The modeling and solution of course of action generation based on IN-DEA

Lujun Wan¹, Wen Li², Jiangbin Dai¹ and Aqian Huang³

¹Air Traffic Control and Navigation College, Air Force Engineering University, No.1# Chang-le East Road Xi’an City, Shanxi Province, China.
²Equipment Management and Safety Engineering College, Air Force Engineering University, No.1# Chang-le East Road Xi’an City, Shanxi Province, China.
³Air and Missile Defense College, Air Force Engineering University, No.1# Chang-le East Road Xi’an City, Shanxi Province, China.
Email: L J.Wan pandawlj@126.com

Abstract. The course of action (COA) generation is the course of designing the task planning flow, which is one of the key supporting techniques while the military organization executing aerial operation. Due to the outstanding advantage of influence nets (IN), describing the causal logic between random variable, this proposal will introduce IN into the fast building mechanism of COA model. The COA generation model utilize influence nets is established. The probability propagation method is designed, which is based on influence intensity. An improved differential evolution algorithm (IDEA) is designed to solve the action policy optimized selected model. At last, the simulated result of aerial arrack campaign case show the action policy optimized selected in variety influence constants can improve the capability of cause-effect modeling, and the improved differential evolution algorithm has fine convergent and optimized capability.

1. Introduction
The Course of Action (COA) is a type of overall orderly behavior that the joint commanding organization of arms of the services conducts rational planning of the combat troops’ action, allocation of the task sequence, scheduling of military resources, to achieve its operational objectives according to the analysis results of the battle-field-related information; it is a crucial component in the joint operational plan. Because the action plan is essentially an optimal process for COA to achieve the overall mission, it can be seen as a directed evolution process with the goal of combat mission, meaning that according to the combat process, select a series of COA so that the strength of the organization or cluster can achieve the desired evolutionary results. Probabilistic network can represent the interrelationships between variables in a simple and effective way, and becoming a powerful tool with broad application prospects in the uncertainty reasoning system [1]. In the application of COA generation, the probability network has prominent advantages in describing the association relations of COA elements, so it has gradually become the main tool for COA modeling and production in recent years. Probably the most representative of the probabilistic network is markov networks (MN), bayesian networks (BN) and influence networks (IN). In the actual practice of COA problem modeling, BN and MN has two limitations; one is probabilistic reasoning difficulties and the other one is the need to specify a large number of conditional probabilities. The former limitation leads to low time efficiency of variable probability calculation and the latter limitation leads to the complex modeling under a small amount of information [2]. IN simplifies the process of...
enlightening knowledge, and is ideal for conditional probability modelling between node variables that require subjective estimation. Standard IN can only describe the static causal relationship between random variables, but fail to give a valid description of the condition that the impact of variables has certain time-dependent relationship. This paper focuses on the generation of uncertain COA of air operations, uses the influence network model to establish its mathematical model, and uses the Differential Evolution Algorithm (DEA) to optimize the model of COA.

2. Modeling of COA Based on Influence Networks

2.1. Course of Action and Model Elements

Enemy action (E): it is the action taken by the enemy according to the changes in the battlefield, associated with the action process; it is the uncontrollable resistance factor in the generation of COA. Suppose that the action set that the enemy may adopt is $E = \{E_1, E_2, \ldots, E_{NE}\}$ and $NE$ = $|E|$ is the action number of the enemy. Under normal circumstances, the enemy action is unpredictable and often randomly occurs; the commander needs to work with staff officers to give its occurrence probability in a certain stage of combat according to their own experience. Suppose that the combat process can be divided into $T$ stages in the implementation of joint air operations and the occurrence probability of enemy action $E_i$ in the battle $t_k$ stage is $P_k(E_i = 1) = P_k(E_i)$, $1 \leq i \leq NE$ and $P_1(E_i)$ is an interval value. To meet the computing needs, set $P_k(E_i)$ to follow the approximate normal distribution.

