

テキストマイニングを用いた航空機騒音の 苦情の原因の同定

森長 誠¹・月岡 秀文²

¹非会員 防衛施設協会（〒105-0014 東京都港区芝3-41-8）

E-mail: morinaga@dfeia.or.jp

²非会員 防衛施設協会（〒105-0014 東京都港区芝3-41-8）

E-mail: tsukioka@dfeia.or.jp

空港や自衛隊等の飛行場周辺で発生している航空機騒音の問題解決に向けた情報共有の手法の提案として、航空機騒音に関する苦情をテキスト解析するためのデータベース作成と、それらを使用した苦情の原因同定に関する分析を試算的に実施した。ケーススタディとして沖縄県宜野湾市がWeb上で公開している苦情を利用したデータベースを作成し、航空機騒音に関する苦情の主成分を抽出した。それらを利用した分析の結果、苦情の原因となっている物理的要因と影響の関連性を確認することができ、それらの原因の地理空間的な分布を確認できる可能性を示した。本稿は、苦情データをテキスト解析するこれらのプロセスについて、実例を踏まえながら報告したものである。

Key Words : aircraft noise, complaint, text mining

1. 研究の背景

空港や自衛隊等の飛行場周辺では航空機騒音が大きな環境問題の一つとなっており、地域住民から自治体や飛行場施設に対して苦情の申し入れがなされている。多くの場合、苦情は電話や電子メール等を通じて申し入れられ、苦情を受けた自治体や施設ごとに記録・管理がなされている。これらの苦情を今後の対策にどのように結びつけば良いか、その活用方法について多くの自治体が思案しているが、一般的には、受け付けた苦情の件数を公開する等、基礎的な情報の呈示にとどまっている。電話や電子メールで受け付けた苦情をテキスト化し苦情の発生日や時間帯などの情報とセットでホームページ上に公開している自治体もあるが、これらがテキストマイニングなどの手法で統計的に整理された構造化データとして公開されれば、苦情の内容について定量的かつ具体的な情報を住民に周知できると期待される。また、苦情の発生位置・発生時期など、飛行経路等の客観的なデータとの関係を定量的に把握することができれば、苦情の発生源の同定につながるため、自治体サイドはそれらの情報をもとに対策の優先順位を付ける、あるいは、より具体的な対策を検討することも可能となる。発生源サイドにとっては、これらの情報公開が騒音低減のモチベーションにつながる可能性もある。

2. 目的

自然言語で書き下されたデータを用いたテキスト解析を行う際に大きな障壁となるのは、同じ内容を意味する語彙であっても、発言の仕方によって表現方法が異なる語彙の“ゆらぎ”の事前処理に膨大な時間と労力を有するという点が挙げられる。本研究は、苦情のテキスト解析を実施することで、飛行場周辺における苦情原因の同定を行うことを目的としているが、本稿では最初に、航空機騒音に関する苦情データを用いた形態素解析を行うと共に、語彙のゆらぎの確認等を行い航空機騒音に関する苦情の語彙のデータベース化を試みた。この作業によって、今後、新たに苦情データを追加した際に、それらの作業を省力化することが可能となる。また、その結果を用いた多変量解析を実施し、航空機騒音の苦情に関する語彙やそれらの共起関係の整理、地理情報システム（GIS）との連動などによる苦情の原因同定のプロセスを提案し、今後の課題について整理した。

3. 航空機騒音の苦情の語彙のデータベース化

インターネット上で苦情のテキスト情報を公開している事例は少なく、今回は沖縄県宜野湾市の航空機騒音に関する苦情のデータ¹⁾のみを使用して、航空機騒音に関

する苦情の語彙の構造化を行った。データ数は2005年～2016年までの苦情、計2,854件である。苦情は24時間の電話対応とE-mailでの受付となっており、電話対応の場合は、受け付けた市の職員がその内容を手入力で記載していると考えられるため、苦情者の発言と一字一句、整合しているとは限らない。データは苦情内容の他に、苦情の受入日・時刻、苦情者の性別、居住者の居住地域の情報が含まれているが、特に居住地域に関しては、必ずしも全ての居住者についての情報が含まれているわけではない。また、町丁目単位まで含まれている場合もある。なお、苦情者の個人を同定できる情報は含まれていないため、同一苦情者による複数回の苦情であるか否かの判断は不可能である。分析にはIBM SPSS Text Analytics for Surveys 4を用いた。

