

欺瞞的対話における判断に関する ゲーム理論に基づく分析について

An Analysis of Judgment in Fraudulent Communication based on Game Theory

山下雅弘^{1*} 明神聖子^{2†} 馬場口登^{2‡}
Masahiro Yamashita¹ Seiko Myojin² Noboru Babaguchi²

¹ 大阪大学工学部

¹ School of Engineering, Osaka University

² 大阪大学大学院工学研究科

² Graduate School of Engineering, Osaka University

Abstract: Fraudulent communication has bad influence on our society and study is needed to consider why people are deceived. We focus on billing fraud as a kind of fraudulent communication. In this paper, we describe the first step of analysis related to victims' judgement based on game theory. We expect that our study also includes suggestion for designing systems with artificial intelligence.

1 はじめに

現在、社会の中で通信技術が発展を遂げており、携帯電話並びにインターネットの普及に伴い、世界中の誰とでもコミュニケーションを取れる時代になった。また、近年の画像や音声の生成技術においても発展を遂げている。例えば、Salimans ら [1] は敵対的生成ネットワークによって、高い精度で、ある人物の顔から別の人物の顔を生成できる手法を提案している。Ping ら [2] は 2000 人以上から 800 時間を超える音声データを集め、これらを基に高い精度の音声生成手法を提案している。画像や音声以外の生成技術に注目してみると、Coburn ら [3] は Twitter 上で任意のアカウント ID とプロフィール画像を用いて、アカウント並びになりすまし投稿の生成を行った。

しかし、このような技術は不正目的で扱われる危険がある。改めて、画像や音声の生成技術について考慮してみると、今後は画像や動画などに出てくる人物などを模倣した偽メディアと接する機会が増えてくると考えられる。本論文ではこのような偽メディアのことを

メディアクローンと呼ぶことにする。メディアクローンの識別が電話や Skype などを行う際に困難な時、それを見る人は本人だと思い会話をする可能性が高い。今後これらの技術を利用した新しいタイプの詐欺の発生が懸念される。画像や音声の生成技術の不正な悪用は既存の詐欺にも影響を及ぼすと考えられ、オレオレ詐欺について考えてみると、日常的に息子の顔や声を認知している女性でさえ分からないメディアクローンが詐欺を仕掛けてくると考えられる。現在の日本では度々詐欺の話題がテレビで取り上げられている程、様々な詐欺の被害が広がりつつあり、これへの対策が急がれる。

このような背景を基に、本研究では日本で発生している詐欺の一つである架空請求詐欺に焦点を当て、なぜ騙されるのかを考察する。特に判断に影響を及ぼす架空請求詐欺の特徴や要因を明らかにすること、および明らかにするための一つの方法を示すことを目的とする。本研究では架空請求詐欺における行動を分析していくが、このような行動を分析するのにしばしば使われる手法としてゲーム理論がある。ゲーム理論とは経済や社会における行動主体の相互に依存する意思決定を数学モデルを用いて研究する学問である [4]。ゲーム理論は金銭的取引の分析手法として幅広く使われている。本研究では架空請求詐欺をゲーム理論に基づいて分析することによって、架空請求詐欺において仕掛けられる側の判断に影響を及ぼす特徴や要因を考察する。

