

EVALUATION OF TRAIN SCHEDULE BY USING INFORMATION THEORY

K. MORO & Y. SUGAI

Graduate School of Engineering, Chiba University, Japan.

ABSTRACT

Convenience is an important requirement of a scheduled transit system. The authors proposed a new method to evaluate a train schedule by using entropy of the information theory. By the aspect of statistical mechanics, the randomness of a schedule is associated with the convenience of the schedule. However, the way of using entropy requires special conditions. In this paper, we show that we can relax the conditions by interpreting the entropy as information.

First, we explain the representation of a schedule by a tree structure. Introducing the tree structure, we recognize the more composite construction of schedules. The tree structure leads to the concept of conditional entropy. Using the concept of conditional entropy, we can evaluate schedules of the states of different places or times. Secondly, we can understand what the information of the schedule is. We can grasp that the information is the benefit of knowing the schedule. The increase of the entropy corresponds to the increase of the benefit. Moreover, it becomes possible to express an unexpected disbenefit by the entropy with the loss of information. In other words, the benefit is a positive entropy and the disbenefit is negative entropy. Finally, we show that we are able to deal with an influence of wasteful time in a schedule, such as an unnecessary waiting time and an extra duration by using the entropy. It is very important that we can evaluate negative effects, namely the discrepancy of supply and demand from the point of view of time.

Keywords: entropy, evaluate timetable, loss of information, tree structure.

1 INTRODUCTION

The schedule of a transit system, such as a railway network, has a great influence on the convenience of the transit system. We sense the transit system is convenient in that a great number of trains run and the distribution of departures is even. In other words, we can get a train in a short time at anytime we come in a station. Moro and Sugai [1] presented the method of evaluating schedules from the viewpoint of the number of trains and the distribution of departures. This method uses Shannon's entropy [2] of the information theory. Nevertheless it assumes that each passenger chooses his/her train. It does not cope with a case that a passenger cannot select any train.

In this paper, we introduce a method that deals with such cases, in which some passengers fail to select their trains, not only in cases in which passengers are able to select their trains. First, we show a way of expressing a schedule by a tree structure. Employing this structure we will be able to evaluate various states of schedules together by using entropy. Secondly, we explain what the information of the schedule is. By using the information of the schedule, we will be able to evaluate the adverse influence of the schedule according to its entropy. Finally, we show that we are able to deal with the influence of wasteful time in a schedule, such as unnecessary waiting time and an extra duration by the entropy.



This paper is part of the proceedings of the 15th International Conference on Railway Engineering Design and Operation (COMPRAIL)

www.witconferences.com

2 ENTROPY AND INFORMATION OF THE INFORMATION THEORY

In the beginning, we will explain Shannon’s entropy and information of the information theory [2].

2.1 Definition

Let X be a discrete random variable on a finite set $\tilde{X} = \{x_1, \dots, x_n\}$, with probability distribution function $p(x) := Pr(X = x)$.

2.1.1 Information

We introduce information as a means to measure a rarity of an event. Information of the event $x \in \tilde{X}$ with probability $p(x)$ is defined as

$$I(x) := -\log p(x). \tag{1}$$

2.1.2 Shannon’s entropy

The entropy $H(X)$ of X is defined as

$$H(X) := -\sum_{x \in X} p(x) \cdot \log p(x). \tag{2}$$

The convention $0 \log 0 = 0$ is adopted in the definition. The logarithm is usually taken to the base 2, in which case the entropy is measured in ‘bits’. This definition shows that we can calculate entropy only if we define the probability.

2.2 Implications of entropy

Shannon’s entropy of X corresponds to the expected value of the information of X .

3 EXPRESSION OF SCHEDULE BY THE TREE STRUCTURE

In this section we introduce the expression of a departure schedule at a station by a tree structure.

3.1 Definition

Let S be a schedule consisting of trains as

$$S = \{tr_1, tr_2, \dots, tr_n\}. \tag{3}$$

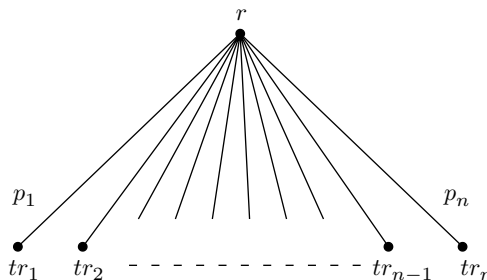


Figure 1: Tree structure of a schedule.