Basic Action (A): targeted at the mission of joint air operations, in the operational conception stage, the decision-makers should determine the set of basic action $A = \{A_1, A_2, \ldots, A_NA\}$ and $NA$ = $|A|$ is the number of basic action. If $P_k[A_1 = 1] = P_k(A_1) = 1$ and $1 \leq i \leq NA$, it shows that the basic action $A_1$ is carried out during the combat $t_k$ phase. In the unconstrained conditions, there will be the basic action $A$ of $2^T$ types of combinations. Set the basic action $A_i$ as $\pi(A_i) = \{a_i | 1 \leq j \leq \pi(A_i)\}$, $|\pi(A_i)|$ is the number of $A_i$, COA is the set of all the basic actions $A_i$ combinations $G = \{a_i | a_i = \pi(A_i), 1 \leq i \leq NA\}$ and the combat COA is $\Pi_G = \pi(A_1) \times \pi(A_2) \times \ldots \times \pi(A_{NA})$.

Desired combat goal: the ultimate operational goal of joint air operations and the combat effect is to make the greatest operational goals achieve joint probability. Set the expected goal set of combat in a certain battlefield situation as $D = \{D_1, D_2, \ldots, D_{ND}\}$ and $ND$ = $|D|$ is the number of combat goals.

Stage combat goal: the stage goal of the combat organization or entity in the process of achieving a combat goal. Suppose that in a certain battlefield situation, the stage goal set is $S = \{S_1, S_2, \ldots, S_{NS}\}$ and $NS$ = $|S|$ is the number of stage goal.

2.2. Generation Model of COA Based on the Influence Network

In order to effectively describe the influencing model based on influence intensity, the enemy-us action in joint air operations is deemed as the action in a unified way. Take the example of the influence of the action set $\{A_i\}_{i=1}^{IAi}$ on Action B to analyse the influence of the implementation of $\{A_i\}_{i=1}^{IAi}$ (Event True/False) on the implementation probability of Action B. Suppose $X^*$ is a random vector of 0 or 1 value of $n$ dimension and $X_j$ is the j component. If the action $A_j$ is implemented, $X_j = 1$. If the action $A_j$ is implemented, $X_j = 0$. Suppose $x^*$ is the real value of the random vector $X^*$, representing the specific state vector of the action set $\{A_i\}_{i=1}^{IAi}$. In order to conduct a quantitative description of the influence of the state vector $x^*$ on Action B, the following expression is used to describe the influence intensity $h_a(x^*)$ of the state vector $x^*$ on Action B:

$$h_a(x^*) = \begin{cases} 
1 & \text{Set } x^*, \text{Action B is bound to occur} \\
-1 & \text{Set } x^*, \text{Action B is bound not to occur} \\
0 & \text{The occurrence probability of Action B is unrelated to } x^* 
\end{cases}$$

(1)
For a given state vector $x^*$, set the occurrence probability of Action B is $P(B|\ x^*)$, then the quantification of the influence intensity is described as 1, $P(B)$ or 0 correspondingly. $P(B)$ is the probability of Action B. If $h_b(x^*) \in [-1,1]$, use the influence intensity $h_b(x^*)$ and $P(B)$ to deduce the conditional probability of the occurrence of Action B as:

$$
P(B|\ x^*) = \begin{cases} 
P(B) + h_b(x^*) \cdot [1 - P(B)] & h_b(x^*) \in [0,1] \\
P(B) + h_b(x^*) \cdot P(B) & h_b(x^*) \in [-1,0] 
\end{cases}
$$

(2)

