

# スマートビルにおけるトイレ個室利用状況の可視化と 予測手法の提案

小林美貴<sup>1</sup> 山下雄一郎<sup>1</sup> 洞井晋一<sup>1</sup>

**概要：** トイレの使用状況を可視化すると利用者の利便性を向上させることはできるが、混雑の緩和にはあまり影響を与えられない。混雑を緩和し設備の有効活用を実現するためには、どのトイレが最も早く空室になるかを利用者へ情報提供することで人員を誘導する必要がある。本研究では空室を予測するため、統計による予測方法と RandomForest を用いた予測方法を提案する。統計による予測方法は 30% 程度の回答率であったが、RandomForest を用いた予測手法の精度は 78% であった。この予測手法を用いれば、人員を誘導し混雑を緩和できると考えられる。

## Prediction and Visualization of Toilet Occupancy for Smart Buildings

MIKI KOBAYASHI<sup>1</sup> YUICHIRO YAMASHITA<sup>1</sup> SHINICHI DOI<sup>1</sup>

### 1. はじめに

商業施設やオフィス等の人が集まるところでは、トイレの混雑により待ち時間が発生することが多い。例えば男女の利用者の比率がフロアによって異なる場合、女性が多いフロアでは女性トイレに行列ができるが、男性の多いフロアでは女性トイレはほとんど利用されていない。混雑したトイレの行列に並ぶのではなく、別のフロアへ移動した方が待ち時間が短くなるというのはよく見られる光景である。個室の数を予め多く設計していればこのような問題は起こらないが、施設の管理者はトイレに与える面積をできる限り小さくして効率的に運用したいと考える。またトイレの増設は容易でないため、商業施設ではテナントの入れ替わりに伴う客層の変化も許容できるだけの個室数が必要となる。

トイレの個室数を可能な限り最小限に抑えながら利用者の利便性を向上するための取り組みとして、トイレの満空情報の可視化サービスがあり、配線・給電不要なセンサーを利用して百貨店や空港等で導入され始めている。トイレのドアに取り付けられた開閉センサーは EnOcean や LPWA 等を利用してサーバーやクラウドに開閉情報を送る。利用者はトイレ付近のサイネージ画面やスマートフォンで満空情報が閲覧でき、混雑を避けて空いているトイレを知ることができる。

しかし現在の満空情報を利用者に開示するに留まるものが多いため、近隣のトイレに空室があれば利用者に目的地を選択させることはできるが、すべて混雑している場合はどのトイレにも誘導することができない。道路や店舗の

混雑状態を予測するようにトイレの混雑状態を予測することができれば、休日の混雑した商業施設であっても利用者の利便性を向上できると考えられる。

そこでトイレの混雑状況を予測し、より効率的なトイレの運用を試みる。予測した混雑情報をリアルタイムの満空状況と合わせて利用者に提供することで、人員の誘導によるトイレの混雑緩和を目指す。フロアやエリアによる利用時間の偏り、外的要因（曜日、天気、気温等）による利用時間の変化を分析して混雑率を予測する方法を提案し、自社ビルにて予測手法の精度を評価する。

### 2. 自社ビルにおける可視化実験

自社ビルでも先に述べたようなトイレの混雑が発生しており、ビルの利便性についてのアンケートでもトイレの混雑についての声が多く挙がった。そこで混雑緩和および利便性向上のため使用状況の可視化の実証実験に取り組んだ。データを取得することで使用状況の現状を把握し、使用状況を利用者に開示することで使用率を均等化することを目指す。

11 階建ての自社ビルには各フロアに男女 1 か所ずつ 4 個室のトイレがあり、各フロアの従業員数は 200 人前後である。9,10,11 階の 3 フロア分の男性トイレを使って約 3 か月間の検証を行った。個室の扉に開閉センサーを取り付けてデータを収集し、センサー固有の ID、開閉状況、収集時刻のデータを蓄積した。収集したデータを分析したところ、フロアや時間帯によって利用回数や利用時間が異なり満室になりやすい時間帯やフロアがあることが分かった。例えば 10 時台は一人あたりの利用時間が長く満室になること

