

Twitter データを用いた COVID-19 状況下の テレワーク時における建築空間の使われ方調査

下村隼生*・沖拓弥**

A survey on how architectural spaces are used during telework under COVID-19 conditions using Twitter data

Toshiki Shimomura*, Takuya Oki**

The purpose of this paper is to investigate how architectural spaces are used during telework under COVID-19 conditions, using Twitter data that can quickly collect information transmitted by various users: (1) We collected 76,049 tweets using the Twitter API based on the specific keywords, time period, and other criteria; (2) As a result, 26,556 tweets related to telework were obtained by filtering based on images and text information in the tweets; (3) We extracted the information on architectural space for telework, emotions, and factors during telework using these tweets; (4) Finally, we analyzed the characteristics of how each space is used during telework and changes in emotions and factors at different times.

Keywords: Twitter (Twitter) , テレワーク (telework) , COVID-19 (COVID-19) , 物体検出 (object detection) , 感情分析 (sentiment analysis) , 使われ方調査 (usage survey)

1. はじめに

2020年、世界規模のパンデミックを引き起こした新型コロナウイルス (COVID-19) を契機に、従来の就業場所以外で業務を行う「テレワーク」と呼ばれる働き方を始める人が増加した。こうした「新しい生活様式」に相応しい建築空間のあり方を検討する上で、テレワークにより生じた問題点を明らかにする意義は大きいと考えられる。

そこで本稿では、人々の生活実態に関する情報を迅速かつ大量に取得可能な Twitter データを用いて、ツイート (つぶやき) に含まれる文章や画像および時間情報からテレワークの実態を分析することを目的とする。具体的には、まず Twitter API を用いて収集したツイートデータに対し、複数のフィルタリング作業を行うことで、テレワークに関係すると考えられるツイートを抽出する。次に、抽出したツイート本文の内容をもとに各ツイートを「ポジティブ」、「ネガティブ」、

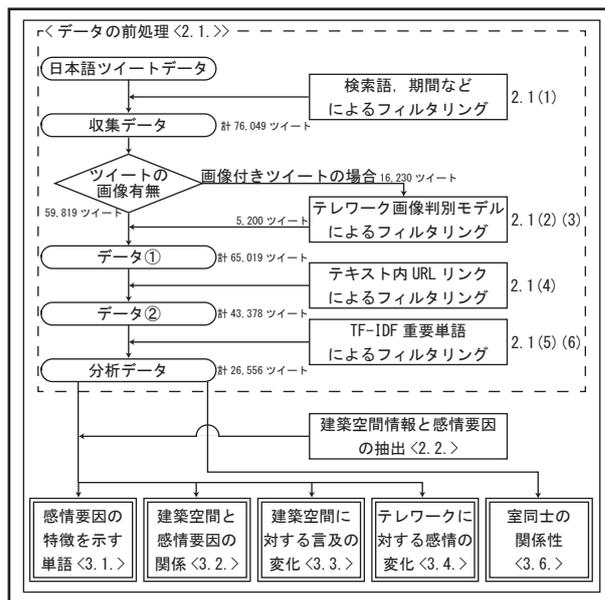


図1 本研究の流れ

「ニュートラル」のいずれかに分類する。そして、「ポジティブ」または「ネガティブ」と分類されたツイートから、(1)テレワーク時の就労スパー

* 学生会員 東京工業大学環境・社会理工学院 (Tokyo Institute of Technology)
〒152-8550 東京都目黒区大岡山 2-12-1 E-mail: shimomura.t.aa@m.titech.ac.jp
** 正会員 東京工業大学環境・社会理工学院 (Tokyo Institute of Technology)

ス、および、(2)テレワークに対する感情とその要因を抽出する。さらに、本文の内容やツイート時期の変化に着目した分析を行うことにより、テレワーク時における建築空間の使われ方を明らかにする。

