Building occupant level estimation based on heterogeneous information fusion

Heng-Tao Wang, Qing-Shan Jia, Chen Song, Ruixi Yuan, Xiaohong Guan

Abstract

It is of great practical interest to estimate the number of occupants at a zonal level in buildings, which is useful for energy-efficient control of air conditioning and lighting systems. We consider this important problem in this paper. First, the occupant level estimation problem is formulated as an information fusion problem with heterogeneous information sources with the criterion of the minimum mean square error (MMSE). Two fusion methods are developed. The first method assumes independent observation noises and the second method exploits the correlation among the multiple information sources to improve the estimation accuracy. The experimental results show that in comparison with individual RFID or video cameras, the two fusion methods improve the accuracy of occupant level estimation by 43% and 73%, respectively, and outperform the linear least mean square error (LLMSE) method. Simulations and theoretical analysis are also conducted to analyze the performance of the two methods under different occupant levels and different correlated observations. It is shown that the second method is more effective for the cases where the multi-sensor measurements are highly correlated.

1. Introduction

Buildings are responsible for nearly 40% of the energy consumption in the United States and in Europe. Measuring the number of occupants at a zonal level in buildings is the basis for energy efficient control of heating, ventilation, and air conditioning (HVAC) systems and lighting systems under normal situations, and for fast evacuation in emergency situations. Many approaches have been developed to estimate the occupant level in indoor environment and their accuracies depend on various factors. For approaches that are based on radio frequency identification (RFID) and the wireless sensor network, the received signal strength indicator (RSSI) is used to locate and to track occupants. Their estimation accuracies are affected by the multipath effect. For approaches that use video cameras, the accuracies...
depend on the background lighting, the density of the occupants, and the angle of the view \[14\]. Therefore it is of practical interest to combine multiple systems to improve the accuracy of the estimation of the occupant level.

Many information fusion methods follow the Bayesian framework to combine the measurements of multiple sensors \[1,6,21,28,41\]. This framework requires the joint distribution of the measurements of multiple sensors. But the RFID and video cameras are usually calibrated independently according to the historical data. And in most of the cases there are not any data to estimate the joint distribution of the observations from the two systems.

We consider this important problem in this paper. First, the occupant level estimation problem is formulated as an information fusion problem with heterogeneous sources under the criterion of the minimum mean square error \( (\text{MMSE}) \). Two approximate fusion methods are developed. The first method assumes independent observations of the multiple systems. The second method exploits the correlation among the multiple information sources. The experimental results show that in comparison with the measurement based on individual RFID or video cameras, the fusion methods can improve the accuracy of occupant level estimation by 43% and 73%, respectively, and outperform the linear least mean square error \( (\text{LLMSE}) \) method \[3\]. Theoretical analysis and simulations are also conducted to analyze the performance of the two methods under different occupant levels and different correlation among the observations. It is shown that the second method is more effective for the cases where the multi-sensor measurements are highly correlated.

The rest of the paper is organized as follows. A brief literature review on existing occupant localization systems in an indoor environment is presented in Section 2. The problem formulation is shown in Section 3. The two approximate fusion methods are presented in Section 4. The experimental results with the fusion of the RFID system and video cameras are shown in Section 5. The advantage of information fusion methods in the occupant level estimation problem, and the impact of the correlation among multiple systems are discussed in Section 6. A brief conclusion is presented in Section 7.

2. Literature review

Many methods have been developed for occupant localization and counting in an indoor environment. These methods can be classified into two groups. The first group, including RADAR \[4\], SpotON \[22\], LANDMARC \[34\], Ekahau \[16\], and UWB \[36\], is based on strength measurements of Radio Frequency (RF) signals. These methods can be significantly affected by indoor electric–magnetic conditions. The other group, including Active Badge \[40\], Cricket \[35\], and Easy Living \[10\], is based on the information from infrared \[39\], ultrasound, or video cameras, and the estimation accuracy may vary significantly as the environment changes \[19\].

Li et al. \[25\] developed an occupant monitoring system based on an RFID. Liao et al. \[30\] proposed an agent-based method with graphical modeling for building occupancy. Casas et al. \[11\] used face detection in video image to estimate the density of occupant. Lin et al. \[29\] estimated the occupant density in crowds by detecting the featured areas of the head contours, and then classified these areas by Support Vector Machine (SVM) method. Guo et al. \[20\] estimated crowd density based on Markov random field. A common issue with the above methods is that the estimation accuracy based on a single sensing sources can be easily affected by the measurement environment.

There are many information fusion methods to improve the accuracy of the estimation, such as the Bayesian method \[8\], the Bayesian network \[27,42\], the Dempster–Shafer evidence combination rule \[7,26,37\], the fuzzy logic \[37\], and the neural network \[27\], SVM \[6,44\], just to name a few. Multiple methods can be combined for further improvement \[27,37\]. For example, Alvarez-Alvarez et al. \[2\] proposed a method for human activity recognition in indoor environment by fusing information from WIFI and accelerations. Meyn et al. \[33\] developed a sensor-utility-network model under the criterion of maximum a posterior (MAP) probability to improve the estimation accuracy for occupant distribution. To locate a robot, Bonci et al. \[9\] proposed an Extended Kalman Filtering method by fusing the information of odometric, video camera, and sonar measures. Acharya et al. \[1\] combined the information of odometer, accelerometer, and GPS, and proposed a new scheme for longitudinal localization for trains. These fusion methods generally require the joint distribution of the measurements of the multiple sensing sources, which is usually not available in our problem.

