Chinese Journal of Chemical Engineering 26 (2018) 1605–1612

Contents lists available at ScienceDirect

Chinese Journal of Chemical Engineering

journal homepage: www.elsevier.com/locate/CJChE

Article

A decision tree based decomposition method for oil refinery scheduling☆

Xiaoyong Gao [1,2], Dexian Huang [2,*], Yongheng Jiang [2], Tao Chen [3]

[1] Institute for Ocean Engineering, China University of Petroleum, Beijing 102249, China
[2] Department of Automation, Tsinghua University, Beijing 100084, China
[3] Department of Chemical and Process Engineering, University of Surrey, Guildford GU2 7XH, UK

ARTICLE INFO

Article history:
Received 26 August 2017
Accepted 12 October 2017
Available online 8 November 2017

Keywords:
Refinery scheduling
Decision tree
C4.5
Decomposition method

ABSTRACT

Refinery scheduling attracts increasing concerns in both academic and industrial communities in recent years. However, due to the complexity of refinery processes, little has been reported for success use in real world refineries. In academic studies, refinery scheduling is usually treated as an integrated, large-scale optimization problem, though such complex optimization problems are extremely difficult to solve. In this paper, we proposed a way to exploit the prior knowledge existing in refineries, and developed a decision making system to guide the scheduling process. For a real world fuel oil oriented refinery, ten adjusting process scales are predetermined. A C4.5 decision tree works based on the finished oil demand plan to classify the corresponding category (*i.e.* adjusting scale). Then, a specific sub-scheduling problem with respect to the determined adjusting scale is solved. The proposed strategy is demonstrated with a scheduling case originated from a real world refinery.

1. Introduction

Refinery scheduling has attracted increasing concerns because of increasing environmental standard, intense global competition, and volatile market demand. Fruitful and valuable reports have been published in the last decade.

The dominant research on oil refinery scheduling focus on the integrated model based optimization. The general framework for refinery planning and scheduling were proposed by Pinto and co-workers [1–3]. Jia and Ierapetritou proposed a continuous time formulation for refinery scheduling problem and spatially decomposed it into three sub-problems [4,5]. In the work of Dogan and Grossmann [6], a decomposition method for simultaneous integrated planning and scheduling problem is proposed. Wu and Ierapetritou [7] proposed a hierarchical approach for production planning and scheduling, in which uncertainty is considered. More recently, Shah and Ierapetritou incorporated logistics into the short term scheduling problem of a large scale refinery and proposed a comprehensive integrated optimization model for the scheduling problem of production units and end-product blending problem [8]. Gao *et al.* considered the impact of variations in crude oil on scheduling for complex reaction processes with a crude classification based multimodal method [9] and also nonlinear characteristics between process model and operating variables for hydro-upgrading processing units using piecewise linear approximation [10], where an integrated model is resulted. Göthe-Lundgren *et al.* [11] proposed a multi-fixed yield model in terms of several predefined operating states, in which too many binary variables are introduced in the scheduling model, and it is hard to be solved in reasonable time. This method was adopted in Luo and Rong's report [12], in which a hierarchical approach to short-term scheduling is proposed and the binary variables in optimization are significantly reduced. In addition to these specific studies, some excellent reviews have been published in this area, such as Floudas and Lin [13], Bengtsson and Nonås [14], Shah *et al.* [15], Joly [16], Harjunkoski *et al.* [17].

However, the formulated integrated large scale mixed integer linear/nonlinear programming problem is really hard-to-solve and seldom successful industrial application report is published. Based on the prior knowledge from field experience, there is no need for plant-wide adjustment in most cases. Moreover, the current scheduling optimization leads to the plant-wide adjustment, which may result in the unreasonable difference between the computed and real yield, due to the existence of process dynamics and long process transition caused by plant-wide arbitrary scheduling adjustments. The arbitrary schedule disturbs the stability of control system, which in turn makes scheduling model inconsistent with the actual process [18]. Though some published reports are concentrated on the seamless integration between scheduling and its subside process control system. However, the solution seamlessly mingles these two parts and formulates it as an integrated mathematical programming problem [19,20], which is computationally expensive and hard to obtain a practicable result for a real industrial scale problem

☆ Supported by the National Natural Science Foundation of China (21706282, 21276137, 61273039, 61673236), Science Foundation of China University of Petroleum, Beijing (No. 2462017YJRC028) and the National High-tech 863 Program of China (2013AA 040702).
* Corresponding author.
E-mail address: huangdx@tsinghua.edu.cn (D. Huang).

https://doi.org/10.1016/j.cjche.2017.10.006
1004-9541/

in tolerable time. The scheduling task is to stabilize the process as fast as possible and satisfy the demand in most economical way when confronted with the unit break-down or demand changes.

To accelerate the computation of MILP and obtain schedule in reasonable time for real world industrial applications, we propose a decision tree based heuristic decomposition method for refinery scheduling, taking decision tree based upper decision making layer and optimization layer. Decision tree determines the adjusting scale to decrease the process transient time caused by resulted schedule. The lower optimization layer solves a corresponding smaller scale (comparing with the integrated plant-wide scheduling problem) sub-scheduling problem, which not only accelerate the solving speed, but also decrease the schedule's dynamic transition time when carrying out (*i.e.* the accumulated process dynamic time of all adjusted units for all time periods) and effectively avoid the more complex computation in terms of the model taking schedule dynamics into consideration.

The paper is organized as follows. Decision tree based heuristic decomposition method is proposed in the following Section 2. The mathematical model in terms of decision output is detailed in Section 3. Case study is given in Section 3.1.4 to validate the effectiveness of the proposed strategy and the conclusion is drawn in Section 3.1.5 at the end of the paper.

2. Decision Tree Based Decomposition Method

2.1. Decision tree based decomposition method for refinery scheduling optimization

As aforementioned in the above analysis, refinery scheduling is always formulated as a large scale mixed integer linear or nonlinear programming problem, which is hard to be solved in reasonable time. Moreover, the resulted schedule from the comprehensive model integrating all the plant processes worsen the performance of control because of the existence of process dynamics, which is discarded in the most current dominant researches. In this paper, we proposed a novel decision tree based decomposition method to accelerate the computation and improve the schedule quality. (See Fig. 1.)