本論文の構成は以下の通りである。第 2 章において関連研究を述べ、第 3 章では、分析対象と提案手法の

*連絡先：大阪大学工学部電子情報工学科

〒 565-0871 大阪府吹田市山田丘 2-1

E-mail: yamashita [at] nanase.comm.eng.osaka-u.ac.jp

†連絡先：大阪大学大学院工学研究科電気電子情報工学専攻

所属機関住所同上

E-mail: myojin [at] nanase.comm.eng.osaka-u.ac.jp

‡連絡先：所属機関名同上

所属機関住所同上

E-mail: babaguchi [at] comm.eng.osaka-u.ac.jp

詳細を説明する。次に第4章で、実際に行った分析と、それに対する考察を述べる。最後に第5章で、結論と今後の課題を述べる。

2 関連研究

ここでは、欺瞞的対話をゲーム理論の枠組みで分析している研究を紹介する。まず、シグナリングゲームをモデルとして用いたものにCarrollら[5]とMohammadiら[6]の研究がある。Carrollら[5]はコンピュータネットワークの攻撃者と防御者の間の作用に対する詐欺効果の調査を行った。調査の結果、防御者が自身(内部ユーザー)への不正アクセスを防げるような条件を導いた。また、Mohammadiら[6]は外部ユーザーによって不正アクセスされたソーシャルネットワーク内のユーザーを特定するために、防御者が偽のアカウントを展開するネットワークをシグナリングゲームモデルで作成した。その結果、偽のアカウントが内部ユーザーに外部からのアクセスを警告する条件を導出した。大倉ら[7]は保険金詐欺発生メカニズムを不完備情報動学ゲームというモデルで分析した。その結果、各保険会社の枠を超えた情報交換の重要性を示した。小田中ら[8]は社会学の観点から、ハイパーゲームを用いてなぜ詐欺は成功するのかを分析した。研究の結果から嘘の誘導スタイルによって詐欺の類型化ができることを示した。以上のようにゲーム理論に基づいて、仕掛けられる側の判断に影響をおよぼす、架空請求詐欺の特徴や要因を明らかにする研究は見つかっていない。

3 提案方法

本章では、ゲーム理論に基づく分析方法を提案する。

3.1 分析対象

架空請求詐欺とは、突然メールを行い利用料金を請求、放置すると裁判に訴えるなどと脅し(本研究では詐欺を仕掛ける側のこれらの行動を指示と呼ぶ)金銭を騙しとる事とされている[9]。つまり犯人からの指示で始まり、各指示を被害者が実行することによって詐欺が成立していることが分かる。本研究ではこれを架空請求詐欺の特徴と捉え、モデル化において考慮する。

3.2 ゲーム理論に基づいた分析方法の提案

ゲーム理論は、二人、あるいはそれ以上の人々の駆け引きの状況を相手の出方を予想しながら自身の打つ手を考えるというのが基本的な考えかたであり、個人

と社会など二者の相互関係を分析するのに適している[10]。

本研究では架空請求詐欺の登場人物をゲーム理論にならってプレイヤーと呼ぶ。何かしらの情報を発信するプレイヤーを送り手(A)、Aから発信された情報を受け取るプレイヤーを受け手(B)と定義する。詐欺を仕掛けられる側の行動基準を分析するためにBの行動を考察する事が本研究の本筋となる。架空請求詐欺は二者のやり取りから成り立つことを踏まえ、プレイヤーが順番に行動を選択する展開型ゲームと考えられる。各プレイヤーの行動であるが、Aは「指示」と「断念」の二つの行動をとり、Bは「妥協」と「中断」の二つの行動をとると本研究では設定する。序論でも説明した背景から、本研究ではAは悪意をもたない真性というタイプと悪意を持った疑似という二つのタイプを持つと考える。BはAがどちらのタイプなのかを知らずに、信念と言われる主観的な確率分布を基に判断する[10]。Bは q という確率でAを真性と判断し、 $1-q$ の確率で疑似と判断する。本研究のように、送り手と受け手の情報に非対称性があるゲームのことを不完備情報ゲームという。したがって、展開型不完備情報ゲームをモデルに採用することによって、架空請求詐欺をモデル化できる。

3.3 不完備情報ゲームのツリー

この節では一般的なゲームツリー、並びに手法として用いるゲームツリーの概要について説明する。ゲームツリーを視覚的に理解してもらうために、その例を図1に示す。まず、ツリーは点および点をつなぐ枝から構成される[11]。枝が分かれる分岐点はプレイヤーの意思決定点で[4]、点は各プレイヤーが行動する順番を表す[11]。ゲームの結果を表す点[11]を終点と呼び、プレイヤーごとの利得を書く[12]。利得とは金銭などの物理的な利益やお得感や満足感などの心理的な利益のことを示す[10]。ツリーは必ずひとつの意思決定点、根とも呼ばれる点から始まる[12]。