最初に、形態素解析によって抽出された語彙の頻度解析と係り受け解析を行い、頻度解析で上位に位置する名詞、動詞、形容詞、形容動詞を抽出し、これらを分類の一次カテゴリとした。頻度解析を行う際、一つの苦情の中に同一の語彙が繰り返し抽出されてもダブルカウントしないこととした。次に、頻度解析で必ずしも上位でなかった語彙を個々に確認し、一次カテゴリと同義の語彙が存在する場合は、それらを同義の一次カテゴリに含めた（二次カテゴリ）。ここでは係り受け解析の結果や苦情本文を確認しながら、文脈全体の中で同義であることを確認した。最終的に、一つのカテゴリに含まれる様々な表現の中で代表的と考えられる語彙でラベリングを行った。全2,854件の苦情の中で、どのカテゴリにも属さない苦情は5件であり、これらは内容が明確でない苦情であることを確認した。

表-1に、カテゴリの構造と出現頻度の結果を示す。苦情が発生するまでのプロセスとして、航空機の飛行による騒音の発生など、「物理的な入力」があり、それらによってうるささなどの「影響が発生」し、その結果として「苦情が発生」とするというメカニズムを想定し²⁾、カテゴリは航空機の種類や飛行形態など苦情の物理的要因と考えられるカテゴリと、うるささや睡眠影響等、苦情の原因となっている影響のカテゴリに大きく分類した。宜野湾市は米軍の普天間飛行場を有する自治体であり、ここでの苦情はほぼ、普天間飛行場を離着陸する航空機に対する苦情であると考えられる。普天間飛行場の常駐機は主としてヘリコプタであり、このことは苦情の原因となっている機種の中でヘリコプタの苦情件数が最大となっていることと整合している。飛行形態を見ると、夜間飛行や低空飛行などの件数が多い。影響についてみると、単にうるさいという苦情の他に、墜落の危険性を感じる苦情や、睡眠影響なども見受けられ、また、音以外の要因として、テレビの電波障害（航空機の飛行による

電波の受信障害に伴う画像の乱れ）に関する苦情も多くなっている。これらは単純集計の結果であり、以降では、物理的要因と影響の関係性など、語彙の共起関係についての検討、さらに、これらの語彙の地域性の有無などについて検討を行う。

表-1 カテゴリの構造と出現頻度

	カテゴリの大分類	カテゴリ	件数
物理的要因	機種	ヘリコプタ	1303
		米軍機	428
		オスプレイ	330
		ジェット機	158
		戦闘機	119
		プロペラ機	47
		その他	431
	飛行形態	飛行している	1431
		夜間飛行	623
		低空飛行	251
		旋回	223
		飛行ルート	71
		タッチアンドゴー	36
		編隊飛行	28
		ヘリモード	8
	地上騒音		113
	低い音・低周波音		18
影響	心理的影響	アノイアンス(うるささ)	1293
		墜落	108
		不安	49
		精神的苦痛・ストレス	48
		驚愕	43
		恐怖	41
		危険	32
	睡眠影響		214
	聴取妨害	テレビ聴取妨害	111
		会話妨害	27
	身体的影響	体調不良	46
		心拍	40
		頭痛	28
		難聴	7
		吐き気	6
		耳鳴り	5
		その他	3
	テレビ電波障害		489
	がたつき・振動		87
	子供への影響		179

4. 苦情原因と考えられる物理的要因と影響の共起性

全2,854件の苦情について、表-1の「その他」を除く35カテゴリを変数とした主成分分析を行い、カテゴリ同士の共起関係を考察した。具体的に、個々の苦情データについて、それぞれのカテゴリに属する場合は1、属さない場合は0とするダミーコーディングを行い相関係数行列による分析を行った。分析ソフトはRを用いた。表-2に分析結果を示す。固有値1以上の主成分（Principal Component: PC）は13種類であり、それらに含まれる主なカテゴリを物理的要因と影響に分けて整理している。これら13の主成分による累積寄与率は50.1%であった。

表-2 主成分分析で抽出された各成分に含まれるカテゴリ

	物理的要因	影響
PC 1	ヘリコプタ	うるささ 電波障害
PC 2	ジェット機 戦闘機	驚愕 恐怖 睡眠影響 子供
PC 3	編隊飛行	墜落 不安
PC 4	ヘリコプタ 低空飛行 旋回	がたつき
PC 5	プロペラ機	タッチアンドゴー
PC 6	オスプレイ ヘリモード	がたつき
PC 7	—	頭痛 吐き気
PC 8	低周波音	ストレス 難聴 がたつき
PC 9	ジェット機 夜間飛行	睡眠影響
PC 10	—	聴取妨害
PC 11	戦闘機 編隊飛行 飛行ルート	—
PC 12	地上騒音	体調不良
PC 13	ジェット機	耳鳴り

