrm{s}$ if not otherwise stated (Table I).

Compared to the realistic social mobility model [6], used in the evaluation, the characteristics of CoMon remain the same with Gaussian [14] or random waypoint [13] mobility models. The random waypoint mobility model results in fewer interconnections between nodes. In contrast, with Gaussian

Nils Richerzhagen, Roland Kluge, Björn Richerzhagen, Patrick Lieser, Boris Koldehofe, Ioannis Stavrakakis and Ralf Steinmetz: **Better Together: Collaborative Monitoring for Location-based Services**, In Proc. 19th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), IEEE, 2018, ISBN 978-1-5386-4725-7



(a) 50 % anchors: cost-state ratio (b) 50 % anchors: offset (c) Average offset (d) Cost-state ratio

Figure 4: Performance and cost comparison of the collaborative location retrieval service compared to the baseline approaches with GPS-assisted or Wi-Fi–assisted position sensing.

movement, nodes tend to gather in the middle of the map, which is beneficial for CoMon, but unrealistic concerning the movement of nodes. The resulting evaluation figures and the improvements achieved with the composition and range learner are omitted here for space reasons.

C. *Potential of Collaborative Location Retrieval*

To assess the potential of our proposed collaborative location retrieval service CoMon, we compare our proposed system with the baseline system as used today. The baseline system describes the localization with either GPS or Wi-Fi triangulation techniques. Thus, we evaluate different sensing error models (with error range ε), as proposed in [30], while omitting the cellular triangulation because it results in high sensing errors of up to 600 m: (i) the no-error sensing model ($\varepsilon_{0\,\mathrm{m}}$), (ii) the GPS-assisted sensing error between zero and 15 m ($\varepsilon_{0\text{–}15\,\mathrm{m}}$), and (iii) the Wi-Fi–assisted sensing error between 0 m and 130 m ($\varepsilon_{0\text{–}130\,\mathrm{m}}$). The sensing errors follow a circular error probable with its radius set to $\varepsilon/2$. In an environment where GPS-assisted sensing is used, CoMon reduces the introduced cost of the location retrieval process significantly compared to the baseline, where all nodes run GPS localization. Figure 4(a) shows that, with CoMon, nodes are performing cost-intensive measurements for only a fraction of the overall measurement intervals. Here, with 50 % anchor nodes and the negative interference selection, the fraction of active measurements over all measurement intervals for the observed simulation time ranges in its extremes between 20 % and 80 %. On average, the needed cost-intensive measurements can be reduced by 50 % using CoMon. However, reducing the resource consumption of the nodes comes at the expense that the achieved offset of the estimated positions increases. Nevertheless, with only 50 % of the nodes selected as anchor nodes CoMon delivers offsets below 22 m for 75 % of the nodes which is only 10 m less accurate than the GPS-assisted baseline approach, as shown in Figure 4(b).

Using Wi-Fi–assisted position sensing is significantly less cost-intensive than GPS-assisted position sensing, but with sensing errors ranging between 0 m and 130 m $\varepsilon_{0\text{–}130\,\mathrm{m}}$ [30]. Thus, such techniques are often neglected as the achieved position accuracy is not meeting required standards. Figure 4(b) shows that, with CoMon, the achieved position accuracy improves considerably compared to the Wi-Fi–assisted baseline approach with sensor error $\varepsilon_{0\text{–}130\,\mathrm{m}}$. Whereas the offset for the Wi-Fi–assisted localization ranges between 30 m and 95 m for 75 % of the nodes, CoMon reduces the offset to a range between 15 m and 35 m—which is an improvement of up to 270 % (see Figure 4(b)). Additionally, by selecting a subset of nodes that use the Wi-Fi–assisted localization, the introduced costs are reduced further even if Wi-Fi–assisted sensing is already less cost-intensive than GPS-assisted sensing. Thus, by incorporating additional available connectivity information into the location retrieval service, as proposed in this work, the achieved sensing accuracy can benefit strongly while the introduced cost decreases significantly.

