Report

Assumptions:

Model will assume that people in this building will walk in a soft and peaceful pace such that within each timestamp interval (15 seconds), they would only go to the neighbours of their current location, or stay unchanged.
The neighbours of a room *r* are considered as those rooms which are directly connected to *r*. For example, ‘r5’ and ‘c3’ are the neighbours of room ‘r6’.

The neighbours of corridors were reduced manually to control the size of transition matrix and it won’t affect too much performance because we don’t have control of the lights in corridors. The detailed information would be explained in next section.

Algorithm explanation:

Modelization:

The core algorithm used to implement this approach is by using Probabilistic Graphical Model, Hidden Markov Model and Markov Chain.

For the room with sensors, they are modelled as Hidden Markov Chain, the information of occupancy of those room is the hidden state, and the information obtained from sensors is the observation of emissions.

For the room without a sensor, they are just normal Markov Chains.

Data Normalization

The information obtained from reliable and unreliable sensor will be mapped into a Boolean value as following: ‘Motion’ = True, ‘No Motion’ = False, indicating whether room is empty or not.

The information obtained from door sensor will be mapped into a Boolean value as following: Data > 0 = True, data == 0 = False, indicating whether someone passed door or not.

The information obtained from robot and the ground truth training data will be normalized as the same way as the door sensor data.

Training:

There are two core parameters need to be established from training, transition matrices, and emission probability table.

The establishment of emission table is straightforward. Knowing that sensors sometimes may fail, by counting the empirical frequency of the instance when there is someone in the room with sensor, or when someone pass through doors, the emission probability of sensors will be derived.

The transition matrix of a room is the state of that room at time *t* transited from time *t-1* by its neighbours and itself. To calculate the transition matrix of a room *r* at time *t*, we treat the state of room *r* at time *t* as the transited output. The transition starting point are the movements of its neighbours (including itself) at time *t-1*. Finally, the probability table will be obtained by counting the empirical frequency of cases.

Initialization:

The initial state is set as all rooms are empty except the 'outside' area. The probability to represent the extreme cases are deliberately set as 0.99 or 0.01 to maintain the consistency of future's multiplication.

Prediction:

Before making predictions, the data obtained from sensors needs to be normalized into the same format as training. Considering sensors may fail at any time, the excepted cases need to be handled as well.

For those room without sensor, the prediction of the state of the room at *t* can be calculated from the formula: B(x_t) = P(x_t|x_t-1) * P(x_t-1). Which can be implemented by the joining of transition matrix with state at *t-1*, corresponding to the miniforward() function.

For the rooms with sensor, the prediction of the state of the room at *t+1* can be calculated from the formula: B(x_t+1) = P(x_t+1|e_1:t+1). This requires two steps. First, we calculate B(x_t), and then introduce the observed emission evidence. The involvement of emission evidence corresponds to forwardOnline() function.

Finally, to make a decision that whether to open lights or not at time *t+1* is just equivalent to answer an mpe query of the Markov Chain at time *t+1*.

Method Justification

It is straightforward to consider the prediction as a probabilistic problem and Bayesian Network is the model dealing with such problems. At the beginning, a static Bayesian Network was experimented. Office was divided into small areas and each area is represented by a static Bayesian Network. The room equipped with sensors are the parents of their neighbours which do not have one, and sensors are the parents of all those equipped rooms. To make a prediction that whether to open lights or not, is required to calculate the concrete probability of all the leaf nodes given the sensors' data. Although it is easy to realize that there is relationship between each rooms and sensors, the relationship is quite obscure and unreliable. Besides, the training data is relatively insufficient. To enforce a study based on this data, overfitting might be invertible.

With considering the autonomous of people's movement and the passage of time, a dynamic Bayesian Network, Markov Chain was finally decided. Although there are approximately 40 rooms, and each room represents a Markov Chain, with the assumptions proposed above, the size of the transition matrix of each chain won't be too large.

Time Complexity Measurement

The major time consuming part of the prediction approach is from the forwardOnline() function. The forwardOnline() function follows the Forward Algorithm which is $O(n|X|^2)$.