mathematics

MDPI

Article

The Expected Competitive Ratio on a Kind of Stochastic-Online Flowtime Scheduling with Machine Subject to an Uncertain Breakdown

Zhenpeng Li [1] and Congdian Cheng [2,*]

[1] School of Electronics and Information Engineering, Taizhou University, Linhai 317000, China; lizhenpeng@amss.ac.cn

[2] School of Intelligence Technology, Geely University of China, Chengdu 641400, China

* Correspondence: chengcongdian@bgu.edu.cn

Abstract: We consider the problem of scheduling jobs on a single machine subject to an uncertain breakdown to minimize the flowtime. Assuming the machine is unavailable during the breakdown, the starting time of the breakdown is a random variable s with distribution function $D(s)$ and the terminating time of the breakdown has no any other information; jobs are non-resumable. Under these assumptions and starting from the perspective of statistical optimization, we first establish the scheduling problem HSONRP, which contains deterministic information, stochastic information, and online information and then define the expected competitive ratio of an algorithm to find the optimized solution of the problem HSONRP. In addition, then, we propose and prove certain results on the expected competitive ratio of the SPT rule. In particular, we prove the expected competitive ratio of SPT rule is less than $\left(1+\frac{\max\{p_i\}}{2P}\right)$ when s is the uniform distribution on interval $(0, P]$, where p_i is the processing time of job i, $P = \sum_{i=1}^{n} p_i$, and show that it is no more than $\frac{5}{4}$ under a quite loose condition. Meanwhile, we also make some discussions about our studies. What we have performed will enrich and improve the research results on the area of scheduling to minimize flowtime and advance the development of online optimization and stochastic optimization.

Keywords: single machine; single breakdown; schedule; flowtime; stochastic optimization; online optimization

MSC: 90B36; 68M20; 60K37

Citation: Li, Z.; Cheng, C. The Expected Competitive Ratio on a Kind of Stochastic-Online Flowtime Scheduling with Machine Subject to an Uncertain Breakdown. *Mathematics* **2023**, *11*, 2413. https://doi.org/10.3390/math11112413

Academic Editor: Amir Elalouf

Received: 27 March 2023
Revised: 26 April 2023
Accepted: 18 May 2023
Published: 23 May 2023

1. Introduction

As is known to all, machines may experience some breakdowns or become unavailable for periods of time. In addition, many breakdowns have a lot of uncertainty. To meet the actual needs, many scholars have paid much attention to the study of the scheduling problem with the machine subject to breakdown. In the first place, in 1984, Glazebrook [1] studied a kind of scheduling problem for a machine with a stochastic breakdown. In 1989, Adiri et al. [2] considered the problem of scheduling jobs on a single machine to minimize the flowtime (total completion time); they mainly discussed the problem that the SPT policy minimizes the flowtime in the case of a single breakdown and the more general case of multiple breakdowns, respectively. In 1990, Birge et al. [3] studied a kind of single-machine scheduling problem with the machine subject to a sequence of stochastic breakdowns. After this, the scheduling work with the machine subject to a sequence of stochastic breakdowns has constantly appeared. For instance, in 1993, Mittenthal and Raghavachair [4] investigated the problem of quadratic early-tardy penalties with the machine subject to a sequence of stochastic breakdowns. In 1996, Cai and Tu [5] made a further investigation into the problem that was considered by the paper [4] under the condition of jobs with random processing times. In 2008, Tang et al. [6] studied a stochastic JIT scheduling problem with a machine subject to a sequence of stochastic breakdowns. In 2008, Cheng et al. [7] studied the problem of scheduling jobs on a machine subject to stochastic breakdowns to minimize absolute early-tardy penalties. Currently, the research on the machine with breakdowns is still in progress. In 2022, Li and Chen [8] explored the problem of minimizing total weighted late work on a single machine with non-availability intervals. In 2022, Choi and Park [9] considered a single-machine scheduling problem with resource-dependent processing times and multiple unavailability periods.

Since this century, the study of online scheduling with machine subject to breakdowns has been gradually thriving. In 2002, Tan and He [10] studied the optimal online algorithm for scheduling on two identical machines with unavailable periods. In 2014, Huo et al. [11] investigated the scheduling problem on a single machine with a single breakdown, or say with a single unavailable period; they considered the online optimization problem of minimizing total weighted completion time under the condition assumed that every job has a weight proportional to its processing time. In 2016, Kacem and Kellerer [12] investigated three single-machine scheduling problems with the so-called semi-online scenario, which are actually the scheduling problem on a single machine with an unexpected breakdown. In 2020, Tian [13] improved a part of the work of literature [12].

Up to now, most research on the online scheduling problem with the machine subject to a single breakdown focuses on the following research setting. The breakdown results in an unavailable period or hole $[s, t]$, during which the machine cannot perform the processing of any job. At the beginning of the scheduling horizon, except for the breakdown, no other information is known; in other words, its state is unknown, or say, the starting time s and the terminating time t, including the distributions of theirs, are unknown and independent; a sequence of the jobs must be determined in advance and the sequence cannot be changed in the whole process of production. Jobs are non-resumable, that is, the job which has been interrupted by the breakdown has to be restarted to process after the machine is recovered. The objective is to find an optimization sequence or schedule.

In 2008, Cheng et al. [14] put forward the following views.

In practice, problems are often not really online or offline, but somewhere in between. This means that, with respect to the online problem, some further information about the jobs is available. This additional information allows the improvement of the performance of the best solution algorithm with respect to the online problem.

