

Received December 7, 2019, accepted December 15, 2019, date of publication December 20, 2019, date of current version February 14, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2961239

A New Markov Decision Process Based Behavioral Prediction System for Airborne Crews

YAOZHONG ZHANG¹, KANGJIA YAO¹, JIANDONG ZHANG¹, FRANK JIANG²,
AND MATTHEW WARREN²

¹School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129, China

²Centre for Cyber Security Research and Innovation (CSRI), Deakin University, Geelong VIC 3220, Australia

Corresponding author: Frank Jiang (frank.jiang@deakin.edu.au)

This work was supported in part by The Aeronautical Science Foundation Under Grant 2017ZC53033.

ABSTRACT In order to ensure the normal and stable flights in the aircraft, a variety of sensors and corresponding instrumentation systems have been applied on the aircraft to monitor/control the current flight status, and the resulted data ensure the flight safety with a heavy burden on the pilot. In views of this, nowadays, the aircraft cockpit automation assistance system has become a hot topic. This paper is based on the pilot's future operational behavior which can be predicted through different stages of flight operations after the automated assistance system is triggered, thus providing the pilot with assistance in accordance with his operating habits. We have established a MDP (Markov Decision Process) model via analyzing and modeling of pilot operational behavior and mission requirements for flight processes, and we also use value iterative algorithm to find the optimal prediction sequence, lastly, we verify the operability of the algorithm by flight operation simulation experiment. It provides a new solution for the safety of pilot operations and the intrusiveness of the cockpit adaptive automation assistance system.

INDEX TERMS Cockpit automation assistance, Markov decision process, behavioral prediction.

I. INTRODUCTION

The crew's flight operations can directly affect the flight safety of the aircraft, the literature [1] counts 310 accidents occurred in China during the landing period from 1996 to 2005 and analyzes the causes of the accident. Among them, there were 155 accidents directly related to the crew's operation, accounting for 50% of the total number of accidents. Figure 1 shows the distribution of relevant flight accidents and accident type statistics.

As shown in Figure 1, the improper operation of the pilots is one of the most direct factors that cause the accident, other factors including pilot's judgment mistakes, illegal operations, unstable approach, etc. When an aircraft system failure occurs, the attention of most flight crews is focused on troubleshooting, ignoring the overall monitoring of factors affecting flight safety, and thus easily leads to flight accidents. In practice, approximately 11.8% of flight crews fail to comply with standard operating procedures during troubleshooting, this becomes the main reason of flight accidents.

The associate editor coordinating the review of this manuscript and approving it for publication was Amr Tolba¹.

The introduction of cockpit automation assistive technology occurs initially in the 1970s, it was defined as a device that accomplishes a function that was previously carried out by the pilot, that is, the equipment in the cockpit can be self-executed before the pilot performs and/or complete the corresponding function. However, it was found that this improvement of automation technology caused the degradation of the pilot's situational awareness and operational skills. Therefore, the concept of adaptive automation is proposed to improve the performance of cockpit automation management from the aspects of information acquisition, analysis and display information, behavior decision-making and task management. Subsequent studies have shown that adaptive automation is a more advanced automation technology than traditional automation. Today's cockpit automation assistance systems are further improved for automated task management, dynamic task assignment, and driver workload.

At present, with the development of artificial intelligence technology, computer technology and modern control theory, engineers have applied more and more intelligent auxiliary systems to the field of flight safety, in particular, the high degree of automation and intelligent technology in the aircraft

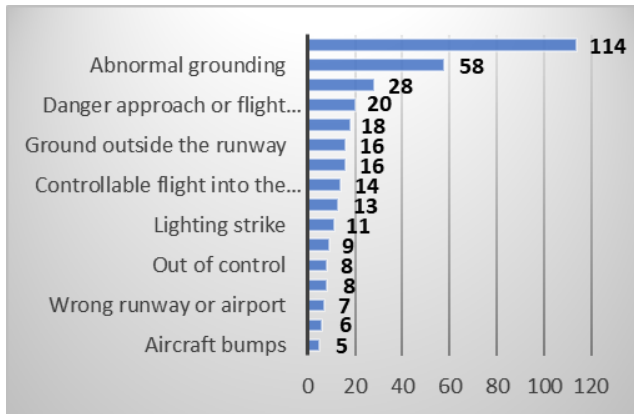


FIGURE 1. Overall stats on various types of domestic flight accidents in China up to 2018.

cockpit has reduced the probability of driver mis-operation, greatly improving the safety of flight. Statistics show that the cockpit automation assisted system can help pilots complete scheduled missions more accurately and reliably, but at the same time, the increase in automation also leads to a decline in pilots' sense of control, that is, the pilot correctly judges the current and predicts future. The flight ability is reduced, which can easily lead to flight accidents. NASA (The National Aeronautics and Space Administration) has set up an IIFDT (Integrated Intelligent Flight Deck Technologies) project [2] for this purpose in order to provide safer flight cockpit assist systems to enhance driver automation's AA (Adaptive Automation) assistance capabilities [3], [10], [11], [17], [18].

The current cockpit automation intelligent control system design lacks of flexibility, resulting in the intervention of the decision-making system of the flight operation process becomes inefficient. This unsuitable interruption and intrusion operation will cause the pilot's cognitive deficit and trust failure, resulting in a flight accident. The cockpit adaptive automation assistance system flexibly adjusts the automation level, redistributes functions and operating rights, enables the crew to get the best automation assistance in specific situations, and guarantees the safety of flight driving. But how to reduce the intrusiveness of automated assistance is a key issue in adaptive automation assistance [4], [12]–[14] [19]–[21].

The automated assistance process of the cockpit is a special decision-making process. The selection of operational action depends only on the current flight state of the aircraft or the flight state in a short period of time. It has a certain Markov property regardless of the premature time state. Markov decision is the optimal decision process of stochastic dynamic system based on Markov process theory. Through the study of state space, behavior space and state transition probability, the future state and change of the system can be predicted to some extent. Therefore, this paper uses the MDP to predict the crew's operational behavior, so that the crew's most urgent operational assistance needs can be fulfilled, and in the future, it can effectively reduce the intrusiveness of automation system assistance.

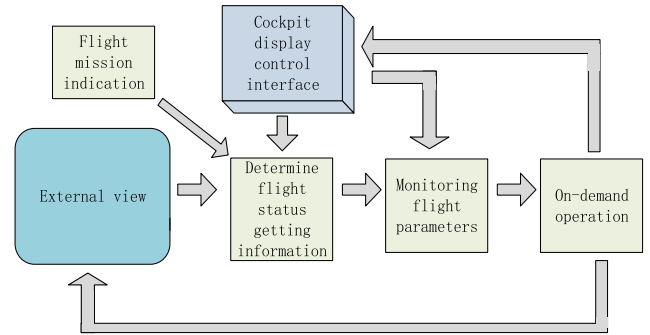


FIGURE 2. Crew working process.

The contributions of this work are mainly threefold, firstly a MDP (Markov Decision Process) model is proposed and validated via analyzing and modeling of pilot operational behavior and mission requirements in flight processes, and secondly, proposed a value iterative algorithm to find the optimal prediction sequence. Thirdly, the decision-making process enables the automated pilot's behavioral prediction model for the practice use for the cabinet in the airborne. Specifically, the direct outcome/contribution is we can determine the behavioral actions of the pilot in the current and foreseeable future for decision making, the experiments accumulate a large amount of data which can be used as the publicly accessible benchmark data for academic purposes.