<sup>1</sup> 西日本電信電話株式会社

が多いが、12 時台は利用者数が多いにもかかわらず一人あたりの占有時間が短いため満室になることはあまりない。

これらの情報を元にトイレの使用状況を可視化し、トイレ前に設置したサイネージや自席の PC を使って図 1 のように利用者に開示したところ、利便性の向上について満足の声を多数得ることができた。しかし、使用状況の開示前後のデータを比較したところ使用時間や使用回数、個室の使用率に大きな差が生じていないことが分かった。



図 1 サイネージに表示したトイレの使用状況

実際に利用者の声をヒアリングすると、自分のいるフロアが満室で上下フロアのトイレに空室がある場合でも、フロアの移動をしている間に上下フロアも満室になる可能性があるためフロアの移動はしない、行列ができるわけではないのでフロアの移動はせず空室ができるのを待っているが、思いのほか待ち時間が長く別のフロアに行けばよかったと思うことがある、との声があがった。

以上のことから使用状況の開示は利便性の向上には一定の効果があるが、混雑時の人流の誘導については大きな効果が得られないことが分かった。これは、各フロアのトイレの使用状況に明確な差がある場合、使用状況を開示するだけで人流を誘導することができるが、どのフロアも比較的混雑している場合はフロア同士を比較し検討する要素がないため自分が向かうトイレを決定づけることができなからだと考えられる。そこで混雑を緩和するためには使用中の個室が空室になる確率を予測し開示すること、更には自分のいるフロアに留まるべきか別のフロアに移動するべきかを助言することができれば混雑の緩和に有効であると考えられる。

### 3. 関連研究

トイレの IoT 化の取り組みは過去に実施されており、扉の開閉情報を用いた使用状況の調査に留まらず、大便器、小便器、水栓、ハンドドライヤー等に搭載されたセンサーの情報を収集・分析する実証実験[1]も行われた。開閉センサーを用いた使用状況の可視化は既にサービス化されており、一部の商業施設等で導入され始めている。例えば KDDI の提供するサービス[2]では、クラウドに開閉データを蓄積することで利用者のパソコンやスマートフォンで空室状況を閲覧することができ、長時間の利用者がいた場合はアラームを上げることもできる。VACAN が提供するサービス[3]は大丸東京店等の大型商業施設に導入されただけでなく、可視化の対象をレストランや喫茶店等トイレ以外の施設にも広げている。これらのサービスでは情報収集のため

のセンサーは無線通信が可能な開閉センサーを利用する機会が多く、通信方式としては省電力な EnOcean や BLE 等が用いられる。これらは大掛かりな設置工事が不要で既設の設備への導入も容易である。

しかしこれらの取り組みは使用状況を可視化することで利便性を向上させ、トイレ設備の使用回数をモニタリングすることで設備の予防保全に活かすことはできるが、混雑したトイレの使用率を平準化させ設備を有効利用することに繋がるわけではない。

## 4. 提案手法

本研究ではトイレの混雑を緩和し効率的に設備を使用するため、トイレに向かう利用者を混雑しているトイレから別のトイレに誘導する方法を検討する。2 章で述べた通り、各フロアのトイレの使用状況に明確な差がある場合は単純な使用状況の開示で利用者を誘導することができるが、どのフロアも比較的混雑している場合はフロア同士を比較し検討する要素がないため利用者を誘導することはできない。

そこで我々は使用中の個室が空室になる確率を予測し、上フロアへの移動、下フロアへの移動、もしくは滞在の 3 択を利用者に開示することを目指す。これにより利用者を誘導し、トイレの混雑緩和・平準化を目指す。空室になる確率を予測するため次の 2 手法を提案する。

