A Cascaded Algorithm Incorporating Knowledge Transfer Q-learning and Interior Point Method for Coordinated Operation of Integrated Energy System

Yuan Jianhua[1], *Li* Haofei[2] , *Ji* Chenyu[1], *Ji* Yu[1]

[1]Nantong Power Supply Company, State Grid Jiangsu Electric Power Company, 226000 Nantong, China
[2]Suzhou Huatian Power Technology Company, 215000 Suzhou, China

Abstract. The recent development of the Energy Internet has urged the conventional inefficient utilization of single energy to change towards the more developed energy usage of optimal dispatch of the integrated energy system. In this context, the joint optimization scheduling framework of integrated energy system is established based on the energy hub. Then a typical integrated energy system model is developed considering carbon emission and energy supply costs with valve point effect. To solve this non-linear problem with non-convex, discontinuously differentiable characteristic, the cascaded algorithm combined with the knowledge transfer based Q-learning algorithm and interior point method is applied on the model. Meanwhile, the efficiency is greatly improved by knowledge transfer. Case studies have been carried out on a 33energy hubs test system to verify the effectiveness of the proposed model and algorithm.

1 Introduction

Power system optimization scheduling refers to finding the optimal combination of controllable variables to meet system security constraints and load requirements [1]. Similarly, the joint optimization scheduling of integrated energy systems is also to achieve integrated energy system operation by adjusting controllable variables under various energy requirements and various energy network topologies.

In recent years, many scholars have studied the joint optimization scheduling of integrated energy systems. The energy hub model is proposed in [2]. Because of its simple principle, scalability and versatility, energy hub is widely used in the research of integrated energy systems. In addition, the probabilistic optimal power flow of the electro-mechanical interconnection system considering the correlation is established and the point estimation method based on the Nataf transform is used to calculate the probability optimal power flow in [3]. A joint economic operation model of electricity-gas interconnection system based on carbon trading mechanism optimizing by the sum of power generation cost and carbon cost as the objective function is established in [4]. The market equilibrium problem of energy hubs participating in the energy market is studied based on game theory in [5]. However, the above document solving methods are based on the interior point method [6], which solves the large-scale integrated energy system in a concentrated way. When the objective function is non-convex, it is likely to converge to the local optimal solution. An integrated energy system optimization model with the goal of minimizing energy

supply cost solved with multi-agent genetic algorithm is established in [7-8], but the solution speed is difficult to meet the requirements for large-scale systems. The existing solution models and methods are mostly based on a combination of power system and natural gas network, and most of the objective functions only consider the operating costs.

This paper establishes an integrated energy system optimization scheduling model taking into account energy costs and carbon emissions. In order to accurately describe the economic cost of the system, the energy supply cost considers the valve point effect of the unit. Besides, multi-objective optimization problem is transformed into single-objective optimization problem by membership function [9]. Secondly, the cascaded algorithm combined with the knowledge transfer based Q-learning algorithm and interior point method is proposed to solve the discontinuous and non-convex nonlinear problem. Finally, the simulation analysis of the 33 energy hubs test system is used to verify the effectiveness of the proposed model and algorithm.

2 Optimal joint dispatch of integrated energy system model

2.1 Objective function

In this paper, the objective function of integrated energy system in a single scheduling period is the energy supply cost target and the carbon emission target. To calculate the energy supply cost accurately, the valve point effect of the unit is considered in this paper.

$$W_e = \sum_{i \in \Omega_{\text{elec}}, \Omega_{\text{gas}}} (a_i P_{\text{in}i}^2 + b_i P_{\text{in}i} + c_i)$$
$$+ \sum_{i \in \Omega_{\text{elec}}} e_i \left| \sin(f_i(P_{\text{G}i} - P_{\text{G}i}^{\text{min}})) \right| \tag{1}$$

$$W_c = \sum_{i \in \Omega_{\text{elec}}, \Omega_{\text{gas}}} (\alpha_i P_{\text{in}i}^2 + \beta_i P_{\text{in}i} + \gamma_i) \tag{2}$$

where Ω_{elec} is the unit injection node set; Ω_{gas} is the gas source injection node set; $P_{\text{in}i}$ is the energy injection power, including unit injection and gas source; $P_{\text{G}i}$ is the active injection of the i-th node generator respectively; a_i, b_i and c_i are the energy cost coefficients; e_i and f_i are valve point effect characteristic parameter of units; α_i, β_i and γ_i are the energy carbon emission parameters.

