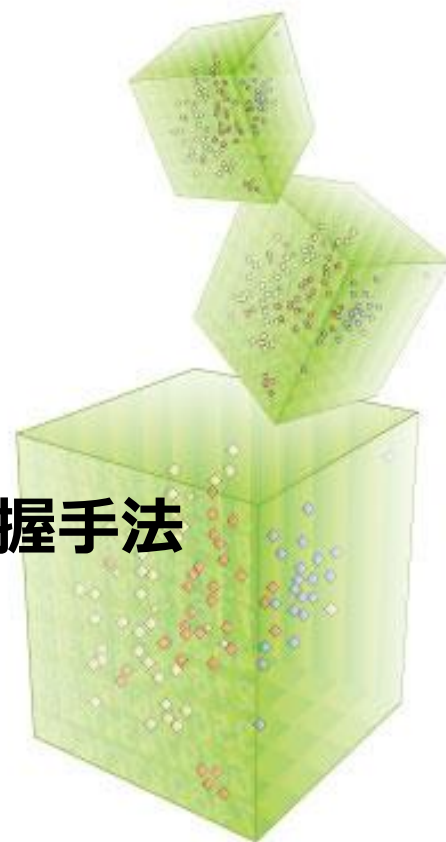




2013年9月4日
日本行動計量学会 第41回大会
ベイズアンネットワークを用いたデジタル広告
大量配信データからの各媒体貢献度の定量的把握手法
(アトリビューションアナリシス)

安達章浩 伊藤徹郎

アルベルト
株式会社ALBERT

〒151-0053 東京都渋谷区代々木2-22-17
TEL: 03-5333-3703 (代表) FAX: 03-5333-3723
www.albert2005.co.jp/

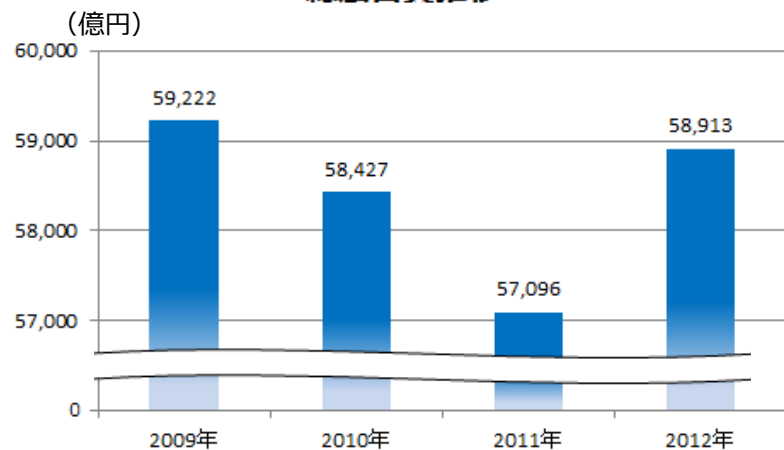
- 1、インターネット広告の歴史と背景
- 2、アトリビューション分析の経緯
- 3、ベイジアンネットワークとアトリビューション
- 4、実配信データを用いた検証結果

巻末、参考文献・参考サイト、
会社概要
創業の経緯

1、インターネット広告の歴史と背景

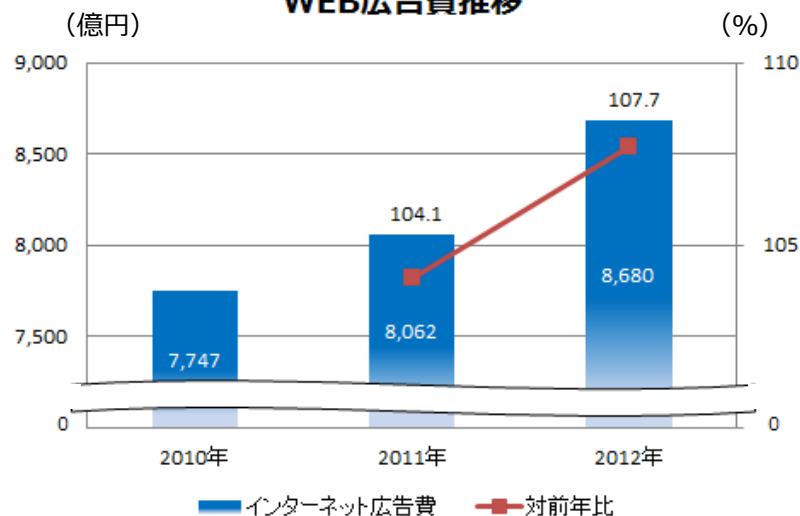
広告費の推移

総広告費推移

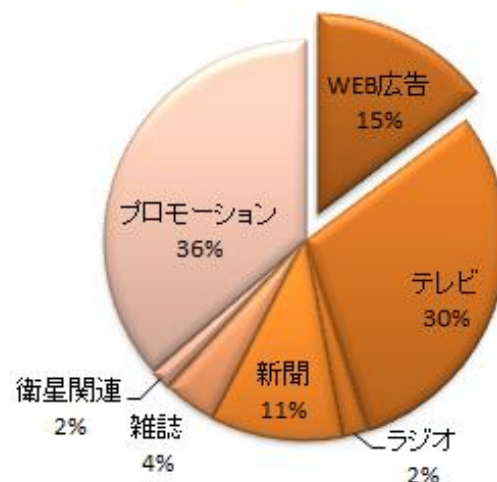


- ✓ 総広告費は2012年には震災前の水準に復活
- ✓ WEB広告費は2012年も対前年比108%と順調に推移、年々市場は拡大している。
- ✓ WEB広告のシェアもプロモーション広告(屋外、車内、チラシ、DM等)、テレビ広告について第3位の地位を確固たるものにしている。
- ✓ このような推移の源泉は配信法の多様化や技術的な進歩にある。

