

戦略ワード抽出手法と財務特異状況抽出手法の統合による企業情報マイニングの提案

Proposal of Enterprise Information Mining by Integrating Method to Extract Strategy-Words and Method to Extract Characteristic Corporate Financial Conditions

峯田誠也 (東京都市大学大学院) 岡田公治 (東京都市大学)

1.はじめに

近年、コンピュータの性能向上によりデータマイニングの研究が盛んになっている。特にマーケティングの分野での研究が多く行われているが、企業の公開している情報（以下IR情報）に対する研究は少なく、膨大なIR情報の中から何か知見を発見しようとしている研究は見当たらない。そこで過去、筆者らが開発したIR情報から戦略ワードを抽出する手法(2章)と他企業とは異なる特徴的な推移をしている財務指標を抽出する手法(3章)を統合し、膨大な企業情報から有益な情報の抽出を行う**企業情報マイニングのコンセプトを提案**する。

2.企業の戦略ワード抽出手法

階層化N-Gram手法

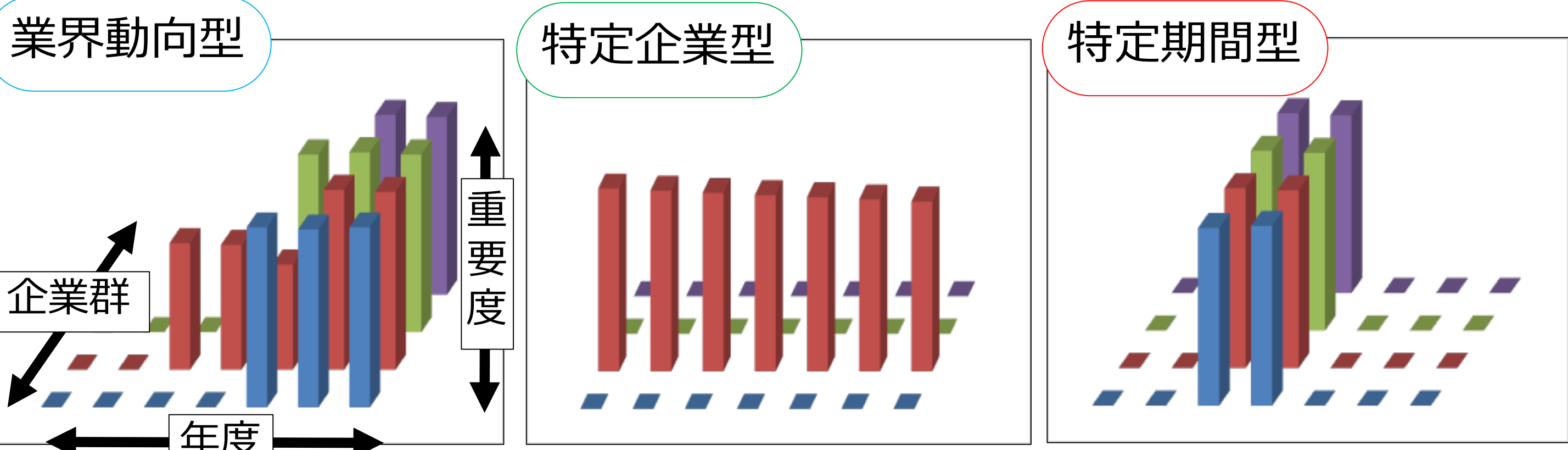
- 2以上N文字列の抽出
- N文字列長とN+1文字列長との包含関係と出現頻度を基に重要度を算出

グラム数	2-Gram	3-Gram	...	5-Gram
ワード	企業 業評 評価 価を を行 行う う上 上で で、 財	企業評 業評価 評価を 価を行 を行う 行う上 う上で 上で、 財、 財務	...	企業評価を 業評価を行 評価を行う 価を行う上 を行う上で 行う上で、 財務 で、財務状 況

