

Title	ODデータおよびSuicaデータの組合せによる駅構内旅客流動分析
Sub Title	The analysis of passenger flow in station by integrating OD data and Suica using data
Author	酒造, 推(Shuzo, Tai)
Publisher	慶應義塾大学湘南藤沢学会
Publication year	2012
Jtitle	交通運輸情報プロジェクトレビュー No.21 (2012.) ,p.36- 45
JaLC DOI	
Abstract	JR東日本では多数の鉄道路線及び駅を有しており,多くの旅客にご利用いただいている。その中で,駅の構内において旅客がどのように移動しているかを俯瞰的に把握することは輸送障害時の現状把握や輸送計画の策定に際して大変重要と考えられる。しかし現在は駅構内の旅客の動きを把握するには防犯カメラの映像を直接調査する方法が中心のため時間と労力を要しており,俯瞰的な把握が難しい。そこで本研究では,旅客が鉄道を利用する際に生成されるデータから駅構内における旅客の動きをシミュレートするアルゴリズムを提案し,実際の駅における旅客流動とアルゴリズムの適用結果を比較することで有用性の評価を行った。
Notes	2012年度慶應義塾大学JR東日本寄附講座報告書 慶應義塾大学交通運輸情報プロジェクト その2 : .JR東日本派遣研究員の活動報告
Genre	Technical Report
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO92001006-00000021-0036

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the Keio Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

OD データおよび Suica データの組合せによる駅構内旅客流動分析

The analysis of passenger flow in station by integrating OD data and Suica using data

酒造 推

慶應義塾大学 SFC 研究所

東日本旅客鉄道株式会社 フロンティアサービス研究所

Tai Shuzo

Keio Research Institute at SFC, Keio University

Frontier Service Development Laboratory, East Japan Railway Company

JR 東日本では多数の鉄道路線及び駅を有しており、多くの旅客にご利用いただいている。その中で、駅の構内において旅客がどのように移動しているかを俯瞰的に把握することは輸送障害時の現状把握や輸送計画の策定に際して大変重要と考えられる。しかし現在は駅構内の旅客の動きを把握するには防犯カメラの映像を直接調査する方法が中心のため時間と労力を要しており、俯瞰的な把握が難しい。そこで本研究では、旅客が鉄道を利用する際に生成されるデータから駅構内における旅客の動きをシミュレートするアルゴリズムを提案し、実際の駅における旅客流動とアルゴリズムの適用結果を比較することで有用性の評価を行った。

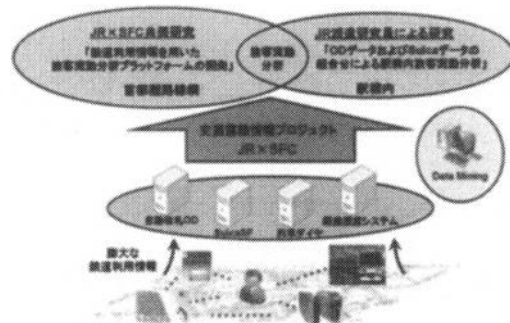
キーワード：旅客流動、OD データ、自動改札機、データベース、データマイニング
Keywords: passenger flow, OD data, automatic ticket gate, database, data mining

1. 研究背景

JR 東日本ではグループ経営ビジョン V において駅のポテンシャルを高めることを重要な経営課題としており、さまざまな施策を進めることとしている。具体的な施策としては「大規模駅のターミナル開発」・「街や周辺エリア全体の顔となる駅の実現」・「駅のホームにおける安全対策の推進」が挙げられる。これらの課題を進める上で共通の課題が、駅の中で常に変化する動的情報の把握が必要であるということである。

一方で、JR 東日本は毎日多くのお客さまにご利用いただいております。自動改札 OD データや SuicaSF データ等の膨大な鉄道利用情報が日々手に入る環境にあり、これらの有効活用が課題となっている。JR 東日本が慶應 SFC と進めている交通運輸情報プロジェクトでは、これらの鉄道利用情報からデータマイニングの手法

で有用な知見を得ることを目的とした研究を立ち上げた。研究概要及びイメージを図 1 に示す。本研究ではこのような状況の中、駅の動的情報として駅構内の旅客流動に着目し、鉄道利用情報を元とした旅客流動分析技術を確立することを目的とした。



2. 本研究の特徴

本研究で提示する分析手法の特徴について2点述べる。1点めは、同時に複数の駅に対しての旅客流動分析が可能であることである。このため、流動結果を基にした駅ごとの比較・検討が容易に行える。2点めは、JR 東日本がすでに所持している鉄道利用情報を利用するため、分析コストが少なくすむことである。

これらの特徴から、本手法は駅構内における事業計画の策定や旅客利用に伴い劣化する設備の取替時期の検討など、同時に複数の駅の旅客流動を分析し比較するケースに対して有用と言える。図2に本手法を用いたシステムの概念図を示す。ユーザが駅名及び分析期間をクエリとして入力すると、システムは鉄道利用情報から当該データを取得して分析を行い結果を返す。ユーザは結果から希望する駅の旅客流動を把握することができる。

