

数理学による問題解決

– IBMにおける取組と事例 –

日本アイ・ビー・エム株式会社 東京基礎研究所 所長
森本 典繁
2012年 4月13日

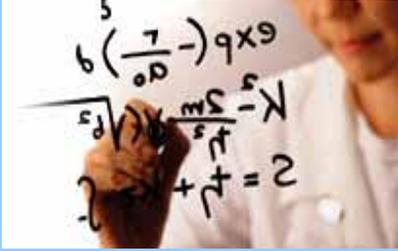
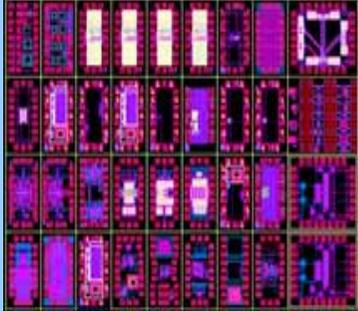


IBM Research (基礎研究部門)

世界7ヶ国に9つの研究所があり、約3,000名が基礎研究に従事しています。



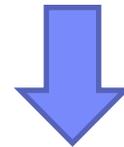
IBM基礎研究部門のストラテジー・エリア

<p>数理学</p> 	<p>インダストリー・ソリューション</p> 	<p>サービス</p> 
<p>ソフトウェア</p> <p>WebSphere</p> 	<p>システム</p> 	<p>テクノロジー</p> 
	<p>基礎科学研究</p>	

数理学の応用先は変わりつつあります

伝統的な応用エリア： 計算機の機能・性能向上

- 暗号理論
- コンパイラ
- 社内システムの性能改善



応用先のシフト

最近の傾向： お客様や社会の問題解決

- 社会的・ビジネス的な問題の解決
- Smarter City, Smarter Planet

システムにおける部分的適用から、問題そのものを解決する手段へ

機械学習という新しいデータ解析の技術が、新しい市場を創造しつつあります

書籍販売の推薦システム

- 商品数 × 顧客数 からなる巨大な行列の行列分解という操作に基づく
- これは商品と顧客のそれぞれを、少ない数の類型に分けることに対応する

検索と広告ビジネス

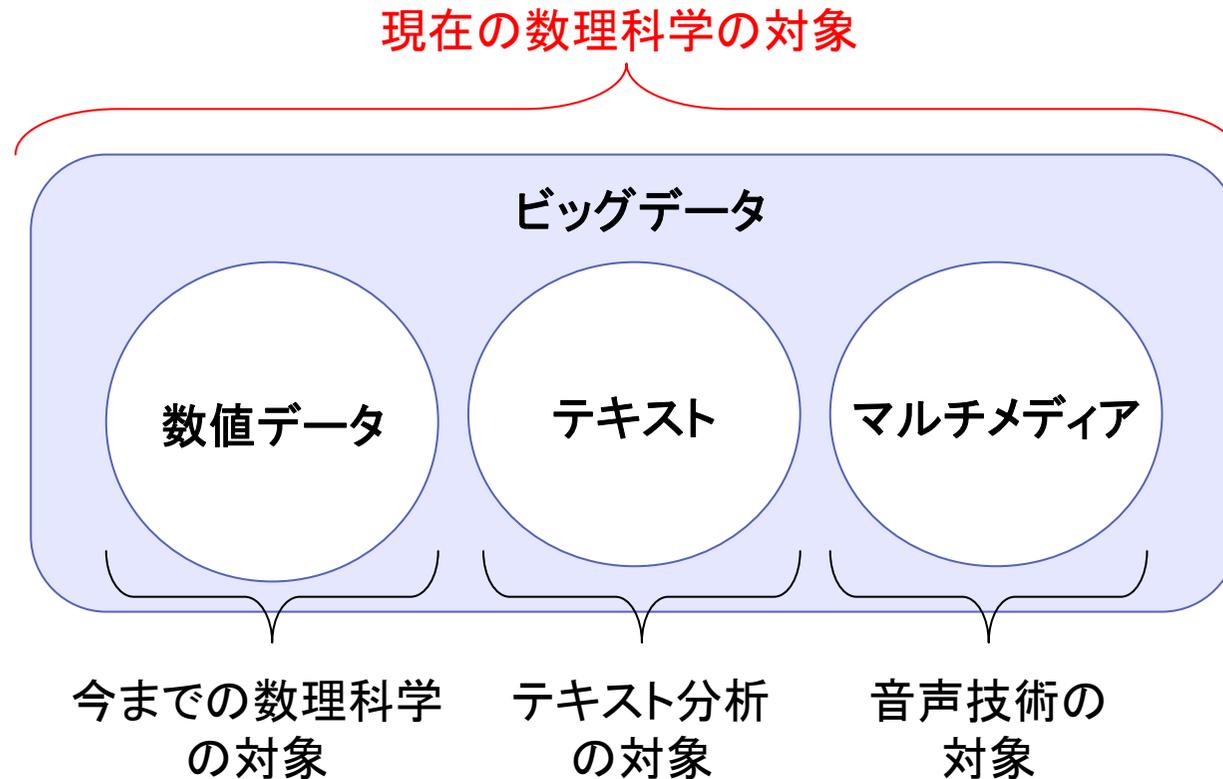
- 引用、被引用を表すネットワークの固有値分解という操作に基づく
- これは、多くのページから参照されるページは権威が高い、という直感に基づく