3. Problem formulation

Consider a large office which is divided into multiple zones according to the positions of air outlets and lights. An example of three zones is shown in Fig. 1. Because the estimation of occupant level in each zone follow similar procedures, we focus...
on this estimation problem for a single zone. Suppose there are m localization systems, each of which can estimate the number of occupants in the zone. Let \( n_k \) be the number of occupants that is observed by system \( k \), and \( n_k \) be the number of occupants that is estimated by system \( k \), where \( k = 1, \ldots, m \). Note that it is possible that \( n_k \neq \hat{n}_k \). For example, the video camera system may observe that there are 10 occupants in the zone, i.e., \( n_k = 10 \). If on the average only 80% of the objects detected by the video cameras are occupants while the rest 20% are not (i.e., false detection), the system may estimate that there are only 8 occupants in the room, i.e., \( \hat{n}_k = 8 \). Define \( \mathbf{n}_s = (n_1, \ldots, n_m)^\top \) and \( \mathbf{n}_o = (\hat{n}_1, \ldots, \hat{n}_m)^\top \), where \( \mathbf{n}_s \) and \( \mathbf{n}_o \) denote the vector of observations and the vector of estimations of the m systems, respectively, and \( \tau \) represents the transpose.

Then, the question is how to find an estimate \( \hat{n} \) of \( n \) given \( \mathbf{n}_o \) to minimize the mean square error, i.e.,

\[
\min_{\hat{n} \in \mathbb{Z}^+} \mathbb{E}[(\hat{n} - n)^2|\mathbf{n}_o = \mathbf{i}] = \min_{\hat{n} \in \mathbb{Z}^+} \sum_{j=0}^\infty \Pr(n = j|\mathbf{n}_o = \mathbf{i}) (\hat{n} - n)^2,
\]

(1)

where \( \mathbb{Z}^+ = \mathbb{Z}^+ \cup \{0\} \) is the set of nonnegative integers; \( \mathbf{i} = (i_1, \ldots, i_m)^\top \) is a vector of constants; \( n \) is the true number of occupants in the zone; and \( \Pr(n = j|\mathbf{n}_o = \mathbf{i}) \) is the conditional probability for the true number of occupants \( n = j \) when the vector of observations of the \( m \) systems \( \mathbf{n}_o = \mathbf{i} \). We take the Bayesian viewpoint that the observations are given and the true number of occupant is a random variable. Solving Eq. (1) is equivalent to solving the following problem

\[
\min_{\mathbf{n} \in \mathbb{Z}^+} J(\mathbf{n}, \hat{n}),
\]

(2)

where \( J(\mathbf{n}, \hat{n}) = \sum_{j=0}^\infty \Pr(\mathbf{n}_s = \mathbf{i}|n = j) \Pr(n = j) (\hat{n} - n)^2 \).

Note that calculating \( J(\mathbf{n}, \hat{n}) \) requires the a priori knowledge of the distribution of the true number of occupants \( \Pr(n = j) \) and the conditional joint distribution of the observations \( \Pr(\mathbf{n}_s = \mathbf{i}|n = j) \).

Ihler [24] showed that in some cases the random arrival and departure of the occupants to a zone can be reasonably approximated by a nonhomogeneous Poisson process, where the arrival rate \( \lambda_t \) is time-dependent. In this paper, we assume that the number of occupants in a zone at a discretized time \( t \) has the Poisson distribution with parameter \( \lambda_t \), i.e.,

\[
\Pr(n = j) = \frac{\lambda_t^j e^{-\lambda_t}}{j!},
\]

(3)

where \( \lambda_t \) is the average number of occupants at time \( t \). To simplify the notations, when there is no confusion, we simply use \( \lambda_t \) as \( \lambda \) in the rest of this paper.

We assume that Eq. (3) holds in the rest of this paper. For the case of uniform distribution, i.e., \( \Pr(n = j) = 1/M, 0 \leq j \leq M \), our methods are still applicable and the results are presented in Appendix A.