### 4.1 統計による予測

使用中の個室の中で最も空室になる可能性が高い個室を統計データを用いて予測し、最も空室になる可能性が高い個室のフロアへの移動を利用者に提案する。予測用のデータは過去の滞在時間の実績データを使用する。度数（人数）を  $y$ 、範囲（60 秒間隔の時間）を  $x$  としたヒストグラムを作成し、そこから滞在時間ごとの退出率を算出する。滞在時間が  $t$  秒以上の利用者数を  $U_t$  とし、退出率  $L_t$  を以下の式で求める。

$$L_t = \frac{U_t - U_{t+60}}{U_t}$$

滞在時間のヒストグラムと退出率を算出した結果、図 2 のようになった。この退出率を使って空室になる確率を算出し、最も確率が高い個室のフロアに利用者を誘導する。また、滞在時間のヒストグラムは状況に応じて変化することが想定できる。例えば自社ビルでの可視化実験で見られたように、10 時台と 12 時台では一人当たりの利用時間の長さが異なっている。そこで、同様の算出方法を用いて、以下の条件ごとに分けて退出率を導出し、条件による正答率の差異を検証する。

1. 全データ
2. 滞在するフロアで分類したデータ

3. 滞在する個室番号で分類したデータ
4. 入室した時間帯ごとに分類したデータ
5. 入室した曜日ごとに分類したデータ

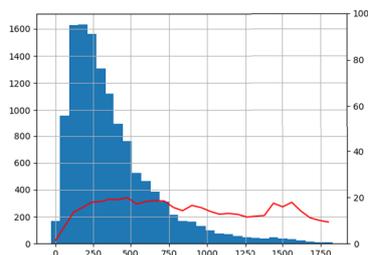


図 2 滞在時間のヒストグラムと退出率。棒グラフがヒストグラム、折れ線グラフが退出率を表す

## 4.2 RandomForest による予測

RandomForest[4]は分類や回帰に用いられる機械学習のアルゴリズムであり、複数の弱学習機器を用いて高度な予測を行うアンサンブル学習の一つである。ランダムに抽出された学習用データから図3のような決定木を複数作成し、それらの結果の多数決によって分類を行う。決定木を複数作成しその結果を組み合わせることで精度の高い分類や回帰を行うことができる。

この手法を用いて使用中の個室が1分以内に空室になるか否かを予測し、空室になる確率の高い個室のフロアに利用者を誘導する。

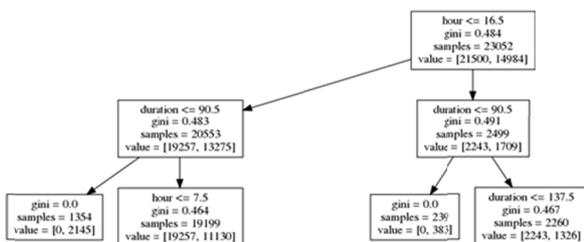


図 3 決定木のイメージ

### 4.2.1 入力データ

RandomForest の入力データセットとして  $x$  と  $y$  を作成する。 $x$  は過去の滞在時間からランダムな秒数を引いて作成した評価用の滞在時間とフロア情報や入室の時間情報、気象情報を付与したもので、表1の通りである。決定木を作成する際の特徴量として、滞在しているフロアや部屋番号、入室した時間帯等に加え、気象庁から収集した1時間ごとの気温や気圧等の気象情報を付け加えた。 $y$  はその個室が評価用の滞在時間から数えて1分以内に実際に空室となったか否かを”1”（空室となった）か”0”（空室とならなかった）で表した。

データ	内容
$x$ :	
Duration	滞在している秒数(評価用滞在時間)
Floor	滞在しているフロア
Room_no	滞在している個室番号
Day	入室した日
Hour	入室した時間帯
Min	入室した分
Weekday	入室した曜日
Press	入室した時間の気圧
Rain	入室した時間の降水量
Temp	入室した時間の気温
Humi	入室した時間の湿度
Win_s	入室した時間の風速
$y$ :	
Ans	1分以内に空室になったか否か