■この例から分かること

- 商品の推薦、というような、これまでは玄人の技だと思われていたものが自動化できるようになってきた。つまり、Analytics技術が新しい市場を創造した
- ビジネスモデルからしてBig Dataの処理が必然になってきた

自然言語や音声データも数理学の分析対象です

- テキストやマルチメディアデータの特徴を発見し、データの傾向をマクロに把握することが問題解決に役立ちます
- テキスト分析や音声技術と組み合わせることで、新しいサービスが創出されます



数理学の応用分野と事例

ロジスティクス

配送経路最適化

物流ネットワーク設計

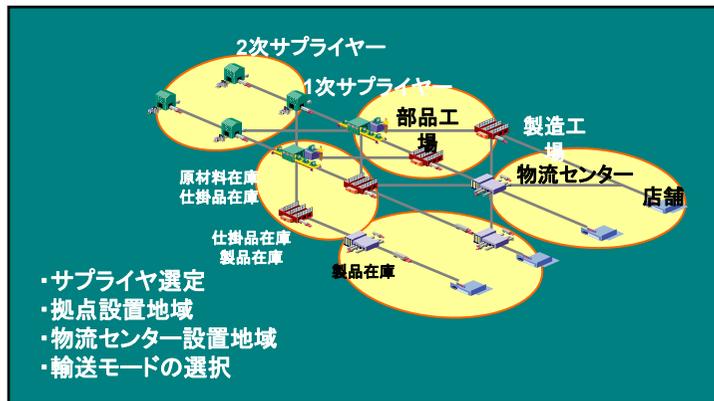
拠点最適配置

製造

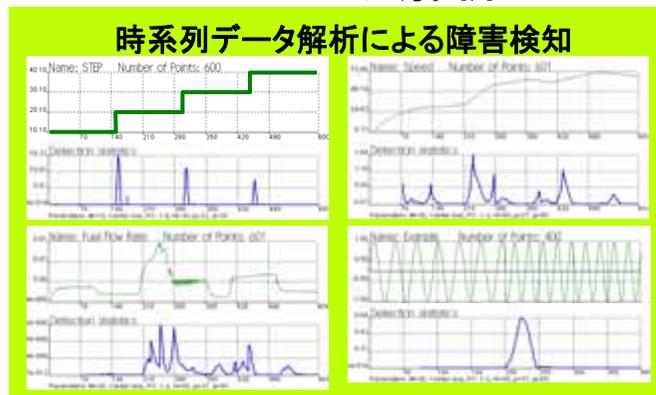
生産スケジューリング

板取り問題

SCM



センサー・データ解析



不正検知

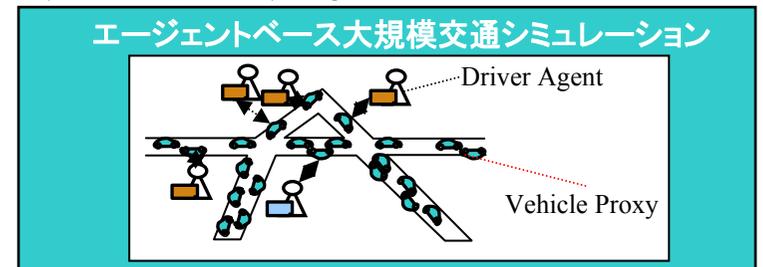


CRM

カスタマーターゲットティング

キャンペーン最適化

シミュレーション

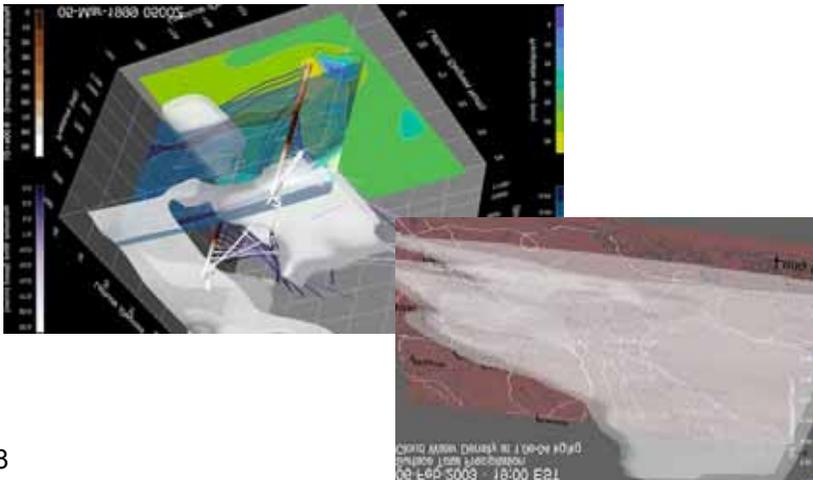


天気予測ソリューション Deep Thunder

- さまざまな業務への活用に向けて開発された、地域メッシュの細かい、比較的短い時間軸での、モデリングに基づく気象予測システム
- 活用性の高い気象予測システム
 - 有益な気象予測データを、さまざまなアプリケーションで必要なときにタイムリーに使用できる形で生成
 - 国立測候所(NWS)などの利用可能な公開情報のデータを活用
 - 柔軟な気象ビジュアル化と、データ配布機能で、様々なカスタマイズが可能

