小山工業高等専門学校　電気電子創造工学科

平成30年度　卒業論文

**広域センサネットワークを用いた**

**トラッキングサービスにおける**

**サービスに不要なプライバシデータ流出比の定量評価**

Quantitative Evaluation of Ability for Privacy Protection Using Simulator of

Local Production for Local Consumption Type Networks

in Wide Area Sensor Tracking Service

電気電子創造工学科

田村　崚

Department of Innovative Electrical Electronic Engineering

Ryo TAMURA

指導教員　干川 尚人

Naoto HOSHIKAWA

|  |  |
| --- | --- |
| 学科長 | 指導教員 |
|  |  |

1. **背景**

近年のInternet of Things (IoT)技術の発展とオープン化により，町中にある様々な機器がサービスへ利用可能になると予想されている．さらに高精度な画像処理技術が合わさることで，センサネットワークを利用した情報処理サービスはより広域かつ高精度になると期待される．一方でこのようなサービスでは無差別なセンシングが頻繁に行われ，サービスには不要なプライバシデータ(Non-related Privacy Data,以下NPDと記載)も収集してしまう課題がある．顔などの生体関連のプライバシデータ提供を不安に感じる利用者は特に多い[1]ため，もし有用なサービスが提案されても社会実装に理解が得られない可能性がある．この課題に対し，本研究チームは広域センサによる見守りサービスにおけるNPD流出を抑制できる地産地消型ネットワークアーキテクチャ(Local Production for Local Consumption Network Architecture，以下LPLCモデルと記載)を提案している[2]．本稿では広域見守りサービスにおけるNPD流出比を算出する数理モデルを示し，それを再現するシミュレータを示す．またそのシミュレータを用いて，LPLCモデルのプライバシ保護能力を定量評価する．

1. **数理モデル提案**
   1. **LPLCのモデル化**

LPLCの動作モデル図を図1に示す．

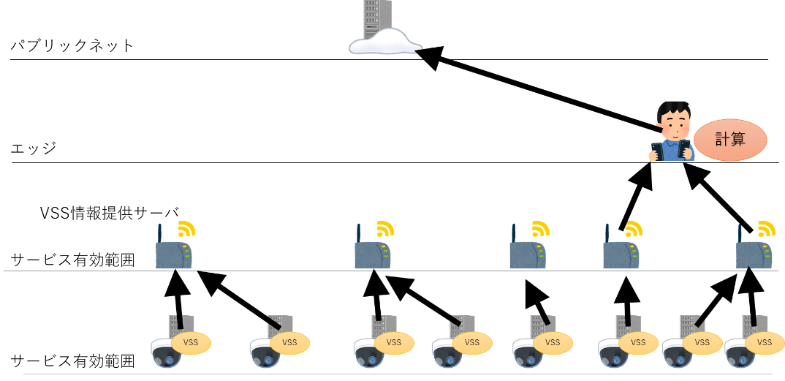
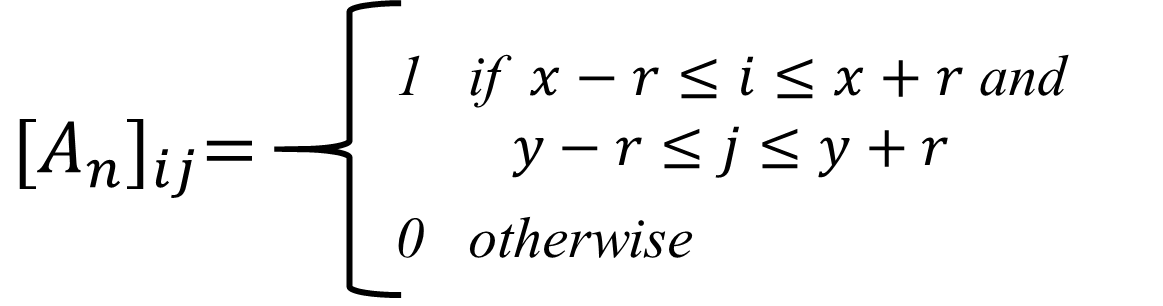