With this thought, they investigated the semi-online scheduling problem with decreasing processing times (SOSDPs), which evolves from a classical problem by increasing the additional information, and obtained excellent results.

In the scheduling research areas, there are many important and interesting problems with uncertainty, as well as many related and excellent research results and methods, see Pinedo [15], and Sotskov and Werner [16]. In addition, this research field is still evolving broadly in the present. For instance, in 2018, for the scheduling problem to minimize the total weighted completion time, Lai et al. [17] established the concept of the optimality box with a fixed permutation and using the optimality box as the stability measure, they investigated the permutation with the largest optimality box; in addition, they also derived an O(n)-algorithm for calculating the optimality box and investigated the properties of the optimality box.

Motivated by the above background information, we note that in many cases, we have known some states of the related processing situation in advance, for example, the probability distribution of the breakdown. It is very meaningful work to discuss the optimization scheme of problems under certain statistical information of the breakdown. In the present work, we address the problem to minimize flowtime (minimize total completion time) with the above setting integrated the condition that the starting time is a random variable with known distribution. We will first establish the scheduling problem HSONRP, which contains deterministic information, stochastic information, and online information, and then define the expected competitive ratio of an algorithm to find the optimized solution with the problem HSONRP. Further we investigate the expected competitive ratio of the SPT rule. In addition, we will also make some discussions on our study.

The remainder is organized as follows. Section 2 gives preliminary notations and knowledge. Section 3 introduces the related works. Section 4 pays attention to our main work. Section 5 makes some discussions on our work. Finally, we draw our conclusions in Section 6.

2. Formulation

Process n jobs on a single machine M subject to a single breakdown. Assume the starting time and the terminating time of the breakdown are s and t, respectively; machine M cannot work in the period of time $[s,t]$ for the breakdown. Jobs are non-resumable, i.e., when a job is interrupted by the unavailable interval, it has to be restarted after the machine is recovered, and all jobs are ready to be scheduled at the time 0. Unless there is a breakdown during the processing of a job, jobs must be processed without interruption. Each job $J_j,\ j = 1, 2, \cdots, n$, has a determined processing time $p_j > 0$. w_j is a positive real number, called the weight of job J_j. The process starts from the time 0 and one job can be processed at a time. A processing schedule, or say sequence, S must be made before the time 0 and the schedule S cannot be changed in the whole process. Let C_j be the completion time of the job J_j. Consider the problem to find a schedule $S : J_{j_1}, J_{j_2}, \cdots, J_{j_n}$, which is also denoted by $S : J_1, J_2, \cdots, J_n$ below for conciseness, before time 0 such that the (weighted) flowtime, or say, the total (weighted) completion time, is minimized, namely $\sum_{j=1}^{n} C_j$ ($\sum_{j=1}^{n} w_j C_j$) is minimized.

When $s = t$, we believe no breakdown occurs, which is called the degenerate case. The problem is just the classical scheduling problem of minimizing the (weighted)flowtime, which is denoted by

$$1|nr - a|\sum w_j C_j. \tag{1}$$

When $s < t$, the problem can be divided into the following three types. If $[s,t]$ is given before time 0, we call the problem as minimizing weighted flowtime for the machine with a non-availability interval or hole. If $[s,t]$ is unknown before time 0, we call the problem as online minimizing weighted flowtime for the machine with a non-availability interval or hole. If a part of distribution on interval $[s,t]$ is known before time 0, we call the problem as the scheduling problem of stochastic-online minimizing weighted flowtime for the machine with a non-availability interval or hole. In the present work, by the three-field notation, we, respectively, denote the three problems as

$$1, h_1|nr - a|\sum w_j C_j, \tag{2}$$

$$1, h_1, online|nr - a|\sum w_j C_j, \tag{3}$$

$$1, h_1, stochastic - online|nr - a|\sum w_j C_j. \tag{4}$$

When $w_j = 1,\ j = 1, 2, \cdots, n$, for convenience, we denote the problems (2), (3), and (4), respectively, as

$$1, h_1|nr - a|\sum C_j, \tag{5}$$

$$1, h_1, online|nr - a|\sum C_j, \tag{6}$$

$$1, h_1, stochastic - online|nr - a|\sum C_j. \tag{7}$$

For problem (7), assume the starting time is a random variable $\boldsymbol{s}$ and there is no any other information on the terminating time t; assume also

$$\Pr(\{\boldsymbol{s} < 0\}) = 0,\ \Pr(\{\boldsymbol{s} = 0\}) = p,\ \Pr(\{\boldsymbol{s} > P\}) = q;$$
$$D(s) = p + \int_0^s \varphi(x)dx,\ 0 < s \leq P,$$

where $P = \sum_{i=1}^{n} p_i$, $D(s)$ is the distribution function of $\boldsymbol{s}$; $\varphi(x)$ expresses the density of $\boldsymbol{s}$ on the interval $(s, P]$; the integral is Riemann integral. In the case, we denote the problem as HSONRP (hole stochastic online non-resumable problem).

Remark 1. *When $p = 1$, the probability that the breakdown occurs is 0, the problem HSONRP degrades into problem (1). This implies that the problem HSONRP is the generalization of the classical minimizing flowtime.*

In brief, we denote the objective function of the schedule S by $F(S)$, that is, $F(S) = \sum w_j C_j$.