Note that it is difficult to obtain the joint conditional distribution.

\[
\Pr(\mathbf{n}_s = \mathbf{i}|n = j) \]

in practice because the \( m \) sensing systems may be installed for different purposes, and are not designed to share information. To resolve this issue, the conditional probability of each individual system \( \Pr(n_k = i_k|n = j) \) is approximated as follows. When there are \( j \) occupants, suppose an individual occupant is detected by system \( k \) with probability \( d_k \). Errors occur if some objects that are counted as occupants by system \( k \) may not be occupants. Since the total number of objects that would be counted as occupants cannot be accurately estimated a priori, we assume that the number of objects that are wrongly observed as occupants by system \( k \) has Poisson distribution with parameter \( \mu_k \), where \( \mu_k \) is the average number of objects that are wrongly detected as occupants by system \( k \). Then we have

\[
\Pr(n_k = i_k|n = j) = \sum_{l=0}^{\min\{i_k, j\}} \binom{j}{l} d_k^l (1 - d_k)^{j-l} \frac{\mu_k^{l-1}}{(l-1)!} e^{-\mu_k},
\]

(4)

where \( \binom{j}{l} d_k^l (1 - d_k)^{j-l} \) is the probability that the number of detected occupants is \( l \leq i_k, l \leq j \), and \( \frac{\mu_k^{l-1}}{(l-1)!} e^{-\mu_k} \) denotes the probability that the number of incorrectly observed occupants is \( i_k - l \).

The question is then that given Eqs. (3) and (4), how we can approximate the solution to the information fusion problem in Eq. (2). We present two approximate fusion methods in the next section.

### 4. Two approximate fusion methods

In this section, two approximate fusion methods for solving Eq. (2) are presented. In Method 1 we assume that the observation noises of different systems are independent. In Method 2 we assume that the solution to Eq. (2) has the same form as the estimation that minimizes the mean square error of an individual system, and the required parameters can be estimated based on the observed data.

#### 4.1. Method 1

Assume that the observations of different systems are independent. Note that \( J(\mathbf{n}, \hat{n}) \) is a convex function w.r.t. \( \hat{n} \in \mathbb{R}^+ \cup \{0\} \) (the proof is shown in Appendix B). Then for \( \hat{n} \in \mathbb{R}^+ \), we have
\[
\frac{dJ(i, \hat{n})}{dn} = 2\sum_{j=0}^{\infty} \Pr(n = i | n = j) \Pr(n = j | \hat{n} = j).
\]

By letting \(dJ(i, \hat{n})/dn = 0\), and with the assumption that the observations of the \(m\) sensing systems are independent, we have

\[
\hat{n}^*(i) \approx \hat{n}_k^*(i) = \sum_{j=0}^{\infty} \prod_{k=1}^{m} \Pr(n_k = i | n = j) \Pr(n = j).
\]  

Combining Eqs. (3)-(5), we then have

\[
\hat{n}_k^*(i) = \frac{A_1}{A_2},
\]

where

\[
A_1 = \sum_{l_k=0}^{l_{\text{max}}(k)} \frac{d_{l_k}^1}{(l_1 - l_k)!} \sum_{l_2=0}^{l_{k}} \frac{d_{l_2}^2}{(l_2 - l_k)!} \ldots \sum_{l_m=0}^{l_{k}} \frac{d_{l_m}^m}{(l_m - l_k)!} \times \prod_{j=0}^{\infty} \frac{\prod_{k=1}^{m} \left( \frac{j}{l_k} \right)^{(1 - d_k)^{-h}}}{j!}.
\]

\[
A_2 = \sum_{l_k=0}^{l_{\text{max}}(k)} \frac{d_{l_k}^1}{(l_1 - l_k)!} \sum_{l_2=0}^{l_{k}} \frac{d_{l_2}^2}{(l_2 - l_k)!} \ldots \sum_{l_m=0}^{l_{k}} \frac{d_{l_m}^m}{(l_m - l_k)!} \times \prod_{j=0}^{\infty} \frac{\prod_{k=1}^{m} \left( \frac{j}{l_k} \right)^{(1 - d_k)^{-h}}}{j!}.
\]

In the experimental and numerical results in Section 5, two sensing systems are considered, i.e., \(m = 2\). We approximate the infinite number of summations in \(A_1\) and \(A_2\) by \(M\) (adequately large) summations. \(A_1\) and \(A_2\) can then be approximated by,

\[
A_1 = \sum_{l_k=0}^{l_{\text{max}}(k)} \frac{d_{l_k}^1}{(l_1 - l_k)!} \sum_{l_2=0}^{l_{k}} \frac{d_{l_2}^2}{(l_2 - l_k)!} \ldots \sum_{l_m=0}^{l_{k}} \frac{d_{l_m}^m}{(l_m - l_k)!} \times \sum_{l_{k+1} \geq m}^{\infty} \frac{\prod_{k=1}^{m} \left( \frac{j}{l_k} \right)^{(1 - d_k)^{-h}}}{j!}.
\]

\[
A_2 = \sum_{l_k=0}^{l_{\text{max}}(k)} \frac{d_{l_k}^1}{(l_1 - l_k)!} \sum_{l_2=0}^{l_{k}} \frac{d_{l_2}^2}{(l_2 - l_k)!} \ldots \sum_{l_m=0}^{l_{k}} \frac{d_{l_m}^m}{(l_m - l_k)!} \times \sum_{l_{k+1} \geq m}^{\infty} \frac{\prod_{k=1}^{m} \left( \frac{j}{l_k} \right)^{(1 - d_k)^{-h}}}{j!}.
\]

4.2. Method 2

If there are correlations among the observations of different sensing systems, assume that the solution of the fusion problem has the same form as the estimation that minimizes the mean square error of an individual system. We first obtain the optimal estimation for a single sensing system, and then the optimal estimation of multiple sensing systems by selecting the best parameters.

