A Mixed Integer Model for Large-Scale New Energy Medium-Term Operation Problem

Tieqiang Wang\textsuperscript{a}, Fang Liu\textsuperscript{b,*}, Xin Cao\textsuperscript{a}, Chenjun Sun\textsuperscript{a}, Zhice Yang\textsuperscript{b} and Jue Wang\textsuperscript{b}

\textsuperscript{a}State Grid Hebei Electric Power Company, Shijiazhuang, China.
\textsuperscript{b}Computer Network Information Center, Chinese Academy of Science, Beijing, China

Abstract

In China, new energy is developing rapidly. In recent years, new energy power generation has been installed with explosive growth. However, the coordination problem between new energy penetration capability and the operation mode of the system has not been solved. Especially in the ‘Three North’ areas, new energy is severely limited. As a result, the large-scale new energy medium-term operation optimization algorithm and its parallelization are very urgent. This paper established a mixed integer model for the large-scale new energy medium-term operation problem, and proposed a new method to simplify the 0-1 constraints. Since the most commonly used software has some limitations on solving our mixed integer programming (MIP) problem, we developed a parallel algorithm library (CMIP) V2.0 of our own intellectual-property rights and exploited the parallelism of the algorithm for better performance. Preliminary numerical experiments show that CMIP V2.0 can solve the new energy medium-term operation optimization problem, at least as well as the commercial software CPLEX and the open source software SCIP.

Keywords: new energy; mixed integer model; CMIP

(Submitted on October 20, 2017; Revised on November 22, 2017; Accepted on November 30, 2017)

© 2017 Totem Publisher, Inc. All rights reserved.

1. Introduction

In the last decade, wind power and photovoltaic (PV) power have been developed rapidly to deal with the worsening energy problem in China, especially in the ‘Three North’ areas. However, the coordination problem between new energy penetration capability and the operation mode of the system remains one of the problems in urgent need to be solved in the field. In order to assure the security of the power system, the power generation from new energy sources should be curtailed by the peak load dispatching the electric network, power system load flow, voltage, and power quality, etc. \cite{4,10,12,19,22}. The aim of our research on large-scale new energy medium-term operation problems is to maximize the use of new energy in a certain amount of installed capacity, constrained by the power balance constraints and a variety of operational constraints related to the conventional thermal units. This paper investigates the optimization techniques for large-scale new energy medium-term operation problems in the development of new-energy industry to fill the blank. The provinces with great quantities of installed capacities of new energy generation units need this kind of technique urgently, especially in the ‘Three North’ areas. In the context of high proportion of the conventional thermal units and large demand for winter heating in these areas, the research on large-scale new energy medium-term operation optimization algorithm and its implementations are very urgent.

The contributions of this paper include:

- A mixed integer model is established for the large-scale new energy medium-term operation problem, and a new method to simplify the 0-1 constraints is proposed.
- The parallel algorithm library (CMIP) V2.0 of our own intellectual-property rights is developed and improved for better performance.

* Corresponding author.

E-mail address: liufang@sccas.cn
The rest of this paper is organized as follows. Section 2 briefly summarizes the main constraints for large-scale new energy medium-term operation problem. Section 3 describes our mixed integer model and presents our new method to simplify the 0-1 constraints. Section 4 describes the implementation details of our parallel algorithm library (CMIP) V2.0 and improvement for better performance. Experimental results and a general discussion on performance are presented in section 5. Finally, we conclude our paper and propose future work in section 6.

2. Influence Factors and Characterization Method for Large-scale Random New Energy Production Simulation Modeling

The optimization of the new energy production simulation model can be considered as a large-scale mixed integer programming problem. The goal of optimization is to maximize the new energy output power, including the capacity of wind power and photovoltaic power, which is equivalent to maximizing the sum of new energy power in all regions of all periods of time. The objective function is depicted by the below formula:

$$\max \sum_{t=1}^{T} \sum_{n=1}^{N} \text{newenergy}(t, n)$$

(1)

