

会話情報から複数の発生事象の関係を推定する 運行状況案内システムの提案

米原 三揮[†] 川崎 健治[†] 山上 拓見[†] 白木 光彦[†] 長谷川 敬祐[†]

[†](株)日立製作所 〒100-8280

E-mail: [†]miki.yonehara.cs@hitachi.com

あらまし 鉄道を安全・安心に利用するためにはタイムリーな旅客案内が重要である。案内担当者は運行担当者から情報を収集、整理して運行情報を手入力をしているが、輸送サービス拡大や災害増加により運行状況が複雑化することで入力に時間を要している。本稿では案内担当者の業務負荷軽減を目的とした運行状況案内システムを提案する。本システムは運行管理者の談話を入力として、音声認識技術と意味構造化技術を用いて運行状況を出力する。意味構造化では、路線や駅の位置関係に着目して談話を数値化することで運行状況のパターンに対し少ないコーパスからも構造の学習と予測を可能にした。本システムのプロトタイプを開発し有用性を検証した。

キーワード 鉄道, Word, 信学技報, テンプレート

Traffic Status Navigation System based on Estimation of Relationship between Multiple Events in Conversation

Miki YONEHARA[†] Kenji KAWASAKI[†] Takumi YAMAGAMI[†] Mitsuhiro SHIRAKI[†] Keisuke HASEGAWA

[†]Hitachi, Ltd.

E-mail: [†]miki.yonehara.cs@hitachi.com

Abstract Timely provision of traffic status is becoming more and more important for passengers to use railway safely. Navigation operators get hung up on manual input of traffic status, because their transportation service is becoming larger and disasters increase. In this paper, we propose traffic status navigation system in order to reduce their work load. This system generates traffic status automatically from conversations between navigation operators and traffic operators, using voice recognition and semantic structure analysis. The system digitizes conversations based on locational relationship of lines and stations so that it can learn and estimate semantic structure from few corpus compared to the huge patterns of traffic status.

Keywords Windows, Word, Technical Report, Template

1. はじめに

鉄道を安全・安心に利用するためにはタイムリーな旅客案内がますます重要になっている。国土交通省の資料をもとに作成した 1989 年度～2018 年度の日本国内の輸送障害件数の推移を Fig.1 に示す[1]。この 30 年で 2 倍～3 倍程度に増加していることがわかる。

上記のような輸送障害も含めた運行状況を案内する業務への負担が増加している。案内担当者は運行管理者から情報を収集、整理して運行状況を手入力をしているが、輸送サービス拡大や災害増加により件数が増加、複雑化することで入力に時間を要している。

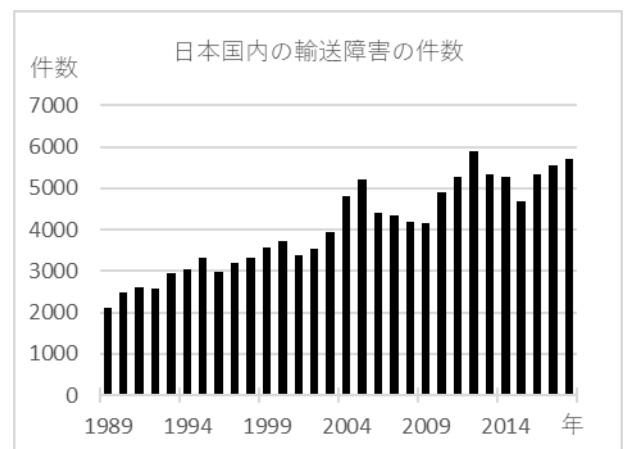


Fig.1 日本における輸送障害件数の推移

本稿では運行状況案内業務の業務負荷軽減を目的

とした運行状況案内システムを提案する。本システムは案内担当者と、運行担当者との会話を入力として、音声認識技術と構造化技術を用いて運行状況を生じ出力する。構造化技術では、文法ではなく発生した事象の内容、場所、時間の関係性に着目することで、文法のあいまいな会話からも精度よく運行情報を生成することを可能にした。特に、場所については路線上的関係性に着目したコーパスを作成することで、少量のデータからでも予測が可能になる。本システムのプロトタイプを開発し有用性を検証した。

本稿の構成を示す。2章では運行状況案内業務の概要と、本稿で対象とするシステム化の範囲、課題について述べる。3章では提案する運行状況案内システムを説明する。特に、特徴となるコーパスと、それによる意味的な構造的学習、予測方式について説明する。4章では評価について述べる。

