Journal of Computational Science 6 (2015) 1–10

Contents lists available at ScienceDirect

Journal of Computational Science

journal homepage: www.elsevier.com/locate/jocs

Analytical and simulation models for collaborative localization

George Kampis[a,b,*], Jan W. Kantelhardt[c], Kamil Kloch[a], Paul Lukowicz[a]

[a] German Research Center for Artificial Intelligence (DFKI GmbH), Kaiserslautern, Germany
[b] ITMO University, St. Petersburg, Russia
[c] Institute of Physics, Martin-Luther-University Halle-Wittenberg, Germany

ARTICLE INFO

Article history:
Received 22 July 2014
Received in revised form
25 September 2014
Accepted 29 September 2014
Available online 13 October 2014

Keywords:
Analytical models
Agent based modeling
Pedestrian dead reckoning system
Automated localization
Interacting mobile computers
Knowledge fusion

ABSTRACT

Collaborative localization is a special case for knowledge fusion where information is exchanged in order to attain improved global and local knowledge. We propose analytical as well as agent based simulation models for pedestrian dead reckoning (PDR) systems in agents collaborating to improve their location estimate by exchanging subjective position information when two agents are detected close to each other. The basis of improvement is the fact that two agents are at approximately the same position when they meet, and this can be used to update local position information. In analytical models we find that the localization error remains asymptotically finite in infinite systems or when there is at least one immobile agent (i.e. an agent with a zero localization error) in the system. In the agent model we tested finite systems under realistic (that is, inexact) meeting conditions and tested localization errors as function of several parameters. We found that a large finite system comprising hundreds of users is capable of collaborative localization with an essentially constant error under various conditions. The presented models can be used for predicting the improvement in localization that can be achieved by a collaboration among several mobile computers. Besides, our results can be considered as first steps toward a more general collaborative (incremental) form of knowledge fusion.

1. Introduction

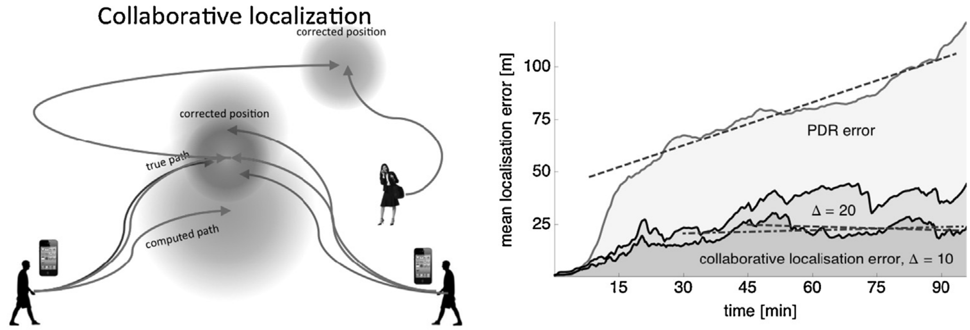
Automated localization of mobile computers is a well studied problem in informatics that is important for a broad range of mobile and ubiquitous computing applications [1]. Unfortunately it is a hard problem that has no reliable general purpose solution for all application domains [2]. Besides, automated localization systems typically require the use of external position signals such as GPS for at least some of the agents involved. The design space of cutting edge localization systems is thus bounded by two extremes (1) expensive systems that provide accurate, reliable location at the cost of extensive instrumentation of the environment and (2) simple systems that rely on existing infrastructure (e.g. WiFi access points for indoor systems) and/or sensors in mobile devices but provide inaccurate and unreliable location only.

The built-in system of the individual user (when tracking the user motion entirely locally) is error prone with the average error monotonously increasing as a function of walked distance, hence its applicability is further limited. This is typically found in the so-called pedestrian dead reckoning systems (PDR, also referred to as inertial navigation systems) that track user position by double integration over acceleration and direction given by an accelerometer and a magnetic field sensor [3]. The method is attractive since most modern smart phones contain such sensors leading to a potentially large user base, however, due to a double integration, even very small errors in the sensor signal quickly accumulate and tend to lead to a large error in position. Thus, the further the user moves, the larger the PDR location error becomes. We explicitly test the accuracy of common PDR systems in Section 2 and derive equations that characterize the corresponding localization error.

A new idea has recently been introduced to improve on the above. When two users come close to each other, their systems can use proximity information to correct their position estimates based on the fact that they occupy closely related positions [4]. From Bluetooth signals (which have a limited range) or from near field communication devices or special purpose proximity sensors (with a higher accuracy but an even lower range) such proximity information can be derived. In short, the knowledge that the systems are within a certain distance of each other, combined with the probability density distributions that each system has with respect to its own location allows the construction of a joint distribution that has a lower variance than the individual estimates (Fig. 1).

* Corresponding author at: German Research Center for Artificial Intelligence (DFKI GmbH), Kaiserslautern, Germany.
E-mail address: kampis.george@gmail.com (G. Kampis).

http://dx.doi.org/10.1016/j.jocs.2014.09.001
1877-7503/



(a) The idea of collaborative localization (b) Empirical results from [4]

Fig. 1. Collaborative localization based on a population of PDR systems.

In Section 3 we derive analytical models that quantify the reduced localization errors obtained by the empirical approach. The models show under which conditions the localization error remains asymptotically finite for very large values of time or continues to grow with time.

In a previous study we presented an *ad hoc* simulation based on empirically determined parameters [5]. The simulation showed that two qualitatively different regimes of location awareness are possible. The system makes a transition from a state where the error of each device is unbounded to a state where the averaged maximum error is constant, i.e., location awareness suddenly emerges even though the individual mobile devices are by themselves not capable of exact location and have a tendency to accumulate error without bounds. Here we ground these initial results in both analytical models and an agent based simulation, both reproducing the main result and providing further insights, in particular on the importance of informed agents, system size, correlations, the role of meeting densities and other parameters.

The paper is organized as follows. In Section 2 we describe a motivating experiment using an individual PDR system which leads to a quantitative description of an increasing localization error characterized by a power law. In Section 3 a corresponding analytical mean-field model for such an individual PDR is developed. The model is first extended to collaborating PDRs in an infinite system, then finite systems with and without additional fixed agents are studied. The agent based simulation model is presented in Section 4. Finally, the conclusion and outlook are presented in Section 5.

2. Description of PDR performance

We have performed an experiment to quantify the accuracy of common PDRs under usual conditions. The main goal was to establish rules than can quantify the accumulation of the localization error of a PDR system as a function of time and traveled distance. In the particular experiment, a PDR tracked the position of an Ambient Intelligence (AmI) device (smartphone) while the carrying person walked inside and outside of buildings for 80 min, without intermediate re-calibration. Simultaneously, the smartphone recorded GPS coordinates for reference at a time resolution of 0.25 s (sampling rate 4 Hz).

Fig. 2(a) shows the registered path according to GPS coordinates [4], and Fig. 2(b) shows the corresponding data from the PDR. Note that, at first glance, the two traces do not resemble each other to a recognizable degree. This is clearly due to accumulating errors in the direction of motion in the PDR system. However, closer inspection reveals that local straight-line motion is captured by the PDR quite accurately, if directions are disregarded. We thus conclude that a single PDR cannot be used for long-term position tracking (unless directionality information, e.g., from a reliable compass, is also taken into account). But how accurate is a single PDR on shorter time and distance scales that do not involve changes in direction? Can one quantify the increasing localization error?

