

E-Commerceにおける不満情報とレビューに基づいた不満解決商品推薦手法の提案

林 利憲¹⁾ 王 元元²⁾ 河合 由起子³⁾ 角谷 和俊¹⁾
1)関西学院大学 2)山口大学 3)京都産業大学

Background

Complaints exist for everything

Ex) Smart Phone

Accessories are too expensive.

Not be able to change the battery.



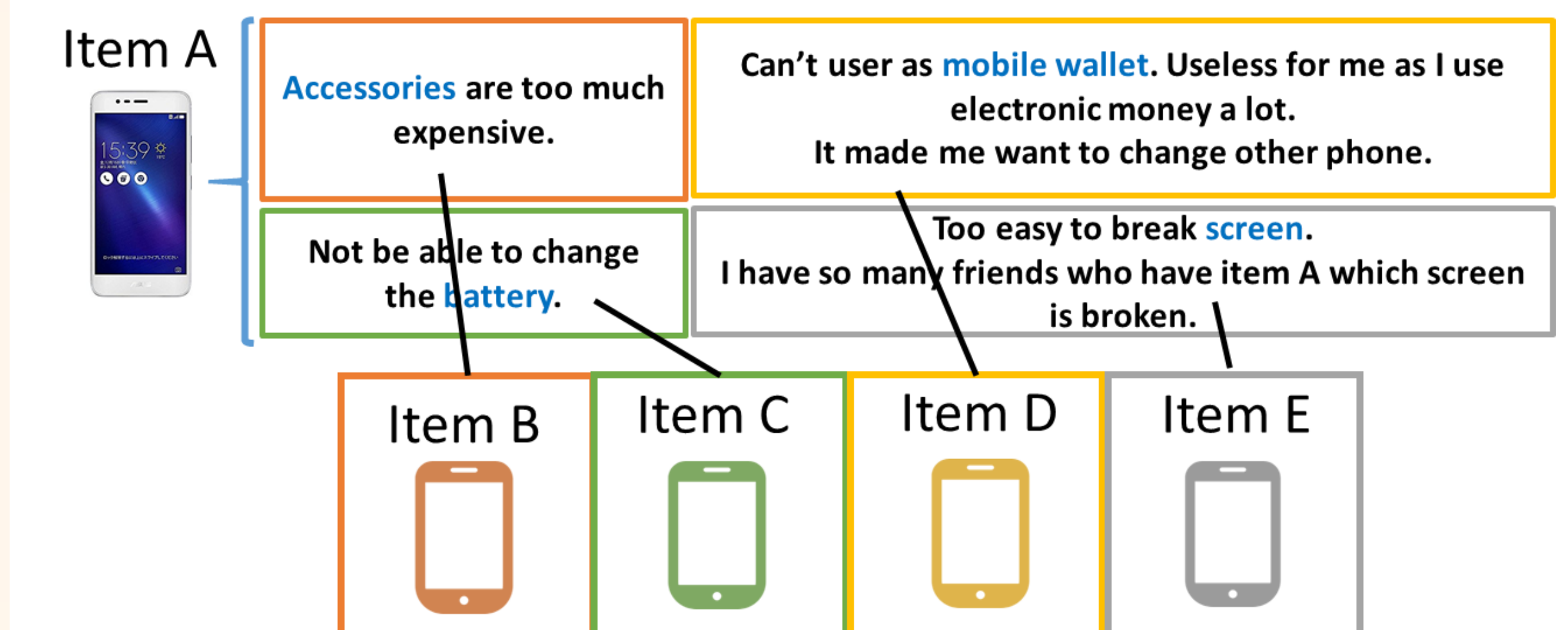
Can't user as mobile wallet. Useless for me as I use electronic money a lot. It made me want to change other phone.

Too easy to break screen. I have so many friends who have itemA which screen is broken.

Only few recommender systems focus on complaints

Proposed Method

Recommending items which solve problems of complaints



Our Approach

- Classify complaints by items
 - Extract nouns by each complaint that belongs to same item
 - Generate vectors by each item using TF-IDF methods
- Extract positive words from reviews
 - Generate vectors of good and bad reviews by each items
 - Calculate vectors of item by subtracting bad from good

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad idf_i = \log \frac{|D|}{df_i}$$