This trend is confirmed by Figure 4(c) and 4(d), which show the average offset after the layout and the ratio of cost-intensive states on the nodes for different anchor fraction configurations of CoMon compared to the respective baseline approach. With no sensing error $\varepsilon_{0\,\mathrm{m}}$ and the GPS-assisted error model $\varepsilon_{0\text{–}15\,\mathrm{m}}$, CoMon reduces the introduced cost, at the expense of reduced sensing accuracy shown in the increased offset (see Figure 4(c)). However, with 75 % of the nodes selected as anchor nodes, CoMon achieves average offsets only 10 m larger compared to the baseline approach, while still reducing the cost by 25 %. With Wi-Fi–assisted position sensing, the potential of CoMon unfolds as the Wi-Fi–based localization baseline approach is outperformed with respect to (i) the achieved sensing accuracy and (ii) the introduced costs. Additionally, with an anchor node fraction of only 40 %, CoMon nearly matches the achieved offset compared to using an anchor node fraction of 75 %, with both configurations doubling the achieved average sensing accuracy as the average offset is reduced by 50 %. Thus, with larger sensor errors, i.e., using techniques such as Wi-Fi–assisted sensing that are less expensive, it is sensible to select fewer anchors and to rely on the position estimations achieved with CoMon. Practically, this further decreases the introduced cost for smaller anchor fractions compared to the baseline approach (see Figure 4(d)).

Nils Richerzhagen, Roland Kluge, Björn Richerzhagen, Patrick Lieser, Boris Koldehofe, Ioannis Stavrakakis and Ralf Steinmetz: **Better Together: Collaborative Monitoring for Location-based Services**, In Proc. 19th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), IEEE, 2018, ISBN 978-1-5386-4725-7

VI. Conclusions

A plethora of applications and services for future Internet scenarios, such as the Internet of Things and mobile social networks, are of the class of location-based services. These share their need for frequent and accurate position updates to provide for functionality. However, accurate location retrieval can be cumbersome with different localization technologies such as GPS of Wi-Fi–assisted triangulation at hand. While technologies based on GPS readings provide for more accurate results, they introduce high costs, which reduce the acceptance of the users to use such services. Wi-Fi–assisted technologies may reduce the cost but lack the needed accuracy.

In this paper, we therefore propose the combination of location and connectivity information between users to provide for accurate position estimations at reduced cost. To this end, we introduce the collaborative location retrieval service CoMon that estimates each user's position by using position measurements of a fraction of selected nodes in combination with connectivity information among users. We analyze the impact of different compositions of CoMon with respect to changing environmental conditions and the potential of CoMon compared to GPS- and Wi-Fi–assisted localization in an in-depth simulation-based evaluation. In doing so, we focus on (i) the achieved sensing accuracy, (ii) the introduced costs of the location retrieval, and (iii) the fairness among users with respect to the cost-intensive measurement task. Our results show that, depending on the location sensing technology, CoMon reduces the introduced cost significantly and improves the achieved sensing accuracy and the fairness among users.

We are currently investigating how movement prediction can improve the location estimation accuracy achieved CoMon. If the movement vector of individual nodes can be predicted, the layout technique used in our approach can benefit from the additional information when estimating the positions of the mobile nodes. Beside this, we are planning a experimental analysis of the proposed service to assess the services performance under real-world influences, such as node churn.

Acknowledgment

This work has been funded by the German Research Foundation (DFG) as part of projects A1, B1, C2, and C3 within the Collaborative Research Centre (CRC) 1053 – MAKI; the LOEWE initiative (Hesse, Germany) within the NICER project; and the German Federal Ministry of Education and Research within the SMARTER project (Support Code: 13N13406).

References

[1] L. A. Tawalbeh, A. Basalamah, R. Mehmood, and H. Tawalbeh, "Greener and Smarter Phones for Future Cities: Characterizing the Impact of GPS Signal Strength on Power Consumption," *IEEE Access*, vol. 4, pp. 858–868, 2016.

[2] L. Hu and D. Evans, "Localization for Mobile Sensor Networks," in *ACM MobiCom*, 2004.

[3] X. Li, "Collaborative Localization with Received-Signal Strength in Wireless Sensor Networks," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 6, pp. 3807–3817, 2007.

[4] K.-F. Ssu, C.-H. Ou, and H. C. Jiau, "Localization with Mobile Anchor Points in Wireless Sensor Networks," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 3, pp. 1187–1197, 2005.

[5] B. Richerzhagen, D. Stingl, J. Rückert, and R. Steinmetz, "Simonstrator: Simulation and Prototyping Platform for Distributed Mobile Applications," in *ACM SIMUTOOLS*, 2015.

[6] N. Richerzhagen, B. Richerzhagen, D. Stingl, and R. Steinmetz, "The Human Factor: A Simulation Environment for Networked Mobile Social Applications," in *IEEE NetSys*, 2017.

[7] P. Zhang and M. Martonosi, "LOCALE: Collaborative Localization Estimation for Sparse Mobile Sensor Networks," in *IEEE IPSN*, 2008.