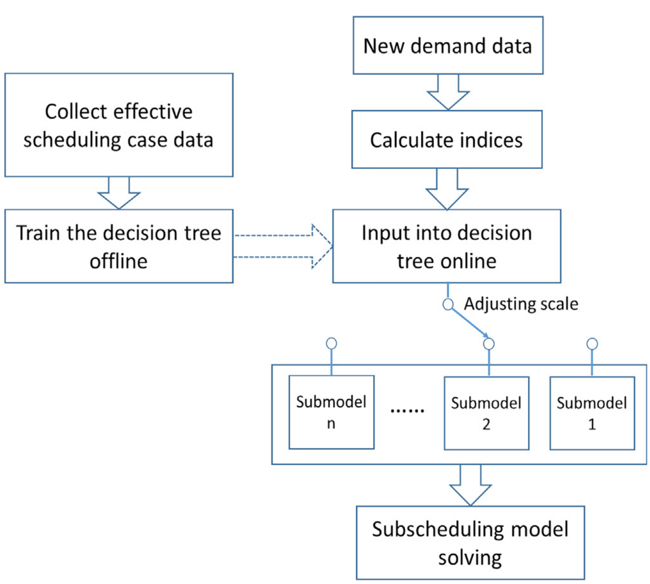

Fig. 1. Schematics of the decision tree based decomposition method.

Based on the field experience, there is no need to adjust all plant-wide units for most cases. When confronted with minor changes in finished product oil demand or a unit's small emergence flaw, to adjust part of downstream units may satisfy the target, which is faster to stabilize from the schedule's operation changes. Fast response and quick stability implies that the plant is fast to supply the on-spec components and final product as anticipated in the schedules. Otherwise, the on-spec components and final products are hard to guarantee, because there is no guarantee for the components yields during the dynamic transition process. Moreover, the schedule obtained from this strategy also benefits the MPC and its bottom control system and it makes MPC's operation more stable.

2.2. Decision tree design

The decision tree structure, of which the input attributes variables are most important, is of vital importance to the classification ability.

1) For fuel refinery, the plant product oil conversion rate, defined as the ratio between the final product and the feed crude, has the global influence on the decision, which means that the whole plant, together with the crude oil blending, should be taken in the scheduling optimization if it exceeds the normal value too far. Hence, based on the intrinsic analysis, the plant product oil conversion rate and its delta-value are factors influencing scheduling decision.
2) From the viewpoint of product oil demand quantity, the gasoline diesel demand ratio and its delta-value, delta gasoline diesel ratio, determine the operation of primary processing units (PPU) and secondary processing units (SPU). For example, the PPU and SPU will be driven to the mode benefiting for gasoline component production if there is an increase in gasoline demand during some period.
3) From the quality viewpoint, the demand changes between high-quality or premium oil products and regular oil products must have some intrinsic impact on the decision for operation units. The blending components from different operation units have their specific quality characteristics. For gasoline components, research octane number (RON) of straight run gasoline from crude oil distillation unit is often low, while gasoline components from FCCU or HC are often high with the same high unsaturated olefins content. The reformate from catalytic reforming unit (CRU) is the ideal component for premium and clean gasoline, which has highest RON, low metal and sulfur content. If the premium gasoline demand increases, then units producing high RON and clean gasoline components should be adjusted to increase their processing capacity to yield more premium gasoline components. Then, the ratio between high quality and ordinary oil product and its delta value are impact factors, which should be taken as input variable of decision tree.

To make a summary, the input variables of decision tree are grouped up in Table 1. One point should be highlighted is the difference between PGR and HGR, PDR and LDR. Some large scale refinery may supply fuels for different regions, where different fuel specification standards are adopted. PGR and PDR denote the ratio between high-spec gasoline and normal-spec gasoline, ratio between high-spec diesel and normal-spec diesel respectively. HGR calculates the ratio of high grade gasoline

Table 1
Schematics of the decision tree based decomposition method

Input attribute variables	Definition or description
PR	Product ratio between product and crude
ΔPR	Delta product ratio between product and crude
GDR	Ratio of gasoline and diesel
ΔGDR	Delta ratio of gasoline and diesel
PGR	Premium gasoline ratio
ΔPGR	Delta premium gasoline ratio
HGR	High grade gasoline ratio
ΔHGR	Delta high grade gasoline ratio
PDR	Premium diesel ratio
ΔPDR	Delta premium diesel ratio
LDR	Low-freezing diesel ratio
ΔLDR	Delta low-freezing diesel ratio

without regards of specifications, while LDR gives the ratio of ultra-low freezing diesel among.

The adjusting scale categories are treated as decision tree output. According to the flow chart of our investigated refinery, the adjusting scale is classified into ten classes, which is listed in Table 2. The detailed classification may be case dependent, and this method is applicable.

Table 2
Tree output

Output classes	Adjusting scale(scope)
C1	Product oil blenders
C2	Product oil blenders, CRU
C3	Product oil blenders, CRU, FCC gasoline hydro-desulfurization, FCC gasoline etherification
C4	Product oil blenders, diesel hydro-treating units
C5	Product oil blenders, diesel hydro-treating units, CRU
C6	Product oil blenders, diesel hydro-treating units, FCC gasoline hydro-desulfurization, FCC gasoline etherification
C7	Product oil blenders, diesel hydro-treating units, CRU, FCC gasoline hydro-desulfurization, FCC gasoline etherification
C8	Product oil blenders, diesel hydro-treating units, CRU, FCC gasoline hydro-desulfurization, FCC gasoline etherification, FCCU fractionator
C9	Product oil blenders, diesel hydro-treating units, CRU, FCC gasoline hydro-desulfurization, FCC gasoline etherification, FCCU (reactor, fractionator)
C10	Product oil blenders, diesel hydro-treating units, CRU, FCC gasoline hydro-desulfurization, FCC gasoline etherification, FCCU (reactor, fractionator), crude distillation unit

Detailed mathematical model based on discrete time representation method is given in Section 3.

3. Mathematical Model

Under decision tree, the output determines the adjusting scale. Each scale is related with a particular sub-model. According to unit type in refinery, five sub-models are resulted and given as follows. Then, the relationship between decision tree output and the corresponding sub-model is given.

