

イベント参加者数を用いた都市における混雑度予測サービス — スポーツイベント（日本プロ野球とJリーグ）の場合 —

鎌柄 拓史† 新井 教広† 一ノ木 繁† 金丸 正憲† 小林 峻† 吉野 悦朗† 中野 美由紀†

†産業技術大学院大学 〒140-0011 東京都品川区東大井 1-10-40

E-mail: † {a152hk, a15z1na, a1505si, a1520mk, a1415tk, a1549ey, miyuki}@aiit.ac.jp

あらまし 都市圏では、さまざまな理由から人が集中することで混雑が発生する。現在の混雑情報を提供する「混雑.com」等のサービスは開始されているが、蓄積された情報からの混雑度予測をする手法としては、わずかな特定のエンターテインメント会場において予測情報が提供されているにすぎない。本論文では、都市圏で生じる混雑を予測することを目的とし、恒常的に生じる混雑（例：朝の主要な JR 駅の混雑）以外に人が集中するイベント、特に短時間に多くの人数が流動するスポーツ観戦を対象として観客動員数の推定を行い、ユーザに混雑度として提示するシステムを開発した。国内で最も観客動員数の多いプロ野球およびJリーグのデータを用い、開発したシステムの有効性を検証する。

キーワード 混雑度, 予測システム, スポーツ観戦

1. はじめに

都市圏では、さまざまな理由から人が集中することで混雑が発生する。現在の混雑情報を提供する「混雑.com」[1]等のサービスは開始されているが、蓄積された情報からの混雑度予測をする手法としては、わずかな特定のエンターテインメント会場における予測情報は提供されているにすぎない。本論文では、都市圏で生じる混雑を予測することを目的とし、恒常的に生じる混雑（例：朝の主要な JR 駅の混雑）以外に人が集中するイベント、特に短時間に多くの人数が流動するスポーツ観戦を対象として混雑度の推定を行う。まずは、国内で最も観客動員数の多いプロ野球のデータを用い、過去の動員数データと球場の位置、周囲の環境と合わせ、観戦者数の推定手法について報告する。

2. 混雑予測システム eyekon

2.1 混雑予測システム eyekon

近年我が国では観光立国に力をいれ外国人観光客は2014年から2015年で1.5倍と飛躍的に伸びている。また、東京2020オリンピック・パラリンピックを控え今後も急増すると予測される。大都市圏、特に東京は混雑が常態ともいえるが、通勤時間帯や曜日により人の動きは異なる。また、大都市圏では大きなイベント会場が点在し、混雑の一因となっている。そこで、我々は東京をターゲットとし、イベント情報を基にユーザの移動、来訪予定に合わせて地域の混雑度情報を提供する「eyekon」システムを開発している（図1）。

eyekonシステムは図1からわかるように、大きくわけて、以下の3つの部分から構成されている。

- ・混雑予測用データベース
- ・混雑予測モデル
- ・混雑予測ウェブサービス

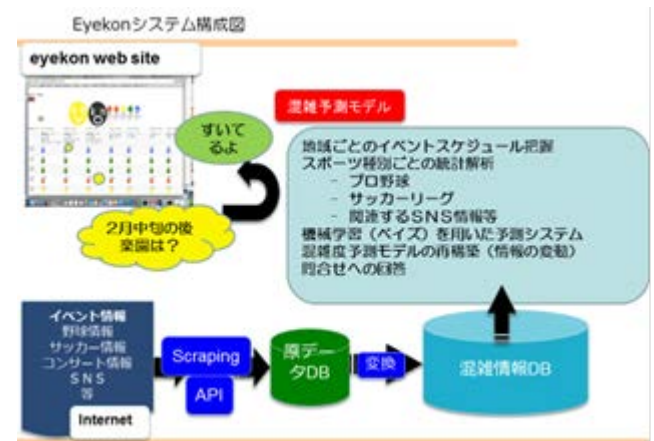


図1 eyekon システム構成

2.2 混雑予測用データベース

混雑予測用データベースはさらに二部構成に分かれている。インターネット上のイベント情報および混雑予測に必要な情報は定期的に原データベースに蓄積される。この中から、混雑度予測に必要なイベント人数等の情報がクレンジングされたのちに混雑情報データベースに格納される。一方、新たに獲得されたイベント情報等は混雑予測モデルのカレンダー上に登録されると同時に、混雑予測ウェブサービス上にもカレンダーを通して反映される。

