Exploring approaches to improve the performance of autonomous monitoring with imperfect data in location-aware wireless sensor networks

Article *in* Advanced Engineering Informatics · October 2014

DOI: 10.1016/j.aei.2014.08.004

4 authors, including:

Xiaowei Luo
City University of Hong Kong

William John O'Brien
University of Texas at Austin

Advanced Engineering Informatics 28 (2014) 287–296

Contents lists available at ScienceDirect

Advanced Engineering Informatics

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

Exploring approaches to improve the performance of autonomous monitoring with imperfect data in location-aware wireless sensor networks

Xiaowei Luo [a], William J. O'Brien [b,*], Fernanda Leite [b], James A. Goulet [c]

[a] Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong
[b] Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, 1 University Station C1752, Austin, TX 78712, United States
[c] Department of Civil and Environmental Engineering, University of California at Berkeley, Berkeley, CA 94720, United States

ARTICLE INFO

Article history:
Received 18 August 2014
Accepted 27 August 2014
Available online 11 October 2014

Keywords:
Wireless sensors
Safety monitoring
Data error
Missing data
Tower crane
Bayesian network

ABSTRACT

In recent years, information and sensing technologies have been applied to the construction industry to collect and provide rich information to facilitate decision making processes. One of the applications is using location data to support autonomous crane safety monitoring (e.g., collision avoidance and dangerous areas control). Several location-aware wireless technologies such as GPS (Global Positioning System), RFID (Radio-frequency identification), and Ultra-Wide Band sensors, have been proposed to provide location information for autonomous safety monitoring. However, previous studies indicated that imperfections (errors, uncertainty, and inconsistency) exist in the data collected from those sensors and the data imperfections have great impacts on autonomous safety monitoring system performance. This paper explores five computationally light-weight approaches to deal with the data imperfections, aiming to improve the system performance. The authors built a scaled autonomous crane safety monitoring testbed with a mounted localization system to collect location data and developed five representative test cases based on a live construction jobsite. Seven hundred and sixty location readings were collected at thirty-eight test points from the sensors. Those location data was fed into the reasoning mechanisms with five approaches to generate the safety decisions at those thirty-eight test points and evaluate system performance in terms of precision, recall and accuracy. The results indicate that system performance can be improved if at least ten position readings from sensors can be collected at small intervals at any location along the moving path. However, by including additional data such as velocity and acceleration that may be read from devices mounted on workers, localization error may be significantly reduced. These findings represent a path forward to improve localization accuracy by mixing imperfect data from the sensed environment with supplemental input.

1. Introduction

This paper reports the results of using five computationally lightweight approaches to improve system performance with imperfect data from location-aware sensors. The analysis indicates that promising results can be obtained using simple approaches, suggesting that viable systems may be developed to remedy common imperfections in data monitoring and support robust applications. The approaches detailed in this paper build from the literature detailing the inaccuracy of sensed data; however, they do not require prior calibration of the sensor network to support flexible deployment.

The motivating context for this research is autonomous job-site safety monitoring – crane safety, in particular. Cranes can perform a great variety of tasks on construction jobsites to increase efficiency and productivity of the construction sites. However, a crane can be very dangerous piece of equipment. CPWR [25] has examined crane-related deaths in construction from 1992 to 2006 and identified there were total 632 crane-related deaths, an average of 42 deaths per year. The study also showed that cranes not only pose risks to construction workers but also to the public since there were approximately 7% of the total crane-related deaths were innocent bystanders. CPWR's findings do not only show the magnitude of the problem but also analyze the possible explanations for the causes of crane-related deaths and injuries,

* Corresponding author.
E-mail addresses: xiaowluo@cityu.edu.hk (X. Luo), wjob@mail.utexas.edu (W.J. O'Brien), Fernanda.Leite@austin.utexas.edu (F. Leite), james_goulet@berkeley.edu (J.A. Goulet).

http://dx.doi.org/10.1016/j.aei.2014.08.004
1474-0346/

for example lack of safety training, inspections, safety plans and traffic control. In addition, OSHA's study identified boom/crane contact with the energized power line (nearly 45% of crane-related accident) and dropped crane loads as some of the major causes of crane accidents [17]. Although there are a large number of safety procedures designed and safety management systems developed for reducing crane-related accidents, the number of injured rate remains high, which further clearly indicated that the existing safety procedures and safety management systems are not effective enough in preventing accidents. Recent advances in sensing and computing technologies offer a potential solution to automate the crane safety monitoring process. Our research envisions an autonomous crane safety monitoring system, which utilizes location data and other information from various sources (e.g., location-aware wireless sensors, accelerometers, building information models). According to CPWR's studies, there were four main types of cranes associated with crane-related fatalities, mobile or truck crane, overhead or gantry crane, tower crane and floating or barge crane, our research will focus on developing tower crane safety monitoring system. The research is divided into three phases: (1) knowledge elicitation, which is a process of extracting safety knowledge and information requirements to serve as a foundation for autonomous safety monitoring system development; (2) development and deployment of the autonomous safety monitoring system in a distributed computing environment; and (3) exploring the impacts of data imperfections situations (e.g., erroneous or missing data) on the safety monitoring system together with approaches to reduce the impacts of the imperfections on the safety system. Prior research has reported on the first two phases [6,12,13,16]. The authors' recent paper [14] has reported the impacts of data imperfections on autonomous jobsite safety monitoring. The working paper indicates that the autonomous jobsite safety monitoring system based on the raw data collected from location-aware sensors is not satisfactory and approaches to deal with imperfect data should be investigated.

The paper is organized as follows: Section 2 summarizes current approaches to autonomous crane safety monitoring and the impacts of data imperfections on such systems; Section 3 describes the implemented autonomous crane safety monitoring system, the scaled testbed, the process of data collection and processing, and the metrics to evaluate the crane safety monitoring system's performance; Section 4 describes the modeling approaches to improve the system performance; Section 5 reports the results of what improvement can be achieved with those approaches; and Section 6 summarized the results and proposed the future research directions.

2. Research background

Tower crane safety monitoring is important on construction jobsites for preventing serious injuries and fatalities. To overcome limitations of manual systems, sensing technologies have been proposed for autonomous crane safety monitoring. One of the most important pieces of information used for construction safety applications is location data. There have been several different technologies available to obtain location information, including GPS, Radio Frequency (RF) sensors, Ultra Wideband (UWB), Ultrasonic, LADAR, laser scanner, Infrared, and video/image-based tracking [24]. For example, Yang [27] used image processing techniques on video streams from surveillance cameras to track if a crane enters predetermined blind lifting areas. Electric field monitoring sensor [18] was also used to monitor if the crane component is getting too close to the power line but it required calibration and the environment might affect the system's performance. Among the techniques mentioned above for the safety alert system, RF-based technology is one of the most promising applications for construction safety. Different RFID technologies have been used to prevent blind lifting for crane operators [10,22] and prevent workers from getting to close to a stationary mobile crane [23] or under the crane [7], and crane collisions [5,28].

The applications demonstrate promising benefits of introducing sensing technologies for autonomous safety monitoring; however, the applications were developed under the assumption that the collected information is perfect, which is not true due to the imperfect nature of sensing devices. Previous studies [1,15,20,21] indicated that imperfect (missing, erroneous, and uncertain) data is common and the imperfect data collected from sensors adversely impacts safety monitoring system performance.