3.1. Mathematical sub-models

3.1.1. Sub-scheduling model I (SSM-1)

Only blenders are adjusted in Step I. The objective function to maximize profit is given in SSM-1(1), and SSM-1(2)~SSM-1(5) represent the constraints of blenders, similar to those established for the single-step decision making approach (Section 3.1). SSM-1(6) and SSM-1(7) are constraints for component oil tanks. In this step, all the upstream units are operated in their steady state. Therefore, the inflow QI_u of a specific component oil tank in SSM-1(6), and the property of component oil $PRO_{s,u,p}$ in SSM-1(2), are all fixed.

Objective:

$$\max \sum_{t\in TP}\left(\sum_{u\in PTK} o_u \cdot \mathrm{QO}_{u,t} - \sum_{u\in TK_I} s_u \cdot \mathrm{INV}_{u,t}\right) \qquad \mathrm{SSM}-1(1)$$

Constraints:

$$QI_{u,t}\sum_{m\in M_u} z_{u,m,t}\cdot PRO^{\mathrm{MIN}}_{s',u,m,p} \le \sum_{s\in IS_u} Q_{s,u,t}\cdot PRO_{s,u,p} \le QI_{u,t}\sum_{m\in M_u} z_{u,m,t}\cdot PRO^{\mathrm{MAX}}_{s',u,m,p} \quad \forall u\in BLD, t\in TP, p\in P, s'\in OS_u \qquad \mathrm{SSM}-1(2)$$

$$r^{\mathrm{MIN}}_{s,u,m}\cdot\sum_{s\in IS_u} Q_{s,u,t} \le Q_{s,u,t} \le r^{\mathrm{MAX}}_{s,u,m}\cdot\sum_{s\in IS_u} Q_{s,u,t} \qquad \mathrm{SSM}-1(3)$$

$$\sum_{m\in M_u} z_{u,m,t} \le 1 \quad \forall u\in BLD, t\in TP \qquad \mathrm{SSM}-1(4)$$

$$QI_{u,t} = \sum_{s\in IS_u} Q_{s,u,t} \quad \forall u\in U_s, t\in TP \qquad \mathrm{SSM}-1(5)$$

$$QO_{u,t} = \sum_{s'\in OS_u} Q_{s',u,t} \quad \forall u\in U_s, t\in TP \qquad \mathrm{SSM}-1(6)$$

$$\sum_{t\in T_{u,b}} QO_{u,t} \ge DD^{\mathrm{MIN}}_{u,b} \quad \forall u\in PTK, b\in B_u, t\in TP \qquad \mathrm{SSM}-1(7)$$

$$INV_{u,t+1} = INV_{u,t} + QI_u - QO_{u,t} \quad \forall u\in \mathrm{BTK_I}, t\in TP \qquad \mathrm{SSM}-1(8)$$

$$INV^{\mathrm{MIN}}_u \le INV_{u,t} \le INV^{\mathrm{MAX}}_u \quad \forall u\in BTK_I, t\in TP \qquad \mathrm{SSM}-1(9)$$

In SSM-1, the decision variables are the inflow and outflow of all the tanks and blenders ($Q_{s,u,t}$ and $Q_{s',u,t}$), and the blending sequence $z_{u,m,t}$.

3.1.2. Sub-scheduling model II (SSM-2)

Blenders and HUPUs are scheduled in sub-model SSM-2. The model SSM-2 is summarized as follows. The objective function in SSM-2(1) includes product oils revenues, inventory cost of tanks in step II, and the operating cost of the adjusted HUPUs.

Objective:

$$\max \sum_{t\in TP}\left(\sum_{u\in PTK} o_u \cdot QO_{u,t} - \sum_{u\in TK_{II}} s_u . INV_{u,t} - \sum_{u\in HUPUs} QI_{u,t}\cdot OPC_{u,t}\right) \qquad \mathrm{SSM}-2(1)$$

Constraints:

The constraints include SSM-1(3–7) because blenders are also considered, and the following:

$$QI_{u,t}\sum_{m\in M_u} z_{u,m,t}\cdot PRO^{\mathrm{MIN}}_{s',u,m,p} \le \sum_{s\in IS_u} Q_{s,u,t}\cdot PRO_{s,u,p,t} \le QI_{u,t}\sum_{m\in M_u} z_{u,m,t}\cdot PRO^{\mathrm{MAX}}_{s',u,m,p} \quad \forall u\in BLD, t\in TP, p\in P, s'\in OS_u \qquad \mathrm{SSM}-2(2)$$

$$Q_{s',u,t} = QI_{u,t}\cdot YLD_{s',u,t} \quad \forall u\in HUPUs \qquad \mathrm{SSM}-2(3)$$

$$YLD_{s',u,t} = YLD^0_{s',u} + f_{s',u}(\Delta PRO_{s',u,t,p}) \quad \forall u\in HUPUs \qquad \mathrm{SSM}-2(4)$$

$$PRO_{s',u,t,p} = PRO_{s,u,t,p} + \Delta PRO_{s',u,t,p} \quad \forall u\in HUPUs \qquad \mathrm{SSM}-2(5)$$

$$OPC_{u,t} = OPC^0_u + f^c_u(\Delta PRO_{s',u,t,p}) \quad \forall u\in HUPUs \qquad \mathrm{SSM}-2(6)$$

$$\Delta PRO^{\mathrm{MIN}}_{s',u,p} \le \Delta PRO_{s',u,t,p} \le \Delta PRO^{\mathrm{MAX}}_{s',u,p} \quad \forall u\in HUPUs \qquad \mathrm{SSM}-2(7)$$

$$QF^{\mathrm{MIN}}_u \le QI_{u,t} \le QF^{\mathrm{MAX}}_u \quad \forall u\in HUPUs \qquad \mathrm{SSM}-2(8)$$

$$INV_{u,t} = INV_{u,t-1} + QI_{u,t} - QO_{u,t} \quad \forall u\in ITK_{II}, t\in TP \qquad \mathrm{SSM}-2(9)$$

$$INV_{u,t} = INV_{u,t-1} + QI_u - QO_{u,t} \quad \forall u\in BTK_{II}, t\in TP \qquad \mathrm{SSM}-2(10)$$

$$INV^{\mathrm{MIN}}_u \le INV_{u,t} \le INV^{\mathrm{MAX}}_u \quad \forall u\in TK_{II} \qquad \mathrm{SSM}-2(11)$$

Since HUPUs are considered for scheduling, the properties of component oils drawn from HUPUs, *i.e.* $PRO_{s,u,p,t}$ in SSM-2(2), are variable. The decision variables are (i) the inflows and outflows of all units (*i.e.* HUPUs, tanks, blenders) being considered ($Q_{s,u,t}$ and $Q_{s',u,t}$), (ii) the blending sequence $z_{u,m,t}$, and (iii) the operating condition of HUPUs, expressed by $\Delta PRO_{s',u,t,p}$.