2.3 混雑予測モデル

混雑予測ウェブサービスから呼び出され、指定された地域、日時から開催されるイベントを見つけ出し、どの程度の混雑（通常の人出と比較して数万人の増加が見込まれる）などの予測を行う。イベントごとに予測モデルが異なるため、コンサートのように繰り返し行われたいイベントに関しては、アーティストと開催会場の容量などを参考にする程度であるが、スポーツ観戦など、観客動員数がある程度正確に見込まれるも

のは、過去のデータから予測を行う。本稿ではプロ野球の観客動員数について後述する。

2.4 混雑予測ウェブサービス

ユーザの指定する日時、地域にあわせ、その後一週間程度の混雑度を天気予報に類似した形で提示する。

3. 日本プロ野球の観客動員数の予測

スポーツ観戦は一試合ごとに数万人の観客を動員し、突発的な混雑の大きな要因となる。我々は代表的なスポーツとして、プロ野球、サッカーを取り上げて観客動員数を予測することにした。本節ではプロ野球における観客数の予測について考察する。

3.1 プロ野球の動員観客数データ

サッカーの観客動員数はJリーグ本部に掲載されており、入手は容易である。また、Jリーグ自身が掲載しているため、データの正確性も高い。一方、サッカーと比べ、プロ野球の観客動員数は主催者によるまとまった報告がなく、プロ野球の趣味サイトでまとめている情報(3)を利用している。2009年から2016年の8年間、プロ野球の公式戦全部、6900試合のデータを収集した。表1にデータ情報を示す。

3.2 プロ野球観客数

図2にプロ野球の全試合の観客数分布を示す。横軸が観客動員数で縦軸が試合数となる。図2からわかるように、2009年から2016年までのプロ野球の観客動員は1万数千人から4万人と変動が大きい。なかでも、球場ごとに収容人数が異なっているため、3万人前後で山が分かれている。つまり、開催される地域、球場の収容人数の制限は観客動員数の大きな要因となる。図3に2016年の東京ドームにおける観客動員数と試合数の関係を示す。図3からわかるとおり、一つの球場においても大きな差がある。4万人を超える試合が多い一方で、一万人前後の動員数しかない試合もある。

3.3 観客数変動の要因

プロ野球はファンも多く、様々な情報がウェブ上にあがっている。東京ドーム等、天候と関係のない球場も多いが、天気は左右される可能性も高い。そこで、天候、開催日(曜日、時間)、球場(ホームチーム)、年棒など、観客動員数に影響を与えると思われるいくつかの要因をとりあげ、検討した。

図3からわかるように、同じ球場でも観客動員数にはばらつきがある。これは東京ドームをホームとするチームとして、巨人と日本ハムの二つの球団があることに起因する。日本ハムは現在札幌ドームをホーム球場としているが、年に数試合を東京ドームで開催しているためである。

表1 データ情報	
期間	2009-04-03~2016-10-05
スタジアム数	57
試合数累計	6900
スタジアムごとの試合数	最小 1、最大 186
観客数累計	181,606,380
観客数	最小 5339、最大 47106

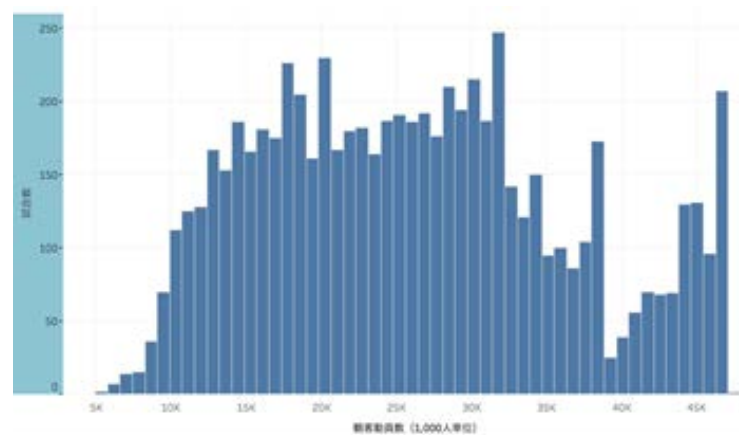
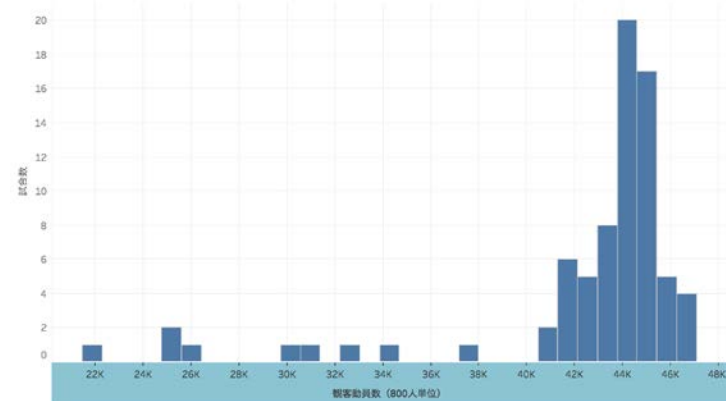


