UAV using Dec-POMDP model for increasing the level of security in the company

Umit Ilhan a, Latafat Gardashova b *, Kenan Kilic c

 a Department of Computer Engineering, Qafqaz University, Baku, Azerbaijan
 b Department of Computer Engineering, Azerbaijan State Oil and Industry University Baku, Azerbaijan
 c Department of Computer Engineering, Qafqaz University, Baku, Azerbaijan

Abstract

One of the most important jobs for every company has always been keeping a high level of security. Various methods of information systems are being applied to ensure and increase the level of security. Unmanned aerial vehicles (UAV) that spread very rapidly in recent years are being applied in various fields. Autonomously controlled UAVs can fulfill almost any job. Markov decision processes on the other hand, play significant role among algorithms that deal with decision-making problems. This article proposes model that uses UAVs and can be used to support and improve information systems security level of a company. The most significant property of drones used in proposed model is that they do their job by directly connecting and sending information to each other. To get the best result decentralized partially observable Markov decision process (Dec-POMDP) was used. To gauge the level of security, calculations of the data were shown with fuzzy data set. In the end, details of the model and proposals are given.

Keywords: Fuzzy Logic; DEC-POMDP; Markov decision processes; unmanned aerial vehicle; information systems; decision making under uncertainty.

1. Introduction

Unmanned aerial vehicles (UAVs), in the light of technological developments of recent years, has increasingly become popular in the fields of academic studies and engineering applications. UAV, in many cases allows fast and safe solutions or analysis to be performed, particularly in military applications, natural disasters, monitoring of...
various sports activities, traffic control, illegal construction, etc. In addition, both companies and ordinary people use UAVs to produce different solutions to many problems in daily life.

Today, many security systems are used to protect some area. Techniques used vary according to the size of the area. In large areas, it is harder to provide security. When we say land security, we understand protection from all hazards that may be in the field, and ability to control the desired location at the desired time. Different solutions are used to solve these problems. Placing security cameras, checking certain areas at certain times are some examples of solutions. However, these solutions do not provide full-time protection of area. Placing lots of cameras in large areas requires very large budgets. Even in areas where camera systems are installed, due to the limitations of the system the risk of blind spots is always high. And solutions that use man power to provide security require lots of resources in large areas as well. For this reason, new solutions with better options are necessary.

In this article, MDP was used as decision-making algorithm. In this solution, multiple UAVs are used. These UAVs can send information to each other. As UAVs states and observations are not certain and partly observable, Dec-POMDP algorithm is used in the model to eliminate this problem. Security level is set according to the results of the observations from the UAVs. Fuzzy logic is used in calculation when retrieving results of these observations. As the values obtained are not final results, the level of security is determined from the fuzzy logic calculations mentioned in the article1-4.

2. Decision-making algorithms

Decision-making is one of the most important and desirable skills of agent. The agent takes input from the environment, makes a decision and fulfils his decision. Decision-making is choosing an action from a pre-determined set of actions established by designer. Agent’s work depends on the action chosen. According to the current state, decision-making algorithm determines the strategy for the choice of the optimal action 5.

2.1. Partially observable Markov decision process

POMDPs provide partially observable area with changeable decision-making system. Partially observable environment means that the agent does not have the right to directly access his own state, and the situation should be solved according to observations. POMDP is well adapted to the area where decision-making is done according to sensor with the participation of the robot. POMDP is natural partially observed model for MDP model. Formally, we define POMDP as multilateral. S, A, T, R, $\gamma$ are derivatives from the MDP model, where S is status group, A is actions group, T is conversion function, R is awards function, $\gamma$ is the discount factor. 6-10. We look at MDP as the MDP highlight of POMDP. The remaining terms are marked as Z and O 13, 15.

2.1.1. Observations

$Z = \{z_0, z_1, ..., z_L\}$ is group of agent’s all possible observations of the environment. Observation is a part of information about the environment. Only $Z_t$ observation may appear at the time step t, but the group of observations $Z$ includes multiple observations flow. $Z$ can be determined by continuous observations, but to do this we have to examine the measurable sustainable environment.