2. 運行状況案内業務

2.1. 運行状況案内業務の概要

運行状況案内業務は鉄道の路線における運転見合わせ、遅延などの現状やその原因となった事故などを旅客にタイムリーに提供することである。タイムリーな運行状況の案内により、旅客の不満や不安を低減することができる[2]。

Fig.1に運行状況案内業務の概要を示す。業務は大きく以下の2つで構成される。

A) 案内担当者は運行状況を運行担当者などから集めて運行状況を整理し、運行状況の案内内容を決定する。運行状況は、案内を行う時刻、その時刻での運転状況、原因、を主とする。

運転状況としては、平常運転、遅延、運転見合わせ、等が一般的である。また、相互直通運転をしている路線では、相互直通運転中止、女性専用車や有料指定座席列車を運行している路線、時間帯では当該サービスの中止、等がある。また、付加情報として、運転見合わせ時には運転再開見込み時刻、他社線での振り替え輸送などの情報がある。

B) 案内担当者は状況を適切な手段で旅客に提供する。適切な手段は、従来は駅や車両での放送が主であった。そのため、指令室の案内担当者は駅係員や乗務員に対し電話、無線などを通じて運行状況を連絡する必要があった。近年では配信システムが導入され、駅や車両に設置されたサイネージ、旅客の所持するスマートフォン向けのアプリやWebサイト、SNSなど多様な提供手段が用いられている[3]。

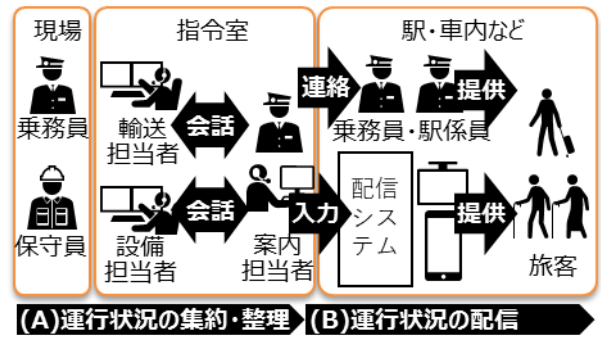


Fig. 2 運行状況案内業務の概要

2.2. システム化の要件

2.1 節で述べた業務 (A) (B) のうち、(B) はシステム化が進んでいる[3]。本稿では (A) をシステム化の対象とする。これを運行状況案内システムと呼ぶ。

運行状況案内システム (以下、本システム) は、案内担当者の業務負荷軽減を目的として、担当者間での会話を収集、分析し、案内すべき運行状況の情報を自動作成する。案内担当者は、自動作成された運行状況を確認して以下の作業を行う。

- (1) 自動作成された運行状況が問題ない場合はそのまま用いる。
- (2) 内容に一部問題がある場合は修正して用いる。
- (3) 必要な運行状況が得られない場合は手作業で新規に作成する。

並行して発生、変化していく複数の運行状況に対応する中で、担当者の会話には複数の運行状況に関する言及が混在している。そのような会話の中から、案内に関する情報を適切に抽出、構造化して運行状況を生じ出力することが求められる。

2.3. 課題

2.2 節で述べた要件に対し、談話構造解析、議論の構造化[4][5]の技術が適用できる。これらの技術には、

- (1) テキストに対し構文解析を適用することで構造化を図る方式、
 - (2) 語彙的結束性に着目して関係性を構築する方式、
 - (3) テキストに直接表れていない意味的な構造をコーパスとして作成して学習し、テキストから予測する方式、
- があるが、それぞれ運行状況案内業務への適用には問題がある。

従来技術 (1) (2) (3) の問題点は以下である。

- (1) この方式ではテキストに対し構文解析を行うが、運行状況案内業務は発話に対するものであり、構文解析が困難である。
- (2) 運行状況案内業務における談話では、語彙的な結束性が必ずしも高くない。例えば事故対応は車両担当者や駅担当者との談話に含まれ、路線の運転再開は路線の運行担当者との談話に含まれ、談話に現れる語彙に直接の関連が見られない場

合がある。

(3) 運行状況案内業務の意味的な構造をコーパス（以下、運行状況案内業務用コーパス）として作成するために必要な談話のデータ件数が不足する。運行状況は多数の路線、駅、運行状況、原因、等の組み合わせで膨大なパターンがあるためである。