WEB広告費推移



2012年媒体別シェア



1996年～

バナー広告（純広告）の登場+ネット広告専門代理店の登場

1996年4月、商用検索サイト「Yahoo! JAPAN」がサービスを開始

1996年6月、「電通」と「ソフトバンク」の合併によるインターネット広告を専門に扱う広告代理店「サイバー・コミュニケーションズ」設立

1996年7月、「Yahoo! JAPAN」がバナー広告の取り扱いを開始

その後、「インフォシーク」、「goo」などの検索サイトを始め、朝日新聞、日本経済新聞などのWebサイトがサービスを開始

「メール広告」が登場したのもこの時期

2002年～

リスティング広告（検索連動型広告）の登場

Google Adwords 2002年9月に日本でサービスを開始

Overture（現Yahoo!リスティング）2002年11月から日本でサービスを開始

これらは従来の広告と全く異なる次世代広告として革命を起こし、「最も優れた広告」と言われている

「能動的で購買に直結、しかも少額予算から可能」という特徴が、多くの広告主にとって大きなメリットとなった

「アフィリエイト広告」が登場したのもこの時期

WEB広告の発展経緯（２）

2005年～

アドテク時代の到来

2005年 行動ターゲティング型広告の登場

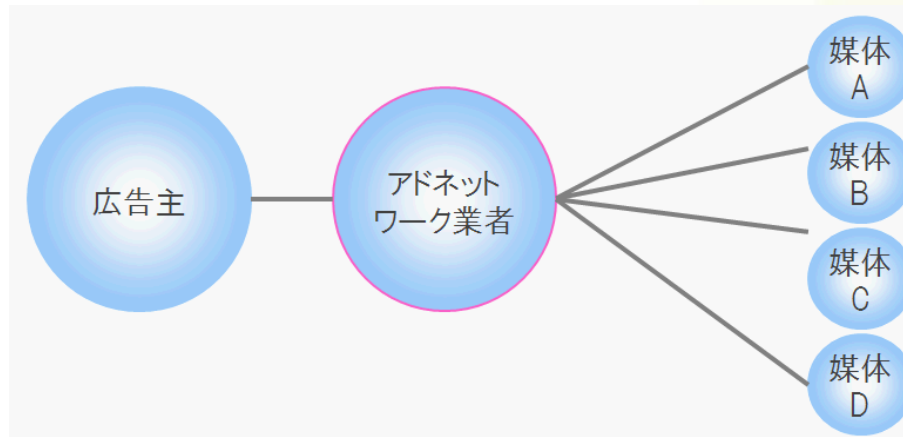
2008年 Yahoo!より興味関心連動型広告「インタレストマッチ」サービス開始

アドネットワーク広告（広い意味での第三者配信）の登場

広告媒体のWebサイトを多数集めて「広告配信ネットワーク」を形成し、多数のWebサイト上で広告を配信する広告配信手法。多くのWebサイトを媒体とすることで、全体では多くのトラフィック量を確保することが可能。

多種多様なジャンルの広告や広告媒体が混在しているため、広告配信効果を最適化する技術として、Cookieのデータをもとにユーザーの傾向を分析する行動ターゲティング広告（BTA）が導入されている場合が多い。

特定の広告について掲載しないようにしたり、ジャンルをあらかじめ限定したりするといった設定が可能。広告主はサイトのジャンルを指定することもできる。



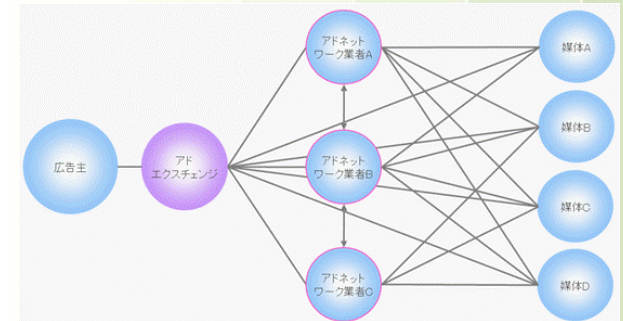
WEB広告の発展経緯（3）

2010年～

「どの単価で出稿するか」 + 「どんな人に広告を出すか」へ

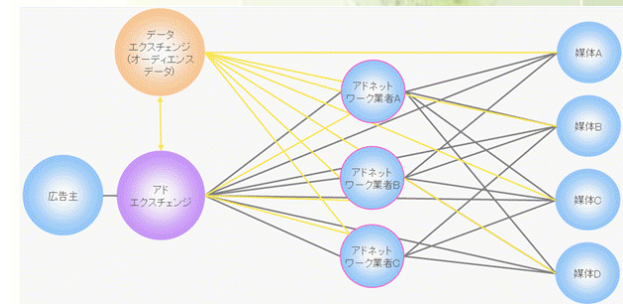
アドエクスチェンジの登場

各媒体・各アドネットワークが抱える広告枠（AD）を交換（EXCHANGE）できる仕組み（第三者配信アドサーバー）。アドエクスチェンジを介して配信プラットフォーム（仕様）が統一された⇒広告の取引市場化。