Therefore, the influence network model can be formally described as $(V,E,C,A,B)$, of which V is the node set of the model, and the single node is a random variable of 0-1 value. $E$ is the directed edge set of the model and is the description of the cause-effect influence between nodes. $C$ is the set of cause-effect influencing intensity between nodes, $E \rightarrow \{[h^{(i)}(x_i = 1), h^{(i)}(x_i = 0)], h_i \in [-1,1]\}$. $A$, as the subset of $V$, represents the state of the influence action set $\{A_i\}_{i \in S}$ and its vector $X^*$. $B$ is the probability distribution that affects the state vector $X^*$ of the action set $\{A\}_{i \in S}$; the child node is the reference probability, and the root node is the prior probability. If $x = 1$, then $h(x) = h$; if $x = 0$, then $h(x) = g$. $h$ and $g$ represents the influencing intensity of Action A on Action B when it is implemented or not (true or false) respectively.

In this way, the cause-effect relationship model of various variables in the course of setting up COA through the influence network is to conduct quantization and modeling of the cause-effect relationship among the basic action, the enemy action, the stage goal and the expected operational target and to use the influence intensity value to quantify and describe the degree of cause-effect influence. In the model, the combat target node and the stage target node will influence each other in the course of the combat, and the current state of the node is associated with the state of the last moment. The parameter $P_r(V^t_i) (i = 1,2,..,T)$, $V_i \in (S|D)$ carried on the dashed line indicate the state value passed at the previous stage to the current stage, namely $P_r(V^t_i) = P_{r_i}(V_i)$.

Under the constraints of the entity resource and the time resource, the best combat goal, which is described as follows:

$$
\Omega = \arg \max_{\Omega \in \Omega} \left\{ \prod_{t=1}^{N_0} P[D(t) | TC(t_0), G] \right\}
$$

(3)

where $R(\Omega) \leq R_{RC}$ & $T(\Omega) \leq T_{RC}$, and $TC(t_0)$ is the initial state of the combat. The best combat goal $\Omega^*$, the initial state $TC(t_0)$ is finally converted to the goal state $TC(t_f)$; if the action program $\Omega$ is followed to implement the combat, the consumption of combat resources should be within the limitations of the entire resource amount $R_{RC}$. If the action program $\Omega$ is followed to implement the combat, the completion time needs to be within the time limit $T_{RC}$.

2.3. Probability Propagation Calculation Influencing the Network

Probabilistic propagation calculation is essentially a probabilistic reasoning mechanism [3]. The use of this reasoning mechanism needs to satisfy the requirement that the nodes of the probability-conduction association are independent. The probability of the root node changes with the process of the combat. In the new stage, the new probability will influence other nodes. Therefore, it is necessary to update the root node probability, calculate and update the probability of other nodes layer by layer.

Step 1: For the constructed influence network $G = (V, E)$, select a node $V_i$; use the Dijkstra method to calculate the minimum distance $W(V_i, V_d)$ to the target sub-node $V_d (d = 1,2,\ldots,D)$. Divide it into $L = \max_{d \in {D}} [W(V_i, V_d)]$ layers;

Step 2: if $t_i \leq t_f$, update the prior probability of the layer number $L$ and root node of the influence network, update the influencing intensity value between the root node and sub-node and turn to Step 3.
Otherwise, the calculation is ended, output the independent probability value and joint probability value of the sub-node;
  Step 3: Calculate the conditional probability distribution of the sub-nodes;
  Given the influence value \( \{h_i^{(x)}(x_i)\}_{i \in \mathcal{A}} \), the influence value of a two-dimensional random state vector \( X^o \) on the target sub-node \( B \) is \( h_i(x^o) \) and the conditional probability \( P(B \mid x^o) \) of the sub-node \( B \) is calculated as follows:

1. Polymerized negative influence intensity value \( I_{\text{negative}} = \prod_{i \in \mathcal{A}} (1 - |h_i^{(x)}(x_i)|) \);
2. Polymerized positive influence intensity value \( I_{\text{positive}} = \prod_{i \in \mathcal{A}} (1 - |h_i^{(x)}(x_i)|) \);
3. Merge the positive and negative influencing intensity values and get the calculation formula of the overall influence value \( h_i(x^o) = \left[ I_{\text{negative}} - I_{\text{positive}} \right] \max \left( I_{\text{negative}}, I_{\text{positive}} \right) \);
4. Calculation of the conditional probability \( P(B \mid x^o) \) of the sub-node \( B \) and the prior probability \( P(x^o) \) of the root node, according to the mutually independent assumption of the occurrence probability of the general formula and the root node, it can get \( P(B) = \sum_{x^o} [P(B \mid x^o) \cdot P(x^o)] \);

Step 4: Calculate the occurrence probability of a sub-node
Given the conditional probability \( P(B \mid x^o) \) of the sub-node \( B \) and the prior probability \( P(x^o) \) of the root node, the influence value of a two-dimensional random state \( x^o \) on the target sub-node \( B \), and return to Step 2; if completed, update the time period \([t_{i-1}, t_i]\), and return to Step 2.

3. Model Optimization and Solution Based on IDEA

3.1. Coding Form of the Individual
The feasible strategy set composed of the basic action \( A_j \) is \( \pi(A_j) = \{a_{j1}, a_{j2}, \ldots, a_{jd} \} \), \( r_d \) is the number of feasible strategies and the combination of the entire action is shown as \( G = \{a_i \mid a_i \in \pi(A_i), 1 \leq i \leq NA \} \).

The experience knowledge of decision-makers should be considered in the single basic action and the action number tends to be integer, so the real number is used to encode the individual; the individual includes \( NA \) real sequences \( \omega = (\omega_1, \omega_2, \ldots, \omega_d, \ldots, \omega_{da}) \), \( \omega_d \in \{1, 2, \ldots, r_d\} \), \( d = 1, 2, \ldots, NA \). Set \( \omega_d = k \) represent the \( k \) basic strategy of the strategy set \( G \) of \( A_j \), set \( p_i \in [0, 1] \) as individual learning probability.

3.2. Fitness Function
For the evaluation of the preferred COA, the ultimate goal of the joint air operations is to achieve the desired target intention. Therefore, the program to make the target effect reach the largest probability is the best COA, \( f = E_{\max}\{P(D(t_f))\} \). In the uncertain environment, in order to accurately estimate the natural state of the enemy action \( E \), the new random test observation can be used to gain the new signal and the simulation data is \( M \). The joint probability value \( f \) can be evaluated by calculating the mean \( \mu \) and variance \( \sigma^2 \) of \( M f \). For the joint air operations, it is desirable to maximize the joint probability of the combat target, that is, the signal-to-noise ratio (SNR) can be defined to integrate the mean \( \mu \) and variance \( \sigma^2 \), replacing the single joint probability \( f \), the fitness function as follows.

\[
\text{SNR} = -10 \cdot \log_{10}(\frac{1}{\mu^2} \cdot (1 + \frac{3\sigma^2}{\mu^2}))
\] (4)

3.3. Differential Variance
Regardless of whether the individual is feasible before and after the mutation, the individual \( \omega = (\omega_1, \ldots, \omega_d, \ldots, \omega_{da}) \) generates a new variant individual according to the linear
combination of the parent individuals, and the basic variation component is the difference variable of the parent individual. The following rules are used to generate new individuals:

\[ v_i = \omega_i + F \cdot (\omega_j - \omega_k) \quad i = 1, 2, \ldots, N_d \]  

(5)

\( v_i \) is the variant of the individual \( \omega_i \); \( \omega_j \), \( \omega_k \) and \( \omega_i \) are three different individuals randomly selected in the parent population, and \( \omega_i \) is not included. \( F \) is a real constant valued between the interval of \([0, 2]\), which can control the influence of the difference vector \((\omega_j - \omega_k)\). The size of the disturbance is positively correlated with the absolute value of the difference vector. When the population approaches the optimal value, the disturbance will also become smaller.