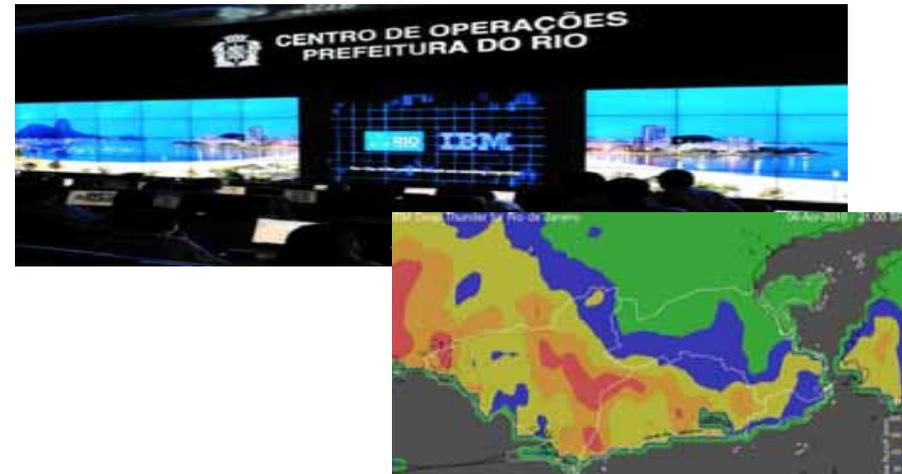
オリンピックでの利用

- 1996年のアトランタオリンピックで使用され、開催期間中、局地数値天気予報を提供
- 北京オリンピックの天気予報システムに採用



Brazil Rioのシティ・コマンド・センター

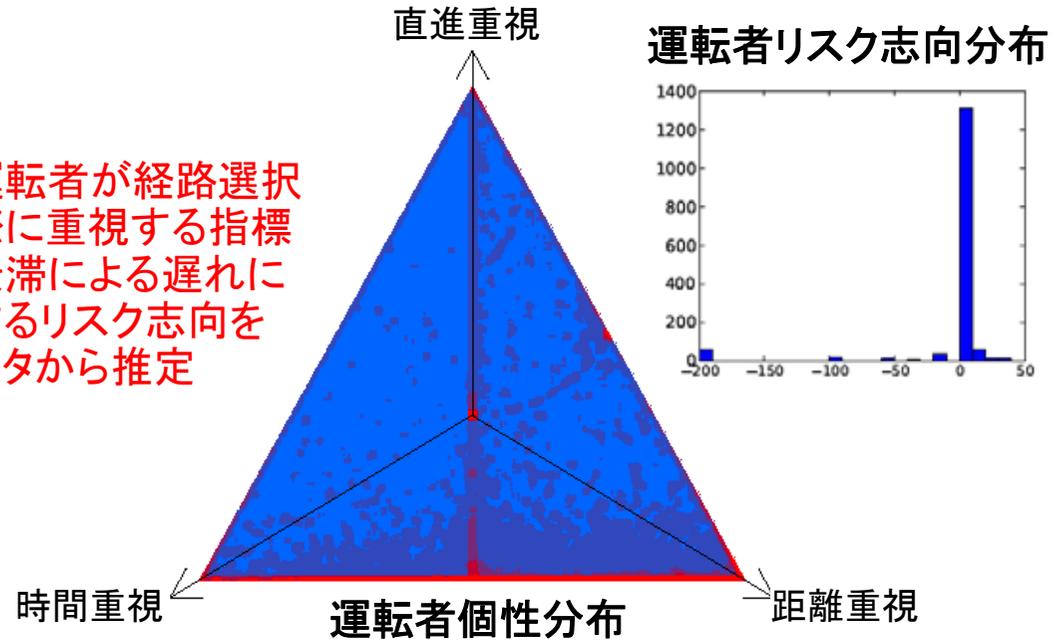
- 自然災害へ迅速に対応するために、48時間後までの天候を1kmおよび12時間ごとに予測
- 雲の動きから、大雨による表面流出量を予測



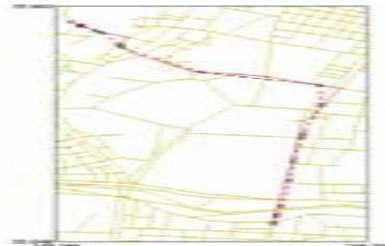
機械学習や最適化等の数理科学を駆使して、意思決定の科学に基づいて、プローブカーデータから運転者の個性を明らかにします