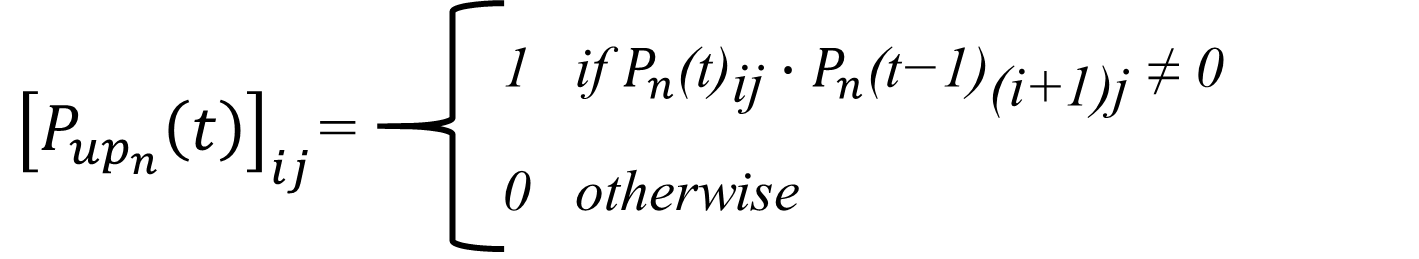
図1.LPLCの動作モデル図

LPLCモデルでは，追跡対象者の持つ計算ノード(以下LPLCノードと記載)がルータアクセス可能範囲内にある時，ルータと同セグメントにあるセンサデータを処理する．そのため全てのデータを収集するクラウドモデルと比べて，NPD流出を制御できる．なお，ここでは計算ノードにデータが流入することを流出と定義する．サービス領域を内包するように長方形(以下マップと記載)を置き，それを外包するように正方形のセルを敷き詰める．それぞれのセルが持つ対象者の有無，センサ有効範囲内といった情報を各行列の要素に対応させモデル化する．まず*n*番目のカメラ集合を定義する．この集合は，各セルがカメラのセンサ有効範囲内か否かを格納した行列と，マップ上でのカメラの撮影方向*up,down,right,left*の逆を要素に持つ．次に*n*番目のルータの集合 を定義する．これは属するカメラ集合の集合とアクセス可能範囲行列を要素に持つ．ここでルータ位置とアクセス可能距離をそれぞれとすると，は次のように定義できる．

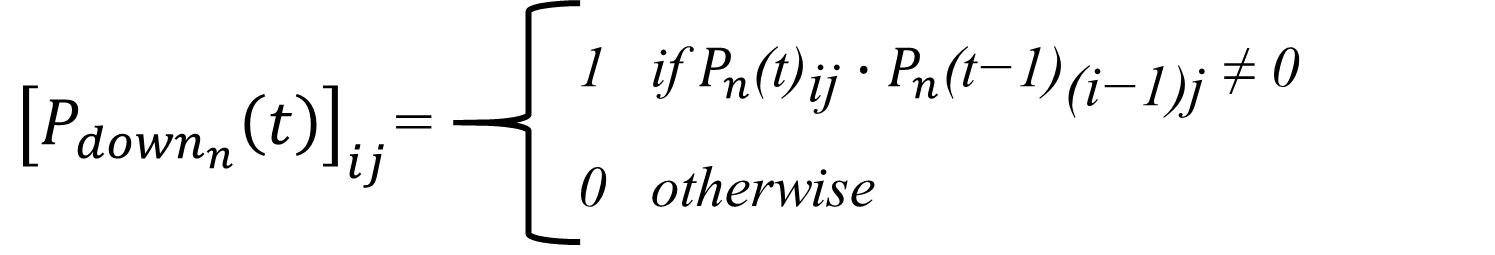


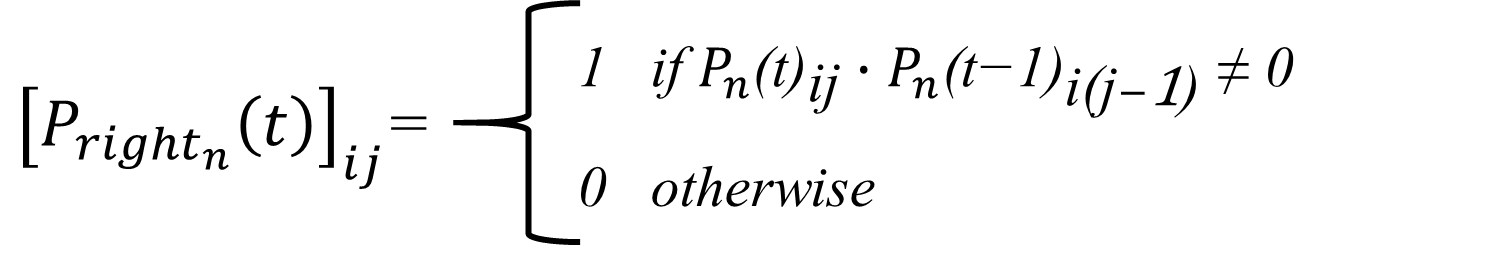
(1)