For the problem (1) and (2), let S be a schedule and S^* be the optimal schedule. Then we call $\rho(S) = \frac{F(S)}{F(S^*)}$ as the competitive ratio of schedule S. Let $\mathcal{A}$ be an algorithm to solve the problem, I be an instance, and S_I be the schedule with I obtained by algorithm $\mathcal{A}$. Then $\rho(\mathcal{A}) = \sup\{\rho(S_I)|I\}$ is called the competitive ratio, or the approximation ratio of algorithm $\mathcal{A}$, where I is any instance of Equations (1) and (2).

For the problem (3), let $\mathcal{A}$ be an algorithm to solve the problem, I be an instance and for the instance I, S_I be the schedule obtained by algorithm $\mathcal{A}$. Let also $\rho(S_I, [s,t])$ be the competitive ratio of the schedule S_I for the problem (2) with hole $[s,t]$. Then we call $\rho(\mathcal{A}) = \sup\{\rho(S_I, [s,t])|I, [s,t]\}$ as the competitive ratio of algorithm $\mathcal{A}$ for the problem.

For the problem HSONRP, let $\mathcal{A}$ be an algorithm to solve the problem I be an instance, and S_I be the schedule obtained by algorithm $\mathcal{A}$ for the instance I. Given $s \in (0, P)$, for any $t > s$, let also $\rho(S_I, [s,t])$ be the competitive ratio of the schedule S_I for the problem (2) with the hole $[s,t]$. Define

$$\rho(s, \mathcal{A}) = \sup\{\rho(S_I, [s,t])|I, t > s\}.$$

When $s = 0$ or $s \geq P$, we define $\rho(s, \mathcal{A}) = \sup\{\rho(S_I)|I\}$, where $\rho(S_I)$ is the competitive ratio of the schedule S_I for the problem (1). Since, if the starting time $s = 0$, we can arrange the schedule after the breakdown, and if the starting time $s \geq P$, the breakdown does not disturb the arranged schedule, we can believe no breakdown occurs in the two cases. That is, when $s = 0$ or $s \geq P$, it is reasonable that we define

$$\rho(s, \mathcal{A}) = \sup\{\rho(S_I)|I\}.$$

When the starting time is the random variable $\boldsymbol{s}$, we call the mean value of random variable $E[\rho(\boldsymbol{s}, \mathcal{A})]$ as the expected competitive ratio of algorithm $\mathcal{A}$ for the problem.

In order to convenience, we also introduce the following notations and assumptions.

Given $[s,t]$, let $S: J_1, J_2, \cdots, J_n$ be a schedule. We use J_k express the interrupted job. Let also

$$P(i,j) = \sum_{l=i}^{j} p_l, P_j = P(1,j),\ 1 \leq i \leq j \leq n; P_0 = 0; FP = \sum_{j=1}^{n} P_j.$$

Denote the schedule obtained by SPT rule as $S' : J'_1, J'_2, \cdots, J'_n$. Since the case that $n = 1$ is trivial, we assume $n > 1$ in the following. Moreover, we also assume that the event $\{\boldsymbol{s} = 0\}$ includes the two cases that the time 0 is in the breakdown period and there is no breakdown.

3. Related Work

In the first place, Equation (1) is the classical scheduling problem of minimizing the total weighted completion time; in 1956, Smith [18] firstly studied this problem and proposed the famous WSPT(weighted shortest processing time) rule, that is, arranging jobs as the order $\frac{w_j}{p_j}$ non-decreasing, the objective is minimized. Formula (2) is a quite difficult problem. Substantial works have discussed this problem over the last couple of decades; see the related work of Huo et al. [11]. In particular, in 1996, Lee [19] showed that the error bound in the WSPT rule can be arbitrarily large. In 2008, Kacem and Chu [20] showed that both the WSPT rule and MWSPT (modified weighted shortest processing time) rules have a tight error bound of 3 under some conditions. In 2008, Kacem [21] proposed a 2-approximation algorithm with $O(n^2)$ time complexity and showed bound 2 is tight. In 2014, Huo et al. [11] proposed the so-called the FF-LPT rule, which is a modified version of the LPT rule; they showed both the LPT rule and FF-LPT rule can give a tight approximation ratio of 2 in the case of $w_j = p_j$.

Formula (5) is the special case of Equation (2), which is called the minimizing flowtime problem. For the problem, in 1989, Adiri et al. [2] showed the problem is NP-complete, and the SPT rule has the approximation ratio $\frac{5}{4}$ (the relative error for the SPT schedule is less than or equal to $\frac{1}{4}$). In 1992, Lee and Liman [22] proved that SPT rule has a tight approximation ratio of $\frac{9}{7}$, instead of $\frac{5}{4}$. Later, in 2005, Sadfi et al. [23] gave the so-called modified SPT police algorithm that has a tight approximation ratio of $\frac{20}{7}$. In 2006, He et al. [24] also proposed a polynomial time approximation scheme. In 2007, Breit [25] developed a parametric $O(n\log n)$ -algorithm, in which better worst-case error bounds can be obtained.

For problem (3), in 2014, Huo et al. [11] first showed, in the case $w_j = p_j$, there is no algorithm $\mathcal{A}$ such that $\rho(\mathcal{A}) > \frac{\sqrt{5}+1}{2}$, and LPT rule has a tight competitive ratio 2. Problem (6) is the special case of Equation (3), which is called online minimizing the flowtime for the machine with a non-availability interval or hole, and closely contact with Equation (5). Since the SPT rule and the above stated results of the competitive ratio on the rule for problem (5) have no relation with specific conditions of s and t, these results are also true for the problem (6).