Since absolute positions are irrelevant for estimating the increases in localization errors with time and traveled distance, we can use each space-time point $(x^{(1)}, y^{(1)}, t^{(1)})$ of the trail as a starting point. Then we can determine, for each time delay $\Delta t = t^{(2)} - t^{(1)}$ relative to these starting points, (a) the distances traveled according to the GPS, i.e. $\Delta r_{\mathrm{GPS}} = \sqrt{(x^{(2)}_{\mathrm{GPS}} - x^{(1)}_{\mathrm{GPS}})^2 + (y^{(2)}_{\mathrm{GPS}} - y^{(1)}_{\mathrm{GPS}})^2}$ (reference distances) and (b) the distances traveled according to the PDR, i.e. $\Delta r_{\mathrm{PDR}} = \sqrt{(x^{(2)}_{\mathrm{PDR}} - x^{(1)}_{\mathrm{PDR}})^2 + (y^{(2)}_{\mathrm{PDR}} - y^{(1)}_{\mathrm{PDR}})^2}$.

Fig. 3 shows the color-coded average $\langle \Delta r_{\mathrm{PDR}} \rangle$ as function of Δt and reference distance Δr_{GPS}. The averaging $\langle \cdot \rangle$ is done over all starting points (all $t^{(1)}$). One can see that a close similarity between the two distances gets lost over large time delays Δt beyond

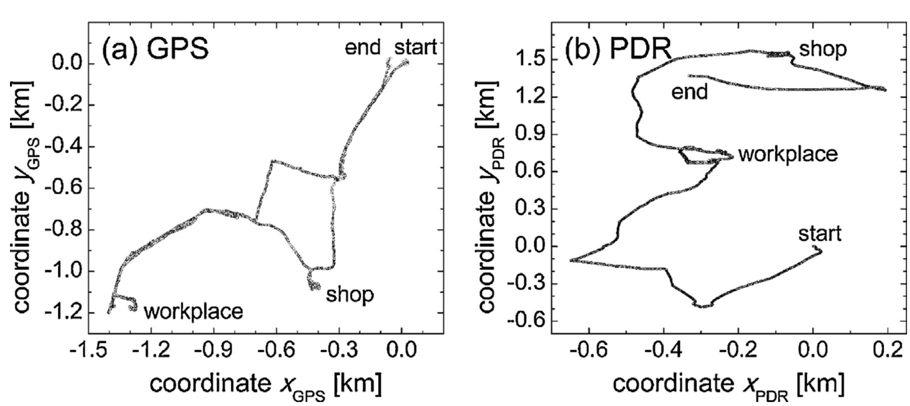


Fig. 2. Traces of motion captured by (a) the GPS sensor and (b) the PDR system of a smart phone during normal daily activities.

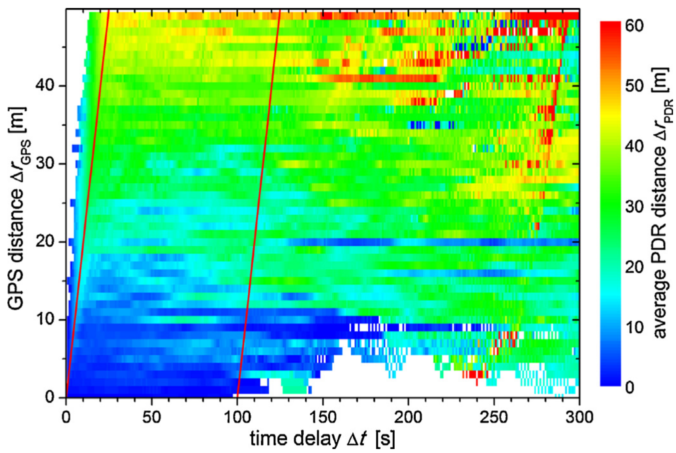


Fig. 3. Average PDR-measured distance $\langle \Delta r_{PDR} \rangle$ (color coded, see legend) as function of time delay Δt and reference distance Δr_{GPS}. Between the two straight red lines at $\Delta r_{GPS} = (2\mathrm{m}/s)\Delta t$ and $\Delta r_{GPS} = (2\mathrm{m}/s)(\Delta t - 100\mathrm{s})$ a systematic relation between the two distances is visible. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

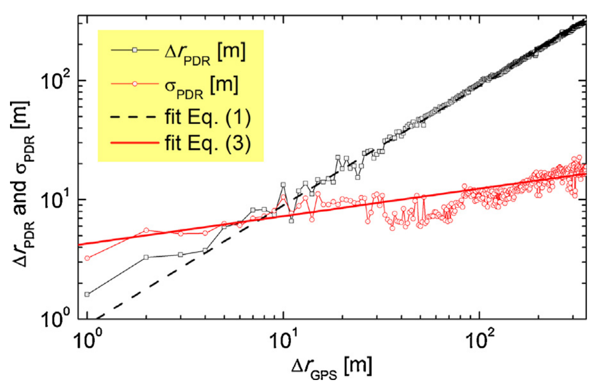


Fig. 4. Double logarithmic plot of the dependence of the average PDR-measured distance $\langle \Delta r_{PDR} \rangle$ and the corresponding PDR standard deviation σ_{PDR} (Eq. (2), red) on the reference distance Δr_{GPS}. The lines are power-law fits to the data (see text). Data points below 5 m should be disregarded, because GPS distances are not accurate in that range. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

100 to 130 s. Traveling for longer times usually involves stops or changes of direction which are not well detected by the PDR system, leading to unreliable PDR distance estimates. However, in time delay frames between an upper bound, $\Delta r_{GPS} < (2\mathrm{m}/s)\Delta t$ (related to a maximal speed of 2 m/s),[1] and a corresponding lower bound $\Delta r_{GPS} > (2\,\mathrm{m}/s)(\Delta t - 100\,\mathrm{s})$ (red lines in Fig. 3), a consistent relation between Δr_{GPS} and $\langle \Delta r_{PDR} \rangle$ is seen. In that time delay range, low Δr_{GPS} correspond to low $\langle \Delta r_{PDR} \rangle$ (blue), intermediate Δr_{GPS} correspond to intermediate $\langle \Delta r_{PDR} \rangle$ (green), and large Δr_{GPS} correspond to large $\langle \Delta r_{PDR} \rangle$ (yellow).

We can thus focus on time delays Δt between Δt_{min} and Δt_{max} for estimating the dependence of the average $\langle \Delta r_{PDR} \rangle$ on Δr_{GPS}. The result is shown in Fig. 4. As expected, $\langle \Delta r_{PDR} \rangle$ increases approximately linearly with Δr_{GPS}, indicating a quite reliable estimation of traveled distances by the PDR up to about 300 m (or about 2 min).

[1] The maximal speed of 2 m/s corresponds to fast walking. Faster motion did hardly occur during the data recording, hence larger GPS-based velocities are attributed to errors in GPS (reference) localization, and the corresponding data are ignored. We do not set a lower bound for the allowed velocities, however (see also Fig. 3, where velocity corresponds to the slope of the line).

The fit in the double logarithmic plot (dashed line) has in fact a unity slope, but a pre-factor slightly smaller than unity:

$$\langle \Delta r_{PDR} \rangle \sim 0.9 \Delta r_{GPS}. \quad (1)$$

This result indicates that the calibration of the distance measurements of the PDR could (using the given devices) be improved by increasing all distances Δr_{PDR} by approximately 10%. However, this slight distance underestimation is surprisingly independent of the reference distance Δr_{GPS}, at least in the considered range from 5 to 300 m. We note that hardly any longer straight-path motion occurred in our experiment (see Fig. 2).