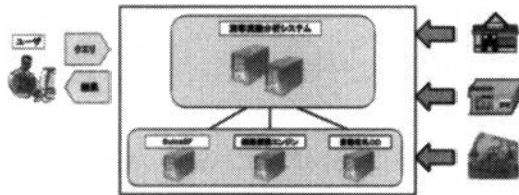


図2 旅客流動分析システム概念図

3. 駅構内の旅客移動モデル

本研究で用いる、駅構内の旅客移動モデルについて述べる。駅構内における旅客流動は、個々の旅客の移動の集合として表される。図3に旅客の移動モデルを示す。旅客はまず出発駅の自動改札機から入場し、お茶や新聞など物販の購入を行った後出発番線へ移動する。列車に乗車後、到着番線についた後、到着駅にてATMやトイレ等のその他の行動をとり、自動改札機から出場する。これが一連の流れである。

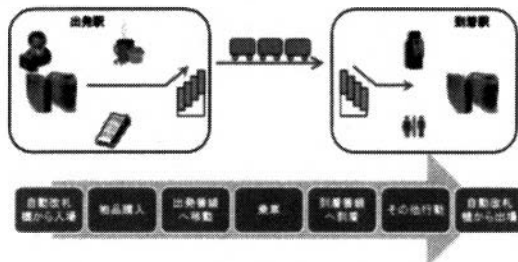


図3 旅客移動モデル

旅客が駅構内を移動する際、自動改札機の通過と出発番線への移動、及び到着番線から自動改札機への移動は必ず行う。したがって、駅構内における個々の旅客移動を「自動改札機と利用番線および移動時間の組合せ」にモデル化することができる。本研究ではこのモデルを駅構内トリップと定義した。駅構内トリップを用いることで、駅構内における個々の旅客の移動を示すことが可能となる。図4では例として吉祥寺駅における旅客の移動を駅構内トリップで示した。旅客が4番線ホームに到着して中央改札を15:25に出場したことや、南口改札から9:10に入場して1番線ホームに向かったことを示すことができる。

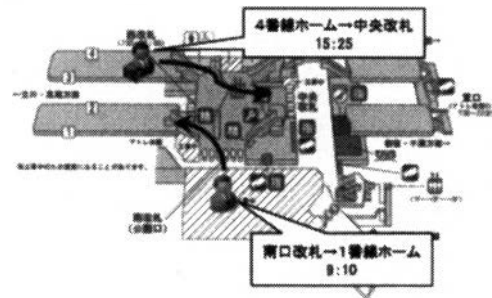


図4 駅構内トリップ

駅構内における個々の旅客の移動を駅構内トリップとしてモデル化すれば、これを合成することで旅客流動を再現することができる。そこで、旅客が自動改札機を通過する際に生成される鉄道利用情報から駅構内トリップを求めるプロセスについて述べていく。旅客流動の再現イメージを図5に示す。

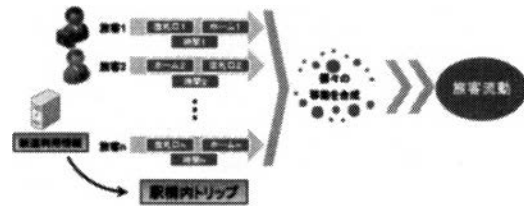


図5 旅客流動再現イメージ

4. 鉄道利用情報

本研究では、鉄道利用情報を「旅客が鉄道利用時にログとして蓄えられる情報」と定義した。概要を図6に示す。鉄道利用情報は車両や自動券売機などさまざまな機器に存在しているが、旅客の移動状態を示すデータを持つ自動改札機に着目する。旅客の移動状態を示すデータとしては自動改札ODデータ及びSuicaSFデータがあり、自動改札ODデータは旅客数・発着駅・利用券種・改札口、SuicaSFデータは発着駅・利用時間・改札口をデータ内容として持っているため、これらのデータから駅構内トリップを求めることが分析の目的となる。

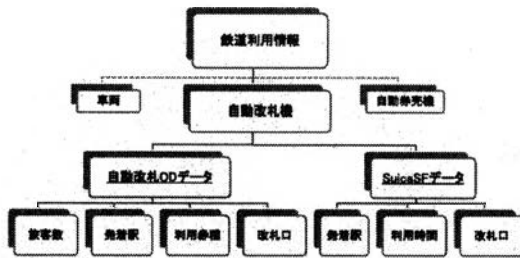


図6 鉄道利用情報

4.1. 自動改札 OD データ

自動改札 OD データは JR 線内における利用旅客数の一覧を示しており、発着駅及び到着時間 1 時間ごとの旅客数合計を所持している。そのほか、利用券種や到着改札口等ごとの人数も所持している。図7にデータ概要を示す。自動改札 OD データは駅構内トリップを求めるベースとなりうるが、そのままでは作成が困難であり、他データとの組み合わせが必要である。