2.1.2. The observation function

We denote this function by O. It defines the probability of observation of action a in state s and is expressed as

$$O(s, a, z) = P(z_{t+1} = z|s_t = s, a_t = a) \forall t$$  (1)

2.1.3. The history and state of belief

Agent in partially observable environment can not enter the current state S directly. However, the agent receives observations that provide direct access to the state at every step. Overall, these observations are not sufficient to guarantee knowledge of the state.

The most obvious way to keep track of the state of the agent is to keep the history of actions and observations of agent. We denote the history with $h_t$. The history of observations is expressed as

$$h_t = \{a_0, z_1, a_1, z_2, ..., a_{t-1}, z_t\}$$  (2)
However, it is not practical to collect the history for some of the tasks. Instead of determining the history, we can note the belief state $B_t$, which is the probability of the division on the states at the given history. The division state is favorable statistics for history. $B_0$ denotes the initial belief state.

The advantage of this is that it allows shortening of the agent history. When using the belief, you do not need to keep the history. Let $S$ denote possible belief groups. It should be noted that for the majority of states $s$ is simple. Therefore, it has size and it is continuous. If the previous state of belief is known its simple calculation gives another advantage. We accept $T$ as function of conversion from old belief state $T_1$ to new belief state.

\[
b_t(s) = P(s_t = s| h_t, b_0) \tag{3}\]

where $b_0$ is the initial state of belief. As we can see from formula

\[
b_t(s') = \tau(b_{t-1}, a_{t-1}, z_t) = \frac{1}{p(b_{t-1}, a_{t-1})} O(s', a_{t-1}, z_t) \sum_{s \in S} T(s, a_{t-1}, s') b_{t-1}(s) \tag{4}\]

the next belief state at $s'$ is simply based on the previous belief, and observation for belief $S'$ is probability normalized by $Z_t$ that is selected by $z$ probability and accepts probability $P(b_{t-1}, a_{t-1})$. This factor is calculated expressed as follows:

\[
P(b_{t-1}, a_{t-1}) = \sum_{s \in S} O(s', a_{t-1}, z_t) \sum_{s \in S} T(s, a_{t-1}, s') b_{t-1}(s) \tag{5}\]

This update process of belief is similar to Bayes filter.

2.1.4. Behaviours and value functions

As in the case of MDP, we can note the behavior of the agent that determines the process in POMDP environment. It should be noted that it is same as for the MDP, but instead of choosing action relying on the state we choose belief state. For behaviour POMDP acts as state in belief MDP state. In fact, POMDP itself is featured by belief MDP. As conversion function is favorable statistics, we can establish MDP belief. Where, set of states POMDP is from belief set, action set $A$ is the same as it was in POMDP, the conversion function is same as belief conversion function, and the award is the same as it was in the MDP.

2.2. Decentralized Partially Observable Markov Decision Process

Decentralized Partially Observable Markov Decision Process (Dec-POMDP) is an extension of POMDP for solution with multiple agents. When decentralized agents act in order to achieve a few points, this problem can be modeled as Dec-POMDP. At each step of time, every agent makes action, accepts local observation and close collaboration award. Local behavior for each representative makes maps from observations order to the actions 11, 12, 15, 16. In this article, designing defined by Bernstein was used 17. Dec-POMDP model consists of 7-tuple $(n; S; A; T; \Omega; Obs; R)$: $n$ is the number of agents; $S$ is a finite set of states; $A$ is Cartesian product of $A_i (i = 1; 2 \ldots; n)$set of joint actions. In other words, it is a set of possible actions of agent, $T$ is the transition function of state that determines possible next state probabilities according to $S$ current situation and $a$ joint action; $\Omega$ is Cartesian product of $\Omega_i (i = 1; 2 \ldots; n)$ set of observations. In other words, it is a set of possible actions of agent, $S$ at any time step agents accept joint observation $o = (o_1; o_2; \ldots; o_n)$ from the environment; $Obs$ is observation function, that specifies probability that accepts $s$ current state and $o$ joint observation given to $a$ current joint action; $R$ is instant award function that is gained by the team of multiple agents with given current state and current action.

3. Preliminaries

Definition 18. A fuzzy set $A$ is defined on a universe $X$ may be given as:

\[
A = \{(x, \mu_A(x)) | x \in X\}
\]

where $\mu_A : X \rightarrow [0,1]$ is the membership function. A membership value $\mu_A(x)$ describes the degree of belongingness of $x \in X$ in $A$. 

