

# テキストマイニングによる陸上無線技術士国家試験問題の内容分析

松田豊稔\*

## Content analysis of the national examination for Technical Radio Operator for On-the-Ground Services through a Text Mining Technique

Toyonori Matsuda\*

Content analysis of the national examination for Technical Radio Operator for On-the-Ground Services in Japan is conducted by using KH Coder which is a powerful free software for text mining: The technical terms necessary for leaning the national examination are extracted as frequent terms appeared in the examination sentences; A co-occurrence network of the technical terms extracted is obtained to clarify the systematic relationship among the technical terms. As a result, the useful information to support leaning of the national examination such as classification of the examination questions is provided.

キーワード：テキストマイニング、KH Coder、共起ネットワーク、陸上無線技術士

**Keywords** : text mining, KH Coder, co-occurrence network, Technical Radio Operator for On-the-Ground Services

### 1. はじめに

テキストマイニング<sup>(1)</sup>は、アンケートの回答や雑誌の記事などの大量なテキストデータの中から自動的に語句を抽出し、その抽出した語句に対して検索・集計、そして種々の統計手法やグラフ理論を用いた計量的な分析を行い、テキストデータが持つ特徴的なパターンや一定のルールなど有用な情報を取出す技術である。近年、様々なテキストマイニング用ソフトウェアが開発され、その信頼性の向上と高機能化に伴いテキストマイニングの有用性が認められ、現在、テキストマイニングが商業目的から学術分野まで広く利用されている<sup>(1)</sup>。

本稿では、テキストマイニング用のフリーソフトウェアの KH Coder 2<sup>(1)・(2)</sup>（以下、KH Coder）を用いて、第一級陸上無線技術士の国家試験科目「無線工学 B」の内容分析を行う。最初に、「無線工学 B」の試験問題の中から指定する語句を出現回数順に抽出し（これを抽出語と呼ぶ）、次にこの抽出語に対して共起性（同じ文の中に、語句が同時に出現すること）を求め、頻出語間の関連性を定量的に調べる。その分析結果から、試験問題に共通して出現する語句、また試験問題の各設問の特徴を示す語

句を抽出し、さらに抽出語間の関連性（共起性）をネットワーク図として視覚的に表示する。

本稿の調査結果は、陸上無線技術士国家試験の学習に必要な予備知識や専門分野（無線工学）の学習範囲を定量的に示し、試験問題の構成を把握するのに有用である。

### 2. 調査対象および調査方法

#### 2.1 陸上無線技術士と試験科目「無線工学 B」

陸上無線技術士は、無線通信に用いる設備の技術操作を行うための国家資格で、通信・放送業界をはじめ ICT 分野で需要が高い資格である。陸上無線技術士には技術操作の範囲により第一級と第二級があり、上位の第一級は大学卒レベルで、無線従事者資格の中で最も難関な資格と言われている。陸上無線技術士の国家試験は、「無線工学の基礎」、「無線工学 A」、「無線工学 B」、「法規」の 4 科目で構成され、年 2 回（1 月期と 2 月期）実施される。

本稿では、平成 25 年から平成 29 年に実施された第一級陸上無線技術士国家試験の科目「無線工学 B」の 10 回の試験問題に対してテキストマイニングによる内容分析を行う。「無線工学 B」の試験問題の設問数は 25 問で、出題形式はすべて多岐選択式である。「無線工学 B」の出題範囲は、(1) 空中線（アンテナのこと）系等の理論、構造及び機能の詳細 (2) 空中線系等のための測定機器の理論、構造及び機能の詳細 (3) 空中線系及び空中線系等のための測定機器の保守及び運用の詳細と定められている<sup>(3)</sup>。

高専や大学の電子通信系の学科では、「無線工学 B」は、

電子情報システム工学系  
〒861-1102 熊本県合志市須屋 2659-2  
Faculty of Electronics and Information Systems Engineering,  
2659-2 Suya, Koshi-shi, Kumamoto, Japan 861-1102

\* Corresponding author:  
E-mail address: tmatsu@kumamoto-nct.ac.jp (T. Matsuda).