これらを見ると、それぞれの主成分には物理的な要因とその影響が同時に含まれているケースが多く、苦情原因となっている物理的要因とその影響の組み合わせを推測できる結果となっていることが分かる。PC 1～PC 13は寄与率の大きい順であり、ヘリコプタによるうるささ、電波障害の苦情が多いと考えられる。またPC 2では、ジェット戦闘機による驚愕や恐怖、睡眠影響に関連しており、これらが子供に及ぼす影響に対する苦情も含まれているようである。普天間飛行場の常駐機はヘリコプタ3機種（AH-1, CH-53, UH-1）とティルトロータ機（MV-22）、プロペラ機（C-12）、小型のジェット機（UC-35）の合計6機種であり、ヘリコプタに関する苦情の寄与率が高いことは当然の結果であるが、外来機としてのジェット戦闘機の飛来も行われており、そのような偶発的な大きい騒音レベルによる飛行が驚愕や睡眠影響を行った苦情の原因となっていることも示唆される。この他にも編隊飛行による墜落等への不安やオスプレイのヘリモードによる飛行に関連したがたつきなど、より具体的な情報を簡潔に確認することができる。ただし、各主成分に含まれるカテゴリは必ずしも共起していることを意味しないため、特に表-1で発生件数が少ないカテゴリ

については、物理的要因と影響の関連が共起していないケースもあると考えられる。この課題については変数間のJaccard係数による階層的クラスタリングなども考えられるが、一般的に階層的クラスタリングは用いる連結アルゴリズムの違いが結果に影響しやすいことなどが知られているため、今回は主成分分析による試算を行った。

5. 地域ごとの苦情原因の特徴

主成分分析で得られた各苦情ごとの主成分得点を用いて、宜野湾市の小地域区分における苦情の特徴を考察する。既に述べたとおり、全ての苦情について地域情報が紐づけられている訳ではないため、ここでは、町単位までの情報が含まれている苦情に限定して分析を行う（N=2,144）。PC 1～PC 13の各主成分の得点を町単位で合計し、この合計値を平成27年度国勢調査³⁾による町別の人口で規準化した。この規準化した得点の大小をGIS上におけるポリゴンの色の濃淡で表現した結果を図-1と図-2に示す。ここではPC 1～PC 13の結果の内、例としてPC 1とPC 2の結果を示している。図-1のPC 1（ヘリコプタ、うるささ、電波障害）の結果を見ると、飛行場の北東部で主成分得点が大きくなっていることが分かる。一方で図-2のPC 2（ジェット機、戦闘機、驚愕、恐怖、睡眠影響、子供）の結果では、飛行場の南西部で得点が大きくなっている。那覇気象台における2005年～2016年の平均データによると⁴⁾、年間のおよそ65%は北～東よりの風であり、35%は南～西よりの風となっている。航空機の運用特性として風に向かって離着陸することに鑑みれば、飛行場の北東側は主として離陸に関連する地域、南西側は着陸に関連する地域であることが推測される。このような運用の特徴が苦情の違いにも表れているかもしれない。例えば、PC 2には驚愕や恐怖といった影響に関する苦情に基づく主成分である。このことは、飛行場