II. CREW WORK PROCESS ANALYSIS

Figure 2 shows the crew's working process. For a specific type of aircraft, the cockpit display control system is fixed and cannot meet the crew's control requirements flexibly. Especially in the case of complicated external scenery such as rain and fog, although the current ISL instrument system greatly improves the safety of blind drop, the poor views will still affect the crew's judgment. In the concept of human-machine integration, the key part is human-computer interaction, crew monitoring and automated monitoring in parallel, human and automation systems can simultaneously observe each other's status information. This is a design that is parallel to trust and distrust. The cockpit's automated assistance system must have the effect of blocking human error while maximizing the crew's flexibility and agility.

To reduce the intrusiveness of the displayed information, the prediction process needs to be carried out according to the crew's operating habits, while ensuring flight safety. Reasonable solution to the problem of crew mis-operation is a more feasible method to improve flight safety. Boeing and Airbus have also been developing intelligent autonomous driving to ensure that automated equipment helps the crew improve the reliability of driving behavior or block the handling of threats to flight safety. Based on the DDS (Display and Decision Support) in NASA's smart cockpit design requirements, we have proposed an operational intent prediction model that is essentially a process of sensing crew behavior and making inferences. The forecasting process must consider not only

the immediate effects of the outcome of the decision, but also the opportunity to create opportunities for future decisions. Since the selection of the crew's operational actions depends only on the current flight status of the aircraft or the flight status within a short period of time, it is independent of the premature moment. Therefore, we apply the MDP theory in the sequence decision model to predict the crew's operational intent.

III. MDP THEORY

A. BASIC COMPOSITION OF MDP

The characteristics of Markov's decision-making process are the set of behavioral spaces which can be used and the specific actions that will be taken only up to the current state of the system and have nothing to do with past history. This property is called Markov. The main part of Markov's decision process is similar to the general sequence decision model, it has: Decision cycle, state, action, transition probability and return [5], [15], [16], [22], [23], namely:

$$\begin{aligned} MDP &= \{T, S, A(s^i), p(s^j|s^i, a_k), r(s^i, a_k)\} \\ MDP &= \{S, A, T(s^i, a_k, s^j), R(s^i)\} \end{aligned} \quad (1)$$

And there also are:

- (1) T represents a set of decision time sets, which are subsets on a non-negative solid line, which can be a finite point set, a list of infinite points, or a continuous set. In this paper, the decision set of the system is $T = \{1, 2, \dots\}$.
- (2) S represents a set of system states, called state spaces, that contain the states that may occur in all system namely $S = \{s^1, s^2, \dots, s^{n_s}\}$, The state of the system in this paper is the flight state of the aircraft.
- (3) A indicates the action space, which is a set of actions that can change the state of the system. Here, the crew's operational behavior can change the flight status of the aircraft. The action space is related to the system status, $A(s^i)$ represents the available action in state i .
- (4) $T(s^i, a_k, s^j)$ represents a set of all state transition probabilities. Any one of the elements $p(s^j|s^i, a_k)$ indicates the probability that the state a_k changes and the state of the system changes to state s^j under state s^i . Hence we have:

$$\sum_{j \in S} p(s^j|s^i, a_k) = 1 \quad (2)$$

- (5) $R(s^i)$ represents the return set. $r(s^i, a_k) \in R(s^i)$ indicates the return of action a_k under state. When $r(s^i, a_k) > 0$ is the return, and $r(s^i, a_k) < 0$ is the fee? Once behavior a_k is selected, the reward can be an exact value or an expected value. At the same time, the calculation method of the return can be a one-time gain to the next decision time, or a cumulative return to the next stage and a random return to the next state. In general, the return on earning depends on the state s^i

at the next moment, namely $r(s^i, a_k, s^j)$. By definition, we have

$$r(s^i, a_k) = \sum_{j \in S} r(s^i, a_k, s^j) p(s^j|s^i, a_k) \quad (3)$$

It can be seen from the definition that the transition probability and reward of Markov decision depends only on the current state and the actions selected by the decision maker and does not depend on the past history.

B. MDP DECISION RULES AND STRATEGIES

The state of the system using the Markov decision process and the corresponding actions can be represented by a trajectory:

$$h_t = (s_0^i, a_0, s_1^i, a_1, \dots, s_{t-1}^i, a_{t-1}, s_t^i) \quad (t \geq 0)$$

is called a sequence of processes from time 0 to time t . Use s_k^i to indicate the state the system is in at time $t = k$, a_k is the action selected from the behavior space $A(s_k^i)$ at time $t = k$

The decision rule indicates that the system determines the behavior to be taken according to the state at any decision time. The decision rule of the MDP is defined as:

If function f on state space S satisfies: For each $s^i \in S$, there is $f(s^i) \in A(s^i)$, that is, $f : S \rightarrow A(s^i)$, then f is called the decision function. The set of all decision functions is denoted as F .

The probability distribution family π_t of state S satisfies: For each t moment $s^i \in S$, there is a probability distribution of $\pi_t(\cdot|s^i) \in Dis(A(s^i))$ on $A(s^i)$, that is, both $\pi_t(a_k|s^i) \geq 0$ and $\sum \pi_t(a_k|s^i) = 1$ are satisfied, Then call π_t is the MDP decision rule. In application, the determined MDP decision process rules are generally used.

Define a sequence of decision functions $\pi = (f_0, f_1, f_2, \dots)$ which is a deterministic MDP strategy, where $f_t \in F, t$ is the decision time. f_t depends only on the current decision time t , The set of all MDP strategies is denoted as Π_m^d , which is called the MDP strategy class.

According to the operating characteristics of the crew in the cockpit, that is, at any time, the control aircraft is on the safety envelope and the mission line. Although it is not guaranteed to be optimal operation at every moment, at least the aircraft must be safely and smoothly. The established requirements complete the missions of each phase. Therefore, in order to meet the requirements of the actual situation, the MDP strategy should be determined, which is: An MDP strategy $\pi = (f_0, f_1, f_2, \dots) \in \Pi_m^d$, if $\forall t f_t \equiv f_0$ is satisfied, is called a smooth strategy. The set of all the stationary strategies is denoted as Π_s^d , which is a stationary strategy class. In this paper, the stationary strategy is adopted in the unit operation intention prediction model.

C. MDP OPTIMAL CRITERIA

It is assumed that after selecting a specific strategy and implementing it, the decision maker obtains a series of rewards at a certain probability at time T , the specific utility function of the model is the accumulation after the discounted reward, This method is called the infinite stage discount model, and its discount rate is the discount factor [7]. Defined as follows:

For strategy $\pi \in \Pi_s^d$ and fixed discount factor β and $0 < \beta < 1$, the reward utility function of the discount model is:

$$V_\beta(s^i, \pi) = \sum_{t=0}^{\infty} \beta^t E_\pi^s [r(s^i, a_k)] \quad (4)$$

Indicating that the discounted expected total reward of the Strategy π system is used under the condition that the start time 0 is triggered from the state s^i . According to the definition of reward $r(s^i, a_k)$, the reward function can be bounded, and the utility function is also bounded.

$$V_\beta^*(s^i) = \sup_{\pi \in \Pi_s^d} V_\beta(s^i, \pi) \quad (5)$$

Denote equation (5) as the optimal value function. When the equal sign is established, the corresponding strategy π^* is the optimal strategy. In the actual situation, the perfect optimal solution cannot be obtained in the calculation, and the error bound parameter ε is introduced. For $\varepsilon \geq 0$, if the strategy π^* makes $V_\beta(s^i, \pi^*) \geq V_\beta^*(s^i) - \varepsilon$ to the state s^i , then π^* is considered as the ε optimal strategy of the discount model.