3.4. Cross Operation

For the time being, regardless of whether the individual can be decoded before and after the cross operation, this paper encodes the individual in integer form and performs the individual cross-operation with the following mechanism: the individuals that perform the cross operation are the mutation vector \( v = \{v_1, \ldots, v_d, \ldots, v_{N_d}\} \) and the target vector \( \omega = \{\omega_1, \ldots, \omega_d, \ldots, \omega_{N_d}\} \) respectively.

Through the cross rules in the formula (6), generate new individual \( \omega_j = \{\omega_{j1}, \ldots, \omega_{jd}, \ldots, \omega_{jN_d}\} \).

\[ \omega_j = \begin{cases} v_i & \text{rand}(0,1) \leq CR \\ \omega_i & \text{rand}(0,1) > CR \end{cases} \]  

(6)

In formula, \( \text{rand}(0,1) \) is the random value between 0 and 1; \( CR \) is a cross constant, the value of which is the real number between \([0, 1]\). The probability of cross increases as the \( CR \) value increases. The cross operation ensures that at least one phase element is exchanged between individuals, thus creating new individuals and preventing the evolution of the population from interruption.

3.5. Selection Operation with Learning

The selection operation in IDEA is a preference model based on the "greedy mechanism", and the size of the adaptive function of the individual is the only reference standard. If and only if the fitness value of the new vector individual \( v_i \) is better than that of the target vector individual \( \omega_i \), the new individual will be selected. Otherwise, the new generation of population \( \omega_j \) is still reserved for iterative calculation, and in the process, it continues to conduct variation and cross-operation of target vector. IDEA’s selection strategy is relatively harsh; parent individual and offspring candidate individual compete against each other and the preference is selected, making the offspring individuals always inferior to the parent individuals and promoting the population to seek for preference towards the best orientation in the shorted path. With the ranking result, seeking common gene of the whole chromosomes to get the fine individual and marked the former-rank chromosome gene bit. While individual learning probability \( p_i \geq \text{rand}(0,1) \), utilize the superiority common gene instead of the latter-rank gene bit, to enhance the algorithm convergence rate.

4. Simulation and Calculation Analysis

4.1. Case Scenario

In order to organize and apply all kinds of air power in a unified way, the scenario destroys most of the important goals on an island reef, forces them to abandon the island and thus gets rid of the barrier for the friendly forces to carry out landing operations. In order to ensure that the operational intent is achieved, the air force will carry out the following actions: \( A_1 \) is air cover on the air attack; \( A_2 \) is air cover on marine attack; \( A_3 \) is air ground assault of high value fixed island goal; \( A_4 \) is air-ground destruction of the ground’s fixed air defense weapons target; \( A_5 \) is air-found destruction of high maneuver air defense target; \( A_6 \) is air-ground assault of high-value island target; \( A_7 \) is air command and control operations; \( A_8 \) is regional cover; \( A_9 \) is electronic interference to the enemy; \( A_{10} \) is
organization of air refueling. The enemy counterattack may include: $E_i$ is the enemy aircraft carries out air interception to us; $E_2$ is the enemy ships organizes the attack of air plane; $E_3$ is the island air defense weapon system conducts attack of air plane; $E_4$ is the enemy organizes electronic anti-interference; $E_5$ is air interception of the enemy’s second echelon. The expected battle effect under the joint function of basic action and enemy action: $S_1$ is air-to-air combat effect; $S_2$ is air-to-sea combat effect; $S_3$ is attack effect of air-to-ground air defense firepower; $S_4$ is gathering of air formation; $S_5$ is destruction degree of island objective; $S_6$ is return and meet. Focus on the generation of COA and $D_1$, the cross probability is battle loss. Combined with staff experience and knowledge of experts in the field, the influence network is used to set up the COA generation model, as shown in Figure 1. According to the operational guidance, under the common analysis of the planning staff and experts in the field [4], divide the combat into six stages for organization and implementation: $[t_0,t_1], [t_1,t_2], [t_2,t_3], [t_3,t_4], [t_4,t_5]$ and $[t_5,t_6]$. 