Besides this systematic deviation of the average $\langle \Delta r_{PDR} \rangle$ from Δr_{GPS}, random (statistical) deviations occur in each case, which are characterized by the standard deviation (fluctuation, statistical error bar)

$$\sigma_{PDR} = \sqrt{\langle \Delta r_{PDR}^2 \rangle - \langle \Delta r_{PDR} \rangle^2}. \quad (2)$$

The quantity σ_{PDR} characterizes the reliability of the PDR, since it measures the variation of the subjective distances Δr_{PDR} obtained for each objective distance Δr_{GPS}. Our result is shown in Fig. 4 by the red data points and the corresponding power-law fit. Specifically, we find

$$\sigma_{PDR} \sim 4.3\,\mathrm{m}(\Delta r_{GPS}/\mathrm{m})^{0.23} \quad (3)$$

(red fit in Fig. 4).

The slope of the fit, i.e. the power-law exponent 0.23, is quite noteworthy. This value indicates that localization errors increase with traveled distance Δr_{GPS} and thus with time (assuming approximately constant velocity). This increase is much less rapid than expected for a random walk, i.e. if random uncorrelated errors were accumulated. If that was the case, an exponent close to 0.5 should be observed. The much smaller experimental result indicates that errors are anti-correlated in time, so that large increases in the total error are partly compensated later. The behavior of the PDR could thus be described by an anti-correlated random walk – a *sub-diffusive* process.

Specifically, random error vectors $\delta\vec{r}_i$ for each time step i are accumulated to obtain the total error vector $\delta\vec{r}(t)$ of the position estimate at time t, so that the difference between the subjective position $\vec{r}_{PDR}$ of the agent (with respect to the origin) and its objective position $\vec{r}_{GPS}$ is

$$\delta\vec{r}(t) = \vec{r}_{PDR}(t) - \vec{r}_{GPS}(t) = \sum_{i=1}^{t} \delta\vec{r}_i. \quad (4)$$

The time series of error increments $(\delta\vec{r}_i)$ must be anti-correlated in time in order to achieve a sub-diffusive behavior of the accumulated error.

In the following, we will focus on the standard deviation $s(t) = \sqrt{\langle (\delta\vec{r})^2(t) \rangle}$ of position estimates, assuming that there is no bias in the position estimates, i.e. no trend in $\delta\vec{r}(t)$, so that the expected value remains $\langle \delta\vec{r}(t) \rangle = 0$. The experimental results for σ_{PDR} shown in Fig. 4 motivate a time dependence of $s(t)$ following approximately

$$s(t) = \sqrt{\langle (\delta\vec{r})^2(t) \rangle} = \sqrt{\langle (\delta x)^2(t) \rangle + \langle (\delta y)^2(t) \rangle} \sim t^{\alpha} \quad (5)$$

with $0 < \alpha < 1$. Here, $\alpha < 1/2$ represents anti-correlations (partly self-compensating errors of the PDR) and $\alpha > 1/2$ represents positive correlations (stronger accumulating errors). Note that the experimental power-law exponent $\alpha = 0.23$ from Fig. 4 may be inaccurate, in part because data from only a single experiment is available. Therefore, we will keep α as a parameter in the following and develop a theory that works for all values of α between 0 and 1.

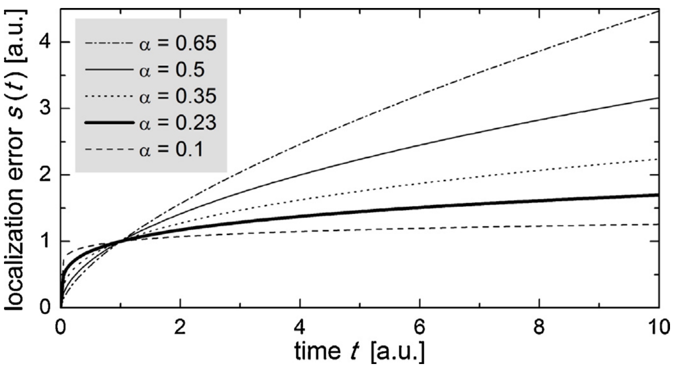


Fig. 5. Time dependence of (mean-field) localization error for individual (not collaborating) PDRs following Eqs. (5) and (6) with $\alpha=0.1$ (dashed line), 0.23 (thick line), 0.35 (dotted line), 0.5 (thin line), and 0.65 (dash-dotted line). In all cases, the localization error does not remain finite asymptotically for $t\to\infty$.

In the agent based model presented later, α will be fixed at 0.5 (regular random walk).

3. Analytical models for individual and collaborative PDR systems

Now we can turn to the characterization of localization systems using PDRs. We define a system of N pedestrians moving randomly in a large restricted area (e.g. a large room) of size $L\times L$ with an average speed of v. Each pedestrian is carrying an AmI device with a PDR system tracking the position of the pedestrian. Initially (at $t=t_0$) each device "knows" the exact position. Due to measurement inaccuracies, the error in the estimated position of each pedestrian increases with time. Since we aim at a mean-field theory for the dependence of these errors on time and other parameters, we disregard the specific characteristics of the measurements of the devices and we also disregard variations among the pedestrians.

Our main quantity of interest is thus the average localization error $s(t)$, which depends on the time t (or, equivalently, on distance traveled, if a constant speed is assumed). We assume $s(t)$ to be identical for all devices. As motivated by the experiment in the previous section, a natural time dependence of $s(t)$ for individual (i.e. not collaborating) PDR systems is given by Eq. (5). The differential equation characterizing the dynamics of $s(t)$ for individual PDRs is thus obtained as

$$\frac{ds(t)}{dt}=\alpha[s(t)]^{(1-1/\alpha)}. \tag{6}$$

Example solutions for different values of α are shown in Fig. 5.

Now we can turn to models for collaborative PDRs. We consider the collaboration of devices of pedestrians when they come close to each other. As before, we assume that all agents start moving (and accumulating errors) simultaneously.[2]

Using a short-range communication, position measurements of both devices are shared between them as soon as the persons come sufficiently close to each other. We assume that the real distance d between the devices at the meeting point is negligible compared with the accumulated position measurement errors $s_1(t)$ and $s_2(t)$ of the PDR devices.

For the first analytical model, let us also assume that the two devices (or agents) never met before. This is an assumption of infinitely large systems (infinite linear dimension L) with infinitely many agents N but finite density $\rho=N/L^2$. In this case, the accumulated measurement errors in both devices, $\delta\vec{r}_1(t)$ and $\delta\vec{r}_2(t)$ as well as their components $\delta x_1, \delta y_1, \delta x_2$, and δy_2, will be independent. When the position estimates of the two agents are combined,

$$x_{col}=\frac{x_1+x_2}{2},\quad y_{col}=\frac{y_1+y_2}{2},\quad \delta x_{col}=\frac{\delta x_1+\delta x_2}{2},$$
$$\delta y_{col}=\frac{\delta y_1+\delta y_2}{2}, \tag{7}$$

the errors of these collaborative position estimates are smaller than the original accumulated measurement errors. Since δx_1 and δx_2 are independent and identically distributed (iid) random variables with zero mean, the variance of their average δx_{col}, i.e. $\langle(\delta x_{col})^2\rangle$, is lower by a factor of 1/2 compared with the variance of δx_1 and δx_2, $\langle(\delta x_1)^2\rangle=\langle(\delta x_2)^2\rangle=\langle(\delta x)^2\rangle$. The same holds for δy and for the two-dimensional variance, so that the standard deviation

$$s_{col}(t)=\sqrt{\langle(\delta x_{col})^2\rangle+\langle(\delta y_{col})^2\rangle}=\sqrt{\frac{1}{2}\langle(\delta x)^2\rangle+\frac{1}{2}\langle(\delta y)^2\rangle}$$
$$=\frac{1}{\sqrt{2}}s(t) \tag{8}$$

decreases by a factor of $1/\sqrt{2}$ upon collaboration.[3]