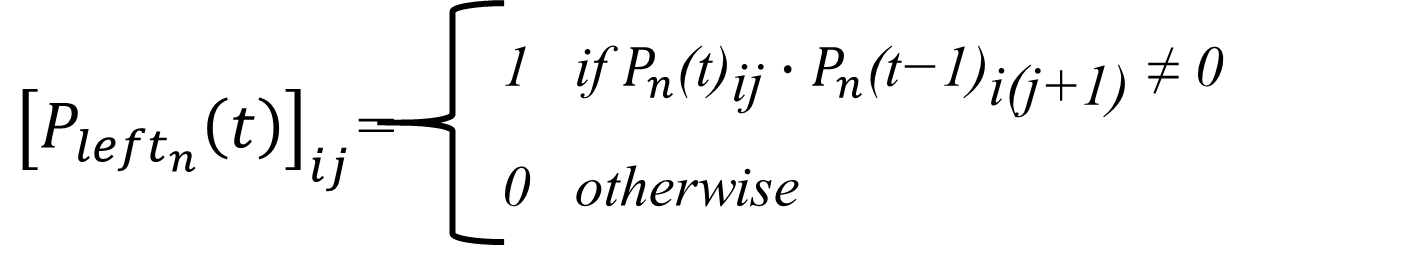
最後に*n*人目の対象者の集合を定義する．この集合は各セルでの対象者の有無を格納した行列と対象者の向いている方向を表現する行列*,,,*を要素に持つ．例として，が１であるとは，*ｎ*人目の対象者が時刻*t*の時にセル*ij*でマップ上方を向いていたことを示している．それぞれの行列の定義を下に示す．また，の総和をと定義する．方向を示す行列も同様である．非対象者集合も同様に定義する．



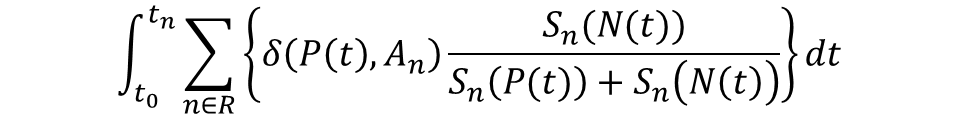
(2)



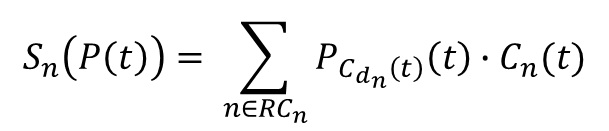


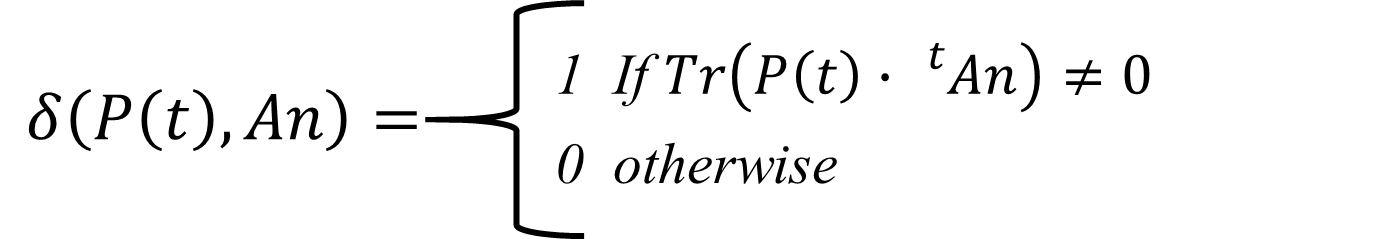


サービス開始時間を，終了時間をとすると，NPD流出比は次式のようになる．



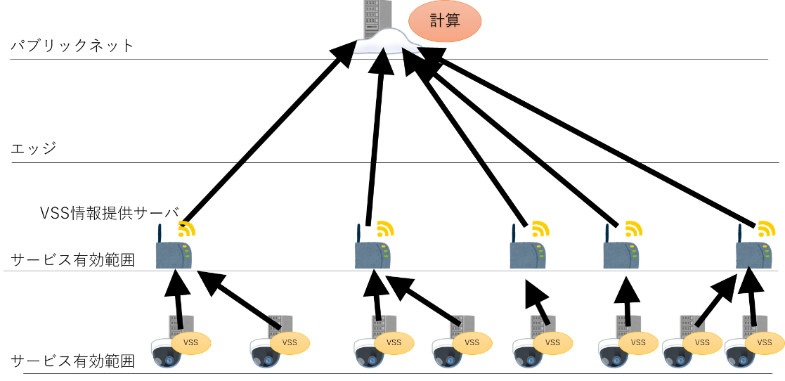
なお，は*n*番目のルータが時刻*t*に対象者全体のセンシング量を返す関数，は*n*番目のルータのアクセス可能範囲内に時刻*t*の対象者が存在するか否かを返す関数である．