![Influence network Model Generated by COA](image)

**Figure 1.** Influence network Model Generated by COA

### 4.2. Simulation and Calculation Analysis

The parameters of IDEA and standard DEA are set as follows: population size $N_p = 30$, calculation iteration number $n_{iterative} = 50$, scaling factor $F = 0.5$, cross constant $CR = 0.5$ and individual characteristic learning probability $p_l = 0.75$. The GA parameters: population size $N_p = 30$, the number of calculation iterations $n_{iterative} = 50$, the cross probability $p_c = 0.8$ and the individual mutation probability $p_m = 0.05$. The electronic interference action in Action $A_i$ has the best effect on $S_1$ and $S_2$ in the stage of $[t_0,t_1]$ and $[t_1,t_2]$. Subsequently, the enemy conducts electronic anti-interference and the interference effect gets worse in the stage of $[t_2,t_3]$ and $[t_3,t_4]$. This action will not be organized in the stage of $[t_4,t_5]$, leading to the same influencing intensity as before. The joint probability of the target functions $D_1$ and $D_2$ is the largest.

The objective function $f_i(s)$ is the independent occurrence probability of the node $D_i$; $f_j(s)$ is the independent occurrence probability of $D_j$, and the objective function $f_i(s)$ is the joint probability that the node $D_i$ occurs and $D_j$ does not occur, as shown in Fig.2 and Fig.3. The final target function value 0.8897 is better than the target function value 0.7875 of $\Omega_1$. Focusing on the generation of COA $\Omega_2$, in the case of operating for 30 times, as shown in Fig 4 and Fig 5, the three algorithms can find
the optimal under certain parameter setting conditions, although IDEA and DEA have more running
time than GA, and when approaching the global optimal solution, it is necessary to go through several
iterations to achieve the optimal, which is decided by DEA's own mechanism\(^5\), but compared to GA,
the global search ability of IDEA and DEA is stronger and they can seek for the global optimal
solution more quickly in the feasible solution space. Compared with the standard DEA, IDEA has
better average convergence efficiency through the feature learning of the dominant individual points.

![Graph](image1)

**Figure 2.** The target node probability of COA $\Omega_1$

![Graph](image2)

**Figure 3.** The target node probability of COA $\Omega_2$
5. Conclusion
Aiming at the characteristics of the generation of air operation COA, such as a large number of influencing factors and complex correlations, in order to solve the generation problem of uncertainty COA based on cause-effect relationship and obtain the best action strategy under different battle periods, an influence network based on influence intensity is used to set up the generation model of COA and the calculation step of corresponding probability propagation is given. Combining with the generation model features of COA, the coding, mutation and crossover strategy are designed based on IDEA. The feasibility and validity of IDEA are verified by simulating the case of air attack.

6. Acknowledgments
This work was financially supported by the National Natural Science Fund (61472443, 61573017), Program of ShanXi Natural Science Foundation (2016JQ6062).

7. References
[1] Liu J W, Li H E, Luo X L Study on Probabilistic Graphic Learning Technology Journal of Software 40 1025-44
[2] Lei T, Zhu C, Zhang W M Dynamic Bayes network based target-attacking course of action generation 2013 Proc. Int. Conf. on Intelligence and Security Informatics(Changsha) pp 1-5
[3] Zaidi A K, Papantoni K P 2010 Theory of influence networks. Journal of Intelligent & Robotic Systems, 60 457-91
[4] Yang D S, Zhang W M, Liu Z 2006 Mathematical Description and Solution Algorithm Research of Battle Task Plan Journal of Systems Engineering and Theory 6 26-34
[5] Gong D W Qin N N, Sun X Y 2010 Study on large-scale evolution of large population based on optimization problem of hybrid performance Control Theory & Applications 27 769-74