For the problem (7), assume the starting time to breakdown is a random variable $\boldsymbol{s}$. Adiri et al. [2] proposed a definition of a schedule to stochastically minimize the flowtime and showed that if the distribution function of the staring time $\boldsymbol{s}$ is concave, then the SPT rule stochastically minimizes the flowtime in the sense of their definition. In addition, for problems (4) and (7), to the best of our knowledge, there are no other works until now.

In the remainder part of present work, we address problem HSONRP, which is a kind of problem (7). We mainly study the expected competitive ratio of the SPT rule, and make some discussions about the work we study. We will establish a few inequalities of $E[\rho(\boldsymbol{s}, \text{SPT})]$ to facilitate the estimation of the expected competitive ratio of SPT rule. In particular, we will prove

$$E[\rho(\boldsymbol{s}, \text{SPT})] \le 1 + \frac{\max\{p_i\}}{2P},$$

When $\boldsymbol{s}$ is uniform distribution on the interval $(0, P]$, and we will also show $E[\rho(\boldsymbol{s}, \text{SPT})] \le \frac{5}{4}$ under a quite loose condition.

Remark 2. *The processing times of jobs are all deterministic in the problems (1)–(7). However, they may be uncertain in some situation. Therefore, it is a quite meaningful topic to explore the stability approach of the scheduling problems that the processing times of jobs are uncertain. Some studies have been conducted in this field, e.g., see [5,17]. We believe to study problems (1)–(7) with the uncertain processing times are an interesting direction to pursue.*

4. Main Work

First of all, given $[s,t]$, let $S : J_1, J_2, \cdots, J_n$ be a schedule. Then we have

$$\begin{aligned}
F(S) &= \sum_{j=1}^{k-1} C_j + \sum_{j=k}^{n} C_j = \sum_{j=1}^{k-1} C_j + \sum_{j=k}^{n} [t + P(k,j)] \\
&= \sum_{j=1}^{k-1} C_j + \sum_{j=k}^{n} [(t - C_{k-1}) + C_{k-1} + P(k,j)] \\
&= \sum_{j=1}^{k-1} P_j + (n-k+1)(t - P_{k-1}) + \sum_{j=k}^{n} [P_{k-1} + P(k,j)] \\
&= \sum_{j=1}^{k-1} P_j + (n-k+1)(t - P_{k-1}) + \sum_{j=k}^{n} P_j \\
&= \sum_{j=1}^{n} P_j + (n-k+1)(t - P_{k-1}) \\
&= \sum_{j=1}^{n} P_j + (n-k+1)[(t-s) + (s - P_{k-1})].
\end{aligned}$$

Here, $(s - P_{k-1})$ is the idle time immediately before the breakdown in the schedule S.

Lemma 1. *For the problems (5), (6) and (7) with given $[s,t]$, let $S' : J'_1, J'_2, \cdots, J'_n$ be the SPT schedule. Then we have the following two conclusions. (1) For any schedule $S : J_1, J_2, \cdots, J_n$, we have: $k \le k'$, where J_k and $J'_{k'}$ are the interrupted jobs, respectively, in the schedule S and the schedule S'. (2) For the optimal schedule $S^* : J_1^*, J_2^*, \cdots, J_n^*$, we have: $P'_{k'-1} \le P^*_{k^*-1} \le s$.*

Proof. For $S' : J'_1, J'_2, \cdots, J'_n$ is the SPT schedule, we have $p'_1 \le p'_2 \le \cdots \le p'_n$. Hence, (1) holds. Now we prove (2) as follows. By the above expressions of $F(S)$, we have

$$\begin{aligned}
F(S') &= \sum_{j=1}^{n} P'_j + (n-k'+1)[(t-s) + (s - P'_{k'-1})]; \\
F(S^*) &= \sum_{j=1}^{n} P^*_j + (n-k^*+1)\left[(t-s) + \left(s - P^*_{k^*-1}\right)\right].
\end{aligned}$$

From the term (1), we have $(n-k^*+1) \ge (n-k'+1)$. Due to S^* is the optimal schedule, we also have $F(S) \le F(S')$. Due to S' is the SPT schedule, we have $\sum_{j=1}^{n} P'_j \le \sum_{j=1}^{n} P^*_j$. Combining these results, we have

$$\begin{aligned}
&\sum_{j=1}^{n} P^*_j + (n-k^*+1)\left[(t-s) + \left(s - P^*_{k^*-1}\right)\right] \le \sum_{j=1}^{n} P'_j + (n-k'+1)[(t-s) + (s - P'_{k'-1})] \\
&\Rightarrow \left(\sum_{j=1}^{n} P^*_j - \sum_{j=1}^{n} P'_j\right) + (k'-k^*)(t-s) + (n-k^*+1)\left(s - P^*_{k^*-1}\right) \le (n-k'+1)(s - P'_{k'-1}) \\
&\Rightarrow (n-k^*+1)\left(s - P^*_{k^*-1}\right) \le (n-k'+1)(s - P'_{k'-1}) \\
&\Rightarrow \left(s - P^*_{k^*-1}\right) \le (s - P'_{k'-1}) \Rightarrow P'_{k'-1} \le P^*_{k^*-1}.
\end{aligned}$$

In addition, $P^*_{k^*-1} \le s$ is obvious. Thus, (2) holds and we have completed the proof. □