各運転者が経路選択の際に重視する指標や渋滞による遅れに対するリスク志向をデータから推定

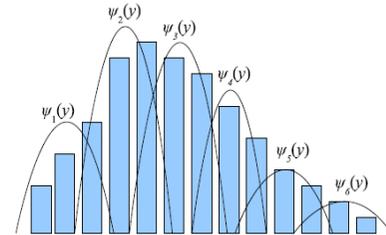


マップマッチング技術



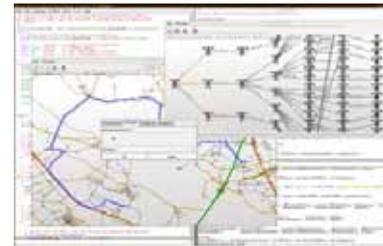
誤差のあるGPS点列から運転経路を復元

旅行時間分布推定技術



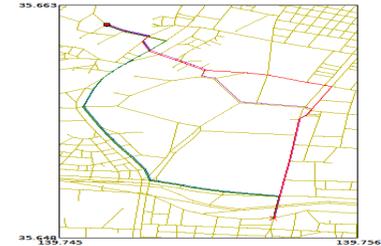
無数の道路の所要時間を限られたデータから推定

リスク考慮型経路選択技術



リスク志向に応じて最適な経路を選択

個性推定技術

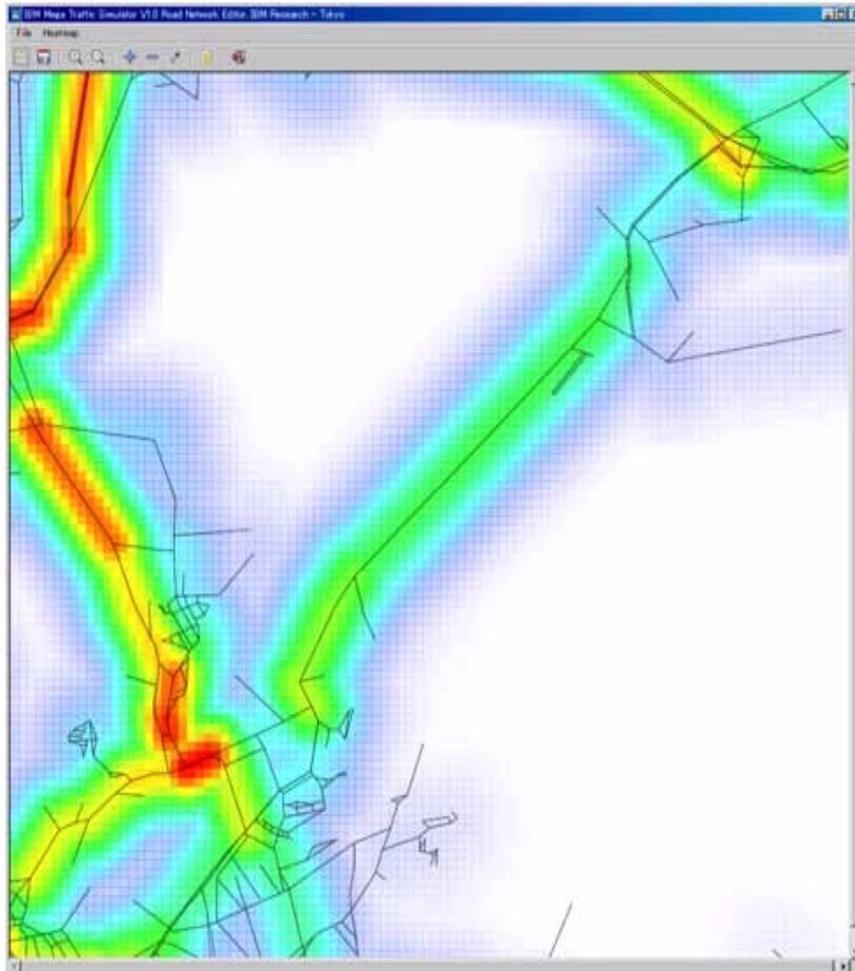


特定の選択をした個人の特性を推定

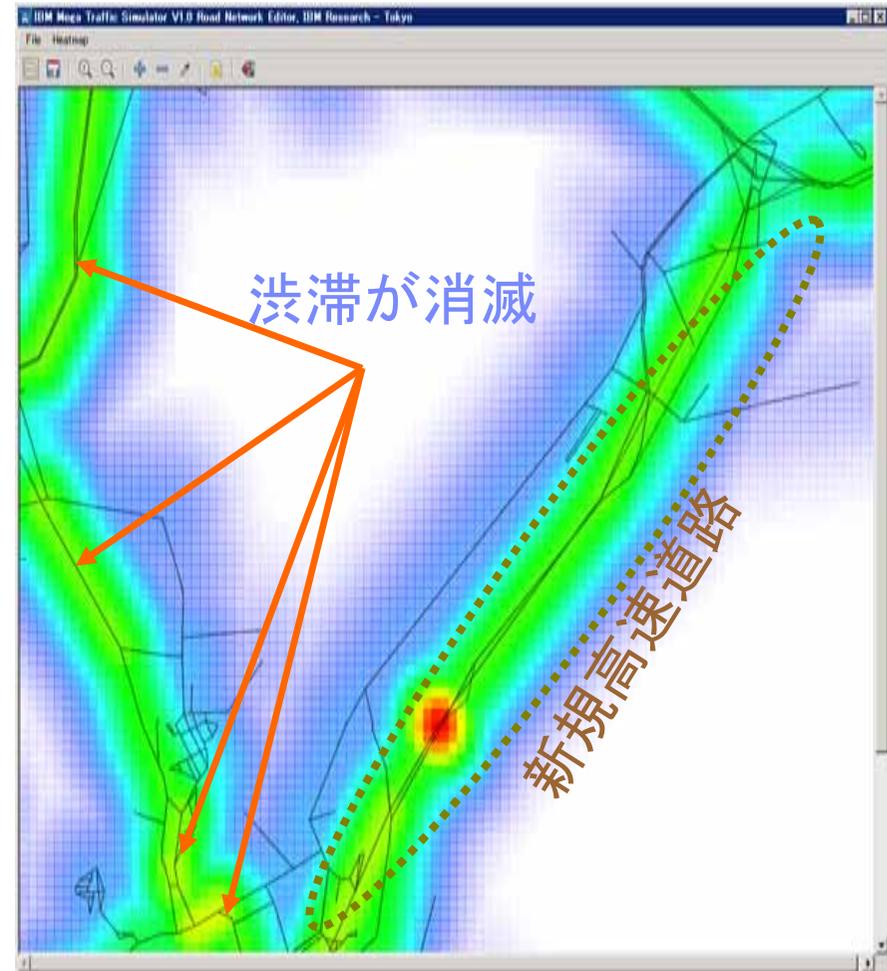
適用例: 広島市における新規高速道路の建設前後の渋滞状況をシミュレーション

新規高速道路をマップに追加し、シミュレーションを高速道路の建設前後の両方の地図に対して実行
 新規高速道路の建設によって渋滞が減ることを確認
 実時間の10倍以上の速度でシミュレーションが実行可能

高速道路建築前



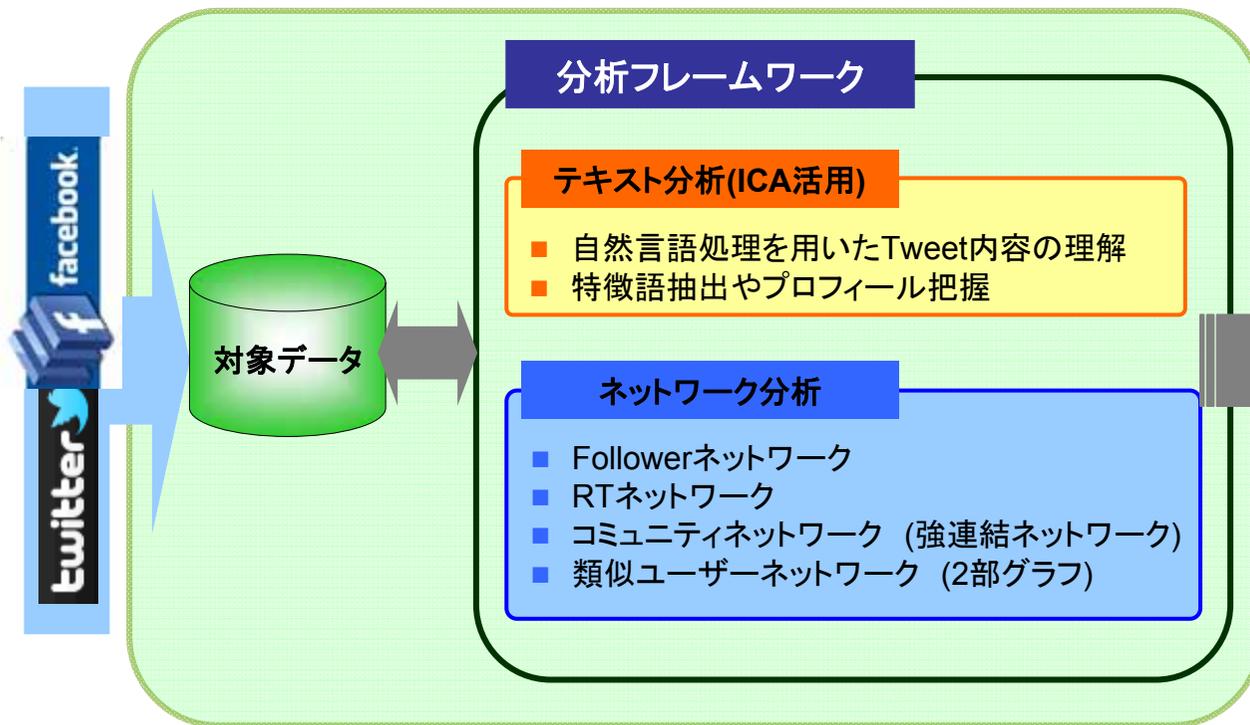