オーディエンスタグティングの登場

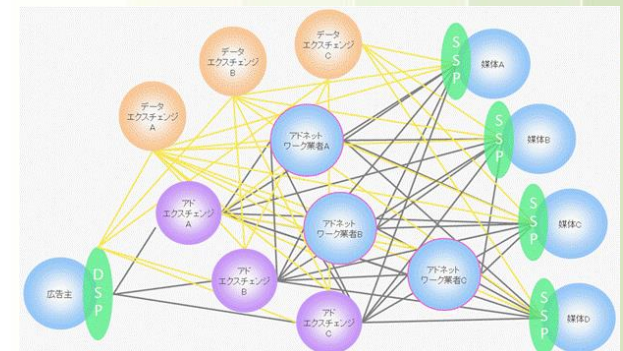
インターネット上の行動データ（Cookie）をもとにした、「人（オーディエンス）」ターゲティング手法。アメリカでは古くから存在していた「行動ターゲティング」をベースに、2007年頃からはデータソースを複雑に絡めたターゲティング手法により、市場の規模を徐々に拡大していった。



DSP (Demand-Side Platform) の登場

複数のアドエクスチェンジやネットワークを一元管理する広告配信の統合プラットフォーム。広告在庫の買い付け、広告配信、掲載面、オーディエンスタグティング等を一括して行うことができる。

複数のSSP (Supply-Side Platform) とRTB接続することで、広告主の求める在庫を必要な時に必要な分だけ確保することが可能。



WEB広告の多様化×相互作用の発生



各媒体個別の効果測定が困難に



各媒体への適正な予算配分も不明



合理的な媒体評価方法へのニーズが発生



アトリビューション分析への期待

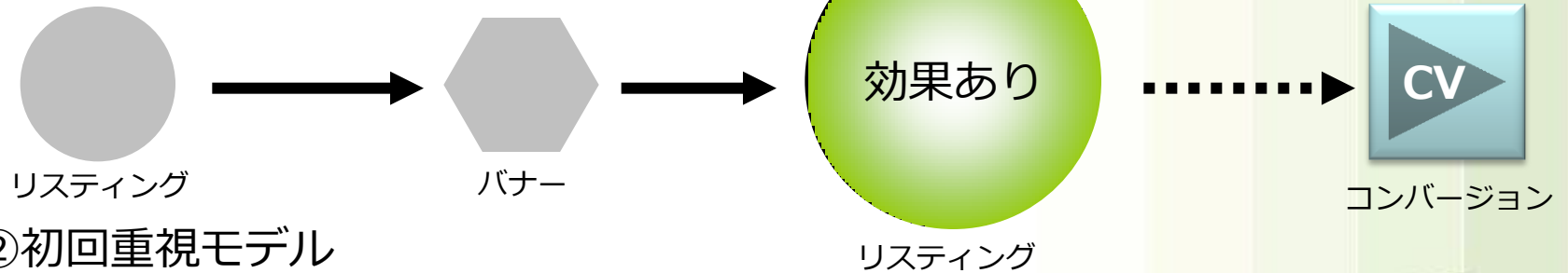
もともとアトリビューション分析は金融業界で発達。

ファンドの運用成績を評価・分析するパフォーマンス評価において、投資元本に対するリターンがどのような要因から発生したのかを明確にし、導き出された結果をベースに先々の運用方針を決定する分析手法を「アトリビューション分析」という。 [3]