S leads a graph $T_S = \{N_S, A_S\}$. N_S is the set of trains and root node r . A_S is the set of arcs. For every $a_i = (r, tr_i) \in A_S$ is a pair of nodes of N_S with probability p_i . The tree structure of the schedule is represented in Fig. 1. In addition, we call the node described at the upper *parent* and at the lower *child*.

3.2 Nested schedule

Now, we think about integrated plural schedules. Consider the integrated schedules of the days of week for example. Suppose that the trains $\{tr_k\}^i$ are scheduled in the i th day of the week d_i . Moreover, the $p_j^{(i)}$ is the probability of train $tr_j \in \{tr_k\}^{(i)}$ in the day d_i . Furthermore, we assume that π_i is the ratio of ridership on the day d_i to the ridership of whole term. Namely, the set of trains $\{tr_k\}^i$ nests under the day d_i , and the set of days makes an intermediate class.

Specifically, we can say that D is a discrete random variable on the finite set $\tilde{D} = \{d_1, \dots, d_n\}$ with probability distribution function $\pi_i = Pr(D = d_i)$. Besides, we can think of d_i as the set of the trains $\{tr_k\}^i$. Additionally, $p_j^{(i)}$ is the conditional probability under d_i , and what is more, if tr_j is not a member of d_i , then $p_j^{(i)} = 0$.

Figure 2 indicates a nested schedule. The intermediate class E contains two elements: η_1 and η_2 . The elements η_1 and η_2 are formed as follows:

$$\eta_1 = \{e_1, e_2, e_3\}, \quad \eta_2 = \{e_4, e_5\}. \tag{4}$$

The entropies of E , η_1 and η_2 are

$$H(E) = - \sum_{i \in \{1,2\}} \pi_i \cdot \log \pi_i, \tag{5}$$

$$H(\eta_1) = - \sum_{j \in \{1,2,3\}} p_j^{(1)} \cdot \log p_j^{(1)}, \tag{6}$$

$$H(\eta_2) = - \sum_{j \in \{4,5\}} p_j^{(2)} \cdot \log p_j^{(2)}. \tag{7}$$

Then the entropy of the entire system S becomes as follows:

$$H(S) = H(E) + \sum_{i \in \{1,2\}} \pi_i \cdot H(\eta_i). \tag{8}$$

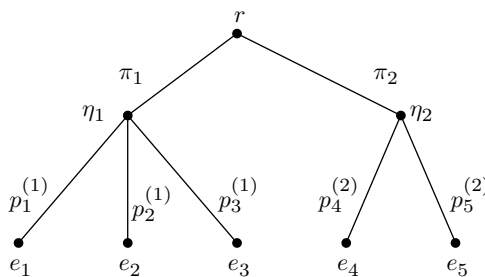


Figure 2: Nested schedule.

4 INTERPRETATION OF THE INFORMATION OF A SCHEDULE

We will discuss the relation between the convenience of a schedule and the information.

4.1 Relation between convenience and information

Moro and Sugai asserted that the entropy increases when the convenience of schedule increases [1]. That is, the entropy of the useful schedule is larger than that of the less-useful schedule. Moreover, the entropy means the expected value of the information in information theory. From this viewpoint, one may say that the more convenient the schedule of transit system is, the more its information grows.

4.2 Missing information

In a previous study [1], only the case in which all passengers can select their trains was explained. However, some passengers fail to catch a train occasionally, e.g. the passenger arrives at the station after the last train has departed.

We take it for granted that the fact that some passengers miss their trains reduces the benefit provided by the transportation system. In other words, when some passengers miss their trains, the entropy must decrease compared to the case when all passengers catch trains successfully. Therefore, we propose an interpretation that the event which is disbenefit has negative information, i.e. the system loses information by worthless events.