高速道路建築後の予測



テキスト分析とネットワーク分析を組み合わせたソーシャルメディア分析

ソーシャルネットワーク上での、情報の拡散や意見クラスタリングを可視化することにより、影響度や信頼性を知る手助けを行います

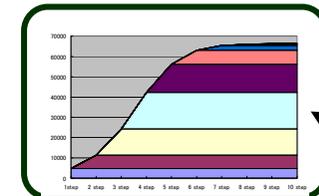
Text解析(ICA)とネットワーク解析の融合



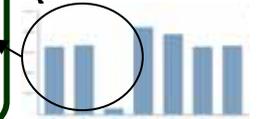
アプリケーション例

■ バースト検出

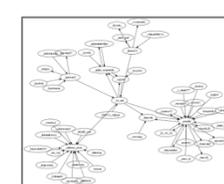
拡散度合い表示



時系列表示

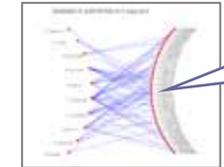


■ コミュニティとTopic検出



Tweetのテキストを解析

■ インフルエンサーの発見



ある話題における中心的な人物を特定

■ Topic遷移

未知のオブジェクト同士のつながりを関連するデータから予測します

- 多量の特徴量データと大規模な関係ネットワークデータを高速かつスケーラブルに解析できる機械学習技術を応用
- 多量の特徴量をコンパクトにメモリに保持し、解析しながらデータのノイズを除くことで、さらなるデータ解析の大規模化・高速化と予測精度の向上が可能