戦略ワードの出現パターンの分類

同一業界複数企業、複数年度にわたる出現パターン
分類による不要なワードの削減

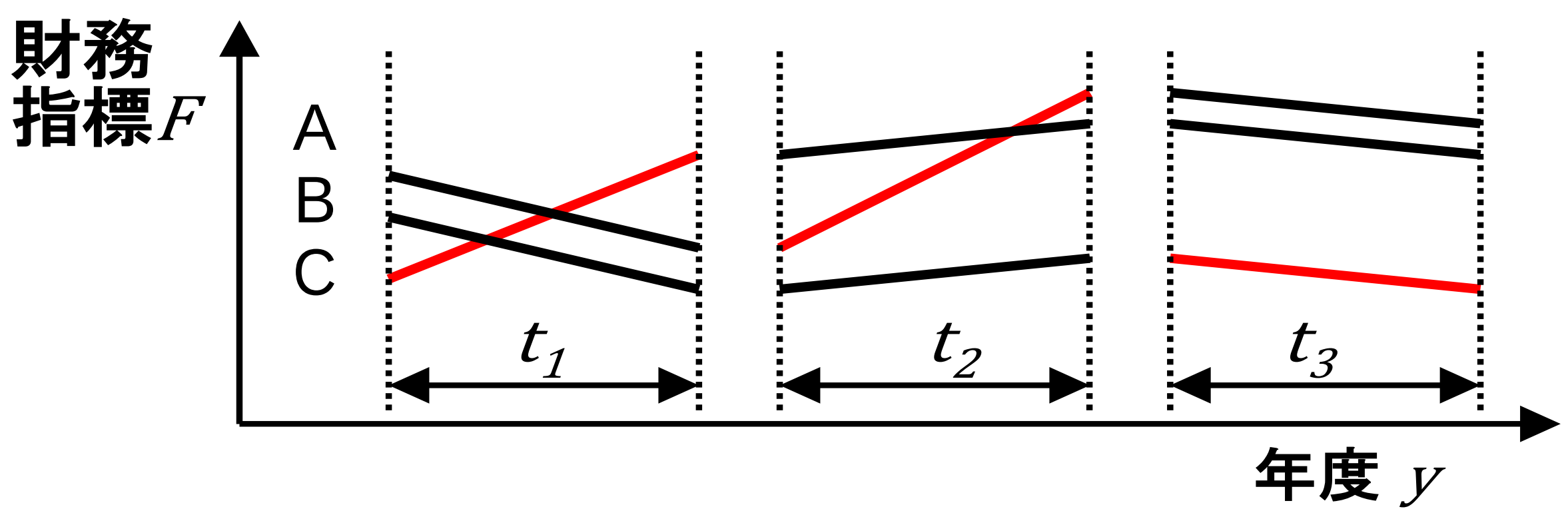
- 業界動向型戦略ワード（例 クラウド）
⇒ある時点から業界全体で使用
- 特定企業型戦略ワード（例 エネファーム）
⇒製品・サービス名等、特定の企業だけで使用
- 特定期間型戦略ワード（例 エコポイント）
⇒政策・災害対策等、特定の時期に業界内外で使用



3.企業の特異な財務指標の抽出手法

抽出する経営状態特異パターン

波形 (推移が逆方向) 振幅 (急激な変動) 水準 (異なる変動域)



- 企業の全ての組合せについて一対比較を行う
- その結果をまとめ、対象企業が他の企業群と異なる度合いを算出
- 異なる期間の特定のため、期間の長さで期間の開始年度を順に変えながら比較を行う

	一対比較法	統合法
波形の特異度	COS類似度	相加平均
振幅の特異度	分散比	対数相加平均
水準の特異度	平均値の差	相加平均

5.おわりに

企業の膨大なテキストデータと財務データから特徴を抽出し統合することで企業の有益な情報を抽出できるという企業情報マイニングのコンセプトを提案した。また、電機業界の財務指標とアニュアルレポートを対象に分析を行った。その結果を、同様の時系列データにし、統合させることで財務的变化と戦略ワード変化の関係を抽出できることを示した。今後は、企業情報マイニングのコンセプトを基により有益な結果を目指す。

*戦略ワードとは：IR文書中に強調され繰り返し記述されるワードやフレーズ
例：「オムニ7」セブン &アイ・ホールディングス 「みんなと暮らす街を幸せにします」ローソン