In the authors' recent paper [14], imperfect information in the autonomous safety monitoring system was studied and the results show that imperfect information results in unsatisfactory system performance in terms of precision and recall to the point where commercial implementation is likely non-viable without improvement. The authors' working paper indicates that future research should be conducted to explore potential approaches to deal with imperfect data for improving the safety monitoring system's performance in a realistic world.

Razavi and Haas [20] have recognized that the acquire date from sensing devices are imperfect and found there are five major frameworks to deal with imperfect data: probabilistic, evidential belief reasoning, soft computing, optimization-based, and hybrid methods. They modified a data fusion model based on the JDL model with the integration of the five major farmworkers for tracking onsite materials among multi fusion levels (low for location detection, high for relocation detection and meta-process for project state). Since safety management requires near real time decision making based on sensed data, dealing with imperfect data for safety monitoring in this paper focuses at the low level (location detection) based on Razavi and Hass's work. In machine learning-based approach, the number of training set data points used for localization is usually very large and the computing time would be large even based on linear computing complexity. To reduce the burden to the system and speed safety decision making, the computing complexity issue needs to be addressed. In this paper, the authors use Bayesian approach based on few data points for safety situation decision making to reduce the processing burden of the system. Besides location information, this paper includes non-location information (velocity, and acceleration) to assist location prediction modeling. Both soft computing and evidential belief reasoning are adopted for improving the system's performance with imperfect data.

3. Autonomous crane safety monitoring system testbed and data collection

In the autonomous crane safety monitoring system, the dangerous areas under a crane load are divided into three areas based on different risk levels, arranging from the high risk to low risk: the red zone, the yellow zone and the green zone [13]. The target is considered to be in dangerous situations when falling in the red zone, to be close to the dangerous situation when falling in the yellow zone and safe in the green zone. Based on the three types of control areas, the autonomous crane safety monitoring system has two derivatives: (I) Red–Yellow–Green (R/Y/G) system, which contains both warnings and alerts and (II) Yellow–Green (Y/G) system, which only contains warnings. In system I, an alert is triggered if the target is in the red zone and a warning is triggered if the target is in the yellow zone. In system II, a warning is triggered if the target is in the yellow zone or the red zone. The purpose of a Yellow–Green system is to place emphasis on entering a warning zone (avoid any entry into a red zone) while potentially reducing the likelihood of false alarms.

To prevent workers being exposed under the risk, the authors built a scaled testbed (Fig. 1) to simulate the real-world conditions and allow exploring approaches to deal with data errors for improving the autonomous safety monitoring system performance under different scenarios. The authors went to an office building construction jobsite in Austin, with two 6-floors buildings and observed the activities on the jobsite. Based on the observation at framework construction phase, five representative construction operation activities were selected to develop the five test cases to study the different approaches to deal with imperfect data. Although there are few techniques (e.g., differential GPS, Ultra Wideband) with higher accuracy at higher cost, lower-cost non-differential GPS and Wi-Fi-based localization systems are more feasible and mature for safety applications in construction due to their lower investment in initial cost as well as setup and operation. Previous studies [26] indicated that the localization error of GPS is usually in the range of 1–13 m and the error of Wi-Fi-based technologies can be reduced down to 1 m [3]. This range guides the choice of a cricket system for use in the testbed to provide location information because other studies [15,19,29] indicated that the localization error of the cricket system ranges from 1 to 10 cm. The scale between the GPS/Wi-Fi-based technologies' error and cricket system's error is similar to the scale between the dimension of real-world jobsite and our scaled testbed. The testbed was built based on the aforementioned jobsite using LEGO bricks in a laboratory environment. The LEGO model tower crane (City Set #7905) was used as the basic of scaled tower crane in the testbed. Lego Mindstorms NXT was also deployed in on the scaled tower crane as the controller control the servo motors for crane load's movement at three degrees of freedom. Using three servo motors, the model crane's boom can rotate 360° clockwise and anticlockwise, the trolley on the crane boom can move forward and backward along the track under the crane boom, and the hook can rise up and down vertically. The Cricket Development Kit [19], based on ultrasonic and radiofrequency, was deployed to track the location information of the crane load. The beacon nodes of the cricket system were mounted at fixed locations on the ceiling and floor with pre-determined locations. Another cricket node was mounted on the tower crane's load and the worker to continuously send/receive radio frequency and ultrasonic packages to/from the beacon nodes and time difference and transmission speed information were used to calculate the crane load and the worker's location information and sent to the autonomous safety monitoring system. The crane hook's location information and crane load's size are used to generate the red and yellow zones based on the extracted safety knowledge [13]. As the crane load and the worker move, the dangerous area would change and the worker's safety situation would also change in real time.

Five test cases (details can be found in [11] were developed based on five representative construction tower crane operations observed on the jobsite to explore various approaches to deal with imperfect data in the autonomous safety monitoring system. Each test case represented a tower crane operation and the moving path of a nearby worker. Fig. 2 shows test case #2 as an example. In this representative crane operation scenario test case #2 derived from, the tower crane is lowering a steel beam for frame work assembly near the north side of the building under construction while a worker is traveling along the north edge on the roof of the same building. Six discrete points along the worker's travel path and the crane load's vertical lifting path were taken. Based on the scale of the testbed, the corresponding points were first marked in the testbed as the ground truth positions. For each test case, the time interval between each point is 0.5 s and twenty readings were obtained using cricket localization system using a sampling frequency of 200 Hz at each ground truth positions. Since the sampling interval of data at each ground truth is small and the worker's/crane load's positions are considered to stay the same for the 20 readings (young pedestrian's average walking speed is 150 cm/s [9] and therefore the possible movement at each sampling interval would be less than 150/200 = 0.75 cm). The same data collection process was repeated for the other four test cases. The location data collected from the cricket sensors indicated that the average location error ranges from 1.48 cm to 10.06 cm with an average error of 4.77 cm.

At each test point in the five test cases, the ground truth safety situation was calculated based on the ground true location information and the measured safety situation was calculated based on the measured location. By comparing the measure safety situation and the ground true safety situation, the calculated safety situation can fall into one type within true positive, true negative, false positive and false negative [14]. To evaluate the performance of the autonomous crane safety monitoring system, three metrics were used: precision, recall and accuracy.

Precision is the ratio of True Positive to the sum of True Positive and False Positive. It measures the percentage of actual alerts and warnings that are triggered by the system.

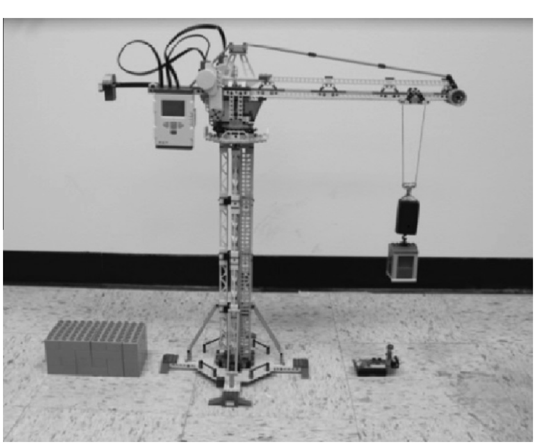

Fig. 1. Model crane in the lab environment.