D. VALUE ITERATIVE ALGORITHM FOR SOLVING THE MDP OPTIMAL PROCESS

Definition in MDP:

$$\exists v^* \quad T v^* = v^* \quad (6)$$

T satisfies:

$$T v \equiv \sup_{f \in F} T_f v \equiv \sup_{f \in F} [r(f) + \beta P(f)v] \quad (7)$$

Equation (6) is the optimal equation for MDP. The solution v^* of the equation is V_β^* . Since v^* is unique, the calculated v^* satisfies $\|v^*\| = \sup |v|$. The decision function f in equation (7) is a deterministic decision rule, and the obtained strategy π^* is an infinite phase stationary strategy.

We use the VI (Value Iteration) algorithm to solve the optimal equation of MDP. The VI algorithm, also known as Successive Approximations, is a simple and easy numerical algorithm [8]. The advantage of the VI algorithm is that it is easy to calculate. As long as the action set and state set are limited, it can quickly converge to a solution set, and the initial set value does not affect the calculation result. The specific workflow of the algorithm is given below:

Step 1: Establish the corresponding optimal equation

$$v(s^i) = \max_{a_k \in A(s^i)} \left\{ r(s^i, a_k) + \beta \sum_{s^j \in S} p(s^j | s^i, a_k) v(s^j) \right\} \quad (8)$$

Step 2: Let v^0 be a bounded set, given the discount factor β and the error bound ε , the number of iterations of $n = 0$.

Step 3: For each state $s^i \in S$, by calculating

$$v^{n+1}(s^i) = \max_{a_k \in A(s^i)} \left\{ r(s^i, a_k) + \beta \sum_{s^j \in S} p(s^j | s^i, a_k) v^n(s^j) \right\} \quad (9)$$

get $v^{n+1}(s^i)$.

Step 4: if

$$\|v^{n+1} - v\| < \frac{\varepsilon(1 - \beta)}{2\beta}$$

Jump to Step 5;

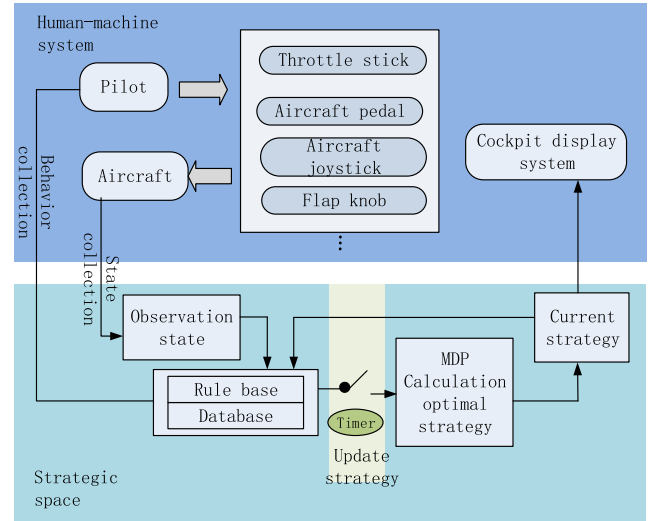


FIGURE 3. MDP prediction model structure.

Otherwise, the number of iterations n is increased by 1, return to step 3.

Step 5: for each state $s^i \in S$, take

$$f_e(s^i) \in \operatorname{argmax}_{a \in A(s^i)} \left\{ r(s^i, a_k) + \beta \sum_{s^j \in S} p(s^j | s^i, a_k) v^{n+1}(s^j) \right\} \quad (10)$$

Then stop

Equations (9) and (10) can be written as $v^{n+1} = T v^n$

$$f_e \in \operatorname{argmax}_{f \in F} \left\{ r(f) + \beta P(f)v^{n+1} \right\} \quad (11)$$

Equation (11) gives the action strategy with the greatest benefit in each state at the current time.

IV. CREW BEHAVIOR PREDICTION MODEL

The crew's operational tasks in the cockpit are complex and varied. To accurately predict the operational intent of the crew (mainly the pilot), it is necessary to understand the pilot's working space. According to the flight manual of FAA (the Federal Aviation Administration), the unit workspace is divided into four levels: flight phase, flight mission, flight state and direction operation, as depicted in Figure 4. This paper focuses on the flight status and flight operation level and predicts the pilot's flight operations through flight status.

A. OPERATIONAL BEHAVIOR PREDICTION MODEL STRUCTURE

In order to establish an MDP model of the man-machine system in the cockpit, it is necessary to ensure that the set of states and actions set are observable. At the same time, the operations performed by the crew at the next moment are only related to the state of the aircraft at the current or previous time, and the state here is also the set of states set by the model. Figure 3 shows the MDP prediction model structure in the cockpit man-machine system. The database and rule

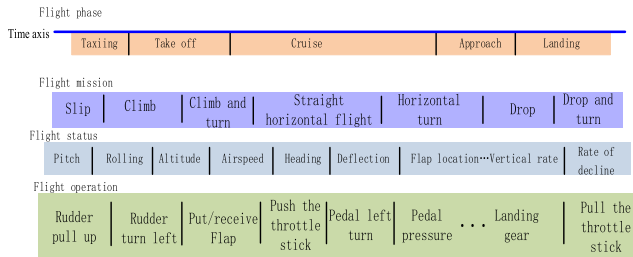


FIGURE 4. Four levels of unit workspace.

base in the diagram need to be pre-set to describe a certain time period or a process mission and mission requirements. When the automated assistance system is triggered, the system provides a set of calculated optimal strategies based on the behavior set and the state set. When the crew’s operation does not match the provisioning strategy, the system updates the strategy according to the crew’s historical operation. Whenever the updates of the strategy occur, the strategy is recalculated according to the optimal criteria, starting from the actions of the crew that do not conform to the historical strategy.

In the flight process, in addition to the necessary operations, the crew also needs to solve some unexpected situations in time, which increases the workload of the crew. In addition, the in-flight state transition may not have an exact correspondence with the achievement of the target. If the target in each state cannot be fully achieved, the effect will continue to the next state.

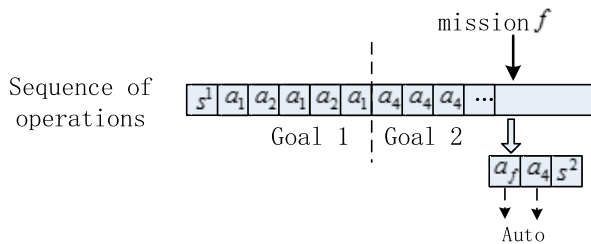


FIGURE 5. The effect of inserting a task on the sequence of operations.

Figure 5 shows the crew’s need to perform a single operational behavior or a sequence of operations to complete the insertion task after the burst task is inserted. Before the state transition, if the target in the state is not fully completed, then the crew does not have enough resources to process the tasks or complete the remaining targets. At this time, the system will help the crew to perform related tasks according to the setting of the criteria library.

By using the MDP prediction model observed in different mission environments and states, a set of behavioral strategies that best meet the crew’s intent can be obtained. When the system can accurately predict the follow-up behavior of the crew, then the crew can provide the corresponding assistance through the sequence of behaviors, or correct the crew before the wrong behavior, or when the crew needs to handle more

Help the crew perform tasks when interrupting and inserting tasks.

