一般社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
IEICE Technical Report
NS2021-138(2022-03)

複数エリアのクラウドセンシングにおける最適ポイゾニング攻撃

藤本　　凛†　　上山　憲昭††

† 福岡大学工学部電子情報工学科　〒 814-0180 福岡福岡市城南区七隈 8-19-1
†† 立命館大学情報理工学部　〒 525-8577 滋賀県草津市野路東 1-1-1
E-mail: †tl181284@fukuoka-u.ac.jp, ††kamiaki@fc.ritsumei.ac.jp

あらまし　近年，低コストで多くの情報を収集できる，モバイル機器を用いたクラウドセンシングが注目されている．しかし不特定多数のユーザからデータを収集する性質上，悪意のあるユーザが誤差の大きなデータを意図的に送信することで，計測値の推定誤差を増大させるデータポイゾニング攻撃の問題が指摘されている．既存の研究では，単一エリアを対象としたデータポイゾニング攻撃法や防御法が研究されているが，複数エリアを対象としたクラウドセンシングに対するデータポイゾニング攻撃は未検討である．そこで本研究では，複数エリアに存在するユーザから収集したデータに対して，エリアごとに測定値を推測するクラウドセンシングにおけるデータポイゾニング攻撃を想定する．そして全エリアの総誤差と，エリアごとの誤差の最小値の二つの観点から，攻撃者が攻撃効果を最大化するよう各エリアに配置する攻撃者数を設計する方式を提案する．そして計算機シミュレーションにより，提案方式を用いることで，多くのエリアで誤差を生じさせ，総誤差が増加することを示す．
キーワード　クラウドセンシング，データポイゾニング攻撃，複数エリア

Optimal Poisoning Attacks on Crowdsensing at Multiple Locations

Rin FUJIMOTO† and Noriaki KAMIYAMA††

† Faculty of Engineering, Fukuoka University
8-19-1, Nanakuma, Jonan-ku Fukuoka, 814-0180 Japan
†† College of Information Science and Engineering, Ritsumeikan University
1-1-1, Nojihigashi, Kusatsu, Shiga, 525-8577 Japan
E-mail: †tl181284@fukuoka-u.ac.jp, ††kamiaki@fc.ritsumei.ac.jp

Abstract　Recently, crowdsensing using mobile devices has gathered wide attention as a way to estimate various environmental data with small cost. However, due to the nature of collecting data from unspecified users, the problem of data poisoning attack in which the estimation error of data is enlarged by intentionally sending data with large errors from malicious users. Existing researches have studied data poisoning attacks and prevention methods considering a single area, and data poisoning attacks for multiple areas have not been studied. In this paper, we propose methods of optimally designing the number of attackers allocated at each area to maximize the effectiveness of the attack in crowdsensing in which the measurement value is estimated at each area independently. Using computer simulations, we show that the proposed method generates error at many areas and increases the total error among all the areas.
Key words　crowdsensing, data poisoning attack, multiple areas

1. はじめに

今や，モバイル端末は私たちの日常生活に欠かせないものである．それらは高度なセンシング能力を有し偏在していることから，モバイル機器は周囲の環境や社会環境に関する膨大な情報が収集できる．そこでモバイル端末をセンサーデバイスとして活用するモバイルクラウドセンシング (MCS: mobile crowdsensing) が，低コストで多くのデバイスから情報を収集し，気温などの様々な環境データの値を推測する方法として注目されている．しかし，データの提供元であるモバイル端末が多数，存在することによって，矛盾したデータが集まり競合が発生することが考えられる．このようなデータの矛盾は，データの記録や送信で発生するエラーや機器の誤作動等が原因で発生する．特にユーザが意図的に誤差の大きなデータを送信することで推定値を歪ませるデータポイゾニング攻撃の問題が指摘されている [3]．データポイゾニング攻撃に対して推定誤差を抑制するために，ユーザごとに報告値と推定値との差異に基づく信頼性を推定し，信頼性を重みとした報告値の加重和を推定値に用いる CRH（Conflict Resolution on Heterogeneous data）法が提案されている [1]．また CRH 法に対し，攻撃者が推定誤差を最大化するよう，各攻撃者の測定報告値を設定する方式である DPA（Data Poisoning Attack）法と，その防御法が考案

This article is a technical report without peer review, and its polished and/or extended version may be published elsewhere.
Copyright ©2022 by IEICE

されている [2]．これらの研究では，単一エリアを考慮したデータポイゾニング攻撃が想定されているが複数のエリアに対するデータポイゾニング攻撃は考慮されていない．そこで本研究では，複数エリアにおいてデータポイゾニング攻撃の効果を最大化するよう攻撃者を各エリアに配置する方式を考察する．