2.2 Constraints

2.2.1 Power System Constrains

In an electricity transmission network, the optimal power flow is subject to both equality and inequality constraints. The equality constraints contain the AC power flow equations.

$$\begin{cases} P_{\text{G}i} - P_{\text{D}i} - V_i \sum_{j \in N_{\text{B}}} V_j(g_{ij}\cos\theta_{ij} + b_{ij}\sin\theta_{ij}) = 0 \\ \\ Q_{\text{G}i} - Q_{\text{D}i} - V_i \sum_{j \in N_{\text{B}}} V_j(g_{ij}\sin\theta_{ij} - b_{ij}\cos\theta_{ij}) = 0 \\ \\ P_{\text{G}i}^{\text{min}} \leq P_{\text{G}i} \leq P_{\text{G}i}^{\text{max}} \qquad i \in N_{\text{G}} \\ \\ Q_{\text{G}i}^{\text{min}} \leq Q_{\text{G}i} \leq Q_{\text{G}i}^{\text{max}} \qquad i \in N_{\text{G}} \\ \\ V_i^{\text{min}} \leq V_i \leq V_i^{\text{max}} \qquad i \in N_{\text{B}} \\ \\ P_l^{\text{min}} \leq P_l \leq P_l^{\text{max}} \qquad l \in N_{\text{L}} \end{cases} \tag{3}$$

where N_{G}, N_{B} and N_{L} are the number of generating units, nodes and branches, respectively; $P_{\text{D}i}$ and $Q_{\text{D}i}$ are the active and reactive loads at bus i, respectively; V_i and θ_i is the voltage magnitude and phase angle of bus i; g_{ij} and b_{ij} are the real and imaginary parts of the i-jth entry of the nodal admittance matrix, respectively.

2.2.2 Natural Gas Network Constrains

The steady state natural gas flow rate through a pipeline can be expressed as follow :

$$f_{mn} = \begin{cases} k_{mn}\sqrt{p_m^2 - p_n^2} & p_m \geq p_n \\ \\ k_{mn}\sqrt{p_n^2 - p_m^2} & p_m < p_n \end{cases} \tag{4}$$

where p_m and p_n are the gas pressures at buses m and n, respectively; f_{mn} is the natural gas flow of the pipeline from bus m to bus n.

Due to the friction in the pipelines, the gas pressure will decrease in gas transmission process. In order to ensure the reliability of gas transmission, a certain number of compression stations are installed in the natural-gas network, the model of which can be expressed as follow :

$$f_{\text{com}} = k_{\text{com}} f_{kn}(p_k - p_m) \tag{5}$$

where f_{com} denotes the power of natural gas consumed by the compression station; k_{mn} is the transmission coefficient of the pipeline from bus m to bus n.
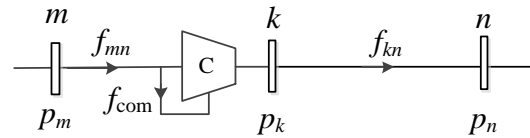
Fig 1. Model of compressor.

As same as the power system flow, the natural gas flow should satisfy the nodal equation.

$$(A + U)f + w - T\tau = 0 \tag{6}$$

where A and U are the natural gas network pipeline-node correlation matrix and the pressure station-node correlation matrix, respectively; f represents the branch flow vector; w denotes the node static flow; T denotes the pressurized station consumption matrix and node's association matrix;

Besides the natural gas transmission process is also subject to several constraints listed as follow :

$$\begin{cases} P_{gi}^{\text{min}} \leq P_{gi} \leq P_{gi}^{\text{max}} & i \in N_{\text{S}} \\ \\ p_i^{\text{min}} \leq p_i \leq p_i^{\text{max}} & i \in N_{\text{N}} \\ \\ R_i^{\text{min}} \leq \dfrac{p_k}{p_m} \leq R_i^{\text{max}} & i \in N_C \end{cases} \tag{7}$$

where R_i^{min} and R_i^{max} denote the upper and lower limits of compression station pressure ratio.