表 1 入力データ

出力はその個室が評価用の滞在時間から数えて1分以内に空室になるか否かを”1”か”0”で予測、分類したものとした。

上記の特徴量をもったデータセットを作成し、学習用データ  $train$  と評価用データ  $test$  に  $train:test=3:1$  にランダムに振り分ける。学習用データを使って RandomForest を構築し、評価用データを使って評価を行う。

### 4.2.2 正答率と精度

機械学習の結果を評価・分析する際、混同行列(Confusion Matrix)の考え方をを用いる。本研究を図4のように混同行列に当てはめて考えると、1分以内に空室になるか否かを正しく分類するのではなく空室になることを正確に予測することが重要となることがわかる。仮に1分以内に空室にならないと予測し多個室が空室になっても利用者に対する影響は少ないが、空室になると予測し多個室が空室にならなかった場合、空室になると予測したトイレに誘導された利用者はトイレを利用することができなくなってしまう。すなわち本研究では評価指数は正答率(Accuracy)ではなく精度(Precision)を重視するべきである。精度  $P$  は以下の式で求められる。

$$P = \frac{TP}{TP + FP}$$

		予測結果	
		「空室になる」と分類	「空室にならない」と分類
実際の結果	実際に「空室になる」	TP: True Positive 予測したい部分	FN: False Negative
	実際に「空室にならない」	FP: False Positive	TN: True Negative

図 4 混同行列に本研究を当てはめた図

### 4.2.3 パラメータ

RandomForest のパラメータのうち、今回調整するもの

は表 2 の通りである。

パラメータ名	内容
max_depth	決定木の深さ
max_features	決定木の特徴量の数,
min_samples_leaf	決定木を分割する際のサンプル数の最小値
n_estimators	決定木の本数

表 2 調整するパラメータ

それぞれの値をグリッドサーチを用いて決定することとする。その際の評価方法は 4.2.2 章で述べた通り精度を重視した評価とする。

#### 4.2.4 特徴量の選択

RandomForest はそれぞれの特徴量が結果に対してどの程度重要であるかを相対的に算出することができる。入力した特徴量について重要度を算出し、重要度の低い特徴を除外することによって結果に差が生じるか検証する。

## 5. 評価

自社ビル 9～11 階の男性トイレを使って評価を行った。2019 年 2 月 28 日～2019 年 4 月 7 日の間の月曜から金曜、8 時から 21 時までのデータを用いて評価を行った。30 分以上個室に滞在している人は極めて少なく、センサーの不具合またはデータの収集漏れも考えられるためデータから省くこととした。

### 5.1 統計による予測

4.1 章で提案した手法で条件ごとに作成したヒストグラムを図 5～9 に示す。条件によって度数の最大値やヒストグラムの尖度等に大きな差が認められた。フロア別で比較した場合、10 階は短時間の利用者の数が多く 9 階は 750 秒以上の利用者が多い。部屋番号で比較した場合、1 番の個室は滞在時間が短い利用者が最も多く、4 番は 1000 秒以上滞在する利用者数が多い傾向にある。