The integration of large-scale new energy by month or year is influenced by three main factors: conventional unit, new energy, and the system. The conventional unit contains condensing steam thermal power unit, back-pressure thermal power unit, exhaust thermal power unit, and reservoir type hydropower unit. The modeling of the conventional units needs to consider the upper and lower bounds of the optimized power, the ramp rate, and the total power as depicted by the below formulas:

$$0 \leq P(t, j) \leq [TP_{\text{max}}(j) - TP_{\text{min}}(j)] \times X(t, j)$$

(2)

$$P(t + 1, j) - P(t, j) \leq \text{UpRamp}(j)$$

(3)

$$P(t, j) - P(t + 1, j) \leq \text{DownRamp}(j)$$

(4)

$$TP(j) = TP_{\text{min}}(j) \times X(t, j) + P(t, j)$$

(5)

where, TP_{\text{max}}(j) and TP_{\text{min}}(j) represents the upper and lower bounds of the output of the jth unit's in the tth time period. UpRamp(j) and DownRamp(j) represent the ramp up and ramp down rate of the jth unit, respectively. X(t,j) is a binary variable which represents the status of the jth unit in the tth time period, where 0 means the halted state of unit and 1 means the running state. Formula (2) constrains the range of the value of P(t,j), which is dependent on the value of X(t,j), and the difference value of TP_{\text{max}}(j) and TP_{\text{min}}(j). Formula (3) and (4) constrain the ramp rate of the power unit, while formula (5) shows how to compute the total power of the power units.

In addition, it is necessary to make a further consideration about the minimum time for each unit to start and stop running. X, Y and Z are three major constraint variables of the starting and stopping stage of the power units, where X(t,j) is defined above, and Y(t,j) and Z(t,j) are defined as below:

- Y(t,j): a binary variable which represents the starting state of the jth unit in the tth time period, where 1 means the units is starting, otherwise, the variable is set to 0.
- Z(t,j): a binary variable which represents the stopping state of the jth unit in the tth time period, where 1 means the units is stopping, otherwise, the variable is set to 0.

The power units cannot start or stop frequently because the constrains of the physical characteristics of the units, and the energy consumption and operation cost of the units. Thus, the minimum time during the starting and stopping stages needs to be further considered. The constrain functions are defined by formula (6) and (7) as follows:

$$Y(t, j) + Z(t + 1, j) + Z(t + 2, j) + \ldots + Z(t + k, j) \leq 1$$

(6)

$$Z(t, j) + Y(t + 1, j) + Y(t + 2, j) + \ldots + Y(t + k, j) \leq 1$$

(7)

where, k is determined by the parameters of the minimum start and stop time of the power unit. The value of k represents the minimum duration time of the starting and stopping stages. The start or stop operation is completed in only one time period, and the starting and stopping stages should be maintained for at least k time periods. Thus, formulas (6) and (7) can be explained in another way as follows: when a power unit starts or stops in the tth time period, then neither start nor stop operation is allowed during time interval [t, t + k]; there are no constrains for the rest time periods.
In addition, the logical constrains of the above three key constraint factors X, Y, Z are depicted by these formulas as below:

\[
\begin{align*}
X(t, j) - X(t-1, j) - Y(t, j) + Z(t, j) &= 0 \\
-X(t, j) - X(t-1, j) + Y(t, j) &\leq 0 \\
X(t, j) + X(t-1, j) + Y(t, j) &\leq 2 \\
-X(t, j) - X(t-1, j) + Z(t, j) &\leq 0 \\
X(t, j) + X(t-1, j) + Z(t, j) &\leq 2
\end{align*}
\] (8)

The above set of equations and inequalities compose the logical constraints of the start-stop operations and the running status, which guarantee that the status of each unit in the power system is logically correct.

Our MIP model is established according to the above seven 0-1 constrains listed in formula (2) - (8) and six other factors: generating capacity of hydropower, cross-regional transmission capacity, regional load balance, spinning reserve constrains, wind power and photovoltaic power [20], and increasing the number of the constraints to thirteen. The original MIP can be transformed to the sum of a subset of problems and be solved using the Zero-One Programming method [7, 8, 17, 18, 21, 23]. During the optimizing process, there are usually tens of thousands of variables and constrains in each iteration. Due to the long simulation time periods, the large number of power units and the corresponding constrain conditions, the model is optimized quite slowly. In this regard, this paper proposed an effective simplified model, which can greatly improve the speed of optimization.

3. Simplifications and Deformations of the Logical Constraints of the Start-Stop Model

As described above, the minimum duration times of start-up and shutdown status are logically constrained by formula (3-1) and (3-2). In addition, the three major constraint variables X, Y and Z about the start-stop and running status are constrained by formula (8). These constrains can be deformed and simplified according to property (1) and (2) as below.