3.1.3. Sub-scheduling model III (SSM-3) under Step III

Profit maximization is the objective function, and the only difference compared with SSM-2 is that the operating costs of SPUs are included.

Objective:

$$\max\sum_{t\in TP}\left(\sum_{u\in PTK}o_u\cdot QO_{u,t}-\sum_{u\in TK_{III}}s_u\cdot INV_{u,t}-\sum_{u\in SPUs}\sum_{m_f\in M_u^f}QI_u\cdot z_{u,m_r,m_f,t}\cdot OPC_{u,m_r,m_f}-\sum_{u\in HUPUs}QI_{u,t}\cdot OPC_{u,t}\right)\quad \text{SSM}-3(1)$$

Constraints:

The constraints include SSM-1(3–7), SSM-2(2–8), and the following:

$$Q_{s',u,t}=QI_u\cdot\sum_{m_f\in M_u^f}z_{u,m_r,m_f,t}\cdot YLD_{s',u,m_r,m_f}\quad\forall u\in SPUs\quad \text{SSM}-3(2)$$

$$\sum_{m_f\in M_u^f}z_{u,m_r,m_f,t}=1\quad\forall u\in SPUs\quad \text{SSM}-3(3)$$

$$PRO_{s',u,p,m_r,m_f}=const\quad\forall u\in SPUs\quad \text{SSM}-3(4)$$

$$OPC_{u,m_r,m_f}=const\quad\forall u\in SPUs\quad \text{SSM}-3(5)$$

$$QF_u^{\text{MIN}}\le QI_{u,t}\le QF_u^{\text{MAX}}\quad\forall u\in SPUs\quad \text{SSM}-3(6)$$

$$INV_{u,t}=INV_{u,t-1}+QI_{u,t}-QO_{u,t}\quad\forall u\in ITK_{III}\quad \text{SSM}-3(7)$$

$$INV_{u,t}=INV_{u,t-1}+QI_u-QO_{u,t}\quad\forall u\in BTK_{III}\quad \text{SSM}-3(8)$$

$$INV_u^{\text{MIN}}\le INV_{u,t}\le INV_u^{\text{MAX}}\forall u\in TK_{III}\quad \text{SSM}-3(9)$$

In this step, the fractionators of SPUs are also included in schedules, formulated in SSM-3(2–7). The inflows of SPUs (QI_u) are fixed in SSM-3(1) and SSM-3(2), because in this step, only main fractionators are considered, while the reactors of SPUs are still in steady state.

The decision variables in this step are (i) the inflows and outflows of HUPUs, tanks and blenders ($Q_{s,u,t}$ and $Q_{s',u,t}$), and the outflows $Q_{s',u,t}$ of SPUs; (ii) the blending sequence $z_{u,m,t}$; (iii) the operating condition of HUPUs, *i.e.* $\Delta PRO_{s',u,t,p}$; and (iv) the fractionation mode m_f of SPUs ($z_{u,m_r,m_f,t}$).

3.1.4. Sub-scheduling model (SSM-4)

In this step, the reactors and main fractionators of SPUs are all adjusted. The variable inflow ($QI_{u,t}$) for SPUs is formulated in SSM-4(1,2); the reaction mode (m_r) and fractionation mode (m_f) are adjusted in SSM-4(2,3). The decision variables are (i) the inflows and outflows of the units (SPUs, HUPUs, tanks, blenders) considered in Step IV: $Q_{s,u,t}$ and $Q_{s',u,t}$, (ii) the blending sequence $z_{u,m,t}$, (iii) the operating condition of HUPUs ($\Delta PRO_{s',u,t,p}$), and (iv) the reaction mode m_r and fractionation mode m_f of SPUs ($z_{u,m_r,m_f,t}$).

Objective:

$$\max\sum_{t\in TP}\left(\sum_{u\in PTK}o_u\cdot QO_{u,t}-\sum_{u\in TK_{IV}}s_u\cdot INV_{u,t}-\sum_{u\in HUPUs}QI_{u,t}\cdot OPC_{u,t}-QI_{u,t}\cdot\sum_{u\in SPUs}\sum_{m_r\in M_u^r}\sum_{m_f\in M_u^f}z_{u,m_r,m_f,t}\cdot OPC_{u,m_r,m_f}\right)\quad \text{SSM}-4(1)$$

Constraints:

The constraints include SSM-1(3–7), SSM-2(2–8), SSM-3(4–6), and the following:

$$Q_{s',u,t}=QI_{u,t}\cdot\sum_{m_r\in M_u^r}\sum_{m_f\in M_u^f}z_{u,m_r,m_f,t}\cdot YLD_{s',u,m_r,m_f}\quad\forall u\in SPUs\quad \text{SSM}-4(2)$$

$$\sum_{m_r\in M_u^r}\sum_{m_f\in M_u^f}z_{u,m_r,m_f,t}=1\quad\forall u\in SPUs\quad \text{SSM}-4(3)$$

$$INV_{u,t}=INV_{u,t-1}+QI_{u,t}-QO_{u,t}\quad\forall u\in ITK_{\text{IV}}\quad \text{SSM}-4(4)$$

$$INV_{u,t}=INV_{u,t-1}+QI_u-QO_{u,t}\quad\forall u\in BTK_{IV}\quad \text{SSM}-4(5)$$

$$INV_u^{\text{MIN}}\le INV_{u,t}\le INV_u^{\text{MAX}}\forall u\in TK_{IV}\quad \text{SSM}-4(6)$$

3.1.5. Sub-scheduling model 5 (SSM-5)

In this sub-model, PPUs are also considered in the objective function in SSM-5(1), and the corresponding constraints are formulated in SSM-5(2–6). The decision variables are (i) the inflows and outflows of all units (*i.e.* PPUs, SPUs, HUPUs, tanks, blenders): $Q_{s,u,t}$ and $Q_{s',u,t}$, (ii) the blending sequence ($z_{u,m,t}$), (iii) the operating condition of HUPUs ($\Delta PRO_{s',u,t,p}$), (iv) the reaction mode m_r and fractionation mode m_f of SPUs ($z_{u,m_r,m_f,t}$), and (v) the operating mode m of PPUs ($z_{u,m,t}$).