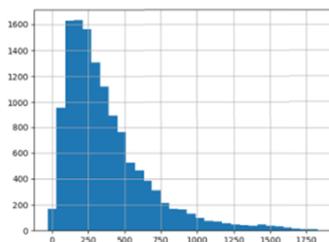


図 5 全てのデータより算出したヒストグラム

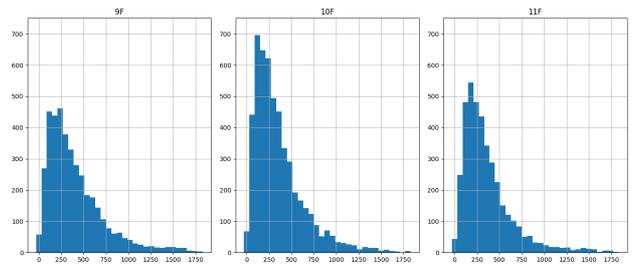


図 6 フロア別に算出したヒストグラム

左から順に 9、10、11 階

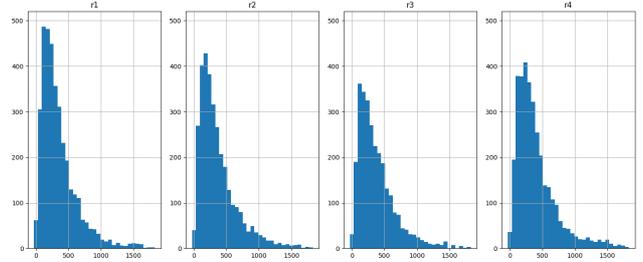


図 7 部屋番号ごとに算出したヒストグラム

左から順に部屋番号 1、2、3、4 番

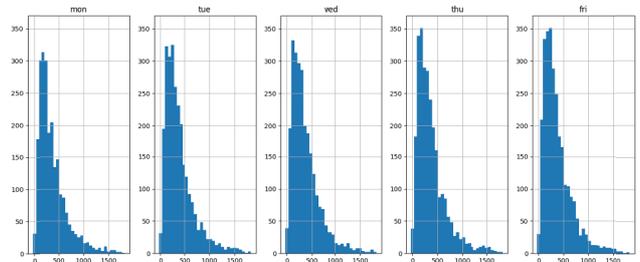


図 8 曜日ごとに算出したヒストグラム

左から順に月曜、火曜、水曜、木曜、金曜

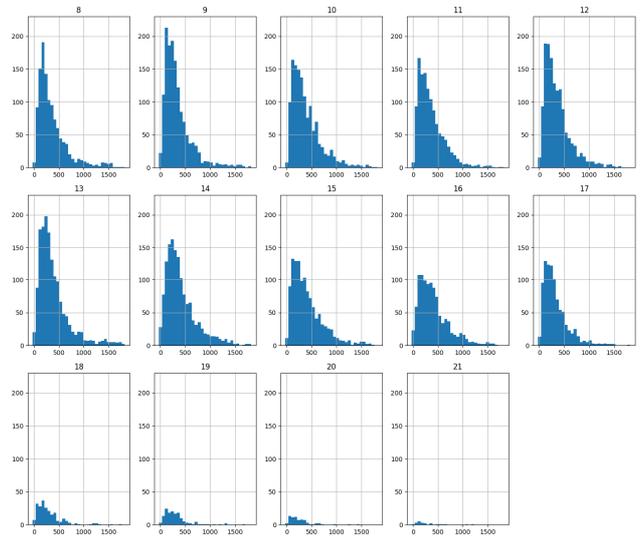


図 9 時間帯別に算出したヒストグラム

左上から順に 8 時から 21 時まで

次にこれらのヒストグラムから算出した退室率のグラフを図 10～13 に示す。ヒストグラムでは条件ごとに特徴があるように見受けられたが、退室率についてはグラフの形状に特筆するほど優位な差は見られなかった。0 秒から

300秒まで上昇を続け、300秒から500秒の間20%程度で一定となる。その後1200秒までは緩やかに減少し続け、1200秒からわずかに上昇する。ここまでの形状はどの条件でもほぼ同様である。ただしデータ数が少ない時間帯または条件については値が発散してしまっている。