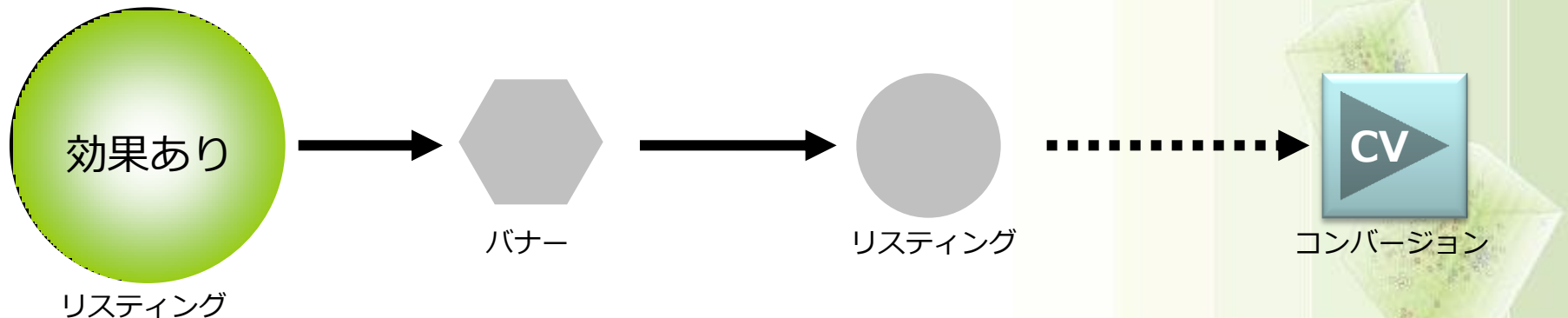
2、アトリビューション分析の経緯

従来のアトリビューション分析

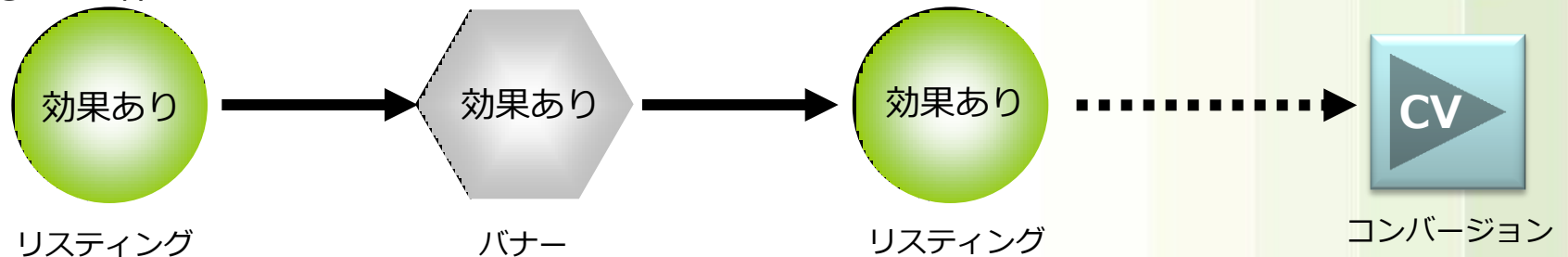
①ラスト重視モデル



②初回重視モデル



③均等配分モデル



貢献度の配分は仮説における定性的なものでしか語られていなかった。

3、ベイジアンネットワークとアトリビューション

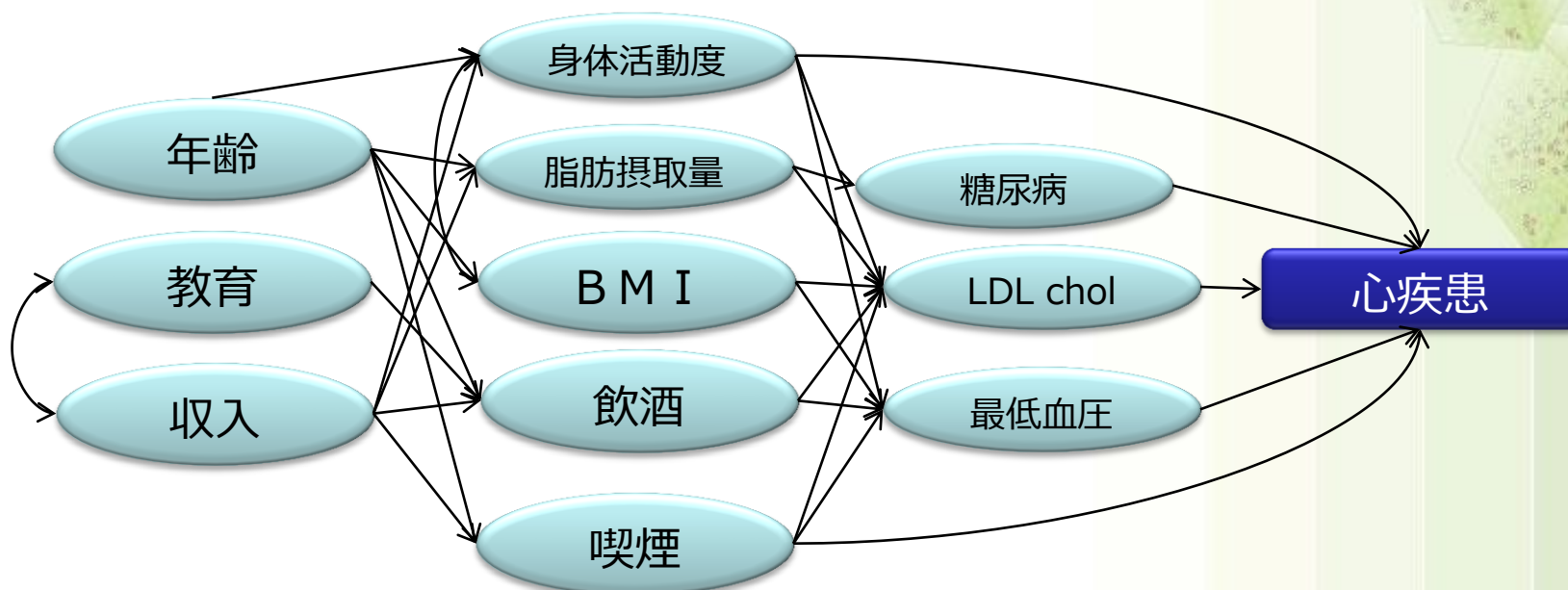
ベジアンネットワークとは

ベジアンネットワークとは…

複数の確率変数の間の定性的な依存関係を**グラフ構造**によって表し、
個々の変数間の定量的な関係を条件付確率で表した確率モデルのこと

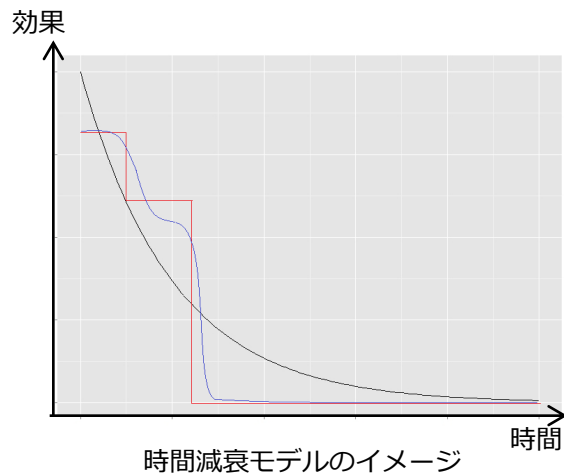
様々な分野で活用されている確率推論モデル

医療分野における心疾患のベジアンネットワークの例



従来のアトリビューション分析との違い

①コンバージョンへの時間効果を考慮している点



広告接触からの時間を考慮した際、その効用関数の真のモデルを知ることはできないが、仮説として左図の青線のような非線形なものであると考えられる。しかし黒線のような指数近似した場合、ロバスト性を担保することができないため、本研究においては、ロバスト性を担保するため赤線のような時間を区切り、その中での当てはめを行っている。

②未CVユーザーの考慮を行っている点

従来のアトリビューションでは、広告に接触したが、CVに至らなかったユーザーに関しては、考慮されていなかったが、本研究ではそうしたユーザーの行動もモデルに反映している。

