

オンライン・コミュニティの社会ネットワーク特性とイノベーションの普及パターン

濱岡 豊 (慶應義塾大学商学部)

hamaoka@fbc.keio.ac.jp

ネットワークの科学が注目されているが、それらの多くはネットワークの構造を分析したものである。これに対して本研究では、ネットワークの構造とパフォーマンス=普及パターンとの関係を分析した。オープンソース・ソフトウェア・プロジェクトにおいて開発されたソフトウェアの月次のダウンロード数に Bass モデルをあてはめ、革新者係数、模倣者係数、潜在市場を推定した。さらに、各プロジェクトでのフォーラムでのコミュニケーションから、社会ネットワーク変数を算出した。推定された Bass モデルのパラメタを社会ネットワーク変数および、プロジェクトの特性変数で説明したところ有意に説明された。

1. 研究の背景

音楽、画像、動画、ソフトウェアなどインターネット上で配布、販売される財が増加しているが、これについての研究は極めて限られている。一方、研究環境としても、コミュニティ・メンバーの発言数、メンバー間のコメント関係、ダウンロード数など、これまで入手が困難であった情報の入手が容易となっている。しかし、インターネット上での普及についての研究は行われていない。

2. 分析対象と方法

1) 分析対象

Sourceforge.net におけるオープンソース・ソフトウェア・プロジェクトからランダムに 2000 を抽出し、それらのうち、ソフトウェアを公開していること、月次データが 4 ヶ月以上あること、フォーラムにおいて 100 以上のメッセージが投稿されているという条件を満たす 236 プロジェクトを分析対象とした。

2) 普及パターンのパラメタ化

Sourceforge.net では、個人別のダウンロードの情報は入手できず、月次のダウンロード数が公開されている。これに Bass(1969)モデルをあてはめてパラメタ p, q, m を推定した。これは他者から影響を受けず採用する革新者と影響を受けて採用する模倣者が存在することを前提としている。

$$\frac{dY(t)}{dt} = p[m - Y(t)] + q \frac{Y(t)}{m} [m - Y(t)]$$

Y(t) 新製品導入後 t までの累積採用者数

t 新製品投入からの時間の長さ

m 潜在市場の大きさ

q 模倣係数 既に採用した人と接触して採用

p 革新係数 他者からの影響を受けず採用

なお、革新係数 p が大きいほど初期から採用者数が多いこと、模倣係数 q が大きいほど、急速に普及曲線が立ち上がることを意味している。

3) 社会ネットワーク指標の算出

各コミュニティへのメンバーシップについてのデータは入手不可能であるため、各プロジェクトのフォーラム(複数あるものについては合算した)への投稿メッセージからソシオマトリクスを構成し、下記の指標を算出した[Butts(2007)]。

・コミュニティの規模

・密度 density

・結合度 connectedness

・階層性 hierarchy

・効率性 efficiency

・Lubness

・多重性 multiplexity

なお、オンラインコミュニティのうち、フォーラムに投稿した者のみしか考慮されていないこと、また、約 3 年間分をプールしておりダイナミクスを無視していることに注意が必要である。

4) プロジェクトの特性

Sourceforge.net では、ソフトウェアの利用対象ユーザー、ソフトウェアのトピック、OS、ライセンスなどの情報が登録されている。これらの変数をダミー変数として定義して用いた。

5) 分析方法

Bass モデルのパラメタについては非線形最小二乗法を用いた。推定された p, q, m を被説明変数、

上述の社会ネットワーク変数およびプロジェクトの特性変数を説明変数としたステップワイズ回帰分析を行った。プロジェクトによって月次データのサンプル数や Bass モデルのあてはまりが異なるため、それぞれのパラメタを従属変数とする際には、推定されたパラメタの標準誤差の逆数を重みとした。