B. CREW BEHAVIOR PREDICTION MODELING

The operational intent prediction of the crew in the cockpit is carried out in an uncertain environment. The model composition is the same as the general form of MDP. The output is the strategy that meets the optimal criteria based on the current state of time. The specific expression is as follows:

$$\begin{cases} MDP = \{T, S, A_s, p(s^j|s^i, a_k), r(s^i, a_k)\} \rightarrow \pi_t(s^i) \\ MDP = \{S, A, T(s^i, a_k, s^j), R(s^i)\} \rightarrow \pi(s^i) \end{cases} \quad (12)$$

The return in the crew behavior prediction model is that the return in the variable crew behavior prediction model is changed. The Timer in Figure 3 starts from the initial decision time. Each interval A needs to re-observe the system state and action set. If the state shifts, and the state transferred to does not match the state expected by the crew, at which point the system will recalculate the strategy. Since the target task is very clear during the flight, the process of MDP calculation should go from a certain state to an optimal state to reach the target state. In fact, the design of Timer is essentially a timely reminder of human error. At the same time, when noise or loss occurs in the communication information, the Timer can continuously reduce the occurrence of input errors by continuously sampling. As shown in Figure 6, due to the complexity and volatility of the crew’s cognition and behavior, as well as the inherent error and delay of the system, the transfer of the system state is not carried out as originally calculated by the model. The strategy is continually updated to meet the needs of the human-machine system.

The actions and states in equation (12) are specifically expressed as:

$$\begin{cases} S = \{s^1, s^2, \dots, s^{n_s}\} \\ A = \{1, 2, \dots, n_a\} \\ s^i = \{G^i, A^i, F^i\} \\ G^i = \{g_1^i, g_2^i, \dots, g_{n_g}^i\}, \quad g_k^i \in \{0, 1\}, \quad k = \{1, \dots, n_g\} \\ F^i = \{f_1^i, f_2^i, \dots, f_{n_f}^i\}, \quad f_k^i \in \{0, 1\}, \quad k = \{1, \dots, n_f\} \\ A^i = \{a_1^i, a_2^i, \dots, a_{n_h}^i\}, \quad a_k^i \in A, \quad k = \{1, \dots, n_h\} \end{cases} \quad (13)$$

Among them, each state s^i in the system state set S is observable. Each state s^i contains three sets of variables: task target G^i , insert task target F^i , and state action set A^i , among them:

1) Task Objective G^i is a set of low priority task target states, which are binary values. Use 1 and 0 to indicate that the task is completed or not completed. The value can be calculated according to the status of the monitoring parameters.

2) Inserting a task target F^i is a set of high-priority task target states that describe the tasks that need to be performed and inserted at a particular moment in the normal course of the system.

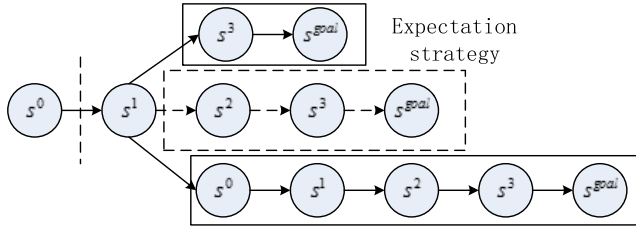


FIGURE 6. MDP policy update process.

3) The state action set A^i represents a set of historical action sets in state s^i , and n_h is the number of set elements, representing the number of operation sequences at the historical time. A^i is a subset of the action set A .

Action Set A is a set of coded vectors. The elements of the vector are natural numbers 1 through n_a , which represent the crew’s normal operations and routine operations. During the flight, the crew selects actions to complete the corresponding tasks or insert mission targets. It is assumed here that at the same time, actions can only be selected to accomplish a goal, that is, the elements in G^i and F^i are independent of each other.

The utility function $r(s^i, a_k)$ in equation (12) indicates that the action has an effect on the state, indicating the effect of the action on the state. Since the task target state is judged based on the flight parameters, the deviation between the parameter α and the expected interval d_α affected by the behavior a_k is used to calculate Behavioral rewards. To ensure a reasonable return, we stipulate that:

$$\begin{cases} r(s^i, a_k) < 0 & \text{if } \alpha(a_k) \notin d_\alpha \\ r(s^i, a_k) \geq 0 & \text{if } \alpha(a_k) \in d_\alpha \end{cases} \quad (14)$$

$\alpha(a_k)$ indicates the possible value of α after taking action a_k . $\alpha(a_k)$ is not necessarily a certain value, its meaning is to describe the direction of change of parameter α , indicating the impact of a_k on the mission target. Suppose $\alpha(a_k) \notin d_\alpha$ means that after action a_k , α changes to d_α ; otherwise, $\alpha(a_k) \in d_\alpha$ means that after action a_k , α changes within d_α . Then, after correcting equation (14), you can get:

$$\begin{cases} r(s^i, a_k) < 0 & \text{if } \alpha(a_k) \notin d_\alpha \text{ and } \alpha \in d_\alpha \\ r(s^i, a_k) \ll 0 & \text{if } \alpha(a_k) \notin d_\alpha \text{ and } \alpha \notin d_\alpha \\ r(s^i, a_k) = 0 & \text{if } \alpha(a_k) \in d_\alpha \text{ and } \alpha \in d_\alpha \\ r(s^i, a_k) > 0 & \text{if } \alpha(a_k) \in d_\alpha \text{ and } \alpha \notin d_\alpha \end{cases} \quad (15)$$

C. AN EXAMPLE MODEL FOR CREW BEHAVIOR PREDICTION

During the full flight of the aircraft, taking off and landings are a relatively dangerous phase and a high level of pilot demand for automated assistance. Approximately 70% of dangerous approaches are caused by unfunded resource calls and allocations by the crew. 40% of them are too slow or too low during the approach, and 30% of the approach is too fast or too high. If the unit can be reminded to improve or correct it before the unit performs improper operation,

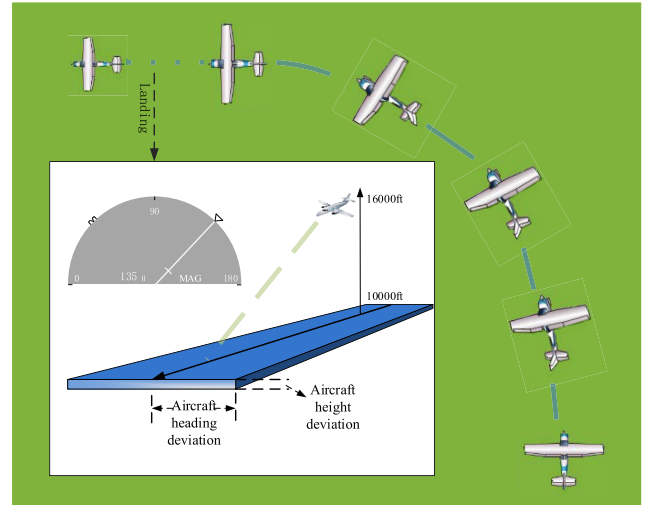


FIGURE 7. Horizontal turning and descent of the aircraft.

the danger can be better blocked as soon as possible, so that flight safety can be better ensured. In order to illustrate the problem, we select the horizontal turn and the height drop as the research stage and establish the MDP crew behavior prediction model. The horizontal turn and the height drop are the two main flight stages in the approaching landing process, with good representation.

In the landing phase, the horizontal turn and the final approach drop during the approach are two of the main missions performed by the aircraft. The operation required for the approach of the aircraft to the landing is shown in the figure below.