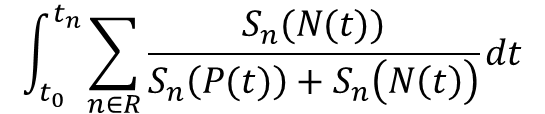
* 1. **クラウドサービスのモデル化**

クラウドモデルの動作モデル図を図に示す．



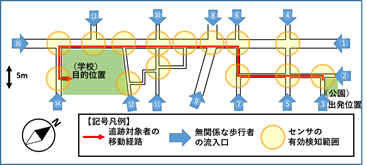
図クラウドサービスの動作モデル図

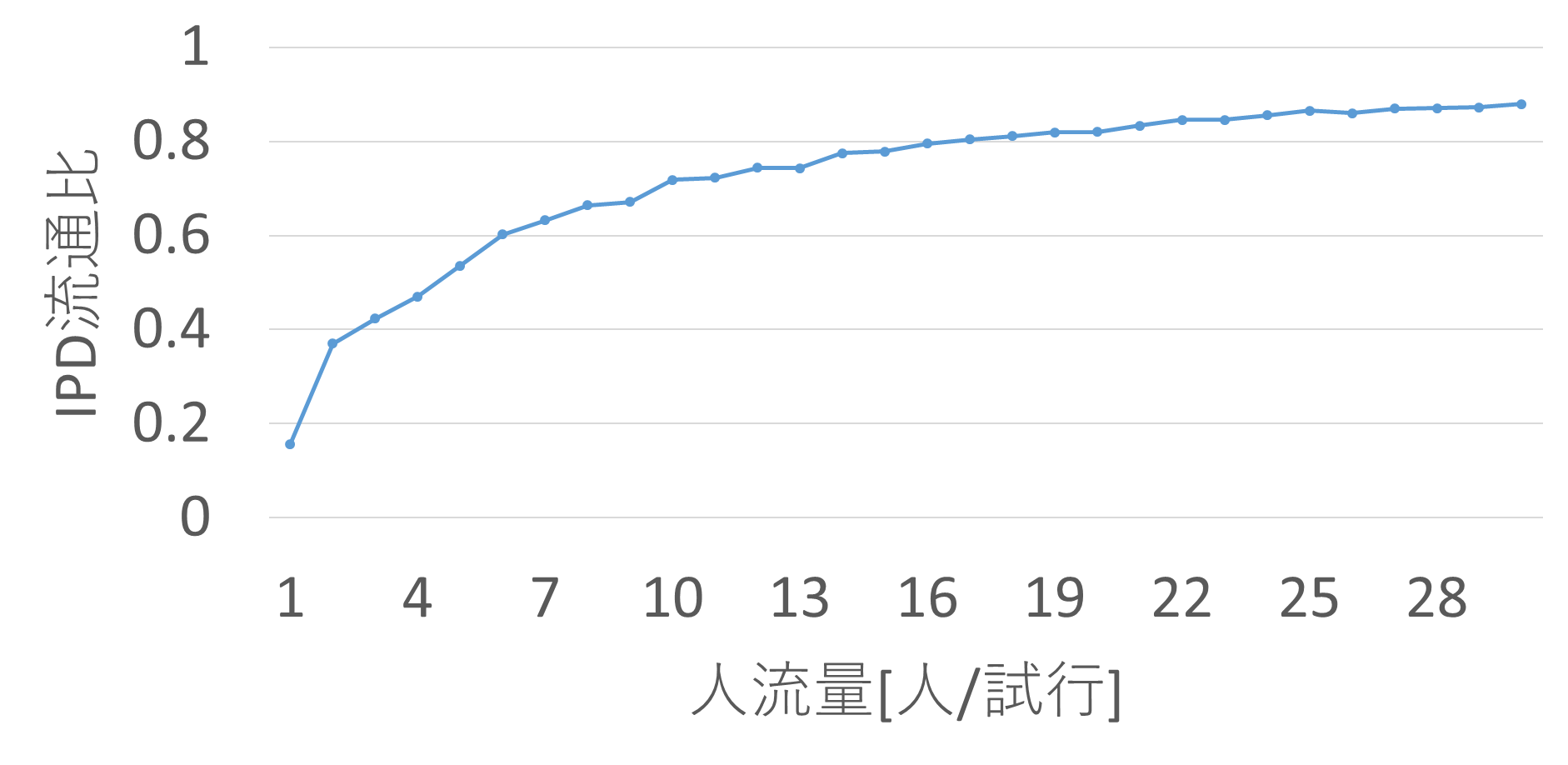
既存サービスモデルは主にクラウドサービスを使って実現されている[3][4]．クラウドサービスではあらゆるデータをクラウド上にあるサーバに送信し，画像処理はそのサーバ上で行っている．よってLPLCと違い収集するデータはセンシングされたデータ全てとなる．クラウドモデルのNPD比は次のような式で表される．



