

変化分析型決定木を用いた製品・サービスの普及予測

Prediction of Product/Service Diffusion based on change analysis decision tree

藤原由希子*¹
Yukiko Fujiwara

富沢伸行*¹
Nobuyuki Tomizawa

井口浩人*¹
Hiroto Iguchi

*¹ NEC サービスプラットフォーム研究所
Service Platform Laboratories, NEC Corporation

Diffusion is the process by which a new product/service is accepted by the market. A new product/service would spread through society in an S curve, as the early adopters select the product first, followed by the majority. Previously, diffusion was predicted by curve fitting to the past diffusion curve. However, at the beginning of diffusion when diffusion curve is flat, the prediction is very inaccurate. To improve the accuracy, we propose the diffusion prediction method based on customer data. In this method, the impact of a product/service is estimated using the past diffusion and is used for the prediction for the future diffusion. This method was applied to the diffusion of iPod® to verify the method. The results show that the proposed method accurately predicted the diffusion curves. Also, the advancement of customer is important element that reflects the diffusion.

1. はじめに

顧客に製品・サービスを提供する場合、費用対効果や調達資源を知るため、製品・サービスの普及の推移を予測することが重要である。普及とは、製品・サービスが、時間的経過の中で情報伝達を通して人々に採用される過程であり、少数の革新的採用者(イノベーター)が購入する導入初期、一般の消費者(マジョリティ)が購入する時期、導入末期ごとに採用者数が変化し、採用者の割合の累積曲線は一般的に S 字曲線となる[Rogers 2003]。そのため、従来は、普及実績を S 字曲線へ適合することにより予測が行われてきた[Mahajan 1990]。しかし、導入初期段階では、将来の普及の推移を正確に予測できなかった。

一方で、近年、各企業は顧客の購買情報を顧客データとして大量に蓄積しつつあり、顧客の多様化する嗜好はデータマイニングを用いて分析される。例えば、協調フィルタリングでは、趣味の似た人からの意見を参考にして、推薦アイテムを選択する。しかし、採用までの時間的経過は予測されていなかった。

本稿では、顧客データを用いて普及の推移を予測する手法を提案する。提案法は、初期の普及実績から製品・サービスの影響度を推定し、将来の普及の推移を予測する。ここで、影響度とは、一定期間中に、製品・サービスの顧客が他の未採用者に製品・サービスの採用を喚起させる度合いと定義する。あるいは、一定期間において、製品・サービスの未採用者が他の顧客からの影響でその製品・サービスを受け入れる程度と言い換えることもできる。このような影響度を用いて、ある時刻のデータから決定木を作成し、次の時刻の予測をするように構成する。このような決定木を変化分析型決定木と呼ぶこととした。変化分析型決定木を用いた提案法は、携帯音楽機器 iPod の普及予測に適用し、有効性を確認した。

2. 提案法

普及におけるコミュニケーションは、参加者が互いに情報を創り出し共有していく過程であり、互いに似ている同類的な人の間で最も繁盛に行われ、採用行動を決定付けると言われている[Rogers 2003]。このコミュニケーション過程を再現するため、一

定期間ごとに分類器を生成する。属性は、年齢、性別、他の製品・サービスの利用履歴、先進性などであり、クラスは採用、未採用の 2 種類である。そして、再び全員のクラスを予測し、未採用者が採用と予測されれば、コミュニケーションの結果、採用に変化したとみなす。分類器には、今回は決定木を用いた。

通常の機械学習では、採用者は採用、未採用者は未採用と概ね予測され、採用者数は増加しない。そこで、採用者への重い加重により学習を行うこととした。採用者数を M 、未採用者数を N 、影響度を I とするとき、購買者への重み W は、以下の式で表す。

$$W = I \times N / M$$

すなわち、採用者への重み W という加重により、決定木に与える採用者と未採用者の比率は、一定の $I:1$ となるようにした。

2.1 普及実績に基づく影響度推定

影響度 I は、普及実績から推定する。今回は、一定期間を 3 ヶ月とし、直前の 6 ヶ月で影響度 I を推定した。現時刻の 6 ヶ月前の採用・未採用者の情報から影響度を変化させながら決定木を生成、新たな採用者を予測し、3 ヶ月前の採用者数との誤差が最も小さい場合の影響度を求めた。同様に、3 ヶ月前の採用・未採用者の情報から現時刻の採用者数との誤差が最小となる場合の影響度を求め、平均を影響度の推定値とした。