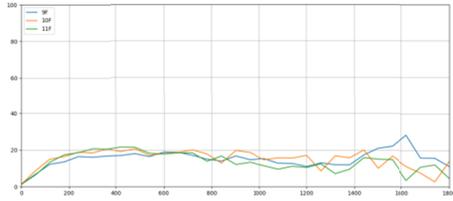


図 10 フロア別に算出した退出率

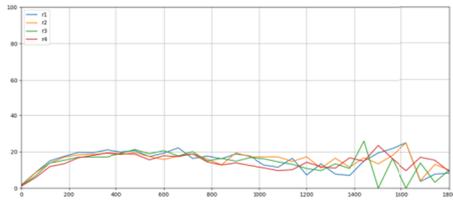


図 11 部屋番号別に算出した退出率

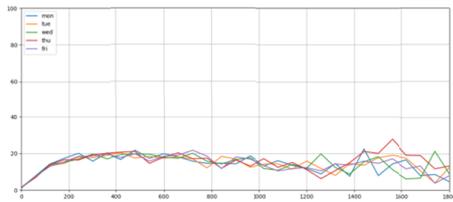


図 12 曜日別に算出した退出率

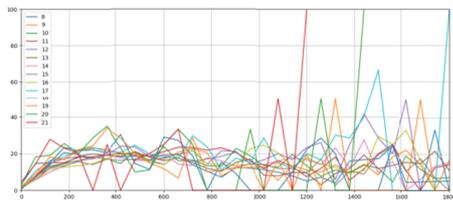


図 13 時間帯別に算出した退出率

この退出率を用いて最も空室になる確率の高い個室を予測する。評価用のデータとして実際にフロアのトイレが満室になった時の滞在時間を使用し、1番早く空室となった個室番号を正解データとする。評価用データの退出率を計算しその割合が最も高い個室が1番に空室になる個室とする。

算出結果とその正答率を表3に示す。表を見てわかる通り、多くの条件で30%前後の正答率となった。条件ごとに分類して予測を行っても明確な差はなく、正答率を上昇させることができないことが分かった。

条件	評価件数	正解数	正答率
全データ	4094	1155	0.2821202
フロア別			
9F	1325	413	0.3116981
10F	1741	515	0.295807
11F	1028	321	0.3122568
All	4094	1249	0.3050806
部屋番号別			
r1	1916	561	0.2927975
r2	832	266	0.1388309
r3	767	221	0.1153445
r4	579	194	0.1012526
All	4094	1242	0.3033708
時間別			
8	453	137	0.3024283
9	519	150	0.2890173
10	447	142	0.3176734
11	260	80	0.3076923
12	513	153	0.2982456
13	725	219	0.302069
14	529	153	0.289225
15	327	92	0.2813456
16	205	65	0.3170732
17	105	29	0.2761905
18	2	0	0
19	1	0	0
20	0	0	0
21	0	0	0
All	4086	1220	0.2985805

表 3 統計による予測結果とその正答率

## 5.2 RandomForestによる予測

4.2章で提案したRandomForestによる予測手法で全ての特徴量を使用して予測を行った結果、正答率と精度、件数は表4の通りだった。空室になると予測した件数が8,367件でそのうち本当に空室になったのが2,674件、空室ならないと予測した件数が11,133件で本当に空室ならなかったのが7,094件であった。精度を重視して評価を行ったため、精度が正答率を上回る結果となった。

Accuracy	0.6557435897435897
Precision	0.680
Matrix:	[[7094 2674] [4039 5693]]

表 4 RandomForestで1分以内に空室になるか否かを予測した結果

4.2.3章の通りグリッドサーチで最適なパラメータを探した結果、表5のパラメータが最適であることが分かった。

パラメータ	値
max_depth	5
max_features	auto
min_samples_leaf	1
n_estimators	100

表 5 グリッドサーチで探索したパラメータ

決定木の本数は多ければ多い方がいいが、決定木の階層は深いほどいいわけではないことが分かった。階層を適切な値より深くすると、過学習の傾向があった。

### 5.2.1 特徴量の選択

特徴量の重要度を算出した結果表 6 の通りだった。

順位	特徴量	重要度
1	Duration	0.925783
2	Min	0.013159
3	Temp	0.012105
4	Press	0.009868
5	Hour	0.008834
6	Win_s	0.006598
7	Humi	0.005364
8	Day	0.004957
9	Room_no	0.004313
10	Floor	0.003943
11	Weekday	0.002736
12	Rain	0.00234

表 6 特徴量の重要度と順位

滞在時間が最も重要な特徴量であり、その他の特徴量の重要度は限りなく小さいことがわかった。

重要度の低い特徴量を順に除外した結果、正答率と成果は表 7 の通りとなった。正答率及び精度はほとんど変化せず、空室になると予測した総数(TP+FP)も変わらなかった。