1. **シミュレーション実験**
   1. **提案手法**
      1. **プロトタイプシミュレータについて**

以前簡易的なクラウドサービスのシミュレータ[5干川先生のやつ?]を作成し、クラウドモデルのNPD比について検討を行った．このシミュレーションでは実在する小学校の周囲をモデルにサービス領域を作成(図)し，サービス内容は「公園から学校までの見守り」がクラウドサービスで提供されるとした．見守り対象者(以下対象者と記載)を児童とし，そうでない人物を見守り非対象者(以下非対象者と記載)とした．カメラは全ての交差点に設置し，このシミュレーションではカメラの方向を考えない~~(数式でフォローを入れる？)~~．人の歩行速度を1.25[m/sec]，カメラの動作周期をt=1[s]，セル１辺を1.25[m]とした．また，追跡対象者は赤線に示す固定のルートを，非対象者は青矢印に示す15カ所の流入口から1つ選びそこからのルートを立ち止まることなく進む．非対象者が増減した際のNPD流出比を図に示す．このシミュレータから，人が快適に移動できると言われる00％未満であってもNPD流出比は00％を超えていることがわかった．





* + 1. **オープンデータについて**

プロトタイプでは見守り対象者の動きや人数について自由に設定したため，現実的なNPD流出比であるかは疑問が残った．そこで本稿ではG空間情報センターの公表している松江駅構内人流センサデータ（西日本旅客鉄道株式会社）[6]（以下オープンデータ）を用いることで，より現実に即したNPD流出比を得られるようにした．オープンデータはセンサ19個ごとに「センサID, 日付, 時刻, In, Out, In累計, Out累計」が定義されている．また，センシング結果は1分毎の合計が記録されている．センサ位置，in/outデータは公開されている次の画像(図)に従う．ここでは2018年6月，平日の5時から24時のデータを扱う．なお，人流センサ15番は動作を確認できなかったため扱っていない．



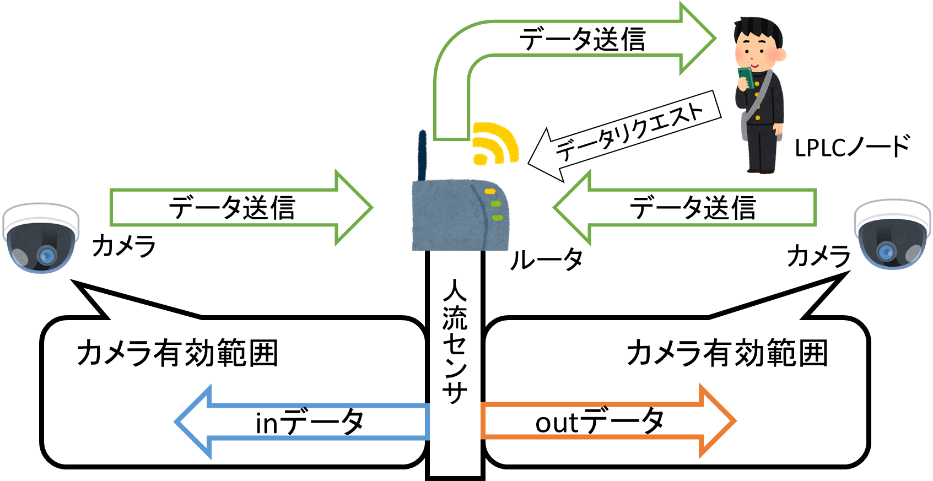
センサ設置位置図



センサin/out定義図

* + 1. **シミュレーション方式**

オープンデータはあくまで人流センサデータであるため，このままではプライバシデータとして扱えない．そこでin/out人流のそれぞれを撮影する仮想のカメラを設定し，人流センサデータをカメラデータへ変換する機構を作成した(図)．カメラはRD-CA213 AHDカメラ220万画素バレット型赤外線搭載屋外カメラ[7]を想定し，同カメラを販売している会社の公表しているデータ[8]からセンシング可能距離を15mとした．またLPLCモデルにはカメラの近くにルータを設置する必要があるため，オープンデータにおける人流センサと同じ位置にルータを設定した．