Property (1). Constraints depicted by formulas (6) - (8) can be simplified to formula (9).

\[
\begin{align*}
Y(t, j) + Z(t, j) + Z(t+1, j) + Z(t+2, j) + \ldots + Z(t+k, j) &\leq 1 \\
Z(t, j) + Y(t, j) + Y(t+1, j) + Y(t+2, j) + \ldots + Y(t+k, j) &\leq 1 \\
X(t, j) - X(t-1, j) - Y(t, j) + Z(t, j) &= 0
\end{align*}
\] (9)

In fact, property (1) can be proven to be correct under the conditions that the values Y(t,j) and Z(t,j) are all correct with four different combinations of values of X(t,j) and X(t-1,j) as below:

- If X(t,j)=1 and X(t-1,j)=1, it can be proven that Y(t,j)=Z(t,j)=1 or Y(t,j)=Z(t,j)=0 according to formula (8-1). Furthermore, according to formula (8-3), Y(t,j)=Z(t,j)=0.
- If X(t,j)=0 and X(t-1,j)=0, it can be proven that Y(t,j)=Z(t,j)=1 or Y(t,j)=Z(t,j)=0 according to formula (8-1). Furthermore, according to formula (8-3), Y(t,j)=Z(t,j)=0.
- If X(t,j)=1 and X(t-1,j)=0, it can be proven that Y(t,j)=1 and Z(t,j)=0 according to formula (8-3).
- If X(t,j)=0 and X(t-1,j)=1, it can be proven that Y(t,j)=0 and Z(t,j)=1 according to formula (8-3).

Property 2. Constraints depicted by formulas (6) - (8) can be further simplified to formula (10):

\[
\begin{align*}
\sum_{t \in \mathcal{U}} Z(t, j) + \sum_{t \in \mathcal{V}} Y(t, j) &\leq 1 \\
X(t, j) - X(t-1, j) - Y(t, j) + Z(t, j) &= 0
\end{align*}
\] (10)

Through the deformations above, the number of constraints is decreased from 13 to 8, and the number of variables in the model is also decreased consequently, making the speed of solution greatly improved.

4. Model Verification and Analysis

In recent years, the provincial power load steady growth, and annual growth rate has been more than 5%. However, the grid peak valley rate is increasing year by year, which makes it difficult to control the power grid operation and make the safety
of power grid operation decrease year by year. The province electric heating unit installed is more than 70% in all capacity. Over the winter period, the capacity of peak load regulation falls sharply. "Keep High Peak" and "pressure trough" are two big difficulties. Low load in power system becomes the main factor of clean energy, which is given in province power grid.

Take the annual input into the model to carry out the simulation, and get the output of the annual model, such as what is shown in Table 1.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>intwindgross</td>
<td>2595934.808</td>
<td>Total acceptance of wind power (MW·h)</td>
</tr>
<tr>
<td>intPVgross</td>
<td>2476645.68</td>
<td>Total acceptance of PV power (MW·h)</td>
</tr>
<tr>
<td>intnewenergygross</td>
<td>5072580.488</td>
<td>Total acceptance of new energy power (MW·h)</td>
</tr>
<tr>
<td>newenergyerrorgross</td>
<td>677.1084</td>
<td>Total new energy limits (MW·h)</td>
</tr>
</tbody>
</table>

The table shows that the new energy limited power is 677.1084 MW. New energy totally includes wind power generation and photovoltaic power generation, and its value is 5072580.488 MW. According to the previous two values, we can calculate that the new energy power limits rate is 0.0133484%. From the result, abandoning wind and light is focused on the 6th and 7th days in the 6th weeks. There is virtually no limit to power generation in the province.

The output model results are calculated, and the output results of the annual model are shown in Table 2.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PVaveragepower</td>
<td>283.4988187</td>
<td>Photovoltaic average power(MW)</td>
</tr>
<tr>
<td>Pvuusetime</td>
<td>1042.556988</td>
<td>Photovoltaic use hours (h)</td>
</tr>
<tr>
<td>windaveragepower</td>
<td>297.1537097</td>
<td>Wind average power(MW)</td>
</tr>
<tr>
<td>windusetime</td>
<td>2008.797511</td>
<td>Wind use hours(h)</td>
</tr>
</tbody>
</table>

Average use of photovoltaic and wind power can be calculated as shown in Formula 11. The meaning is in the annual model and annual average power of new energy generation per hour. As shown from the results of wind and solar power average power:

\[
\text{New\_energy\_average\_power} = \frac{\text{New\_energy\_acceptance}}{\text{Total\_hours\_per\_year}} \quad (11)
\]

The calculations of hours of photovoltaic and wind power use the formula that is shown in Formula 12. The meaning of it is under the current wind power and photovoltaic capacity, the number of hours required for a result's electricity, reflecting the photovoltaic and wind power can actually use of time:

\[
\text{New\_energy\_use\_hours} = \frac{\text{New\_energy\_acceptance}}{\text{New\_energy\_installed}} \quad (12)
\]

According to historical data, in 2016, wind power utilized 1746 hours in a year; compared to the wind forecast in 2017, wind power utilization increased by 262 hours and the forecast error is 13 percent.