以上のことから、本稿では案内担当者の談話から、運行状況のパターンに対して教師データが少ない場合にも構造を効率よく学習、予測できる方式の開発が課題である。

3. 運行状況案内システム

3.1. アーキテクチャ

本稿で提案する運行状況案内システム（以下、本システム）のアーキテクチャを Fig.3 に示す。業務中に動作する予測系と、予測モデルを最適化する学習系がある。予測系は指令室で会話がなされるたびに動作する。学習系は運用にもよるが月～年単位で運行状況案内業務用コーパスを更新するたびに動作し、予測モデルを最適化する。

予測系の入力データは指令室内で案内担当者と運行担当者との談話を収集した音声データである。出力データは運行状況である。機能は音声認識、事象自動作成、構成案自動作成で構成される。音声認識の出力はテキスト化された談話である。談話は発言者、発言単位で入力され、発言を開始した時刻と内容を文字列かしたテキストで構成される。事象自動作成は、談話単位で事象を作成する。構成案自動作成は、事象のリストから運行状況を作成し、案内担当者に提示する。案内担当者は、提示された運行状況を選択、修正して案内する内容を確定する。

学習系の入力データは指令室で収集された案内担当者や運行担当者の音声データである。出力は音声認識、事象自動作成、構成案自動作成それぞれの予測モデルである。機能は教師データ作成およびモデル最適化である。教師データ作成はデータ作成者向けの分析・データ作成画面を提供し、データ作成者が運行状況案内業務用コーパスを作成する。モデル最適化は運行状況案内業務用コーパスをもとに予測モデルを最適化する。

運行状況案内業務用コーパス（以下、コーパス）の構成を Fig.4 に示す。コーパスは以下の 4 項目から構成される。

- (1) 談話ごとに分割されたテキスト。図では矩形で区切った範囲が談話である。
- (2) テキスト中の運行状況に関連するタグ。タグは、事象の内容を示すタグ（事象タグ）、事象の発生場所を示すタグ（場所タグ）、に分類さ

れる。図では事象として異音検知、運転見合わせ、場所は B 線、B 線 A 駅、を例示している。

- (3) 事象タグと場所タグ間の構造関係を示すタグ（事象内構造タグ）。タグは、事象タグと場所タグの関係を定義する Where タグがある。図では、異音検知が B 線 A 駅で発生したことを表現している。
- (4) 事象間の関係を示すタグ（事象間構造タグ）。事象間構造タグは、原因と状況の関係を定義する Incident-Operation タグなどを定義する。

