

混雑を予測し情報提供することで心地良い賑わいを

NTTコムウェアではユーザ個人の属性・時間・位置・周辺状況等の情報を基にニーズや行動を先読みし、観光・移動・宿泊・買い物・食事のあらゆる場面で利用者1人ひとりに最適な情報を提案するサービスLIKEUPを提供しています。ここでは、その中で活用している人の混雑度を予測する基礎技術への取り組みや今後の展望について紹介します。

混雑が事前に予測できることの価値

ショッピングモールやスポーツ施設、テーマパークなどへ出かける際に、お出かけ先が混んでいるのか気になることが多いのではないのでしょうか。新型コロナウイルスの感染リスクを下げる意味はもちろんのこと、小さい子ども連れのご家族であれば「ベビーカーが通るのに困らないかな」や「子どもが長い行列は待てないからな」など、気になる点がさらに増えることと思います。

そんなとき、事前に目的地の混雑具合が分かったらどうでしょうか。例えばショッピングモールに出かけて「子ども連れだから、空いているうちに早めにお昼ご飯に行こうかな」と思ったときに、飲食店が混みだすまでに少し時間があることが分かれば、しばらく買い物に時間を使うことができます。また、イベントやスポーツ観戦に出かけた帰りに駅が混んでいる場合に、あと1時間で駅が空いてくることが分かれば、友人とカフェに寄って混雑が緩和されてから帰宅することも可能でしょう。このように混雑情報を

予測できることで、これまでよりも快適な時間を過ごすことができる可能性が高まるのではないかと思います。

一方、世界に目を向けるとMaaS (Mobility as a Service)* という言葉と概念が少しずつ浸透してきており、自動車やバスに加えてタクシーやレンタル自転車等を組み合わせ利用できる世界が訪れようとしています。このような交通の自由度が上がった世界では、ほんの少し空いた時間を活用できる選択肢が増え、移動先の混雑情報を知ることによって快適な過ごし方を実現できる可能性が高まっていくのではないかと考えています。

NTTコムウェアでは3つのManaging Valueを活かし、新たな社会Smart Worldの実現に向けて活動していますが(図1)、特に今回は予測や分析などのData Management領域での強みを活かした人の流れや混雑度を予測す

* MaaS: ICTを活用して交通をクラウド化し、公共交通が否か、またその運営主体にかかわらず、マイカー以外のすべての交通手段によるモビリティ(移動)を1つのサービスとしてとらえ、シームレスにつなぐ新たな「移動」の概念です(国土交通省より引用)。

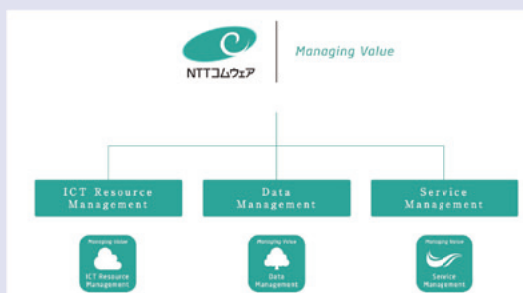


図1 NTTコムウェアのManaging Value

る技術について紹介します。

エリア汎用型予測モデル

■予測における学習データ取得期間の課題

B2Bモデルでターゲットと想定している鉄道事業者や商業施設、自治体などに利用していただく場合には費用や予測精度、機能性などに加えて、導入の容易性もポイントになってきます。

導入の煩雑さを上げる要因については、まず物理的なセンサ設置の手間が大きな要因となりますが、もう1つ学習用データの準備という要因もあげられます。具体的には機械学習や時系列モデリングなどの手法を利用する場合には予測モデルをつくるために過去のデータが必須となるため、実際に予測を始めたい期間に先立って1、2カ月前からセンサを設置して予測に必要な学習データを収集しておく必要があります。弊社データサイエンス推進室では、この事前の学習データの収集が不要になるような技術開発を行っています。

■エリア汎用型予測モデルのアプローチ

時系列予測を行う場合、混雑の増加の勢いなどの時系列特徴量を作成したうえで機械学習の各種アルゴリズムやARIMAなどの時系列アルゴリズムを用いて予測を実施することになりますが、通常は図2の「レベル1」に示すよ

うに当該の場所で取得した人流データは他の場所で利用できないので、新しい場所へ適用するには学習データの取り直しを行います。さらに、人流の傾向が大きく違う場合には、利用する特徴量の設計を変更する対応も必要となってきます。