4.2.1 General definition

How shall we calculate the loss of information? We consider the schedule as in Fig. 3. There are n trains planned in this schedule. The probability of train tr_i is p_i . And a child node labelled *failure* exists. This node corresponds to the passengers who catch no train. Its probability is p_f . When we suppose that the all-child nodes are useful choices, the entropy is as follows:

$$H_F^\dagger = - \sum_{i=0}^n p_i \cdot I_i + p_f \cdot I_f. \tag{9}$$

The value I_i denotes information of an event with probability p_i as

$$I_i = -\log p_i. \tag{10}$$

However, the information of *failure* is added in eqn (9), i.e. the disbenefit choice improves the convenience. Hence, we define that we do not add the information of unusable choices when we calculate the entropy of a schedule; namely, we dismiss the information of worthless choices. For instance, the entropy of the schedule as in Fig. 3 is as follows:

$$H_F^* = \sum_{i=1}^n p_i \cdot I_i. \tag{11}$$

4.2.2 Representation by nested tree

We will represent a schedule which is the same as Fig. 2 by a nested tree structure as in Fig. 4.

There is an intermediate class consisting of *success* and *failure*. The probability of *success* is π_s and the probability of *failure* is π_f . The conditional probability under *success* of the train tr_i is \hat{p}_i .

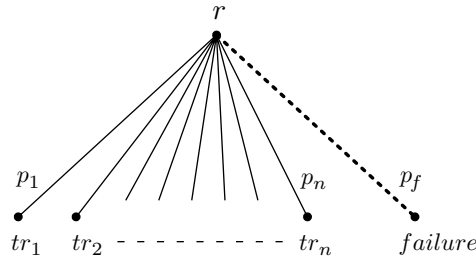


Figure 3: Tree structure with failure.

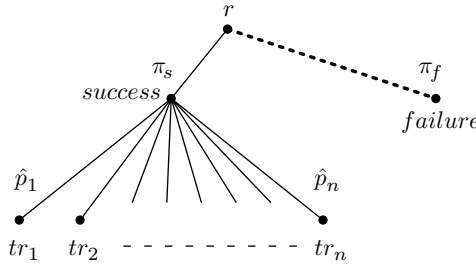


Figure 4: Nested tree structure with failure.

4.2.3 Example of calculation

We will calculate the entropy as follows in Table 1. Probabilities of the event of *success* and *failure* are as follows:

Table 1: Number of passengers: with failure.

	tr_1	tr_2	tr_3	tr_4	tr_5	tr_6	tr_7	tr_8	Failure	Total
Number of passengers	250	255	235	246	242	253	257	234	17	1,989
Subtotal of <i>success</i> 1,972										

$$\pi_s = 0.9915, \tag{12}$$

$$\pi_f = 0.0085. \tag{13}$$

Then the conditional probability \hat{p}_i of tr_i under *success* and information I_i of tr_i calculated by \hat{p}_i are as in Table 2.

The entropy of *success* is

$$H_s = \sum_{i=1}^8 \hat{p}_i \cdot I_i = 2.9992. \tag{14}$$

The entropy of the whole is obtained as

$$\begin{aligned} H &= -\pi_s \cdot \log \pi_s + \pi_s \cdot H_s \\ &= 0.0123 + 0.9915 \cdot 2.9992 = 2.9860. \end{aligned} \tag{15}$$

Table 2: Probabilities and information: *success*.

Probability							
\hat{p}_1	\hat{p}_2	\hat{p}_3	\hat{p}_4	\hat{p}_5	\hat{p}_6	\hat{p}_7	\hat{p}_8
0.1268	0.1293	0.1192	0.1247	0.1227	0.1283	0.1303	0.1187
Information							
I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8
2.9797	2.9511	3.0689	3.0029	3.0266	2.9625	2.9398	3.0751
Probability \times Information							
$\hat{p}_1 * I_1$	$\hat{p}_2 * I_2$	$\hat{p}_3 * I_3$	$\hat{p}_4 * I_4$	$\hat{p}_5 * I_5$	$\hat{p}_6 * I_6$	$\hat{p}_7 * I_7$	$\hat{p}_8 * I_8$
0.3777	0.3816	0.3657	0.3746	0.3714	0.3801	0.3831	0.3649

Including disbenefit phenomena reduces entropy.

5 EVALUATION OF WASTEFUL TIME

Nobody expects to waste time. Unexpected unuseful occurrences must reduce the entropy of a transit system. In this section, we explain the influence of wasteful time on the entropy.

5.1 Assumption

Previous studies of the evaluation of a transit system by its entropy have focussed only on the case in which the capacity of each train is infinity. Now, we would like to focus on a case in which the limited capacity of a train creates unexpected wasted time.