At this point, the two tasks that the crew needs to perform are:

- 1) Control the aircraft to complete the heading adjustment. If the current heading is 135° under the magnetic reference, it needs to be adjusted to 45° and the error is controlled at $\pm 0.5^\circ$;
- 2) After the heading adjustment is completed, the flying height needs to be lowered to 10000ft ± 300 ft and then turned to level flight.

Taking Boeing’s B737-800NG aircraft as an example [6], it is assumed that the crew has ten kinds of executable operational behaviors, as shown in Table 1 below:

The reset operation in Table 1 corresponds to an executed operation in A^i . It is assumed that the current teletype device of the aircraft can directly measure the angle of rotation or push of each operating device and transmit it to the task computer, which will be rewritten into the following matrix form:

$$A^i = \begin{bmatrix} 3 & 7 & 1 & 6 & 9 \\ -0.3 & -0.35 & 0.75 & -0.12 & 0 \end{bmatrix} \quad (16)$$

The row definition of A^i is the same as in equation (13). The columns are respectively the operation behavior codes in Table 1 and the corresponding operation degrees. For example, in equation (16), the first column corresponds to the

TABLE 1. Crew operational behavior sequence.

A	Flight operation	Description
1	Push the throttle stick	Increase engine power and increase thrust
2	Pull the throttle stick	Reduce engine power and reduce thrust
3	Turn left joystick	The left wing is lowered relative to the pilot, forming the left turn of the aircraft
4	Turn right joystick	The right wing descends from the pilot, forming the right turn of the aircraft
5	Push lever	The elevator is deflected downwards and the nose is lowered relative to the pilot
6	Pull lever	The elevator deflects upwards and the nose rises relative to the pilot
7	Left-hand flight pedal	The rudder is deflected to the left and the nose is deflected to the left relative to the pilot
8	Right-hand flight pedal	The rudder is deflected to the right and the nose is deflected to the right relative to the pilot
9	Maintain	Maintain the status of each manipulator in each behavior A^i
10	Reset	Reset the action A^i device that has been executed

left turn joystick and the position of the turn. The meaning of equation (16) is that it is possible to directly determine which operation the unit has performed based on changes in the position of the operating device. The amount of operations is the value of a set of $[-1, 1]$. And the left full bias is -1 , the right full bias is 1 ; the rear pull to the maximum indicates -1 , and the forward push to the front indicates 1 . Equation (16) eliminates the corresponding uncertainty in the historical behavior A^i of the reset operation, and the angle at which the reset needs to be rotated is different depending on the model. Here, the reset operation corresponds to the value 0 in the manipulated variable.

1) HORIZONTAL TURN TASK

During the execution of a horizontal turning mission, the aircraft needs to continuously monitor several parameters such as slope, heading, airspeed and altitude. The execution procedure for horizontal turns in the FAA’s Aircraft Flight Manual is shown in Figure 8:

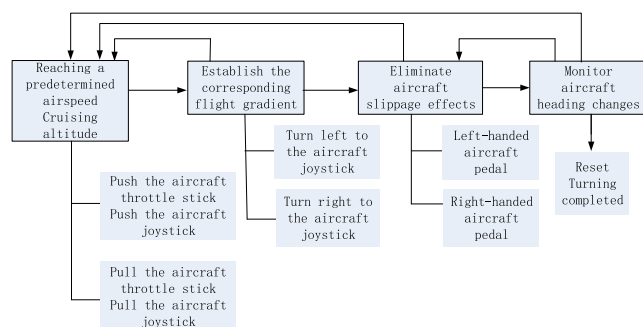


FIGURE 8. Horizontal turning operation flow in FAA.

According to the description of equation (13), we can get: According to the state conditions in Table 2, the effects of operational behavior are different in different states. Since the different operations of the horizontal left turn and the horizontal right turn have different effects on the state in the equation (13), in order to fully and objectively represent the influence of the state and behavior of the entire horizontal turn with a generalized model, the states in Table 2 Simplified into discrete state sets $\{-1, 0, 1\}$. -1 indicates that the current state is left or too small; 1 indicates that the current state is right or too large; 0 indicates that the target state satisfies the task requirement. It should be noted that in order to meet the general model of horizontal turning, it is stipulated that the slope along the longitudinal axis of the machine is positive, the slope established by clockwise is positive, the slope established by counterclockwise is negative, and the value is floated by the change process to determine the left. Partial or right deviation.

TABLE 2. Horizontal turn task target set.

G^i	Parameter	Description
θ_1	slope	The passenger aircraft requires a turning slope at $[-30^\circ, 30^\circ]$, here choose 15°
θ_2	course	Under MAG, turn left from 135° to 45°
f_1	airspeed	Airspeed is maintained at 200kt-215kt
f_2	altitude	The horizontal turn phase height position is assumed to be 16000ft
f_3	side slip	Monitor PFD at all times to eliminate aircraft slippage as quickly as possible

According to the description of the aircraft horizontal turning mission described in Figure 8, ideally, that is, the unit operation without any error, the state deflection should be:

$$s^0 = [1, 1, 0, 0, 0] \rightarrow s^1 = [0, 1, 0, 0, 0] \rightarrow s^2 = [0, 0, 0, 0, 0]$$

In the actual flight mission execution process, human error will be generated at every moment, and because of the hysteresis of the aircraft power system, the error will accumulate to a certain extent, and will be detected after a period of time, and the deviation will be transmitted to the cockpit. Will greatly increase the workload of the crew. If the cockpit system can report the current level of the aircraft’s relatively perfect completion of the mission and provide the necessary information for the pilot to correct the operation in time, then the crew will be safer and more reliable when performing complex manual tasks.

In this paper, we propose a MP (Mission Pressure) indicator to describe the timing of the decision and the satisfaction of the target G^i with respect to the overall mission requirements. For example, the mission requires the aircraft to establish a reasonable slope within $6s \sim 8s$ and complete

the heading transition within 60 ~ 65s. We can assume that the initial value of the target parameter is θ^0 at this time, and the expected arrival value is θ^T , and the current value is θ^t . The initial time is 0 and the task request time is T . Then the current task urgency can be expressed as:

$$MP(\theta)_t = \frac{\frac{|\theta^T - \theta^t|}{|\theta^T - \theta^0|}}{(T - t)/T} \tag{17}$$

When the pilot performs the task perfectly, the calculated value of $MP(\theta)_t$ should approach 1 at any time. Any greater than 1 or less than 1 indicates that the current task execution is too early or too late. Especially in the approaching landing stage, when the altitude $MP > 1$ indicates that the rate of decline is too fast, the aircraft passes the target altitude too early; when $MP < 1$, the aircraft passes the target height too late, and the distance from the runway entrance is too close.

The calculation of MP can discretize a state with a large span very well. At the same time, according to the characteristics of human operation, the acceptance domain of MP at each moment is also related to the execution time. Generally, when the task is first started, the limitation of human error is low, and the hysteresis of the system itself, MP can be flexibly maintained a wider range. When the task is executed to the end, in order to ensure the safety of the human-machine system, it is necessary to strictly control MP to ensure that the flight mission is completed with the specified flight parameters.

According to the definition of state set s^i , the influence of operational behavior on flight state can be obtained, as shown in Table 3.