2. 建築空間の使われ方調査の流れ

2.1. テレワーク関連ツイートデータの抽出

以下の(1)～(6)に示す手順で段階的にツイートデータをフィルタリングし、テレワーク環境に関する情報が十分に含まれるツイートを抽出する。

(1) 図2(a)に示す条件のもと、Twitter APIを用いてツイートデータを収集する(収集データ: 計76,049ツイート(図2(c)))。

(2) 収集データのうち画像を含むツイートの中から、テレワークに関連する画像と関連しない画像を500枚ずつ無作為に抽出し、VGG16 (Simonyan. et al., 2015) とサポートベクター分類(SVC)を用いた「テレワーク画像判別モデル」(図3)を構築する。テストデータでの正解率やF値は約97%と高い判別性能を示す。

(3)(2)で構築したモデルを用いて、収集データのうち画像を含むツイートから、テレワーク関連画像を含むツイートを抽出する(16,230ツイート→5,200ツイート)。

(4) データ①((3)のフィルタリング後のツイート)と、収集データのうち画像を含まないツイート)のエンティティ情報をもとに、外部URLリンクを含まないツイートを抽出する¹⁾(データ①:65,019ツイート→データ②:43,378ツイート)。

(5) データ②のうち、本文からテレワークに対する感情とその要因(図4(a))が読み取れるツイートを、無作為に388ツイート抽出した。要因ごとにツイート内で共通して用いられる特徴的な単語が存在すると考えられる。それらの単語を抽出するために、抽出したツイートを各感情要因ごとに一つの文書としてまとめ