First, for system \(k\), giving the observation \(n_k\), the optimal estimation that achieves the MMSE is

\[
\arg\min_{\hat{n}_k} \mathbb{E}[(\hat{n}_k - n)^2 | n_k = i_k],
\]

which is equivalent to

\[
\arg\min_{\hat{n}_k} J_k(i_k, \hat{n}_k),
\]

where

\[
J_k(i_k, \hat{n}_k) = \sum_{j=0}^{\infty} \Pr(n_k = i_k | n = j) \Pr(n = j) (\hat{n}_k - j)^2.
\]

Because \(J_k(i_k, \hat{n}_k)\) is convex w.r.t. \(\hat{n}_k\), by letting \(dJ_k(i_k, \hat{n}_k)/d\hat{n}_k = 0\), we then have

\[
\hat{n}_k(i_k) = \hat{i} (1 - d_k) + \frac{\hat{i} d_{i_k}}{\hat{d} + \hat{\mu}_k}.
\]

Assume that when there are multiple sensing systems, the optimal estimation \(\hat{n}^*\) has the similar form, i.e.,

\[
\hat{n}^*(i) \approx \hat{n}_g^*(i) = \hat{i} (1 - d) + \frac{\hat{i} d}{\hat{d} + \hat{\mu}}.
\]

where \(d, \mu\), and \(i\) need to be determined.

We have

\[
\mathbb{E}[n_k | n = j] = d_k j + \mu_k, \quad \text{Var}[n_k | n = j] = d_k j + \mu_k - d_k j.
\]
Thus we have
\[ \mu_k = E[n_k|n = j] - \sqrt{(E[n_k|n = j] - \text{Var}[n_k|n = j])^2,} \]
\[ d_k = \sqrt{(E[n_k|n = j] - \text{Var}[n_k|n = j])^2}. \]
We may use the observation \( i_k \) of system \( k \) to replace \( E[n_k|n = j] \) and obtain the following estimation of \( \mu_k \) and \( d_k \), i.e.,
\[ \hat{\mu}_k = i_k - \sqrt{(i_k - \text{Var}[n_k|n = j])}/j. \]
\[ \hat{d}_k = \sqrt{(i_k - \text{Var}[n_k|n = j])}/j. \]

When there are \( m \) sensing systems, \( i_k \) and \( v_k(j) \) depend on their observations and variances. If the observation noises among the \( m \) sensing systems are jointly Gaussian and have covariance \( \Sigma \), and the probability density function of the a priori distribution of the true \( n \) is also Gaussian, the optimal estimation for a single sensing system is \( \hat{n}_k(n_k) = g(n_k, \sigma_k^2) \), and that for the \( m \) systems is
\[ \hat{n}^* = g \left( \frac{\text{e}^{\Sigma^{-1}n_0}1}{\text{e}^{\Sigma^{-1}e}} \right), \]
where \( e = (1, \ldots, 1)^T \). \( g(\mu, \sigma^2) = \frac{1}{\sqrt{(2\pi)\sigma^2}} \int_{-\infty}^{\infty} p(j) \exp \left\{ -\frac{(j-\mu)^2}{2\sigma^2} \right\} dj \), and \( p(j) \) is the probability density function for the a priori distribution of the true value of \( n \).
We replace \( j \) in Eq. (7) by
\[ i = \frac{\text{e}^{\Sigma^{-1}n_0}}{\text{e}^{\Sigma^{-1}e}}, \]
and replace \( i_k \) and \( \text{Var}[n_k|n = j] \) in Eq. (9) by Eq. (10) and \( 1/(\text{e}^{\Sigma^{-1}e}) \) respectively and thus obtain the estimate of \( \mu \) and \( d \), as follows
\[ \hat{\mu} = \frac{\text{e}^{\Sigma^{-1}n_0}}{\text{e}^{\Sigma^{-1}e}} \hat{n} \]
and \( \hat{d} = \left( \frac{\text{e}^{\Sigma^{-1}n_0}}{\text{e}^{\Sigma^{-1}e}} - \frac{1}{\text{e}^{\Sigma^{-1}e}} \right)/i \).
\[ \hat{n} = \sum_{j=0}^{\infty} \text{Pr}[n = j] \hat{\mu}, \quad \text{and} \quad \hat{d} = \sum_{j=0}^{\infty} \text{Pr}[n = j] \hat{d}. \]

By combining Eqs. (7) and (11), \( \hat{n}_k(i) \) can then be calculated.

5. Experiments and performance analysis

In this section, two groups of experiments are shown. In the first group, a real system with RFID and video camera is considered. Physical experiments are used to show how the aforementioned two fusion methods improve the accuracy of the estimation of occupant level. In the second group, simulations are used to analyze how the performance of our methods are affected by occupant level and correlated observations.