**Definition 2**: The trapezoid membership function is defined as

\[
\mu_{\text{Tra}}(x, \alpha, \beta, \gamma, \delta) = \begin{cases} 
0, & x < \alpha \\
\frac{x - \alpha}{\beta - \alpha}, & \alpha \leq x < \beta \\
\frac{\gamma - x}{\delta - \gamma}, & \gamma \leq x < \delta \\
0, & x \geq \delta
\end{cases}
\]

**Definition 3**: The operation of fuzzy equality is widely used to calculate the truth-value of fuzzy rules in expert systems and fuzzy control systems: \(a \approx b\),

where \(a\) and \(b\) are linguistic values; \(\approx\) denotes the operation «is close to». This operation is defined as a possibility measure for \(a\) to have the same value as \(b\).

Calculate \(\text{Poss}(a/b)\) if \(a\) and \(b\) are trapezoidal fuzzy numbers (Fig. 1, a, b):

\[
\mu_a(x) = \begin{cases} 
1 - \frac{a_1 - x}{\alpha_1}, & \text{if } a_1 - \alpha_1 \leq x \leq a_1 \\
1, & \text{if } a_1 \leq x \leq a_2 \\
1 - \frac{x - a_2}{\alpha_r}, & \text{if } a_2 \leq x \leq a_2 + \alpha_r \\
0, & \text{otherwise}
\end{cases}
\]

\[
\mu_b(x) = \begin{cases} 
1 - \frac{b_1 - x}{\beta_i}, & \text{if } b_1 - \beta_i \leq x \leq b_i \\
1, & \text{if } b_1 \leq x \leq b_2 \\
1 - \frac{x - b_2}{\beta_r}, & \text{if } b_2 \leq x \leq b_2 + \beta_r \\
0, & \text{otherwise}
\end{cases}
\]

\[
\text{Poss}(a/b) = \max \min(\mu_a(x), \mu_b(x)) = \begin{cases} 
1 - \frac{a_1 - b_2}{\alpha_r + \beta_r}, & \text{if } 0 < a_1 - b_2 < \alpha_1 + \beta_1 \\
1, & \text{if } \max(\alpha_1, \beta_1) \leq \min(a_2, b_2) \\
1 - \frac{b_1 - a_2}{\beta_r + \alpha_r}, & \text{if } 0 < b_1 - a_2 < \beta_r + \alpha_r \\
0, & \text{otherwise}
\end{cases}
\]

4. **Proposed model**

In this article, we developed (Decentralized Partially Observable Markov Decision Process, POMDP) partially observable model with variable security problems as a solution to problem of inability to provide full information system security in the large area companies. Details on the model are described in the following sections.

4.1. **Model Parameters**:

\(T = \) number of periods of states used in finite time; \(n = \) total number of UAVs; \(C = \) situations with the security
deterioration; \( M \) = maximum safety level demand \( 0 \leq M \leq C \); \( \pi_t \) = distribution of security demands in the period \( t \), in the form of \( \pi_t(\lambda_t) = \{ \pi_t(\lambda_t) \} \), where \( \lambda_t = k \) average demand expression, refers to the state \( \pi_t(\lambda_t) = f(\tilde{d}t = i; \lambda_t = k) \), so that \( i = 0, \ldots, M \); \( l_t \) = state of the safety level in normal times; \( \tilde{d}t \) = security level demand \( 0 \leq \tilde{d}t \leq M \) in period \( t \) caused by \( \pi_t \) state, \( \tilde{d}t \) calculate by using possibility measure, which is given definition 3; \( a_t \) = company’s desired security level in period \( t \).

4.2. Model Information:

Security system has at least 2 UAVs. System security will be achieved through camera system, image processing and some sensors. Due to the UAV decision-making system UAVs may make decisions by themselves. All UAVs can send information to each other with UAV decision-making system. The system will operate in an integrated manner with the company’s security system. The safety level of the system is divided into three. If we make the assumption, based on the criteria given by the company, we assess the corruption of security level out of 100. Between 0-30 system security level will be intact. Between 30-80 security level will be a less disturbed, and between 80-100 security level will be more disturb.