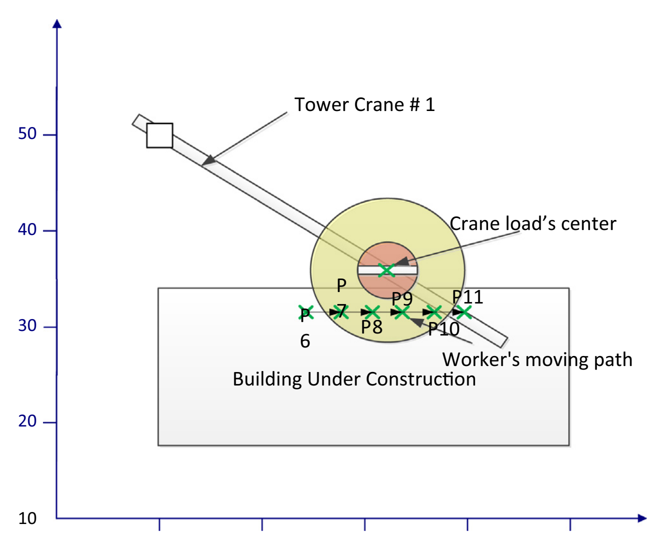

Fig. 2. Test case #2.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (1)$$

Recall is the ratio of True Positive to the sum of True positive and False Negative. It measures the percentage of actual dangerous situations that are reported (when an alert or alarm is triggered by the system).

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (2)$$

Accuracy: the accuracy is the percentage of true results in the population.

$$\text{Accuracy} = \frac{\text{number of true positives} + \text{number of true negatives}}{\text{number of true positives} + \text{true negatives} + \text{false positives} + \text{false negatives}} \quad (3)$$

4. Approaches used to improve localization information with imperfect data

To aid robust implementation, ideally approaches to improve prediction of localization information under imperfect data will be lightweight in terms of computational speed and associated power usage of the distributed sensor and computing network. We introduce five approaches below, including a base case.

4.1. Decision making based on raw data without considering data imperfection (base case)

The raw location data was fed into the autonomous safety monitoring system assuming that the data has no error. The system's performance was calculated based on the raw data and served as the base for performance benchmarking.

The worker's position at time T_n in each test case, based on the kth sample/reading is $\overrightarrow{P_{w,T_n,\text{kth}}}$.

The crane load's position at time T_n in each test case, based on the kth sample/reading is $\overrightarrow{P_{l,T_n,\text{kth}}}$.

Therefore, the distance between the crane load and the worker $D_{T_n,\text{kth}}$ at T_n in each test case can be calculated as:

$$D_{T_n,\text{kth}} = \begin{cases} \text{missing (if either } \overrightarrow{P_{w,T_n,\text{kth}}} \text{ or } \overrightarrow{P_{l,T_n,\text{kth}}} \text{ is missing)} \\ |\overrightarrow{P_{w,T_n,\text{kth}}} - \overrightarrow{P_{l,T_n,\text{kth}}}| \text{(if neither } \overrightarrow{P_{w,T_n,\text{kth}}} \text{ nor } \overrightarrow{P_{l,T_n,\text{kth}}} \text{ is missing)} \end{cases} \quad (4)$$

Using the calculated distance between the crane load and the worker at time T_n in each test case, based on the kth sample/reading, the worker's safety situation can be determined as:

$$\text{Safety}\left(T_n, k^{\text{th}}\right) = f(D_{T_n,\text{kth}}) \quad (5)$$

4.2. Simple averaging

The approach uses multiple location data readings collected in a short period of time (near simultaneously) at the same location to calculate the average and use it to determine the safety situation. By changing the N value, the performance varies and a number of samples required at each test point can be suggested.

The worker's average position at time T_n in each test case, using the first k samples/readings is

$$\overrightarrow{\widetilde{P_{w,T_n,k}}} = \sum_{j=1}^{k} \overrightarrow{P_{w,T_n}} \Big/ k \quad (6)$$

The crane load's average position at time T_n in each test case, using the first k samples/readings is

$$\overrightarrow{\widetilde{P_{l,T_n,k}}} = \sum_{j=1}^{k} \overrightarrow{P_{l,T_n}} \Big/ k \quad (7)$$

Based on the average position using k samples, the distance between the crane load and the worker at T_n in each test case can be calculated as

$$D_{aver,T_n,k} = \left| \overrightarrow{\widetilde{P_{w,T_n,k}}} - \overrightarrow{\widetilde{P_{l,T_n,k}}} \right| \quad (8)$$

Therefore, the worker's safety situation at T_n in each test case can be determined by

$$\text{Safety}(T_n, k) = f(D_{aver,T_n,k}) \quad (9)$$

4.3. Prediction using location readings of the previous two points and a vector for prediction

The approach uses location data at the previous two points to calculate the vector of the traveling distance and uses this vector to predict the possible location at the next time stamp. Then the safety monitoring system would use the predicted location to determine the system's performance.

For each test case,

The worker's location $\overrightarrow{Pp_{w,T_n,k}}$ at T_n (for $n > 2$) can be predicted based on the previous two sensed locations as:

$$\begin{aligned} \overrightarrow{Pp_{w,T_n,k}} &= \overrightarrow{\widetilde{P_{w,T_{n-1},k}}} + \left(\overrightarrow{\widetilde{P_{w,T_{n-1},k}}} - \overrightarrow{\widetilde{P_{w,T_{n-2},k}}} \right) \\ &= 2 * \overrightarrow{\widetilde{P_{w,T_{n-1},k}}} - \overrightarrow{\widetilde{P_{w,T_{n-2},k}}} \end{aligned} \quad (10)$$

The crane load's location $\overrightarrow{Pp_{l,T_n,k}}$ at T_n (for $n > 2$) in each test case can be predicted based on the previous two predicted locations as:

$$\begin{aligned} \overrightarrow{Pp_{l,T_n,k}} &= \overrightarrow{\widetilde{P_{l,T_{n-1},k}}} + \left(\overrightarrow{\widetilde{P_{l,T_{n-1},k}}} - \overrightarrow{\widetilde{P_{l,T_{n-2},k}}} \right) \\ &= 2 * \overrightarrow{\widetilde{P_{l,T_{n-1},k}}} - \overrightarrow{\widetilde{P_{l,T_{n-2},k}}} \end{aligned} \quad (11)$$

Based on the average position using k samples, the distance between the crane load and the worker at T_n ($n > 2$) in each test case can be calculated as

$$D_{pred,T_n,k} = \left| \overrightarrow{Pp_{w,T_n,k}} - \overrightarrow{Pp_{l,T_n,k}} \right| \quad (12)$$

Therefore, the worker's safety situation at T_n ($n > 2$) in each test case can be determined by

$$\text{Safety}(T_n, k) = f(D_{pred,T_n,k}) \quad (13)$$

4.4. Prediction using predicted location data at the previous point with velocity and acceleration

The approach uses the velocity and acceleration collected at the previous point to calculate the traveling distance between the previous timestamp and current time. The location at current point can be predicted by adding the traveling distance on the top of the location at the previous point. Then the safety monitoring system would use the predicted location to determine the system's performance.