図 1 に示すように，データ収集領域は N 個のエリアから構成され，各エリアにおいてセンシングデータをモバイル端末のユーザから収集し，エリアごとに測定値を推定するサービスを想定する．例えば市街を 100 メートルの正方領域に分割し，各エリアの二酸化炭素濃度を推定するようなサービスが該当する．そして M 人の攻撃者を N 個のエリアに配置する問題を考える．文献 [2] で提案されている DPA 法をエリア s に適用したときの，正常ユーザの報告値のみで計算した推定値と，攻撃者を含む全ユーザの報告値で計算した推定値との差の絶対値を誤差 e_s と定義する．攻撃者は，正常ユーザの測定値やエリア毎の人数についての情報を知っているものとし，攻撃者を一人ずつ各エリアに配置していく近似解法を想定する．攻撃者の最適化目標として，全エリアの誤差の総和 E の最大化だけではなく，各エリアの誤差 e_s の最小値 e_{min} の最大化の 2 つを考える．

以下，2. 節で関連研究をまとめ，3. 節で既存の CPA 法と DPA 法の概要をまとめ，さらに 4. 節で DPA 法を単に複数エリアに拡張した場合の結果について考察する．そして 5. 節で本稿で新たに提案する複数エリアの MCS における最適攻撃者割当法を述べ，6. 節で数値評価結果を示し，最後に 7. 節で全体をまとめる．

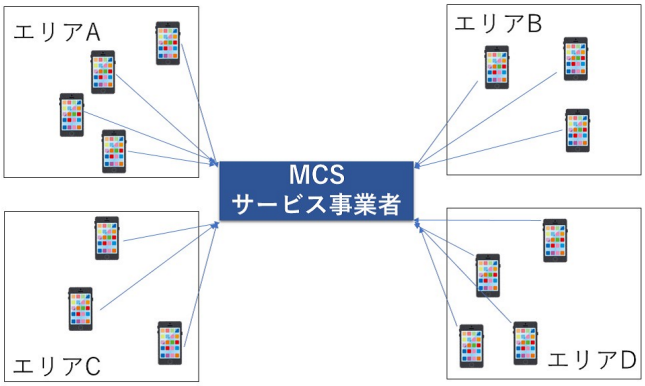

図 1　複数エリア MCS

2. 関 連 研 究

これまでに様々な MCS サービスに対する Poisoning 攻撃手法や，その防御法が提案されている．連続量を扱う MCS サービスを対象としたものとして，例えば Quintal らは，MCS で送信される正常ユーザの報告データを観測し，攻撃者のデータパタンと似たパタンの正常ユーザデータに，malicious データを加算することで，検知される確率を下げた Poisoning 攻撃を提案している [5]．また Huang らは，複数の各オブジェクトに対し真値を推測する場合に，誤差最大化を目的とした Poisoning 攻撃者に対し，推測値からの誤差が閾値を超えたユーザを除外する防御法を提案している [2]．

また複数のカテゴリから該当するものを選択する MCS を対象としたものとして，例えば Varshney は複数の報告値から多数決でカテゴリを選択する場合を対象とした Worker 間の結託に対する防御法を提案している [6]．また Miao らは複数の各オブジェクトに対し，カテゴリ選択型のデータを複数のユーザから集め，各オブジェクトの真のカテゴリを推測する場合に，ユーザの信頼度に応じた重みを考慮する Dawid-Skene モデルをクラウドが用いている場合に，攻撃者の最適攻撃法を提案している [3]．さらに Miao らは複数の各オブジェクトに対し，カテゴリ選択型のデータを複数のユーザから集め，各オブジェクトの真のカテゴリを推測する場合に，(1) 推測結果の誤差を最大化，(2) 設定したターゲットに推測結果を誘導，の二つのタイプの Poisoning 攻撃の最適攻撃法を提案している [4]．

しかしこれらの研究はすべて，単一のエリアにおける MCS を想定しており，複数のエリアにおける MCS を対象とした研究は見られない．本稿では新たに，複数のエリアを対象とした MCS における Poisoning 攻撃を想定する．

3. 既存方式: CRH 法・DPA 法

3.1 CRH 法

CRH 法は，複数の情報源から真実を推測することを目的としたアルゴリズムである [1]．推定値と入力データとの間の全体的な加重偏差を最小化する．その際にユーザごとの信頼性を推定し，信頼性を重みとした測定値の加重和を推定値に用いる．MCA に適用した場合，測定値の推定値と報告値との差異が小さいユーザの信頼性は高く，大きいユーザの信頼性は低くなるように各ユーザの信頼性を設定する．推定値とユーザごとの信頼性を交互に更新することで，推定値と真値の差を最小化する．以下に CRH 法のアルゴリズムを示す．ただし正常ユーザの集合を $\boldsymbol{N}$ とし，攻撃者の集合を $\boldsymbol{A}$ とする．

（i） 各ユーザ k の信頼性 (重み)w_k を 1 に初期化

（ii） 式 (1) で，各ユーザ k の報告値 v_k と w_k から，推定値 $\tilde{v}$ を計算

$$\tilde{v} = \frac{\sum_{k \in \boldsymbol{N} \vee \boldsymbol{A}} v_k w_k}{\sum_{k \in \boldsymbol{N} \vee \boldsymbol{A}} w_k} \tag{1}$$