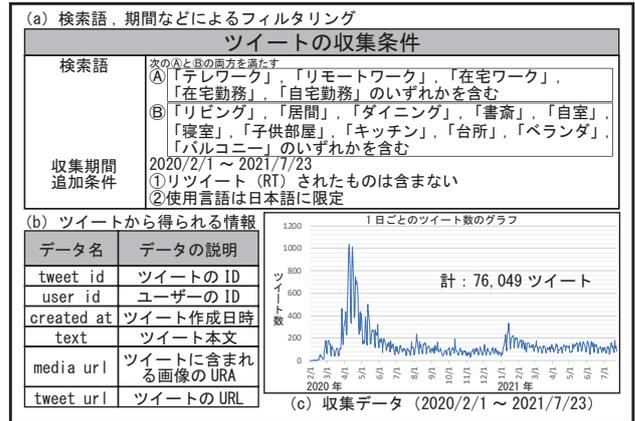


図2 ツイートの収集条件

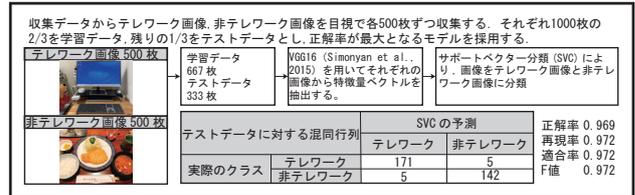


図3 テレワーク画像判別モデル

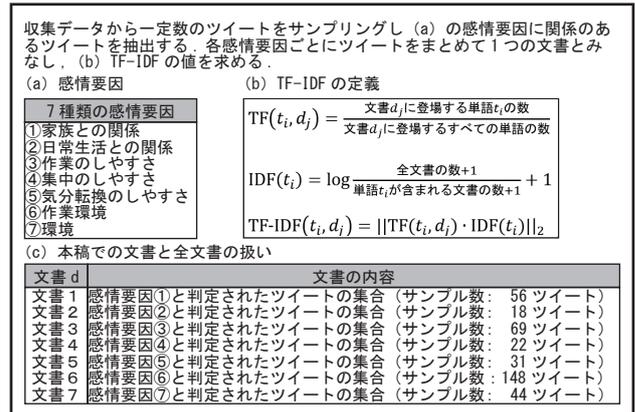


図4 TF-IDFによる重要単語の抽出

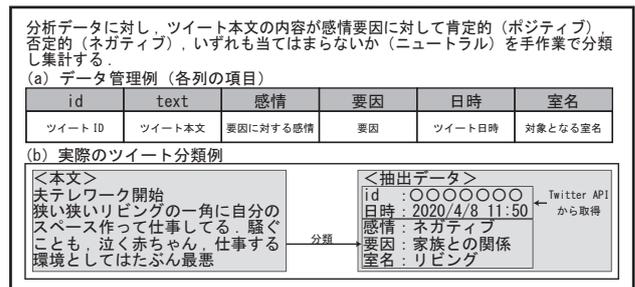


図5 肯定的・否定的言及の集計

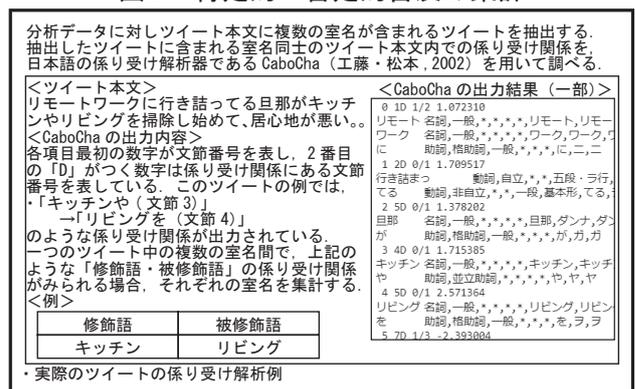


図6 室同士の関係性の分析

表 1 要因ごとの TF-IDF 上位 20 単語

順位	①家族との関係	TF-IDF	②日常生活との関係	TF-IDF	③作業しやすさ	TF-IDF	④集中のしやすさ	TF-IDF	⑤気分転換のしやすさ	TF-IDF	⑥作業環境	TF-IDF	⑦環境	TF-IDF
1	邪魔	0.232	ご飯	0.310	セット	0.184	誘惑	0.313	転換	0.392	チェア	0.301	ライト	0.186
2	子供	0.224	カメラ	0.290	パソコン	0.179	集中	0.250	気分	0.332	モニター	0.265	オフィス	0.177
3	パソコン	0.203	片付け	0.186	オフィス	0.149	避難	0.188	ヨガ	0.168	イス	0.207	はで	0.163
4	家族	0.178	アベニュー	0.175	出張	0.138	寝たきり	0.176	うち	0.166	購入	0.159	イス	0.163
5	会議	0.153	レンタル	0.175	不便	0.118	漫画	0.176	リズム	0.135	座り	0.159	電球	0.163
6	感じ	0.153	データ	0.145	効率	0.118	音楽	0.176	住人	0.135	ディスプレイ	0.145	気分	0.161
7	移動	0.134	共有	0.145	イメージ	0.111	間食	0.176	息抜き	0.135	会社	0.140	弁当	0.139
8	ブチ	0.134	背景	0.145	余裕	0.111	騒音	0.176	空気	0.135	一角	0.136	キュウリ	0.131
9	無理	0.134	階段	0.145	大事	0.111	可能	0.163	部長	0.135	パソコン	0.135	ゴーヤ	0.131
10	占領	0.117	食事	0.145	拡充	0.111	雑音	0.146	エアコン	0.125	サブ	0.132	トマト	0.131
11	奥さん	0.117	パソコン	0.141	無線	0.111	感じ	0.143	生活	0.125	整備	0.113	ナス	0.131
12	大変	0.116	家族	0.141	父上	0.111	コタツ	0.125	コーヒー	0.112	写真	0.113	体制	0.131
13	ノート	0.102	会社	0.124	粘土	0.111	音楽	0.095	外出	0.112	姿勢	0.113	光量	0.131
14	キーボード	0.100	サイズ	0.124	緩和	0.111	すずめ	0.088	オフィス	0.109	完成	0.113	うち	0.121
15	ゲーム	0.100	大変	0.108	遠隔	0.111	そよ風	0.088	移動	0.096	ノート	0.110	カウンター	0.109
16	位置	0.100	準備	0.108	仕事場	0.105	はめ	0.088	おかげ	0.096	感じ	0.098	占領	0.109
17	ガソリン	0.094	邪魔	0.108	勉強	0.105	まめ	0.088	オンオフ	0.096	スタンド	0.098	転換	0.109
18	ケタ	0.094	感じ	0.094	子供	0.105	もろもろ	0.088	最高	0.096	設置	0.098	天気	0.106
19	要望	0.094	うどん	0.087	仕様	0.102	アダプタ	0.088	マット	0.083	移動	0.097	おかげ	0.093
20	場所	0.090	きれい	0.087	フル	0.092	アレクサ	0.088	週間	0.083	コタツ	0.097	一緒	0.093