「電気回路」や「電磁気学」など基礎科目を履修した後の高学年で「アンテナ工学」等の科目名で学習する。このように「無線工学B」の学習には、基礎科目の予備知識は必須であり、予備知識の内容を具体的に特定することは「無線工学B」の学習において重要である。

## 2.2 調査方法及び作業手順

本調査で行った KH Coder によるテキストマイニングの作業手順について述べる。

### (1) テキストデータ作成

平成 25 年から平成 29 年の「無線工学 B」の試験問題の pdf ファイルを入手し<sup>(4)</sup>、テキストファイル（以後、テキストデータと呼ぶ）に変換する。本調査では、試験問題の文章を構成する単語、語句、記号、式を分析の対象とし、図面はテキストデータから削除する。

KH Coder では、テキストデータを文・段落・節・章といった階層構造として認識させることができ、集計や分析の単位として指定することができる。本調査では、平成 25 年から平成 29 年まで実施された 10 回の各試験を「章」に、その中の各設問を「節」に対応させる。また、「段落」は改行により、「文」は句点により認識される。

### (2) 前処理（形態素解析<sup>(1)</sup>）

前処理では、KH Coder に取り込んだテキストデータに対して、文を品詞単位（名詞、動詞、助詞など）の単語で区切り、その語を品詞別に抽出する形態素解析を行う。本調査では、KH Coder に用意されている「茶釜（ちやせん）」<sup>(5)</sup>を用いて形態素解析を行った。

形態素解析において、ユーザ辞書として利用者が指定する語句の一覧（CSV ファイル）を登録すれば、その語句は“タグ”という品詞で認識され、集計や分析の単位として指定できる。本調査では、ユーザ辞書として“基礎用語”と“専門用語”の二つの辞書を作成した。“基礎用語”は「無線工学 B」を学習するのに必要な基礎科目で習う語句の一覧（後述の表 2 参照）であり、“専門用語”は「無線工学 B」で学習する専門用語の一覧（同表 3 参照）である。

## 2.3 抽出語の分析

形態素解析により求められた抽出語に対して、KH Coder では、種々の検索・集計、統計解析、グラフ理論などの分析手法が用意されている。以下に、本調査で行う分析手法を簡単に説明する。なお、KH Coder で利用できる分析手法の詳細については、文献(1)を参照されたい。

### (1) 抽出語リスト<sup>(1),(2)</sup>（頻出語の検索）

KH Coder のツール[抽出語リスト]は、全ての抽出語が頻度（出現回数）順に並べられた抽出語のリストを出力する。テキストマイニングにおいて語句の出現回数は重要なパラメータであり、抽出語リストからテキストデータの内容分析に有用な語句を把握することができる。

### (2) 共起ネットワーク<sup>(1),(2)</sup>（語と語の共起性）

抽出語間の関連性を表す指標として、一つの文中に同時に出現する抽出語を調べる共起性がある。KH Coder では、

共起性の尺度として、抽出語“X”と“Y”に対して次式で定義される Jaccard 係数

$$\text{Jaccard} = n(X \cap Y) / n(X \cup Y) \quad (1)$$

が用いられる。ここで、 $n(X \cup Y)$ は“X”か“Y”のどちらか一方が出現する文の数で、 $n(X \cap Y)$ は“X”か“Y”の両方が同時に出現する文の数である。共起ネットワークは、指定する抽出語の Jaccard 係数を求め、抽出語間の共起関係をネットワーク図にして表示したもので、抽出語間の関連性や系統性を視覚的に把握することができる。

### (3) 共起ネットワーク<sup>(1)</sup>（語と見出しの共起性）

KH Coder では、語と見出しの共起ネットワークも求めることができ、例えば、テキストデータの章を見出しとすれば、各章で特徴的な語句が調べられる。語と見出しの共起性の尺度を表す Jaccard 係数は、

$$\text{Jaccard} = n(N \cap X) / n(N \cup X) \quad (2)$$

で計算される。 $n(N \cap X)$ は第 N 章中で抽出語“X”が出現する文の数であり、 $n(N \cup X)$ は第 N 章の文の数を  $n(N)$ 、テキストデータ全体で“X”が出現する文の数を  $n(X)$ とすると  $n(N \cup X) = n(N) + n(X) - n(N \cap X)$ で計算される。語と見出しの共起ネットワークは、抽出語と見出しがどのように結びついているかを示している。