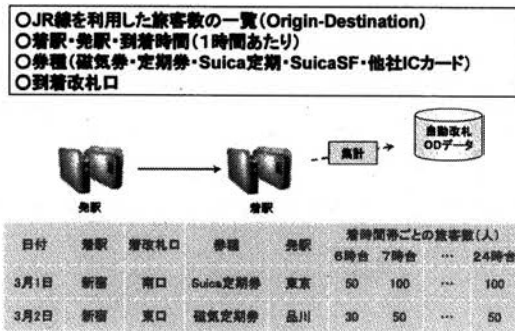


図7 自動改札 OD データ

自動改札 OD データから駅構内トリップを得る際の課題を図8に示す。課題の1点めは駅での利用番線及びホームが不明であることである。そのため、本研究では経路探索エンジンを元に利用番線の推定を行う。例えば東京駅から新宿駅まで向かう場合、経路探索エンジンから東京駅 2 番ホームと新宿駅 12 番ホームを利用することがわかる。課題の2点めは、自動改札 OD データが1時間単位の累計の旅客数であるため個々の旅客の移動時間が不明であることである。そのため、SuicaSF データ及び駅時刻表データを元に詳細な移動時間の推定を行う。

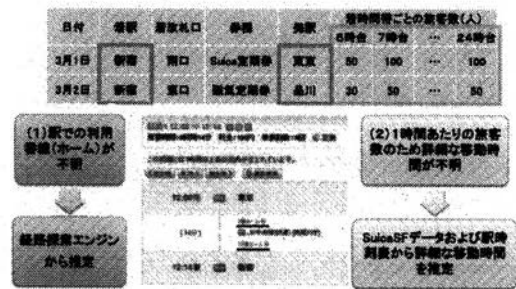


図8 自動改札 OD データの課題

4.2. SuicaSF データ

SuicaSF データは電子マネーSuica の利用履歴である。データ概要を図9に示す。SuicaSF データはSuicaを持つ個々の旅客の移動ごとに入場駅・出場駅について1分単位の利用時間を持っているが、SuicaSF 利用時のみ生成され、Suica 定期利用時には生成されない。

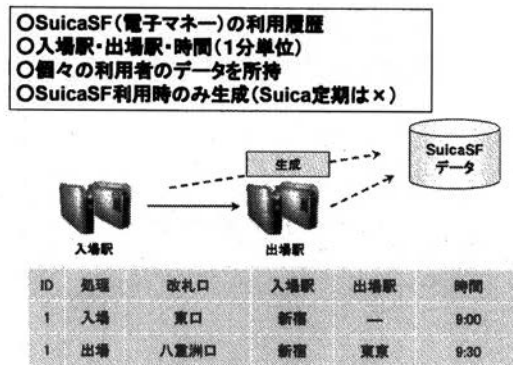


図9 SuicaSF データ

SuicaSF データは個々の移動における入出場駅及び利用時間を所持しているため駅構内トリップと意味上は同値の情報であるが、SuicaSF を利用した旅客の流動しか示せない課題がある。これらのことから、分析を行うには自動改札ODデータとSuicaSFデータの組合せが必要となる。

5. 駅構内旅客流動分析

本研究における駅構内旅客流動分析イメージを図10に示す。

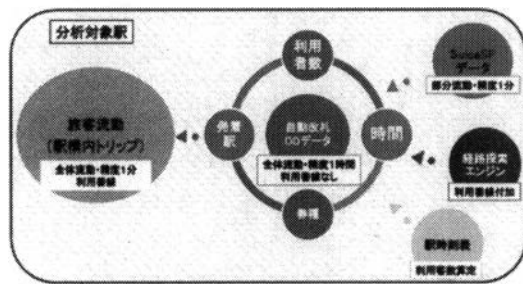


図10 旅客流動分析イメージ

ベースとなる鉄道利用情報が自動改札ODデータであり、駅の旅客流動全体を示しているが、駅における利用番線の情報を持たず、また1時間ごとの累計人数であるため第3章で述べた駅構内トリップを単独で求めることはできない。そのため、他データから情報の補完を行う。具体的には自動改札ODデータの旅客数をSuicaSFデータで1分あたりの移動に分割し、経路探索エンジンを用いて利用番線を付加し、分析駅における時刻表を用いて個々の列車における利用客数の推定を行う。

本研究では流動分析の具体例として吉祥寺駅に着目し、上記で述べた旅客流動分析プロセスを適用することで駅構内の旅客流動の定量化を行った。データ分析にはリレーショナルデータベース管理システムであるPostgreSQL9.0を用いた。

5-1. 吉祥寺駅概要

吉祥寺駅の概要を図11に示す。吉祥寺駅は中央線の主要駅であり、中央快速線および中央緩行線が複々線として平行に走っている。分析駅への選定理由としては、駅の各ホームと改札

口を結ぶルートが一意に定まること、構内店舗がなく旅客の物品購入等で生じる滞留を考慮しなくてよいこと、中央快速線と中央緩行線の2路線が平行方向に走っているため構内乗換が少ないこと、などから旅客流動分析の最も単純なモデルが適用できるためである。

吉祥寺駅	
ホーム	2面4線(中央快速線・緩行線)
改札口	4箇所
乗車人員	137,555人(2011年)



