

# 混合分布の最適化による交通量の推定

Estimation of Traffic Volume by Optimizing Mixture Distributions

○ 太刀岡 勇気<sup>1</sup>

○ Yuuki Tachioka<sup>1</sup>

<sup>1</sup> デンソーアイティラボラトリ

<sup>1</sup>Denso IT Laboratory

**Abstract:** The evolutionary computation competition in 2022 introduced an estimation task of traffic volume of pedestrians by using simulation. This task requires participants to estimate the departure time and the number of people of three types of pedestrians (slow, busy, guided), respectively. In this report, we propose to model traffic volume as a mixture distribution and to optimize the center point, width, strength, distribution type, and the width of Gaussian filter by using evolutionary computation algorithms. For one of the single objective optimization problems, our methods came second, which shows the effectiveness of our proposed method.

## 1 はじめに

大規模イベントでの事故防止などを目的として、デジタルツイン技術を用いた高精度な交通量の推定が求められている [1]。進化計算コンペ 2022[2] では、関門海峡花火大会<sup>1</sup>をとりあげて、人流のマルチエージェントシミュレーションを用いた交通量の推定が課題となっていた。これは高速なマルチエージェントシミュレーター CrowdWalk [3] により、それぞれの観客をエージェントとして、その動きをシミュレーションしたもので、MAS-Bench [4] を元としている。マルチエージェントシミュレーションでは、エージェントごとに異なる特性を与えることができる。ここではエージェントの特性は、誘導指示に従うエージェント (Guided) に加え、誘導指示に従わない、駅への最短経路を選択するエージェント (Busy)、お祭り会場となっている長い経路を選択するエージェント (Slow) の 3 種類とした [5]。Guided を基準として、Busy は移動速度が 1.1 倍、Slow は 0.5 倍である。シミュレーションの精度はエージェントの情報の入力精度に依存しているため、正しくエージェントの情報「いつ」「どこからどこへ」「どのような人が」移動したのかを入力する必要がある。進化計算コンペ 2022 では経路 (「どこからどこへ」) は固定のため、「いつ」「どのような人が」移動したのかを推定する。すなわち、上述の 3 つの異なるタイプ (Guided, Busy, Slow) の歩行者が、いつ何人出発するかを推定することが求められる。ユーザーが決定変数を指定するとブラックボックスでシミュレーターが動き、目的関数値が返される。その際にカメラと GPS トラッカーにより得られた限られた観測データとの誤差が最小となるシミュレーション結果を出力できるパラメータを求める単一目的関数タスクが出題された。本報では、混合分布を用いてエージェントの交通量をモデル化し、その中心点、幅、強度、分

布の種類、ガウシアンフィルタの幅を進化計算アルゴリズムにより最適化する方法を提案する。単一目的関数タスクで 19 チーム中 2 位となり、提案法の有効性が示された。

## 2 進化計算による混合分布の最適化

### 2.1 混合分布

時刻  $t (= 1, \dots, T)$  に出発するエージェント数を推定することが求められている。ここでは任意の関数  $f_{i,k}^*(t) (i = 1, 2, \dots)$  からなる混合分布 (1) から、エージェントのヒストグラムを生成することとする<sup>2</sup>。

$$f_k^*(t) = \sum_i f_{i,k}^*(t) \quad (1)$$

ここで  $k$  はエージェントのタイプ毎の属性で  $k \in \{Slow, Busy, Guided\}$  である。式 (1) の\*としては以下の分布が使用できる。

中心点  $c_{i,k}$ 、幅  $w_{i,k}$ 、強度  $\eta_{i,k}$  をパラメータとして、分布として一様分布を使う場合には式 (2) のようになる。

$$f_{i,k}^u(t) = \begin{cases} \eta_{i,k} & (c_{i,k} - w_{i,k} \leq t < c_{i,k} + w_{i,k}) \\ 0 & (\text{otherwise}) \end{cases} \quad (2)$$