カメラ，ルータ位置図

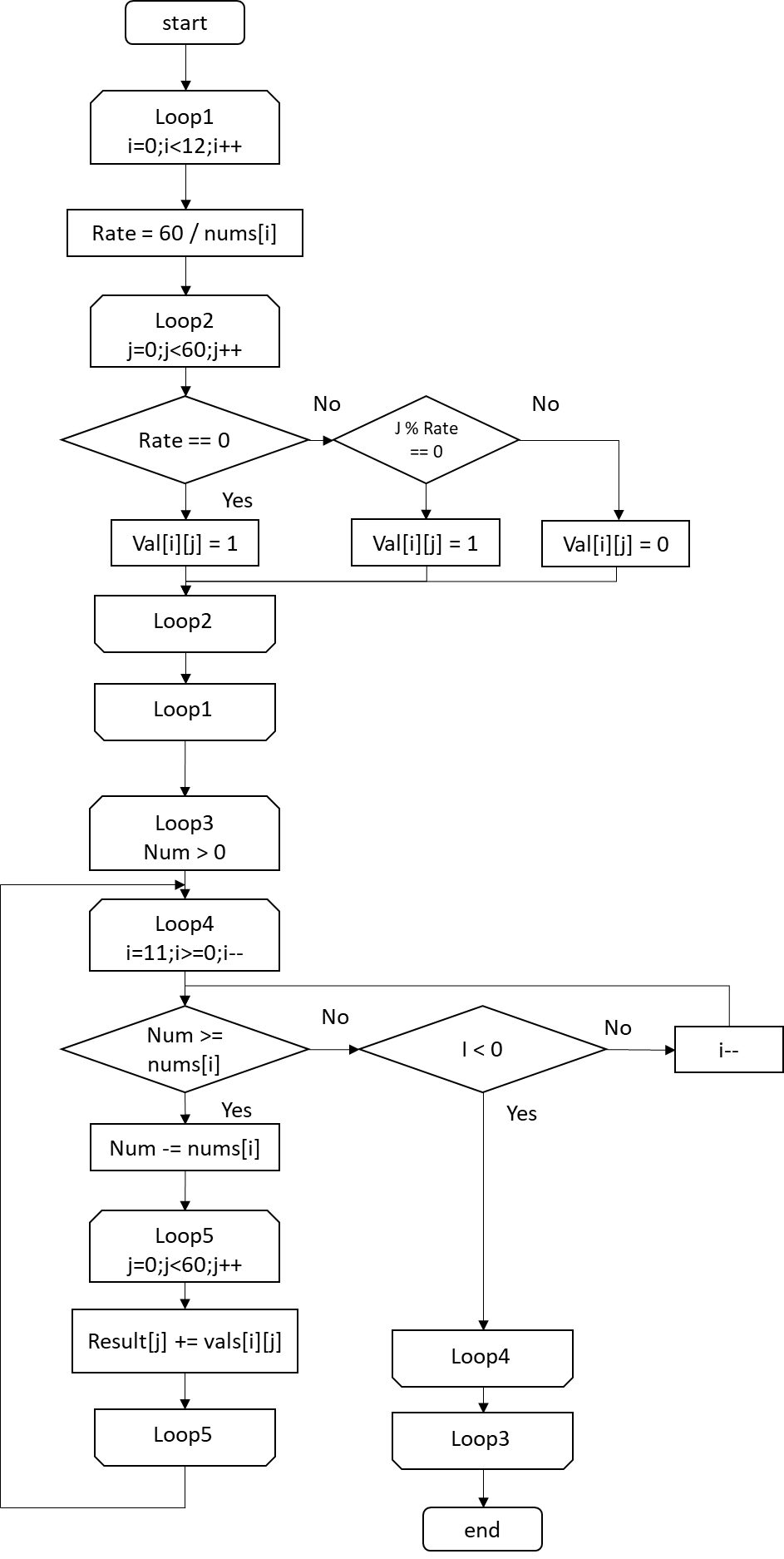
本稿ではカメラのセンシング周期を1秒とする．オープンデータは1分毎に更新されるので，次に示すアルゴリズムを用いてオープンデータを1秒毎のデータへ変換した．

表１．アルゴリズムに用いる変数の定義

|  |  |
| --- | --- |
| 60の約数 | *nums*[12]={1,2,3,4,5,6,10,12,15,20,30,60} |
| 約数毎の60分割 | *vals*[12][60] |
| 入力データ | *input* |
| 出力データ | *result*[60] |

*vals[i]*には*nums[i]*を60分割したデータが入っている．例えば 20の*vals*には{1,0,0,1,0,0,1,…,1,0,0}と1,0,0が20回繰り返されるデータを格納している．このように*vals*には60個の要素の合計が*nums*となりつつ，数字が偏らないように格納されている．*input*が0になるまで次の処理を繰り返すことで*result*に*input*を60分割した数が格納される．