③因果推論を用いて分析をしている点

ベイジアンネットワークは確率推論モデルであるが、従来行われてきたアトリビューション分析において、因果を推論できるような分析を行った事例はなかった。

ベイズアンネットワークの作成方法

ベイズアンネットワークを生成する方法として、下記の2つの方法があります。

- 1、条件付独立性検定および
グラフ理論に基づいて生成する方法
- 2、スコア（情報量規準など）に
基づいて生成する方法

本研究では2を採用しています。

条件付独立性とは

変数 a, b, c があるとする。 b と c が与えられたときに、 a の条件付分布が b の値に依存しない。

$$P(a|b, c) = p(a|c)$$



c が与えられたもとで、 a は b に対して条件付独立である。

C で条件づけられた a および b の同時分布を考える

$$\begin{aligned} P(a, b|c) &= p(a|b, c)p(b|c) \\ &= p(a|c)p(b|c) \end{aligned}$$



c が与えられたとき、 a および b が統計的に独立である $a \perp\!\!\!\perp b|c$

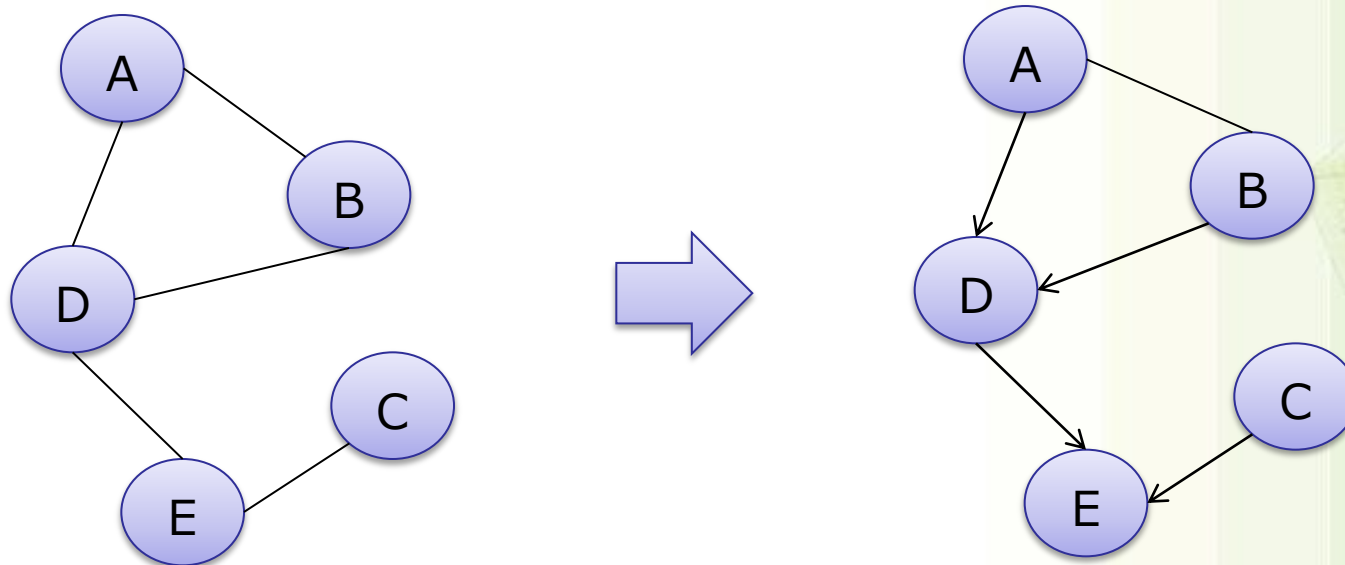


各ノード間において、条件付独立を考え、その独立性を検定します。
条件付独立であるとなった場合、そのノード間にエッジは引かれず、
あり得るすべての条件で条件付独立でない場合は、
従属であるとみなし、エッジを引きます。