ガウス分布を使う場合には式 (3) のようになる。

$$f_{i,k}^g(t) = \eta_{i,k} \exp\left(-\frac{(t - c_{i,k})^2}{2w_{i,k}^2}\right) \quad (3)$$

正規分布を使う場合には式 (4) のようになる。

$$f_{i,k}^n(t) = \frac{1}{\sqrt{2\pi}w_{i,k}} f_{i,k}^g(t) \quad (4)$$

### 2.2 閾値によるスパース化

連続な分布を用いるとヒストグラムがスパースにならない。エージェントの出発数はスパースであることが知

<sup>2</sup> デモ問題で複数の分布を用いる有効性を確認した。

<sup>1</sup> <https://kanmon-hanabi.love/modules/guide/index.php/jikkoiin.html>

られているため、閾値  $\theta$  による閾値処理でこれを行う。

$$f_{i,k}^*(t) = \begin{cases} f_{i,k}^*(t) & (f_{i,k}^*(t) \leq \theta_k) \\ 0 & (\text{otherwise}) \end{cases} \quad (5)$$

閾値  $\theta_k$  は  $f_k^*$  の最小値

$$\theta_k = \min_t f_k^*(t) \quad (6)$$

とした。最後に、分布を滑らかにするために、 $f_k^*$  に幅  $\sigma$  のガウシアンフィルタを適用する。

図 1 に一様分布を用いた場合の流れを例に示す。それぞれのパラメータに対応する一様分布を 3 つ用意する (図 1(a))。これを加算し、閾値処理を行う (図 1(b))。ガウシアンフィルタを適用して滑らかにしたのち整数化し、ノイズを加算して次項に述べる総数の制約 (2.3.1) と最大数の制約 (2.3.2) を満たすようにすることで最終的なヒストグラムが求まる (図 1(c))。

### 2.3 制約条件の考慮

$f_k^*$  は実数値であるため、ヒストグラム  $h_k(t)$  を生成するには整数化する必要がある。その際に、ヒストグラム  $h_k(t)$  は以下の 2 つの制約条件を満たさなければならない。

#### 2.3.1 総数の制約

エージェントの総数が  $N_A$  であることの制約であり、式 (7) であらわされる。

$$\sum_t \sum_k h_k(t) = N_A \quad (7)$$

これを満足するために

$$\alpha \sum_t \sum_k f_k^*(t) = N_A \quad (8)$$

を満たす  $\alpha$  を求め、 $f_k^*$  を  $\alpha$  倍する。これを

$$h_k(t) = \lfloor \alpha f_k^*(t) \rfloor \quad (9)$$

のように整数化して、ヒストグラムに変換する。 $\lfloor \cdot \rfloor$  はフロアリング関数である。これにより

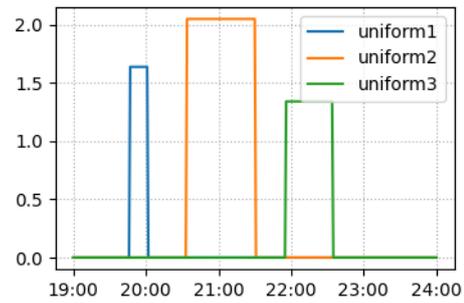
$$N_A - T \sum_k 1 < \sum_t \sum_k h_k(t) \leq N_A \quad (10)$$

が満たされる。さらに、不足分

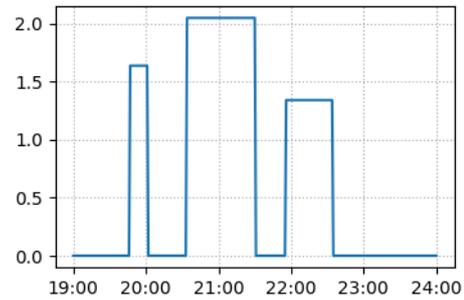
$$Q = N_A - \sum_t \sum_k h_k(t) \quad (11)$$