（iii） 式 (2) でユーザごとの信頼性 w_k を更新

$$w_k = -\log \frac{(v_k - \tilde{v})^2}{\sum_{k \in \boldsymbol{N} \vee \boldsymbol{A}} (v_k - \tilde{v})^2} \tag{2}$$

（iv） $\tilde{v}$ 及び w_k が収束するまで (ii)(iii) を反復

3.2 DPA 法

DPA 法は，CRH 法による推定値及びユーザごとの信頼性の更新ステップ，攻撃者の報告値の更新ステップの 2 ステップを繰り返すことで，推定誤差を最大化する攻撃者の最適報告値設定アルゴリズムである [2]．攻撃者の最適化目標は，正常ユーザのみで推定した値 $\tilde{v}$ と，攻撃者を含めた全ユーザで推定した値 $\hat{v}$ の絶対差を最大化することである．単一エリアを対象とした DPA 法のアルゴリズムを以下に示す．

（i） 各攻撃者 k の報告値 v_k の初期化

（ii） 正常ユーザのみで CRH 法を用いて推定値 $\tilde{v}$ を算出

（iii） 全ユーザを対象に CRH 法を用いて推定値 $\hat{v}$ を算出

（iv） 各攻撃者 K に対し，式 (3) で報告値 v_k を更新

$$v_k = v_k + 2 \times (\hat{v} - \tilde{v}) \times \frac{w_k}{\sum_{k \in \boldsymbol{N} \vee \boldsymbol{A}} w_k} \tag{3}$$

（v） v_k が収束するまで (ii)〜(iv) を反復

4. DPA 法の複数エリアへの拡張

4.1 アルゴリズム

本節では，複数エリアを対象に DPA 法を適用する．以下に各攻撃者を順次，攻撃の効果が最大となるエリアに一人づつ配置する貪欲算法をベースとしたアルゴリズムを示す．

（i） 全攻撃者を未配置状態に初期化

（ii） 未配置攻撃者の中から一人，攻撃者 m を選択
（iii） 各エリア s を選択し，m を仮配置
（iv） エリア s に DPA 法を用いて，正常ユーザのみを用いて算出した推定値 $\tilde{v}_s$ と，全ユーザを対象に算出した推定値 $\hat{v}_s$ を算出
（v） 式 (4) でエリア s に攻撃者 m を配置したときのエリア s の誤差 e_s を算出

$$e_s = |\hat{v}_s - \tilde{v}_s| \tag{4}$$

（vi） e_s が最大となるエリア s に攻撃者 m を配置
（vii） 全攻撃者を配置し終わるまで (ii)〜(vi) を反復

4.2 想 定 条 件

エリア数を $N = 10$ に設定し，各エリアに正常ユーザ $k_s = 32$ 人を配置する．正常ユーザの報告値は正規乱数で与えるが，平均は全エリアで 50 に統一し，標準偏差をエリアごとに最小が 10，最大が 55 の 5 刻みで設定する．攻撃者の初期報告値は全て 50 とし，10 回の試行における平均値で評価する．

4.3 結果と考察

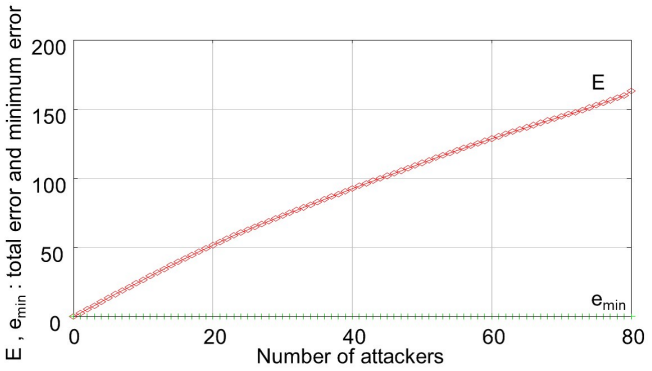

図 2 攻撃者数に対する総誤差と最小誤差

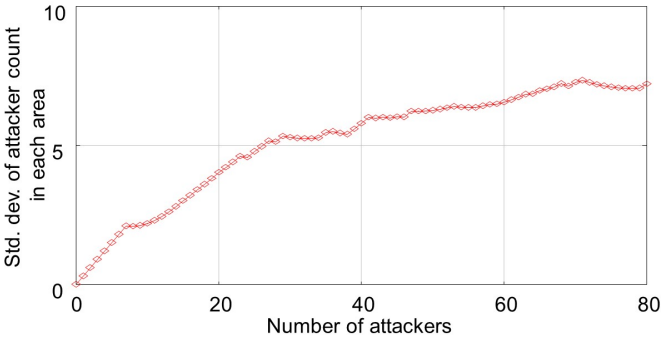

図 3 攻撃者数に対する攻撃者数の標準偏差

図 2 に攻撃者数を変化させたときの総誤差 E と最小誤差 e_{min} を，また図 3 には各エリアの配置攻撃者数の標準偏差を各々示す．攻撃者数の増加に伴い総誤差 E は単調に増加するが，最小誤差 e_{min} は常に 0 のまま変化しない．正常ユーザの報告値の分散が大きなエリアほど，攻撃者を一人配置することで増加する誤差が大きいため，正常ユーザの報告値の分散が大きなエリアに偏って攻撃者が配置される．そのため攻撃者数の増加に伴い配置攻撃者数の標準偏差が増加する．このように，エリアの中には攻撃者が配置されやすいエリアと配置されにくいエリアが存在し，配置されやすいエリアにはどんどん攻撃者が配置されているので総誤差 E は増加するが，配置されにくいエリアには攻撃者が一切配置されないため最小誤差 e_{min} は常に 0 となる．