Lemma 2. *For problem (5), we have*

$$\begin{aligned}
\rho(\mathrm{SPT}) &\le 1 + \tfrac{(n-k'+1)\left(s - P'_{k'-1}\right)}{FP' + (n-k'+1)(t-s)} \\
&\le 1 + \tfrac{(n-k'+1)\left(s - P'_{k'-1}\right)}{FP'};
\end{aligned} \quad (8)$$

$$\begin{aligned}
\rho(\mathrm{SPT}) &\le 1 + \tfrac{(n-k'+1)p'_{k'}}{FP' + (n-k'+1)(t-s)} \\
&\le 1 + \tfrac{(n-k'+1)p'_{k'}}{FP'} \le 1 + \tfrac{(n-k'+1)p'_n}{FP'}.
\end{aligned} \quad (9)$$

For problem (6), we have

$$\rho(\mathrm{SPT}) \le 1 + \frac{1}{FP'} \max_{1 \le k' \le n} \{(n-k'+1){p'}_{k'}\}. \tag{10}$$

Here SPT denotes the algorithm of SPT rule, sequence S' : ${J'}_1, {J'}_2, \cdots, {J'}_n$ is the SPT schedule.

Proof. Let S^* be the optimal schedule. Since S' be the SPT schedule, we have

$$\sum_{j=1}^{n} P_j^* \ge \sum_{j=1}^{n} {P'}_j.$$

By Lemma 1, we also have $k \le k'$. Therefore, we obtain

$$\begin{aligned} &\sum_{j=1}^{n} P_j^* + (n-k^*+1)\left[(t-s) + \left(s - P_{k^*-1}^*\right)\right] \\ \ge &\sum_{j=1}^{n} {P'}_j + (n-k'+1)(t-s) + (n-k'+1)\left(s - P_{k^*-1}^*\right) \\ \ge &\sum_{j=1}^{n} {P'}_j + (n-k'+1)(t-s). \end{aligned}$$

Hence,

$$\begin{aligned} \frac{F(S')}{F(S^*)} &= \frac{\sum_{j=1}^{n} {P'}_j + (n-k'+1)\left[(t-s) + \left(s - {P'}_{k'-1}\right)\right]}{\sum_{j=1}^{n} P_j^* + (n-k^*+1)\left[(t-s) + \left(s - P_{k^*-1}^*\right)\right]} \\ &\le \frac{\left[\sum_{j=1}^{n} {P'}_j + (n-k'+1)(t-s)\right] + (n-k'+1)(s - {P'}_{k'-1})}{\sum_{j=1}^{n} {P'}_j + (n-k'+1)(t-s)} \\ &= 1 + \frac{(n-k'+1)(s - {P'}_{k'-1})}{\sum_{j=1}^{n} {P'}_j + (n-k'+1)(t-s)}. \end{aligned}$$

Equation (8) holds. For ${J'}_k$ is the interrupted job, we have ${P'}_{k'-1} \le s < {P'}_{k'}$. This leads to

$$0 \le s - {P'}_{k'-1} < {P'}_{k'} - {P'}_{k'-1} = {p'}_{k'}.$$

Bring the result in Equation (8), we obtain Equation (9). Formula (10) is straight obtained from the definition on the competitive ratio of the SPT rule with problem (6) and the formula (9). The proof is finished. □

Theorem 1. *For problem HSONRP, we have*

$$E[\rho(\mathbf{s}, \mathrm{SPT})] \le 1 + \frac{1}{FP'} \left[\sum_{k=1}^{n} (n-k+1) \int_{{P'}_{k-1}}^{{P'}_k} \left(s - {P'}_{k-1}\right) \varphi(s) ds\right]. \tag{11}$$

When $\mathbf{s}$ is uniform distribution on the interval $(0, P]$, we further have

$$E[\rho(\mathbf{s}, \mathrm{SPT})] \le 1 + \frac{(1-p-q)}{2P \cdot FP'} \left[\sum_{k=1}^{n} (n-k+1){p'}_k^2\right]. \tag{12}$$

$$E[\rho(\mathbf{s}, \mathrm{SPT})] \le 1 + \frac{(1-p-q){p'}_n}{2P}. \tag{13}$$

Moreover, if $p'_n \le \frac{1}{2}P$ *or* $p'_n \le \sum\limits_{k=1}^{n-1}(n-k+1)p'_k$*, then we have*

$$E[\rho(\mathbf{s},\mathrm{SPT})] \le 1 + \frac{(1-p-q)}{4}.$$

Proof. We can easily know $\rho(s,\mathrm{SPT}) = 1$ when $s = 0$ or $s \ge P$ and in terms of (8), we have

$$\rho(s,\mathrm{SPT}) \le 1 + \frac{(n-k+1)(s-P'_{k-1})}{FP'},\ P'_{k-1} < s \le P'_k,\ k = 1,2,\cdots,n. \tag{14}$$