図2 日本プロ野球観客動員数(2009年-2016年)

プロ野球—東京ドーム観客動員数(2016年)



Attend(ビン) [1] の attend のカウントの傾向です。データは year および the_stadium でフィルターされます。year フィルターの範囲は 2016 ~ 2016 です。the_stadium フィルターは 東京ドーム を保持します。

図3 東京ドームの観客動員数(2016年)

図4に東京ドームに観客数の分布のうち、曜日毎に色を変えたものを示す。図からわかるとおり、金、土、日に開催される試合の観客動員数は多い傾向にあり、逆に観客動員数の少ない試合はほとんどが平日に開催されていることが分かる。

観客動員数に影響を与えると思われる要因と試合の観客数の相関係数（今回はAICを用いた[2]）を調べた。その結果を表2に示す。球場、曜日、ホームチームとの相関関係が高いことがわかった。

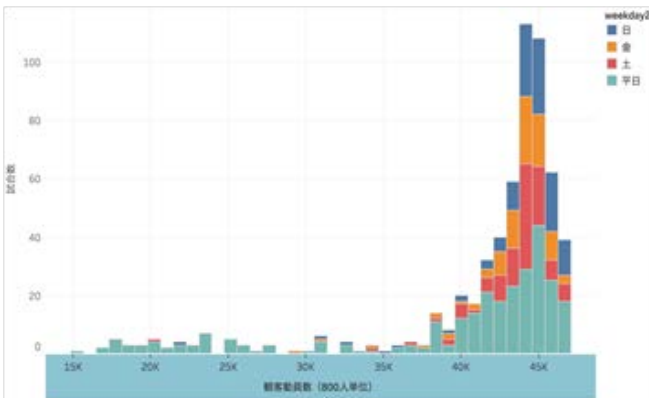


図4 東京ドームの観客動員数と開催曜日

表2 観客動員数と特徴量の相関関係

説明変数	AICの値
スタジアム、曜日2	-7771.16
スタジアム、時間帯	-7713.33
スタジアム、曜日2、ホームチーム	-7521.76
曜日2、ホームチーム	-7518.15
スタジアム、曜日2、時間帯	-7492.39
スタジアム、曜日	-7153.93
スタジアム、曜日、曜日2	-7153.93
スタジアム、ホームチーム	-6990.50
スタジアム	-6980.62
ホームチーム	-6384.24

3.4 プロ野球観客動員数の予測モデル

収容人数およびホームチームが異なるため、球場ごとに予測器を作成することとした。表1のデータを球場ごとに分けて使い、複数の要因を組み合わせ単純ベイズ器で学習させることとした。2009年から2015年までのデータを学習データとして使い、2016年の試合を予測してみた。その結果を表3に示す。試合の観客動員数を一万人単位で予測することとし、東京ドーム球場ではホームチームと曜日用いることで、最低でも60%の精度で動員数を予測できることが分かった。

図4から明らかなように、ほとんどの試合が4万人以上であり、4万人の試合の予測はできるが、一方で2万人代、3万人代の観客動員数は、試合数が限られているために予測は難しい。今回の混雑度では、F1値の大きなホームチームと曜日（曜日2は金曜日と土、日を区別している。曜日は金、土、日とそれ以外の曜日で区別している）を採用する。

