

エージェントモデルシミュレーションと 機械学習を用いた顧客解約予測

林 眞司† 金盛 克俊‡ 大和田 勇人‡

東京理科大学大学院理工学研究科† 東京理科大学理工学部‡

1 はじめに

製品の解約を行う顧客の傾向を理解することは、企業が顧客取得から顧客保持へ経営戦略を移すための課題である。現在まで顧客解約予測を行うために様々な機械学習方法での成果が報告されており、これらの予測方法は離反予測分析に有効である。しかし、機械学習の結果は、顧客解約に関する決定で必ずしも適用されるわけではなく、機械学習の結果は戦略を実行すべき理想的な時期がわかっていない意思決定を導くだけである[1]。

ランドら[2]では機械学習だけでなくエージェント・ベース・モデルを組み合わせることで予測精度の向上を目指す実験が行われており、有効な意思決定を行うには、顧客保持プロセスの内に社会構造にもとづく動的なシミュレーションが必要であると考えられる。

2 本研究の目的

本研究では、顧客解約予測を分析するために、顧客の解約行動をモデル化し、動的構造のシミュレーションを行い、その結果から、シミュレーションベースで顧客の解約行動を予測する手法を提案する。

3 本研究の対象とする問題

本研究で顧客の解約行動を分析する対象は自動延長サービスとする。自動延長サービスは利用製品の利用期間が切れる直前に自動的に更新手続きを行い、更新料金を顧客から回収するサービスである。更新手続きは1年ごとに行われ、顧客はいつでも解約することができる。また製品の有効期限が切れる数日前から、上位製品の紹介メール、更新お知らせメールや期限切れ通知等のメールを送信され、更新お知らせ時に解約数が増えるなどのことがわかっている。

4 提案手法

顧客の解約行動をモデル化しシミュレーションを行う。シミュレーションでは将来顧客がどのような行動をとるか予測が必要となる。

4.1 エージェント・ベース・モデル

エージェント・ベース・モデルを作成し、顧客の解約行動のシミュレーションを行う。各顧客を表す Agent は $g = \{g_1, g_2, \dots, g_n\}$ とし、顧客がとる行動 ActionSet は $A = \{\text{Cancellation}, \text{Rennew}\}$ で表し、各顧客の状態 StateSet を $S = \{s_1, s_2, \dots, s_n\}$ で表す。またある経過日数でのある状態を $s_i = (t_{pc}, t_{ar}, pri)$ とし、 t_{pc} はある経過日数での PC 故障率、 t_{ar} をある経過日数でのメール受信による解約率、 pri を購入製品の価格によって変化する解約確率とする。上記の解約確率をふまえ、PC 故障率を $troubleRate$ 、メールでの解約率を $mailRate$ 、価格帯での解約率を $priceRate$ とし、本研究ではある経過日数での解約確率 P を

$$P(s) = troubleRate(t_{pc}) + mailRate(t_{ar}) + priceRate(pri) \quad (1)$$

と定義する。各エージェントがとる ActionSet により StateSet が変更されるようモデルを構成し、各エージェントそれぞれの現在の状態により実行される行動が変更されることとする。またシミュレーションは1日毎に行われる。

4.1.1 PC の故障確率

PC の故障確率は PC を利用している期間が長くなるほど高くなると考え、PC の使用期間に応じた故障確率をハザード関数で求める。ハザード関数は以下の式で表される。

$$F(t) = 1 - \exp(-(\lambda t)^p) \quad (2)$$

図1ではハザード関数を用いた PC の故障率を表す。

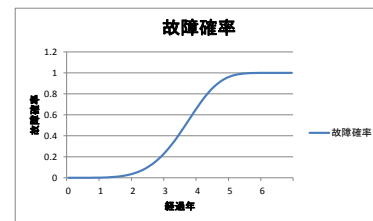


図1：PC 故障確率

4.1.2 メール受信後の解約確率

メール受信後の解約率の変化は、特定の日数でメールを受信してから契約期間が終わる間に顧客が解約する傾向があると仮定する。メールを受け取った日の解約確率が最大になり、契約終了日に近づくにつれて徐々に解約確率が低下する。

4.1.3 価格帯による解約確率

購入された製品価格も解約に影響があると仮定する。購入された製品の平均価格を基準に平均以下であれば、乗り換えなどの行動をとる可能性があると考え契約期間の前半に解約確率が増え、平均以上であれば継続期間は長くなり契約期間の後半に解約確率が増えると考え。

5 評価実験

5.1 本研究で用いるデータ

本研究では顧客情報と WEB アクセス履歴データを用いて分析を行う。WEB アクセス履歴データは2013年4月からデータの取得を開始したため、本研究では顧客情報・WEB アクセス履歴データともに2013年4月～2013年10月までの半年間のデータを用いて分析を行う。

5.2 実験結果

価格帯が解約行動を予測するのに有効か検証するため priceRate を用いる実験と用いない実験を行い比較する。mailRate と troubleRate での実験結果は図2,表1に示す。図2より、メール受信日の解約率の合計が利用者の1%になると仮定して実験を行ったが、契約期限切れ前後で実際の解約者数とシミュレーション結果に大きく差があり、解約傾向を表すことができていない。相関係数も低い値となっている。そこで利用者全体の5%が解約すると仮定し、再度実験を行うと契約期限切れ前後の予測が改善され、相関係数が0.77と上昇した。

次に価格による解約率をふまえて実験を行った。図3,表1が実験結果となる。mailRate と troubleRate を用いた結果と比較すると製品の利用を継続した場合の初期解約、契約期間終了間際の解約行動を捉えることができ、0.84と高い相関が得られた。

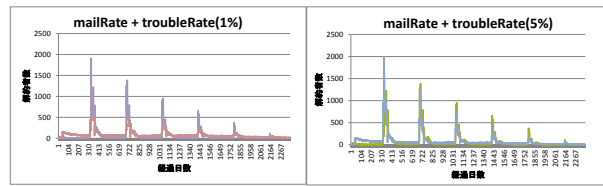


図2:シミュレーション結果と解約者の推移

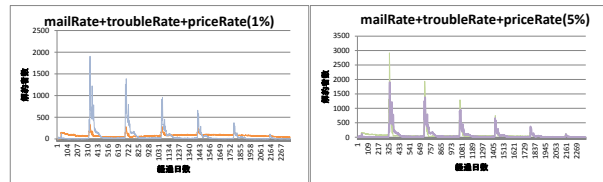


図3:シミュレーション結果と解約者の推移(priceRate)

表1:相関係数と平均二乗誤差

	解約者推移との相関係数	平均二乗誤差
troubleRate +mailRate(1%)	0.5636	21703.34
troubleRate +mailRate(5%)	0.7753	17366.99
troubleRate +priceRate +mailRate(1%)	0.7978	11359.21
troubleRate +priceRate +mailRate(5%)	0.8405	8354.80

また表1より、それぞれの実験の平均二乗誤差を比較しても、priceRate を用いた実験が低い値を示しており、解約予測に価格による解約確率を考慮することで、より実際の解約者推移に近づいていることがわかる。

6 結論

本研究では顧客解約行動を理解するためにエージェント・ベース・モデルを用いた解約行動のモデル化しシミュレーションを行った。シミュレーション結果は高い精度を得ることができ、さらに製品価格帯が解約行動を予測するのに有効であることが示された。

参考文献

- [1] Prasasti, N. Analysis Applicability of machine-learning techniques in predicting customer defection, IEEE, 157 – 162,2014.
- [2] Rand, W. Machine Learning Meets Agent based Modeling: When Not to Go to A Bar. Agent 2006, Chicago,