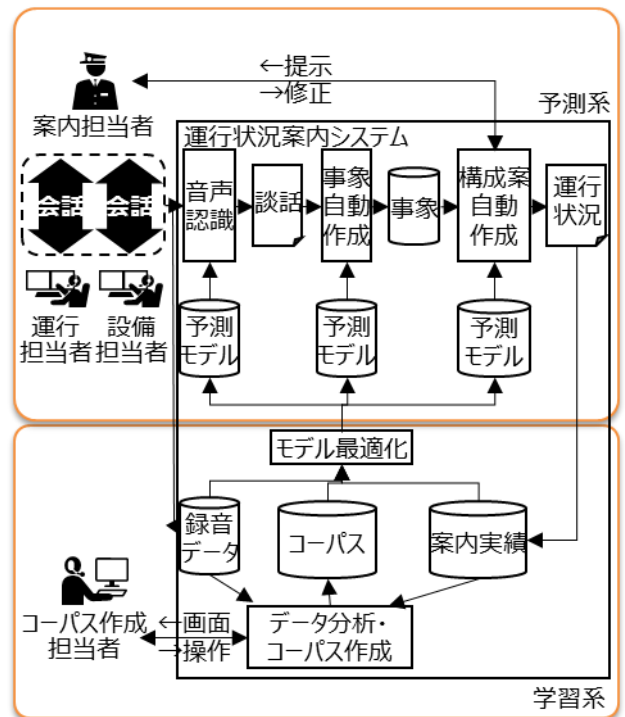


Fig.3 アーキテクチャ

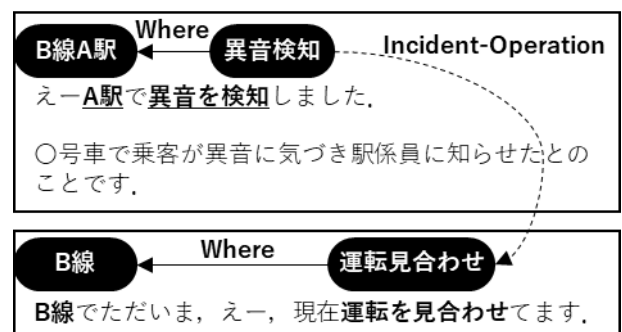


Fig.4 運行状況案内業務用コーパス

3.2. 事象自動作成

3.1 節で示したように、構造化には、事象自動作成機能が行う談話から事象への構造化と、構成案自動作成機能が行う事象から運行状況への構造化の 2 段階がある。談話から事象への構造化は、(1) 談話から事象タグ・場所タグの予測と、(2) 事象内構造タグの予測

により行う。

(1) の学習は談話と、事象タグ・場所タグのセットを用いた教師あり学習を行う。予測は、談話と学習済みの予測モデルを用いて事象タグと場所タグを予測する。談話の文字ごとにタグの有無や種別を付与する系列ラベリング問題であり、条件付確率場 (CRF) [6] などが適用できる。

(2) の学習は事象タグ・場所タグと、事象内構造タグを用いた教師あり学習を行う。予測は、事象タグ・場所タグと学習済み予測モデルを用いて事象内構造タグを予測する。事象タグと場所タグの組み合わせごとに事象内構造タグの有無や種別を付与する分類問題であり、サポートベクトルマシン (SVM) [7] などが適用できる。

3.3. 構成案自動作成

構成案自動作成の学習と予測を Fig.4 に示す。予測では、事象タグ、場所タグ、事象内構造タグと、事象間構造タグとを用いて教師あり学習を行う。予測では、事象タグ、場所タグ、事象内構造タグと予測モデルを用いて事象間構造タグを予測する。事象タグ、場所タグ、事象内構造タグの組み合わせごとに事象間構造タグを予測する分類問題であるため、SVM などが適用できる。

2 章で述べた通り、運行状況のパターンに対してコーパスが十分でない。図に示すように、B 線に関するコーパスはあるが、E 線に関するコーパスがない、などである。このようなコーパスで場所タグをそのまま用いて学習すると B 線に対し過学習となり、E 線での予測ができないと考えられる。

そこで、路線や駅の位置関係に着目して場所タグを数値化することで少ないコーパスからも学習を行う。例えば駅は路線の順序を考慮して数値化する。図に示すように B 線は A 駅を 1, C 駅を 2, D 駅を 3 と数値化して学習する。予測時にも学習時と同様の数値化を行うことで、コーパスにないパターンでも予測が可能になる。図では、E 線の F 駅, G 駅, H 駅も B 線と同様に 1, 2, 3 と数値化することで、B 線のコーパスで学習した予測モデルでの予測が可能になる。

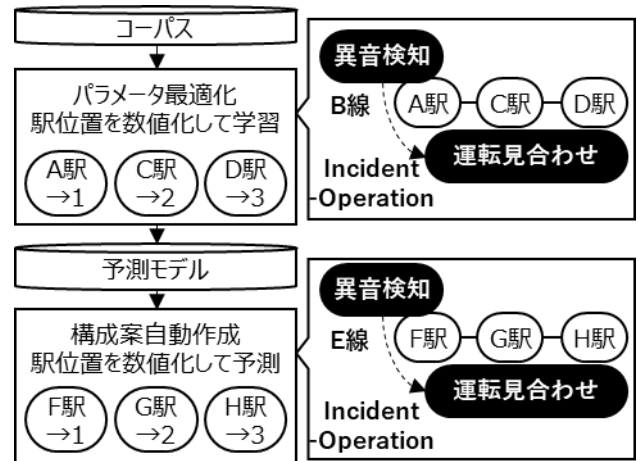


Fig.4 駅の位置関係にもとづく学習と予測

4. 評価

4.1. 評価用プロトタイプの開発

3 章で示したシステムの有用性を検証するために評価用プロトタイプを開発した。音声認識機能は汎用的な製品・サービスがあることから本プロトタイプでは音声認識機能は対象外と、事象自動作成機能と、構成案自動作成機能を開発した。開発言語は Python, 各タグの学習と予測には scikit-learn を使用した。コーパス作成には Brat Annotation Tool を用いた。ベースとなる談話は有識者へのヒアリングや各社の公開情報などを参考に模擬データを作成した。