We consider that all passengers arrive at the station according to uniform distribution. We suppose some trains are closely packed before their departure.

5.2 Amount of waiting time

Under the conditions as stated above, we can calculate the amount of waiting time as in Fig. 5. We denote c is the capacity of the train. An empty train fills up with term $l^* > 0$. This train leaves station with term l later after filling up. The total waiting time of each passenger before the train fills up becomes w^* and after filling up becomes w . The value w is just equivalent to an idle time.

Now we can express w , w^* and $\exists \beta > 0$ as follows:

$$w = l \cdot c = \beta \cdot l^* \cdot c, \tag{16}$$

which denotes $l = \beta \cdot l^*$. Similarly, there exists $\alpha > 0$ such that

$$w^* = \alpha \cdot l^* \cdot c. \tag{17}$$

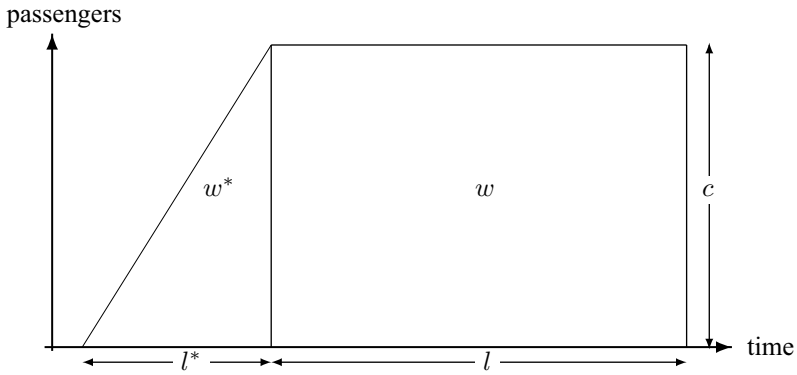


Figure 5: Amount of waiting time.

5.3 Ratio of effective time

We define the effective time rate r as

$$r := \frac{w^*}{w^* + w} \tag{18}$$

If the train left before filling up, then we define $r := 1$. For the following discussions, we consider the condition to be limited to the prior condition.

Due to eqns (16) and (17), we can modify the effective time rate as follows:

$$r = \frac{\alpha \cdot l^* \cdot c}{\alpha \cdot l^* \cdot c + \beta \cdot l \cdot c} = \frac{\alpha}{\alpha + \beta} \tag{19}$$

5.4 Information of train with unexpected wasted time

We can express the part of train tr_i in a tree structure of schedule with probability p_i as in Fig. 6. The information of tr_i without useless time is

$$I_i = -\log p_i \tag{20}$$

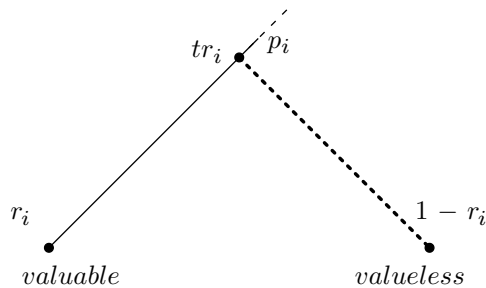


Figure 6: Example: with wasteful time.

We define that the information-reflected wasted time is calculated as

$$I_i^\dagger := -\log p_i^\dagger, \tag{21}$$

which denotes $p_i^\dagger := r_i \cdot p_i$. Furthermore, we define the entropy-reflected wasted time as

$$H^\dagger := \sum_i p_i^\dagger \cdot I_i^\dagger. \tag{22}$$

5.5 Influence of wasted waiting time

We calculated entropy of schedules with some variations of wasted waiting time.

5.5.1 Preconditions

We assume some preconditions to calculate entropy.

- Three trains are planned. The capacity of each train is the same.
- All passengers arrive at the origin station for first 15 min.
- The number of passengers equals 3 times the capacity of a train.
- Distribution of passengers arriving at the origin station is uniform, i.e. $\alpha = \frac{1}{2}$.

We will calculate the entropy about the schedules (a) to (f) as in Fig. 7.

5.5.2 Case (a): optimal schedule

The times of departure are {5, 10, 15} and the rates are {1, 1, 1}. All trains planned in (a) depart immediately when each train fills up. That is the optimal schedule.