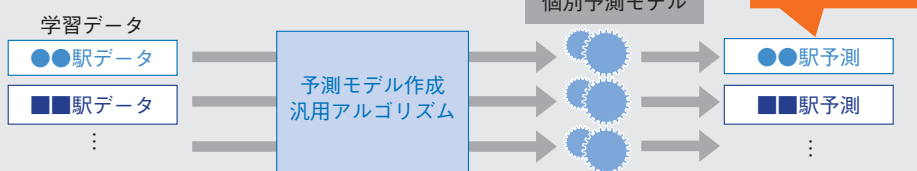
ここで人流特性（例えば昼に人が多く集まり夜は人がいなくなるなど）を似た場所に絞ってみると、ある場所でも取得したデータに最適化した特徴量と予測モデルを利用して、新しいデータを取得しなくても、ほかの類似地域に最適化した特徴量とデータを利用してある程度の混雑度を予測できるのではないか、というのがこのエリア汎用型予測モデルの発想です（図2「レベル2」特許出願中）。

例えば駅やショッピングモール、スポーツ施設など特定の種類の場所を集めて予測する場合であれば1週間の周期的な傾向やイベントなど突発的事象による増減度合いなどが似てくることが予想されます。例えば、住宅地にある駅であれば通勤時間帯の影響が大きいので過去の曜日やカレンダー、時間帯を重視して設計し、より精度の高いモデルを作成することができ、逆にスタジアムやコンサート会場などがある場所では、直近の人の増減率や、近隣施設に置いたセンサデータなどを使ってイベントをとらえるように設計することで、別々の駅であっても同じ予測モデルを共用することができます。

さらなるメリットとしては例えば同じような住宅地の駅

【レベル1】個別予測モデル作成の汎用アルゴリズム

◆ センサ設置後学習データがたまったフェーズ



【レベル2】未学習の駅を予測可能な汎用予測モデル

◆ センサ設置後の学習データがたまる前のフェーズ

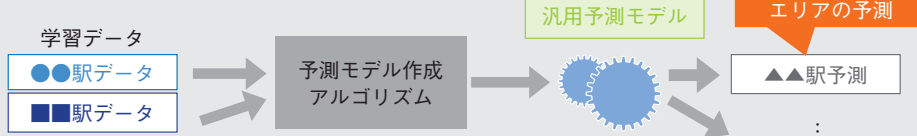


図2 混雑予測の仕組み

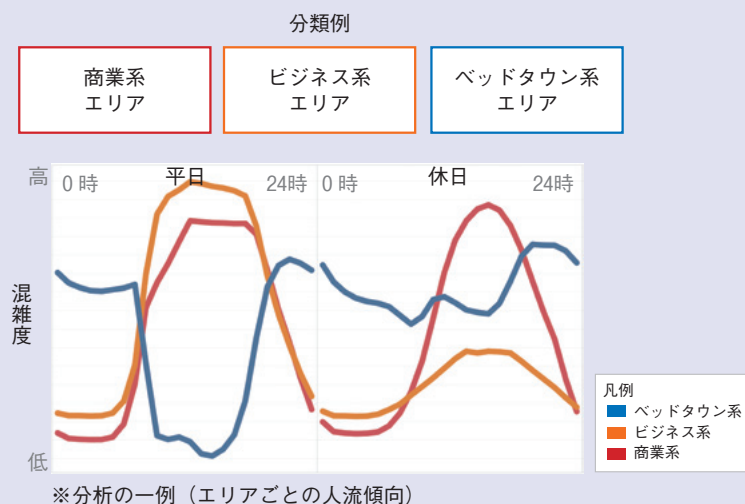


図3 エリア分類例

であれば、最初から類似エリアの予測モデルを使うことで、データ取得する期間を設けなくても、最初から類似エリアの予測モデルを使って予測を開始できるようになります。

■エリア分類仮説

では、さまざまな場所を類似エリアにどのように分類するかを鉄道駅周辺のエリアで検討していきます。当初の仮説としては、以下の4つのエリアに分類できるのではないかと想定したうえで検証を進めました。

(1) 商業系エリア

ショッピングやレジャー、友人との食事などに出かける多くの人が集まるような場所を想定しています。首都圏だと新宿や横浜など地域を代表するような場所をイメージしています。これらの場所では平日、休日通じて昼間帯から夜まで人が多くなることが想定されます。