2.2.3 Energy hub Constrains

A typical energy hub is illustrated in the figure 2. The imported energy of an energy hub will be injected stored and converted to meet the requirement of energy loads. Taking the simple energy hub in Fig. 2 as an example, the energy flows between the input and output ports can be represented as:

$$\begin{cases} L_{\text{e}} = \eta_{\text{trans}}^{\text{e}} P_{\text{e}} + \eta_{\text{CHP}}^{\text{e}} v_{\text{ge}} P_{\text{g}} \\ \\ L_{\text{h}} = \eta_{\text{CHP}}^{\text{h}} v_{\text{ge}} P_{\text{g}} + \eta_{\text{Fur}} v_{\text{gh}} P_{\text{g}} \\ \\ L_{\text{g}} = (1 - v_{\text{ge}} - v_{\text{gh}}) P_{\text{g}} \end{cases} \tag{8}$$

where P_{e} and P_{g} are the electricity and natural gas input flows, respectively; L_{e}, L_{g} and L_{h} are the electricity, gas and heat output flows, respectively; v_{ge} and v_{gh} are the proportions of natural gas through the CHP and the gas furnace ; η is the efficiency of different energy conversion devices.
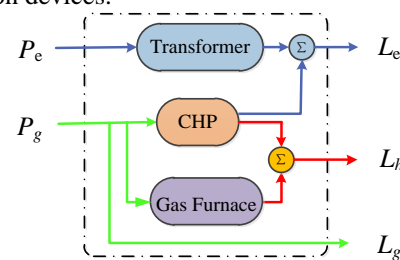
Fig 2. A typical energy hub.

2.3 Multi-objective Conversion

This paper uses the membership function to solve the multiple objective function problems. For the optimal target $W(x)$ minimizing the optimization goal, its membership can be expressed as

$$\mu= \begin{cases} 1 & W(x) \leq W_{min} \\ \dfrac{W_{max} - W(x)}{W_{max} - W_{min}} & W_{min} < W(x) < W_{max} \\ 0 & W(x) \geq W_{max} \end{cases} \quad (9)$$

In this paper, the middle part of the piecewise function is used to describe the relationship between μ and $W(x)$. According to the principle of maximum and minimum satisfaction, the multi-objective optimization problem is transformed into a single-objective problem for solving the maximum and minimum membership degree. The single-objective problem is described as:

$$\begin{cases} \max \quad \lambda \\ s.t. \quad \mu(W_i(x)) \geq \lambda \\ \qquad 0 \leq \lambda \leq 1 \\ \qquad \text{formula(1) and (4)~(10)} \end{cases} \quad (10)$$

This optimization model is a discontinuous and non-convex optimization model. In this paper, the cascading algorithm of knowledge transfer based Q-learning and interior point method is used to solve this model. That is, the active power of generators is taken as an action variable of Q-learning in the upper structure and solving the integrated energy system model with the interior point method in the lower structure. Besides, transfer learning of historical optimization information is introduced to accelerate the convergence speed. Since each interior point method uses the unit injection determined by the upper Q-learning as a constant, the lower interior point method can be directly solved.

3 Knowledge transfer based Q-learning algorithm

3.1 Action space discretization

The traditional Q-learning can only be used for discrete variable optimization, but the active power of the units in this model is continuous. In order to enable the Q-learning algorithm to optimize the continuous variable problem, this paper converts continuous variables into binary numbers as follow:

$$\begin{cases} (D_0^i, D_1^i, ..., D_{m-1}^i) = f\left(\dfrac{x_i - x_i^{min}}{x_i^{max} - x_i^{min}} \times 2^m\right) \\ x_i = \sum_{j=0}^{m-1} (2^{m-j} D_j^i)(x_i^{max} - x_i^{min}) + x_i^{min} \end{cases} \quad (11)$$

where m is the variable binary digits; f is a function converting decimal to binary; x_i is the i-th component of solution vector X; D_j^i denotes the j-th binary code of variable x_i.

3.2 Action space discretization

The traditional state-action $Q_{S \times A}$ matrix is mainly implemented by 'lookup'. As the variables increase, the number of actions grows exponentially, making it difficult for computers to store. Since each variable is represented by a plurality of interrelated binary code numbers, the 0-1 action selection of the previous binary coded bit can be used as the state of the next binary coded bit, so that the high-dimensional state-action Q matrix can be converted into multiple interrelated low-dimensional state and action chains. The state and action chains of all variables constitute the Q matrix of this problem.