The entropy is as follows:

$$H_{(a)} = \left(-\frac{1}{3} \cdot \log \frac{1}{3}\right) \cdot 3 = \log 3 = 1.585. \tag{23}$$

5.5.3 Case (b)

The times of departure are {5, 10, 20} and the rates are $\left\{1, 1, \frac{1}{2}\right\}$.

The entropy is as follows:

$$\begin{aligned} H_{(b)} &= \left(-\frac{1}{3} \cdot \log \frac{1}{3}\right) \cdot 2 - \left(\frac{1}{3} \cdot \frac{1}{3}\right) \cdot \log \left(\frac{1}{3} \cdot \frac{1}{3}\right) \\ &= \frac{8}{9} \cdot \log 3 = 1.409. \end{aligned} \tag{24}$$

5.5.4 Cases (c) to (f)

The entropies of case (c) to (f) are as follows:

$$\begin{aligned} H_{(c)} &= 1.371, & H_{(d)} &= 1.266, \\ H_{(e)} &= 1.141, & H_{(f)} &= 1.049. \end{aligned} \tag{25}$$

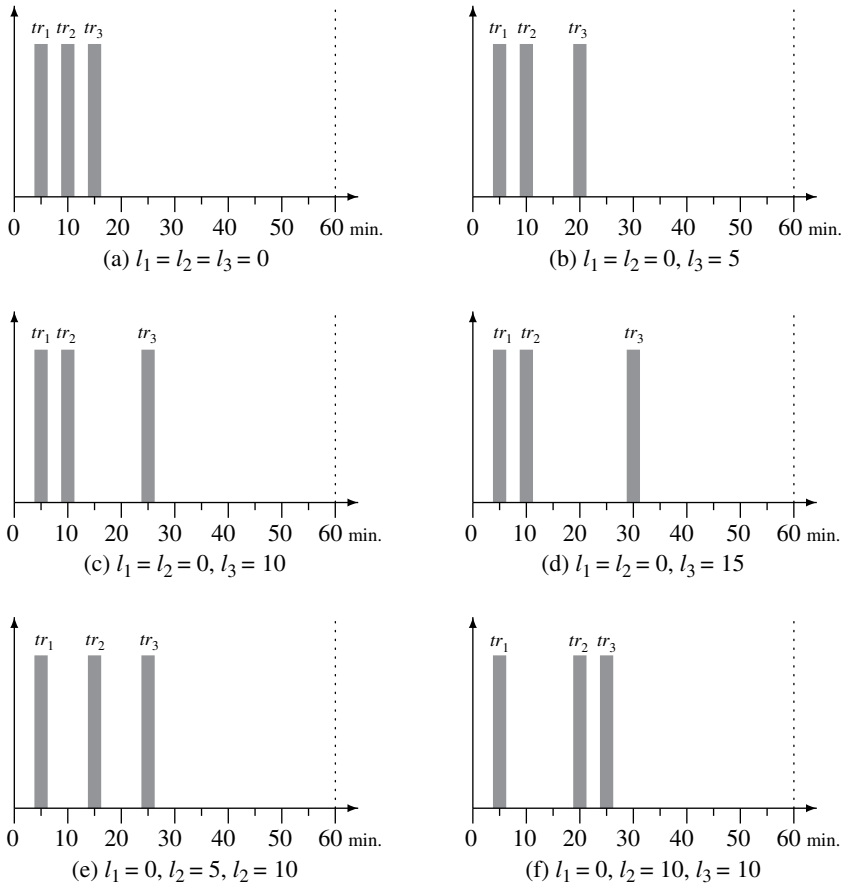


Figure 7: Timetables, for example.

5.6 Conclusions concerning the influence of wasted waiting time

Comparing with the results of (a), (b), (c) and (d), the entropy diminishes as the waiting time increases. Furthermore, comparing with the results of (d) and (e), when more passengers have to wait, the entropy decreases when the total waiting time equals each other. Thus we conclude that the proposed method is reasonable to evaluate the unexpected wasted time of a schedule.

5.7 Nested schedule with loss of information

We show a calculation example of a nested schedule with loss of information. We will consider the root with transfer at a station. Three trains are planned on each section. All passengers boarding tr_i of the predecessor section transfer to tr_i of the successor section as in Fig. 8.