In a general situation, an object's location at T_i can be predicted based on its predicted location, velocity and acceleration at T_{i-1} using the equation below:

$$\overrightarrow{P_{T_i}} = \overrightarrow{P_{T_{i-1}}} + \overrightarrow{v_{T_{i-1}}} * (T_i - T_{i-1}) + 0.5 * \overrightarrow{a_{T_{i-1}}} * (T_i - T_{i-1})^2 \quad (14)$$

In the author's testbed, only location information is collected in this phase. Based on the observation on jobsite, the worker's velocity value and the crane load's angular velocity are approximately constant in the middle of certain activities. Therefore, the author assumes that the value of the worker's linear velocity and the value of the crane load's angular velocity are constant; the worker's linear acceleration and the crane load's angular acceleration are zero. The direction of the simulated velocity reading is set with the value of the previous two sensed locations. Therefore,

The worker's location $\overrightarrow{Pp_{w,T_n,k}}$ at T_n (for $n > 2$) can be predicted based on the previous sensed locations as

$$\overrightarrow{Pp_{w,T_n,k}} = \overrightarrow{P_{T_{n-1}}} + \overrightarrow{v_{T_{n-1}}} * T \tag{15}$$

where T is the timer interval between T_n and T_{n-1} (0.5 s in our test cases)

$$\overrightarrow{v_{T_{n-1}}} = * \left(\overrightarrow{P_{T_{n-1}}} - \overrightarrow{P_{T_{n-2}}}\right) \Big/ \left|\overrightarrow{P_{T_{n-1}}} - \overrightarrow{P_{T_{n-2}}}\right| \tag{16}$$

The crane load's location $\overrightarrow{Pp_{l,T_n,k}}$ at T_n (for $n > 2$) can be predicted based on the previous sensed locations as

$$\overrightarrow{Pp_{l,T_n,k}} = \overrightarrow{P_{T_{n-1}}} + \overrightarrow{v_{T_{n-1}}} * T = \overrightarrow{P_{T_{n-1}}} + \overrightarrow{\omega_{T_{n-1}}} * \frac{\left|\overrightarrow{r_{T_{n-1}}}\right|^2}{\overrightarrow{r_{T_{n-1}}}} * T \tag{17}$$

where T is the timer interval between T_n and T_{n-1} (0.5 s in our test cases).

Based on the average position using k samples, the distance between the crane load and the worker at T_n ($n > 2$) in each test case can be calculated as

$$D_{pred,T_n,k} = \left|\overrightarrow{Pp_{w,T_n,k}} - \overrightarrow{Pp_{l,T_n,k}}\right| \tag{18}$$

Therefore, the worker's safety situation at T_n ($n > 2$) in each test case can be determined by

$$\text{Safety}(T_n, k) = f\left(D_{pred,T_n,k}\right) \tag{19}$$

4.5. Bayesian approach with velocity and acceleration

Bayesian conditional probability [2] is well known and widely applied in almost every field of science. In the crane safety scenario, the hazard state of the worker can be categorized and her probability of being in any of these states can be described. The three variables describing the worker's safety conditions are:

1. In the green zone.
2. In the yellow zone.
3. In the red zone.

The alarm (if the worker is in the red zone) or warning (if the worker is in the yellow zone) triggering is based on the probability $P(\alpha)$ of the worker being in the safety condition. If $P(\alpha)$ is larger than a threshold $\varnothing$ where $\varnothing \in [0, 1]$ (in our test, 50% was chosen as the threshold), a signal would be sent to the worker for notifying her safety situations. In practice, $P(\alpha)$ can be estimated by updating the probability of being in a hazardous situation. $P(\alpha|E)$ conditionally on observations (E) made of the worker's position relative to the hazard.

The Bayesian approach consists of defining and continuously updating the probability $P(\alpha|E)$ in a Bayesian manner, based on the worker's monitored behaviors. The position, velocity and acceleration of both the worker and the crane load are monitored at a frequency $f\left(\frac{1}{T}\right)$. Based on the captured data, three quantities can be obtained at any time step i: the worker's distance to the dangerous area's center ($D_{w,i}$), the worker's velocity towards the dangerous area's center ($v_{w,i}$), and the worker's acceleration towards the dangerous area's center ($a_{w,i}$). Defining $P(\alpha|E)$ based on these three quantities allows the transient characteristics of a worker's behavior to be included. For instance, if someone is working close to the red zone and is conscious of the danger, no alarm or signal should be given. In another case, if a worker is still relatively far from the red zone, but walking at constant speed towards it with no sign of slowing down, an alarm should be triggered. This heuristic knowledge is encapsulated in a likelihood function $P(E|\alpha) = g(D_{w,i}, v_{w,i}, a_{w,i})$ and is used to update previous knowledge about the hazard state $P(\alpha|E)$. Based on the theory of experimental design [4,8] a set of 45 events $x = [x_1, x_2, \ldots, x_{45}]^T$ where $x_n = [D_{w,n}, v_{w,n}, a_{w,n}]$ for $n \in [1, 45]$ are defined in a central-composite-design arrangement. The probability of being in an imminent hazardous situation $P(\alpha|E)$ has to be quantified for each point based on a mutual agreement.

The probability $P_n(E|\alpha)$ has to be evaluated for each point n in order to define a second-order polynomial function that can return $P(\alpha|E)$ in the entire domain using Eq. (20).

$$\vec{P}(E|\alpha) = \beta_0 + \sum_{i=1}^{3}\beta_i x_i + \sum_{i=1,j=1}^{3}\beta_{3i+j-1}x_i x_j + \epsilon_n \tag{20}$$

Eq. (21) is the matrix representation of Eq. (20) where β_i are parameters determined using Eq. (23). $\vec{\beta}$ is a vector containing the parameters β_i that minimizes the approximation error ϵ_n.

$$\begin{bmatrix} P_1(E|\alpha) \\ P_2(E|\alpha) \\ \vdots \\ P_{45}(E|\alpha) \end{bmatrix} = \begin{bmatrix} 1 & x_{1,1} & x_{1,2} & \cdots & x_{1,3}^2 \\ 1 & x_{2,1} & x_{2,2} & \cdots & x_{2,3}^2 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{45,1} & x_{45,2} & \cdots & x_{45,3}^2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_9 \end{bmatrix} + \begin{bmatrix} \epsilon_0 \\ \epsilon_1 \\ \vdots \\ \epsilon_{45} \end{bmatrix} \tag{21}$$

$$\vec{P}(E|\alpha) = X\vec{\beta} + \vec{\epsilon_r} \tag{22}$$

$$\vec{\beta} = (X^T X)^{-1} X^T \vec{P}(E|\alpha) \tag{23}$$

The probability of being in an imminent hazardous situation at a time i, $P(\alpha|E)|_{t=i}$ can be obtained by updating the previous state of knowledge $P(\alpha|E)|_{t=i-\xi}^{t=i-1}$ obtained in the last ξ ($\xi \in N$) iterations with the likelihood function $P(E|\alpha)|_{t=i}$ computed from monitored data obtained at time i. ξ can either be defined deterministically or it can be updated recursively using the state of $P(\alpha|E)|_{t=i-1}$ as shown in Eq. (24). In this equation, κ ($\kappa \in N$) is weighted by $P(\alpha|E)|_{t=i-1}$. When the hazard is low, it is useful to include a large number of previous points in order to avoid potentially misleading outliers. When the hazard is high, the window size should be reduced in order to adapt quickly to changing situations.