● 関係ネットワークと予測対象の関係(リンク)

実世界での人とコンテンツの関係

ウェブ・ブログスペースの場合:
 ノード: ユーザー, コンテンツ, 商品など
 関係: 聞く, 読む, コメント
 特徴量: ユーザーの年齢層, 性別など, コンテンツのテキスト内容など

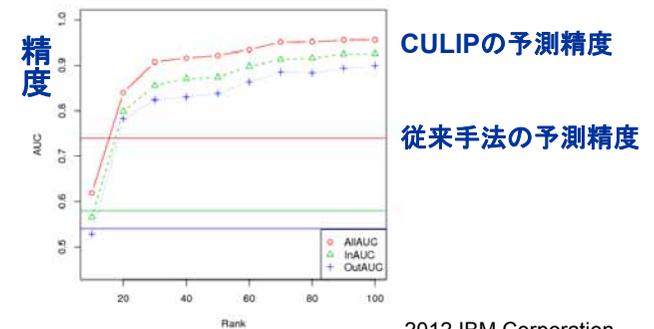
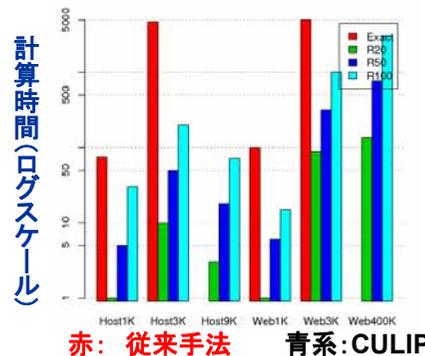
製薬会社の場合:
 ノード: たんぱく質, 遺伝子など
 関係: たんぱく質の反応促進の順序, 破壊的関係など
 特徴量: たんぱく質の構成と機能

たんぱく質A 類似 たんぱく質a
 たんぱく質B 類似 たんぱく質b
 たんぱく質C 異なる たんぱく質c
 たんぱく質D 異なる たんぱく質d

関係がある (類似)
 関係がない (異なる)

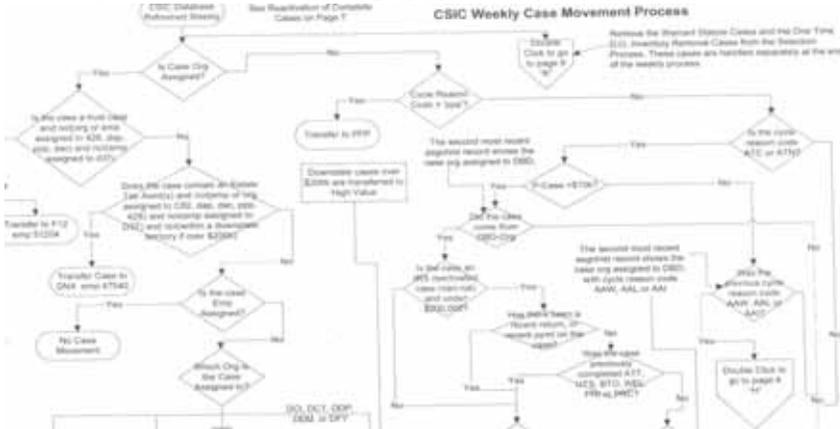
● テスト実験から40万以上のウェブ・ページ間のURLリンクを高速・高精度で予測を確認

ウェブページネットワークの場合
 ノード: ページ・広告
 関係: URLリンク
 特徴量: ページのテキストコンテンツ



TACOS (TAX COLlectionS optimizer)は、マルコフ決定過程に基づいて、各納税者に対して取るべきアクションを推薦します

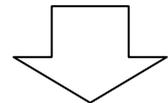
以前の徴収プロセスは非常に複雑で
リソースの有効活用が困難



徴税プロセスの課題

- 特定の納税者から徴収可能か？
 - 徴収のためにどのアクションをとるべきか？
 - いつアクションをとるべきか？
 - 誰が徴収するべきか？
 - 様々な制約下で収入を最大にするには？