Objective:

$$\max\sum_{t\in TP}\left(\sum_{u\in PTK}p_u\cdot QO_{u,t}-\sum_{u\in TK_{IV}}s_u\cdot INV_{u,t}-\sum_{u\in HUPUs}QI_{u,t}\cdot OPC_{u,t}-\sum_{u\in SPUs}\sum_{m_r\in M_u^r}\sum_{m_f\in M_u^f}QI_{u,t}\cdot z_{u,m_r,m_f,t}\cdot OPC_{u,m_r,m_f}-\sum_{u\in PPUs}\sum_{m\in M_u}QI_{u,t}\cdot z_{u,m,t}\cdot OPC_{u,m}\right)\quad \text{SSM}-5(1)$$

Constraints:

SSM-1(3–7), SSM-2(2–8), SSM-3(4–6), SSM-4(2,3)

$$Q_{s',u,u'',t}=QI_{u,t}\cdot\sum_{m\in M_u}z_{u,m,t}\cdot YLD_{s',u,m}\quad\forall u\in PPUs\quad \text{SSM}-5(2)$$

$$\sum_{m\in M_u}z_{u,m,t}=1\quad\forall u\in PPUs\quad \text{SSM}-5(3)$$

$$PRO_{s',u,u'',p,m}=const\quad\forall u\in PPUs\quad \text{SSM}-5(4)$$

$$OPC_{u,m}=const\quad\forall u\in PPUs\quad \text{SSM}-5(5)$$

$$QF_u^{\text{MIN}}\le QI_{u,t}\le QF_u^{\text{MAX}}\quad\forall u\in PPUs\quad \text{SSM}-5(6)$$

$$INV_{u,t}=INV_{u,t-1}+QI_{u,t}-QO_{u,t}\quad\forall u\in ITK_{\text{V}}\quad \text{SSM}-5(7)$$

$$INV_{u,t}=INV_{u,t-1}+QI_u-QO_{u,t}\quad\forall u\in BTK_{V}\quad \text{SSM}-5(8)$$

$$INV_u^{\text{MIN}}\le INV_{u,t}\le INV_u^{\text{MAX}}\quad\forall u\in TK\quad \text{SSM}-5(9)$$

3.2. Decision tree outputs and sub-models

It is clear that there is a particular sub-model related with every adjusting scale, *i.e.* the decision tree's output. The detailed relationship between the decision tree outputs and sub-models is listed in Table 3. Clearly, for adjusting scale C1, only blenders are optimized. So, the corresponding sub-model SSM-1 is to be solved. What to be highlighted is that the adjusted unit set is different for tree outputs from C2 to C7 because different HUPUs are included for optimization, though the same sub-model form, *i.e.* SSM-2 is used.

Table 3
Relationship between decision tree outputs and sub-models

Decision tree outputs (adjusting scale)	Sub-models to be solved
C1	SSM-1
C2	SSM-2
C3	SSM-2
C4	SSM-2
C5	SSM-2
C6	SSM-2
C7	SSM-2
C8	SSM-3
C9	SSM-4
C10	SSM-5

4. Case Study

To verify the effectiveness of the proposed decision tree based decomposition method, a scheduling case study is provided here.

4.1. Refinery flowchart

A case refinery originated from a real world refinery in china is provided here, whose flow chart is depicted as in Fig. 2. Several crude oils are blended as mixed feed for crude distillation units (*i.e.* atmospheric and vacuum distillation unit, CDU and VDU in Fig. 2). Light gasoline from CDU is further upgraded in CRU to produce reformate for premium and high grade gasoline production. The first side draw, the CDU kerosene, is partly deep processed in DHT and partly distributed as diesel blending component. The second side draw of CDU, the CDU diesel, together with the partly CDU kerosene and light vacuum gas oil (LVGO in Fig. 2), is upgraded in DHT to decrease the heavy mental and sulfur content. CDU residue is processed in VDU. High vacuum gas oil (HVGO in Fig. 2) is further processed in FCCU, where heavy fraction is cracked into two light fractions, FCC gasoline and diesel. FCC gasoline is upgraded in FCC light gasoline etherification and FCC heavy gasoline desulfurization separately. The gasoline components and diesel components are then blended into the final on-spec gasoline and diesel respectively in blenders.

4.2. Scheduling case study

4.2.1. Decision tree training

Two hundred eighty-five cases are collected to train the decision tree, where 200 cases are used for train and the rest 85 cases are for test. In this research, C4.5 algorithm, is adopted, in which confidence factor is set as 0.25, and ten-fold cross validation method is used to train the tree classifier. The training statistic result is listed in Table 4, which implies that the resulted tree classifier satisfies the accuracy demand, because there are only three error cases among 200 cases. The root mean squared error on test cases is 0.026.

Table 4
C4.5 tree training result statistics

Items	Value
Correct cases	197(98.5%)
Wrong cases	3(1.5%)
Kappa statistics	0.9822
Mean absolute error	0.0046
Root mean squared error	0.0481
Relative absolute error	2.737%
Root relative squared error	16.569%

The well-trained decision tree is then utilized to make schedule decision determining the optimal schedule adjusting scale and the detailed case study is given in the following section.