図11 吉祥寺駅概要

5-2. 吉祥寺駅構内図

吉祥寺駅の構内図を図12に示す。1番線・2番線が中央緩行線ホーム、3番線・4番線が中央快速線ホームであり、コンコースからは階段とエスカレータで結ばれている。また改札口は中央口、西口、南口、東口の計4箇所である。本分析では改札口4箇所と4番線を結ぶ計16ルートの駅構内トリップを求め、駅構内の旅客流動を定量化する。

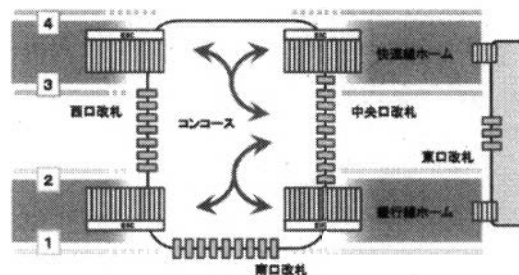


図12 吉祥寺駅構内図

5-3. 吉祥寺駅旅客流動概要

吉祥寺駅の自動改札ODデータから求めた吉祥寺駅の2013/1/23の利用人数グラフを図13に示す。このグラフは旅客流動の全体概要を示しており、朝と夕方に通勤ラッシュによるピークがあるほか、南口と西口の利用人数が多いことがわかる。一方、時間軸が1時間と粗いため、流動の概要把握に留まり、さらなる詳細な分析は困難である。

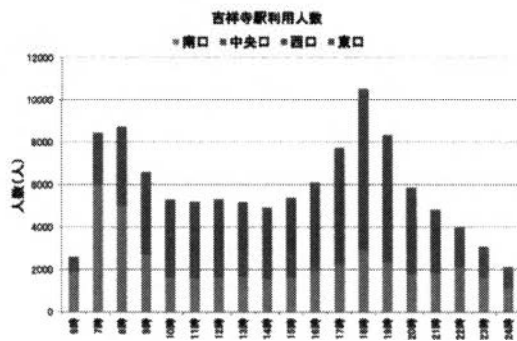


図 13 吉祥寺駅利用人数

5-4. 旅客流動分析プロセス

自動改札ODデータを基とした旅客流動分析プロセスを図14に示す。STEP1では、経路探索エンジンから吉祥寺駅における利用番線とホームを決定する。STEP2においてSuicaSFデータを解析し1分単位の旅客流動の概況を調査する。STEP3で自動改札ODデータと駅時刻表による旅客数割り当てを行う。STEP4でこれらの結果を元に駅構内トリップを決定し、旅客流動を求める。

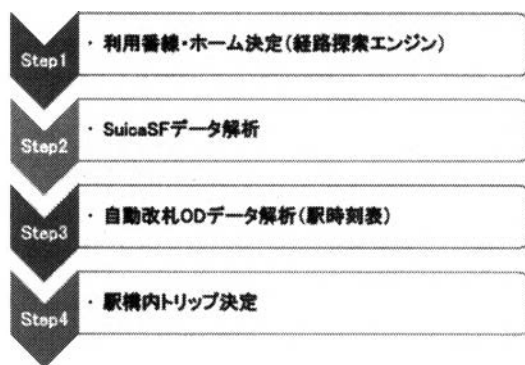


図 14 旅客流動分析プロセス

5-5. 利用番線・ホーム決定 (STEP1)

STEP1では吉祥寺駅の各番線と首都圏全駅を1:1対応させるため、経路探索エンジンで首都圏全駅から吉祥寺駅へのルートを検索することで利用番線の決定を行った。経路探索エンジンは「駅探 (<http://ekitan.com/>)」を用い、JR線のみ・乗換回数優先条件で検索を行った。例として東京駅で検索した結果を図15に示す。

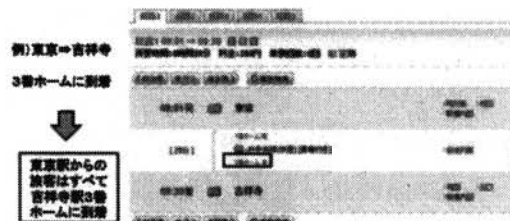


図 15 東京駅検索結果

東京駅から吉祥寺駅まで向かう場合、中央快速線を利用するため、吉祥寺駅での到着番線は3番線であることがわかる。この結果から、東京駅から吉祥寺駅へ向かった旅客はすべて吉祥寺駅3番ホームへ到着すると推定できる。

この調査を首都圏全域の約400駅に対して行い、各駅に対する吉祥寺駅到着ホームを決定した。結果の一部を図16に示す。図より、中央快速線ホームへの到着が多く、総武緩行線ホームへの到着は同線内の駅にはほぼ限られていることがわかった。なお、吉祥寺駅から出発する際は同じホームの逆の番線を利用するとした。