このように DPA を単に複数エリアに適用した場合，攻撃者の最適化目標としてあげた総誤差 E の最大化に関しては達成できるが，各エリアの誤差 e_s の最小値 e_{min} は常に 0 となる．そのため本稿では，誤差の上限値，下限値を設定することで，より多くのエリアに攻撃者を配置し，全エリアで誤差が発生することを目指す．

5. 複数エリアにおける Crowdsensing の最適データポイゾニング攻撃

本節では，本稿で新たに提案する，複数エリアの MCS サービスを対象とした最適データポイゾニング攻撃法について述べる．1. 節で述べたように，攻撃者の最適化目標として，全エリアの誤差の総和 E と，各エリアの誤差 e_s の最小値 e_{min} の，2 つの最大化を考える．

5.1 誤差上限法: ULM (upper limit method)

誤差 e_s の上限値 η_u を設定し，誤差 e_s が η_u 未満の各エリア e に攻撃者を一人，追加配置していく．このとき，誤差 e_s が上限値 η_u に達しているエリアは追加配置する対象エリアからは除外する．一人の攻撃者を各エリア s に追加配置した際の e_s の増加量を算出する．そして $e_s < \eta_u$ の全エリアの中で，e_s の増加量が最大となるエリアに攻撃者を一人，追加配置する処理を全エリアの誤差 e_s が上限値 η_u に達するまで反復する．全エリアの誤差が上限値 η_u に達した場合は，e_s の増加量が最大となるエリアに攻撃者を配置する．この処理を全攻撃者の配置が完了するまで反復して行う．

5.2 誤差下限法: LLM (lower limit method)

誤差 e_s の下限値 η_l を設定し，誤差 e_s が η_l 未満のエリアの中から，誤差 e_s が最小のエリアに攻撃者を一人配置する処理を全エリアの誤差 e_s が下限値 η_l に達するまで反復する．全エリアが下限値 η_l に達した場合は，e_s の増加量が最大となるエリアに攻撃者を配置していく．この処理を全攻撃者の配置が完了するまで反復する．

6. 数 値 評 価

6.1 全エリアの特性が同一の場合

エリア数を $N = 10$ に設定し，各エリアに正常ユーザ $k_s = 32$ 人を配置する．正常ユーザの報告値は，平均が 50，標準偏差が 40 の正規乱数で与える．攻撃者の総数は $M = 80$ 人とし，攻撃者の初期報告値は 50 とする．結果は全て 10 回の施行における平均値で評価する．

図 4 は上限値，及び下限値に対する，総誤差 E と最小誤差 e_{min} を，図 5 は上限値，及び下限値に対する，各エリアの配置攻撃者数の標準偏差を，ULM と LLM について各々示す．

誤差上限法 (ULM) においては，η_u が小さい領域では，すぐに全エリアの e_s が η_u に達するため，攻撃者は e_s の増加量が大きなエリアに集中的に配置され，E が大きく e_{min} が小さい．また η_u の増加に伴い，より e_s の増加量が小さいエリアに攻撃者が多く配分される結果，E は減少していく半面 e_{min} は増加し，攻撃者の偏りは減少する．さらに η_u を増加させていくと，再度，より多くの攻撃者を e_s の増加量が大きなエリアに割り当て可能となり，η_u の増加に伴い E が増加するが，e_{min} は減少し，配置攻撃者数の偏りがより大きくなる．

一方，誤差下限法 (LLM) においては，η_l が小さい領域では，すぐに全エリアの e_s が η_l を上回るため，攻撃者は e_s の増加量が大きなエリアに集中的に配置され，E が大きく e_{min} が小

さい．また η_l の増加に伴い，より e_s の増加量が小さいエリアに攻撃者が多く配分される結果，E は減少する半面 e_{min} は増加し，攻撃者の偏りは減少する．さらに η_l を増加すると，より多くの攻撃者が必要な e_s の増加量が小さなエリアに多くの攻撃者を割り当てる．そのため η_l の増加に伴い E が減少するが，e_{min} が増加し，配置攻撃者数の偏りも増加する．しかし，さらに η_l を増加させると E と e_{min} は η_l が変化しても一定となる．これは，どのエリアも η_l の値に届かなくなり，攻撃者の配置状況が変わらなくなるためである．