※d denotes the document that is integrated by all Fuman of one item

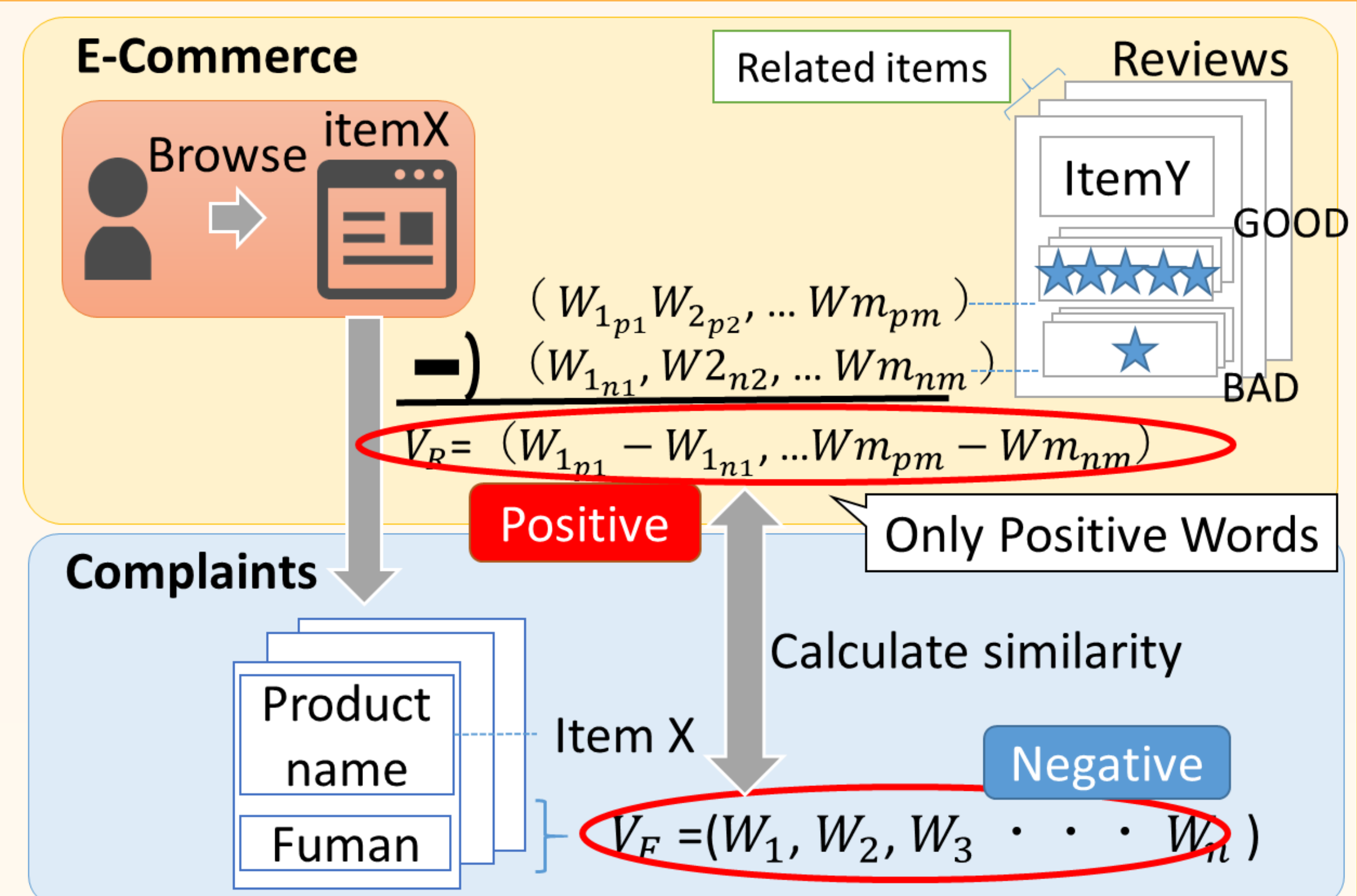
$$Sim(V_F, V_R) = \frac{\sum_{i=1}^{|V|} W_n \cdot W_{p_i} - W_{n_i}}{\sqrt{\sum_{i=1}^{|V|} (W_n)^2} \cdot \sqrt{\sum_{i=1}^{|V|} (W_{p_i} - W_{n_i})^2}}$$

V_F is the vector generated from complaints, V_R is the vector generated from reviews

- Match items of complaints and reviews

Calculate similarity between complaints and reviews

- Recommend items to users



Evaluation

Dataset

I. Complaints (Fumankaitori Centre^{*1})

- 4 items (itemA, itemB, itemC, and itemD)
- All items are categorized as smartphone
- 200 Complaint data

*1 <http://fumankaitori.com/>

II. Reviews (Amazon^{*2})

- 7 Items (item1 to item7)
- All Items are categorized as smartphone except Item6 and item7
- 350 Review data