条件付独立性検定およびグラフ理論に基づく生成方法

前述の条件付独立性検定によって、ノード間のネットワークが構築されました。この状態では、ネットワークの向きがない無向グラフのため、V字合流の性質を用いて、ネットワークの向きを決定します。

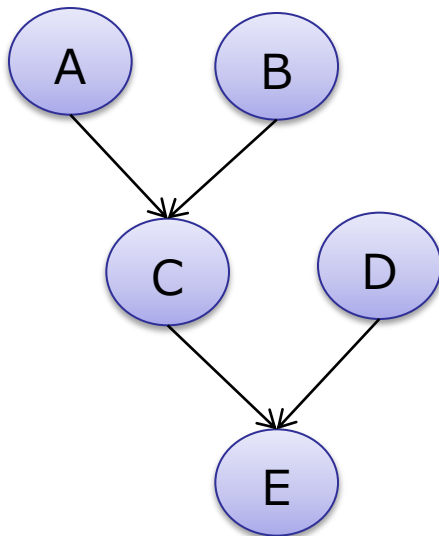
※V字合流の性質における説明は割愛します。



しかし、この方法では検定で誤ったネットワークが生成される危険があったり、一般的にすべてのエッジに向きがつけられない欠点が存在します。また、この方法はノイズに弱いことが知られています。そのため、本研究では別の方法を採用しています。

スコアに基づいて生成する方法

前述の方法では必ずしもグラフの向きを決定することができないので、
ベイジアンネットワークの情報量規準を用いて、ネットワークを生成します。



左図のネットワークの同時確率分布 $P(A,B,C,D,E)$ は
 $P(A,B,C,D,E) = P(A)P(B)P(D)P(C|A,B)P(E|C,D)$
と定義できます。

このネットワークの条件付確率表から情報量規準
を算出し、その最小値となるネットワークを
採用いたします。

$$AIC = -2 \log L(\tilde{p}; A, B, \dots, E) + 2d$$

$$BIC = -2 \log L(\tilde{p}; A, B, \dots, E) + d * \log(N)$$

他のモデルとの比較について

アトリビューション分析を行うにあたって、他の分析手法についても検討し、以下のような理由でベイジアンネットワークを用いることに決定しました。

- ①回帰モデル：各変数間の関係性はモデリングできるが、第3の変数の影響などを分析することができない。
- ②因子分析：変数間の関係性を線形に分析することができるが、非線形の共変主成分分析 関係をモデリングしたい場合、適切ではない。
- ③SEM：分析者が仮説として因果構造を与えるため、そもそもの仮説が誤っている場合の担保ができない点。
(共分散構造分析) また、目的変数にはガウス分布を仮定した上でモデル化するため、自由度が低い。

ベイジアンネットワーク

条件付確率表によって離散化し、個々の確率値を割り当てた不連続な確率分布によるモデル化が可能で、線形から非線形まで幅広く近似することが可能。
仮説ありきのネットワークではなく、観測値から因果関係を推測可能。

4、実配信データを用いた検証結果

アトリビューション分析で取り扱うデータ

今回アトリビューション分析を行うにあたって、取り扱ったデータは下記の通りです。

分析対象データ

某社の1か月のデジタル広告配信データ
(バナー広告、DSP広告、リスティング広告)