(2) ビジネス系エリア

商業系エリアと同じように多くの人が集まる駅でもどちらかというとオフィスビルが建ち並ぶ場所をイメージしています。首都圏だと大手町や品川などでしょうか。これらの場所では平日の昼間帯に人が多くなることが想定されます。

(3) ベッドタウンエリア

主にマンションや戸建てなどの住宅が多い場所をイメージしています。首都圏であれば都心まで30～40分程度かかる私鉄・メトロの沿線駅をイメージしています。これらの

場所では夜間帯に滞在人口が多くなることが想定されます。

(4) イベントエリア

スタジアム、コンサートホール、イベントホールなどのイベント施設がある場所をイメージしています。これらの場所ではイベント開催時に急に混雑が発生すると想定されます。

■エリア分類の技術

分類の技術として機械学習における複数のクラスタリング技術を併用・比較しながら実施していますが、ここでの分類のために利用する元データを2種類利用することにしました。

1番目は商業系・ビジネス系・ベッドタウン系の人流傾向をとらえるものですが、これらは仮説からすると昼夜や平日によって混雑度が変わるサイクリックな周期性の特徴があると想定できます。このため、7日×24時間の平均的な混雑度をクラスタリングの元データとして各種クラスタリング技法で分類を試みました。この結果として分類したエリアごとの平均的な混雑度の変動を集計したものを模式的に図3に示しています。ベッドタウンエリアにおける平日の昼間に人口が減るのに対して、ビジネス系エリアでは平日昼間に人口が増加し、商業系エリアについては平日、休日とも人口が増加するかたちになっていることが見てとることができ、明確に各エリアの人流傾向の違いエリアの分類ができていることが分かりました。

	エリアグループ			
	ベースの人流傾向			イベント特性
エリア差異がある観点	商業系	ビジネス系	ベッドタウン系	日々の変動
平日 朝→昼の人数	増加	増加	減少	(エリアに依存)
平日 夕方→深夜の人数	減少	減少	増加	(エリアに依存)
休日 朝→昼の人数	増加	横ばい	横ばい	(エリアに依存)
休日 夕方→深夜の人数	減少	横ばい	横ばい	(エリアに依存)
平日と休日の 昼間人口の差	少ない	大きい	少ない	(エリアに依存)
日々の夜間人数バラツキ	(イベント特性に依存)	(イベント特性に依存)	(イベント特性に依存)	大きい

図4 エリアごとの人流傾向

2番目はイベント的要素になりますが、こちらは1番目とは逆に1週間または24時間のサイクリックな人の変動からどの程度外れた要素が強いかをとらえるものとなります(もちろん、野球やサッカーなどのオンシーズンには一定の曜日や間隔でイベントを開催する要素もありますが、必ずしも固定的な周期ではないのでアルゴリズムでとらえにくいものとなります)。したがって、ここでは統計値としてばらつきの指標である分散や変動率などをクラスタリングの元データとして分類を実施しました。こちらの結果としてイベントエリアでの変動率が特に大きくなっており、イベントがある日とない日での混雑度の違いが大きい一方、イベントエリア以外ではある曜日ある時間での混雑度の違いが日によってあまり変わらないことが分かりました(図4)。

社会貢献のための技術開発

以上のとおり、現在取り組んでいる混雑度を予測する技術について紹介をしてきました。NTTコムウェアでは単なる混雑予測にとどまらずユーザがより良い行動を選択する支援ができるように、さまざまな技術の開発やサービスの企画・開発に取り組んでいます。

例えば混雑度の情報だけを単に提供するのではなく、ユーザの行動特性や嗜好に合わせた近隣の飲食店やカフェなどの情報を提供することで、混雑回避の時間を快適な時間に

変えられるようなアプリケーションの提供も行っています。一口に快適な時間といっても1人ひとりのユーザには思考・好み・行動特性などの違いがあるため、このようなさまざまな観点から見た個性に基づいたユーザエクスペリエンスにより心地良さを演出していく仕組みづくりが重要となります。

こういった仕組みは単なるユーザ個人としてのメリットだけではなく、そのエリアやコミュニティにとっても、集中的に混雑するエリアができないようコントロールしながら、全体として賑わいを保つという点で持続的な社会成長に貢献するソリューションであると考えています。

NTTコムウェアはこれからもData Management領域のさまざまな技術の活用を通じて社会に貢献できる技術開発に取り組んでいきます。

◆問い合わせ先

NTTコムウェア

ビジネスインキュベーション本部 データサイエンス推進室

E-mail bi-bid-bd-service-g@srv.cc.nttcom.co.jp