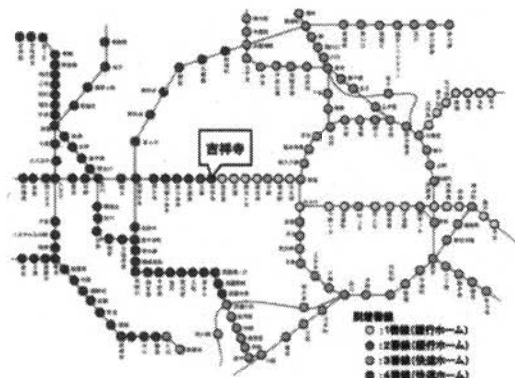


図 16 吉祥寺駅到着ホーム

5-6. SuicaSF データ分析 (STEP2)

STEP2では吉祥寺駅を発着駅とするSuicaSFデータの分析を行った。分析概要を図17に示す。

SuicaSF データ全体から吉祥寺駅に関するものを抽出し、別々に存在する入場及び出場データを接続した上で先ほど求めた利用番線を付加することで、SuicaSF データを駅構内トリップへ変換した。

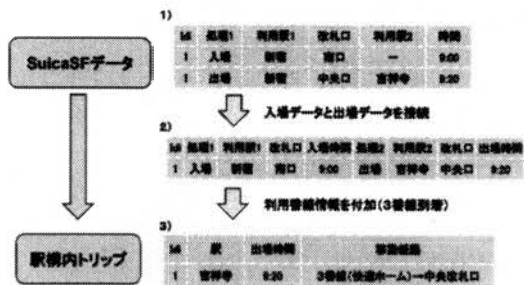


図 17 駅構内トリップ変換

SuicaSF データ(1)は入場と出場でそれぞれ別のデータが作成されるため、両データを結合して1つのトリップデータ(2)とした。さらに、5-5 から新宿駅からの旅客は吉祥寺駅3番線に到着と決定したので、その情報を付加することで駅構内トリップ(3)を求めることができる。図17の駅構内トリップはある旅客が吉祥寺駅快速ホーム3番線に到着して中央改札口に移動し、9:20に改札口を通過したことを示している。

SuicaSF から求めた吉祥寺駅の全改札口の駅構内トリップを集計した結果を図18に示す。グラフの縦軸は人数、横軸は1日を分単位で示している。

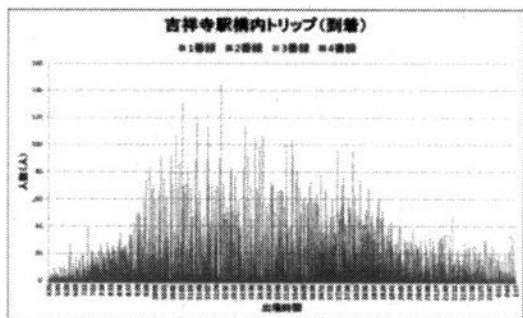


図 18 吉祥寺駅構内トリップ (SuicaSF)

グラフから 12 時頃に流動のピークを迎えており、全流動(図13)で見られたような朝や夕方方のラッシュは見られない。これは SuicaSF による流動であるからである。

そこで図18のグラフをさらに考察するため、10時~11時の範囲で4番線のみ抽出し、電車の到着時刻を時間軸上に矢印で付記した結果を図19に示す。

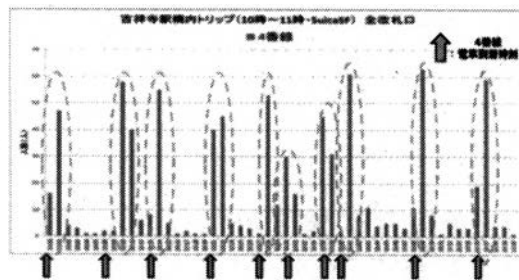


図 19 4番線抽出結果(10時~11時)

図19から、4番線からの流動のピークは電車到着時間とほぼ1:1に対応しており、電車到着後1分遅れ程度で生じることがわかる。そのため、旅客流動のピークが SuicaSF 以外の券種でも電車到着時間と対応していると仮定すれば、自動改札 OD データに電車到着時間を重ねることで流動のピークを求めることができる。

5-7. 自動改札 OD データ分析 (STEP3)

STEP3 では5-6で得られた結果を元に、自動改札 OD データ分析を行った。具体的には、駅時刻表を元に、吉祥寺駅の各番線における1時間あたりの旅客数を各電車へ分解した。分析の概要を図20に示す。

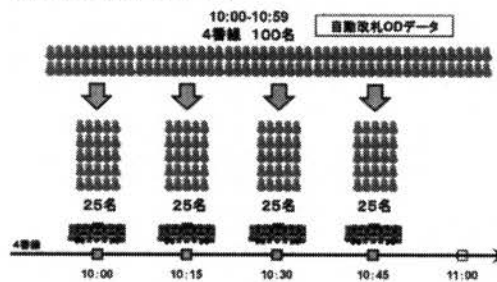


図 20 自動改札 OD データ分析

自動改札 OD データから、吉祥寺駅4番線に10時台で100名の旅客が到着したとする。ここで4番線に電車が10:00、10:15、10:30、10:45の4回到着するならば、各電車に旅客を25名ずつ配分することで、4回流動のピークが生じると推定できる。同様の手続きを全時間・全番線について行うことで、駅構内の旅客流動を1分単位で明らかにすることが可能となる。本分析では電車の到着と流動のピークの遅れ時間を1~2分と設定した。