3,100万
レコード超

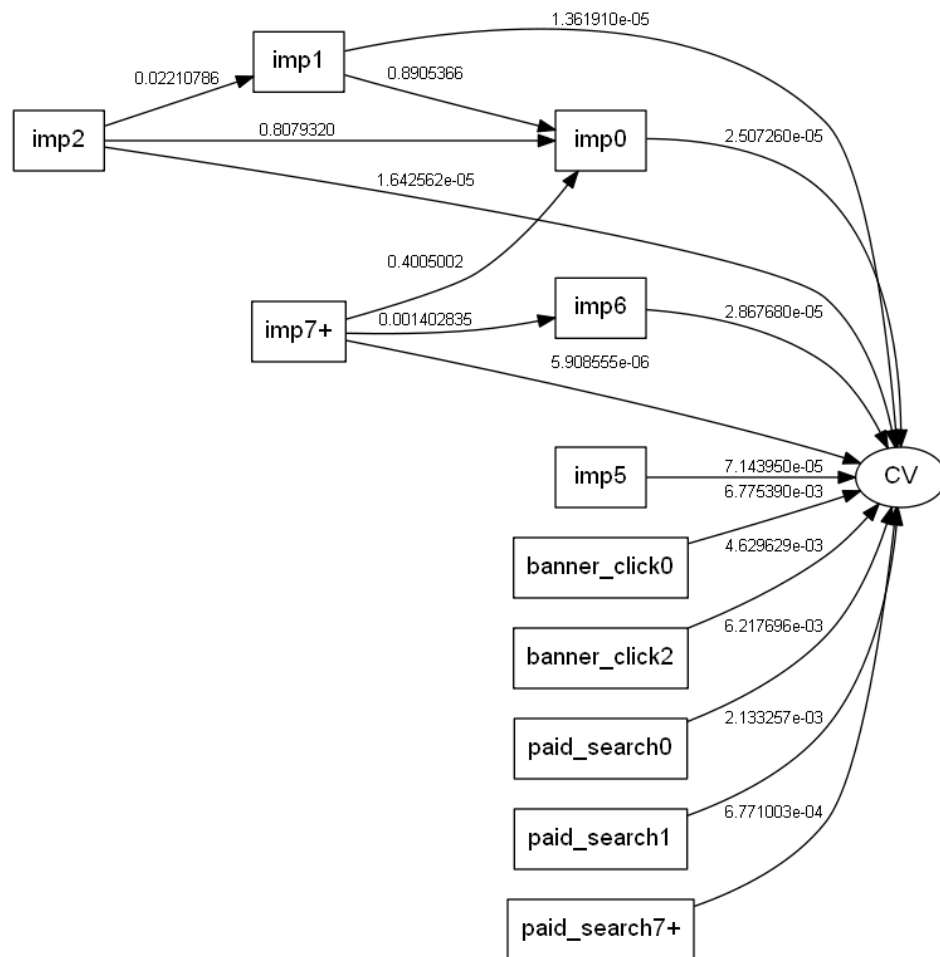
対象となるID数

400万ID超

上記のデータを分析用に加工し、ベイジアンネットワークにかけ、CVにかかる因果推論を行いました。

コンバージョンパス図の事例

ある広告商材を1か月配信した際のCVにおける各媒体と時間効果の影響をパス図にすると、下記のような結果となりました。



0 : cvから24時間以内

1 : cvから24～48時間以内

2 : cvから48～72時間以内

⋮

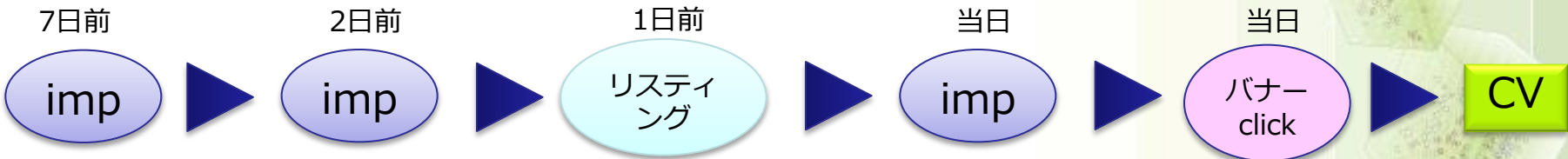
7+ : cvから168時間以上後

前述のパス図からアトリビューションスコアを割り振ります。
CVへのパス係数は下記の通りです。

b0	b2	p0	p1	p7+	i0	i1	i2	i5	i6	i7+
6.775390e-03	4.629629e-03	6.217696e-03	2.133257e-03	6.771003e-04	2.507260e-05	1.361910e-05	1.642562e-05	7.143950e-05	2.867680e-05	5.908555e-06

このままでは適用できないので、これを重みとしてアトリビューションスコアを算出します。

重みづけの例

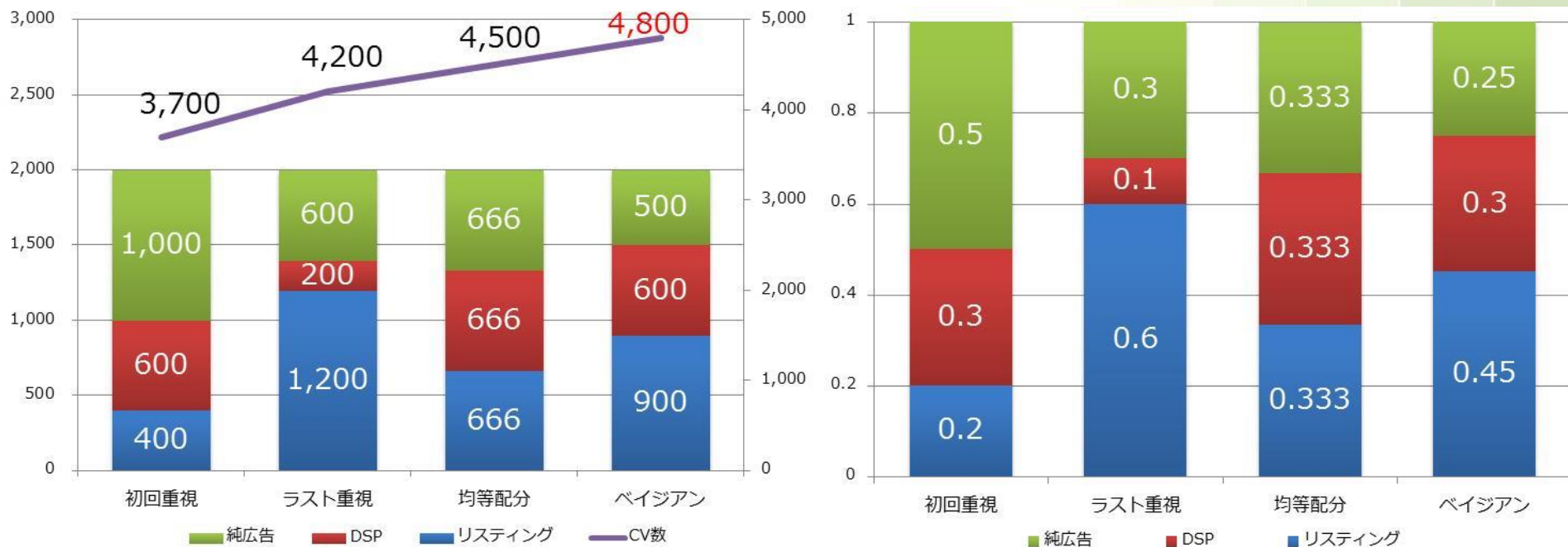


	i7+	i2	p1	i0	b0
頻度	1	1	1	1	1
重み×頻度	5.908555e-06	1.642562e-05	2.133257e-03	2.507260e-05	6.775390e-03
アトリビューションスコア	0.0006597275	0.001834024	0.2381916	0.002799514	0.7565151

