

不動産情報検索支援のための情報推薦

大坪五郎

株式会社 LIFULL

ohtsubogoro@lifull.com

概要 不動産情報検索は、多くのユーザにとって馴染みのないタスクである。そのため検索に便利な「こだわり条件」について知られておらず、せっかくの機能が十分に活用されているとは言えない。そのため、音声認識技術を使ってユーザと不動産エージェントの会話からこだわり条件を推薦するシステムを試作した。そのアルゴリズム、インターフェースおよび初期段階の評価結果について述べる。

キーワード Web インタフェース, 情報探索, 不動産物件探索

1 はじめに

インターネットの普及に伴い、Web サイト上で不動産物件情報を検索することが可能となった。しかしながら日本人は生涯の間に平均6回程度しか転居しない。つまり不動産物件を探索するというタスクはほとんどのユーザにとって馴染みのないものである。ある不動産情報サイトでは、「ペット可」「楽器可」などのこだわり条件が71種類使用可能であり、ユーザニーズに合わせて利用すれば効率的に情報探索を行うことができる。しかしながらほとんどのユーザはその存在を知らず、結果として有効に使われているとは言い難い。

そこで本研究ではユーザにこだわり条件を自動で推薦することを目的とする。この際入力としてユーザと不動産エージェント間の会話を用いる。近年音声認識により動作可能なスマートスピーカーが広く使われるようになってきている。しかし現状の音声認識、対話制御技術は人間の不動産エージェントを完全に置き換えるには程遠い。そのため、本システムはユーザと不動産エージェントの脇に位置し、会話をモニターする。その会話の中に、こだわり条件につながる言葉が含まれていた場合にユーザに対して関連するこだわり条件を推薦することを目標とする。

2 システム概要

開発したシステムの画面を図1に示す。ユーザはあらかじめGUIを用いて、概略の検索条件を設定する。その上で自由にエージェントと会話を行う。システムは会話を継続的にモニターする。システムが音声認識したテキストは画面の下部に表示される。システムは認識したテキストから自動的に「コンビニ800m以内」、「ペット可」などのこだわり条件を検索、表示する。またそれらのこだわり条件を設定した際に検索される不動産情報も

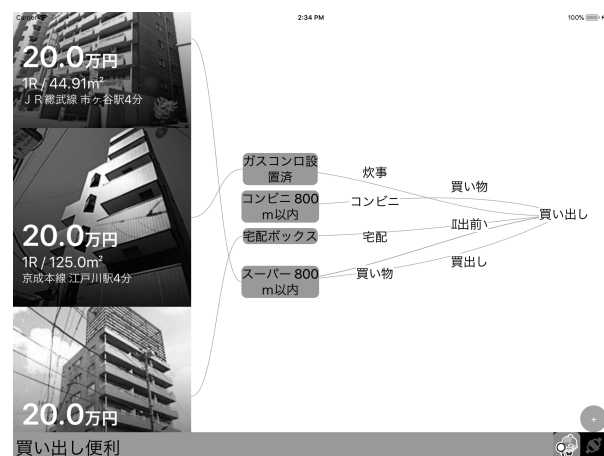


図1 Screen shot of the developed system.

表示する。ユーザが発声したどのような単語からどのようなキーワードを経てこだわり条件に結びつけたかは全て画面上に表示されている。

次に、ユーザの会話データを用いてこだわり条件を検索するアルゴリズムについて説明する。

3 こだわり条件検索アルゴリズム

既存の音声インターフェースに対して多くのユーザが感じる不満は、発声した言葉がシステムによって正しく認識されないことである。システムはプログラムされたキーワードのみを認識するが、ユーザはシステムが認識可能な単語を予め知ることができない。またユーザがシステムが認識するキーワードと類似の意味を有する別の単語を発音しても、システムは単語間の類似性を理解することができず人間が期待する応答を返すことができない。こうした場合ユーザは「なぜこの単語では認識されないのか」と大きなフラストレーションを感じるようになる。質疑応答 (Q & A retrieval) [3] の分野でも同様の困難が存在する [4]。

この研究では、そうした問題を解決するため Word2Vec [2] を使用することを試みた。Word2Vec を使用すると、

各単語がベクトル化されれば単語間の類似度が計算可能となる。本システムでは単純な単語間のマッチングではなく、ユーザが発声した単語と、こだわり条件に関連する単語を Word2Vec を用いて拡張し、関連する単語間でのマッチングを行なっている。アルゴリズム概要を以下に示す。