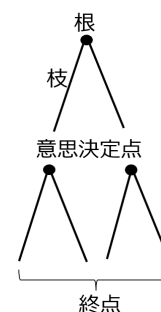


図1: 一般的なゲームツリー

一方、展開型の不完備情報ゲームでは一般的なゲームツリーとは少し形が変わり、根として自然という偶然手番を設ける [10]. 不完備情報ゲームのツリーを図 2 に示す. 点線は B が属する情報集合を表す. B は自分が情報集合のどの枝にいるのかわからない [10]. 同じ情報集合に属している枝から選択され得る行動は、同じでなければならない [10]. 図 2 の p は A が二つあるタイプのどちらかであるかの確率で、 q は B の信念を意味している. 本研究では図 2 のような自然を根とするツリーを基に詐欺の分析を行っていく.

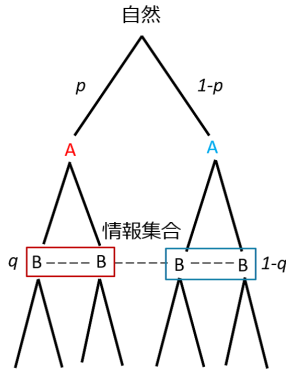


図 2: 不完備情報ゲームのツリー

4 架空請求詐欺のモデル化

架空請求詐欺をモデル化した結果を図 3 に示す. ツリーの終点に記載されているのは利得であり、左側はプレイヤー A (A) の利得、右側はプレイヤー B (B) の利得である. 変数は全て 0 より大きいとする. 架空請求詐欺のシナリオを設定し、その中で利得を説明する. 架空請求詐欺の登場人物は女性 (B) と携帯会社職員 (真性 A) と詐欺師 (疑似 A) とする. 携帯会社職員または詐欺師がアクションを起こす事から始まる. A のタイプに関わらず A (携帯会社職員または詐欺師) が女性に何らかの要求を出すことを断念した場合、お互い何も得ていないので両者の利得は 0 とする. 次に携帯会社職員が女性に携帯の利用料金を要求し、女性が応じれなかった場合、職員は支払ってもらえると思っていた料金が支払ってもらえなかった損失 (W_a) を被り、女性は支払えなかった事への不安な気持ちという損失 (W_b) を被る. 女性が職員の要求に応じた場合、職員は支払ってもらえた事への嬉しい気持ちという利得 (R_a) を獲得でき、女性は金銭は損失したが安心感を得られるので中断した時より少ない損失 ($-R_b$) を被る. そして、詐欺師が不当な要求を行い、女性が中断した時の両者の利得は、されるべきでない取引が阻止されたので 0 である. しかし、女性が実行した場合、

女性は金銭を失い損失 ($-X_b$) を被る. 一方、詐欺師は利得 (X_a) を得る事ができる.

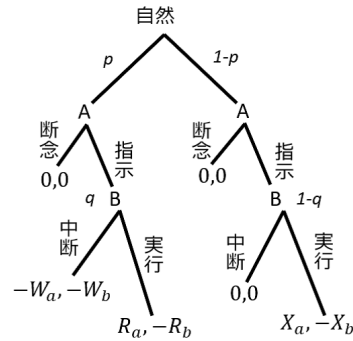


図 3: 架空請求詐欺のゲームツリー

B の行動の分析を行うために、B が中断を選択した時と実行をした時の期待利得をそれぞれ求める. 期待利得とは、ある行動を選ぶ確率とその場合の利得をかけたものを全ての行動について足し合わせたものである. なお、B は自身の信念 (q) を基に A のタイプの確率を決め、その結果自分が得られる期待利得を予測し、自分の行動を選択する. この事を踏まえ、まず、中断した時の期待利得は以下の式で表せる.

$$q \times (-W_b) + (1 - q) \times 0 = -qW_b \quad (1)$$

一方、B が実行した時の期待利得は以下の式で表せる.

$$q \times (-R_b) + (1 - q) \times (-X_b) \quad (2)$$

中断の期待利得より実行の期待利得の方が大きくなっている時、B は実行する. ゆえに式 (1)(2) から次の式が成り立つ.

$$q \times (-R_b) + (1 - q) \times (-X_b) > -qW_b \quad (3)$$

式 (3) を q について解くと、($W_b - R_b + X_b > 0$) の時、以下の通り変形できる.