$$\xi = \lfloor \kappa / P(\alpha|E)|_{t=i-1} \rfloor \tag{24}$$

The Eq. (6) presents the Bayesian updating procedure where previous knowledge is updated using the likelihood function (Eq. (20)) translating the relative position, speed and acceleration of a worker into a probability of being in a hazardous state. In this equation, $P(E)$ is a normalization constant itself evaluated in Eq. (26).

$$P(\alpha|E)|_{t=i} = \frac{P(E|\alpha)P(\alpha|E)|_{t=i-\xi}^{t=i-1}}{P(E)} = \frac{1}{P(E)}\prod_{j=i-\xi}^{i} P(E|\alpha)|_{t=j} \tag{25}$$

$$P(E) = \left(\left(\prod_{j=i-\xi}^{i} 1 - P(E|\alpha)|_{t=j}\right) + \prod_{j=i-\xi}^{i} P(E|\alpha)|_{t=j}\right)^{-1} \tag{26}$$

5. Results

Using each individual approach, the safety situation at each test point can be calculated and used to evaluate the autonomous

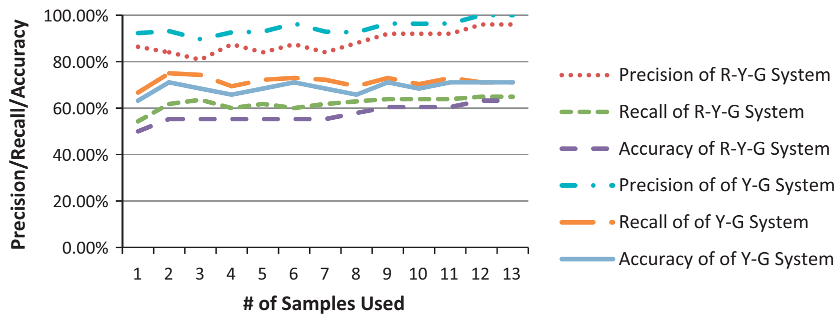

Fig. 3. Performance of simple averaging with number of used samples changing.

Table 1
Performance comparison of various approaches to deal with imperfect data.

	R/Y/G system			Y/G system		
	Precision (%)	Recall (%)	Accuracy (%)	Precision (%)	Recall (%)	Accuracy (%)
Raw location data	87.6	53.0	49.3	94.9	62.7	60.7
Simple averaging using 10 location readings	92.0	63.9	60.5	96.3	70.3	68.4
Prediction using location reading of the previous two points and a vector for prediction	100.0	35.7	35.7	100.0	39.3	39.3
Bayesian approaches with velocity and acceleration	76.7	74.2	60.5	93.5	80.6	76.3
Prediction using predicted location data at previous point with velocity and acceleration	93.8	96.8	90.9	100.0	97.0	97.0

safety monitoring system performance. Fig. 3 shows the number of samples used to obtain the average location data for safety situation calculation. By increasing the number of samples used for averaging, the performances of both systems get improved. Based on the figure, when a number between 10 and 13 is chosen as the number of samples used for averaging, the performance reaches a flat platform.

The performances of various approaches are summarized in Table 1 and Fig. 4. Because the localization error is large based on the data I collected, as well as the relative path of the worker might not be a straight line, the approach of simple prediction using the vector based on the previous two sensed location does not reach a better performance than simple used the raw data without any data processing approach. Simple averaging using the first ten readings also averages the localization error down. Therefore, the simple averaging using the first ten readings gets a little better performance than just using raw data for decision making. If additional information (e.g., velocity and acceleration) other than location can be obtained and used to support decision making, the performances of both systems get improved and both systems have performance metrics with value of over 80%.

Fig. 5 shows the detailed safety decisions made with the Bayesian approach in the R/Y/G system and the Y/G system. In both systems, the calculated safety situations at points in the green zone

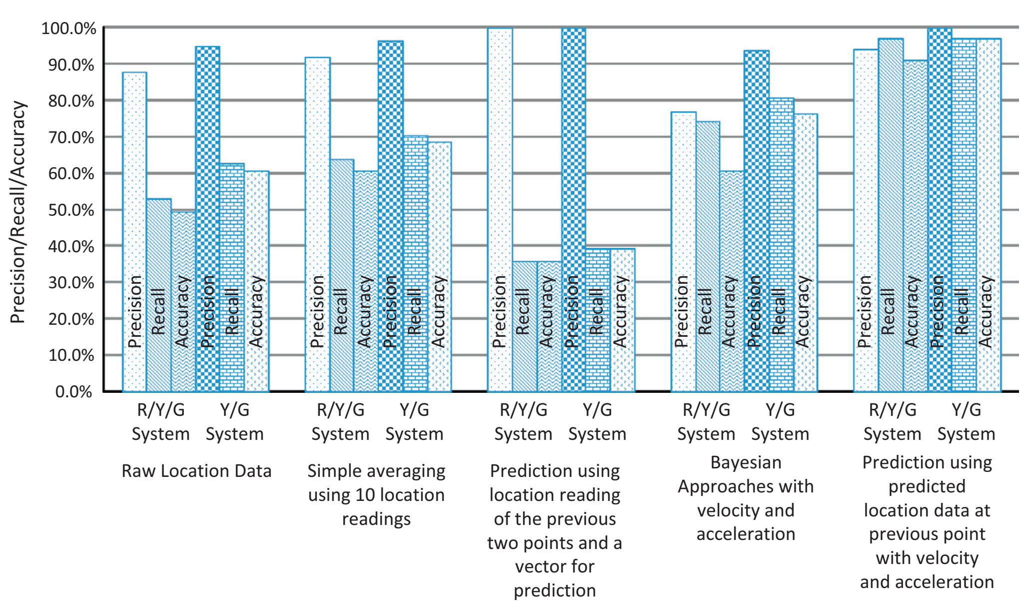

Fig. 4. Performance comparison of various approaches.