Further,

$$\begin{aligned} E[\rho(\boldsymbol{s},\mathrm{SPT})] &= \int_{-\infty}^{+\infty}\rho(s,\mathrm{SPT})dD(s) \\ &= \int_{-\infty}^{0}\rho(s,\mathrm{SPT})dD(s) + \int_0^P\rho(s,\mathrm{SPT})dD(s) + \int_P^{+\infty}\rho(s,\mathrm{SPT})dD(s) \\ &= p + q + \int_0^P\rho(s,\mathrm{SPT})dD(s), \end{aligned} \tag{15}$$

where the integrals are all the Lebesgue–Stieltjes integral of the function $\rho(s,\mathrm{SPT})$ on the function $D(s)$. For $D(s) = p + \int_0^s\varphi(x)dx,\ 0 < s \le P$, by Equation (14), we have

$$\begin{aligned} \int_0^P\rho(s,\mathrm{SPT})dD(s) &= \int_0^P\rho(s,\mathrm{SPT})\varphi(s)ds \\ &= \sum_{k=1}^{n}\int_{P'_{k-1}}^{P'_k}\rho(s,\mathrm{SPT})\varphi(s)ds \\ &\le \sum_{k=1}^{n}\left[\int_{P'_{k-1}}^{P'_k}\varphi(s)ds + \int_{P'_{k-1}}^{P'_k}\frac{(n-k+1)(s-P'_{k-1})}{FP'}\varphi(s)ds\right] \\ &\le \int_0^P\varphi(s)ds + \sum_{k=1}^{n}\int_{P'_{k-1}}^{P'_k}\frac{(n-k+1)(s-P'_{k-1})}{FP'}\varphi(s)ds \\ &= 1-p-q+\sum_{k=1}^{n}\frac{(n-k+1)}{FP'}\int_{P'_{k-1}}^{P'_k}(n-k+1)(s-P'_{k-1})\varphi(s)ds. \end{aligned} \tag{16}$$

Combining Equations (15) and (16), we obtain Equation (11). When $\boldsymbol{s}$ is uniform distribution on the interval $(0,P]$, we have

$$\varphi(s) = \frac{1-p-q}{P},\ s \in (0,P].$$

This leads to

$$\int_{P'_{k-1}}^{P'_k}(s-P'_{k-1})\varphi(s)ds = \frac{(1-p-q)}{P}\int_{P'_{k-1}}^{P'_k}(s-P'_{k-1})ds = \frac{(1-p-q)p'^2_k}{2P}.$$

Bring it in Equation (11), we obtain

$$E[\rho(\mathbf{s},\mathrm{SPT})] \le 1 + \frac{(1-p-q)}{2P\cdot FP'}\left[\sum_{k=1}^{n}(n-k+1)p'^2_k\right].$$

That is, Equation (12) holds. Note that $p'_k \le p'_n,\ k = 1,2,\cdots,n$, and

$$FP' = \sum_{k=1}^{n}P'_k = \sum_{k=1}^{n}(n-k+1)p'_k.$$

We know that Equation (13) is true from Equation (12). If $p'_n \le \frac{1}{2}P$, from Equation (13), the inequality

$$E[\rho(\mathbf{s},\mathrm{SPT})] \le 1 + \frac{(1-p-q)}{4}$$

is obvious. If $p'_n > \frac{1}{2}P$ and $p'_n \leq \sum_{k=1}^{n-1}(n-k+1)p'_k$, we have

$$\begin{aligned}
&E[\rho(\mathbf{s},\mathrm{SPT})] \\
&\leq 1+\tfrac{(1-p-q)}{2P\cdot FP'}\left[\sum_{k=1}^{n}(n-k+1){p'}^2_k\right] \\
&= 1+\tfrac{(1-p-q)}{2P\cdot FP'}\left[\sum_{k=1}^{n-1}(n-k+1){p'}^2_k+{p'}^2_n\right] \\
&= 1+\tfrac{(1-p-q)}{2P\cdot FP'}\left[\sum_{k=1}^{n-1}(n-k+1){p'}^2_k+p'_n\left(p'_n-\tfrac{1}{2}P\right)+p'_n\left(\tfrac{1}{2}P\right)\right] \\
&\leq 1+\tfrac{(1-p-q)}{2P\cdot FP'}\left\{\sum_{k=1}^{n-1}(n-k+1){p'}^2_k+\left[\sum_{k=1}^{n-1}(n-k+1)p'_k\right]\left(p'_n-\tfrac{1}{2}P\right)+p'_n\left(\tfrac{1}{2}P\right)\right\} \\
&\leq 1+\tfrac{(1-p-q)}{2P\cdot FP'}\left\{\left[\sum_{k=1}^{n-1}(n-k+1)p'_k\left(p'_k+p'_n-\tfrac{1}{2}P\right)\right]+p'_n\left(\tfrac{1}{2}P\right)\right\} \\
&\leq 1+\tfrac{(1-p-q)}{2P\cdot FP'}\left\{\left[\sum_{k=1}^{n-1}(n-k+1)p'_k\left(P-\tfrac{1}{2}P\right)\right]+p'_n\left(\tfrac{1}{2}P\right)\right\} \\
&\leq 1+\tfrac{(1-p-q)}{2P\cdot FP'}\left\{\left[\sum_{k=1}^{n}(n-k+1)p'_k\right]\left(\tfrac{1}{2}P\right)\right\}=1+\tfrac{(1-p-q)}{4}.
\end{aligned}$$

The proof is completed. □

Remark 3. *Normally,*

$$\lim_{n\to\infty}\left\{\frac{(1-p-q)}{2P\cdot FP'}\left[\sum_{k=1}^{n}(n-k+1){p'}^2_k\right]\right\}=0.$$

So, Equation (12) implies that the larger the scale n of problem HSONRP, the better the SPT rule is.