- If the level of security is intact, between 0-30, UAV will continue to fulfill the tasks assigned without any new decisions.
- If the level of security is less disturbed, between 30-80, UAV will inform the central security system about criteria that deteriorated security level. UAV will continue to fulfill the tasks assigned unless told otherwise.
- If the level of security is more disturbed, between 80-100, UAV will leave all given tasks and will focus on criteria that brought down the level of security. In this case, all other UAVs will be informed and they will leave their tasks and will follow this problem.
- In this paper, a concrete models using fuzzy data are considered. Fuzzy logic is effective tool to describe uncertainty. Strengths of fuzzy logic are\(^{20} \): interpretability, transparency, plausibility, graduality, modelling, reasoning, tolerance to imprecision.
- It is known, that the advantages of the fuzzy Method are: a) It is intuitive; b) It has widespread acceptance; c) It is simple.
- However, the reasons are the need in precise input information and also a loss of information in defuzzification process. From this viewpoint possibility measure is more effective\(^{20} \).
- Let us describe fuzzy parameters of model:

\[
\begin{align*}
\mu_{\text{undisturb}}(x, \alpha, \beta, \gamma, \delta) &= \begin{cases} 
0, x \leq 0 \\
1, 0 \leq x < 30 \\
\frac{100 - x}{70}, 30 \leq x < 100 \\
0, x \geq 100
\end{cases} \\
\mu_{\text{lessdisturb}}(x, \alpha, \beta, \gamma, \delta) &= \begin{cases} 
0, x \leq 0 \\
\frac{30 - x}{30}, 0 \leq x < 30 \\
1, 30 \leq x \leq 80 \\
\frac{100 - x}{20}, 80 < x < 100 \\
0, x \geq 100
\end{cases} \\
\mu_{\text{moredisturb}}(x, \alpha, \beta, \gamma, \delta) &= \begin{cases} 
0, x \leq 0 \\
\frac{80 - x}{80}, 0 \leq x < 80 \\
1, 80 \leq x \leq 100 \\
0, x > 100
\end{cases}
\end{align*}
\]

The values received in the security level calculations will be calculated with fuzzy logic and degree of deterioration of the security level will be decided. For example, if values received are close to the less disturb level, fuzzy logic will decide if the level of safety is disturb or not.

- In period \( t \) security system has security level \( l_{t-1} \), the security level from the previous period. Monitors the current safety requirements \( \tilde{d}t \). This observed demand will determine the status of the security level. \( \tilde{d}t \) is observed depending on the probability distribution \( \pi_t \) derived from the current average level of safety \( \lambda_t \). According to the probabilities based on observations sent to the system by UAV, two cases appear: If \( l_{t-1} \geq \tilde{d}t \),
UAV decision-making system calculates that according to the observations systems security status is not disturb. UAV-1 receives information that \( \Delta t \) security level is not disturb and UAV-1 returns to fulfill previously given tasks.

- If \( l_{t+1} < \Delta t \), security system calculates that according to the observations systems security status is disturb. System sends the information about disturb security level to the UAV-1, and UAV-1 starts following the problem until given different task. Other UAVs (UAV-2, UAV-3, ... ..., UAV-N) also get involved.

According to the safety problems coming from UAV-1, UAV decision-making system has to determine new tasks for other UAVs. Meanwhile, we assume that areas of other UAVs are totally secure. As tasks to be given to other UAVs by UAV decision-making system are limited to the \( 0 \leq l_t \leq C, t = 0,1,...,T \), the number of tasks that will be given should not exceed the number of other UAVs, which means expression should be as following \( 0 \leq a_t \leq C - \max(0,l_{t+1} - \Delta t), 0,1,...,T \).

4.3. Components of proposed in this article Dec-POMDP model

4.3.1. State: State of the system in period \( t \), vector containing average tasks of the next period, requests generated in the current period and the initial state of the UAV decision-making system; \( s = (\lambda_{t+1}, \Delta t, l_{t+1}) \). In the system, task level \( l_{t+1} \) of the state and the tasks of the current state \( \Delta t \) are fully observable. While, average tasks of the next state \( \lambda_{t+1} \) are partially observable.

Action: As the system includes multiple decision makers, UAV decision-making system will decide on new tasks at.