表3 予測モデルの東京ドーム2016年のデータを用いた評価

説明変数		精度、再現率、F1			
ホームチーム、曜日2	20000-29999	4	0.6	0.75	0.67
	30000-39999	5	1	0.4	0.57
	40000-49999	67	0.96	0.99	0.97
ホームチーム、曜日	20000-29999	4	0.6	0.75	0.67
	30000-39999	5	1	0.4	0.57
	40000-49999	67	0.96	0.99	0.97
ホームチーム、時間帯 曜日2	20000-29999	4	0.5	0.75	0.6
	30000-39999	5	1	0.2	0.33
	40000-49999	67	0.96	0.99	0.97
ホームチーム、 ホームチーム順位	20000-29999	4	0.5	0.75	0.6
	30000-39999	5	1	0.2	0.33
	40000-49999	67	0.96	0.99	0.97

4. Jリーグの観客動員数の予測

本節では、サッカーを取り上げ、Jリーグの観客動員数の予測について考察する。

4.1 サッカーJリーグのデータ

サッカーの観客動員数はJリーグ本部に掲載されており、入手は容易である。また、Jリーグ自身が掲載しているため、データの正確性も高い表4にJリーグの観客動員データを示す。

表4 Jリーグデータ情報

	全データ	補足
期間	2009-02-28~ 2016-12-03	2009年-2016年リーグ ディビジョン1の全試合
スタジアム数	44	
試合数累計	2874	明治安田生命J1リーグ (Jリーグディビジョン1) 明治安田生命Jリーグチャンピオンシップ YBCルヴァンカップ ゼロックススーパーカップ
スタジアムごとの試合数 最小・最大	1~159	
観客数累計	47624929	
観客数 最小・最大	0 62632	無観客試合数: 1

4.2 サッカー観客数

図5に2009年から2016年までの2874試合のJリーグ観客動員数を示す。図からわかるように、サッカー観客数は数千人から4万人以上と変動が大きい。また、野球と異なり、サッカースタジアムの収容人数がそれ

ほど影響を与えておらず、2万人から3万人の間に山ができています。一方で、いくつかの試合は5万人を超えている。

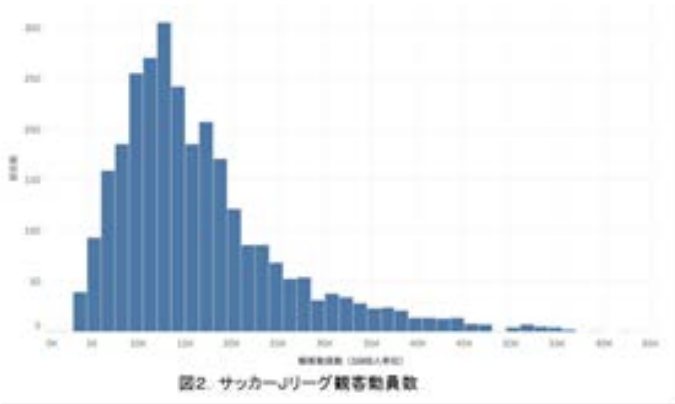


図5 Jリーグの観客動員数グラフ

4.3 サッカー観客動員変動の要因

図5に示したデータを用い、動員数に相関関係(AICを用いた[2])の高い特徴量を求めた。結果を表5に示す。

表5 サッカーにおける観客動員数と特徴量の相関関係

説明変数	AICの値
スタジアム、試合種別	-1897.91
スタジアム、試合種別、ホームチーム	-1763.83
試合種別、ホームチーム	-1713.63
スタジアム	-1618.88
スタジアム、ホームチーム	-1492.06
スタジアム、時間帯	-1371.05
スタジアム、試合種別、時間帯	-1297.46
ホームチーム、	-1270.20
スタジアム、曜日	-1114.28
スタジアム、ホームスタジアム同士の距離	-847.78

4.4 サッカー観客動員数の予測

開催スタジアム、試合種別、時間帯を用い、単純ベイズ器で学習した。この結果、66%の精度で動員数の予測が可能となった。現在、他の特徴量を用いて学習した場合の結果と比較を行っているが、サッカーの場合、5万人を超える試合が極めて数が少なく、クラブ決勝戦など特別な試合であるため、他の試合の学習からは良い結果が得られていない。