TABLE 3. Behavioral influence matrix during horizontal turn mission.

status	Unit operating behavior							
	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
g_1	0	0	-1	1	0	0	0	0
g_2	0	0	-1/2	1/2	0	0	-1/2	1/2
f_1	1/2	-1/2	0	0	1/2	-1/2	0	0
f_2	0	0	0	0	-1	1	0	0
f_3	0	0	0	0	-1	-1	1	1

The elements of each row in the table indicate what needs to be done to complete the corresponding task target. For example, the correction of the heading g_2 requires both the joystick a_3a_4 to establish the slope and the pedal a_7a_8 to eliminate the side slip. The coordinated use of the two can stably change the heading. The positive and negative values corresponding to each element value indicate the direction in

which the behavior changes state. When the task is executed, the slope g_1 target is first completed, and the status of the high priority target f_1, f_2, f_3 is monitored and corrected in time. When g_1 is reached, the operation action (return position) that affects g_1 is eliminated, and g_2 is monitored while the target is completed.

Since the pilot's operational behavior will take some time to affect the flight status of the aircraft, and the system will feed the parameters back to the cockpit display interface, the pilot will often perform his own estimated operation and then wait for the table speed. Operate or maintain the operation. When determining the state transition probability $p(s^j|s^i, a_k)$, it is necessary to consider whether the current pilot is at the same speed or in operation. We stipulate that if the current operation in A^i is to maintain the behavior, it means that the pilot is waiting for the pace. If it is not satisfied, record the current near-current maintenance behavior in A^i to the most recent operation behavior as \bar{A}^i . It is assumed that the transition of the state cannot jump across the level, that is, the transfer of each task target in Table 2 can only be from -1 to 0 , and not from -1 to 1 . When \bar{A}^i exists and the unit performs the behavior of , the corresponding $p(s^j|s^i, a_k)$ will increase.

According to the description of Equation (15) and Table 3, the state change and the corresponding reward value that may be caused after the execution of the operation behavior a_k can be obtained, as shown in the following table:

The value of $p(s^j|s^i, a_k)$ in Table 4, in G^i, f_1, f_2 , the former indicates the state transition probability when there is no a_k in \bar{A}^i ; The latter indicates the state transition probability that a_k has already in \bar{A}^i . In f_3 , the previous probability of $p(s^j|s^i, a_k)$ content is expressed as the state transition probability that the inner slide down pedal operation a_7 and a_8 do not exist in \bar{A}^i and the state transition probability that the pedal operation exists in \bar{A}^i , The outer slide-down lever rotation operation a_3 and a_4 state transition probability that does not exist in \bar{A}^i and the state transition probability that the joystick rotates in \bar{A}^i .

The adjustment of the aircraft's heading can be seen as a correction to the slope g_1 and the airspeed f_1 . When the MP of the aircraft heading g_2 is too large, it indicates that the current turning rate of the aircraft is too low; on the contrary, the turning rate is too large:

$$R = \frac{562 \tan(arf)}{v} \tag{18}$$

$$r = \frac{v^2}{g \tan(arf)} \tag{19}$$

Equations (18) and (19) represent the general calculation formulas for the airplane's turning rate and turning radius, respectively. Where arf represents the established slope; v represents the current aircraft airspeed; g represents the gravity acceleration of the area in which it is located.

It can be seen from the formula that adjusting the aircraft airspeed or changing the turning flight gradient can change the current turning rate of the aircraft. When the value of MP

TABLE 4. Horizontal turn task strategy calculation parameter table.

s^i	A	s^j	$p(s^j s^i, a_k)$	$r(s^i, a_k)$
g_1	a_3	-1	1	-2
	a_4	-1	0.5/0.1	1
		0	0.5/0.9	2
	a_3	0	0.5/0.9	2
1		0.5/0.1	1	
g_2	a_3	-1	1	-2
	a_4	-1	0.5/0.25	1/0
		0	0.5/0.75	2
	a_3	0	0.5/0.75	2
1		0.5/0.25	1/0	
f_1	a_4	1	1	-2
	a_7	0	0.5/0.75	2
		1	0.5/0.25	1/0
	a_8	1	1	-2
f_2	a_1	-1	0.5/0.1	1/0
	a_2	-1	1	-2
		0	0.5/0.9	2/0
	a_5	-1	0.5/0.1	1/0
0		0.5/0.9	2/0	
f_3	a_6	-1	1	-2
	a_1	1	1	-2
		0	0.5/0.1	1/0
	a_2	1	0.5/0.1	1/0
0		0.5/0.9	2/0	
f_3	a_5	1	1	-2
	a_6	1	0.5/0.1	1/0
		0	0.5/0.9	2/0
	a_5	-1	1	-2
f_3	a_6	-1	0.5/0.1	1
	a_5	0	0.5/0.9	2
		1	0.5/0.9	1
	a_6	1	1	-2
a_3		-1	1	-3
f_3	a_4	-1	1/0.1	-2/1
	a_7	0	0/0.9	-2/2
		1/0.5	2/1	
	a_8	-1	1/0.5	-3/0
f_3	a_3	0	1/0.5	2/1
	a_4	1	1	-3
		a_7	1	1
	a_8	0	0/0.9	-2/2
a_8	1	1/0.1	-2/1	

is too large, it means $g_2 = 1$. At this time, increasing the slope and reducing the airspeed can correct the heading deviation. In civil aviation aircraft flight, excessive slope will affect the passenger's comfort, so choose to reduce the airspeed preferentially, that is, set $f_1 = 1$. Conversely, MP is too small, indicating $g_2 = -1$. At this time, the slope deviation can be corrected by reducing the slope and increasing the airspeed. In order to ensure passenger comfort, the operation of reducing the slope is often selected, that is, $g_1 = 1$ is set.

According to the MDP model of the horizontal turn task established earlier, the corresponding simulation flow chart can be obtained, as shown in Figure 9.

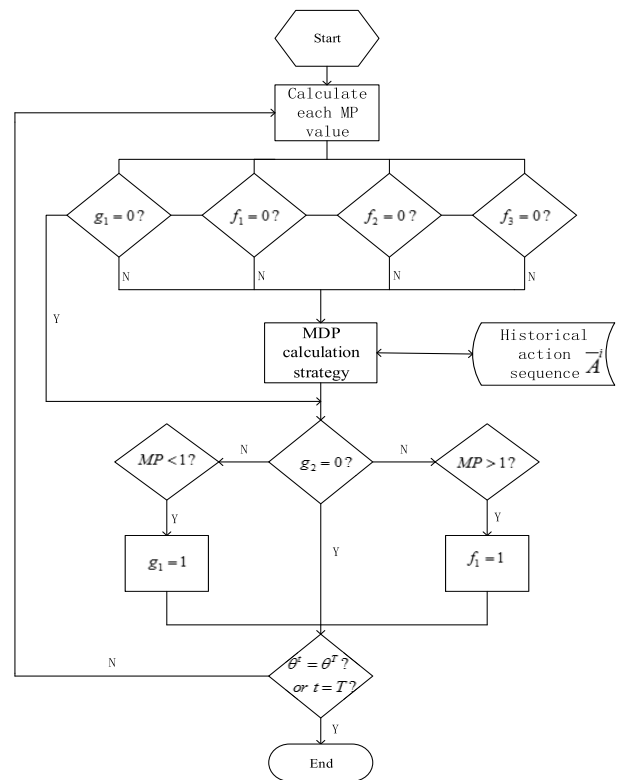


FIGURE 9. Aircraft horizontal turning task simulation flow chart.

2) MISSION OF AIRCRAFT FALLING

When the aircraft Falling, the performance of the aircraft mainly shows the rate of decline, the angle of decline and the distance of the drop. The falling angle is the angle between the falling trajectory and the horizontal line. The descending distance is the horizontal distance through which the aircraft descends at a certain height. The MDP model of the descent process is similar to the horizontal turn. Figure 10 shows the power reduction operation procedure.