Hence, the change in the standard deviation $s(t)$ upon each collaboration event is $s_{col}(t)-s(t)=(1/\sqrt{2}-1)s(t)$. To include the collaboration events in the theory, we introduce the parameter $t_{new\ col}$, which represents the average time interval between events of two PDRs coming close to each other. This parameter can be approximated by $t_{new\,col}\approx\ell/v$, where $\ell=1/(\rho d)$ is the mean free path, v is the velocity, d is the devices' radio range for proximity detection (or, equivalently, the diameter of the devices), and $\rho=N/L^2$ is the surface density of people (or devices). The approach thus yields

$$t_{new\ \ col}=1/(\rho vd)=L^2/(Nvd). \tag{9}$$

We note that $m=1/t_{new\ col}=vd\rho=vdN/L^2$ is the frequency of proximity events for a given device, i.e., the meeting frequency. Upon each collaboration event the localization error $s(t)$ is reduced by $(1-1/\sqrt{2})s(t)$; this occurs at a rate of $1/t_{new\ col}$. The following differential equation thus describes the time evolution of the error of collaborative localization in an infinite system:

$$\frac{ds(t)}{dt}=\alpha[s(t)]^{(1-1/\alpha)}-\left(1-\frac{1}{\sqrt{2}}\right)\frac{s(t)}{t_{new\ col}}. \tag{10}$$

This equation can still be solved analytically. Fig. 6 shows plots of solutions with different parameters α and $t_{new\ col}$. If the collaboration between the PDRs is switched on later (at $t=3$ in the example), curves like those shown in Fig. 7 are obtained.

The asymptotic value s_{max} of $s(t)$ for large values of time t (saturation level) can be defined and analytically calculated by setting the time derivative to zero in Eq. (10):

$$\alpha s_{max}^{(1-1/\alpha)}=\left(1-\frac{1}{\sqrt{2}}\right)\frac{s_{max}}{t_{new\ col}}, \tag{11}$$

so that

$$s_{max}=\left(\frac{\alpha t_{new\ col}\sqrt{2}}{\sqrt{2}-1}\right)^{\alpha}. \tag{12}$$

The asymptotic values s_{max} are also shown in Fig. 7. Note again, however, that this result holds for infinite systems only.

Finally, then, we have to deal with the finite systems, where the same two PDRs will meet several times. For convenience, we employ a second time parameter $t_{rep\ col}$, which represents

[2] If this condition is relaxed and some agents enter later with a zero error, the situation is similar to that of fixed agents discussed below.

[3] Note that $\langle\delta x_i\rangle$=0, since we assume the absence of a bias, and all agents have accumulated errors of the same magnitude, i.e., $\langle(\delta x_i)^2\rangle=\langle(\delta x)^2\rangle$. The standard deviation, which is the square root of the variance, decreases by the factor $1/\sqrt{2}$.

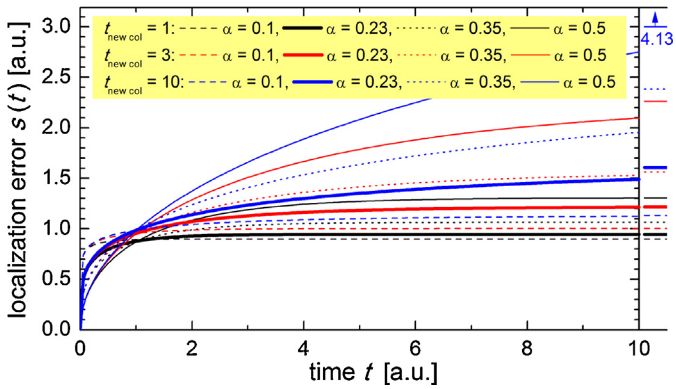


Fig. 6. Time dependence of (mean-field) localization error for collaborating PDRs following Eq. (10) with $\alpha = 0.1$ (dashed lines), 0.23 (thick lines), 0.35 (dotted lines), and 0.5 (thin lines) as well as $t_{new\,col} = 1$ (black), 3 (red), and 10 (blue). In all cases, the localization errors are smaller than those without collaboration (see Fig. 5) and remain finite asymptotically. The values for $t \to \infty$ are shown by short line segments to the right-hand side of the right axis. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

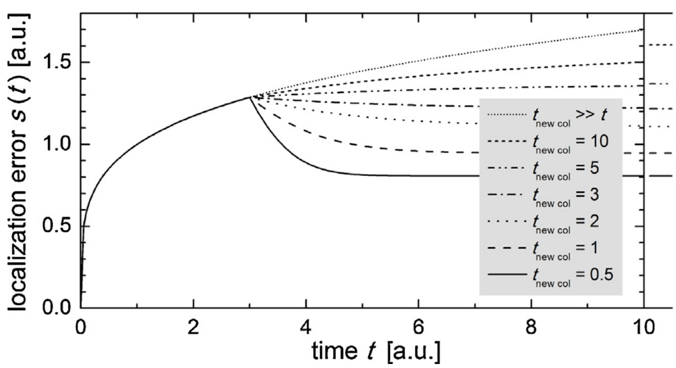


Fig. 7. Time dependence of (mean-field) localization error for collaborating PDRs following Eq. (10) with $\alpha = 0.23$ and $t_{new\,col} = 0.5$, 1, 2, 3, 5, 10, ∞ (from bottom to top; $t_{new\,col} = \infty$ is the case without collaboration). The parameter $t_{new\,col}$ from Eq. (9) includes the density of agents ρ, the speed of the agents v, and the radio range of the devices d. The communication between the PDRs was suspended before $t = 3$.

the average time interval between repeated meetings (and hence collaboration) of the same two persons (and PDRs). Usually, $t_{rep\,col} \gg t_{new\,col}$, and thus

$$t_{rep\,col} \approx N t_{new\,col} \quad \text{or} \quad t_{rep\,col} \approx (N/2) t_{new\,col} \tag{13}$$

seem to be reasonable approximations.

If the same two PDRs meet again, their errors are not completely independent. Instead, only additional errors accumulated after the last meeting will be independent. Therefore, we must replace $s(t)$ in the second term of Eq. (10) by a term approximating this additional accumulated error, i.e. $s(t) - s(t - t_{rep\,col})$. This is a rather conservative approximation, becoming more exact for smaller values of $t_{rep\,col}$. It yields the following time-delayed differential equation for the collaborative localization in a finite system:

$$\frac{ds(t)}{dt} = \alpha[s(t)]^{(1-(1/\alpha))} - \frac{\sqrt{2}-1}{t_{new\,col}\sqrt{2}}[s(t) - s(t - t_{rep\,col})], \tag{14}$$

where the initial condition $s(t) = 0$ holds for all $t \le 0$.

Note that this equation can no longer be solved analytically. A constant level s_{max} is no longer reached in this case, and the phase transition thus disappears. The growth of $s(t)$ as function of time is delayed (compared with the individual PDR system) but not asymptotically limited in the finite system. Fig. 8 compares a numerical solution of Eq. (14) with the corresponding exact solutions for the previous model versions.