[8] B. Zhang, Fengqi, and Z. Zhang, "Collaborative Localization Algorithm for Wireless Sensor Networks using Mobile Anchors," in *PACIIA*, 2009.

[9] G. Kampis, J. W. Kantelhardt, K. Kloch, and P. Lukowicz, "Analytical and Simulation Models for Collaborative Localization," *Journal of Computational Science*, vol. 6, pp. 1–10, 2015.

[10] L.-w. Chan, J.-r. Chiang, Y.-c. Chen, C.-n. Ke, J. Hsu, and H.-h. Chu, "Collaborative Localization: Enhancing WiFi-based Position Estimation with Neighborhood Links in Clusters," *Perv. Comp.*, pp. 50–66, 2006.

[11] L. Doherty, K. S. J. Pister, and L. E. Ghaoui, "Convex Position Estimation in Wireless Sensor Networks," in *IEEE INFOCOM*, 2001.

[12] J. Ott, L. Kärkkäinen, E. A. Walelgne, A. Keränen, E. Hyytiä, and J. Kangasharju, "On the Sensitivity of Geo-based Content Sharing to Location Errors," in *IEEE/IFIP WONS*, 2017.

[13] T. Camp, J. Boleng, and V. Davies, "A Survey of Mobility Models for ad hoc Network Research," *Wireless Communications and Mobile Computing*, vol. 2, no. 5, pp. 483–502, 2002.

[14] B. Liang and Z. J. Haas, "Predictive Distance-based Mobility Management for PCS Networks," in *IEEE INFOCOM*, 1999.

[15] F. Rebecchi, M. D. de Amorim, V. Conan, A. Passarella, R. Bruno, and M. Conti, "Data Offloading Techniques in Cellular Networks: A Survey," *IEEE Comm. Surv. Tut.*, vol. 17, no. 2, pp. 580–603, 2015.

[16] M. V. Barbera, A. C. Viana, M. D. de Amorim, and J. Stefa, "Data Offloading in Social Mobile Networks through VIP Delegation," *Ad Hoc Networks*, vol. 19, pp. 92–110, 2014.

[17] L. Qing, Q. Zhu, and M. Wang, "Design of a Distributed Energy-efficient Clustering Algorithm for Heterogeneous Wireless Sensor Networks," *Computer communications*, vol. 29, no. 12, 2006.

[18] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient Communication Protocol for Wireless Microsensor Networks," in *IEEE HICSS*, 2000.

[19] L. C. Freeman, "Centrality in Social Networks Conceptual Clarification," *Social Networks*, vol. 1, no. 3, pp. 215–239, 1978.

[20] P. Kang and S. Cho, "k-Means Clustering Seeds Initialization based on Centrality, Sparsity, and Isotropy," *Intelligent Data Engineering and Automated Learning*, pp. 109–117, 2009.

[21] A. Bentaleb, A. Boubetra, and S. Harous, "Survey of Clustering Schemes in Mobile Ad Hoc Networks," *Communications and Network*, vol. 5, no. 02, p. 8, 2013.

[22] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," in *KDD*, vol. 96-34, 1996.

[23] A. K. Jain, "Data Clustering: 50 Years Beyond k-Means," *Pattern Recognition Letters*, vol. 31, no. 8, 2010.

[24] M. de Berg, M. van Kreveld, M. Overmars, and M. Schwarzkopf, *Computational Geometry*. Springer-Verlag, 2000, ch. Quadtrees.

[25] T. M. Fruchterman and E. M. Reingold, "Graph Drawing by Force-directed Placement," *Software: Practice and experience*, vol. 21, no. 11, pp. 1129–1164, 1991.

[26] J. O'Madadhain, D. Fisher, S. White, and Y. Boey, "The JUNG (Java Universal Network/Graph) Framework," *University of California*, 2003.

[27] T. Kamada and S. Kawai, "An Algorithm for Drawing General Undirected Graphs," *Information Proc. Letters*, vol. 31, no. 1, pp. 7–15, 1989.

[28] T. R. Henderson, S. Roy, S. Floyd, and G. F. Riley, "ns-3 Project Goals," in *ACM WNS2*, 2006.

[29] R. Jain, *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling*. John Wiley & Sons, 1991.

[30] P. A. Zandbergen, "Accuracy of iPhone Locations: A Comparison of Assisted GPS, WiFi and Cellular Positioning," *Transactions in GIS*, vol. 13, no. 1, pp. 5–25, 2009.