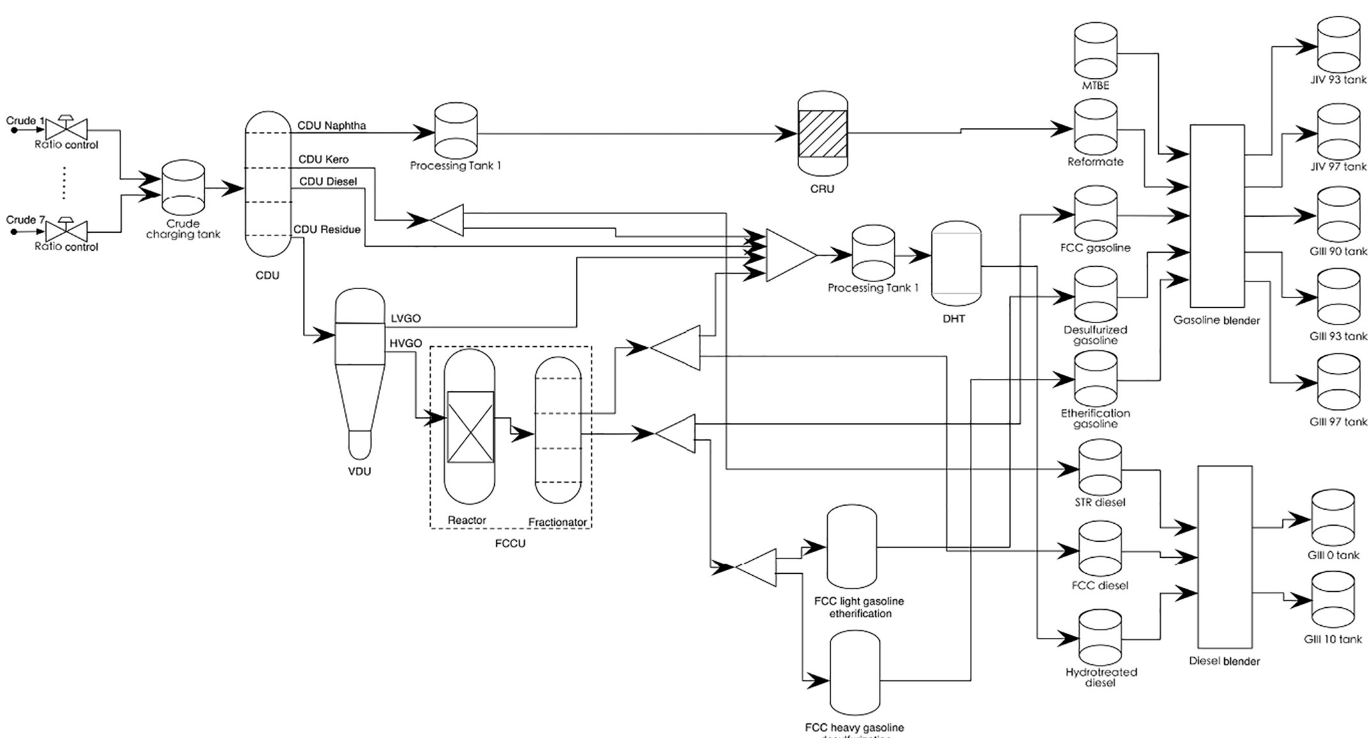

Fig. 2. The case refinery flowchart.

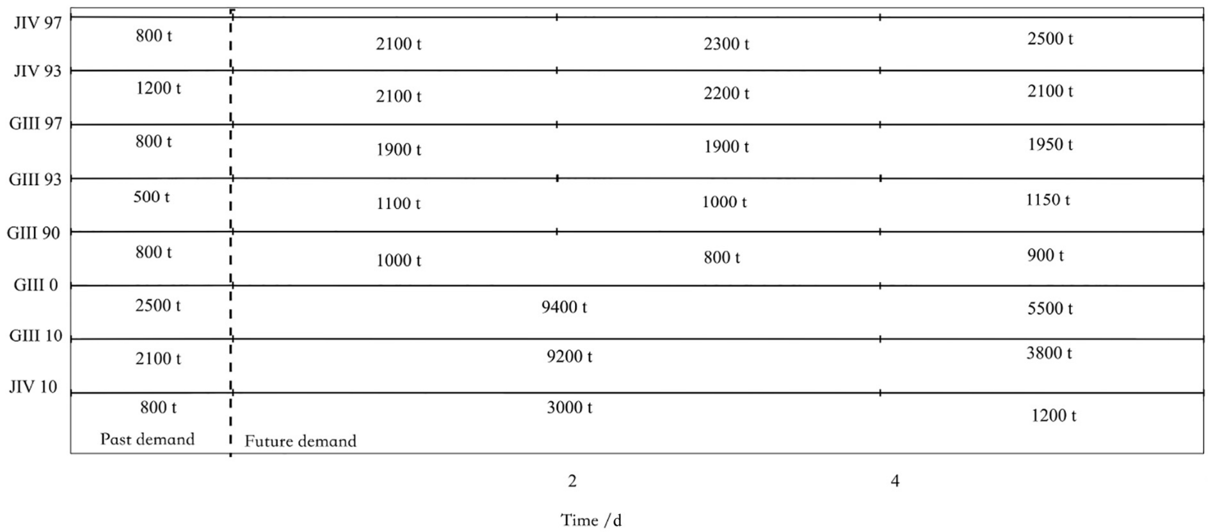

Fig. 3. The Gantt chart of the final oil product demand.

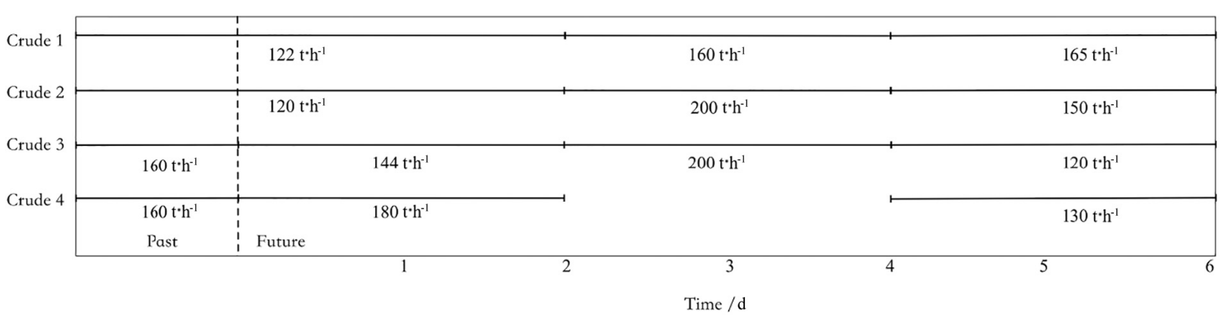

Fig. 4. The crude oil supply plan.