(図 4(c))，各要因ごとと文書ごとに単語の TF-IDF の値を算出する(表 1)．さらに，TF-IDF の値が大きな単語の中からそれぞれの要因の特徴を表す語を抽出する．

(6) データ②に対して，形態素解析器の一つである McCab (参照) を用いて，各ツイートを形態素単位に分割する．そして，形態素に (5) で抽出した特徴を表す語を含むツイートを抽出する²⁾ (データ②：43,378 ツイート→分析データ：26,556 ツイート)．

2.2. 建築空間情報と感情要因の抽出

テレワークに対する肯定的または否定的意見を分析するために，分析データ (26,556 ツイート) から各ツイート内容の感情を抽出する必要がある．本稿では，目視で肯定的または否定的言及を含むツイートを抽出した³⁾．このとき，図 4(a) に示す 7 種類の感情要因をベースに，テレワーク従事者の視点あるいはその家族の視点から，テレワークに対する肯定的言及か否定的言及かを判断し，図 5 のようにデータを整理した．

2.3. 室同士の関係性の抽出

分析データの中には，一ツイート中に複数の室名が含まれるツイートが含まれる．こうしたツイートを対象に室を表す語同士の係り受け関係を調べ (図 6)，テレワーク時における室同士の関係性を抽出した．

3. 建築空間の使われ方分析

3.1. 感情要因に関係の強い語

要因 (文書) ①～⑦ごとに TF-IDF の値が大きい上位 20 単語を表 1 に示す．「①家族との関係」は，上位に「子供」，「家族」，「奥さん」など，家族に関係する語が並んでいる．同様に，「⑥作業環境」は，「モニター」，「イス」，「ディスプレイ」など，テレワーク時の作業環境がうかがえる単語が上位を占める．このことから，各文書の特徴を踏まえた単語が抽出できていることがわかる．

一方，「⑦環境」については，ツイートのサンプル数が少ない (図 4(c)) ため，特定のツイートに含まれる一見無関係な単語が上位に現れている．

3.2. 建築空間と感情要因の関係

表 2 に，分析データから抽出した肯定的，否定的言及を含むツイートをもとに，建築空間と感情要因の関係をクロス集計した結果を示す．(a) ポジティブについてみると，要因として「⑤気分転換のしやすさ」のツイート数が多い．特に，リビング・居間，キッチン・台所，ベランダ・バルコニーの値がとりわけ大きい．ベランダ・バルコニーは，外気に触れることができるスペースであることが関係している可能性がある．また，リビング・居間，キッチン・台所は，家族全員が利用しやすい空間であり，気分転換に使用されたものと推察される．

(b) ネガティブについてみると，「リビング・居間×①家族との関係」の値が，他と比べて明ら

かに多い。リビング・居間は家族と一番長く過ごす空間であり、些細なことでも不満がたまりやすく、共感も得やすいため、多くのツイートがなされた可能性がある。

3.3. 建築空間に対する肯定的・否定的言及の推移

図7に、分析データから抽出した肯定的、否定的言及を含むツイートに対し、建築空間に関するツイート割合を月ごとに集計した結果を示す。

(a) ポジティブについては、2020年、2021年いずれも、ベランダ・バルコニーの4月・5月の割合が大きい。例年この時期は、過ごしやすい気候であることが関係している可能性がある。また、リビング・居間の2020年と2021年の同時期の割合を比較するといずれの場合も後者の割合が高い。テレワークの経験期間が長くなるにつれて、テレワークの生活に順応している可能性がある。