1. *input*より小さい中で，最も大きい*nums*を選ぶ．選ばれた*nums*を仮に*nums[i]*とする．
2. *input*から*nums[i]*を引き，*result*に*vals[i]*を足し合わせる．



図

* 1. **シミュレータの設計**
     1. **サービスモデル**

ここでは広域見守りサービスの例として，松江駅登校者見守りサービスを設定する．サービス領域はオープンデータで示された領域（図）の内，壁に囲まれていない部分とし，サービス内容を「駅を利用する高校生の登校見守り」とした．サービス時間は曜日や学校に依存しにくい登校を想定し，7時から8時とした．対象者は松江駅を利用して登校する高校生とし，そうでない人物を非対象者とした．オープンデータ全てを非対象者のデータとし，対象者のデータについては次の項で述べる．カメラの位置，方向は図で示したとおりである．人の歩行速度を1.25[m/sec]，カメラの動作周期をt=1[s]，セル１辺を1.25[m]とした．モデル化されたJR松江駅を図に示す．サービス領域外を灰色で，センシング可能範囲を緑色で可視化したものを図に示す．



（a）シャミネ



（b）



（c）

図．JR松江駅モデル化図



（a）



（b）



（c）

図．監視範囲図

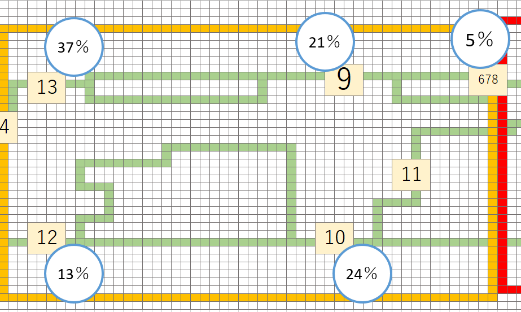
* + 1. **対象者設定**

サービス時間の人流センサデータ数の一日当たりの平均を表に示す．「センサ通過のべ人数」とはin/outデータ両方を加算した値である．

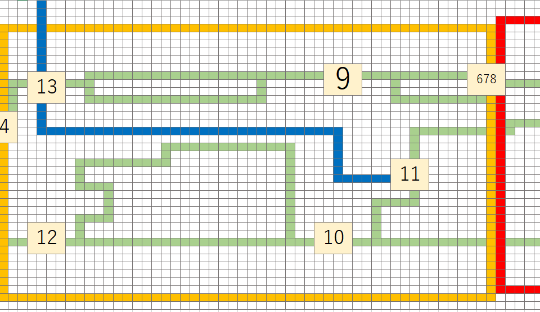
表：登校時人流センサ毎の総データ量

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| センサ番号 | 1 | 2 | 3 | 4 | 5 |
| センサ通過のべ人数 | 0 | 0 | 2 | 0 | 2 |
| センサ番号 | 6 | 7 | 8 | 9 | 10 |
| センサ通過のべ人数 | 94 | 14 | 68 | 452 | 514 |
| センサ番号 | 11 | 12 | 13 | 14 | 15 |
| センサ通過のべ人数 | 1181 | 277 | 811 | 0 | × |
| センサ番号 | 16 | 17 | 18 | 19 | **総和** |
| センサ通過のべ人数 | 0 | 0 | 0 | 0 | **3415** |

ある人物が駅を通過する際には少なくとも駅への流入時と流出時の2つのデータを生成すると考えられる．この仮定と表から，一日当たりの7時から8時の駅利用者をおおよそ1700人分とする．鉄道利用者に占める学生の割合は全体の7％程度[9]であるから，対象者の最大人数を全体の10％程度とし，200人と設定した．対象者が駅を通過する際に使う流入口は人流センサ11番(改札)で固定とし，流出口は物理的に駅の外側に接している人流センサから表に従い確率で選択した．利用される流出口の利用率を図に示す，人流センサ11番から流入し，人流センサ13番から流出する際の対象者のルートを図に青色で示す．