4.2.2. Scheduling case study

There is two types fuel demand. Their specifications are quite different, whose details are referred to our previous papers [9,10]. In this paper, we can simply take them premium fuel (such as JIV 97, JIV 93, JIV 10 in Fig. 3) and regular fuel (*i.e.* GIII 97, GIII 93, GIII 90, GII 0, GIII 10 in Fig. 3). The detailed demand data is depicted as Gantt chart in Fig. 3, and the crude supply plan is given in Fig. 4. Based on these data, the input attribute variables value can be inferred, which are summarized in Table 5. The obtained value is then input into the well-trained C4.5 decision tree to determine the adjusting class output. Based on this priori knowledge, the scheduling model is fine organized then. Compared with the conventional integrated optimization without these priori knowledge, this method decrease the problem scale and so speed up the optimization, which is meaningful for real world application when confronted with emergency for a timely schedule.

Table 5
The input attribute variable value of case study

Input attribute	Present batch	1st batch	2nd batch	3rd batch
PR	0.704	0.699	0.707	0.704
ΔPR	—	−0.005	0.008	−0.003
GDR	0.759	0.667	0.759	0.819
ΔGDR	—	−0.092	0.092	0.06
PGR	0.952	1.050	1.216	1.150
ΔPGR	—	0.098	0.166	−0.066
HGR	0.640	0.952	1.050	1.072
ΔHGR	—	0.312	0.098	0.022
PDR	0.174	0.161	0.161	0.129
ΔPDR	—	−0.013	0	−0.032
LDR	1.160	1.298	1.298	0.909
ΔLDR	—	0.138	0	−0.389

The resulted scheduling model based on the discrete time representation is then optimized by LINGO 11.

To further demonstrate the effectiveness of the proposed strategy, the comparison case study is also provided. Two other methods are involved in the comparisons. One is the conventional method, in which an integrated plant-wide model is computed. The other is the heuristic decomposition method, where a from downstream to upstream stepwise trial method is adopted. Clearly, this heuristic method has little knowledge. The comparison results are listed in Table 6. The results reveal that the decomposition strategy largely reduces the computation burden and decreases the problem scale, comparing with the last two columns and the first one, and C4.5 decision tree based method further accelerate the optimization comparing with the heuristic method because it incorporates the knowledge and avoids some meaningless trials. Moreover, the conventional method needs more transient time to reach the schedule goals. Clearly, the longer transition caused by the schedule adjustment results in more profit loss or even the failure to realize the

Table 6
Comparison results

Items	Integrated method	Heuristic decomposition method	Proposed method
Profit/CNY	69027881	68351037	68439812
Time /s	3285.6	792.0	341.5
Equality number	12708	5688	1896
Inequality number	3744	5364	1788
Continuous variables	11196	6840	2280
Binary variables	468	522	174

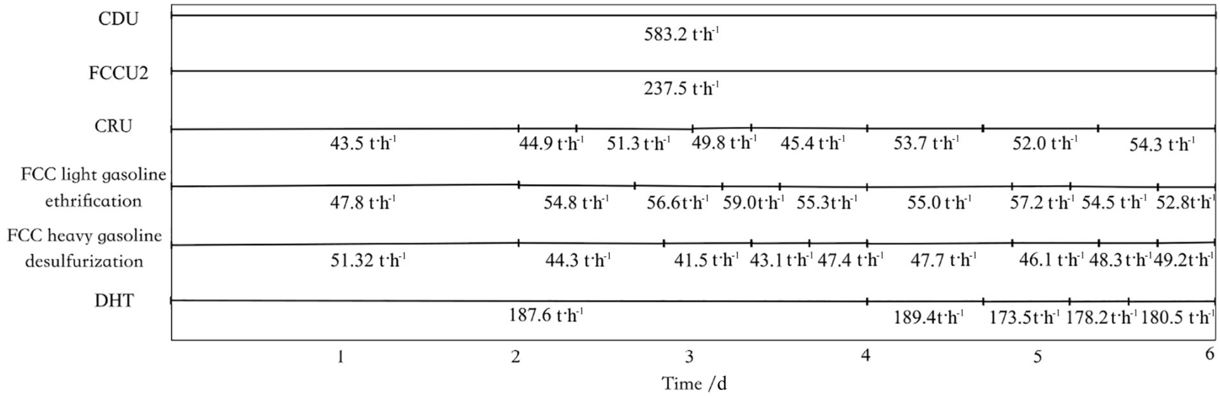

Fig. 5. The Gantt chart of the resulted schedule.

schedule goal in the extreme case if the plant undergoes dynamic transitions during the entire scheduling time horizon. (See Fig. 5.)

5. Conclusions

In this paper, a decision tree based decomposition method is proposed for fuel oriented refinery scheduling optimization to accelerate the computation and improve the decision quality. Based on the deep analysis of mechanism, the decision tree input attributes are well-selected. The trained decision tree is utilized to determine the adjusting scale and guide the scheduling optimization. The case study reveals the effectiveness of the proposed strategy, which speed up the optimization.