に関しては、これを補償するようにランダムにバイナリ値  $\delta_k(t) = \{0, 1\}$  を付加する。ここで

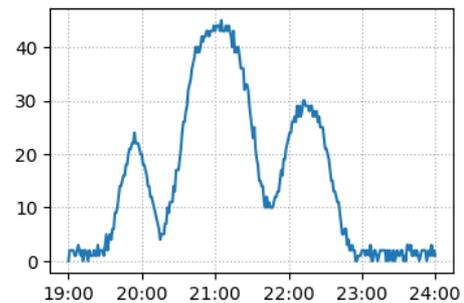
$$Q = \sum_t \sum_k \delta_k(t) \leq T \sum_k 1 \quad (12)$$



(a) 個々の関数  $f_i^u(t)$



(b) 加算値  $f^u(t)$



(c) ガウシアンフィルタ+正規化+ノイズ加算により得られたヒストグラム  $h(t)$

図 1: 混合分布 (強度を変えて加算).

である。これにより最終的にヒストグラムを  $h_k(t) = h_k(t) + \delta_k(t)$  のようにすることで、制約 (7) が満たされる。

#### 2.3.2 時刻当たりの最大数の制約

時刻当たりの最大数が  $N$  以下であることの制約であり、式 (13) であらわされる。

$$\max_t \sum_k h_k(t) \leq N \quad (13)$$

これを満足するためにはヒストグラムを  $\beta (< 1)$  乗して平滑化する。

$$\max_t \sum_k (h_k(t))^\beta \leq N \quad (14)$$

その後再び、2.3.1 の制約を満たすようにする。

### 3 実験

#### 3.1 実験条件

単一目的の最適化問題であり、以下に示す目的関数を最小とするような決定変数を進化計算により発見する。

##### 3.1.1 決定変数

1分ごとの各エージェントの出発人数(ヒストグラム)を決定変数とし、19時から24時までの300分間における出発人数を整数で指定する。決定変数の数は $300 \times$  エージェントタイプ数となる。その際に2.3.1と2.3.2の制約を満たす必要があり、制約を満たさない解は評価されない。

##### 3.1.2 目的関数

正解となるデータは、実際の大会でカメラとGPSにより観測されたデータを用いた。カメラによる計測では門司港駅に到着した人数を時間ごとに観測できる。到着人数から経路内に滞在するエージェントの数を推定できると同時に、門司港駅は全エージェントの最終的な到着点であるため、エージェントの総数も求めることができる。GPSによる計測では10分ごとに各ルートに1名ずつの計測員がGPSトラッカーを持ち周りのエージェントのペースに合わせて歩行することで時間当たりの積算移動距離を計測した。

カメラ計測に基づく到着人数に関する誤差

$$\epsilon^d = \sqrt{\frac{1}{T} \sum_{t=1}^T |D(t) - \hat{D}(t)|^2} \quad (15)$$

とGPS計測に基づく移動軌跡に関する誤差

$$\epsilon^p = \frac{1}{A} \sum_{a=1}^A \sqrt{\frac{1}{t_a^d - t_a^o + 1} \sum_{t=t_a^o}^{t_a^d} |P_a(t) - \hat{P}_a(t)|^2} \quad (16)$$

を算出し、それらの平均 $\epsilon = 0.5(\epsilon^d + \epsilon^p)$ を目的関数とした。ここで $D$ は時刻 $t$ に実際に到着した人数、 $\hat{D}$ はシミュレーション上で到着した人数である。また、 $P_a$ は時刻 $t_a^o$ に出発したエージェント $a$ が $t_a^d$ に到着するまでに移動した距離、 $\hat{P}$ はシミュレーション上での移動距離である。計測員は一人ずつ10分ごとに3ルートに出発するため、 $A = 90$ である。