1. 各こだわり条件に対して、2つから5つの関連キーワードを手動で設定する。
2. 手動で設定した関連キーワードごとに、Word2Vec を使用して関連語を検索し、それらを拡張キーワードとして保存する。
3. 会話音声テキストを文字列に変換する。
4. 認識されたテキストから名詞と動詞を抽出する。
5. 抽出された単語に対して、Word2Vec を使用して関連語を検索する。それらを拡張キーワードとする。
6. 発話された単語の拡張キーワードと、こだわり条件より定義された拡張キーワードで同じ単語が見つかった場合、発話された単語とこだわり条件を関連づける。

本アルゴリズムによるマッチングの例を図 2 に示す。

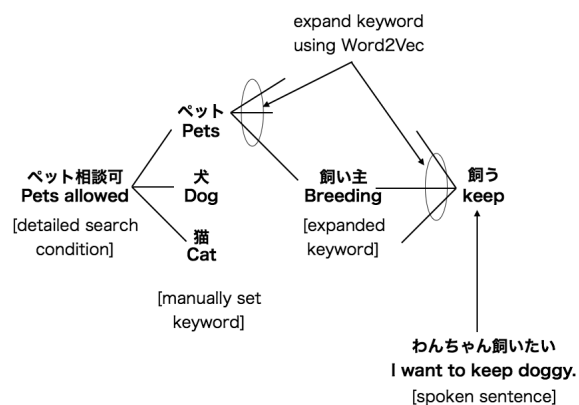


図 2 Concept diagram of matching spoken words and detailed search conditions.

4 評価結果

開発したシステムに対して、まずこだわり条件検索アルゴリズム単体の評価を実施した。入力として不動産に関する Q&A サイト [1] から住まいに関する質問を 50 件集めた。その上で、それぞれの質問に対して最大 5 つのこだわり条件を検索する。そのうち元の質問に関連があると思われるものが一つでもあれば正解とした。この評価では音声認識による誤りの影響を排除するため、入力としてテキストデータを用いている。

50 件の質問のうち、正しいこだわり条件が検索されたものは 34 件だった (68%) このうち 9 件の質問には、手動で設定したキーワードが含まれていたため、Word2Vec による拡張キーワードを用いなければ検索できなかった割合は 41 件中 25 件 (61%) である。

次に開発したシステムを用いて、不動産情報を探しているユーザの支援が可能かどうかを評価した。被験者は 5 名である。最初にシステムの概要を説明した上で、自由に会話してもらった。試験後にインタビューしたところ、すべてのユーザがこのシステムの可能性について肯定的な評価をした。しかし実際に使用した結果は効果的とは言えなかった。

本システムの目的、機能はユーザの会話からこだわり条件を検索する事であると説明したにも関わらず、ほとんどのユーザはシステムに対してなんでも発声できると解釈し、例えば「茨城市のアパート」のような発声を行なった。そうした検索条件を設定するための GUI はメニューから呼び出すことができるのだが、ユーザはそれを忘れて発声を行う場面が多数観察された。このことから本システムには改良が必要なることが明らかとなった。

5 むすび

不動産情報検索において、ユーザの会話に基づきこだわり条件を推薦するシステムの試作を行なった。評価の結果ユーザ発話をこだわり条件に結びつけるアルゴリズムの有効性はある程度認められたものの、インタラクティブシステムとして運用するためには課題があることも明らかになった。今後さらに評価を行いよりよいシステムの開発につなげたい。

参考文献

- [1] OKWeb - 『住まい』に関する質問・疑問一覧. (2017). <https://okwave.jp/c622.html>.
- [2] Thomas Mikolov. 2013. word2vec: Tool for computing continuous distributed representations of words. (2013). <https://code.google.com/word2vec/>.
- [3] Jiwoon Jeon, W. Bruce Croft, and Joon Ho Lee. 2005. Finding Similar Questions in Large Question and Answer Archives. In Proceedings of the 14th ACM International Conference on Information and Knowledge Management (CIKM '05). ACM, New York, NY, USA.
- [4] Xiaobing Xue, Jiwoon Jeon, and W. Bruce Croft. 2008. Retrieval Models for Question and Answer Archives. In Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '08). ACM, New York, NY, USA.