## 3. 調査結果および考察

### 3.1 テキストデータファイル

表 1 は、本調査で分析するテキストデータの章の内容を示す。平成 25 年 1 月期の「無線工学 B」の全問題（設問 25 問）を第 1 章に保存し、その見出し（章名）を“H25-01”とする。第 1 章は（試験問題の設問数が 25 問なので）25 節からなり、第 1 章中の段落数が 173、文の数が 211 である。以下、同様に平成 29 年 7 月期までの試験問題を第 2 章から第 10 章に保存し、テキストデータは全体で、章数 10、節数 250、段落数 1703、文の数 2147 となり、テキストファイルのサイズは 171KB である。

表 1 テキストデータの章の構成

| 試験        | 章    | 見出し    | 節   | 段落   | 文    |
|-----------|------|--------|-----|------|------|
| 平成25年度1月期 | 1    | H25-01 | 25  | 173  | 211  |
| 平成25年度7月期 | 2    | H25-07 | 25  | 169  | 205  |
| 平成26年度1月期 | 3    | H26-01 | 25  | 185  | 216  |
| 平成26年度7月期 | 4    | H26-07 | 25  | 165  | 211  |
| 平成27年度1月期 | 5    | H27-01 | 25  | 203  | 237  |
| 平成27年度7月期 | 6    | H27-07 | 25  | 154  | 230  |
| 平成28年度1月期 | 7    | H28-01 | 25  | 190  | 229  |
| 平成28年度7月期 | 8    | H28-07 | 25  | 165  | 201  |
| 平成29年度1月期 | 9    | H29-01 | 25  | 167  | 200  |
| 平成29年度7月期 | 10   | H29-07 | 25  | 132  | 207  |
| テキストデータ   | 全10章 |        | 250 | 1703 | 2147 |

### 3.2 頻出語検索

表 2 は、ユーザ辞書“基礎用語”で指定した語句の上位 20 位までの抽出語リストであり、各期の試験（H25-01 から H29-07）と全試験のトータル（H29-H25）で出現回数が多い

抽出語の上位 20 位までを示している。“基礎用語”の頻出語は、各試験での抽出語とその出現順位が類似している傾向があり、毎回の試験でだいたい同じ基礎用語が使われている。従って、表 2 に示される抽出語は、「無線工学 B」学習のための予備知識を表す用語と考えることができる。

表 3 は、ユーザ辞書の“専門用語”の語句に対する上位 20 位までの抽出語リストである。“アンテナ”と“電波”という言葉は「無線工学 B」の問題を作成する上で不可欠であり、全ての試験で用いられ、出現回数も多い。一方、3 位以下の抽出語は各期の試験で異なる傾向にあり、出現回数

も少ない。これは、各期の試験で出題される設問の内容が違っていることを示している。そこで、各期の試験を一定期間トータルした抽出語（H29-H25）は、「無線工学 B」の学習範囲を表す語句と考えてよい。

3.3 共起ネットワーク

図 1 は、テキストデータ全体（全 10 章）でユーザ辞書“専門用語”として抽出された語句の共起ネットワークである。円の一つ一つが抽出語を表し、その大きさは出現回数を、円を結ぶ線の太さは抽出語間の共起性の尺度を示す。また、抽出語の中心性（他の抽出語と多く結びついている