5.1. Information fusion of RFID and video cameras

In this section, we apply methods 1 and 2 to fuse the information from RFID and video cameras in a room whose layout is shown in Fig. 1. According to the positions of the air outlets and lights, the room is divided into three zones. There are two occupant counting systems in the room, namely the active RFID and the video cameras. The active RFID system is composed of four readers and multiple tags. The positions of the readers are shown in the figure (denoted as solid nodes). Each occupant in the room wears a tag. When the occupant is in the room, the four readers receive the RSSI of the tag, estimate the position of the occupant, and then estimate the number of occupants in each zone. The second system is composed of two cameras, as shown in the figure (denoted as trapezoids), and the number of occupants is estimated in each zone. We focus on estimating the number of occupants in zone 3.

The experiments were conducted on Fridays from 9:15AM to 9:45AM for four consecutive weeks. The true occupant number and the estimations given by the RFID, the video cameras, and the two approximate fusion methods are given in Fig. 2. From Fig. 2 we can see that at each time moment, there are 7 or 8 occupants in the zone on the average. Both the observation of the RFID system and that of the video camera system are less than the true number, while both estimations of the two approximate fusion methods are closer.

For comparison, the linear least mean square error (LLMSE) method [3] is applied with the assumption of Gaussian and independent observations. Let \( X \) be the true occupant level, which is unknown. There are two sensing systems with observations \( Y = [y_1, y_2]^T \). Assume the observation noises of the RFID system and the video cameras have distribution \( N(0, \sigma^2_1) \) and \( N(0, \sigma^2_2) \), respectively. Suppose the observation equation is as follows,
The optimal estimation of $X$, denoted as $\hat{X}$, is

$$\hat{X} = (H^T R^{-1} H)^{-1} H^T R^{-1} Y,$$

where $R = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$. Based on historical data, the values of $H$ and $R$ are regressed as $H = \begin{bmatrix} 0.7923 \\ 0.5421 \end{bmatrix}$ and $R = \begin{bmatrix} 11.0333 & 0 \\ 0 & 1.5792 \end{bmatrix}$.

The standard deviation of the observation errors and the estimation errors of the RFID system, the video camera system, Methods 1 and 2, and the LLMSE method are shown in Fig. 3. We can see that on the average the RFID system counts two less occupants than the true number, while the video camera systems counts four less occupants. The LLMSE method counts two less occupants. Both fusion methods only count one occupant incorrectly. In Fig. 4 the error rate of the RFID and the video camera defined as

$$e_k = \frac{|\hat{n}_k - n|}{n},$$

(12)

is shown where $n$ and $\hat{n}_k$ are the true occupant number and the estimation of system $k$, respectively. In Fig. 4, the error rate is averaged over the same time moment over the four days. We also show the error rate of the estimations of the two approximate fusion methods and LLMSE method. From Fig. 4 we can see that the two approximation methods count only one occupant incorrectly, proving to be much more accurate than the RFID and video camera systems, as well as the LLMSE method.

In particular, from Fig. 3, we can see that the average standard deviation over the entire period of time, is 1.9130 for RFID, 3.9743 for video cameras, 1.3642 for LLMSE method, 1.0131 for Method 1, and 1.0872 for Method 2. Method 1 improves the...
accuracy of the RFID system by \((1.9130 - 1.0131)/1.9130 = 47.04\%\) and that of the video camera system by \((3.9743 - 1.0131)/3.9743 = 74.51\%\). Similarly we can calculate that Method 2 improves the accuracy of the RFID system and that of the video camera system by 43.17% and 72.64%, respectively. The LLMSE method only improves the accuracy of the RFID system and that of the video camera system by 28.69% and 65.67%, respectively.

The performance of fusion methods are related to the correlations of multiple sensing system. In this experiment, Methods 1 and 2 perform similarly, which implies that the correlation of the RFID system and video camera system is very small. In fact, based on the historical data we find that the correlation of the data of RFID and video camera systems is \(\rho = 0.21\). As will be shown in subsection 5.2, the reason is that the correlation \(\rho\) is close to \(\sigma_1/\sigma_2\), where \(\sigma_1 = 1.91\) and \(\sigma_2 = 3.97\) are the standard deviations of the two systems, respectively. We also find that the fusion results are slightly better than individual system estimation.

5.2. Performance analysis via simulation

In this subsection we evaluate the performance of Methods 1 and 2 by simulation under different occupant levels and correlated observations. We start the discussion from generating the correlated observations of the multiple systems.

5.2.1. Generating correlated observations of two sensing systems

Consider the case of two sensing systems. Their observations \(X\) and \(Y\) follow the probability density functions \(f_X(x)\) and \(f_Y(y)\), respectively. The correlation between the two systems is defined as \(\rho\). It is difficult to generate data with a specific value of \(\rho\) but it is much easier to generate data with different values of \(\rho\)’s by adding correlated noises. For this purpose, we adjust the seeds in pseudorandom number generations as shown below.