##### 3.1.3 進化計算

分布は3混合とし、パラメータは中心点、幅、強度、分布の種類、ガウシアンフィルタの幅の最適化とした。ただし混合分布間で、順序が入れ替わるだけの冗長な最適化を避けるため、 $c_{1,k} < c_{2,k} < c_{3,k}$ となるように中心点をサンプリングした。これらのパラメータを

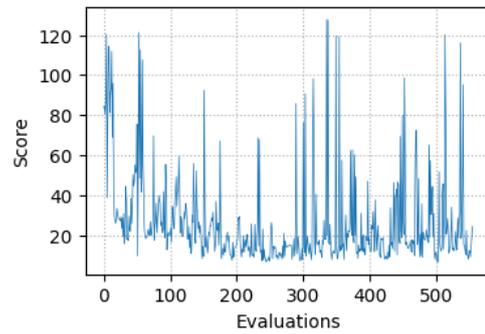


図 2: 目的関数値の推移 (sop1)

表 1: 最適解 (sop1)

評価数	目的関数値	分布	$\sigma$
502	6.573	normal	10
$c_1$	$w_1$	$\eta_1$	
54	7	1.637	
$c_2$	$w_2$	$\eta_2$	
122	28	2.049	
$c_3$	$w_3$	$\eta_3$	
195	19	1.341	

Optuna [6] の TPE sampler を用いて最適化した。ノイズ  $\delta(t)$  はランダムな時刻に 1 を付加して総数を調整した。最大数  $N$  が 45 を超えた場合は、ヒストグラムの  $\beta = 0.9, 0.9^2, 0.9^3, \dots$  を条件を満たすまで行った。

### 3.2 結果と考察

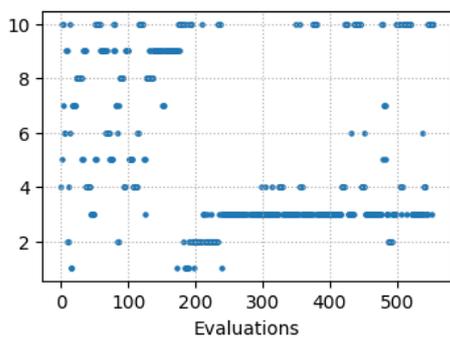
#### 3.2.1 sop1

sop1 は guided エージェントのみが存在する条件 ( $k \in \{Guided\}$ ) でのシミュレーションである。総数  $N_A = 4,660$  であり、時刻ごとの上限は  $N = 45$  である。目的関数値を図 2 に示す。スパイクしながらも徐々に誤差が低下し、502 回目に最小の目的関数値を得ている。

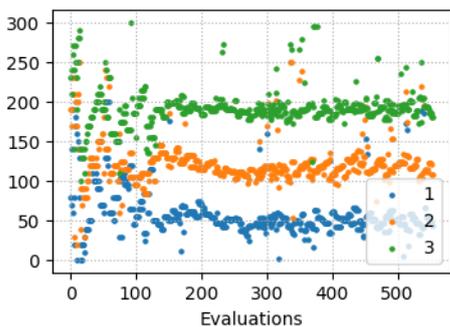
最適化された結果を図 3 に示す。図 3(a) のガウシアンフィルタの幅は、ほぼ 3 から 10 が最適であることがわかる。中心点・幅に関しても図 3(b)(c) に示すように、ほぼ収束していることがわかる。図 3(d) の強度に関しては、多少のばらつきがみられる。