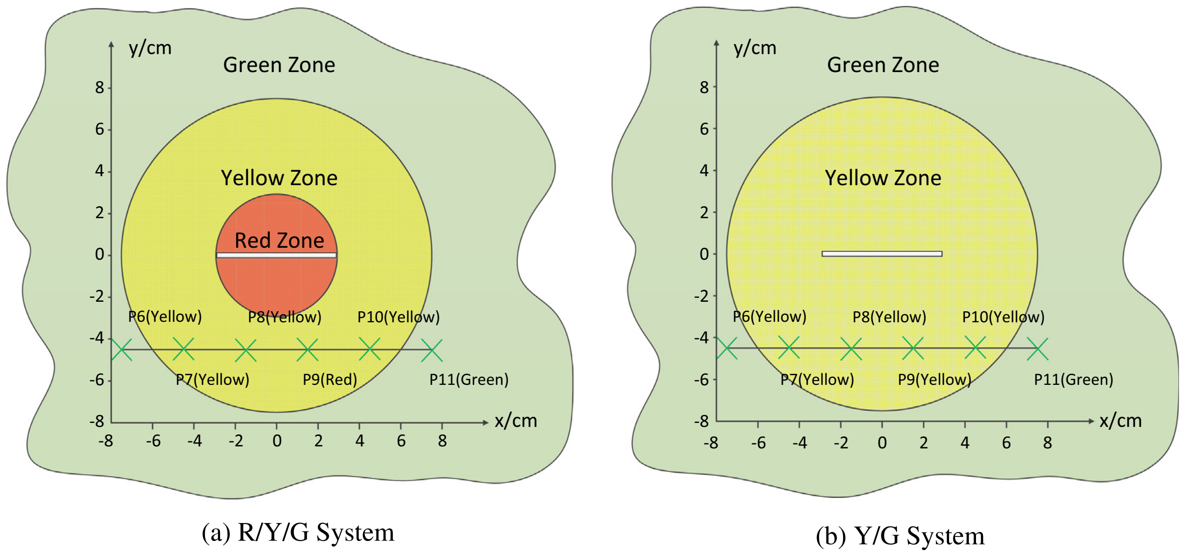

Fig. 5. Safety situation along the worker's moving path in test case #2 made by bayesian approach.

Table 2
Systems performance of Y/G system and R/Y/G system in green zone.

Applied approach	Y/G system and R/Y/G system		
	Precision (%)	Recall (%)	Accuracy (%)
Raw location data	86.1	100.0	86.1
Simple averaging using 10 location readings	88.9	100.0	88.9
Prediction using location reading of the previous two points and a vector for prediction	100.0	100.0	100.0
Bayesian approaches with velocity and acceleration	77.8	100.0	77.8
Prediction using predicted location data at previous point with velocity and acceleration	100.0	100.0	100.0

are the same. The performance would be the same for the same approach at the points in the green zone. False alarms would be raised at some points (e.g., P6) in this area and we want to keep the false alarm rate as low as possible. In this way, the worker's would trust the system and would not turn the system off. In test case #2, the Bayesian approach achieved 100% accuracy at points inside for the yellow zone for the Y/G system (Fig. 5b) and only one occurrence (at P9) of false alarm is observed in the R/Y/G system (Fig. 5a). At P9 in the R/Y/G system, the system exaggerates the worker's dangerous situations. It is more acceptable than failing to give any warning to the worker and telling her that she is in a safe situation. The result indicates that switching the system from the R/Y/G system to the Y/G system improves the autonomous safety monitoring system's performance.

Table 2 and Fig. 6 show the performance for both systems based on data I collected from the testbed for points in the green zone.

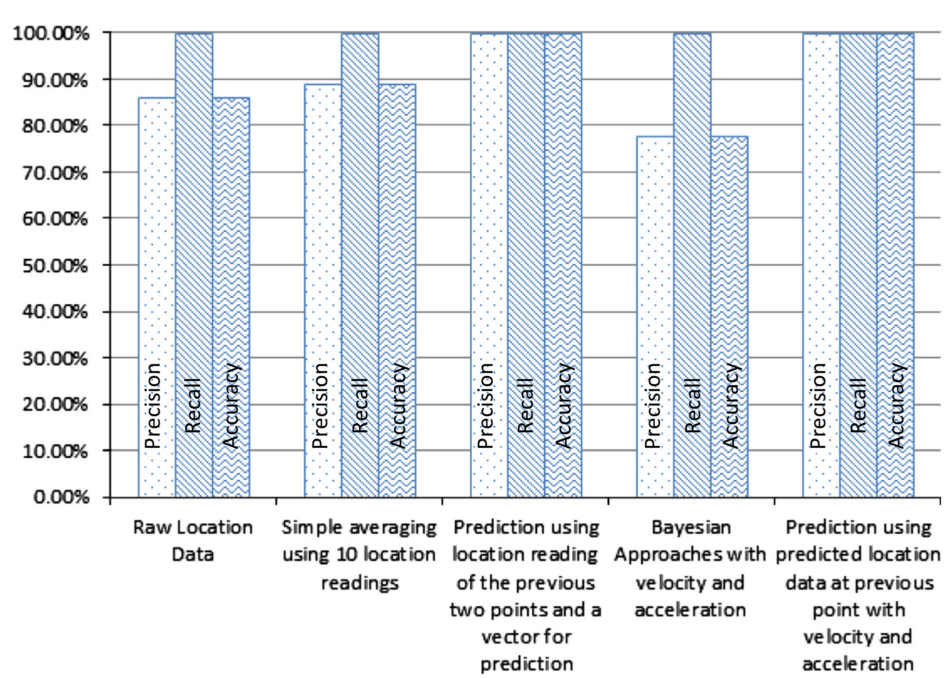

Fig. 6. Systems performance of Y/G system and R/Y/G system in green zone.

Table 3
Systems performance of Y/G system and R/Y/G system in yellow zone.

Applied approach	Y/G system			R/Y/G system		
	Precision (%)	Recall (%)	Accuracy (%)	Precision (%)	Recall (%)	Accuracy (%)
Raw location data	100.0	43.9	43.9	85.5	40.1	37.5
Simple averaging using 10 location readings	100.0	54.6	54.6	91.7	52.4	50.0
Prediction using location reading of the previous two points and a vector for prediction	100.0	25.0	25.0	100.0	25.0	25.0
Bayesian approaches with velocity and acceleration	100.0	72.7	72.7	68.8	64.7	50.0
Prediction using predicted location data at previous point with velocity and acceleration	100.0	95.0	95.0	89.5	94.4	85.0

Both the Y/G and R/Y/G systems have the same performance in this green zone. For the points in the green zone, the decision making based on raw data reach nearly 86% accuracy and precision, which means the calculated safety situation is also green. The prediction using simple vector and the prediction using predicted data and additional information reach accuracies and precisions of 100%, in which cases all the points have been correctly identified in green zone. For all the other three approaches other than those two with 100% accuracies, although the system incorrectly identifies some points, the decisions are not so bad since no points have been incorrectly identified to be in the red zone. The recall rates of all the approaches are 100% because all the safe situations were reported. Based on the results, the system might get better performance when applying the approach of prediction with simple vector and the approach of prediction using predicted location and additional information on the sensed data. The other two approaches and decision making with raw data also get an acceptable performance.

Table 3 and Fig. 7 show the performance for both systems based on data I collected from the testbed for points in the yellow zone. The results are quite consistent with the overall result when considering all the points in different zones as a whole set. For the points in the green zone, the prediction using predicted location data and additional information has the best performances for both the Y/G and R/Y/G systems in terms of precision, recall and accuracy. The Bayesian approach has the 2nd best performance. The results indicate that by introducing additional information (e.g., acceleration and velocity) to support decision making, the performances of both systems get improved and most of the performance metrics have values of over 70%.