Define \(R_X\) and \(R_Y\) as the seeds of the pseudorandom number generators of the two systems, respectively, such that

\[
R_X = |x| R' + (1 - |x|) R'_X, \tag{13}
\]

\[
R_Y = \begin{cases} 
|x| R' + (1 - |x|) R'_Y & x \geq 0 \\
|x|(1 - R') + (1 - |x|) R'_Y & x < 0 
\end{cases}, \tag{14}
\]

where \(-1 \leq x \leq 1\) and \(R' \sim U(0, 1)\), \(R_X \sim U(0, 1)\), \(R'_X \sim U(0, 1)\). In Eqs. (13) and (14), \(R'\) is the common seed for \(R_X\) and \(R_Y\), and \(R_X\) and \(R_Y\) are the private seeds. \(R_X\) and \(R_Y\) are positively correlated when \(x \geq 0\) and negatively correlated otherwise.

Define \(U_X = F(R_X)\), \(U_Y = F(R_Y)\), where \(F(R_X)\) and \(F(R_Y)\) as the cumulative distribution functions (CDFs) of \(R_X\) and \(R_Y\), respectively. Denote \(F_X(X)\) and \(F_Y(Y)\) as the CDFs of \(X\) and \(Y\) respectively. Then we have,

\[
X = \{x|F_X(X = x) = U_X\}, \quad Y = \{y|F_Y(Y = y) = U_Y\}.
\]

Parameter \(x\) is the weight of the common seeds for generating the correlated \(X\) and \(Y\). By adjusting \(x\), we can obtain the pseudorandom numbers of \(X\) and \(Y\) with different correlations. When \(x\) approaches \(-1\), \(\rho_{X,Y}\) approaches \(-1\), and when \(x\) approaches \(1\), \(\rho_{X,Y}\) approaches \(1\). With the Eq. (3) and Eq. (4), a series of correlated pseudorandom numbers from the observations of sensing systems under a special \(j\) can be easily generated.

As an example, Fig. 5 shows the probability mass functions from the observations of two systems that are generated by simulation. We use the following parameter: \(j = 20\), \(d_1 = 0.9\), \(\mu_1 = 0.1\), \(d_2 = 0.8\), \(\mu_2 = 0.2\) and \(x = -0.5\). The correlation of the generated observations is \(-0.52\).
5.2.2. Fusion method performance under different occupant levels

Consider the RFID system and the video system with $d_1 = 0.9$, $\mu_1 = 0.1$, $d_2 = 0.8$, $\mu_2 = 0.2$ and $a = -0.5$. We generate data for $k = 10, 15, 20, 25, 30, 35, 40$ and $a = 0.5$.

Figs. 6 and 7 show the observations of RFID and video cameras and the estimations of RFID, video cameras, and methods 1 and 2, versus the average numbers of occupants $k$. We can see that the error rates of methods 1 and 2 decrease as $k$ increases, which implies that the two methods become more accurate when the occupant level is large. This can also be concluded theoretically.

Combining Eqs. (8) and (12), we have

$$
\epsilon_k = \frac{|\hat{n}_k - n|}{n} \propto \sqrt{\text{Var}[n_k|n = \lambda] \lambda} = \sqrt{d_k \lambda + u_k - d_k \lambda} \propto \frac{O(\lambda)}{O(\lambda)},
$$

where $\propto$ means proportional. So we have

$$
\lim_{\lambda \to \infty} \epsilon_k = 0.
$$

5.2.3. Fusion method performance under different correlation levels of two sensing systems

We generate data for $\lambda = 20$ and $x = -0.9, -0.5, -0.3, 0.3, 0.5, 0.7, 0.9, 0.9999$ with the correlation $\rho = -0.92, -0.55, -0.11, 0.15, 0.47, 0.76, 0.94, 0.97$, respectively. The performance of the aforementioned various methods are shown in Figs. 8 and 9.
We can see that Method 1 performs well except when $q$ approaches 1. Method 2 performs well for all cases, and outperforms Method 1 significantly when $j/q$ approximates 1. Methods 1 and 2 perform similarly around $q = 0$. Method 2 does not perform well around $q = r_{RFID}/r_{Video}$, and its performance is similar to the best performance of the individual RFID and video camera. We will analyze this observation in the next section.

6. Discussions

In this section, we first show in theory why fusing the observations of multiple systems can indeed achieve higher estimation accuracy. Second, we discuss how the correlation among multiple systems helps improve estimation accuracy.
6.1. The advantage of information fusion

Let \( \mathbf{n}_0(k) = (n_1, \ldots, n_k)^T \) and \( \mathbf{i}(k) = (i_1, \ldots, i_k)^T \). Giving \( \mathbf{n}_0(k) = \mathbf{i}(k) \), the MMSE is

\[
\min_{\mathbf{n}(k) \in \mathbb{R}^d} E \left[ (\hat{n}(k) - n)^2 | \mathbf{n}_0(k) = \mathbf{i}(k) \right].
\]

Denote the corresponding optimal estimate as \( \hat{n}(\mathbf{i}(k)) \). Then

\[
J(\hat{n}(\mathbf{i}(k))) = J(\mathbf{i}(k), \hat{n}(\mathbf{i}(k))) = \min_{\mathbf{n}(k) \in \mathbb{R}^d} J(\mathbf{i}(k), \mathbf{n}(k)).
\]