最終的に得られた最適解を表 1 に示す。正規分布が選択されたことがわかる。この目的関数値は参加 19 チーム中 2 位の結果であり、提案法の有効性を示せた。

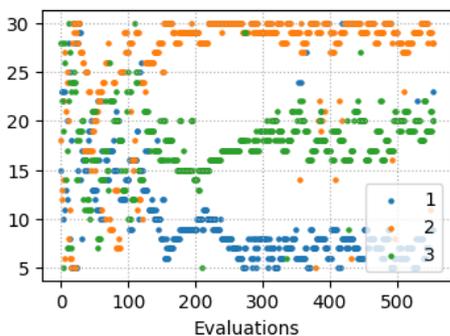
図 4(a)(b)(c) には、最適解のパラメータにおける分布ごとの生成されたヒストグラムの違いを示す。一様分布がもっともピークがはっきりしており、正規分布は幅の狭い分布ほど高さが高くなっている点がガウス分布と異なる。図 4(d) には評価回数に応じてどの分布が選択



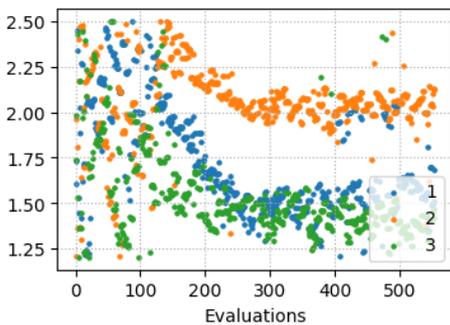
(a) ガウシアンフィルタの幅  $\sigma$



(b) 中心点  $c$



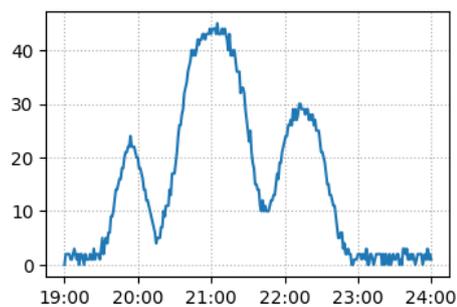
(c) 幅  $w$



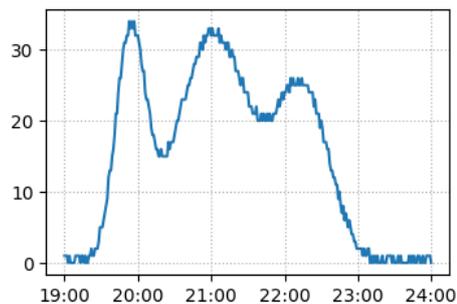
(d) 強度  $\eta$

図 3: 混合分布の最適化 (sop1)

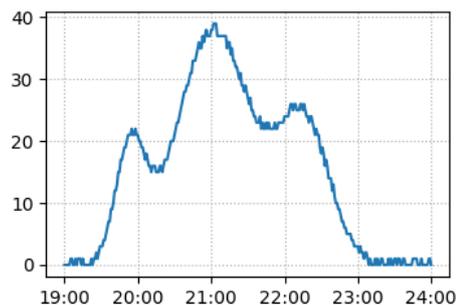
されたかを 20 回の評価ごとに集計して示している。このようにはじめのうちは 3 つの分布が等しく表れるが、



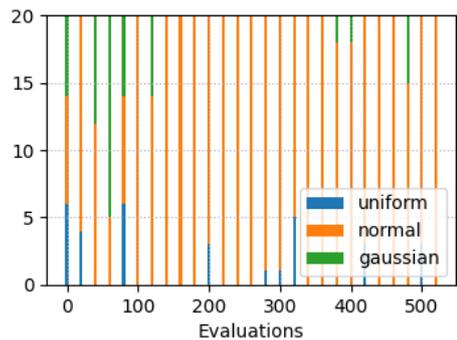
(a) 一様分布  $f^u$



(b) 正規分布  $f^n$



(c) ガウス関数  $f^g$



(d) 割合

図 4: 分布の種類を選択 (sop1)

最適化が進むにつれて正規分布が支配的となることがわかる。生成されたヒストグラムから、ピークはそれぞれ 20 時、21 時、22 時にあり、キリの良い時間に移動を開始する観客が多いことと、花火の打ち上げが 19:50