分析プロセスの適用前後として、南口改札における分析プロセス前の1日分の旅客流動を図21に、分析プロセス後の流動を図22に示す。

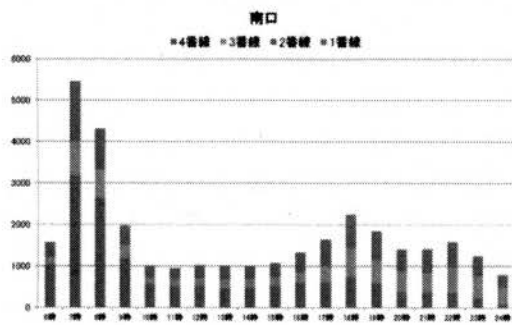


図21 南口旅客流動 (分析前)

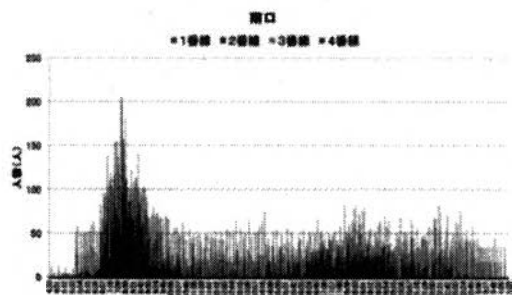


図22 南口旅客流動 (分析後)

分析前(図21)では時間軸が1時間単位であるため、全体的な旅客流動の傾向は把握できるものの、流動のピーク値などの詳細な情報は求めることができない。

一方、分析後では時間軸が1分単位であるため、1日の流動状況を細かく表せている。この結果から、朝ラッシュのピーク時の利用者数が最高で1分あたり約200人であることがわかったほか、朝ラッシュは中央快速線上りの4番線の利用が多く、夕方ラッシュは中央快速線下りの3番線の利用が多いことがわかる。

同様に、東口に旅客流動分析プロセスを適用した結果を図23に示す。東口は駅ビル連絡改札のため利用者数が少なく、流動のピークが1分あたり約40人であった。一方、流動が7:30から22:00までに限られており、これは東口改札の利用可能時間と等しい。

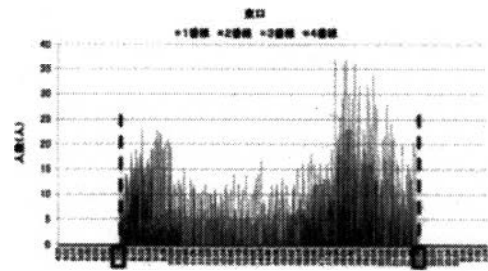


図23 東口流動分析結果

次に駅構内の流動状況把握として、南口について求めた結果を図24に示す。8:00-8:10の間で南口に到着した人数を矢印の太さで表した。同時刻は朝ラッシュ帯であり、4番線の中央快速上り線からの旅客が最も多いことがわかる。

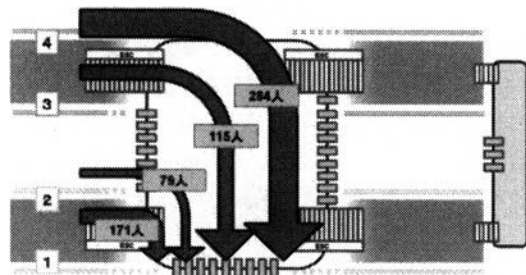


図24 南口流動状況

6. 旅客流動分析評価

旅客流動分析の精度を評価するため、3番線・4番線ホームから出口へ向かう実際の旅客流動を駅構内防犯カメラより測定し、分析結果(シミュレーション)との比較を行った。評価概要を図25に示す。測定日時は分析日時と同じ2013/1/23とし、8:00-8:30と15:00-15:30の2回実施した。流動測定位置は図25の星2箇所であり、階段及びエスカレータ監視用の防犯カメラ映像から通過人数をカウントし、2箇所の合計を取った。比較用のシミュレーション結果としては3番線・4番線から中央口・西口・南口へ向かう流動人員の合計とした。東口へ向かう流動は別ルートを通るため、比較対象に含めていない。

測定日時	2013/1/23
測定時間	8:00~8:30 15:00~15:30
シミュレーション 比較対象	3・4番線⇒中央口・西口・南口 の流動人員合計

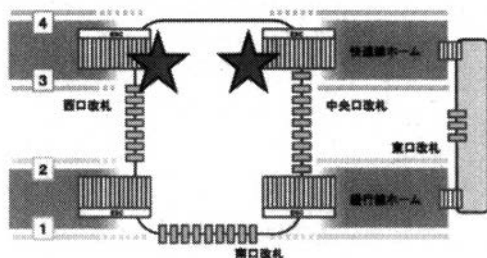


図25 実流動とシミュレーションの比較

6-1. 15時における比較

15:00-15:30 における旅客流動の比較結果を図26に示す。流動の実測値及びシミュレート結果をグラフで示している。グラフ中の点線は両結果に生じている時間のずれである。シミュレーションは改札口を通過した時間をベースとして行っているため、階段を通過した時間を測定している実測値とは時間の遅れが生じている。グラフのピークを取る時間の差から、約2分のずれが生じていることがわかる。