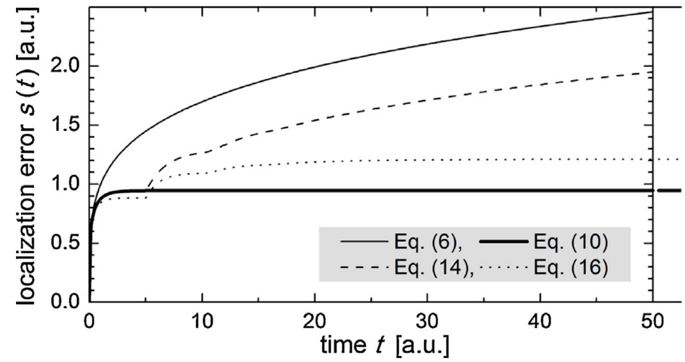


Fig. 8. Comparison of the time dependences of (mean-field) localization errors for an individual PDR (thin line, Eq. (6) with $\alpha = 0.23$), and collaborating PDRs in an infinite system (thick line, Eq. (10) with $t_{new\,col} = 1$), in a finite system (dashed line, Eq. (14) with $N = 10$, $t_{new\,col} = 1$, $t_{rep\,col} = 5$), and a finite system with one fixed agent (dotted line, Eq. (16) with $M = 1$ and $t_{fixed} = 10$). A finite asymptotic value is reached in the second and in the forth case. The wiggling of the dashed and dotted lines is caused by the initial conditions of the corresponding time-delayed differential equations, since a fixed time delay $t_{rep\,col} = 5$ is assumed. Under a more natural assumption of randomly varying time delays for each agent, the curves would be smoothly increasing also around $t = 5$ and 10. The differential equations have been solved analytically and/or numerically using Mathematica [6].

The phase transition toward an asymptotically constant regime can however be restored if we assume that there are a few immobile devices in the room. Since immobile devices do not change their position, they do not need any PDRs and always keep an exact knowledge of their position. If the number of such fixed devices is M, the time

$$t_{fixed} \approx (M/N) t_{new\,col} \tag{15}$$

can describe the mean time interval between two events that a specific person (or PDR) comes close to such a device. Each time that happens, the localization error of that person can be reset to 0. We can thus assume an additional mean error decrease rate of $s(t)/t_{fixed}$, yielding

$$\frac{ds(t)}{dt} = \alpha[s(t)]^{(1-(1/\alpha))} - \frac{\sqrt{2}-1}{t_{new\,col}\sqrt{2}}[s(t) - s(t - t_{rep\,col})] - \frac{s(t)}{t_{fixed}}. \tag{16}$$

This differential equation surely yields a constant asymptotic value for $s(t)$ again, as it can be solved analytically in full analogy with Eq. (10) if the second term on the right hand side is disregarded. However, disregarding a negative term can only increase the localization error. Therefore, we see that even a single immobile device (one single fixed point) is sufficient for restoring the phase transition toward the regime with a constant asymptotic localization error in the limit. Fig. 8 also compares a numerical solution of Eq. (16) with the corresponding solutions for the previous model versions.

4. An agent based model for collaborative PDR systems

We also wanted to approach the problem from a disaggregated, low level perspective, to understand:

1 The behavior of finite systems at finite times under various further assumptions (in particular, to map the area between the dotted and dashed lines of Fig. 8, which is inhabited by most systems).
2 The effect of realistic meeting conditions, i.e. of inexact meetings represented by meeting diameter as a parameter.

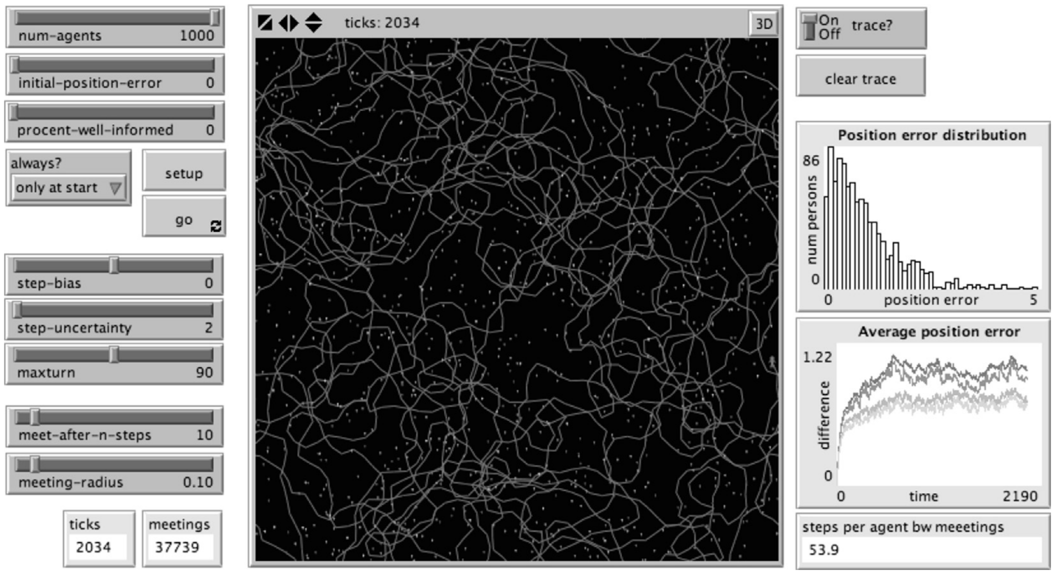


Fig. 9. The model GUI.

3 The effect of correlations and their change as a function of system size.

We developed an agent based simulation model for the collaborative PDR localization process. Agent based models tend to offer a rich, entity-level, causal perspective where parameters can be heterogeneously, individually defined and manipulated. This makes them an ideal counterpart to higher level aggregate models such as differential equations to test the underlying theories in a minimalist, reality-close setting. Our model was written in the freely available agent based modeling (ABM) and simulation environment NetLogo ([7], provided by Northwestern University) that provides high level tools for fast model development and intuitive experimentation. The model can be run interactively from the GUI or in a "headless" mode from command line for parameter sweeps and analysis. The program code has been made publicly available in the open source NetLogo Models Library (http://ccl.northwestern.edu/netlogo/models/, for the model GUI see Fig. 9).

The model uses a population of agents (of size N) that navigate in continuous two-dimensional space[4] of size $L \times L$. Time goes in discrete steps. Each agent possesses a global position vector ($x_g(t)$, $y_g(t)$) (known to the observer) and a local "subjective position" vector ($x_l(t)$, $y_l(t)$) (known to the agent). Agents perform regular random walk as allowed by the topology (which can be a torus, square, or a walled structure of streets or offices – the interactive model includes hand-tools for the drawing of these, or to import a maze). Note that this is a simplification that however (according to Sections 2 and 3) overestimates localization error.

When an agent is moving, its global position will be updated accordingly. Subjective (local) position is also updated, using two kinds of errors in every step: a systematic bias b (positive or negative) of the step size estimate (as a model of PDR based step calculation) and a random step uncertainty p (both defined as a variable and testable parameter). When performing random motion, agents can turn by a degree specified as a parameter. Although different values of this parameter yield visually different behaviors, the model is essentially non-sensitive to this parameter so we omit its further discussion.

Agents can meet. Meeting is understood here as the event that agents are "close to each other", i.e. detected within a range of each other determined by a parameter specifying meeting diameter (or radio range). Note that if the meeting diameter d (from Section 3) is set to $d = 0$, then agents can never meet as their positions are continuous and represented as real numbers that have a negligible chance to match up exactly. Upon meeting, agents exchange information on their subjective position, and update their estimate by forming the average of the two subjective positions. If two agents i and j meet, the new subjective coordinates will be computed as $x_{li}(t+1) = (x_{li}(t) + x_{lj}(t))/2$, and similarly $y_{li}(t+1) = (y_{li}(t) + y_{lj}(t))/2$.

We have built and tested two variants of the model, where the subjective position is either updated only for the active agent or for both agents. The active agent is the one having a current CPU slice, which is usually a well-defined entity in any ABM model. The process was essentially non-sensitive to this difference. Agents may also begin with a nonzero initial position error $I = s(t_0)$ (using the notation from Section 3). A certain number M of agents is allowed to "know" its exact position in every step (fixed agents). Agents can be initiated at the same initial point or randomly dispersed in space.