There are three main reasons for the analysis of the results:

1. due to the units in the province power grid, the new energy sources proportion is very small such as wind power and photovoltaic power units, so the new energy's electricity overall power ratio is small. The new energy power can be used basically the whole year.

2. because there are 4 units of 250 MW of pumped storage power unit, when the user electricity trough, the extra electricity through the water potential energy can be stored in the form of pumped storage power plant. When the user peaks, the potential energy of the water power reduces the electricity load frequent start-stop conventional unit when there is a change, maintaining security and stability of the unit.

3. the causes of the prediction error are mainly the change of wind power capacity in 2017, uncertainty, and wind power installed capacity prediction on the high side, causing wind power prediction to increase and ultimately affect the prediction of wind power use hours on average.
5. Implementation and improvement on the Parallel Algorithms Library CMIP for Mixed Integer Programming Problem

The mixed integer programming (MIP) model[1] can be built according to the characterization of the piecewise linear functions listed above using a week-by-week model or rolling model. Mixed integer programming problem is NP hard, and the difficulty lies in the exponential explosion of the integer variables [2,3]. The most commonly used software for mixed integer programming problems include the commercial software CPLEX (IBM ILOG CPLEX Optimization Studio)[5], GUROBI[9], open-source software SCIP[3,14], and the domestic software CMIP[20] with independent intellectual property rights, which is developed based on SCIP[15].

There are some problems to solve the mixed integer programming model described above using the softwares mentioned above. We analyzed the processing of the middle 19th weeks using CPLEX as an example (total time is 19.28s):

- Input the parameters of the model (0.31s);
- Preprocess the model (7.17s);
- Search for the feasible solution x (0.16s);
- Use heuristic algorithm[11] linear programming (LP) relaxation[6,13,16] in the root node to get a set of solutions X (8.30s);
- Find the result using RIN and Crossover heuristic algorithm (including increasing cuts) [11] based on feasible solution x and relaxation solution X in the root node;
- Branch & cut until the final solution is obtained.

The above processing procedure using CPLEX has some problems:

- Though the processing procedure is fast, the result might be not quite satisfactory. Taking the same 19th week discussed above as an example, the difference between the relaxation solution and the target solution is 534.62%, which will lead to an increase in the cost of the heuristic algorithm;
- The RIN and Crossover heuristic algorithm is effective, but neither of them can decide whether the added cutting edge is truly effective;
- It is nearly useless for our model to use the native automatic tuning function of CPLEX.

To solve the MIP problem using our model, the optimization of the domestic self-developed CMIP software is based mainly on the following model characteristics:

- The effective pre-process and heuristic algorithm is similar every week, and the fluctuation only appears in some specific locations (such as at the beginning and end of the heating period);
- Constraints on the power units are similar between each week, and the number of variables overlap between columns is small. So once the variables are fixed, the problem size will be reduced dramatically;
- The change of our model is great only at the beginning and end of the heating period;
- Due to the dependence between weeks, the solving process of each week is not suitable to be parallelized.

Similarly, we still take the 19th week discussed above as an example to show how to improve and implement solving processes using the steps as below (total time is 12.85s):

- Input the parameters of the model (0.42s, including the copy time);
- Preprocess the model (5.00s);
- Only use the linear programming (LP) relaxation in the root node to get a set of solutions X (5.3s);
- Use the improved RENS method to get a simplified subproblem subMIP;
- Preprocess and solve the subMIP model (1.27s, including the above step);
- Substitute the result of the subMIP to the original problem to solve it again, and validate the result before output of the final solution (0.86s).

The computing time of the process can be reduced by automatic tuning according to the first characteristic of the model.

The RENS algorithm in CMIP can be improved according to the second characteristic of the model. The solving process can break ahead after a near-optimal solution is found. In this way, those slow processes for the optimal solution in some situations can be prevented and thus can accelerate the process to solve the problem. The pseudo-code of the program can be listed as follows:
Algorithm 1 The pseudo-code of the improved RENS algorithm

```plaintext
1. for (i = 0; i < nbinvars + nintvars; i++) {
2.   /* Get the current LP solution for each variable */
3.   solval = CMIPVarGetLPSol(vars[i]);
4.   if (IsFeasIntegral(solval)) {
5.     /* Change the bounds within numerical tolerance for integral sub-problem */
6.     lb = floor(solval+0.5);
7.     ub = lb;
8.   } else if (binarybounds) {
9.     /* Change the bounds to the nearest integers for binary sub-problem */
10.    lb = FeasFloor(solval);
11.   ub = FeasCeil(solval);
12. } else {
13.    /* Otherwise just copy the original bounds */
14.    lb = VarGetLbGlobal(vars[i]);
15.   ub = VarGetUbGlobal(vars[i]);
16. }
17. /* Change the bounds */
18. ChgVarLbGlobal(subvars[i],lb);
19. ChgVarUbGlobal(subvars[i],ub);
20. }
```

The model can be solved using the week-by-week model and the rolling model. In the former mode, the data of each week is processed one by one, while in the latter mode, the data of each day and the following two days is processed day by day. To improve the performance, the most time-consuming parts of the software are paralleled using pthread and MPI, which can support thousands of threads at most.