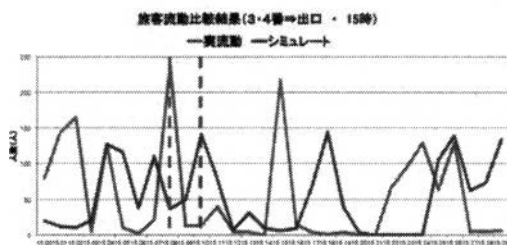


図26 旅客流動比較結果 (15時)

遅れの影響をなくすため、シミュレートの時間を補正した流動の比較結果を図27に示す。波形については実流動を近似できているが、ピーク値においては実流動より低い結果が出ていることがわかる。

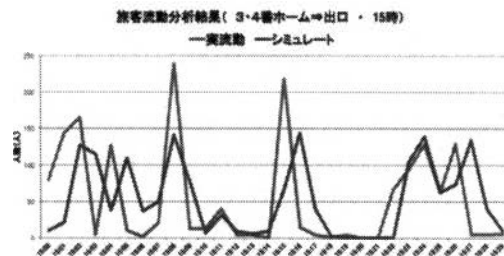


図27 旅客流動比較結果 (15時・補正後)

6-2. 8時における比較

8:00-8:30 における比較結果を図28に示す。点線で囲んだ8:20-8:30の間では近似できているが、それ以外の時間では正しい値を示せていない。その理由としては、朝ラッシュによるダイヤ乱れによる遅れの影響が出ていると考えられる。シミュレーションは時刻表をベースに行っていることから、ダイヤ乱れが発生した場合正しい結果を示すことが困難である。

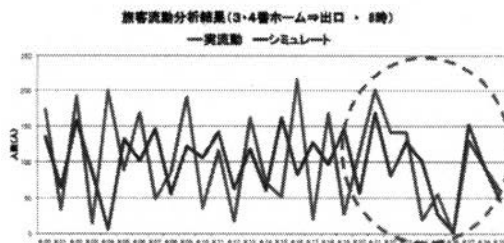


図28 旅客流動分析結果 (8時・補正後)

6-3. 考察

シミュレーション結果が旅客流動の実績値と異なる点について考察する。本研究の旅客流動分析プロセスにおいては、自動改札ODから求めた旅客数を電車本数に対して均等に配分しているため、ピーク時の人数でずれが生じている。改善策としては、各電車の乗客数から旅客数配分を行うプロセスを組み込むことで、より精度の高い旅客流動分析が可能となると考えられる。図29は列車荷重データより適切な旅客配分を行う概念図を示している。

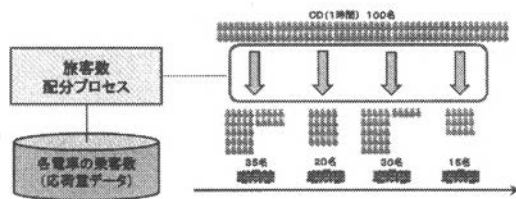


図 29 列車荷重データ適用概念図

7. 他駅への展開

本旅客流動分析手法の課題として、他駅への展開について述べる。本研究においては吉祥寺駅にて分析を行ったが、当駅は乗換需要・経路候補・構内店舗がいずれも少ないため、旅客流動分析の最も単純なモデルが適用できた。しかし、JR 東日本管内には様々な駅があり、すべての駅に同様の分析プロセスを適用することはできない。適用するには駅の特성에応じた新しい分析プロセスを組み込む必要がある。図 30 に他駅への展開に向けた課題を示す。

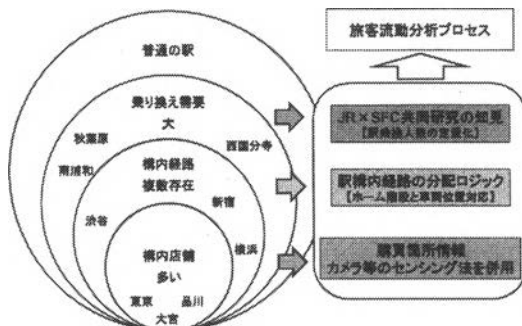


図 30 他駅への展開

まず、乗換需要が大きい駅については、現在 JR と慶應 SFC で進めている共同研究の知見を用いて、駅乗換人数を定量化する分析プロセスの導入が必要であると考えられる。次に、構内経路が複数存在する駅については、駅構内経路候補に旅客数を分配するプロセスを組み込むことでより細かい分析が可能である。構内店舗が多数存在する駅については、購買箇所の情報や防犯カメラ等、他のセンシング法を併用する方法が考えられるが、完全な旅客流動の再現には多くの課題が残る。