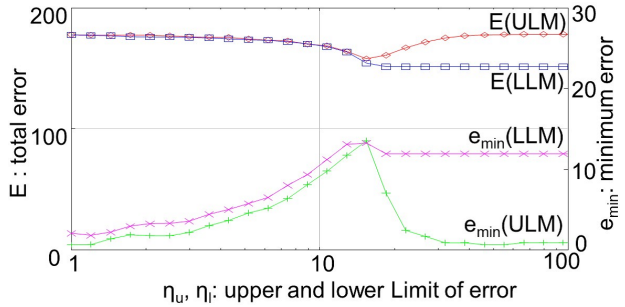

図 4　誤差の上限・下限値に対する総誤差と最小誤差【全エリア同一条件】

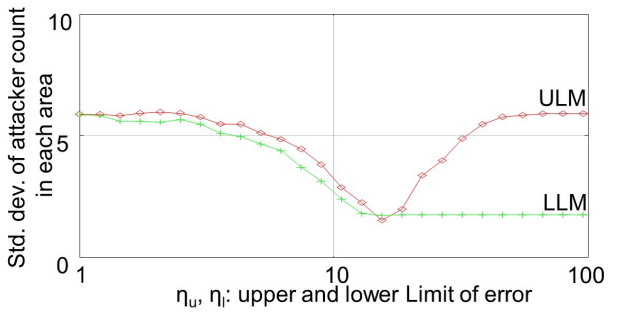

図 5　誤差の上限・下限値に対する各エリアの配置攻撃者数の標準偏差【全エリア同一条件】

以上の結果から今回の評価条件においては，E と e_{min} の両方を大きくできる，η_l を 12 程度に設定した LLM が攻撃者にとって望ましい攻撃者配置法といえる．

6.2　正常ユーザ数がエリアごとに異なる場合

エリアごとの正常ユーザ数 k_s を，最小が 5，最大が 59 で，6 刻みでエリアごとに変化させた場合で評価する．他の条件は 6.1 節の条件と同一である．

図 6 に上限値，及び下限値に対する総誤差 E と最小誤差 e_{min} を，図 7 に上限値，及び下限値に対する各エリアの配置攻撃者数の標準偏差を，ULM と LLM の各々について示す．また図 8，図 9 には上限値 η_u 及び下限値 η_l が 2.0，12.0，70.0，の三つの場合における，各エリアの配置攻撃者数の平均値をプロットする．

総誤差 E と最小誤差 e_{min} といった数値の増減の原因は 6.1 節の時と 2 つの手法とも変わらない．6.1 節と異なる点としてエリアごとの正常ユーザ数 k_s に差異が生じたことによって，k_s が多いエリアでは誤差が出にくいため上限値 η_u，下限値 η_l に達するまでに必要な攻撃者数が増加したことや，k_s が少ないエリアでは誤差が出やすいため攻撃者 1 人が与える影響が大きくなったことがあげられる．