5. Discussion

(1) When the system possesses some information on statistics, such as the distributions of certain variables, it is quite useful in practice to find the high probability laws. Hence, facing the era of big data, we should work hard to develop statistical optimization, whose meaning is mainly to catch the optimization law of large probability. On the other hand, when a system, which has a large amount of uncertainty, does not possess the relative enough and simple information of statistics, it is impossible or quite difficult to find an effective optimization solution by statistical methods alone. Hence, facing the era of big data, we should also work hard to develop the methods of combining statistical optimization with online optimization. The present work is conducted under the promotion of the above idea. The definition of the expected competitive ratio is proposed in the light of the above idea. It provides a new stability measure for us to optimize the problem HSONRP and other similar problems. In addition, according to this measure we can select the schedule (or solution) of the related problem which is more likely superior compared with other schedules (or solutions). However, the previous optimization indicators did not have this function.

(2) Sometimes, considering a problem is under appropriate conditions, it is easy to find a good method. For example, under some conditions, we can obtain the result

$$E[\rho(\mathbf{s},\mathrm{SPT})]\leq 1+\frac{(1-p-q)}{4}.$$

The present work fully demonstrates this viewpoint.

(3) For problem HSONRP, when $\boldsymbol{s}$ is uniform distribution on the interval $(0, P]$, by the related result of Lee and Liman [22], that is, SPT has a tight approximation ratio of $\frac{9}{7}$, we have

$$\rho(\boldsymbol{s}, \mathrm{SPT}) \leq \frac{9}{7},\ 0 < s \leq P.$$

In terms of Equations (15) and (16), we can easily obtain the following conclusion.

$$E[\rho(\mathbf{s}, \mathrm{SPT})] \leq 1 + \frac{2}{7}(1 - p - q).$$

(4) For problem (5), in 1989, Adiri et al. [2] showed the SPT rule has approximation ratio $\frac{5}{4}$. In 1992, Lee and Liman [22] proved the SPT rule has a tight approximation ratio of $\frac{9}{7}$, instead of $\frac{5}{4}$. When $p + q = 0$, which is the setting of the previous work in this research line, from the result

$$E[\rho(\boldsymbol{s}, \mathrm{SPT})] \leq 1 + \frac{1}{4}(1 - p - q),$$

we can obtain $E[\rho(\boldsymbol{s}, \mathrm{SPT})] \leq \frac{5}{4}$. For the condition of this conclusion is quite loose and $\frac{5}{4} < \frac{9}{7}$, to some extent, the result demonstrates the advantage of the present work.

(5) Moreover, from Equation (13), we can obtain

$$E[\rho(\boldsymbol{s}, \mathrm{SPT})] \leq 1 + \frac{p'_n}{2P} = 1 + \frac{\max\{p_i\}}{2P}.$$

In terms of the result, we can quickly estimate $E[\rho(\boldsymbol{s}, \mathrm{SPT})]$ by $\frac{\max\{p_i\}}{P}$. This is also an advantage of the present work.

(6) In practice, we can approximately calculate the expected competitive ratio $E[\rho(\boldsymbol{s}, \mathcal{A})]$ by the method of numerical evaluations. This will fully develop the efficiency of big data and statistical optimization, widen the investigating horizon of the scheduling area and strengthen the effect of the scheduling research. The formulae (8)–(13) can strongly support the numerical evaluations. In the actual estimate, we can choose the formulae from (8)–(13), and make calculations according to the specific needs and conditions of the problem. Take the following as an instance. For the problem HSONRP, let $p = q = 0$, and $\boldsymbol{s}$ is uniform distribution on the interval $(0, P]$. According to the accuracy requirements of the actual problem, choose a $\delta \in (0,\ \min\{p_i\})$. In addition, also, for each j, let $s(i,j) = p'_{j-1} + (i-1)\delta,\ i = 1, 2, \cdots, k_j$, such that $0 < (p'_j - s(k_j, j)) < \delta$. Then $E[\rho(\boldsymbol{s}, \mathrm{SPT})]$ equals approximatively to

$$\left(\sum_{j=1}^{n} k_j\right)^{-1} \left\{\sum_{j=1}^{n} \left(\sum_{i=1}^{k_j} \rho(s(i,j),\ \mathrm{SPT})\right)\right\}.$$

From Equation (8), for each j, we have

$$\begin{aligned}\rho(s(i,j),\ \mathrm{SPT}) &\leq 1 + \tfrac{(n-j+1)\left(s(i,j) - P'_{j-1}\right)}{FP'} \\ &= 1 + \tfrac{(n-j+1)(i-1)\delta}{FP'},\ i = 1, 2, \cdots, k_j.\end{aligned}$$

Hence, we can obtain the approximate evaluation

$$E[\rho(\boldsymbol{s}, \mathrm{SPT})] \leq \left(\sum_{j=1}^{n} k_j\right)^{-1} \left\{\sum_{j=1}^{n} \left(\sum_{i=1}^{k_j} \left[1 + \frac{(n-j+1)(i-1)\delta}{FP'}\right]\right)\right\}.$$

6. Conclusions

In the present work, we develop and study the problem HSONRP. We have first established the problem HSONRP and defined the expected competitive ratio of an algorithm to find the optimized solution of the problem HSONRP. In addition, we mainly studied the expected competitive ratio of the SPT rule. We have proposed and proved a few inequalities to estimate the expected competitive ratio of SPT rule. Through this work, we have developed the approaches to study the scheduling problem which includes determinate information, stochastic information, and online information. For the expected competitive ratio, which can be extended to other models, there is much work to do. In the future, this is an interesting direction to pursue.

Author Contributions: Z.L.: Conceptualization, and writing—original draft. C.C.: Conceptualization, project administration, formal analysis, and writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of China under the grant (71661001).

Data Availability Statement: We declare data availability is not applicable to this article as no new data were created or analyzed in this study.