If the system \((k+1)\) has the observation \(n_{k+1} = i_{k+1}\), after fusing this piece of information, the MMSE is

\[
\min_{\mathbf{n}(k+1) \in \mathbb{R}^d} E \left[ (\hat{n}(k + 1) - n)^2 | \mathbf{n}_0(k + 1) = \mathbf{i}(k + 1) \right].
\]

Denote the corresponding optimal estimate as \( \hat{n}(\mathbf{i}(k + 1)) \). Similarly we can define \( J'(\hat{n}(k + 1)) \). The following theorem shows that fusing system \((k + 1)\) helps improve the counting accuracy.

Theorem 1.

\[
\sum_{k_1 = 0}^{\infty} \min_{\mathbf{n}(k_1) \in \mathbb{R}^d} E \left[ (\hat{n}(k + 1) - n)^2 | \mathbf{n}_0(k + 1) = \mathbf{i}(k + 1) \right] \times \Pr(n_{k+1} = i_{k+1} | \mathbf{n}_0(k) = \mathbf{i}(k))
\leq \min_{\mathbf{n}(k) \in \mathbb{R}^d} E \left[ (\hat{n}(k) - n)^2 | \mathbf{n}_0(k) = \mathbf{i}(k) \right].
\]

where the left hand side (LHS) is the MMSE after fusing system \((k+1)\), which is averaged over all possible value of \(n_{k+1}\), and the right hand side (RHS) is the MMSE without fusing system \((k+1)\). \(\square\)

Proof. Based on the definition of conditional probability, we have

\[
E \left[ (\hat{n}(k + 1) - n)^2 | \mathbf{n}_0(k + 1) = \mathbf{i}(k + 1) \right] \times \Pr(n_{k+1} = i_{k+1} | \mathbf{n}_0(k) = \mathbf{i}(k)) = \frac{f(\mathbf{i}(k + 1), \hat{n}(k + 1))}{\Pr(\mathbf{n}_0(k) = \mathbf{i}(k))}.
\]

We also have

\[
E \left[ (\hat{n}(k) - n)^2 | \mathbf{n}_0(k) = \mathbf{i}(k) \right] = \frac{J(\mathbf{i}(k), \hat{n}(\mathbf{i}(k)))}{\Pr(\mathbf{n}_0(k) = \mathbf{i}(k))}.
\]

Note that

\[
\sum_{k_1 = 0}^{\infty} J'(\mathbf{i}(k + 1)) \leq \sum_{k_1 = 0}^{\infty} \sum_{n = 0}^{\infty} \Pr(\mathbf{n}_0(k + 1) = \mathbf{i}(k + 1) | n = j) \times \Pr(n = j) \times (\hat{n}(k) - n)^2 = J'(\mathbf{i}(k)),
\]

where the inequality is due to the optimality of \( J'(\mathbf{i}(k + 1)) \). Combining Eqs. 16, 17, we have Eq. (15). \(\square\)

Note that Theorem 1 holds regardless of the prior distribution \(\Pr(n = j)\) and the probability distribution of the observation noise. When \(\mathbf{n}_0 \in \mathbb{R}^n\) and \(\mathbf{i} \in \mathbb{R}^n\), we can similarly show that

Corollary 1.

\[
\sum_{k_1 = 0}^{\infty} \min_{\mathbf{n}(k_1) \in \mathbb{R}^d} E \left[ (\hat{n}(k + 1) - n)^2 | \mathbf{n}_0(k + 1) = \mathbf{i}(k + 1) \right] \times \Pr(n_{k+1} = i_{k+1} | \mathbf{n}_0(k) = \mathbf{i}(k))
\leq \min_{\mathbf{n}(k) \in \mathbb{R}^d} E \left[ (\hat{n}(k) - n)^2 | \mathbf{n}_0(k) = \mathbf{i}(k) \right],
\]

where \(\Pr(n_{k+1} = i_{k+1} | \mathbf{n}_0(k) = \mathbf{i}(k))\) is the conditional probability density function.

Theorem 1 and Corollary 1 show that information fusion generally improves estimation accuracy and does not decrease the accuracy.

6.2. Impact of correlations

In order to see how the correlation among multiple sensing systems helps improve the accuracy, we focus on the fusion method for the two systems. Suppose the observation noise of system \( k \) follows Gaussian distribution \( N(\mathbf{n}_k, \sigma_k^2) \). Without loss of generality, assume \( \sigma_1 \geq \sigma_2 \geq 0 \). The correlation between the two systems is
Under the Gaussian assumption, assume \( q(i) \) is independent from \( i \). So we simply write as \( q \). Assume that 
\[
\Pr(n_j = j \mid f_g) = \frac{1}{2M}
\]
for all \(-M \leq j \leq M\), where \( M > 0 \) is a very large constant. Then we can show that the RHS and the LHS of Eq. (18) are