4.2. 機能検証

4.1 節で説明したプロトタイプを用いて談話から運行状況への性能を評価した。CRF と SVM のハイパーパラメータチューニングを行い、予測モデルを最適化した結果を Fig.5 に示す。精度、再現率が 0.8 を超える結果となっていることがわかる。

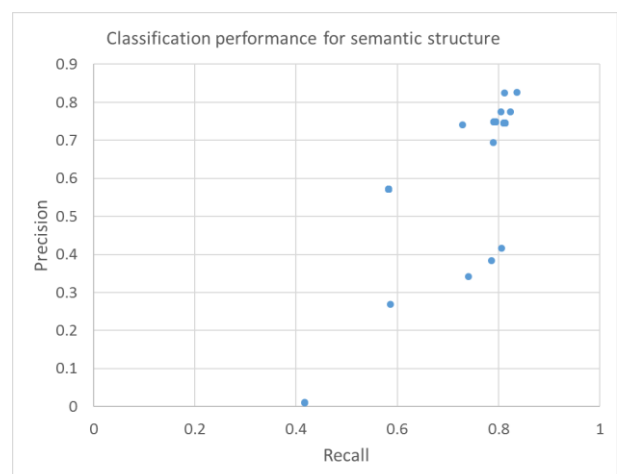


Fig.5 CRF, SVM のハイパーパラメータチューニングによる構成案の精度と再現率の変化

コーパスには同一路線、駅の談話は含まれていないため、3.3 節で示した路線や駅の数値化による効果が

大きいと考えられる。

4.3. 業務への適用に向けた課題

4.2 節で本システムが作成する運行状況の精度，再現率が 0.8 を超えることを示した。しかしながら業務への適用においては，以下の課題がある。

(1) 業務効率化への寄与の評価

今回の評価では精度，再現率といった指標でシステムの評価を行った。これらの指標が業務効率化にどのように寄与するかを評価することが今後の課題となる。例えば，今回は精度と再現率が両立するような結果となったが，業務においてはどちらかを優先したほうが業務効率化になることが考えられる。精度を優先すれば，案内担当者は本システムが提示した運行状況はそのまま配信し，不足している分のみを手入力すればよい。再現率を優先すれば，案内担当者は本システムが提示した運行状況の中から適切なものを選択して配信すればよい。現状では両方の作業が必要になるため作業が複雑化する可能性がある。業務効率化の観点に立った評価が必要である。

(2) 実データでの評価

今回の評価では模擬データを作成し評価した。有識者へのヒアリングなどをもとに作成したデータではあるが，実際の業務はさらに複雑であると考えられる。

5. 結論

本稿では鉄道の案内業務における担当者の業務負荷軽減を目的として，運行担当者と案内担当者の談話を収集し意味構造を学習・予測することで運行状況に変換する運行状況案内システムを提案した。パターンが膨大な運行状況に対し少ないコーパスからでも意味構造の学習・予測を可能とするため，路線や駅の位置関係を数値化する方式を提案した。プロトタイプと模擬データを用いて評価し精度・再現率が 0.8 を超えることを示した。今後は業務効率化への寄与評価，実データでの評価を行いシステムのブラッシュアップを推進していく。

文 献

- [1] 国土交通省鉄道局，“鉄軌道輸送の安全に関わる情報（平成 30 年度）”，Sep.2019
- [2] 山内香奈，“鉄道利用者のダイヤ乱れ遭遇時における不満足の規定要因”，心理学研究，第 83 巻，第 2 号，pp.117-125，Nov.2012
- [3] 加藤 寛四郎 稲坂 翔，“異常時案内用ディスプレイの高機能化について”，鉄道と電気技術，第 26 巻，第 4 号，pp.16-19，Mar.2015
- [4] 杉浦 純 井之上 直也 乾 健太郎，“説明生成に基づく談話構造解析の課題分析”，言語処理学会第 18 回年次大会 発表論文集，pp.115-118，Match 2012
- [5] 三浦 寛也 平田 圭二，“議事録生成技術に関する

サーベイ”，人工知能学会研究会資料 SIG-SLUD-B507-06，pp.34-38

- [6] Andrew McCallum，“Efficiently inducing features of conditional random fields”，Proceedings of 19th Conference on Uncertainty in Artificial Intelligence，pp.403-410 Aug.2003
- [7] Vapnik, V. “Pattern recognition using generalized portrait method”，Automation and Remote Control，24，pp.774-780，1963