誤差上限法については，数値の増減は図 4 と変わらない．しかし誤差の出やすいエリアと出にくいエリアで差が生まれたこ

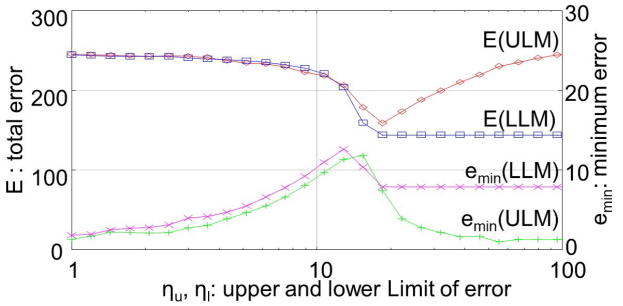

図 6　誤差の上限・下限値に対する総誤差と最小誤差【各エリアのユーザ数が異なる場合】

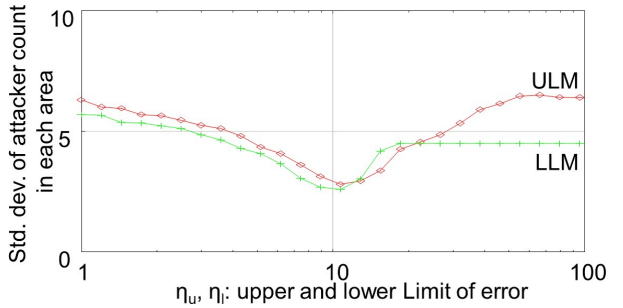

図 7　誤差の上限・下限値に対する各エリアの配置攻撃者数の標準偏差【各エリアのユーザ数が異なる場合】

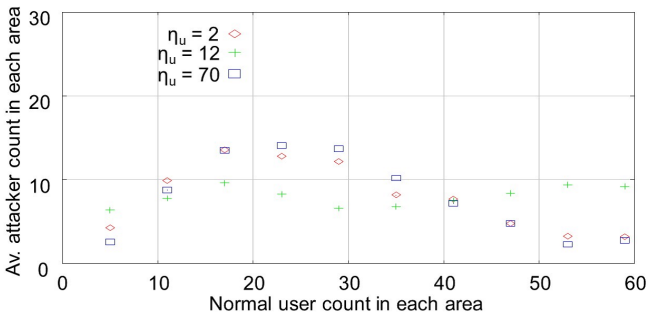

図 8　誤差上限法の各エリアの平均攻撃者配置数【各エリアのユーザ数が異なる場合】

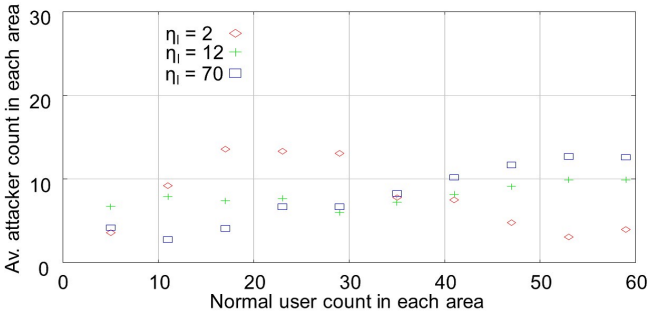

図 9　誤差下限法のエリアの平均攻撃者配置数【各エリアのユーザ数が異なる場合】

とで値の増減量が大きくなっている．攻撃者の配置状況も 8 より，上限値 $\eta_u = 2.0$ や $\eta_u = 70.0$ のように攻撃者の偏りが大きく，正常ユーザ数が少ないエリアに多く配置できる場合は負の相関が見れる．一方，上限値 $\eta_u = 12.0$ のような攻撃者の偏りが小さく，正常ユーザが多いエリアに配置される場合には弱い正の相関が見える．

誤差下限法についても，誤差上限法と同様に，数値の増減は図 4 と変わらないが数値の増減幅が大きくなっており，下限値

η_l が大きくなった際の攻撃者の偏りが大きくなっている．これは誤差が出にくいエリアにおいて，必要となる攻撃者の数が増えたためである．攻撃者の配置状況も 9 より，下限値 $\eta_l = 12.0$ や $\eta_l = 70.0$ のように正常ユーザ数が少ないエリアに多く配置できる場合は弱い正の相関が見れる．上限値 $\eta_u = 2.0$ のような正常ユーザが多いエリアに配置される場合には負の相関が見える．

以上の結果から，今回の評価条件においては，E と e_{min} の両方を大きくできる，η_l を 12 程度に設定した LLM が攻撃者にとって望ましい攻撃者配置法といえる．

6.3 正常ユーザの報告値の標準偏差がエリアごとに異なる場合

本節では 6.1 節で行ったシミュレーション条件をベースに，正常ユーザの報告値の標準偏差をエリアごとに最小が 10，最大が 55 の 5 刻みで設定した場合の評価を行う．図 10 に上限値，及び下限値を増加させた際の総誤差 E と最小誤差 e_{min} の変化について，図 11 に上限値，及び下限値を増加させた際の各エリアの配置攻撃者数の標準偏差の変化について同様に示す．また，図 12 と図 13 に，上限値 η_u 及び下限値 η_l が 2.0，12.0，70.0，の三つの場合の，各エリアの割当攻撃者数の平均値をプロットする．

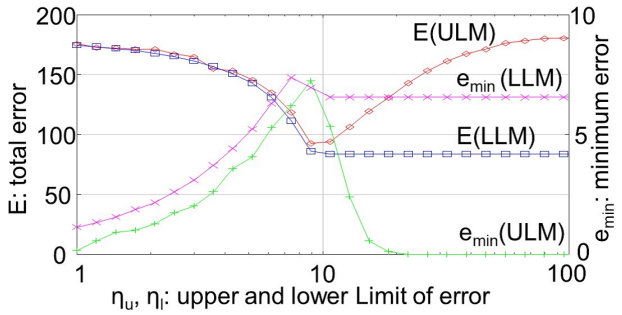

図 10　誤差の上限・下限値に対する総誤差と最小誤差【各エリアの報告値の標準偏差が異なる場合】

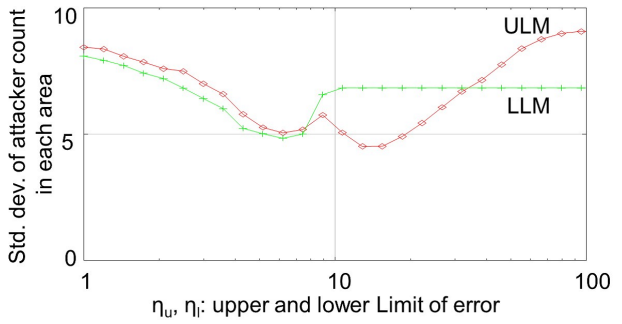

図 11　誤差の上限・下限値に対する各エリアの配置攻撃者数の標準偏差【各エリアの報告値の標準偏差が異なる場合】

総誤差 E と最小誤差 e_{min} といった数値の増減の原因は 6.1 節の時と 2 つの手法とも変わらない．6.1 節と異なる点としてエリアごとの正常ユーザ報告値の標準偏差に差が生まれたことによって，標準偏差が低いエリアでは誤差が出にくいため上限値 η_u，下限値 η_l に達するまでに必要な攻撃者数が増加したことや，標準偏差が大きいエリアでは誤差が出やすいため攻撃者 1 人が与える影響が大きくなったことがあげられる．