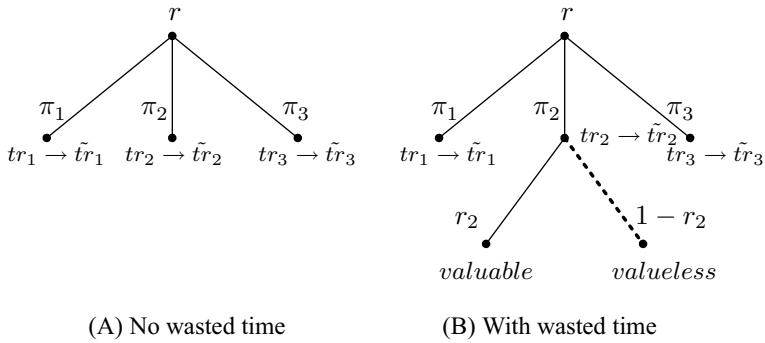


Figure 8: Nested tree structure with wasteful time.

5.7.1 Assumptions

We assume some conditions of the schedule.

- The minimum transfer time at the station is 1 min, i.e. $l^* = 1$.
- The value α in eqn (17) is $\frac{1}{2}$.
- The number of passengers on each train is equal, i.e. $\pi_i = \frac{1}{3}$.

5.7.2 Optimal transfer

We consider the schedule (A) in Fig. 8 as the optimal connection. That is, all passengers retain at the station for only the minimum transfer time. The entropy is

$$H_{(A)} = \left(-\frac{1}{3} \cdot \log \frac{1}{3}\right) \cdot 3 = \log 3 = 1.585. \tag{26}$$

5.7.3 With wasted waiting time

See the schedule (B) in Fig. 8. Suppose the transfer time from tr_2 to tr_2 has a valueless time that equals the minimum transfer time, i.e. $l = l^*$. The value β in eqn (16) becomes 1 and the effective time rate r_2 becomes $\frac{1}{3}$.

We calculate the entropy of wasted time as

$$\begin{aligned} H_w &= -\pi_2 \cdot r_w \cdot \log(\pi_2 \cdot r_w) \\ &= r_w \cdot (-\pi_2 \cdot \log \pi_2) + \pi_2 \cdot (-r_w \cdot \log r_w) \\ &= r_w \cdot (\pi_2 \cdot I_{\pi_2}) + \pi_2 \cdot (r_w \cdot I_{r_w}), \end{aligned} \tag{27}$$

which denotes $r_w := 1 - r_2$. The entropy is

$$\begin{aligned}
 H_{(B)} &= H_{(A)} - H_w \\
 &= \log 3 - \frac{2}{3} \cdot \frac{1}{3} \cdot \left(-\log \frac{1}{3}\right) - \frac{1}{3} \cdot \frac{2}{3} \cdot \left(-\log \frac{2}{3}\right) \\
 &= \log 3 - \frac{2}{9} \cdot \log 3 - \frac{2}{9} \cdot (\log 3 - 1) \\
 &= \frac{5}{9} \cdot \log 3 + \frac{2}{9} = 1.103.
 \end{aligned} \tag{28}$$

This result shows that we must not forget reduction in the entropy of the predecessor trains.

6 CONCLUSIONS

In this paper, we propose the interpretation that the entropy of a schedule is a metric of the usefulness of the schedule. Moreover, we can evaluate the disutility of schedule by using the proposed concept. In other words, we can express the benefit of a schedule as positive entropy and the disbenefit as negative entropy. In addition, the interpretation of a schedule as information raises the possibility that the techniques of the information theory are applicable to the evaluation of schedules and timetable planning. For instance, we will be able to identify the disruption of a schedule with erasure rate in a channel in information theory. And this identification leads to the explanation that we can consider the recovery of a schedule as correcting errors of received message.

REFERENCES

- [1] Moro, K. & Sugai, Y., A relation between a robustness of the disruption of the train schedule and the entropy of information theory. *6th International Conference on Railway Operations Modelling and Analysis*, International Association of Railway Operation Research, pp. 074–1–19, 2015.
- [2] Shannon, C.E., A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, **5(1)**, pp. 3–55, 2001. DOI: [10.1145/584091.584093](https://doi.org/10.1145/584091.584093).