使用した特徴量	Accuracy	Precision	Matrix:
全て	0.655744	0.68	[[7094 2674] [4039 5693]]
11 位以上	0.656051	0.68	[[7084 2684] [4023 5709]]
10 位以上	0.655795	0.68	[[7086 2682] [4030 5702]]
9 位以上	0.655641	0.68	[[7091 2677] [4038 5694]]
8 位以上	0.655795	0.68	[[7080 2688] [4024 5708]]
7 位以上	0.655897	0.68	[[7082 2686] [4024 5708]]
6 位以上	0.655897	0.68	[[7086 2682] [4028 5704]]
5 位以上	0.656	0.68	[[7086 2682] [4026 5706]]
4 位以上	0.655846	0.68	[[7086 2682] [4029 5703]]
3 位以上	0.655692	0.68	[[7091 2677] [4037 5695]]
2 位以上	0.655538	0.68	[[7088 2680] [4037 5695]]
1 位のみ	0.65641	0.678	[[7019 2749] [3951 5781]]

表 7 使用した特徴量と予測結果

## 6. 考察

統計と RandomForest を用いて空室予測を行った結果、統計を用いた予測では平均して 30% 程度の正答率となった。4 つの個室の中からランダムで個室を選出しても 25% の正答率となるため 5% しか優位な差がみられなかった。以上のことからこの統計による手法では予測の精度が不十分であることが分かった。

RandomForest を用いた予測方法では正答率は 65%、精度は 68% で、支配的な特徴量は滞在時間であった。しかし分類の選択方法を調節することで現在の算出方法より精度を高められることが分かった。現在の手法では多数決で分類結果を導き出しているが、閾値を 5 割(半数以上の一致)から 6 割に変更することで精度を 78.4% まで上昇させることができた。

この精度で実際に利用者に情報提供すると考えると、仮に 9 階の個室が 2 個室空くと予測した場合は 2 個室とも空く確立が 61%、どちらか片方の個室が空く確立が 34%、どちらも空かない確率が 5% となる。利用者は 1 個室でも空けば良いため、9 階に移動すれば 95% の確率で 1 分以内にトイレを使用することができる。3 個室空くと予想した場合はさらに確率が上がる。これだけの精度であれば利用者をほとんど間違いなく最適なトイレに誘導することができ、トイレの混雑緩和および有効利用に貢献できると考えられる。

本研究では予測方法の提案および検証を実施したが、今後は実際に混雑緩和に効果があることを検証する必要がある。また比較的混雑するが待ち行列は少ない自社ビルにて検証を行ったが、ショッピングセンター等の行列ができる環境でも利用可能にするためには行列を含めた空室の予測も重要となる。プライバシーの問題からカメラの設置は好ましくないため、センサー等の手法を用いて行列の数や動きを把握し、待ち時間等を予測する方法を検討する必要がある。

## 7. まとめ

トイレの使用状況を可視化するだけではトイレの混雑を緩和できないことが判明したため、本研究では混雑緩和を目指して空室の予測に取り組んだ。

統計的な予測と RandomForest による予測を実施した結果、統計による予測では正答率は 30%程度で実用に耐えないことが分かった。一方で RandomForest による予測の結果精度は 68%であったが、回答の算出方法を半数以上の一致から 6割以上の一致に変更することで 78%まで上昇させられることが分かった。

今後はこの予測方法を使った混雑緩和の実証実験や行列も含めた混雑状況の予測に取り組む必要がある。

## 参考文献

- [1] 電気設備学会:電気設備学会誌 BACS/BEMS における IoT の活用 J. IEIE Jpn. Vol. 36 No. 7
- [2] KDDI IoT クラウド～トイレ空室管理～,  
<https://iot.kddi.com/services/iot-cloud-toiletdoor/>
- [3] Throne- スローン ,<https://www.throneservice.com/>
- [4] Breiman, Leo (2001). "Random Forests". Machine Learning 45 (1): 5–32. doi:10.1023/A:1010933404324.