誤差上限法においては，数値の増減は図 4 と変わらない．しかし，誤差の出やすいエリアと出にくいエリアで差が生まれた

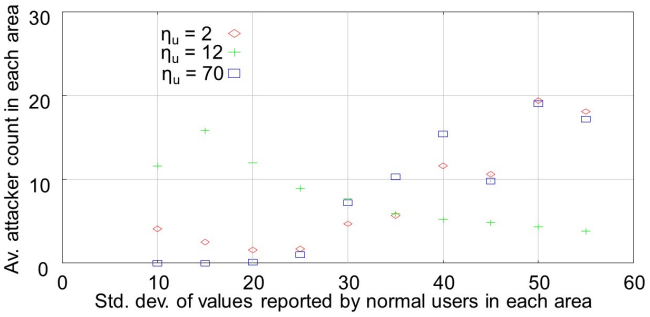

図 12　誤差上限法の各エリアの平均攻撃者配置数【各エリアの報告値の標準偏差が異なる場合】

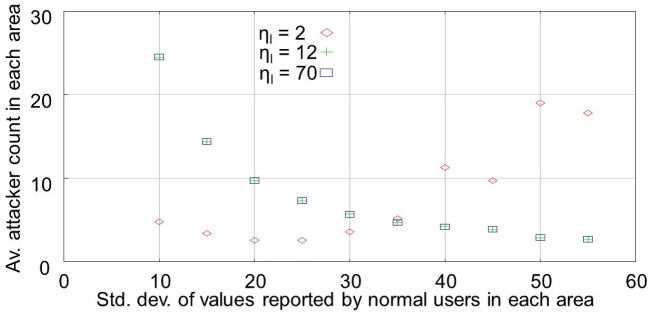

図 13　誤差下限法の各エリアの平均攻撃者配置数【各エリアの報告値の標準偏差が異なる場合】

ことで値の増減量が大きくなっている．特に最小誤差 e_{min} は η_u が 10 付近で減少し，最終的には 0 で変わらなくなる．また，攻撃者の偏りは常に高い水準にまとまっている．攻撃者の配置状況も 12 より，上限値 $\eta_u = 2.0$ や $\eta_u = 70.0$ のように攻撃者の偏りが大きく，正常ユーザ数が少ないエリアに多く配置できる場合は強い正の相関が見れる．上限値 $\eta_u = 12.0$ のような攻撃者の偏りが小さく，正常ユーザが多いエリアに配置される場合には強い負の相関が見える．

誤差下限法においても同様に，数値の増減は図 4 と変わらないが数値の増減幅が大きくなっている．特に下限値 η_l が大きくなった際の攻撃者の偏りが大きくなっている．これは誤差が出にくいエリアにおいて，必要となる攻撃者の数が増えたためである．誤差が出にくいエリアで攻撃者の数が増えたことで最小誤差 e_{min} は大きくなっている．攻撃者の配置状況も 13 より，下限値 $\eta_l = 12.0$ や $\eta_l = 70.0$ のように正常ユーザ数が少ないエリアに多く配置できる場合は強い負の相関が見れる．上限値 $\eta_u = 2.0$ のような正常ユーザが多いエリアに配置される場合には強い正の相関が見える．

以上の結果から，今回の評価条件においては，E と e_{min} の両方を大きくできる，η_l を 7 程度に設定した LLM が攻撃者にとって望ましい攻撃者配置法といえる．

6.4 2 つの尺度を考慮した目的関数による最適化

1.2 節で述べたように，攻撃者の最適化目標として，本稿では全エリアの誤差の総和 E 最大化と，各エリアの誤差 e_s の最小値 e_{min} の最大化の 2 つを想定している．そして 5 節と，6.1〜6.3 節では，各エリアの誤差に上限値や下限値を設け，誤差の総和 E の最大化の観点から攻撃者の割当を行うことで，これら 2 つの最適化目標のバランスをとることを検討した．

本節では別のアプローチとして，誤差の総和とエリアごとの誤差の偏りの両方を考慮した目的関数を定義し，その値の増加量が最大となるエリアに攻撃者を 1 人つづ配置していく方法を検討する．その際，目的関数式に最低誤差 e_{min} を用いると，あ

るエリアに 1 人の攻撃者を追加配置したときに最低誤差 e_{min} は変化しない場合があり，e_{min} を考慮した攻撃者割当アルゴリズムの構築が難しいことから，最低誤差 e_{min} の代わりに各エリアの誤差の標準偏差 e_{sd} を用いる．目的関数として次式を定義する．

$$f(\alpha) = \alpha \times E - (1-\alpha) \times e_{sd} \tag{5}$$

式 (5) で使用している α は 0 から 1 の範囲の実数パラメタである．攻撃者が 2 つの最適化目標のどちらをより重視するかに応じて α の値を設定することで，2 つの最適化目標のバランスをとることが可能である．

