複数地点 Crowdsensing におけるポイゾニング攻撃法

Poisoning Attacks on Crowdsensing at Multiple Locations

藤本 凛 [1]　　上山 憲昭 [2]
Rin Fujimoto　　Noriaki Kamiyama
福岡大学 工学部 電子情報工学科 [1]
Faculty of Engineering, Fukuoka University
立命館大学 情報理工学部 [2]
College of Information Science and Engineering, Ritsumeikan University

1. はじめに

多数のモバイル機器から報告されたデータから真の測定値を推定するクラウドセンシングの利用が拡大している．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，センサーの故障等で誤ったデータの送信，特に悪意のあるユーザが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 [1]．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，ユーザごとの信頼度を推定し，信頼度を重みとした測定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 [2]．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 DPA (Data poisoning attack) と，その防御法が提案されている [3]．これらの研究では，単一地点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，複数の地点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，複数地点に存在するユーザから収集したデータに対して，地点ごとに真値を推測する Crowdsensing におけるデータポイゾニング攻撃を想定し，攻撃者が攻撃効果を最大化するよう各地点に配置する攻撃ユーザ数を設計する方式を考察する．また，今回考察した攻撃者の配置方式が，推定値の誤差に与える影響を分析する．

2. データポイゾニング攻撃の攻撃者配置法

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 [3] で提案されている DPA 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 v_s^n と，攻撃ユーザを含む全ユーザの報告値で計算した推定値 v_s^a との差の絶対値を誤差 e_s と定義する．攻撃者は，通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 の最大化と，各エリア s の誤差 e_s の最小値 e_{min} の最大化の 2 つを考える．

攻撃者を分配する基準を「攻撃者を一人，追加配置することで得られる e_s の増加量が最大のエリアに配置する」とした場合，全エリアの誤差の総和 E は最大となるが，誤差が極端に小さいエリアが発生する場合があり，2 つ目の目的を達成できない．その為，誤差の上限値，もしくは下限値を設定することで，より多くのエリアで誤差が均一に発生することも同時に目指す．以下に，これら 2 つの攻撃者配置法を説明する．

誤差上限法 ULM(upper limit method): 誤差 e_s が事前に定めた上限値 η_u 未満の各エリア s に攻撃者を一人，追加配置したときの e_s の増加量を算出し，その値が最大となるエリアに攻撃者を配置する処理を反復する．全エリアが上限値 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．本処理を全攻撃者の割当が完了するまで反復する．

誤差下限法 LLM(lower limit method): 誤差 e_s が事前に定めた下限値 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する．本処理を全攻撃者の割当が完了するまで反復する．

3. 性能評価

エリア数を $N = 10$ に設定し，各エリアに通常ユーザを 32 人配置する．通常ユーザの報告値を平均が 50 の正規乱数で与える．ただし正規乱数の標準偏差をエリアごとに最小が 10，最大が 55 の 5 刻みで設定する．攻撃ユーザを $M = 80$，すべての攻撃ユーザの初期報告値を 50 とする．結果は全て 10 回の試行における平均値で評価する．図 1 に上限値，もしくは下限値を増加させた際の総誤差 E と最小誤差 e_{min} をプロットする．また図 2 に各エリアの配置攻撃者数の標準偏差を同様に示す．

誤差上限法から分析する．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るため，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_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る．そのため 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増加する．

次に誤差下限法を分析する．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る．その為，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．

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 70 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．

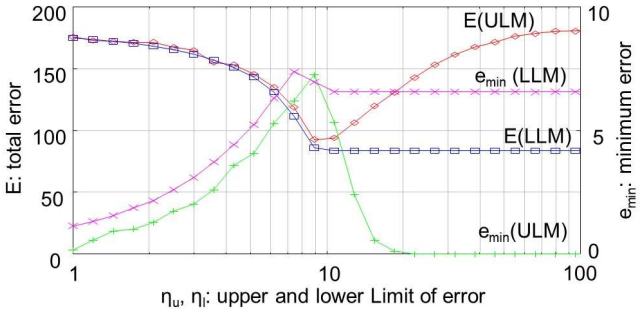

図 1: 総誤差と最小誤差

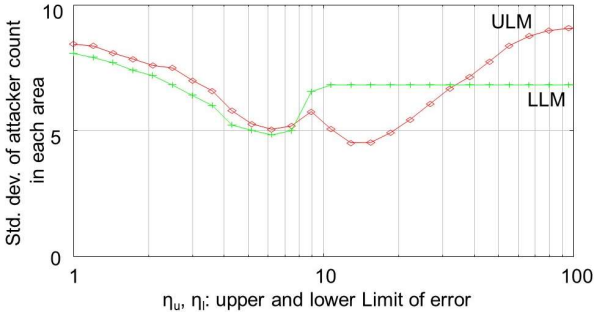

図 2: 攻撃者数の偏り

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．

[1] C. Miao, et al., Attack under disguise: An intelligent data poisoning attack mechanism in crowdsourcing, WWW 2018
[2] 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
[3] Z. Huang, et al., Robust Truth Discovery Against Data Poisoning in Mobile Crowdsensing, GLOBECOM 2019