(b) ネガティブについては、リビング・居間のみ割合の変化が大きい。リビング・居間は十分なスペースがあるため、テレワーク開始時に利用した人が多いかったが、ネガティブな意見の増減は「緊急事態宣言」等で新規にテレワークを始める人の影響がある可能性がある。

3.4. テレワークに対する感情と要因の推移

分析データから抽出した肯定的、否定的言及を含むツイートに対し、感情要因ごとのツイート割合を集計した(図8)。全体的に(a) ポジティブよりも(b) ネガティブのツイート数が多く、Twitterではテレワーク時のネガティブな感情がツイートされやすいものと考えられる。

(a) ポジティブは、第一回目の緊急事態宣言(2020年4月7日)と同時期に、いずれの要因も突発的にツイート数が増加しているが、それ以降はほぼ横ばいで推移している。

(b) ネガティブについてみると、ポジティブと同様に第一回目の緊急事態宣言の前後でツイート数が増加している。また、2020年8月には「⑦環境」の感情要因のツイート数が増加している。これは、夏場の猛暑で、住宅でのテレワーク時に暑さに対する不満が増加した可能性がある。

表2 建築空間とその要因のクロス表

(a) ポジティブ (肯定的言及)		①	②	③	④	⑤	⑥	⑦	合計	凡例
リビング・居間	70	30	103	97	203	145	54	702	200~	
ダイニング	4	3	39	34	41	92	14	227	150~200	
書斎	16	12	28	39	39	83	13	230	90~100	
自室	9	15	56	30	42	58	37	247	80~90	
寝室	25	11	19	24	42	21	16	158	70~80	
子供部屋	6	4	1	2	3	5	1	22	60~70	
キッチン・台所	17	31	57	41	121	44	24	335	50~60	
ベランダ・バルコニー	2	1	77	39	178	55	149	501	30~40	
合計	149	107	380	306	669	503	308	2422	20~30	

(b) ネガティブ (否定的言及)		①	②	③	④	⑤	⑥	⑦	合計	凡例
リビング・居間	533	71	126	213	250	123	309	1625	500~	
ダイニング	34	9	67	40	60	309	47	566	300~500	
書斎	66	40	36	55	26	31	49	303	200~300	
自室	50	72	38	71	90	43	145	509	100~200	
寝室	185	15	12	20	78	43	87	440	90~100	
子供部屋	14	9	2	3	5	6	6	45	80~90	
キッチン・台所	111	43	36	38	88	47	32	395	70~80	
ベランダ・バルコニー	13	0	28	61	64	23	63	252	60~70	
合計	1006	259	345	501	661	625	738	4135	50~60	