2.2 将来の予測

将来の普及は、推定された影響度を変化させずに予測した。まず、現時刻では、実際の採用者・未採用者を用いて分類器を生成した。以降は、未採用者のうち採用と予測された人を採用者とみなして普及予測を続行した。

3. 検証方法

3.1 データ

検証には、携帯音楽機器 iPod の普及の推移を用いた。顧客データは、Web アンケート調査を実施して収集した。アンケートは、2007 年 9 月 7 日に開始し、約 1 日間で 2,458 件を回収した。対象地域は全国、男女はおおよそ同数、年齢は iPod のターゲットと想定される 10 代から 40 代までとした。アンケートでの質問は、携帯電話やパソコンの利用履歴、先進性、メディアへの接触頻度などである。

連絡先: 藤原由希子, NEC サービスプラットフォーム研究所,
〒108-8557 東京都港区芝浦 2-11-5, TEL:03-5476-4387, FAX:03-5476-1083, y-fujiwara@db.jp.nec.com

携帯電話の利用履歴の質問(15種)は、持っているかどうかや機種、携帯電話の様々なサービスの利用頻度などである。パソコンの利用履歴の質問(7種)も同様である。なお、携帯電話やパソコンを持っていない場合、利用頻度は得られないため、属性値は欠損値とした。先進性は、デジタル機器一般に対して新しい製品・サービスの採用の相対的な早さや、他人の行動に影響を与える度合いである。先進性の質問(8種)については、例を以下に示す。

- 携帯電話やパソコンなどの製品やサービスについて人からよく聞かれるほうだ
- 周囲の人が新しい製品やサービスを利用していると気になるほうだ
- 自分の価値観に合わなければ周囲が全て利用していても、新製品を買ったり新サービスを利用したりしないほうだ

メディアへの接触頻度の質問(7種)は、新聞、テレビ、家族・友人との話、Webなどのそれぞれについて情報を得る頻度とした。年齢、性別を加え、これら39種を顧客データの属性とした。

普及実績もアンケートから求めた。iPod発売からアンケート実施時点までの6年間で採用者は582人、未採用者は1,876人であり、3,4,5,6年目の普及率は3.2%, 8.4%, 14.6%, 23.7%であった。

3.2 S字曲線による予測

比較のため、従来法であるS字曲線への適合による普及予測を行なった。予測には、統計分析ソフトSTATA(version 10)を用い、最大普及率Kを0から100の範囲で変化させ、誤差が最も小さくなるS字曲線を予測結果とした。

4. 実験結果

3,4,5,6年目までの情報を用いて将来の普及の推移を予測した結果を、図1の(A),(B),(C),(D)にそれぞれ示す。図1の(A),(B),(C)に示すように、従来法(点線)より提案法(実線)の予測結果の方が実績に近かった。例えば、従来法に比べ提案法は、3年目での普及率の急増や4年目での普及率の飽和を予測できた。また、顧客データから先進性の属性8種を除いた提案法(細い実線)は、予測精度が低下した。