Observation: UAV decision-making system’s thoughts on average tasks of the next time period, present time tasks and initial task observations of present time; \( o = (\Omega_{t+1}, \Delta t, l_{t+1}) \)

Transition Function:

\[
P(s'\mid s, a_t) = (\lambda_{t+1}, \Delta t, l_{t+1}), S' = (\lambda_{t+2}, \Delta t, l_{t+1}) t=0,1,...,T
\]

Issues to be considered when constituting the transition function; Calculation of transition probabilities to next new status \( s' \) containing average task \( \lambda_{t+2} \); when moving from current state to state \( s' \), the probabilities of obtaining average task in new state are equal; Calculation of transition probabilities to next new status containing current term tasks \( \lambda_{t+1}, \Delta t, l_{t+1} \); current term task \( \Delta t, l_{t+1} \) contained in state \( s' \) is calculated from distribution formed by average task \( \lambda_{t+1} \) of \( s \) state. Therefore, \( \Delta t, l_{t+1} \) probability contained in distribution created by \( \lambda_{t+1} \) in \( s \) state, will give the probability of a new state transition containing this demand; Calculation of the transition probability of new state: transition probability of the new state is calculated by multiplying all the probabilities obtained. But as sums of rows have to be equal to 1, matrix is normalized by dividing elements in each row by row total.

Observation Function:

\[
O(o'\mid s', a_t) = (\Omega_{t+2}, \Delta t, l_{t+1}), o' = (\Omega_{t+2}, \Delta t, l_{t+1}) t=0,1,...,T
\]

Matters to be considered when constituting observation function; As current term task \( \Delta t, l_{t+1} \) and pre-assigned tasks \( l_{t+1} \) of new state \( s' \) are fully observable, in this case, current term task \( \Delta t, l_{t+1} \) and pre-assigned tasks \( l_{t+1} \) of observation obtained in this state must have same values. Probability of obtaining different observation is 0; We assume that UAV generally has high demands and pre-assigned tasks are shared with UAV. Therefore, if tasks \( l_{t+1} \) calculated from new state \( s' \) are high, probability for average task \( \lambda_{t+1} \) to be large is high. This means that probability of obtaining observation with large \( \Omega_{t+2} \) is high. So, the big \( \Omega_{t+2} \) containing observations will be more likely to get. Likewise, if tasks \( l_{t+1} \) calculated from new state \( s' \) are small, probability for average task \( \lambda_{t+1} \) to be small is high. Which means that probability of obtaining observation with small \( \Omega_{t+2} \) is higher. Transition probabilities to other states is lower.

5. Results and Recommendations

In this study, security problems in companies with large areas were considered, and Dec-POMDP model that uses partially observable UAVs with variable tasks was established. POMDP and other sequential decision making models take into consideration long-term effects of actions. For this reason, they will provide great benefits in the future UAV applications in industrial environments. The biggest disadvantage of sequential decision-making model is the quick increase of complexity. DEC-POMDP is more complex than POMDP, while POMDP is more complex
than MDP. Due to increase of the number of uncertainties and decision-makers, transaction complexity increases. Accepting of the results of the calculations during the assessment of the security level by decision-makers with precision might sometimes provide wrong results. Therefore, the level of security was divided into three parts as intact, less disturb and more disturb. Performing the task of calculation with fuzzy logic provided better results. In this model, tasks given to UAVs may always vary. When UAVs decision-making system detects this kind of situation, it decides whether to continue old tasks or not based on previously given criteria. If in some situations, based on some criterias, the security system does not want the UAV decision-making system to make a decision, it must be determined in advance and in cases of such observations, the system or the person who will make this decision must be notified.

To improve the level of security in companies with large areas our model provides different solutions for different situations. In case of one UAV in observable state, our model proposes MDP algorithm. In case of one UAV in partially observable state POMDP algorithm is advised. In case of multi-agent partially observable state Dec-POMDP algorithm is applied. In future research, application of proposed model on real world problems with UAVs is planned. UAVs with necessary sensors and cameras will be designed and results of model will be analyzed.

References

5. Abdullayev TS, Alakberov RB, Kilic KS. Fuzzy query processing in corporative information system. Journals of Qafqaz University, Mathematics and Computer Science 2012; 34: 105-113