\[
\text{RHS} = \int_{-\infty}^{+\infty} \frac{J_k(i_k)}{\Pr(n_k = i_k)} \, di_1,
\]
and

\[
\text{LHS} = \int_{-\infty}^{+\infty} J'(i) \Pr(n_1 = i_1) \, di_1,
\]
where \( \Pr(n_1 = x) \) is the probability density function of \( n_1 \). After dropping the common factors in \( J_k(i_k) \) and \( J'(i) \), we have

\[
\bar{J}_k(i_k) = \sigma^2_k, J'(i) = \frac{\sigma^2_1 \sigma^2_2 (1 - \rho^2)}{\sqrt{2\pi(\sigma^2_1 - 2\rho \sigma_1 \sigma_2 + \sigma^2_2)^3}} e^{\left\{ \frac{-(n_1 - \rho n_2)^2}{2(\sigma^2_1 - 2\rho \sigma_1 \sigma_2 + \sigma^2_2)} \right\}}
\]
and

\[
\theta = \int_{-\infty}^{+\infty} J'(i) \, di_1 = \int_{-\infty}^{+\infty} J'(i) \, di_2 = \frac{\sigma^2_1 \sigma^2_2 (1 - \rho^2)}{\sigma^2_1 - 2\rho \sigma_1 \sigma_2 + \sigma^2_2}.
\]

From Fig. 10, we find that the correlation of multiple sensing systems affects the performance of fusion, and that the performance of the fusion method is worst at the point \( \rho = \sigma_2 / \sigma_1 (\sigma_2 \leq \sigma_1) \), better at other points, and best at \( |\rho| = 1 \) (\( \sigma_1 \neq \sigma_2 \)). This gives us a theoretical justification for the results in the last section.

7. Conclusion

In this paper we formulated the occupant level estimation problem in an indoor environment as an information fusion problem with RFID signal and the video camera image as sensing sources. Two fusion methods are developed with independent and correlated measurements respectively. Numerical experiments show that both methods substantially improve the estimation accuracy in comparison with the individual measurement sources, and outperform the LLSME method. The estimation results are also obtained and shown under different occupant levels and correlated observations. Both methods perform better when the true occupant level increases and the measurements of multiple sources are negatively correlated.

Incorporating the human movement model and adjusting the sensor configuration may further improve the accuracy of the estimation [31]. In this case the fusion problem becomes a multistage decision making problem. It is our future work to further improve the estimation of zonal occupant levels in buildings by solving dynamic decision making problems for discrete event systems [12,43] and for hybrid systems [32].
Appendix A. Fusion methods for uniformly distributed occupants

Assume the occupant level has a uniform distribution $U[0,M], M > 0$, i.e., $\Pr\{n = j\} = 1/M, j \in [0,M]$. The two methods can be described as follows.

A.1. Fusion Method 1

\[ n_1^*(i) = \frac{A_1}{A_2}, \]

where

\[ A_1 = \sum_{i_1=0}^{\min(M,i)} \sum_{i_2=0}^{\min(M,i)} \ldots \sum_{i_M=0}^{\min(M,i)} \prod_{j=1}^{M} \left( \frac{j}{d_j} \right)^{i_j} (1 - d_j)^{i_j}, \]

\[ A_2 = \sum_{i_1=0}^{\min(M,i)} \sum_{i_2=0}^{\min(M,i)} \ldots \sum_{i_M=0}^{\min(M,i)} \prod_{j=1}^{M} \left( \frac{j}{d_j} \right)^{i_j} (1 - d_j)^{i_j}. \]

A.2. Fusion Method 2

\[ n_2^*(i) \approx n_0^*(i) = \frac{B_1}{B_2}, \]

where

\[ B_1 = \sum_{i_1=0}^{\min(M,i)} \frac{d_1}{i_1!} \sum_{i_2=0}^{\min(M,i)} \frac{d_2}{i_2!} \ldots \sum_{i_M=0}^{\min(M,i)} \frac{d_M}{i_M!} \prod_{j=1}^{M} \left( \frac{j}{d_j} \right)^{i_j} (1 - d_j)^{i_j}, \]

\[ B_2 = \sum_{i_1=0}^{\min(M,i)} \frac{d_1}{i_1!} \sum_{i_2=0}^{\min(M,i)} \frac{d_2}{i_2!} \ldots \sum_{i_M=0}^{\min(M,i)} \frac{d_M}{i_M!} \prod_{j=1}^{M} \left( \frac{j}{d_j} \right)^{i_j} (1 - d_j)^{i_j}. \]

$\mathbf{d}$, $\mu$, and $i$ need to be estimated.

Appendix B. The convexity of $J(i, n)$ and $J_k(i_k, n_k)$

\[ \frac{1}{2} J(i, n_1) + J(i, n_2) - J(i, \frac{n_1 + n_2}{2}) = \frac{1}{4} \sum_{j=0}^{\infty} \Pr(n_j = i|n = j) \Pr(n = j) (n_1 - n_2)^2 \geq 0. \]

Similarly, we can show that $J_k(i_k, n_k)$ is also a convex function w.r.t. $n_k$.

References