3.3 Q-learning process

Firstly, according to the size of the Q matrix element, each unit of the corresponding binary code is selected, and the action selection is only 0-1 variable. After the action selection is completed, the code is converted into a continuous unit active integrated energy system optimization model, and the target is obtained by the interior point method. The value is converted into a corresponding action reward to update the Q matrix until the optimal strategy is obtained to maximize the reward return. When the integrated energy system optimization model converges to the infeasible solution, the action reward is zero. The action selection strategy is to choose in the binary space by roulette:

$$D_j^i = \begin{cases} 1 & r \geq P_j^i(D_{j-1}^i, 1) \\ 0 & r < P_j^i(D_{j-1}^i, 1) \end{cases} \quad (12)$$

$$P_j^i(s,a) = \frac{Q_j^i(s,a)}{\sum_{b \in D_j^i} Q_j^i(s,b)} \quad (13)$$

where r is a random number between [0,1].

The update strategy of State-action Q matrix element $Q_j^{i,(k)}$ is as follow:

$$Q_j^{i,(k+1)}(s^{(k)}, a^{(k)}) = (1 - a^{(k)})Q_j^{i,(k)}(s^{(k)}, a^{(k)}) + \\ a^{(k)}(R_j^{i,(k+1)} + \sigma \max_{b \in D_j^i} Q_j^{i,(k)}(s^{(k+1)}, b)) \quad (14)$$

where $R_j^{i,(k)}$ is the reward value of the state transition from $s^{(k)}$ to state $s^{(k+1)}$ after the $a^{(k)}$ action in the kth iteration; σ is the discount factor; $a^{(k)}$ is the learning factor of the kth iteration.

3.4 Q-learning process Knowledge transfer

The cascaded method combined with Q-learning algorithm and interior point method is often slower in solving joint optimization of integrated energy systems. This paper introduces knowledge transfer to improve the solution speed. Firstly, the algorithm obtains the optimal Q matrix under different sample loads by pre-learning the samples, and then uses the commonly used neural network data fitting method to obtain the relationship between the sample load and the optimal Q matrix. In the optimization process, the initial Q matrix under this load can be obtained by inputing the load information of the system to the neural network, and Q-learning can be used to optimize based on the initial Q matrix. Since the

system topology is unchanged and the load has similarity, the difference between the initial Q matrix and the optimal Q matrix is small. Therefore, knowledge transfer can be used to accelerate the convergence of the algorithm. The algorithm flow chart is presented in Fig. 3.
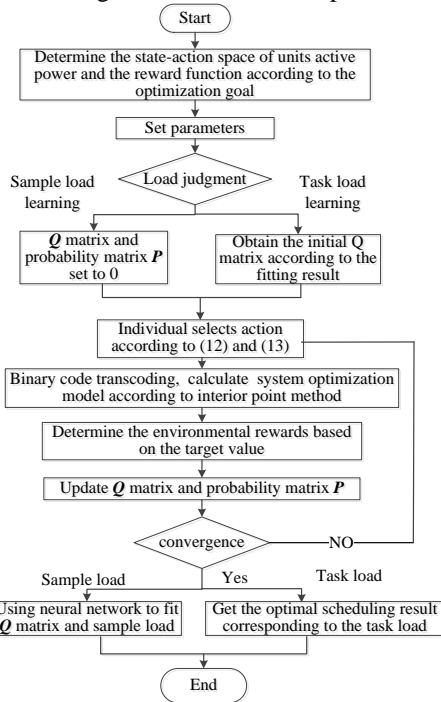

Fig 3. Process for optimization of integrated energy system dispatch with cascade algorithm.

4 Simulation and Analysis

4.1 Simulation model

The test system with 11 energy hubs includes a 14-nodes power system, a 20-nodes natural gas system, and 11 energy hubs shown in Figure 4. The energy hub is a typical model as shown in Figure 2. To illustrate the versatility of the model and algorithm, this paper expands the 11 energy hubs test system into 33 energy hubs test system. Each sub-area is connected by a tie line, and the load and unit position of each sub-area are different. The remaining topologies and parameters are the same.