4.適用事例

適用対象：電機業界8社

シャープ・ソニー・東芝・日本電気
日立製作所・富士通・三菱電機・パナソニック

財務分析結果

財務指標	特徴量	'96	'97	'98	'99	'00	'01	'02	'03	'04	'05	'06	'07	'08	'09	'10	'11	'12	'13	'14	'15
固定資産	形状	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
有形固定資産比率	形状	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
買入債務	水準:小	0.38	0.69	0.92	1.00	0.92	0.77	0.54	0.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
負債	水準:小	0.50	0.83	1.00	1.00	0.83	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
固定負債	水準:小	1.00	1.00	1.00	1.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
売上高営業損益率	水準:大	0.00	0.12	0.29	0.53	0.71	0.88	1.00	1.00	0.94	0.71	0.41	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
売上高営業C/F率	水準:大	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
売上高経常損益率	水準:大	0.00	0.00	0.00	0.00	0.00	0.00	0.67	1.00	1.00	1.00	0.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
負債比率	振幅:大	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
純資産比率	振幅:大	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
棚卸資産回転率	振幅:大	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.29	0.43	0.57	0.71	0.86	1.00	1.00	1.00	1.00	1.00	1.00
自己資本比率	振幅:大	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.50	0.83	1.00	1.00	1.00	1.00	1.00	0.67
固定長期適合率	振幅:大	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.15	0.15	0.23	0.31	0.62	0.85	0.92	1.00	0.92	0.77	0.62
売上高売上原価率	水準:大	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.10	0.14	0.19	0.24	0.33	0.52	0.71	0.86	0.95	1.00	0.90	0.76	0.52
固定長期適合率	水準:大	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.15	0.30	0.45	0.60	0.75	0.85	0.95	1.00	1.00	0.80	0.45
流動負債比率	水準:大	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.40	0.80	1.00	1.00	1.00	0.80
売上高総損益率	水準:小	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.10	0.14	0.19	0.24	0.33	0.52	0.71	0.86	0.95	1.00	0.90	0.76	0.52
売上高販管費率	水準:小	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.67	1.00	1.00	0.67

テキスト分析結果

(64グラムまで抽出)

ワード	'06	'07	'08	'09	'10	'11	'12	'13	'14	'15
資本剰余金	0.98	0.85	0.90	0.00	0.00	0.00	0.00	0.00		
有形固定資産	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00		
亀山工場	0.93	0.61	0.71	0.00	0.59	0.00	0.00	0.00		
「AQUOS」	1.00	0.99	0.55	0.00	0.00	0.00	0.00	0.00		
剰余金	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00		
オンリーワン	1.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00		
進化	0.64	0.90	0.60	0.00	0.00	0.95	0.00	0.00		
液晶技術	0.62	0.97	0.71	0.60	0.00	0.00	0.52	0.00		
ハイビジョン	1.00	1.00	0.98	0.00	0.00	1.00	0.00	0.00		
亀山第2工場	1.00	0.96	0.52	0.00	0.65	0.56	0.00	0.97		
競争の激化	0.00	0.00	0.00	0.64	0.00	0.00	0.78	0.89		
有利子負債	0.00	0.00	0.00	0.00	0.00	0.00	0.99	1.00		
再生と成長を実現	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.81		
内部監査部門	0.00	0.00	0.00	0.00	0.00	0.00	0.60	0.58		
財務体質の改善	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.92		
構造改革ステージ	0.00	0.00	0.00	0.00	0.00	0.00	0.93	1.00		
製造物責任に	0.00	0.00	0.00	0.00	0.00	0.00	0.60	0.58		
買付行為を	0.00	0.00	0.00	0.64	0.63	0.53	0.90	0.89		
ビジネスリスク	0.00	0.00	0.00	0.89	0.00	0.00	0.62	0.60		
成長に向けて	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.58		
マーケティング	0.00	0.00	0.00	0.89	0.00	0.00	1.00	0.60		
再成長ステージ	0.00	0.00	0.00	0.00	0.00	0.00	0.93	1.00		
買収防衛策	0.00	0.00	0.00	0.86	0.86	0.79	0.90	1.00		
内部統制システムの整備	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.64		
大量買付行為を行おうとする者	0.00	0.00	0.00	0.00	0.00	0.00	0.68	0.66		
黒字化	0.00	0.00	0.00	0.00	0.00	0.00	0.92	0.91		

形態素分析では抽出困難なフレーズも抽出

財務データの分析により得られた業績の変化点に対応して、アニュアルレポートに出現するワードの性質も変化している。