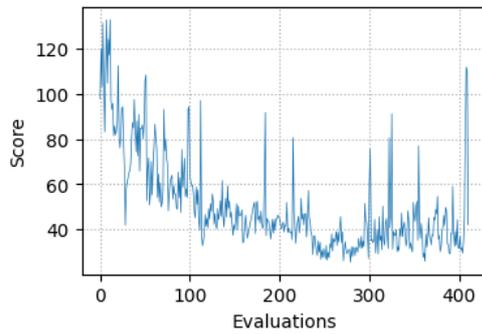


図 5: 目的関数値の推移 (sop2)

表 2: 最適解 (sop2)

評価数			目的関数値			$\sigma$		
279			25.427			8		
$c_1$	$w_1$	$\eta_1$	s	b	g	s	b	g
74	68	127	30	6	7	1.652	1.255	1.200
$c_2$	$w_2$	$\eta_2$	s	b	g	s	b	g
89	82	137	12	27	22	1.422	1.388	1.789
$c_3$	$w_3$	$\eta_3$	s	b	g	s	b	g
139	119	188	23	26	12	1.509	2.059	1.828

から 20:20 に行われるため、その付近で出発人数が急減していることがわかる。

### 3.2.2 sop2

sop2 は上述の 3 種類のエージェントが存在する条件 ( $k \in \{Slow, Busy, Guided\}$ ) でのシミュレーションである<sup>3</sup>。総数  $N_A = 4,079$  であり、上限は同じく  $N = 45$  である。 $k \in \{Slow, Busy, Guided\}$  のそれぞれに 3 分布ずつ割り当てて (計 9 分布)、sop1 と同じように最適化した。目的関数値を図 5 に示す。こちらでもスパイクしながらも目的関数値は低下している。sop2 で提案法は 9 位となった。最終的に得られた最適解を表 2 に示す。 $\{s, b, g\}$  はそれぞれ Slow, Busy, Guided に対応している。sop2 でも、はじめは sop1 と同じように分布の種類も含めて最適化したが、最適化の回数制限内には最適化が終わりそうになかったため<sup>4</sup>、分布は sop1 で最もよかった正規分布に固定し、そのほかのパラメータの最適化を行った。

最適化された結果を図 6 に示す。図 6(a) より、ガウシアンフィルタの幅はほぼ 8 に収束している。図 6(b)

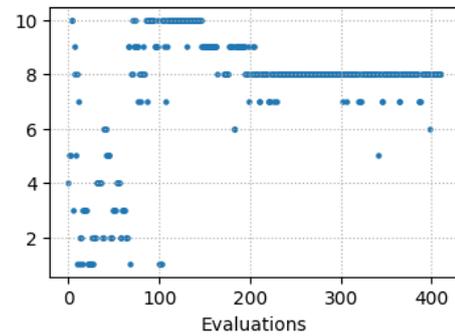
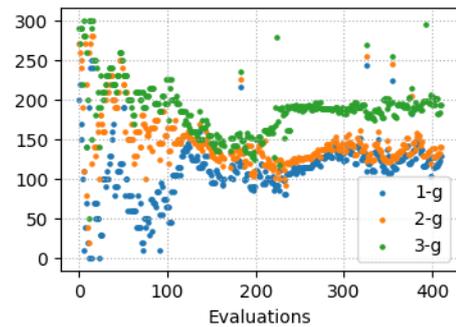
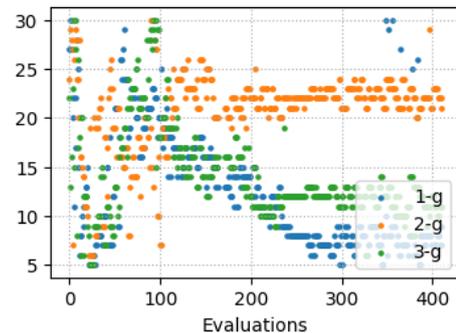
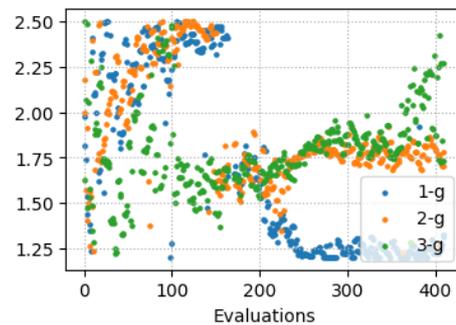
(a) ガウシアンフィルタの幅  $\sigma$ (b) 中心点  $c$ (c) 幅  $w$ (d) 強度  $\eta$ 