エリア数を $N = 10$ に設定し，各エリアに正常ユーザ $k_s = 32$ 人を配置する．正常ユーザの報告値は，平均が 50，標準偏差がエリアごとに最小が 10，最大が 55 の 5 刻みで設定した正規乱数で与える．攻撃者の総数は $M = 80$ とし，攻撃者の初期報告値は全て 50 とする．上限値 η_u 及び下限値 η_l は考慮しない．図 14 に α を変化させた際の総誤差 E 及び，最低誤差 e_{min} を，図 15 に α を変化させた際の各エリアの誤差の標準偏差 e_{sd} を各々プロットする．

α の増加に伴い，総誤差 E が誤差の標準偏差 e_{sd} や最低値 e_{min} よりも相対的に優先されるため，α の増加に伴い E が増加し，e_{min} が減少して e_{sd} が増加することが予想される．実際に図 14，図 15 より，これらのことが確認され，α の値を変化させることで E と e_{min} の間のトレードオフの調整が可能なことが確認できる．

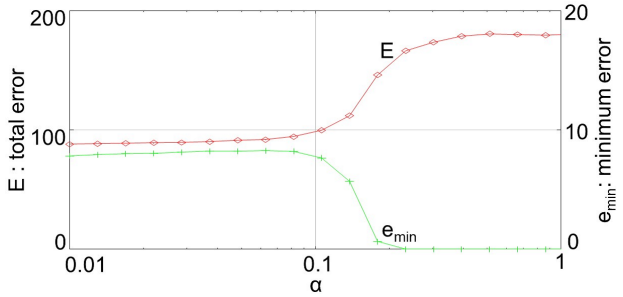

図 14　誤差の上限・下限値に対する総誤差と最低誤差

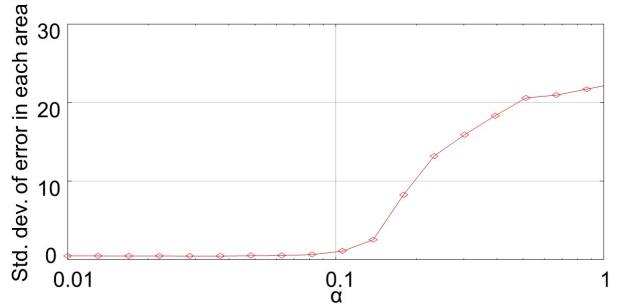

図 15　誤差の上限・下限値に対する各エリアの誤差の標準偏差

7. ま と め

モバイルクラウドセンシングが注目されているが，悪意のあるユーザが誤差の大きなデータを意図的に送信することで，計測値の推定誤差を増大させるデータポイゾニング攻撃の問題が指摘されている．本研究では，複数エリアに存在するユーザから収集したデータに対して，エリアごとに測定値を推測するクラウドセンシングにおけるデータポイゾニング攻撃を想定し，そして全エリアの総誤差と，エリアごとの誤差の最小値の二つの観点から，攻撃者が攻撃効果を最大化するよう各エリアに配置する攻撃者数を設計する方式を提案した．すなわち誤差に閾値を設けることで総誤差を大きくしつつより多くのエリアで誤差を生み出せる方式を 2 つ提案し，複数の条件下で計算機シミュレーションによる評価を行い，より多くのエリアで誤差を生みだし最小誤差を大きくすることが可能であることを示した．また 2 つの最適化目標を含む目的関数を提案し，本目的関数を用いることで 2 つの最適化目標の間のトレードオフが実現できることを確認した．

謝辞　本研究成果は，KDDI 財団研究助成寄付金 190051 の助成を受けたものである．ここに記して謝意を表す．

文　献

[1] Q. Li, et al., Conflicts to Harmony: A Framework for Resolving Conflicts in Heterogeneous Data by Truth Discovery, IEEE Trans. Know. Data Eng., 28 (8), Aug. 2016

[2] Z. Huang, M. Pan, and Y. Gong, Robust Truth Discovery Against Data Poisoning in Mobile Crowdsensing, IEEE GLOBECOM 2019

[3] C. Miao, Q. Li, L. Su, M. Huai, W. Jiang, and J. Gao, Attack under Disguise: An Intelligent Data Poisoning Attack Mechanism in Crowdsourcing, WWW 2018

[4] C. Miao, Q. Li, H. Xiao, W. Jiang, M. Huai, and L. Su, Towards Data Poisoning Attacks in Crowd Sensing Systems, ACM Mobihoc 2018

[5] K. Quintal, E. Kara, M. Simsek, B. Kantarci, and H. Viktor, Sensory Data-driven Modeling of Adversaries in Mobile Crowdsensing Platforms, IEEE GLOBECOM 2019

[6] L. Varshney, Privacy and Reliability in Crowdsourcing Service Delivery, SRII 2012