TACOS (TAX COLlectionS optimizer)
マルコフ決定過程に基づく徴収アクションの最適化



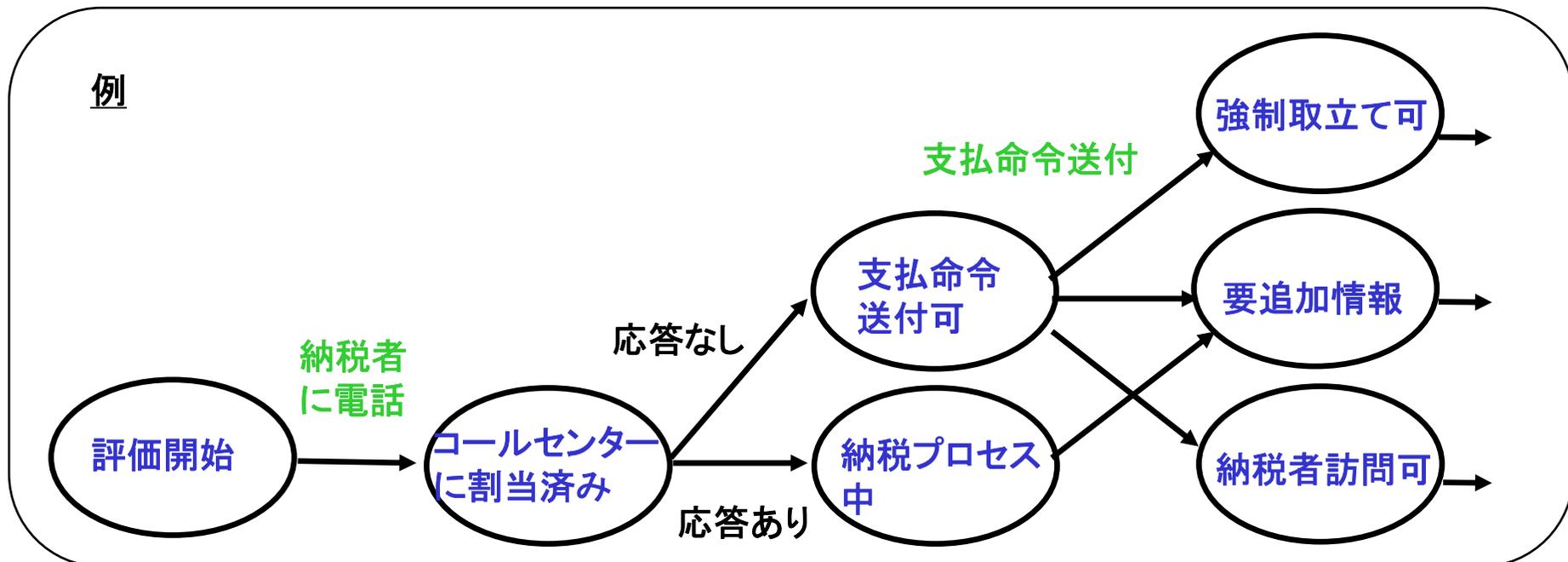
NY州政府は、3年間で1億ドルもの巨額の追加税収を得る見込み*

*プレスリリース: <http://www-03.ibm.com/press/us/en/pressrelease/34304.wss>

徴税支援のためのマルコフ決定過程



- **状態**: 徴収プロセスにおける当該納税者のステージに関する情報（徴収アクションの履歴、納税履歴など）をまとめたもの
- **アクション**: 徴収アクション
- **報酬**: 当該納税者から徴収された税



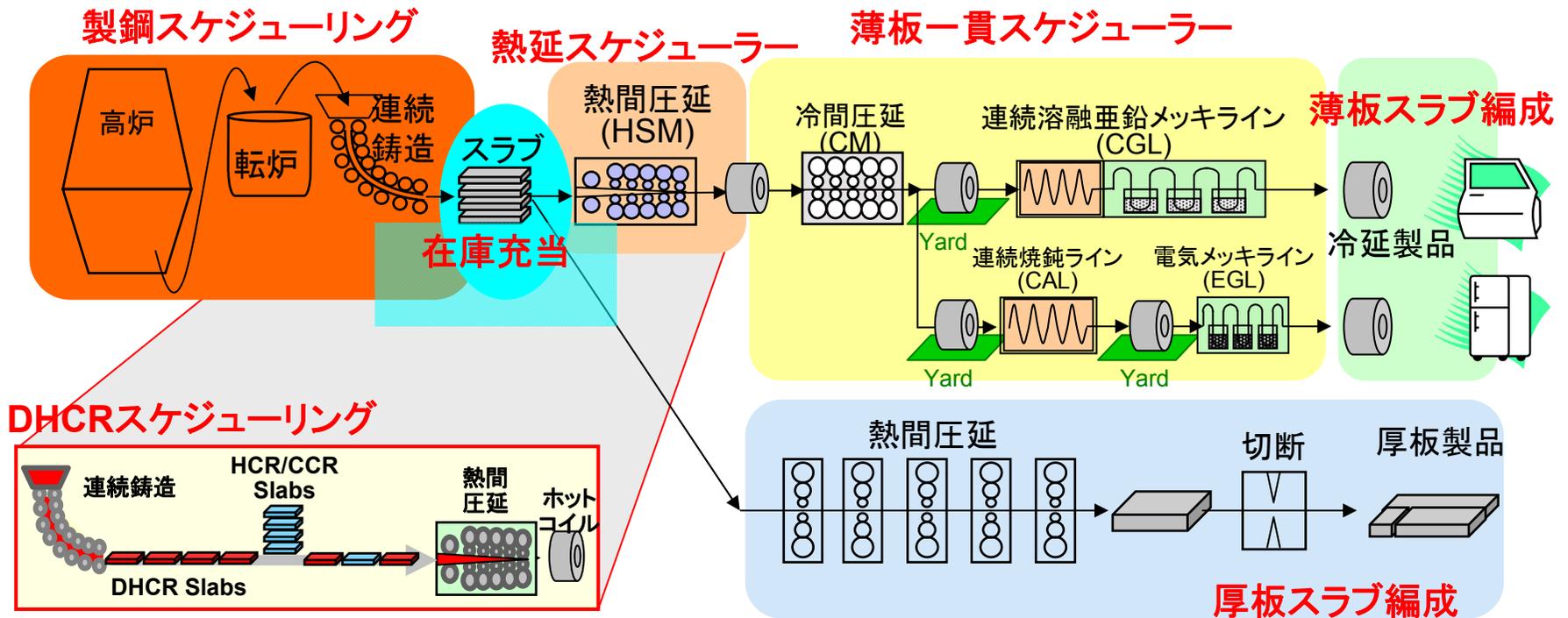
中長期の効果の例:

支払命令送付によって直接税が徴収できることはほとんどないが、将来の強制取立てや差し押さえを可能にする。

PDOS: 鉄鋼業向けにカスタマイズされた最適化アルゴリズム・ライブラリー

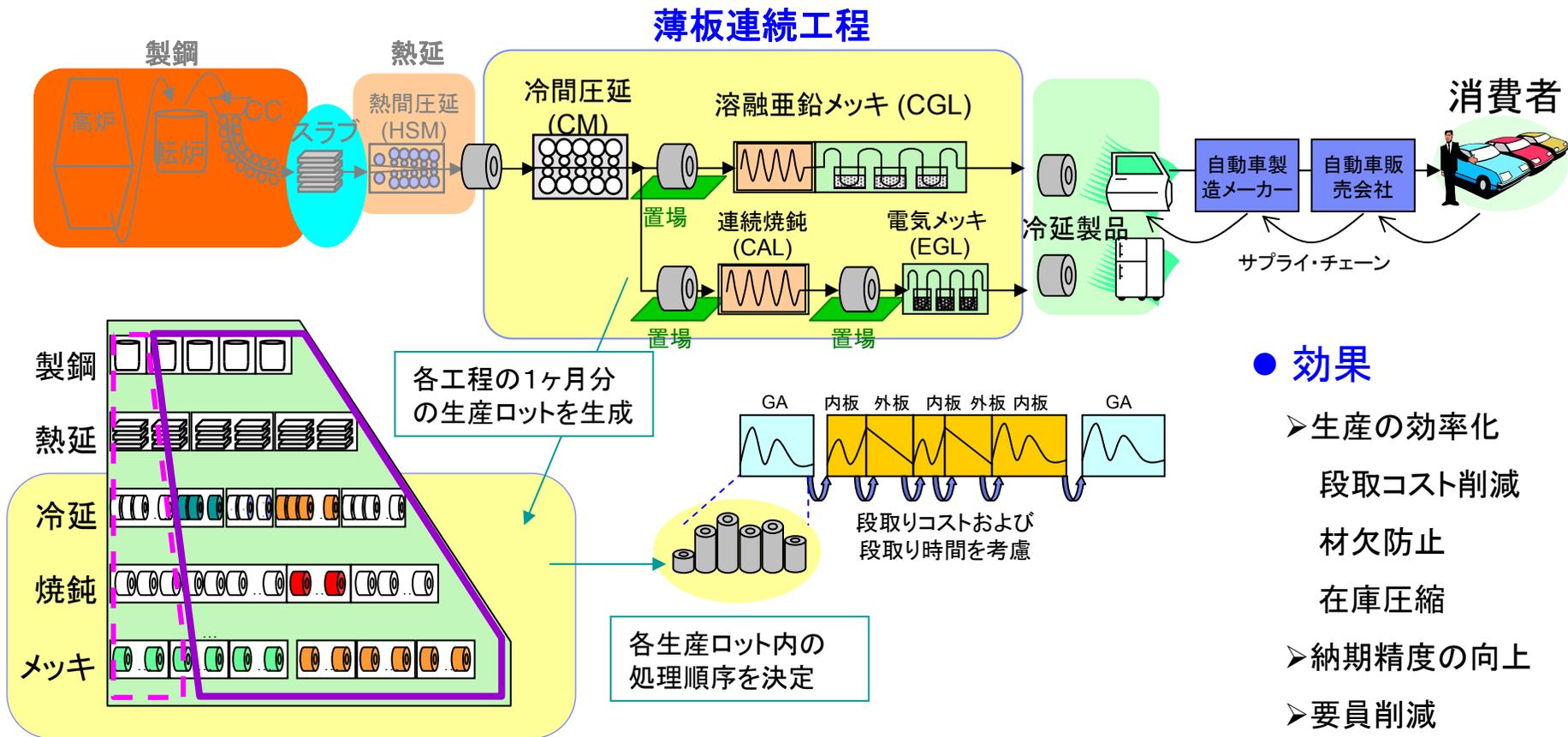
日本の先進的な鉄鋼メーカーをはじめ海外でも実績を上げているソリューションです。

- 在庫充当・材料引当
- 薄板スラブ編成
- 厚板スラブ編成
- キャスト編成／製鋼スケジューリング
- 熱延スケジューラー
- 薄板一貫スケジューラー



PDOS薄板一貫スケジューラー

- 薄板連続工程の1ヶ月分の生産スケジュールを生成
 - これまで専門家による手作業で行っていたスケジューリングを自動化
 - 生産ロット順序の制約、コイルのつながりに関する制約（幅・厚・温度...）などを考慮



● 効果

- 生産の効率化
 - 段取コスト削減
 - 材欠防止
 - 在庫圧縮
- 納期精度の向上
- 要員削減

数理学の適用における今後の課題

適用における課題

1. 複雑な解析結果を実社会における意思決定に役立てるのが困難
2. 数理学と社会における問題の両方に精通した人材の不足
3. 解析結果の数学的な価値と実社会における価値の乖離

対策の方向性

- 対策1: 技術のパッケージ化とプラットフォームの構築(解析結果の可用性)
- 対策2: 専門のコンサルタントの養成、技術者との交流促進
- 対策3: 産業界とアカデミアの協業による実社会の問題解決

IBM