5. 混雑度予測システム・プロトタイプ

混雑度の予測において、単純にイベントの観客動員

数がわかっただけでは、十分ではない。本来であれば、大規模イベントが行われない日中、夜間などの平均滞在人数(例:混雑.comなどで提供されている情報)を元に検討が必要となる。残念ながら、本論文執筆時点において、東京の主要な場所における平均滞在人数などの統計情報は容易には手にはいかない。そこで、我々はプロトタイプとして、野球およびサッカースタジアム周辺を対象とした混雑予測システムを開発している。

(図6)

平均滞在人数としては、東京ドームが情報提供している後楽園の乗降人数(1日およそ11万人)から、イベントがない場合はおよそ5割の人が滞留していると想定する。乗降であるため、半分の人が下りて後楽園周辺に滞在し、半分の人が地下鉄などを乗ること仮定する。

この通常滞在人数に野球等のイベントが開催されると日常的な滞在人数に加えてイベントを目的とする人が増えることになる。野球が開催される場合、4万人以上の試合であれば「混雑++」、2万人以上の試合であれば「混雑+」試合が開催されない、周辺にイベントがないのであればいわゆる平均滞在人数となるため「混雑0」と提示する。また、時間単位の予測はイベントの動員数だけでは予測が難しいが、一方でイベント開催時間の予定はわかることが多い。この場合は、イベントの前後に「混雑注意」などの情報が付加することができる。(図7)

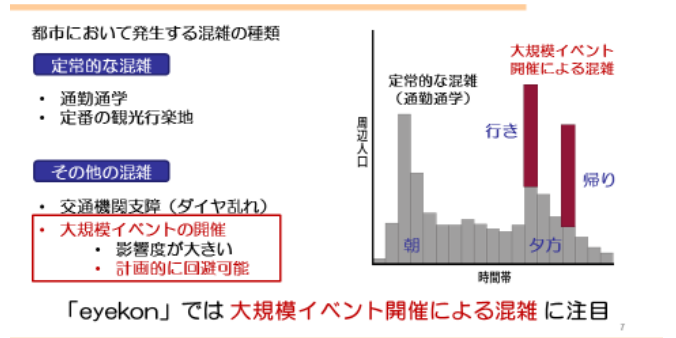


図6. 大規模イベントが関係する混雑度



図7. 混雑度表示画面

6. 関連研究

スポーツ観客動員に関しては、観客動員数を増やすなどの経営、広報的な視点による論文は出ている[4]が、混雑度と結びつけた研究は行われていない。また、混雑度という観点からは、携帯のGPSを用いたIoT的な現在の混雑度[3]、SNS情報を情報源とした集合知による現在の混雑度はあるが、混雑予測という観点からの研究は筆者の知るかぎりない。

7. まとめ

都市圏では、さまざまな理由から人が集中することで混雑が発生する。東京2020オリンピック・パラリンピックを控え、東京では大都市圏に特有の混雑に関し、様々な観点から対応が求められている。我々は、大都市圏における混雑に着目し、現在の混雑ではなく、大都市圏の住人にとっても訪問者にとっても有益な混雑予測情報を提供することで、大都市圏における移動、訪問場所の選択、スケジュール調整等への利便性を提供する。混雑度を絶対値ではなく、大規模イベントなどの情報を基にした相対値とすることで、天気予報と類似した混雑予測システムの開発を行っている。

今回は大人数の移動が想定されるスポーツ観戦を対象に、観客動員数の予測について検討を行った。プロ野球、サッカーを対象として、データを収集、解析を行い、ある程度の予測が可能であることを示した。今後は、スポーツ観戦数のみならず、コンサートイベント、祭りなどの大規模イベントのデータを収集することで、混雑予測システムとして精度をあげていく予定である。

参 考 文 献

- [1] <http://www.konderu.com/>
- [2] CATDUP <http://jasp.ism.ac.jp/ism/catdap/>
- [3] 西村 友洋, 樋口 雄大, 山口 弘純, 東野 輝夫: "スマートフォンを活用した屋内環境における混雑センシング", 情報処理学会論文誌 55(12), 2014年12月
- [4] 河合他, "J-リーグ観客数に影響を与える要因に関する研究", スポーツ産業学研究, Vol.18, pp.11-19, 2008