Conflicts of Interest: The authors have no financial or proprietary interests in any material discussed in this article.

References

1. Glazebrook, K.D. Scheduling stochastic jobs on a single machine subject to breakdowns. *Nav. Res. Logist.* **1984**, *31*, 251–264. [CrossRef]
2. Adiri, I.; Bruno, J.L.; Frostig, E.; Kan, A.H.G.R. Single Machine Flow-Time Scheduling with a Single Breakdown. *Acta Inform.* **1989**, *26*, 679–696. [CrossRef]
3. Birge, J.R.; Frenk, J.B.G.; Mittenthal, J.; Kan, A.H.G.R. Single-machine scheduling subject to stochastic breakdowns. *Nav. Res. Logist.* **1990**, *37*, 661–677. [CrossRef]
4. Mittenthal, J.; Raghavachari, M. Stochastic Single Machine Scheduling with Quadratic Early-Tardy Penalties. *Oper. Res.* **1993**, *41*, 786–796. [CrossRef]
5. Cai, X.; Tu, F.S. Scheduling jobs with random processing times on a single machine subject to stochastic breakdowns to minimize early-tardy penalties. *Nav. Res. Logist.* **1996**, *43*, 1127–1146. [CrossRef]
6. Tang, H.Y.; Zhao, C.L.; Cheng, C.D. Single machine stochastic JIT scheduling problem subject to machine breakdowns. *Sci. China Ser. A-Math* **2008**, *51*, 273–292. [CrossRef]
7. Cheng, C.D.; Tang, H.Y.; Zhao, C.L. Scheduling jobs on a machine subject to stochastic breakdowns to minimize absoluteearly-tardy penalties. *Sci. China Ser. A-Math* **2008**, *51*, 864–888. [CrossRef]
8. Li, S.S.; Chen, R.X. Minimizing total weighted late work on a single-machine with non-availability intervals. *J. Comb. Optim.* **2022**, *44*, 1330–1355. [CrossRef]
9. Choi, B.C.; Park, M.J. Single-machine scheduling with resource-dependent processing times and multiple unavailability periods. *J. Sched.* **2022**, *25*, 191–202. [CrossRef]
10. Tan, Z.Y.; He, Y. Optimal online algorithm for scheduling on two identical machines with machine availability constraints. *Inform. Process. Lett.* **2002**, *83*, 323–329. [CrossRef]
11. Huo, Y.M.; Reznichenko, B.; Zhao, H.R. Minimizing total weighted completion time with an unexpected machine unavailable interval. *J. Sched.* **2014**, *17*, 161–172. [CrossRef]
12. Kacem, I.; Kellerer, H. Semi-online scheduling on a single machine with unexpected breakdown. *Theor. Comput. Sci.* **2016**, *646*, 40–48. [CrossRef]
13. Tian, J.; Zhou, Y.; Fu, R. An improved semi-online algorithm for scheduling on a single machine with unexpected breakdown. *J. Comb. Optim.* **2020**, *40*, 170–180. [CrossRef]
14. Cheng, T.C.E.; Kellerer, H.; Kotov, V. Algorithms better than LPT for semi-online scheduling with decreasing processing times. *Oper. Res. Lett.* **2012**, *40*, 349–352. [CrossRef]
15. Pinedo, M. *Scheduling: Theory, Algorithms, and Systems*; Prentice-Hall: Englewood Cliffs, NJ, USA, 2002.
16. Sotskov, Y.N.; Werner, F. *Sequencing and Scheduling with Inaccurate Data*; Nova Science Publishers: Hauppauge, NY, USA, 2014.
17. Lai, T.C.; Sotskov, Y.N.; Egorova, N.G.; Werner, F. The optimality box in uncertain data for minimising the sum of the weighted job completion times. *Int. J. Prod. Res.* **2018**, *56*, 6336–6362. [CrossRef]
18. Smith, W.E. Various optimizers for single-stage production. *Nav. Res. Logist. Q.* **1956**, *3*, 59–66. [CrossRef]
19. Lee, C.Y. Machine scheduling with an availability constraints. *J. Glob. Optim.* **1996**, *9*, 395–416. [CrossRef]
20. Kacem, I.; Chu, C. Worst-case analysis of the WSPT1 and MWSPT2 rules for single machine scheduling with one planned setup period. *Eur. J. Oper. Res.* **2008**, *187*, 1080–1089. [CrossRef]
21. Kacem, I. Approximation algorithm for the weighted flow-time minimization on a single machine with a fixed non-availability interval. *Comput. Ind. Eng.* **2008**, *54*, 401–410. [CrossRef]
22. Lee, C.Y.; Liman, S.D. Single-machine flow-time scheduling with scheduled maintenance. *Acta Inform.* **1992**, *29*, 375–382. [CrossRef]

23. Sadfi, C.; Penz, B.; Rapine, C.; Błazewicz, J.; Formanowicz, P. An improved approximation algorithm for the single machine total completion time scheduling problem with availability constraints. *Eur. J. Oper. Res.* **2005**, *161*, 3–10. [CrossRef]
24. He, Y.; Zhong, W.Y.; Gu, H.K. Improved algorithms for two single machine scheduling problems. *Theor. Comput. Sci.* **2006**, *363*, 257–265. [CrossRef]
25. Breit, J. Improved approximation for non-preemptive single machine flow-time scheduling with an availability constraint. *Eur. J. Oper. Res.* **2007**, *183*, 516–524. [CrossRef]