In interactive experiments using the model GUI, a few simple facts can be readily observed. An all-important system parameter is meeting frequency m, which is an emergent result of the number of agents N, the meeting diameter d and other factors including the exact topology and similar to $1/t_{\text{new col}} = vdN/L^2$. Using intuition, meeting frequency can be expected to give the most sensitive contribution in the model. Further, we can readily note that a nonzero initial position error can never disappear, unless a proportion of agents is well-informed. This is also expected, because in lack of external information the subjective positions can never improve beyond the best available initial estimate. Increasing the meeting diameter d increases meeting frequency (and hence the chance to collaboratively improve the position estimates) but lowers the precision of every meeting at the same time, hence compromising the improvement of the new estimate, and thus leading to a trade-off. Understandably, higher population numbers tend, in general, to also lead to higher meeting frequency and thus (all other things equal) to a better localization. Mapping these trade-offs was one of the important tasks of parameter sweeps. Further parameters such as turning, topology, the initial positions of the agents, and other

[4] NetLogo functions as a cellular automaton composed of discrete patches yet permitting continuous positions [8,9].

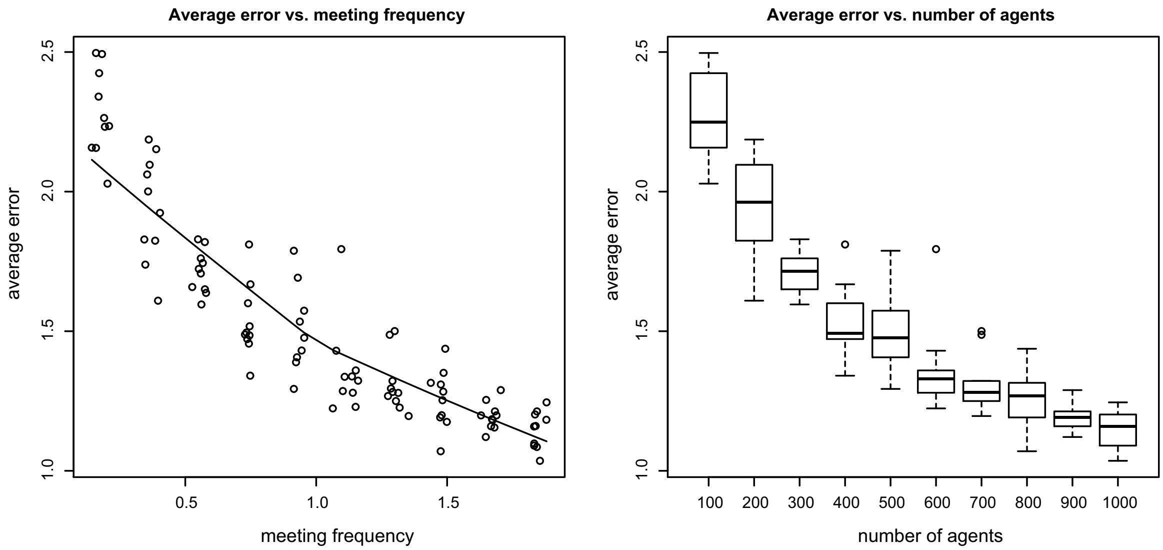


Fig. 10. Parameter sweep results: average error $s(t)$ at $t = 5000$ *versus* (a) meeting frequency m and (b) number of agents N.

Table 1
Parameter intervals and values tested.

Variable	Initial value	Step size	End value
Meeting diameter d	0.08	0.02	1.0
Step uncertainty p	1	1	10
Num. of agents N	100	100	1000

side factors have a negligible or zero effect and will thus be omitted from the detailed analysis below (Table 1).

We performed parameter sweeps using NetLogo's built-in *BehaviorSpace* module in the headless mode using 24 AMD Opteron 6348 cores at 2.8 GHz and 64 GB of RAM. Results were analyzed (obtaining statistics and visualization) using the R statistical programming language/environment. Each test was run for 1000 consecutive time steps (our statistical analysis used all generated data; the time plots show the first 5000 steps), and for each tested parameter combination 10 different random seeds were applied. We tested the system with and without collaborative information exchange, the former represented by the finite nonzero meeting diameter $d = 0.08$ and similar. In the following we use $M = 0$ and $I = 0$ to test the most interesting case where the asymptotic behavior is not well understood in the analytical model.

For the meetings regime (i.e. when the collaborative localization was turned on at various m, p and N values) Fig. 10 shows the quantitative relations between the average localization error and meeting frequency (Fig. 10(a)) viz. the number of agents (Fig. 10(b)). As also expected, there is a gradual yet significant decrease in the localization error $s(t)$ (in both cases taken after 5000 steps) as a function of both meeting frequency and the number of agents.

The main result is best presented in time plots. Fig. 11 shows 1000 different runs using values taken from the parameter intervals indicated above and $N = 400$. Colors (from ochre to pink to blue) indicate distance from a virtually "non-meeting" situation ($m = 0.04$, ochre) to that of frequent meetings due to a large meeting diameter ($d = 1.0$, blue). Note that many intermediate curves are masked out by the overplotted curves that represent the higher values (pink and blue; with blue being the highest). Nevertheless this does not interfere with the result. Although some outliers exist, it is well observable that the upper end of the meeting diameter scale, represented as pink (and blue) values, yield "flat" lines in the time plots, indicating that average localization error stays nearly constant and does not grow after reaching a (relatively low)

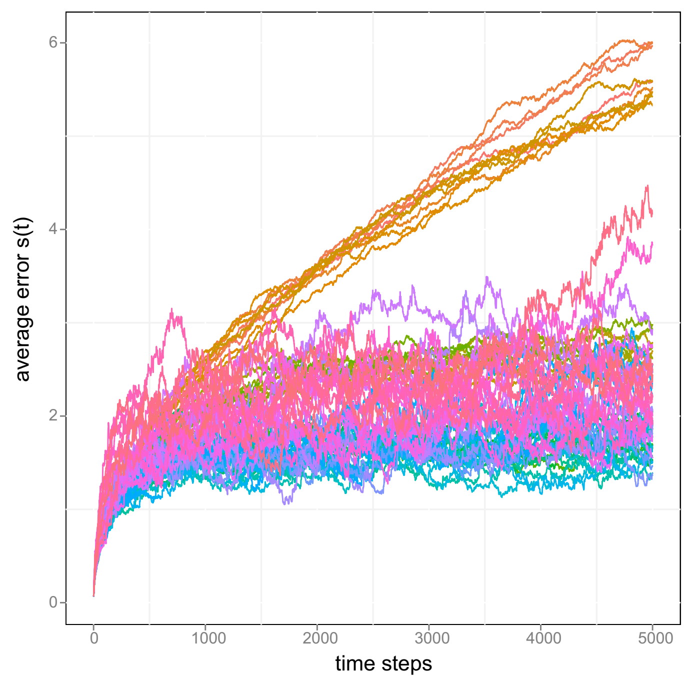


Fig. 11. 1000 time plots at various values of the meeting diameter d at $N = 400$. For explanation of colors, see text.

maximum in the studied time interval. By contrast, in those situations (ochre) where meetings hardly, if at all, occur we observe a continual divergence of the average error term $s(t)$.

For a more advanced visualization, we also produced an interactive 3D plot of the results, an OpenGL object that can be mouse rotated as well as zoomed. Below we include a snapshot (Fig. 12(a)) showing average localization error $s(t)$ versus meeting diameter d and step uncertainty p. The transition is visualized on another 3D plot showing average error $s(t)$ *versus* time for the meeting/no-meeting regimes and parameter m (Fig. 12(b)).