5. 考察

提案法は、従来法より高い精度で予測することができ、予測結果は、概ねS字曲線となった。提案法は、3年目に普及率の急増が予測できており、その理由は、同じ影響度を用いているがインベータに比べマジョリティの人数が多いためと考えられる。提案法は、4年目には普及率の飽和を予測できており、その理由は、顧客の嗜好性を用いるためと思われる。実際、4年目に得られた決定木を分析したところ、例えば、「パソコンOSがWindowsで、パソコンでの音楽の利用頻度が月に1日以上で、パソコンメールの利用頻度が月に3日以下で年齢が48歳以上なら、採用」というルールが得られており、一般的なユーザは音楽への興味から採用したことが示唆された。一方、「パソコンOSがMacで新しいもの好きなら、採用」、「パソコンOSがわからなく、家族・友人から情報をよく得るなら、採用」、「パソコンOSがわからなく、家族・友人から情報をよく得るわけではなければ、未採用」などのルールも得られており、Macユーザは音楽に関係なく物珍しさから採用し、デジタル機器に興味が高い顧客は周囲の影響で採用を決めたことも示唆された。このような顧客の嗜好性を用いたことにより、早い時点で普及率の飽和が予測されたと思われる。

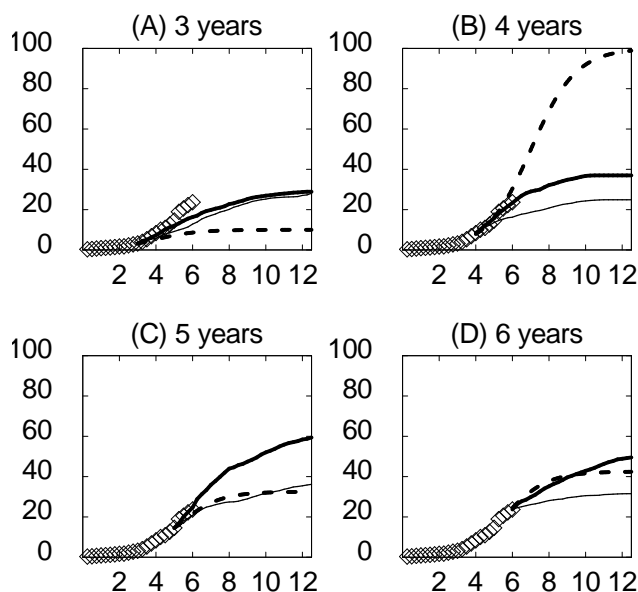


図1 3,4,5,6年目までの情報を用いて将来の普及の推移を予測した結果。x軸は時間(年),y軸は普及率(%)。は6年間の普及実績,実線は提案法の予測結果,点線は従来法の予測結果,細い実線は先進性をデータから除いた場合の提案法の予測結果

また、顧客データから先進性の属性を除くと、予測精度は低下した。ある製品・サービスを採用するまでの時間は、同じカテゴリの先進性と深い関係があることが示唆される。

提案法で予測された最大普及率は、4年目より5,6年目の時点の方が高かった。従来法により予測された最大普及率も、5年目より6年目までの情報を用いた方が10%程度高かった。5年目頃には、iPodの新製品nanoが発売されており、市場が活性化したのではないと思われる。このように、直前の普及実績に基づく提案法は、製品・サービスの時系列的变化を捉えることができる可能性がある。

今回、幅広い対象に短期間で調査できるという利点から顧客データとしてWebアンケートを用いたが、Webを利用できる人というサンプルの偏りがある。また、回答者が過去を振り返って製品についての時間的経過を測定しており、不正確である。発売時の調査の方がより望ましい。

6. まとめ

変化分析型決定木を用いた普及予測法を提案し、iPodの普及の推移に適用して手法の有効性を確認した。また、顧客の先進性が製品・サービスの普及に重要であることを確認した。現在、先進性は回答者の自己評価で抽出しているが、他の製品・サービスの利用履歴による抽出を目指したい。また、提案法は、同類性のある顧客間のコミュニケーションのみを用いている。異質性のあるコミュニケーションやマスメディアの影響を導入することは、今後の課題である。

参考文献

- [Rogers 2003] E. M. Rogers: Diffusion of Innovations, 5th Edition, New York, Free Press, 2003年
- [Mahajan 1990] V. Mahajan, E. Muller, F. M. Bass: New Product Diffusion Models in Marketing: A Review and Directions for Research, Journal of Marketing, 1990年