According to the requirements of the operating procedures, determine the descending task target set, as shown in Table 5.

Compared with the horizontal turning task, the MDP model of the descending task is relatively simple. The specific

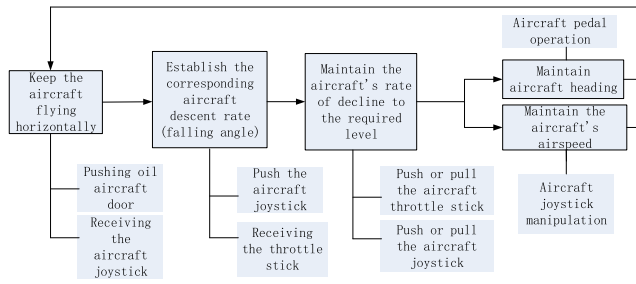


FIGURE 10. Aircraft falling operation procedure.

TABLE 5. Falling task target set.

G^1, F^1	Parameter	Description
g_1	Aircraft descent rate (falling angle)	Adjust the aircraft to the required rate of decline (750ft/min)
g_2	altitude	Monitor the aircraft down to 10000ft
f_1	airspeed	Airspeed is maintained at 200kt
f_2	Aircraft heading	The heading of MAG is 45°

process can refer to the model of the horizontal turning task. The corresponding return parameters of the behavior space and behavior are the same as those of the horizontal turning task.

V. EXPERIMENTAL SIMULATION AND METHOD VALIDATION

In order to verify the theoretical methods and related technologies proposed in this paper, we completed the human loop experiment using virtual cockpit equipment. The MDP method is used to predict the crew’s operational behavior under specified tasks.

A. EXPERIMENTAL PLATFORM INTRODUCTION

The experiment uses the virtual cockpit experimental platform of the School of Electronic Information of Northwestern Polytechnical University. The platform has the functions of pilot data recording and evaluation of pilot mission performance and overall task demand measurement of human-machine system.

1) FLIGHT ENVIRONMENT GENERATION AND FLIGHT DATA GENERATION

The virtual cockpit uses the MSFS (Microsoft Flight Simulator) to generate a virtual driving environment. Microsoft Flight Simulator is a software developed by Microsoft for flight simulation. It is a flight simulator that runs on the Windows operating system. There are a variety of models of civil aircraft models, which can simulate the flight quality of the aircraft more realistically. The software comes with a 3D digital map, with geographic information and runway information for most of the world’s major airports, and some

famous airports have detailed airport facility information. The flight environment generated by the software is very realistic and can realistically simulate the effects of natural light and different climates on flight. The mission can be set freely by using the mission planning function of the software. The software also simulates several typical special flight conditions, such as engine failure. By setting the system failure time, the aircraft will automatically fail to set the system time during the mission. Based on these characteristics of the software, the virtual simulation system uses MSFS to generate the flight environment and generate flight status data. Figure 11 shows the cockpit display interface of the experimental platform. The numbers in the figure correspond to PFD, ND, UDU, flap instrument and landing gear status indicator.



FIGURE 11. Experimental platform virtual cockpit display interface.

2) OPERATING DEVICE

The experimental platform of the virtual cockpit is equipped with a flight joystick, throttle stick and pedal. During the flight, the flight operation is recorded by reading the corresponding rudder, elevator, etc. rotation angle. At the same time, it is also possible to perform the operation of releasing and receiving the flaps, which is a necessary operation in the take-off climbing phase and the approach landing phase. The operation of the landing gear is also placed on the joystick, and the operating mechanism of the entire experimental platform is very close to the real flight cockpit environment.

3) EXPERIMENTAL DATA RECORDER

During the experiment, the experimental data recorder records flight parameter data, pilot report records, and operational behavior sequences according to the set period. All data records will be stored in the flight test data management database for easy reading and subsequent research.

B. EXPERIMENTAL VERIFICATION

The aircraft model was selected as B737-800, and the time was summer daytime. The corresponding flight simulation experiments were carried out in clear weather and foggy weather conditions. The flight environment is shown in Figure 12 and Figure 13, Figure 12 shows the airport and



FIGURE 12. Sunny weather.



FIGURE 13. Foggy weather.

runway environment in fine weather, and Figure 13 shows the airport and runway environment in foggy weather.

The MDP is used to predict the crew’s operational behavior during mission execution. Set the mission as follows:

- (1) Complete the take-off and climb mission and enter the cruise state at the specified level (16,000ft). The heading angle of the MAG is 135°.
- (2) The horizontal turning task is required to be turned to 45° according to the mission requirements, and the heading adjustment is completed within 1 minute, and the leveling attitude is entered.
- (3) Falling to 10,000 ft as required and enter the cruise state.

C. TYPES ANALYSIS OF RESULTS

When using the VI algorithm to iterate the optimal strategy, you need to choose the appropriate discount factor and error bound. Referring to the research results of McGhan C et al. In the calculation of the optimal strategy, the discount factor $\beta = 0.95$ and the error bound $\epsilon = 1 \times 10^{-5}$ [9] are selected. Figure 14 shows the heading curve and the corresponding task urgency MP curve under the horizontal turning task. The red part is the area where the turning rate is too slow, and the task urgency is more than the balance value. Calculate its MDP strategy on the two Skip points in the figure. Figure 15 is a

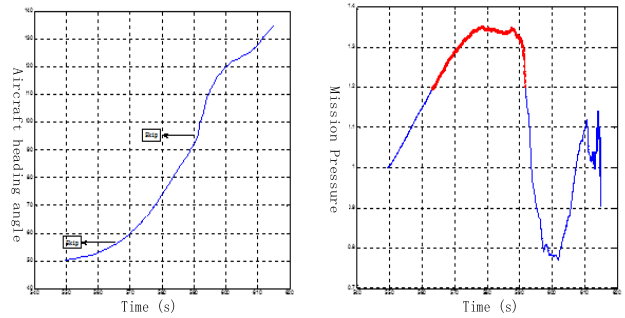
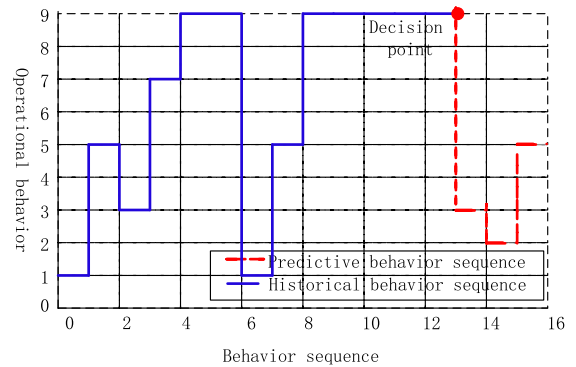
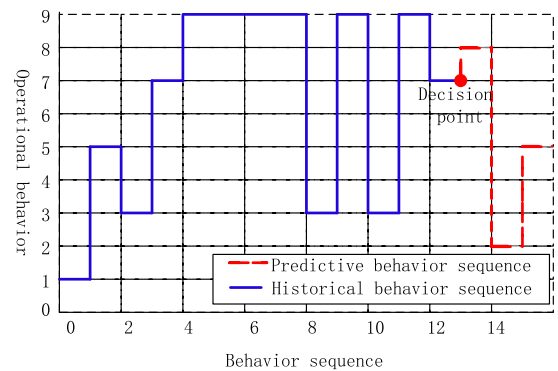


FIGURE 14. Heading angle change and task urgency.