6. Performance Test

The test on the improved CMIP (Version 2.0) software described above is performed on the LAN environment with one Linux server and one test computer. The SCIP and CPLEX software were also tested on the same server which was accessed by the test computer via SSH. The test criterion is that if more than 90% of test cases are successful without any serious system defects while running, it means the test has passed. Otherwise, the test has failed.

On the platform described above, we design 8 test cases for both the week-by-week model and rolling model (averaged each three days) according to three main performance metrics of the solution: performance of the heuristic algorithm, the accuracy of the optimal solution, and the performance of preprocessing. The software was tested in the MPI mode, and the number of processes was set to 8. The results are listed in Table 3.

The results listed in Table 3 shows that all of the test cases are executed successfully. For the week-by-week model, CMIP V2.0 has gained more than 30% improvement compared to CPLEX and SCIP considering the metrics of algorithm performance. For rolling model, CMIP V2.0 has similar or slightly better general performance compared with CPLEX. In addition, CMIP V2.0 can get the feasible solution accurately, which exactly coincides with the default solution considering the metrics of the accuracy of the optimal solution.

<table>
<thead>
<tr>
<th>Performance Metrics</th>
<th>Test Description</th>
<th>Test Model</th>
<th>Other Software</th>
<th>CMIP V2.0</th>
<th>Improvement</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heuristic Algorithm</td>
<td>Solving MIP</td>
<td>Week-by-week model</td>
<td>CPLEX</td>
<td>784.4s</td>
<td>481.19s</td>
<td>38.64%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rolling model</td>
<td>CPLEX</td>
<td>1740.05s</td>
<td>693.7s</td>
<td>2.68%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Week-by-week model</td>
<td>SCIP</td>
<td>54253.7s</td>
<td>481.19s</td>
<td>112.75</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Solving LP problems</td>
<td>Rolling model</td>
<td>Preset Solution</td>
<td>Gain 49 solutions</td>
<td>Pass</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Week-by-week model</td>
<td>CPLEX</td>
<td>271.28s</td>
<td>119.38s</td>
<td>55.99%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Week-by-week model</td>
<td>SCIP</td>
<td>172.37s</td>
<td>30.74%</td>
<td>Pass</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rolling model</td>
<td>CPLEX</td>
<td>271.85s</td>
<td>269.95s</td>
<td>0.69%</td>
</tr>
</tbody>
</table>

7. Conclusions

This paper analyzed the theory and arithmetic MIP problems for large-scale new energy medium-term operations. Simplifications and deformations of the seven logical constraints of the start-stop model can significantly reduce the scale of
the model. The model was verified to be effective by the test results of the province power grid. The model was implemented based on the CMIP library, and it was accelerated according to the MIP model which is set up by the characterizations of piecewise linear functions.

Acknowledgments
This work is jointly supported by State Grid Corporation Technology Research Program (SGHE0000DKJS1600191) and grant 2017YFB0202202 in the National Major Research High Performance Computing Program of China.

References

Tieqiang Wang received the Ph.D. degree from North China Electric Power University. His research interests include high performance computing, and the optimization of the new energy operation problem.

Fang Liu received the Bachelor degree from Nanjing University, Nanjing, China, in 2005, and the Ph.D. degree from the Institute of Software, Chinese Academy of Science. Her research interests include high performance computing, and general purpose computing on GPUs.

Xin Cao received the Master degree from Tsinghua University. His research interests include automation of electric power systems.

Chenjun Sun received the Ph.D. degree from North China Electric Power University. His research interests include high performance computing, and the optimization of the new energy operation problem.
Zhice Yang received the Bachelor degree from Beijing Jiaotong University, and now is a postgraduate student in Computer Network Information Center, Chinese Academy of Science. His research interests include high performance computing and automation of electric power systems.

Jue Wang received the Bachelor degree from North China Electric Power University, and the Ph.D. degree from University of Science and Technology Beijing. His research interests include high performance computing, machine learning, and automation of electric power systems.