In short, for different assumptions our agent based model is still found to justify the claim that there is an important transition between the meeting and non-meeting regimes. Without meetings the associated updating of the subjective position estimates is lost, and the error $s(t)$ for the position estimate diverges in the

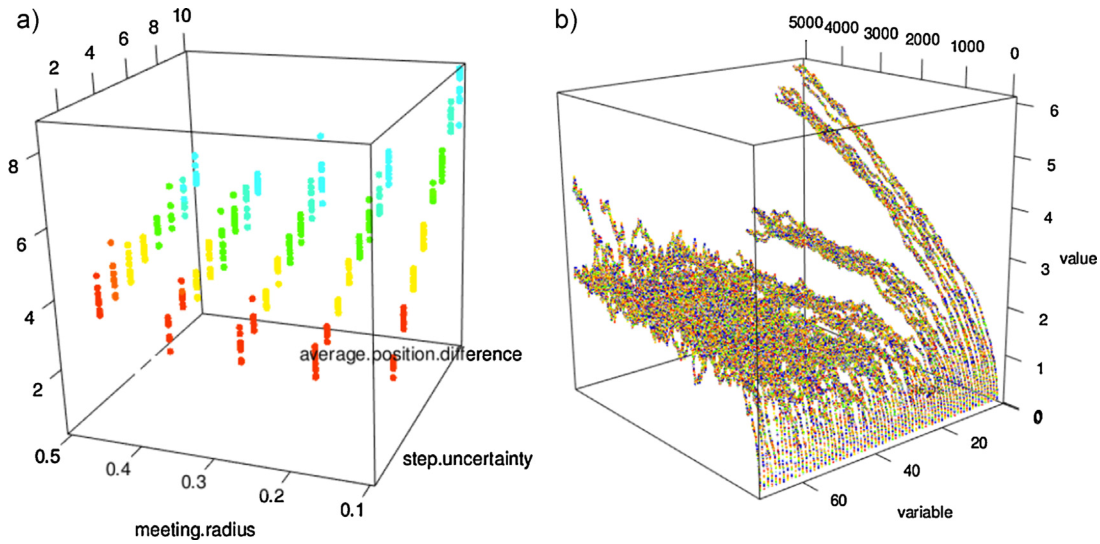


Fig. 12. OpenGL snapshots. Average localization error $s(t)$ in different comparisons: (a) *versus* a meeting radius $R = d/2$ and step uncertainty p at $t = 5000$ and (b) *versus* d and t at $N = 400$.

ABM model, whereas by introducing the meetings (together with their information exchange of the subjective positions and the consecutive computing of the new estimates) we can keep this error essentially constant despite the finite size of the system for large N. The exact error value $s(t)$ depends on many parameters including meeting frequency and the number of agents. A choice of N in the hundreds is realistically high for a population of pedestrians using an application that provides a collaborative localization service.

Finally, we turn to the study of correlations. How is it possible that despite the finiteness of the systems we obtain good results? Note that – after clarifying the effect of N on $s(t)$ in Fig. 10(b) – we have mostly studied large, if necessarily finite, systems.

Do correlations decrease with systems size? Indeed they do, as shown in Fig. 13(a). The figure shows entire-run correlations (Pearson correlation, results represented as points) between the localization error terms $s(t)$ in the ABM model, averaged for 10 runs, between $t = 0$ and $t = 1000$. For small values of N, fluctuations can be seen despite the averaging as random effects are amplified due to the few meetings among few agents. But the overall tendency is clearly visible, especially if Lowess interpolation (solid line) is added. The correlations decrease with increasing N and reach very low levels; the value at $N = 1000$ is 0.004. Therefore, with large N the chance that two error terms are correlated becomes negligible.

This phenomenon further raises the question of meeting frequency and "re-meeting" as the source of correlations. How often is re-meeting expected? The frequency r of re-meeting between two selected agents can be empirically approximated as the fraction of realized meetings in a given interval against the number of possible meetings N^2. We note that r corresponds to $N/t_{\text{rep col}}$ with $t_{\text{rep col}}$ from Eq. (13), since $t_{\text{rep col}}$ is the average time interval between the re-meetings for a selected agent. While we used a constant value (independent of N) for the re-meeting frequency r in the analytical model, Fig. 13(b) shows that r is linearly decreasing with N (from 0.64 to 0.55) in the agent-based model simulations. In other words, we see that uncorrelated random walks do not produce "enough" meetings for them to be correlated as N grows, since the number of possible meetings goes up radically (with N^2) and thus the fraction of meetings decreases. As a result, even for systems of realistic size

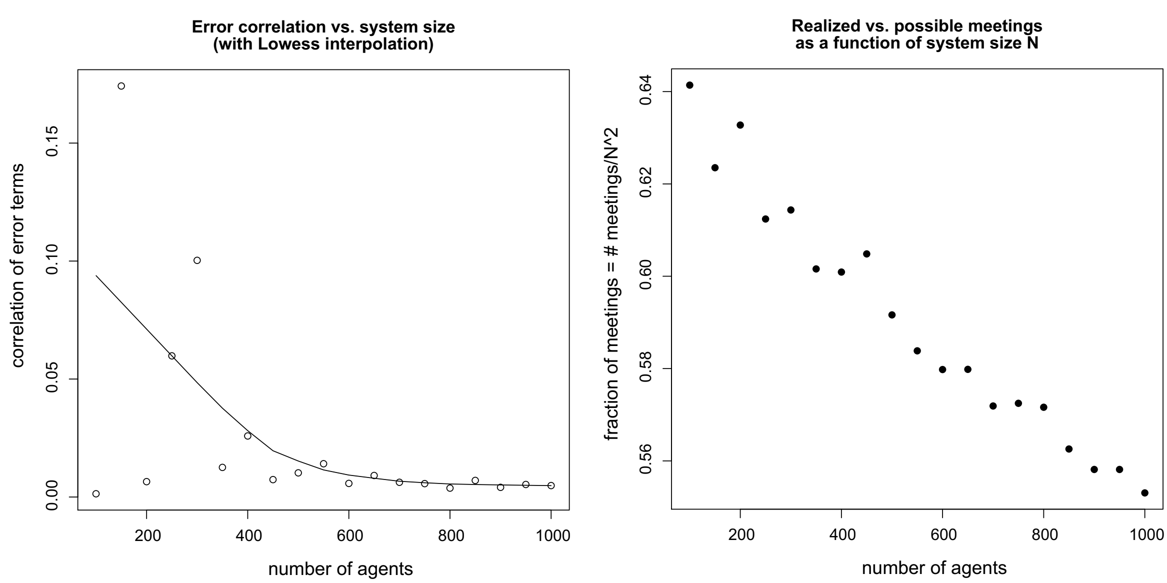


Fig. 13. Error correlations *versus* (a) number of agents N and (b) fraction r of realized meetings over possible meetings N^2. Each point summarizes 10 different runs.

(with hundreds of agents) the correlation problem can safely be neglected and thus the behavior of the system approximates that of the infinite case.

5. Conclusion and outlook

In this paper we have derived different models for collaborative localization. In particular, using a mean-field model we have characterized the increase of the localization error of an individual PDR by an experimentally obtained scaling exponent. When the AmI systems of several agents can interact, they can increase the accuracies of their PDR systems by collaboration. We have shown that the localization error remains asymptotically finite if the system is infinitely large (i.e. agents always continue to meet other agents they never met before) or if there is at least one stationary agent with constantly zero localization error. While the analytical solution of the mean-field model was possible in the case of infinite systems, the time-delayed differential equation for the finite system could only be solved numerically. It was found that, in these time-delayed finite systems, as was the case with individual PDRs, the localization error does not approach a constant value if no fixed (well-informed) agents are applied.