表 2 ユーザ辞書“基礎用語”で指定した語句の抽出語リスト

| 順位 | H25-01  | H25-07  | H26-01  | H26-07  | H27-01  | H27-07  | H28-01  | H28-07  | H29-01  | H29-07  | H29-H25 | 出現回数 |
|----|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|------|
| 1  | 利得      | 周波数     | 波長      | 波長      | 波長      | 利得      | 雑音      | インピーダンス | 波長      | 波長      | 波長      | 206  |
| 2  | 波長      | 波長      | 周波数     | 電力      | 利得      | 波長      | 利得      | 波長      | 周波数     | 周波数     | 周波数     | 139  |
| 3  | 損失      | 雑音      | 電力      | 磁界      | 電界強度    | 電力      | 波長      | 電流      | インピーダンス | 利得      | 利得      | 139  |
| 4  | 電力      | インピーダンス | 電流      | インピーダンス | インピーダンス | 周波数     | インピーダンス | 電力      | 利得      | 電力      | インピーダンス | 132  |
| 5  | インピーダンス | 利得      | 導体      | 電界      | 帯域幅     | インピーダンス | 周波数     | 帯域幅     | 位相      | 導体      | 電力      | 131  |
| 6  | 電界強度    | 位相      | インピーダンス | 電流      | 導体      | 損失      | 電界強度    | 磁界      | 電力      | インピーダンス | 電界強度    | 85   |
| 7  | 周波数     | 電界      | 雑音      | 周波数     | 電力      | 位相      | 電界      | 電圧      | 電界強度    | 電界強度    | 雑音      | 70   |
| 7  | 導体      | 電力      | 電圧      | 電界強度    | 周波数     | 金属      | 電力      | 利得      | 電流      | 位相      | 電流      | 70   |
| 9  | 真数      | 金属      | 電界強度    | 利得      | 真数      | 出力      | 金属      | 周波数     | 帯域幅     | 真数      | 電界      | 60   |
| 10 | 帯域幅     | 真数      | dB      | 抵抗      | 磁界      | 真数      | 電流      | 効率      | 雑音      | 振幅      | 磁界      | 58   |
| 10 | 電界      | 磁界      | 位相      | 効率      | 誘電体     | 電流      | 導体      | 出力      | 出力      | 電流      | 導体      | 58   |
| 12 | dB      | 抵抗      | 利得      | dB      | 位相      | 効率      | 磁界      | 電界強度    | 誘電体     | 抵抗      | 位相      | 56   |
| 13 | 雑音      | 電流      | 磁界      | 屈折率     | 雑音      | 電界      | 出力      | 位相      | dB      | 誘電体     | 真数      | 47   |
| 14 | 磁界      | 電子      | 出力      | 電圧      | 電界      | 電圧      | 真数      | 屈折率     | 抵抗      | dB      | 損失      | 41   |
| 15 | 位相差     | 比誘電率    | 電界      | 位相差     | フィルタ    | 電界強度    | 電圧      | 電界      | 電界      | 位相差     | 帯域幅     | 38   |
| 16 | 回路      | フィルタ    | 誘電体     | 損失      | 周期      | dB      | 無損失     | 回路      | 比誘電率    | 屈折率     | 抵抗      | 36   |
| 17 | 周期      | 周期      | 回路      | 導体      | dB      | 回路      | 誘電率     | 無損失     | 屈折率     | 電圧      | 出力      | 34   |
| 17 | 負荷      | 誘電体     | 屈折率     | 負荷      | 位相差     | 磁界      | インダクタンス | 周期      | 位相角     | 電子      | 電圧      | 34   |
| 19 | 誘電体     | インダクタンス | 周期      | 無損失     | 逆位相     | 抵抗      | 位相      | 導体      | 回路      | 比誘電率    | dB      | 32   |
| 20 | 効率      | リアクタンス  | 振幅      | アンペア    | 抵抗      | 周期      | 周期      | 抵抗      | 磁界      | 効率      | 誘電体     | 31   |