*2 <https://www.amazon.com/>

- Extraction of feature words from complaints and reviews

Items	Top 20 Feature words from complaints
Item A	交換, G, 機, 最近, 量, 写真, 入力, 文字, 容量, 電池, 変, 電波, 携帯, 設定, 修理, SoftBank, ショップ, カード, データ, アプリ
Item B	カバー, 傷, 閲覧, 型, 内蔵, 途中, 動作, 処理, 簡単, ブック, マーク, 作動, 表面, シート, フラッシュ, プレイヤー, 常態, テンキー, 晴れ, キャッシュ, グリッドセンサー
Item C	音楽, ステレオ, メモリ, 電池, カー, 音, OFF, 通話, 液晶, 改善, 場所, 不思議, Xperia, パッケージ, 説明, PC, 画像, フィルム, USB, 設計
Item D	番号, 発信, 電話, ブラウザ, 限り, 相手, 心配, 目的, 指先, 履歴, 着信, タッチ, 指, 反応, 音楽, ステレオ, メモリ, 電池, カー, 音

Items	Top 15 Feature words from reviews
Item 1	包装, めちゃくちゃ, 運用, アマゾン, 満足, クリーニング, 省略, 一瞬, 贈物, オーバースペック, 一掃, 目視, 領域, 店頭, 現役
Item 2	アクティベーションロック, 格安, 無事, ロック, IJ, モバックス, プロファイル, 安心, シム, 接続, 予定, iOS, カーナビ, 娘, ナビ
Item 3	フルセグ, フル, テレビ, アンテナ, デジ, インチ, ネクサス, 端, 車, タブ, 風呂, ペン, タブレット, ドコモ, ガラケー
Item 4	Z, AC, 認証, Xperia, ケータイ, コンパクト, 指紋, サイフケータイ, GPS, touch, 逆, 輸入, 音質, サイズ, 日本語
Item5	Mate, Lite, 指紋, 認証, ガラス, 撮影, 指, 写真, 高級, 最高, lazer, ワイドアパーチャ, ビント, 爆速, 感動

Result

Most of feature words are extracted as subjective of complaints and positive reviews

- Similarity calculation

between complaints and reviews

Complaints					
Reviews		ItemA	ItemB	ItemC	ItemD
Smart phone	Item1	0.28	0.20	0.18	0.10
	Item2	0.63	0.25	0.33	0.12
	Item3	0.46	0.17	0.40	0.20
	Item4	0.46	0.16	0.60	0.05
Non smart phone	Item5	0.60	0.27	0.43	0.06
	Item6	0.53	0.17	0.21	0.04
	Item7	0.23	0.06	0.03	0.02

Result

- High similarity were found after normalization
- It has high similarity with different category

- Qualitative Evaluation

Evaluating proposed system by questionnaire asking 6 people

Q1: Evaluate if products item1-5 resolves each dissatisfaction below

- Complaint about battery
- Complaint about internet connection
- Complaint about charger
- Complaint about photo and memory
- Complaint about reaction of screen

Result

Compared averages of evaluation with result of proposed system

- Pearson's Correlation coefficient : 0.49
- System evaluated ok, but need modification

Q2: Do you think Item 6 and 7 resolve these dissatisfaction? Choose a number below

1.Stronglv disagree 2.disagree 3.Neither disagree nor agree 4.agree 5.Stronglv agree

Result

Average of this question: 3.6

Recommending other categories' item could also resolve the dissatisfaction

Average of evaluation

Adv	Item1	Item2	Item3	Item4	Item5
①	0	0	0.5	0.5	0.5
②	0.5	0.75	0.25	0.5	1
③	0.25	0.25	0	0	0
④	0	0.75	0.75	1	1
⑤	0.25	1	1	0.25	1

Proposed System

System	Item1	Item2	Item3	Item4	Item5
①	0	0	0.5	0.25	1
②	1	0.6	0.8	0.4	0.2
③	0.5	0	0.25	0	1
④	0	0.5	0.25	0.75	1
⑤	0	1	0.75	0.25	0.5

※1 means the item resolve the dissatisfaction, 0 means the item doesn't resolve dissatisfaction.

- Comparison of complaints and negative review

- Compared feature words extracted from both data
- 10 examiner evaluated negative words

Average precision	Complaint	Review
ItemA	0.49	0.16
ItemB	0.46	0.41
ItemC	0.53	0.65
ItemD	0.36	0.27

Result

Feature words extracted from complaints are recognized as more negative

Complaint	Evaluation	Review	Evaluation
Stereo	No-N	Initial	No-N
Music	No-N	Friend	No-N
Xperia	Negative	Download	No-N
Battery	Negative	Car navigation	No-N
Car	No-N	Self	No-N
Sound	No-N	Sim	No-N
Memory	Negative	Beauty	No-N
Movie	Negative	Initially	No-N
USB	Negative	Packaging	Negative
Sebum	No-N	Necessary	No-N

Future Work

- Propose a new extraction methods
- Validate complaints data with many items
- Extracting feature words with other part of speech
- Recommending items from all kinds of dissatisfaction

謝辞 本研究では、株式会社 Insight Tech が国立情報学研究所の協力により研究目的で提供している「不満調査データセット」を利用した。