Nomenclature

$DD_{u,b}^{MIN}$ minimum demand of batch b for blender u
$INV_{u,t}$ inventory of unit u in time interval t
INV_u^{MAX} maximum inventory of tank u
INV_u^{MIN} minimum inventory of tank u
$OPC_{u,m}$ operating cost of unit u (belonging to PPUs) under mode m
OPC_{u,m_r,m_f} operating cost of unit u (belonging to SPUs) under reacting mode m_r and fraction mode m_f
$OPC_{u,t}$ operating cost of unit u (belonging to HUPUs) in time interval t
o_u price of product oil in tank u
$PRO_{s'},u,p,m$ property p of output stream s' of unit u under mode m
$PRO_{s'},u,p,m_r,m_f$ property p of output stream s' of unit u under reacting mode m_r and fraction mode m_f
$PRO_{s',u,m,p}^{MAX}$ maximum property p of blender under mode m
$PRO_{s'}^{MIN},u,m,p$ minimum property p of blender under mode m
$PRO_{s'},u,p,t$ property p for output stream s' of unit u (belonging to HUPUs) in time interval t
$\Delta PRO_{s'}^{MAX},u,p$ maximum delta-change of property p
$\Delta PRO_{s'}^{MIN},u,p$ minimum delta-change of property p
$\Delta PRO_{s'},u,p.\ t$ delta change of property p for output stream s' of unit u (belonging to HUPUs) in time interval t
Q_s,u,t input stream s flow of unit u in time interval t
$Q_{s'},u,t$ output stream s' flow of unit u in time interval t
QF_u^{MAX} maximum processing capacity of unit u
QF_u^{MIN} minimum processing capacity of unit u
QI_u,t inputs flow of unit u in time interval t
QO_u,t outputs flow of unit u in time interval t
r_s^{MAX},u,m maximum receipt of component s for unit u under mode m
r_s^{MIN},u,m minimum receipt of component s for unit u under mode m
s_u storage cost of tank u
$YLD_{s'},u,m$ output stream s' yield of PPU unit under mode m
$YLD_{s'},u,m_r,m_f$ output stream s' yield of SPU unit u under reacting mode m_r and fraction mode m_f
$YLD_{s',u,t}$ output stream s' of unit u (belonging to HUPUs) in time interval t
z_u,m,t binary variable indicating whether unit u for time interval t operated under mode m
z_u,m_r,m_f,t binary variable indicating whether unit u for time interval t operated under reacting mode m_r and fraction mode m_f

Subscripts

b demand batches
m operating mode
o product oil
p property
s, s' flow stream
t time period
u,u',u'' units

Sets

B_u set of demand batches of unit u
BLD set of blenders
BTK_I~BTK_V set of tanks connecting with upstream units under step I to V
HUPUs set of hydro-upgrading processing units
IS_u set of input streams of unit u
ITK_{II}~ITK_V set of tank under adjusting scale II to V
M_u set of operating modes of unit u
M_u^r, M_u^f set of reacting modes and fraction modes for secondary processing unit u
OS_u set of output streams of unit u
P set of properties
PPUs set of primary processing units
PTK set of product oil tank
SPUs set of secondary processing units
TK set of all tanks
TP set of scheduling periods
U set of units

References

[1] J.M. Pinto, M. Joly, L.F.L. Moro, Planning and scheduling models for refinery operations, *Comput. Chem. Eng.* 24 (2000) 2259–2276.

[2] P. Smania, J.M. Pinto, Mixed integer nonlinear programming techniques for the short term scheduling of oil refineries, *Comput. Aided Chem. Eng.* 15 (2003) 1038–1043.

[3] M. Joly, L.F.L. Moro, J.M. Pinto, Planning and scheduling for petroleum refineries using mathematical programming, *Braz. J. Chem. Eng.* 19 (2) (2002) 207–228.

[4] Z. Jia, M. Ierapetritou, J.D. Kelly, Refinery short-term scheduling using continuous time formulation: crude-oil operations, *Ind. Eng. Chem. Res.* 42 (13) (2003) 3085–3097.

[5] Z. Jia, M. Ierapetritou, Efficient short-term scheduling of refinery operations based on a continuous time formulation, *Comput. Chem. Eng.* 28 (2004) 1001–1019.

[6] M.E. Dogan, I.E. Grossmann, A decomposition method for the simultaneous planning and scheduling of single-stage continuous multiproduct plants, *Ind. Eng. Chem. Res.* 45 (1) (2006) 299–315.

[7] D. Wu, M. Ierapetritou, Hierarchical approach for production planning and scheduling under uncertainty, *Chem. Eng. Process.* 46 (11) (2007) 1129–1140.

[8] N.K. Shah, M.G. Ierapetritou, Short-term scheduling of a large-scale oil-refinery operations: incorporating logistics details, *AIChE J.* 57 (2011) 1570–1584.

[9] X.Y. Gao, C. Shang, Y.H. Jiang, T. Chen, D.X. Huang, Refinery scheduling with varying crude: a deep belief network classification and multimodel approach, *AICHE J.* 60 (2014) 2525–2532.

[10] X.Y. Gao, Y.H. Jiang, T. Chen, D.X. Huang, Optimizing scheduling of refinery operations based on piecewise linear models, *Comput. Chem. Eng.* 75 (2015) 105–119.

[11] M. Göthe-Lundgren, J. Lundgren, J. Persson, An optimization model for refinery production scheduling, *Int. J. Prod. Econ.* 78 (3) (2002) 255–270.

[12] C.P. Luo, G. Rong, Hierarchical approach for short-term scheduling in refineries, *Ind. Eng. Chem. Res.* 46 (2007) 3656–3668.

[13] C.A. Floudas, X. Lin, Continuous-time versus discrete-time approaches for scheduling of chemical processes: a review, *Comput. Chem. Eng.* 28 (11) (2004) 2109–2129.

[14] J. Bengtsson, S.L. Nonås, Refinery planning and scheduling: an overview, *Energy Nat. Resour. Environ. Econ.* (2010) 115–130.

[15] N.K. Shah, Z.K. Li, M.G. Ierapetritou, Petroleum refining operations: key issues, advances, and opportunities, *Ind. Eng. Chem. Res.* 50 (2010) 1161–1170.

[16] M. Joly, Refinery production planning and scheduling: the refining core business, *Braz. J. Chem. Eng.* 29 (2) (2012) 371–384.

[17] I. Harjunkoski, P. Maravelias, Bongers, *et al.*, Scope for industrial applications of production scheduling models and solution methods, *Comput. Chem. Eng.* 62 (2014) 161–193.

[18] Xiaoyong Gao, Jiang Yongheng, Huang Dexian, Process modelling based on integration of unitwide optimal process control and plantwide scheduling, *J. Chem. Ind. Eng. (China)* 67 (12) (2016) 5105–5111 (In Chinese).

[19] S. Engell, I. Harjunkoski, Optimal operations: scheduling, advanced control and their integration, *Comput. Chem. Eng.* 47 (2012) 121–133.

[20] Y. Chu, F. You, Integration of scheduling and control with online closed-loop implementation: fast computational strategy and large-scale global optimization algorithm, *Comput. Chem. Eng.* 47 (2012) 248–268.