※丸数字①~⑦は図4(a)に対応
 ※一つのツイートに複数の建築空間に関する言及が含まれる場合や、肯定的・否定的言及が含まれる場合には、重複して集計している。

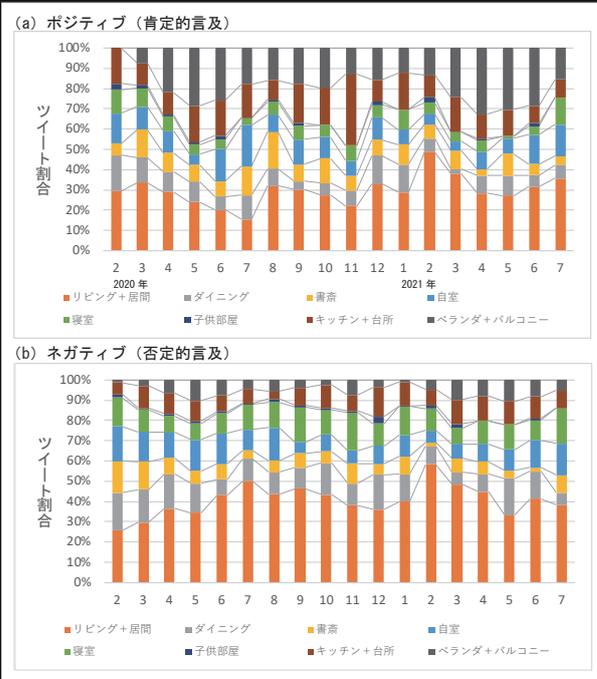


図7 室空間ごとのツイート割合の推移

3.5. 感情要因に関する詳細な考察

図9に、建築空間と感情要因に対し、具体的などのような言及が含まれていたかを例示している。リビング・居間やキッチン・台所といった家族の共有スペースでは、テレワーク従事者よりもその家族が不満を抱えていることがわかる(図9(a)(b))。一方、テレワークの際は、子育てや日常生活に積極的に参加していなくても肯定的言及につながっているケースが見られ興味深い。図9(c)からは、テレワーク時におけるプライバシー

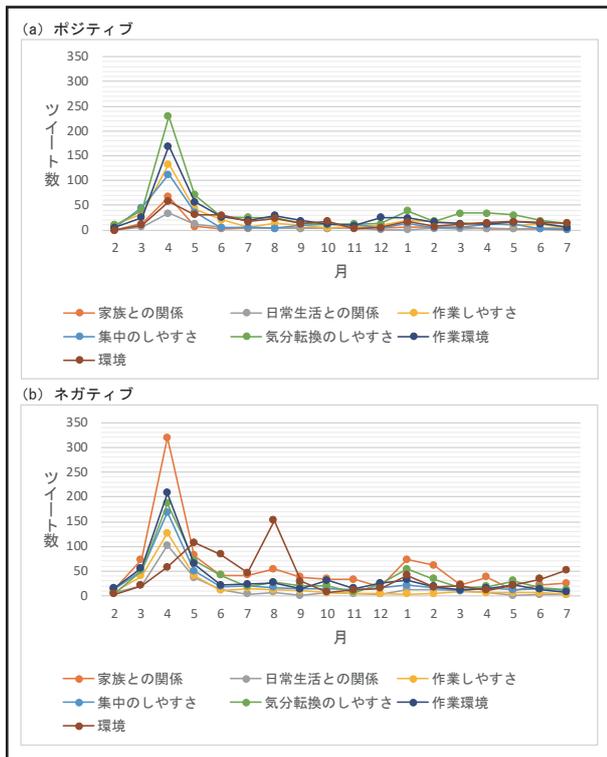


図8 感情要因ごとのツィート数の推移

の問題が、図9(i), (j)からは、多くの人が在宅勤務を始めたことにより生じた問題点が抽出されている。

3.6. 室同士の関係性の分析

分析ツィートの中で、1つのツィートに複数の室名が含まれるツィートから、テレワーク時の複数の室の利用され方の関係を調べ、グラフで表した(図10)。テレワーク時に複数の室を日ごとに変えるテレワーク従事者は、気分転換の目的や作業内容に応じて適した空間を選択している可能性がある。特に、複数の室を利用する場合は、リビングとその他の室が併用される可能性が高いのではないかと推察される(図10(a))。また、テレワークの場所を変更するテレワーク従事者は、ダイニングやリビングといった家族共有の空間から、キッチンや寝室、ベランダといった特定の個人が利用する空間へと変化している様子が見て取れる(図10(b))。