To test how serious this restriction is in practice, we have performed additional numerical experiments in an agent-based model. The model studies the effect of system size (i.e. the number of agents), the effect of realistic (imprecise) meetings and the role of correlations. We found that in finite yet sufficiently large systems ($N \approx$ several *hundred*) under realistic conditions, the localization error remains sufficiently small over time. We can thus conclude that the approach is viable and can produce good localization results under a wide range of conditions. Further experiments, not discussed in this paper because they add no relevant details, indicate that (despite a different spatial scaling [10]) the same applies if agents navigate in mazes or on maps, as opposed to free navigation in barrier-free two-dimensional space as in the results presented here.

The collaboration approach might be very useful not only for pedestrian localization in buildings and in crowds, but also for increasing the accuracy of car localization on roads with intense traffic, where accurate localization information is essential for individual traffic guidance systems that should know in which lane a driver is found (where accuracy required beyond that of GPS information).

Further, we maintain that collaborative localization also presents a baseline case for more general collaborative knowledge fusion (where both present a ground case for complex adaptive systems). What essentially happens in collaborative localization from a more abstract viewpoint is that the agents exchange their *knowledge* in order to achieve an advanced knowledge status together, collectively, one that would not be achievable individually, and achieve this status without any prior planning or coordination of the agent-to-agent (i.e. peer-to-peer) information flow. Knowledge in this simplest case is signified by that of the belief in a (subjective) position. We believe the same principle can be generalized to include more general knowledge systems, the exploration of which is left to future work.

Acknowledgments

We thank Mirko Kämpf and Krzysztof Kulakowski for discussions and Mark Jelasity for valuable suggestions. We acknowledge partial funding from the European Community's Seventh Framework Program (FP7/2007–2013) under Grant Agreement #600854 Smart Society – hybrid and diversity-aware collective adaptive systems: where people meet machines to build smarter societies (http://www.smart-society-project.eu/), by the CoCoRec (Collaborative Context Recognition in Dynamic, Multimodal Smart Environments) project supported by the German Federal Ministry of Education and Research, and by the German Research Society (DFG, Grant KA 1676/4). One of the authors (G.K.) gratefully acknowledges the partial support by the Russian Scientific Foundation, proposal #14-21-00137, "Supercomputer simulation of critical phenomena of complex social systems".

Appendix

See Table 2.

Table 2
Table of symbols.

Symbol	Meaning	Symbol	Meaning
N	Number of agents	L	Linear system size
x, y	Spatial coordinates	$\vec{r}$	Position vector
$t, t^{(i)}$	Time, time instance	Δr	Distance
$s(t)$	Average localization error at t	σ	Standard error of Δr
l	Mean free path bw. meetings	m	Meeting frequency
ρ	Density of agents ($= N/L^2$)	M	No. of fixed ("well-informed") agents
p	Step uncertainty	d	"Radio range", i.e. diameter of sensing proximity

References

[1] M. Youssef, A. Agrawala, A. Udaya Shankar, Wlan location determination via clustering and probability distributions, in: Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, 2003 (PerCom 2003), 2003, pp. 143–150, http://dx.doi.org/10.1109/PERCOM.2003.1192736.

[2] J. Hightower, G. Borriello, Location systems for ubiquitous computing, Computer 34 (8) (2001) 57–66, http://dx.doi.org/10.1109/2.940014.

[3] C. Fischer, H. Gellersen, Location and navigation support for emergency responders: a survey, Pervas. Comput. IEEE 9 (1) (2010) 38–47, http://dx.doi.org/10.1109/MPRV.2009.91.

[4] K. Kloch, P. Lukowicz, C. Fischer, Collaborative PDR localisation with mobile phones, in: Proceedings of the 2011 15th Annual International Symposium on Wearable Computers, ISWC'11, IEEE Computer Society, Washington, DC, USA, 2011, pp. 37–40.

[5] K. Kloch, G. Pirkl, P. Lukowicz, C. Fischer, Emergent behaviour in collaborative indoor localisation: an example of self-organisation in ubiquitous sensing systems, in: Proceedings of ARCS 2011 – Architecture of Computing Systems, Lecture Notes in Computer Science, Vol. 6566, Springer-Verlag, 2011, pp. 207–218, http://dx.doi.org/10.1007/978-3-642-19137-4_18.

[6] Wolfram Research, Inc., Mathematica, Version 8.0, Champaign, IL, 2010.

[7] U. Wilensky, Netlogo, Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL, http://ccl.northwestern.edu/netlogo/

[8] A. Hoekstra, J. Kroc, P. Sloot, Simulating Complex Systems by Cellular Automata, Springer, 2010.

[9] S. Karabasov, D. Nerukh, A. Hoekstra, B. Chopard, P.V. Coveney, Multiscale modelling: approaches and challenges, Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci. 372 (2021) (2014) 20130390.

[10] J.A. Hołyst, J. Sienkiewicz, A. Fronczak, P. Fronczak, K. Suchecki, Scaling of distances in correlated complex networks, Phys. A: Stat. Mech. Appl. 351 (1) (2005) 167–174.

Dr. George Kampis (Mr.) has been acting as founding chairman and Professor (since 1994) of the department of History and Philosophy of Science at Eotvos University in Budapest (http://hps.elte.hu). Holds a Ph.D. and a Habilitation in Biology and a D.Sc. in Philosophy of Science. Main research interests in Artificial Life, cognitive science (Director of the Budapest Semester in Cognitive Science, http://hps.elte.hu/BSCS), complex systems (www.petabyte-research.org) and evolutionary modeling (www.evotech.hu). He was guest professor at Hokkaido University, Fujitsu Chair of Complex Systems at JAIST (Japan Advanced Institute for Science and Technology, in 2002/3), Wayne G. Basler Chair of Excellence at East Tennessee State University in 2007, and Fulbright fellow at Indiana University in 2009.

Dr Kampis has about 150 scientific publications, and he is the author or editor of several books with international publishers Elsevier, Kluwer, and Springer. He has joined DFKI in 2012 as a Senior Researcher.

Dr. Jan W. Kantelhardt (Mr.) was born in 1970 in Giessen, Germany. He studied physics and mathematics in Giessen and Berlin, graduated in 1996, and received his doctorate in theoretical physics in 1999. He spent postdoc years in Israel (Physics Department, Bar-Ilan University) and the USA (Center for Polymer Studies, Boston University, MA). Since 2003 he is working at the Institute of Physics at the Martin-Luther University of Halle-Wittenberg, Germany. His research interests include the physics of disordered materials and applications of statistical physics in computer science, physiology and medicine.

Dr. Kamil Kloch (Mr.) has been a Senior Researcher at the German Research Center for Artificial Intelligence (DFKI GmbH). His current work is focused on the collaborative indoor localization on mobile devices. He received Ph.D. in Computer Science in 2008 from the Jagiellonian University in Krakw, Poland. In his Ph.D. he investigated algorithms for the online scheduling problem.

Prof. Dr. Paul Lukowicz (Mr.) is Full Professor of AI at the Technical University of Kaiserslautern in Germany where he will head the Embedded Intelligence group at DFKI. From 2006 till 2011 he has been full Professor (W3) of Computer Science at the University of Passau. He has also been a senior researcher ("Oberassistent") at the Electronics Laboratory at the Department of Information Technology and Electrical Engineering of ETH Zurich Paul Lukowicz has MSc. (Dipl. Inf.) and a Ph.D. (Dr. rer nat.) in Computer Science a MSc. in Physics (Dipl. Phys.). His research focus is context aware ubiquitous and wearable systems including sensing, pattern recognition, system architectures, models of large scale self organized systems, and applications. Paul Lukowicz coordinated the recent SOCIONICAL projects, is Associate Editor in Chief IEEE Pervasive Computing Magazine, and has served on the TPCs (including TPC Chair of all main conferences in the area.