Table 4 and Fig. 8 show the system's performance (in terms of percentage of points considering to be in a specific safety situation) based on data I collected from the testbed for points in the red zone. For the R/Y/G system, both the simple averaging approach and the approach of prediction using predicted location with velocity and acceleration have the best performance. The Bayesian approach has a similar performance as decision making simply using raw data from the sensors. For the Y/G system, the approach of prediction using predicted location with velocity and acceleration also has the best performance and the Bayesian approach has the 2nd best performance. The prediction approach using simple vector had the worst performance for both systems.

6. Discussion

The simple averaging approach used to feed location for safety situation calculation provides a more stable system's performance as the location errors fluctuate less. The system's performance tended to be more stable if more sample location data were used to calculate the average location. Ten to thirteen would be a *satisfactory* number to use as the sample size of location data that should be collected at a specific location along the moving path in a short period of time (in this paper, we used 50–65 ms). This approach was a simple and effective way to improve the system's performance for a little bit without a need to deploy extra sensors for other types of data.

Deploying extra sensors to collect velocity and acceleration data can greatly improve the system's performance. Both the Bayesian approach and the prediction using predicate location data, velocity & acceleration utilize that information to provide a more accurate safety situation of the workers. The predication approach using velocity and acceleration in addition to location data achieved

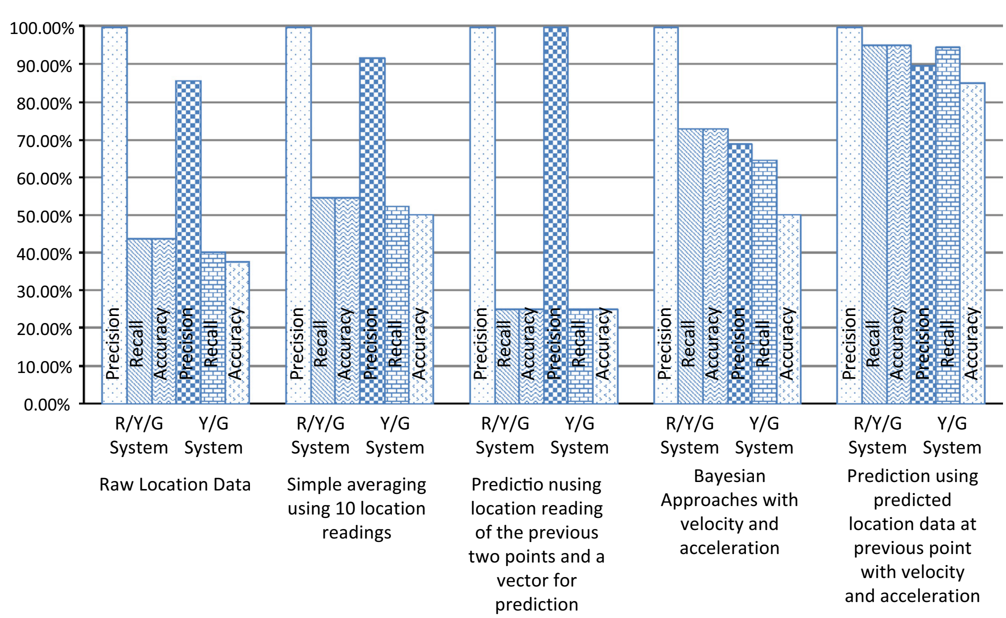

Fig. 7. Systems performance of Y/G system and R/Y/G system in yellow zone.

Table 4
Systems performance of Y/G system and R/Y/G system in red zone.

Applied approach	Y/G system			R/Y/G system		
	Precision (%)	Recall (%)	Accuracy (%)	Precision (%)	Recall (%)	Accuracy (%)
Raw location data	100.0	80.7	80.7	100.0	39.3	39.3
Simple averaging using 10 location readings	100.0	100. 0	100.0	100.0	57.1	57.1
Prediction using location reading of the previous two points and a vector for prediction	100.0	28.6	28.6	100.0	14.3	14.3
Bayesian Approaches with velocity and acceleration	100.0	85.7	85.7	100.0	71.4	71.4
Prediction using predicted location data at previous point with velocity and acceleration	100.0	100.0	100.0	100.0	100.0	100.0

the best performance in the test cases used in this paper. The performance was dependent on the path of worker and crane load. In the test cases described in this paper, the worker was walking in straight line and the crane load was moving along a circle. This simple pattern of the worker and the crane load's moving path was a possible explanation of why this approach had a very high accuracy in deciding the worker's safety situations. The system's performance might degrade if the worker and crane load's moving path did not have such a clear pattern. Future study would include some test cases with irregular moving paths to test this approach's performance.

The Bayesian approach had the 2nd best performance in the test cases described in this paper. The Bayesian approach was dependent on its prior knowledge, the size of time windows use to collect prior knowledge for updating, the confidence level used as probability threshold. Therefore, the Bayesian approach approach's performance was variable based on all these factors. An appropriate set of prior knowledge and time windows can improve the autonomous safety monitoring system's performance using the Bayesian approach.

7. Conclusions

Localization through sensed data is a goal of many research and development efforts. The level of error in the sensed environment on construction sites is relatively high, calling into question the suitability of many technologies for location awareness that requires precision beyond very coarse location accuracy. This paper provides a path forward where basic approaches are used to improve accuracy, from simple averaging of multiple readings to the use of supplemental velocity and acceleration data. All of these approaches show some improvement over use of raw data. Supplemental data in particular appears to show very promising results, whether using a Bayesian approach to update beliefs about location or by simple prediction from prior sensed location. Importantly, all the approaches work without detailed calibration and setup of the sensor network, and, as such, can be applied to networks where the error is unstable due to environmental conditions or transient placement of sensors.

The authors believe their work takes an important step to move the research literature away from analysis of error to development of approaches that actively improve the information available from the sensed environment. However, the work is but a first step. Future research should investigate other approaches and further refine the approaches presented above, particularly in different environmental contexts and conditions. Another next step of the research would include mounting the localization sensors on tower cranes and deploying the autonomous safety monitoring system in a real jobsite with the proposed approaches to test the system performance in a full-scale environment. The research discussed in this paper was implemented on a scaled-down testbed, which might not achieve the same performance as in a full-scale and real-world environment. Companion research would also better specify accuracy requirements for different applications. Ideally, research would then provide developers and practitioners robust technologies with known errors and the appropriate applications for use.