※アトリビューションスコアは重みを足して1となるように正規化しています。

予算の最適化

前述のようにして、各広告媒体に対して、アトリビューションスコアを割り振ることで、適切な広告予算の配分が可能となり、同一予算での獲得数を最大化させることが可能となる。



従来の手法におけるアトリビューションの配分とベイジアンネットワークによるアトリビューションにおける予算配分のイメージ図

今後の課題について

今回の研究においては、各広告媒体での接触とその貢献度を明らかにした。

今回の研究でできなかった点

ディスプレイ広告とリスティング広告などの複数媒体における相互作用を明らかにすることができなかった。

今後の改善点について

本研究ではインターネット広告の配信データを元にアトリビューションを算出したが、現実のマーケティングにおいて無視することのできないマス広告のデータも加味し、企業のマーケティングゴールへの各施策の貢献度を明らかにしていく予定である。

- [1] http://www.dentsu.co.jp/books/ad_cost/2012/index.html
- [2] <http://dmlab.jp/web/history.html>
- [3] 「アトリビューション 広告効果の考え方を根底から覆す新手法」 田中 弦, 佐藤 康夫, 杉原 剛, 有園 雄一 : 著、インプレスジャパン
- [4] 「ベイジアンネットワーク技術」 本村陽一 岩崎弘利 : 著、東京電機大学出版局
- [5] 「Bayesian Networks in R: with Applications in Systems Biology」
Radhakrishnan Nagarajan , Marco Scutari, Sophie Lbre : 著、 Springer

会社概要

社名 株式会社ALBERT

設立 2005年7月1日

資本金 3億3,900万円

株主 デジタル・アドバタイジング・コンソーシアム株式会社、
IVP Incubator, L.P、オリックス・キャピタル株式会社、
株式会社ジャフコ、バーチャレクス・コンサルティング株式会社、
三生キャピタル株式会社、東洋キャピタル株式会社、
ニュー・フロンティア・パートナーズ株式会社、
SMBCベンチャーキャピタル株式会社、信金キャピタル株式会社
PE&HR株式会社、大和企業投資株式会社、
株式会社シーエー・モバイル、役員および従業員

役員 代表取締役会長 山川 義介
代表取締役社長 上村 崇
取締役 徳久 昭彦 (DAC取締役CTO)
執行役員 安達章浩
木野英明
佐藤めぐみ
平原昭次
監査役 谷本 篤彦
非常勤監査役 保月 英機

事業内容 CRMソリューションの開発・提供

- ・ 分析・コンサルティング
- ・ データマイニング・ソフトウェア
- ・ マルチチャネルOne to oneマーケティングソリューション

行動ターゲティング広告システムの開発・提供

- ・ レコメンド特化型DSP
- ・ 広告配信の最適化
- ・ 広告クリエイティブの最適化

レコメンドエンジンの開発・提供

- ・ Webレコメンドエンジン
- ・ モバイルレコメンドエンジン
- ・ 感性検索システム

事業概要



2005年7月設立。事業コンセプトは『分析力をコアとするマーケティングソリューションカンパニー』。高度なCRMソリューションをカジュアルに提供するテクノロジーとして、統計解析、データマイニング、テキスト解析、マーケティングリサーチに加え、画像解析、豊富な導入実績に裏付けられた信頼のWeb、モバイル、ITインフラ技術を保有。

これらのキーテクノロジーをベースに独自開発のビッグデータ対応データマイニング・ソフトウェア『smarticA!データマイニングエンジン』や、クラウド型キャンペーンマネージメントシステム『smarticA!キャンペーンマネージメント』、行動履歴を使ったレコメンドエンジン『おまかせ！ログレコメnder』、対話型意思決定システム『Bull's eye』、さらに、レコメンド特化型DSP『ADreco』、広告配信の最適化を実現する『i-Effect』など、分析力を強みとしたマーケティングソリューションを提供しています。

2000年

2005年

2010年



インターネットリサーチの草分けとして2000年に現ALBERT代表取締役会長の山川と数名が創業。高度な調査分析手法やデータマイニングを強みとし、高付加価値なリサーチコンサルティングで急成長を遂げる。NSX（ネットスタイルインデックス）プロジェクトを創業間もなく発足し、消費者の行動予測モデルの開発に着手する

2007年M&AによりYahooグループ入り。インターネットリサーチのインフラプラントと共にヤフー・バリューインサイトとなる。2010年8月マクロミルと経営統合。



レコメンデーションの専門企業

インタースコープのコア技術を引き継ぐ形で、「レコメンデーションの専門企業」として2005年7月に設立。国内300サイト以上にレコメンドエンジンを提供するリーディングカンパニーに成長