8. まとめ

本研究では、旅客の鉄道の利用時に生じる情報

である自動改札 OD データ及び SuicaSF データを組合せて駅構内の旅客流動の分析を行う方法を提示した。この方法を用いることで、既存の鉄道利用データから旅客流動という今後の施策を進める上で重要な知見が得られることがわかった。続いて、上記手法を用いて吉祥寺駅構内の旅客流動分析を行い、実際の旅客流動と比較することで、分析プロセスの評価を実施した。また、考察として JR 東日本管内の他駅へ展開する際の課題について述べた。

今後、駅構内における旅客流動を知るニーズは高まると確信しており、その際に本研究の知見が一助になれば幸いである。

謝辞

清木先生をはじめ、交通運輸情報プロジェクトの各先生方、慶應義塾大学清木研究室の皆様方には的確なご指導を頂き、研究を進めていくことができました。JR 東日本研究開発センターフロンティアサービス研究所の皆様には温かいご助言とサポートを頂きました。また、IT・Suica 事業本部および吉祥寺駅の皆さまには研究に関する様々なデータをご貸与頂きました。この場をお借りしてご支援、ご協力頂きました皆様に深く感謝申し上げます。

参考資料

- [1] 清木 康, 倉林 修一, 『OD データと経路データを活用した旅客流動の分析』, 慶應 SFC, 2012 年
- [2] 加瀬 史朗, 佐藤 敏彦, 坂本 圭司, 『ホーム上旅客流動の変化を予測できるシミュレーションシステムの開発』, JR EAST Technical Review No.41, 2012 年
- [3] 斎藤 武, 『首都圏ネットワーク上の旅客流動分析とその応用』, 交通運輸情報プロジェクト, 2008 年
- [4] ミック, 『SQL ゼロから始めるデータベース操作』, 翔泳社, 2010 年
- [5] ミック, 『達人に学ぶ SQL 徹底指南書』, 翔泳社, 2008 年
- [6] 石井 達夫, 『PostgreSQL 徹底入門第 3 版』, 翔泳社, 2011 年
- [7] 増永 良文, 『データベース入門』, サイエンス社, 2006 年

付録

本研究を進めるにあたり作成したプログラムの一部を以下に示す。使用言語は PostgreSQL9.0.1 である。

```
--ODデータ スキーマ定義
CREATE TABLE          oddata
(
  Tdate      character varying(10),      --日付
  dest       character varying(15),     --到着駅
  gate       character varying(15),     --改札口
  tic        character varying(15),     --きっぷ種類
  ori        character varying(20),     --出発駅
  t6 integer, t7 integer, t8 integer, t9 integer,
  t10 integer, t11 integer, t12 integer, t13 integer,
  t14 integer, t15 integer, t16 integer, t17 integer,
  t18 integer, t19 integer, t20 integer, t21 integer,
  t22 integer, t23 integer, t24 integer, tsum integer
);
```

```
--Suicaデータ スキーマ定義
CREATE TABLE suicamaster
(
  idi character varying(20) NOT NULL,   --idi
  shori character varying(10) NOT NULL, --処理
  kiki character varying(10),          --会社名
  kiki_id character varying(16),       --機器ID
  riyou_eki1 character varying(10),    --利用駅1
  riyou_eki2 character varying(10),    --利用駅2
  riyougaku integer,                   --利用額
  date_time timestamp without time zone NOT NULL, --時間
  zangaku integer                       --残額
);
```

```
--SuicaデータとODデータ結合における名寄せ駅名マスタ スキーマ定義
CREATE TABLE ekimaster
(
  eki character varying(30),           --駅名
  form character varying(6),          --利用ホーム
  arr      integer,                    --到着駅
  dep      integer,                    --出発駅
  time_d   integer,                    --到着時
  change   integer,
  direc    character varying(6)
);
```

```
--Suicaデータ→トリップデータ変換
CREATE TABLE trip AS
SELECT
  S1.idi,
  S1.riyou_eki1,
  S1.kiki_id as kiki_id1,
  S1.date_time as nyujo,
  S2.riyou_eki2,
  S2.kiki_id as kiki_id2,
  S2.date_time as shutsujo,
  S2.date_time - S1.date_time as shoyo
FROM Suicamaster S1, Suicamaster S2
WHERE S1.idi = S2.idi
  AND S1.shori = '入場' AND S2.shori = '出場'
  AND S1.riyou_eki1 = S2.riyou_eki1
  AND S2.riyou_eki2 <> ''
  AND STRPOS(S2.kiki_id, substr(S2.riyou_eki2, 3)) <> 0
  --'JE'の駅名部分と機器IDを比較して一致するか
  AND S2.date_time = (SELECT MIN(date_time)
  FROM Suicamaster S3
  WHERE S1.idi = S3.idi
  AND S1.shori = '入場' AND S3.shori = '出場'
  AND S1.riyou_eki1 = S3.riyou_eki1
  AND S3.riyou_eki2 <> ''
  AND S1.date_time < S3.date_time);
--条件を満たす出場トリップのうち最小時間を抽出
```