表 3 ユーザ辞書“専門用語”で指定した語句の抽出語リスト

| 順位 | H25-01  | H25-07  | H26-01  | H26-07  | H27-01  | H27-07  | H28-01 | H28-07  | H29-01  | H29-07   | H29-H25 | 出現回数 |
|----|---------|---------|---------|---------|---------|---------|--------|---------|---------|----------|---------|------|
| 1  | アンテナ    | アンテナ    | アンテナ    | アンテナ    | アンテナ    | アンテナ    | アンテナ   | アンテナ    | アンテナ    | アンテナ     | アンテナ    | 400  |
| 2  | 電波      | 電波      | 電波      | 給電線     | 電波      | 電波      | 電波     | 給電線     | 電波      | 電波       | 電波      | 158  |
| 3  | 受信機     | 逆 F 形 ↑ | 指向      | 電波      | 受信機     | 指向      | 導波管    | 開口面     | 受信機     | モノポール ↑  | 給電線     | 63   |
| 4  | 絶対利得    | 導波管     | 反射板     | 導波管     | 送信アンテナ  | 半波長ダイ ↑ | 反射板    | 電波      | 指向      | 受信アンテナ   | 指向      | 57   |
| 5  | パラボラ ↑  | 送信アンテナ  | フェーシング  | モノポール ↑ | 回転放物面   | 絶対利得    | 給電線    | 給電点     | 半波長ダイ ↑ | 絶対利得     | 導波管     | 56   |
| 6  | 指向      | 減衰      | 給電線     | 電磁波     | 開口面     | パラボラ ↑  | 指向     | 反射板     | 給電線     | 実効面積     | 半波長ダイ ↑ | 48   |
| 7  | 自由空間    | 吸収      | 減衰      | 電波      | 実効面積    | フレネルゾーン | 電磁波    | フェーシング  | TE      | 同軸給電線    | 絶対利得    | 45   |
| 8  | 受信アンテナ  | 実効面積    | 散乱      | 微小ダイ ↑  | 受信アンテナ  | モード     | 電波吸収体  | 逆 F 形 ↑ | バラン     | 一次放射器    | 受信アンテナ  | 44   |
| 9  | 送受信     | 受信アンテナ  | 自由空間    | 指向      | 給電線     | 自由空間    | 媒質     | 散乱      | フェーシング  | 送信アンテナ   | 反射板     | 44   |
| 10 | 送信アンテナ  | 半波長ダイ ↑ | 定在波     | 絶対利得    | 指向      | 受信アンテナ  | 励振     | 干渉      | 回転放物面   | アンテナ素子   | 自由空間    | 43   |
| 11 | 導波管     | 偏波      | 電波吸収体   | 放射抵抗    | 絶対利得    | 送信アンテナ  | 一次放射器  | 屈折率     | 減衰      | 給電線      | 受信機     | 40   |
| 12 | 円偏波     | 放射素子    | 電離層     | TE      | 偏波      | 反射板     | 自由空間   | 自由空間    | 自由空間    | 送信電力     | 送信アンテナ  | 39   |
| 13 | 電磁波     | TE      | 半波長ダイ ↑ | モード     | 一次放射器   | 同軸給電線   | 主反射鏡   | 定在波     | 地板      | 定在波      | 偏波      | 39   |
| 14 | 偏波      | デジタル移相器 | 平面波     | 給電点     | 自由空間    | 偏波      | 受信機    | 導波管     | 導波管     | 半波長ダイ ↑  | 電磁波     | 36   |
| 15 | フレネルゾーン | 開口面     | バラン     | 散乱      | 垂直偏波    | TE      | 反射係数   | 半波長ダイ ↑ | 偏波      | アンテナ利得   | 開口面     | 30   |
| 16 | 半波長ダイ ↑ | 給電点     | 干渉      | 主反射鏡    | 導波管     | 減衰      | VSWR   | 電磁波     | デジタル移相器 | インテリジェント | 電離層     | 30   |
| 17 | 同軸給電線   | 送信アンテナ  | 地板      | 電離層     | 半波長ダイ ↑ | 受信電力    | ビーム幅   | 電離層     | パラボラ ↑  | ハイパターンの  | 給電点     | 28   |
| 18 | 減衰      | 短絡      | TE      | 副反射鏡    | 反射板     | 送受信     | 指向性利得  | 反射係数    | モード     | パラボラ ↑   | 減衰      | 28   |
| 19 | 指向性利得   | 電磁波     | マイクロ波   | フェーシング  | 表皮厚さ    | 電離層     | 実効面積   | フラウン ↑  | 回折      | フェーシング   | 散乱      | 28   |
| 20 | 受信電界強度  | 電離層     | モード     | 一次放射器   | 平面波     | 半値幅     | 主ビーム   | マジック T  | 干渉      | モード      | パラボラ ↑  | 27   |
| 20 | 受信有能電力  | 微小ダイ ↑  | 回折      | 遠方界     | フェーシング  | 反射係数    | 受信アンテナ | 印加電流    | 管内波長    | 屈折率      | フェーシング  | 27   |
| 20 | 大地反射波   | 平面走査法   | 給電線     | 近傍界     | マイクロ波   | デジタル移相  | 電力束密度  | 屈折      | 屈折率     | 携帯機      | 送受信     | 27   |

半波長ダイ ↑: 半波長ダイポールアンテナ  
 フラウン ↑: フラウンホーファ領域  
 逆 F 形 ↑: 逆 F 形アンテナ  
 微小ダイ ↑: 微小ダイポール  
 パラボラ ↑: パラボラアンテナ  
 モノポール ↑: モノポールアンテナ