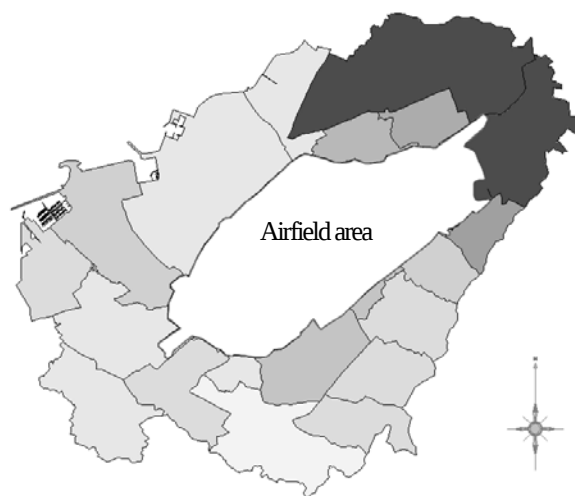


図-1 主成分1(PC 1)の得点の地域分布

南西部における着陸直前の非常に高度の低い飛行に関連している可能性がある。

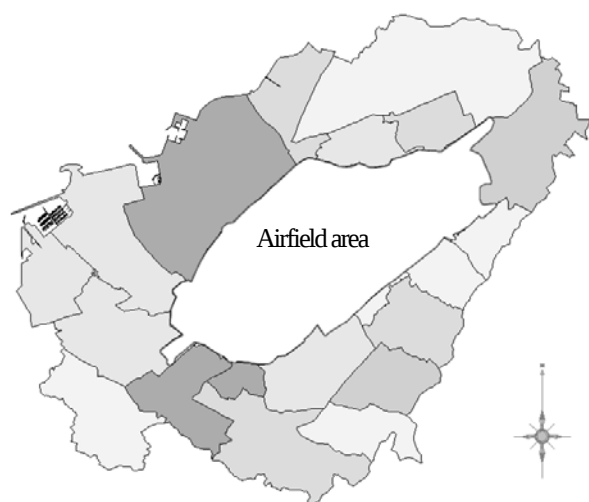


図-2 主成分2(PC2)の得点の地域分布

5. まとめ・今後の課題

自治体等に寄せられる航空機騒音に関する苦情の有効利用の提案として、苦情原因の同定に関するアプローチをケーススタディとして示した。今回は宜野湾市のホームページで公開されているデータを用いて、異なる表現による同義語の苦情を整理する等、航空機騒音の苦情に関するキーワードのデータベース化を行った。当然ながら、民間空港や自衛隊等飛行場の違いなどによって苦情に用いられる固有名詞等が異なると考えられ、今回作成したデータベースのみで汎用的な活用ができる訳ではない。今回は基本的な部分でのデータベース化と考えており、引き続き、他の空港・飛行場の事例を収集して更新していく予定である。

本稿では、苦情の原因の同定を目的として、キーワードの共起性の分析や地域性の分析を試算的に実施した。今回用いたような主成分分析等の多変量解析を行うに当たって、モデルの妥当性を高めていくためにはデータの

蓄積が必要不可欠である。このためにも更なる事例の収集が必要であると考えている。苦情受付のシステム化も重要であると考えている。今回の宜野湾市の事例のように、Web上で苦情データを公開しているケースは非常に少ないという課題が一番であるが、苦情者の地域情報が完全には把握できなかったという課題もある。諸外国の空港では苦情受け付けのフォームがWeb上に作成されている事例が見受けられる（例えば、サンフランシスコ空港⁵⁾）。このような取り組みによって、苦情に関する時間的・空間的な情報を体系的に入手し、それらが自動的にテキスト化されて保存される仕組みがあれば、苦情の解析がより行いやすいのではないかと考えている。

一般的に、航空機の運用サイドは、苦情が発生していることは把握していても、その詳細は分かっていないと考えられる。これらの発生源サイドに対する要請として、こういった地域で、どのような行為に対する影響が発生して苦情につながっているのか、これらの情報を共有することが重要である。

参考文献

- 1) 宜野湾市ホームページ：基地被害 110 番，
http://www.city.ginowan.okinawa.jp/sisei/base/05/kichihigai110ban_499.html
- 2) 矢野隆，横島潤紀，森長誠：環境騒音に関するアノイアンス，苦情，対処の関係，2017 年 3 月 騒音・振動研究会資料，N-2017-12，2017.
- 3) E-Stat：平成 27 年度国勢調査，小地域集計，沖縄県，
http://www.e-stat.go.jp/SG1/estat/GL08020103.do?_toGL08020103_&tc=las-sID=000001082889&cycleCode=0&requestSender=estat
- 4) 気象庁ホームページ：
<http://www.data.jma.go.jp/obd/stats/etrn/index.php?sess=6ef525a9cdef28cea634ce58ca736e68>
- 5) SFO Aircraft Noise Office：
<https://www.flysfo.com/community/noise-abatement/file-a-complaint>

(2017. 8. 25 受付)

TEXT MINING OF AIRCRAFT NOISE COMPLAINTS FOR IDENTIFYING THE CAUSES

Makoto MORINAGA and Hidebumi TSUKIOKA

Aircraft noise is one of the serious environmental issues in the vicinities of airport and/or airfields. Using the reports of aircraft noise complaints from the communities, which are described by natural language, we have been developing the database of aircraft noise complaints to identify the causes of the complaints. In this paper, text mining was conducted, for a case-study, using the complaining report published in the web site by Ginowan city in Okinawa prefecture and the tentative database was shown.

Additionally, co-occurrence between the keywords including the tentative database was examined by the principle component analysis (PCA) and the area characteristics of aircraft noise complaints were examined by the distribution of the PCA score in the subregion block in the city.