4. まとめ

本稿では、Twitterからテレワークに関するツィートデータを抽出し、建築空間ごとのテレワークに対する肯定的・否定的言及とその要因を分析

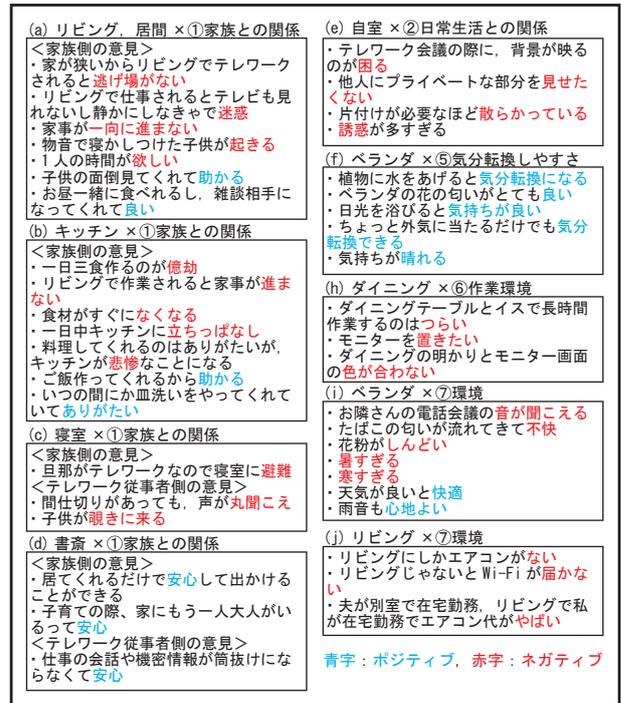


図9 ツィートの言及内容の例

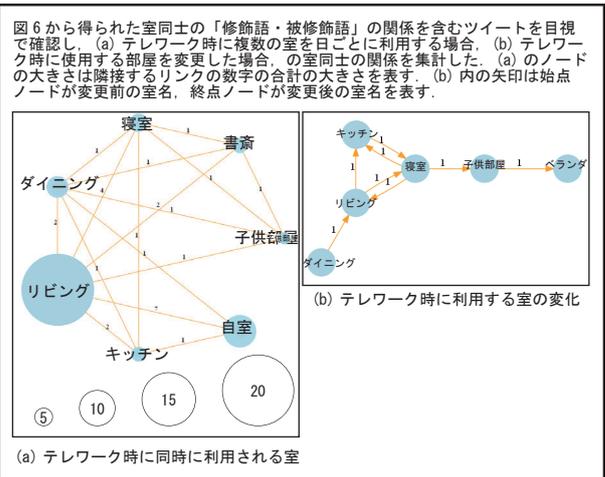


図10 室同士の関係性のグラフ

した。また、建築空間ごとに、テレワークに対する肯定的・否定的言及の特徴を示した。さらに、一つのツィートの中で複数の室名に言及のあるツィートを調べることで、テレワーク時の室同士のつながりの強さを評価した。

謝辞

ツィートデータを分析するにあたり協力いただいた高井峻さん、山中理沙さん(東京工業大学環境・社会理工学院 沖研究室)に謝意を表します。

注釈

1. 外部 URL リンクを含むツイートは広告ツイートやツイート本文に有用な情報が含まれていない可能性が高いと考えた.
2. 各ツイートの形態素の中に 2.1.(5) で抽出した各要因の特徴を表す語が含まれるツイートはツイート本文で各感情要因に対しての言及がある可能性が高いと考えた.
3. 機械学習による感情分析 (例えば, ML-Ask や Google Cloud Natural Language API) も試みたが, 精度は十分高いとは言えなかった.

参考文献

- 工藤拓・松本裕治 (2002) チャンキングの段階適用による日本語係り受け解析. 「情報処理学会論文誌」, 43(6), 1834-1842.
- Kudo, T., Yamamoto, K. and Matsumoto Y. (2004) Applying Conditional Random Fields to Japanese Morphological Analysis. Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP-2004), 230-237.
- Simonyan, K., Zisserman, A. (2015) Very Deep Convolutional Networks for Large-Scale Image Recognition. In Proc. of ICLR.