$$q > \frac{1}{\frac{W_b - R_b}{X_b} + 1} \quad (4)$$

5 考察

詐欺の実例と式 (4) を比較した一例を示す. まず、式 (4) は、A が偽物だった場合に、B の実行で生じる B の損失が大きければ騙されにくいということと、A が本物だった場合に、B の中断で生じる B の損失が大きければ実行するという意味を意味する. 以下の架空

請求詐欺の事例で考えると、訴状にBが屈することの理由付けになっている。

架空請求詐欺の事例 1[13]

被害者：50代女性
公的機関のような団体から「契約会社、運営会社から契約不履行により民事訴訟の訴状が出された」等と記載されたハガキが届いた。取下げ期日が翌日だったので驚き電話をすると、「自分で弁護士に委任するか、紹介する国選弁護士に連絡するか選ぶように」と言われ、国選弁護士を名乗る者の紹介を受けた。その者から「保証金10万円を支払えば取下げ手続きを取り、後日精算し返金する」と説明された。支払いに必要な番号を聞き、指示に従いコンビニの情報端末で10万円を支払った。支払後に団体に連絡したら、「明日以降、返金方法を連絡する」と言われたが、不審だ。

また、架空請求詐欺におけるBの利得の中では主に、焦りや心配などによる心理的な利益が影響することを意味している。これは本研究が、利得に心理的な利益を含めたことで得られた示唆である。

6 むすび

本研究では、架空請求詐欺において、指示してきた側の要求に対し「実行」するか「中断」するかの判断について分析した。ゲーム理論に基づいた本研究のモデルに事例との対応が見られた。本研究は、人の認知プロセスの構造を表現した意味もあり、将来的に人と共生する知的システムのための初歩的な示唆を含む。今後、詐欺のコミュニケーションをより詳細にモデル化することや詐欺師の利得にも着目する事で本研究を改良していける可能性がある。

謝辞

本研究はJSPS 科研費JP16H06302の助成を受けたものである。

参考文献

- [1] Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X.: Improved techniques for training GANs, *Neural Information Processing Systems conference*, 16 pages (2018)
- [2] Ping, W., Peng, K., Gibiansky, A., Arik, S. O., Kannan, A., Narang, S.J., Miller, J. R.: Deep voice 3: scaling text-to-speech with convolutional sequence learning, *International Conference on Learning Representations*, 9 pages (2016)
- [3] Coburn, Z., Marra, G.: Realboy:believable twitter bots, <http://ca.olin.edu/2008/realboy/index.html> (2011)
- [4] 岡田章: ゲーム理論・入門: 人間社会の理解のために, 有斐閣アルマ (2019)
- [5] Carroll, T. E., Grosu, D.: A game theoretic investigation of deception in network security, *International Conference on Computer Communications and Networks*, 6 pages (2009)
- [6] Mohammadi, A., Manshaei, M. H., Moghadam, M. M.: A game-theoretic analysis of deception over social networks using fake avatars, *GameSec*, Vol. 9996, pp. 382-394 (2016)
- [7] 大倉真人: 保険会社のモニタリングに関する一考察: 保険金詐欺発生メカニズムに関するゲーム論的分析, *生命保険論集*, pp. 231-270 (2001)
- [8] 小田中悠: なぜ詐欺は成功するのか: ハイパーゲームによる分析, 慶應義塾大学大学院社会学研究科紀要, No. 83, pp. 15-34 (2017)
- [9] 岡本真一郎: なぜ人は騙されるのか, 中公新書 (2019)
- [10] 佐藤嘉倫: ゲーム理論: 人間と社会の複雑な関係を解く, 新曜社 (2008)
- [11] 渡辺隆裕: ゼミナール ゲーム理論入門, 日本経済新聞出版社 (2017)
- [12] 鎌田雄一郎: ゲーム理論入門の入門, 岩波新書 (2019)
- [13] 独立行政法人 国民生活センター: 架空請求の相談が20万件を突破: 身に覚えがないと思ったら絶対に相手に連絡しないこと!, 報道発表資料 (2019)