3.分析結果

推定結果を表に示す。3つのパラメタともにプロジェクトの特性変数および、社会ネットワーク変数によって有意に説明されていることがわかる。

1)プロジェクトの特性

Takada and Jain (1991)は、米日韓台での耐久財について Bass モデルをあてはめ、後発国ほど q が大きいことを示している。本研究でもプロジェクトの登録時期が模倣係数について正で有意となっており、彼らの知見と一致した結果となった。また、外部へのリンクとして Sourceforge.net の外部にホームページをもっているかというダミー変数も定義したが、これは革新係数、模倣係数について正となり、普及を早めることがわかる。一方、潜在市場規模については負で有意となった。これは外部のホームページでダウンロードをさせているためと考えられる。

2)社会ネットワーク変数

密度が革新者係数と正で有意となっている。Valente(1995)では、負の関係があることが示されており、逆の結果となった。本研究では普及プロセスで投稿されたメッセージをプールしたため、ホームページで活発なメッセージのやりとりを見た者が採用しているためと考えられる。ネットワークの冗長さの低さを表す効率性については、革新係数 p については正、模倣係数 q については負という直感的に納得できる結果が得られた。この他、社会ネットワーク変数もパラメタを有意に説明することが示された。

4.まとめ

本研究では、ネットワークの構造とパフォーマンスについて、オープンソース・ソフトウェアの普及という観点から分析を行い、社会ネットワーク構造が普

及のパターンというパフォーマンスに影響を与えることを明らかにした。

ただし、もっとも基本的な Bass モデルで推定を行ったこと、メッセージをプールしたため、ネットワークのダイナミズムを無視したという限界がある。今後、ページビューやソフトウェアの品質改善などの説明変数の導入、期間毎に社会ネットワーク変数を算出した動態分析を行う予定である。

表 推定結果(ステップワイズ回帰)

		log (p)	log (q)	log (m)
切片		-56.234 ***	-0.602	-12.03 ***
対象ユーザー	End user			0.35 **
	Developer		0.152 *	0.20
	System Administrator	-1.342 ***		-0.56 ***
ライセンス	GPL			0.33 **
	BSD	-2.040 ***	-0.447 **	-0.71 ***
	Independent	-0.774 ***		0.54 ***
	Linux		0.189 **	
プログラム言語	C			0.66 ***
ソフトウェア	Software	0.759 **		0.31
	Multimedia			-0.73 ***
トピック	Desktop			1.52 ***
	Database	-0.912 *		1.43 ***
	Text processing	-1.009 **		
	Internet	0.882 ***		
	Security			1.83 *
	Science			1.24 ***
	System	1.285 ***		1.41 ***
	Office		-0.395 **	
	SForge外でのホームページ開設	0.469 *	0.215 **	-0.22 *
	プロジェクトの登録年月日		4.3E-05 ***	
社会ネットワーク	log(密度)	44.347 ***		17.47 ***
	log(結合度)	-1.464		-4.32 ***
	log(コミュニティの規模)			0.48 ***
	log(階層度)			
	log(効率性)	70.136 ***	-2.780 *	23.72 ***
	log(Lubness)	1.940 ***		
	log(多重性)	5.110 ***	-0.669 *	5.61 ***
AdjR2	0.96	0.24	0.81	

注)***:1%水準で有意 ** :5%水準で有意 * :10%水準で有意

空欄はステップワイズの過程で除外された変数。

参考文献

- Bass, Frank M. (1969), "A New Product Growth Model for Consumer Durables", *Management Science*, 15, pp.215-27.
- Butts, Carter T. (2007), "sna," in *Tools for Social Network Analysis*. (<http://cran.us.r-project.org/>)
- Takada, Hirokazu and Dipak Jain (1991), "Cross-national analysis of diffusion of consumer durable goods in Pacific Rim countries," *Journal of Marketing*, 55 (2), 48.
- Valente, Thomas (1995), *Network Models of the Diffusion of Innovations*: Hampton Press.