流出口使用割合



対象者移動ルート例

* 1. **シミュレーション内容**

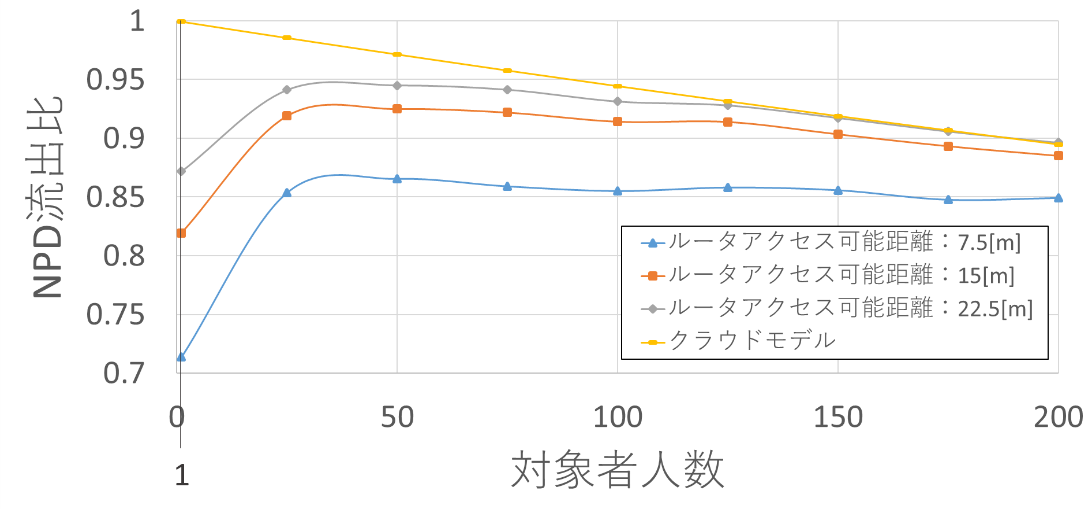
長崎

(福岡)ルータアクセス可能距離の値を設定し，対象者数を1人に加え，25人から200人まで25人刻みで変化させシミュレーションを行った．また，対象者を1人とした時にルータアクセス可能距離を変化させた場合のNPD流出比をシミュレーションした．

* 1. **シミュレーション結果**

長崎

ルータアクセス可能距離毎の対象者人数とNPD流出比の関係，およびクラウドモデルのNPD流出比を図にまとめた．



シミュレーション結果

表

NPD流出比が小さいほどプライバシ保護能力が高いといえるので，ルータアクセス可能距離に依らず，クラウドモデルよりもLPLCモデルの方が優れている．

1. **考察**

ルータアクセス可能距離がセンサ有効距離より短いと追跡サービスとして不十分であるが，反対にアクセス可能距離があまりに長いとNPDセンシング過多の原因となる．よって，センサ有効範囲を内包しつつルータアクセス可能距離を短くすることにより，プライバシに配慮した追跡サービスが実現できる．また図から，対象者数が少ないときのLPLCモデルのプライバシ保護能力が特に優れている．そのため対象者人数に対してLPLCノードの数を減らすことが出来れば，よりNPD流出比を抑制できる．

センシング可能範囲とNPD収集量の間にはトレードオフの関係がある．ただNPD比の削減のみを目指し続けると，そもそもセンシングをしなければよいという支離滅裂な回答が得られかねない．そこでサービス領域カバー率という指標を用意する．これはサービス領域の面積の内何％がセンシング可能であるかを示す指標である．これが100％に近いほど，見守りサービスとして十分であると言える．本シミュレータは図と図より，サービス領域カバー率は46.1％であった．

1. **まとめ**

本稿ではLPLCモデル及びクラウドモデルのNPD比を数理モデルで表現しし，シミュレータからLPLCモデルはクラウドモデルと比べプライバシ保護能力に優れていると定量評価した．今後はより多人数を見守る際のNPD比の削減方法及びLPLCモデルの実証実験等を行い，LPLCモデルについての研究を進めていく．

謝辞

・G空間情報センターの人

・渡邉先生

参考文献

1. 総務省. 特集 データ主導経済と社会変革. 平成29年版 情報通信白書 ICT白書2017, 第1部, 第2章, pp. 79.
2. 干川尚人，下馬場朋禄，伊藤智義，“地産地消型アーキテクチャによるセンサネットワークデータのプライバシ保護，”情報処理学会論文誌，vol.59，No.12，pp.2180-2190，Dec，2018.
3. “AWS IoT プラットフォーム”.   
   https://aws.amazon.com/jp/s/dm/optimization/sem/iot-signup/ (参照 2018-12-20)
4. “PaaSで誰もがデータ分析”. 日経コンピュータ. http://tech.nikkeibp.co.jp/it/atclact/active/15/010700161/010700003/ (参照 2018-12-20)
5. 干川先生の奴？
6. 社会基盤情報流通推進協議会，人流解析チーム，“松江駅構内人流センサデータ”
7. “株式会社アルコム RD-CA213 AHDカメラ220万画素バレット型赤外線搭載屋外カメラ”https://www.arucom.ne.jp/store/g/g250-011409/ (参照 2018-12-20)
8. “株式会社アルコム 人相確認距離から防犯カメラを選ぶ。”https://www.arucom.ne.jp/sample/ (参照 2018-12-20)
9. 関東交通広告協議会，“交通広告調査レポート 2009，”鉄道利用者プロフィール，pp39