Node number	2	3	4	5	6	9	10	11	12	13	14
Energy Hub Number	1	2	3	4	5	6	7	8	9	10	11
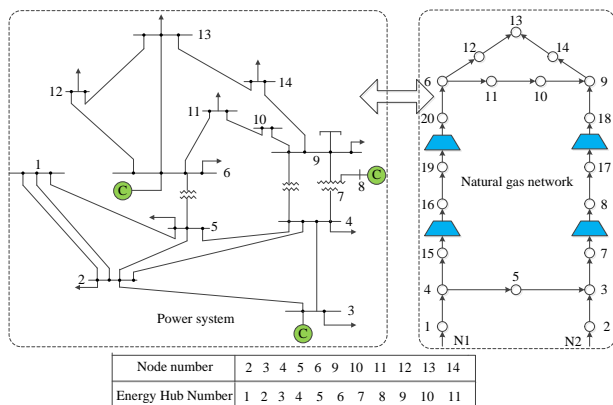
Fig 4. Configuration of 11energy hubs test system.

4.2 Study of optimization model

In this paper, three scheduling patterns are studied to illustrate the advantages of comprehensive consideration of energy supply costs and carbon emission target scheduling: pattern 1 is to schedule only with energy supply cost; pattern 2 is to schedule only with carbon emissions; pattern 3 is to schedule with energy supply cost and carbon emissions. It is assumed that the natural gas of each energy hub is used for gas supply and heat supply in this paper, and the proportion of electricity supply is zero. The sum of energy costs and carbon emissions are shown in Table 1. Optimization results under different patterns are shown in Figure 5.

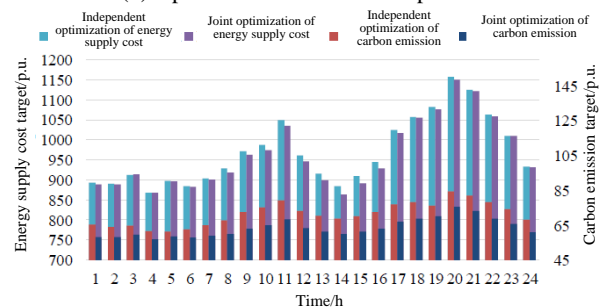**Table 1.** Objective sum of separate and joint optimization

Optimization method	Pattern1 W_e	Pattern2 W_e	Pattern3 W_e	W_e
Independent optimization	2.3201×10^4	1.7093×10^3	2.3259×10^4	1.7111×10^3
Joint optimization	2.2799×10^4	1.4679×10^3	2.3084×10^4	1.5210×10^3
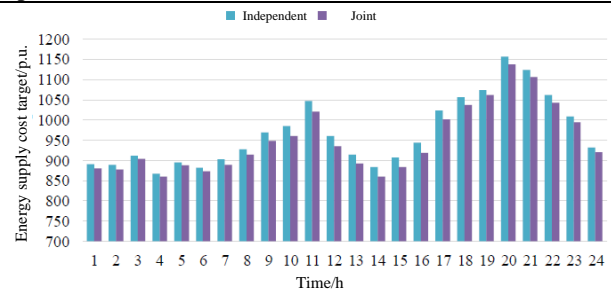
(a) Optimization results under pattern 1
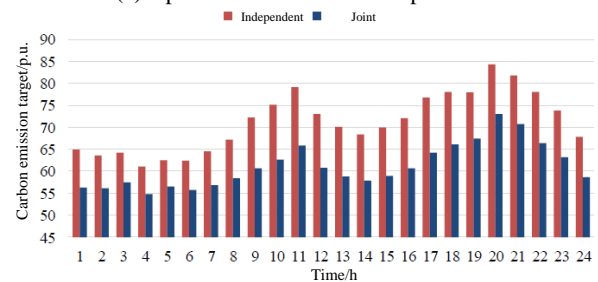
(b) Optimization results under pattern 2

(c) Optimization results under pattern 3

Fig 5. Optimization results under three patterns

Table.1 shows that the joint scheduling optimization target value of integrated energy system is smaller than that of the independent optimization in each pattern. It's clear that the integrated energy system optimization scheduling is

superior to the independent optimization scheduling of each energy network.

When considering a single target, joint optimization is better for each optimization period than independent optimization Therefore, the energy cost of joint optimization under pattern 1 is 1.73% lower than that of independent optimization. The energy supply cost of joint optimization under pattern 2 is 14.12% lower than that of independent optimization. But when considering multiple targets comprehensively, the pattern 3 optimization results show that the energy supply cost of joint optimization is higher than that of independent optimization at different time periods. Pattern 3 is aimed at maximizing the minimum membership. Therefore, the system will sacrifice carbon emissions at a certain cost, which will reduce its energy supply cost by only 0.75%. However, the proportion of carbon emissions fell by 11.11%.