(a) Forecast results starting at Skip1



(b) Forecast results starting at Skip2

FIGURE 15. Horizontal turn task behavior prediction result.

sequence of ideal operational behaviors calculated from the crew’s historical sequence and current state.

In the historical behavior sequence of the crew in Figure 15(b), the operation of the left-turn joystick is performed twice, and finally the Skip appears in the body. This indicates that the current joystick and the pedal coordinate operation are deviated, and the rudder deflection needs to be reduced in time. Angle, since this is a left turn operation, all priority execution behavior 7, Right-hand pedals are used to eliminate side slip. Unlike figure 15(b), in the history sequence observed in figure 15(a), there is no operation of the joystick, and when Skip occurs, the joystick is preferentially deflected to eliminate the side slip, If the side slip is not eliminated, the crew should take the action of Figure 15(b) to eliminate the side slip based on the predicted results. The latter behavior 2 and behavior 6, that is, pulling the

throttle lever and pulling the joystick, is to reduce the airspeed without increasing the altitude, increase the turning rate, and control the MP within the limit error bound. Here, behavior 2 and behavior 6 are synchronized, and can also be performed simultaneously.

Through the analysis of the experimental results, the pilot's operational behavior sequence predicted by the MDP method can effectively guide the current flight state, change the state of the aircraft to a predetermined state, and predict the resulted behavior sequence and the pilot's cognitive process. The same, can well predict the future behavior of the crew, with good feasibility and accuracy.

VI. CONCLUSION

Based on the characteristics of the aircraft cockpit automation assistance system, this paper focuses on reducing the intrusiveness of the automated assistance system. Through the operational characteristics of the crew (pilot) and the mission requirements, in combination with the Markov decision process, the crew's behavior prediction model is established. Meanwhile, by use of two representative task examples of horizontal turn and height drop, the corresponding Markov behavior prediction model is further improved with the experimental verifications, and good experimental results are obtained. By predicting the crew's behavior, we can determine the behavioral actions of the pilot in the current and foreseeable future for decision making, so that the automated assistance system can understand the pilot's needs and provide the best automation assistance at the right time. We conclude the proposed cockpit automation assistance prediction system can reduce the burden on the pilot and ensure the safety and effectiveness of the mission.

REFERENCES

- [1] F. H. Li and Z. N. Xie, "Analysis and preventive measures for the flight accidents and accidents of the national aircraft approaching the landing phase Research," *J. Civil Aviation Flight Univ. China*, vol. 19, no. 2, pp. 3–7, Mar. 2008.
- [2] NASA & The Next Generation air Transportation System (NexTGen), NASA, Washington, DC, USA, 2006.
- [3] R. Parasuraman, M. Mououa, and B. Hilburn, "Adaptive aiding and adaptive task allocation enhance human-machines interaction," in *Automation Technology and Human Performance: Current Research and Trends*. Mahwah, NJ, USA: Erlbaum, 1999, pp. 119–123.
- [4] Q. X. Zhou and Q. Jiang, "Discussion on calculation model of human-machine function allocation and distribution in long-term manned space flight," *Aerosp. Med. Med. Eng.*, vol. 10, no. 6, pp. 417–420, 2007.
- [5] K. Liu, *Perturbed Markov Decision and Hamilton Circle*. Hefei, China: China Univ. of Science and Technology Press, 2009.
- [6] *Aircraft Flight Manual*, Shanghai Airlines, Shanghai, China, 2014.
- [7] Q. X. Zhu, *Optimal Control of Discrete Time Markov Processes—New Optimality Conditions and New Methods*. Guangzhou, China: Zhongshan Univ., 2005.
- [8] H. L. Wang, *Design and Application of Successive Approximation for Optimal Control of Discrete Time-Delay Nonlinear Systems*. Qingdao, China: Ocean Univ. of China, 2007.
- [9] C. L. R. Mcghan, A. Nasir, and E. M. Atkins, "Human intent prediction using Markov decision processes," *J. Aerosp. Inf. Syst.*, vol. 12, no. 5, pp. 393–397, May 2015.
- [10] M.-H. Kim, H. Baik, and S. Lee, "Resource welfare based task allocation for UAV team with resource constraints," *J. Intell. Robot. Syst.*, vol. 77, nos. 3–4, pp. 611–627, Mar. 2015.
- [11] J. Brumbaugh-Smith and D. Shier, "An empirical investigation of some bicriterion shortest path algorithms," *Eur. J. Oper. Res.*, vol. 43, no. 2, pp. 216–224, Nov. 1989.
- [12] J. Granat and F. Guerriero, "The interactive analysis of the multicriteria shortest path problem by the reference point method," *Eur. J. Oper. Res.*, vol. 151, no. 1, pp. 103–118, Nov. 2003.
- [13] A. Skriver and K. Andersen, "A label correcting approach for solving bicriterion shortest-path problems," *Comput. Oper. Res.*, vol. 27, no. 6, pp. 507–524, May 2000.
- [14] M. Iori, S. Martello, and D. Pretolani, "An aggregate label setting policy for the multi-objective shortest path problem," *Eur. J. Oper. Res.*, vol. 207, no. 3, pp. 1489–1496, Dec. 2010.
- [15] E. Machuca, L. Mandow, J. P. De La Cruz, and A. Ruiz-Sepulveda, "A comparison of heuristic best-first algorithms for bicriterion shortest path problems," *Eur. J. Oper. Res.*, vol. 217, no. 1, pp. 44–53, Feb. 2012.
- [16] W. Yan, W. Longcheng, and J. Zhicheng, "Multi-objective vulcanization workshop scheduling based on improved harmony search algorithm," *J. Syst. Simul.*, vol. 30, no. 1, p. 139, 2018.
- [17] Z. Ning, Y. Li, P. Dong, X. Wang, M. S. Obaidat, X. Hu, L. Guo, Y. Guo, J. Huang, and B. Hu, "When deep reinforcement learning meets 5G-enabled vehicular networks: A distributed offloading framework for traffic big data," *IEEE Trans. Ind. Informat.*, vol. 16, no. 2, pp. 1352–1361, Feb. 2020.
- [18] Z. Ning, Y. Feng, M. Collotta, X. Kong, X. Wang, L. Guo, X. Hu, and B. Hu, "Deep learning in edge of vehicles: Exploring trirelationship for data transmission," *IEEE Trans. Ind. Informat.*, vol. 15, no. 10, pp. 5737–5746, Oct. 2019.
- [19] Z. Al-Makhadmeh and A. Tolba, "An Intelligence-based recurrent learning scheme for optimal channel allocation and selection in device-to-device communications," in *Circuits, Systems, and Signal Processing*. 2019, pp. 1–22.
- [20] Z. Ning, P. Dong, and X. Wang, "Deep reinforcement learning for vehicular edge computing: An intelligent offloading system," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 6, p. 60, 2019.
- [21] Z. Ning, J. Huang, X. Wang, J. J. P. C. Rodrigues, and L. Guo, "Mobile edge computing-enabled Internet of vehicles: Toward energy-efficient scheduling," *IEEE Netw.*, vol. 33, no. 5, pp. 198–205, Sep. 2019.
- [22] A. Tolba, "Content accessibility preference approach for improving service optimality in Internet of vehicles," *Comput. Netw.*, vol. 152, pp. 78–86, Apr. 2019.
- [23] A. Tolba and A. Altameem, "A three-tier architecture for securing IoT communications using vehicular dependencies," *IEEE Access*, vol. 7, pp. 61331–61341, 2019.

• • •