図 6: 混合正規分布の最適化 (sop2)

<sup>3</sup>なお sop1 とは別の年度の花火大会を対象としている。

<sup>4</sup>今回の進化計算コンペでは評価回数が全部で 1000 回までに制限されている。

の中心点に関しては収束しているものの、(c) の幅はまだばらついており、更なる最適化が必要であったと思わ

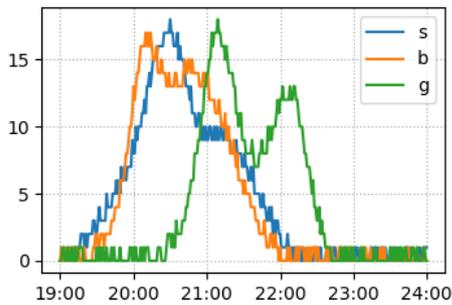


図 7: ヒストグラム  $h_k(t)(sop2)$

れる<sup>5</sup>。

図 7 にはエージェントタイプごとのヒストグラムを示す。Guided は 21:00 と 22:00 に出発しており、キリの良い時刻を好むことがわかる。Slow はそれよりも早く出発しゆっくりと出店を見ながら帰ることを意図したと思われる。Busy は花火の打ち上げ終了のタイミングにピークが来ており、花火の打ち上げを見終わるとともに帰路についたと考えられる。

#### 4 おわりに

混合分布により、エージェントの出発時刻ごとのヒストグラムを生成する問題に対して、分布の中心点と幅、強度、種類を進化計算で最適化する方法を提案した。提案法は sop1 で 2 位、sop2 で 9 位となり、汎用的で簡便な手法ながら実データのシミュレーションにも有用であることが分かった。

#### 参考文献

- [1] M. Zhou, H. Dong, B. Ning and F.-Y. Wang: “Recent development in pedestrian and evacuation dynamics: Bibliographic analyses, collaboration patterns, and future directions”, IEEE Transactions on Computational Social Systems, **5**, 4, pp. 1034–1048 (2018).
- [2] 谷垣, 重中, 鷹見, 大西, 濱田, 原田: “群集シミュレーションを用いた歩行者交通情報の推定”, 進化計算学会論文誌, to appear (2023).
- [3] T. Yamashita, T. Okada and I. Noda: “Implementation of simulation environment for exhaustive analysis of huge-scale pedestrian flow”, SICE Journal of Control, Measurement, and System Integration, **6**, 2, pp. 137–146 (2013).

<sup>5</sup>9 つの分布の結果を図示すると見づらくなってしまうため、guided のみの結果を示した。

- [4] S. Shigenaka, S. Takami, S. Watanabe, Y. Tanigaki, Y. Ozaki and M. Onishi: “MAS-Bench: Parameter optimization benchmark for multi-agent crowd simulation”, Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems, pp. 1652–1654 (2021).
- [5] R. Nishida, M. Onishi and K. Hashimoto: “Construction of a route choice model for application to a pedestrian flow simulation”, 2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), pp. 614–619 (2019).
- [6] T. Akiba, S. Sano, T. Yanase, T. Ohta and M. Koyama: “Optuna: A next-generation hyperparameter optimization framework”, Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 2623–2631 (2019).

#### 連絡先

太刀岡 勇気

E-mail: tachioka.yuki@core.d-itlab.co.jp