Acknowledgements

The authors wish to thank the anonymous reviewers as well as the personnel at the job sites who provided examples of crane

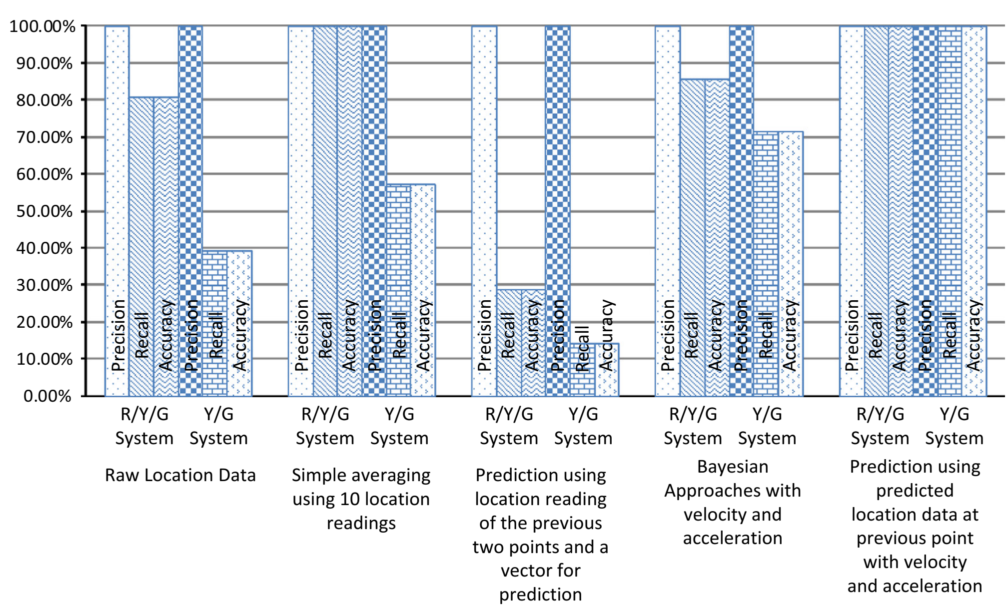

Fig. 8. Systems performance of Y/G system and R/Y/G system in red zone.

positioning and movement that the testbed scenarios were based on. This material is based upon work supported in part by the National Science Foundation under Grant No. OCI-0753360.

References

[1] S. Chi, C.H. Caldas, Image-based safety assessment: automated spatial safety risk identification of earthmoving and surface mining activities, J. Construct. Eng. Manage. 138 (3) (2011) 341–351.
[2] A.P. Dempster, A generalization of bayesian inference, J. Roy. Stat. Soc.: Ser. B (Methodol.) 30 (2) (1968) 205–247.
[3] Ekahau, How does Ekahau RTLS Work, 2014. available from: <http://www.ekahau.com/real-time-location-system/technology/how-rtls-works>.
[4] R.A. Fisher, Design of experiments, Oliver and Boyd, Edinburgh, 1935.
[5] S. Hwang, Ultra-wide band technology experiments for real-time prevention of tower crane collisions, Autom. Construct. 22 (2012) 545–553.
[6] C. Julien, J. Hammer, W. O'Brien, A dynamic architecture for lightweight decision support in mobile sensor networks, in: Proceedings of the Workshop on Building Software for Pervasive Comp, 2005
[7] K. Kim, M. Kim, RFID-based location-sensing system for safety management, Person. Ubiquit. Comput. 16 (3) (2012) 235–243.
[8] R.O. Kuehl, Design of Experiments: Statistical Principles of Research Design and Analysis, Duxbury Press, 1999.
[9] J.N. Laplante, T.P. Kaeser, The continuing evolution of pedestrian walking speed assumptions, ITE J. (2004) 32–40.
[10] G. Lee, H.-H. Kim, C.-J. Lee, S.-I. Ham, S.-H. Yun, H. Cho, B.K. Kim, G.T. Kim, K. Kim, A laser-technology-based lifting-path tracking system for a robotic tower crane, Autom. Construct. 18 (7) (2009) 865–874.
[11] X. Luo, Requirements, Specifications and Deployment Models for autonomous Jobsite Safety Proximity Monitoring, Ph.D. The University of Texas at Austin, Austin, TX, 2013.
[12] X. Luo, W.J.O. Brien, F. Leite, What is a safe working zone? Design for proactive fall prevention, in: Proceedings of the 18th Annual Workshop of the European Group for Intelligent Computing in Engineering (EG-ICE), 2011, pp. 1–9.
[13] X. Luo, F. Leite, W.J. O'Brien, Requirements for autonomous crane safety monitoring, Proc. ASCE Comput. Civil Eng. (2011) 331–338.
[14] X. Luo, W.J. O'Brien, F. Leite, Evaluating the impact of location-aware sensor data imperfections on autonomous jobsite safety monitoring, in: 2013 ASCE International Workshop on Computing in Civil Engineering, Los Angeles, CA, USA, 2013, p. 8.
[15] X. Luo, W.J. O'Brien, C.L. Julien, Comparative evaluation of Received Signal-Strength Index (RSSI) based indoor localization techniques for construction jobsites, Adv. Eng. Inform. 25 (2) (2011) 355–363.
[16] W.J. O'brien, C. Julien, J. Hammer, I. Hassan, S. Kabadayi, X. Luo, J. Trujillo, An architecture for local decision support in ad hoc sensor networks, in: Proceedings of the 2007 ASCE International Workshop on Computing in Civil Engineering, 2007.
[17] OSHA, Crane and Hoist Safety, 2014 <https://www.osha.gov/archive/oshinfo/priorities/crane.html> (21.05.2014).
[18] R. Parr, L. Van Kampen, W. Dimopoulos, S. Sood, Power Line Proximity Monitoring System and Method, Google Patents, 2013.
[19] N.B. Priyantha, A. Chakraborty, H. Balakrishnan, The cricket location support system, in: Proceedings of the 6th Annual International Conference on Mobile Computing and Networking, ACM, 2000, pp. 32–43.
[20] S.N. Razavi, C.T. Haas, Multisensor data fusion for on-site materials tracking in construction, Autom. Construct. 19 (8) (2010) 1037–1046.
[21] S.N. Razavi, C.T. Haas, Using reference RFID tags for calibrating the estimated locations of construction materials, Autom. Construct. 20 (6) (2011) 677–685.
[22] A. Shapira, Y. Rosenfeld, I. Mizrahi, Vision system for tower cranes, J. Construct. Eng. Manage. 134 (5) (2008) 320–332.
[23] J. Teizer, B.S. Allread, C.E. Fullerton, J. Hinze, Autonomous pro-active real-time construction worker and equipment operator proximity safety alert system, Autom. Construct. 19 (2010) 630–640.
[24] J. Teizer, D. Lao, M. Sofer, Rapid automated monitoring of construction site activities using ultra-wideband, in: Proceedings of the 24th International Symposium on Automation and Robotics in Construction, Kochi, Kerala, India, 2007, pp. 19–21.
[25] The Center for Construction Research and Training, Crane-Related Deaths in Construction and Recommendations for Their Prevention, 2009 <http://www.cpwr.com/sites/default/files/CPWR%20Crane%20Rept%20Recmmdtns%20Nov-2009-BLS%20UPDATED.pdf> (01.11.13).
[26] R.C. Weih, M. Gilbert, J. Cross, D. Freeman, Accuracy assessment of recreational and mapping grade GPS receivers, J. Arkan. Acad. Sci. 63 (2009) 163–168.
[27] J. Yang, O. Arif, P. Vela, J. Teizer, Z. Shi, Tracking multiple workers on construction sites using video cameras, Adv. Eng. Inform. 24 (4) (2010) 428–434.
[28] C. Zhang, A. Hammad, S. Rodriguez, Crane pose estimation using UWB real-time location system, J. Comput. Civil Eng. 26 (5) (2012) 625–637.
[29] W. Zhou, J. Whyte, R. Sacks, Construction safety and digital design: a review, Autom. Construct. 22 (2012) 102–111.