4.3 Algorithm performance analysis

The iteration process of algorithm during the 11th period is as shown in Figure 5 to illustrate the convergence of knowledge transfer based Q-learning algorithm. Because the objective function is to maximize the minimum membership, the optimization curve is on the rise. The first picture in Figure 6 shows the ΔQ curve of the 10th generator.

The conclusions can be drawn in figure 5 as follow:① After initializing the Q matrix, the algorithm can obtain the optimization result of 0.5446 at the beginning. This shows that after the initial Q matrix is obtained by the knowledge migration method, the algorithm starts searching in a better action space ；② After the knowledge transfer, the convergence speed of the algorithm is greatly improved, and the algorithm can converge in 38 cycles taking 289s, indicating that the algorithm converges quickly and meets the calculation requirements after the knowledge transfer.
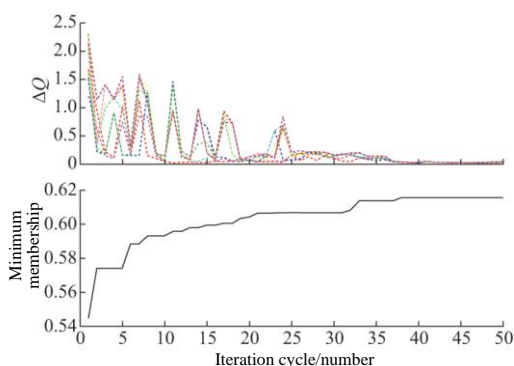
Fig 6. Convergence curves of objective value.

The cascade algorithms combined with MAGA, PSO and the interior point method are introduced in this paper [8] to compare with the knowledge transfer based Q-learning. Each algorithm runs 10 times. The number of MAGA and PSO populations is 100, the number of iteration cycles is 50, the MAGA crossover probability is 0.8 and the mutation probability is 0.9. The PSO learning factor is 1.5 and 1, inertia weight.is take as 0.5. The optimization results of the algorithms in the 11th period are shown in Table 2.

It can be seen that the cascade algorithms combined with artificial intelligence and interior point method can obtain good solutions in Table 2. The three optimization target values differ by 2.32%. MAGA algorithm works best, knowledge migration Q learns the second, and the PSO is slightly worse. The runtime of knowledge transfer based Q-learning is significantly less than other algorithms. It can be seen that knowledge transfer based Q-learning has acquired optimization results that are similar to other algorithms in a shorter iteration cycle due to the introduction of knowledge migration.

Table 2. Optimization results of all algorithms at period 11.

Algorithm	λ^{avr}	W_e^{avr}	W_c^{avr}	t^{avr}/s
MAGA	0.6244	1035.55	68.458	2123
PSO	0.6012	1036.58	68.586	2035
Knowledge transfer based Q-learning	0.6173	1035.90	68.457	297

5 Conclusion

In this paper, the joint optimization scheduling pattern of integrated energy system based on energy hub is proposed. To solve this non-linear problem with non-convex, discontinuously differentiable characteristic, the cascaded algorithm combined with the knowledge transfer based Q-learning algorithm and interior point method is applied on the model. The simulation results illustrate that the proposed model shows its effectiveness and the run time of cascade algorithm combined with knowledge transfer based Q-learning and interior point method is significantly less than other algorithms to the problem solving for non-continuous and non-convex problems.

References

1. C. Peng, H. Sun. CSEE,**29**,34 (2009)
2. M. Geidl, G. Koeppel, P. Favre-perrod. IEEE Power Energy Mag, **5**,1(2007).
3. G. Sun, S. Chen, Z. Wei, AEPS, **39**,21(2015)
4. Z. Wei, S. Zhang, G. Sun, AEPS, **40**,15(2015)
5. Y. Wang, J. Zhao, F. Wen, AEPS, 39,21(2015)
6. N. Karmarkar, COMBINATORICA, **4**, 4(1984)
7. M. Moeini-aghtaif, A. Abbaspour, M. Fotuhifiruza-bad, PWRS, **29**,2(2014)
8. M. Moeini-aghtaif, P. Dehghanian, M. Fotuhifir-uzabad, STE, **5**,2(2014)
9. Y. Mao, W. Liu, P. Zhang, Proceedings of the CSU-EPSA, **12**,6(2000)