

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2020年9月17日(17.09.2020)



(10) 国際公開番号

WO 2020/184716 A1

(51) 国際特許分類:
G06Q 50/10 (2012.01) *G06F 16/9536* (2019.01)
G06F 16/9035 (2019.01)

(21) 国際出願番号: PCT/JP2020/011225

(22) 国際出願日: 2020年3月13日(13.03.2020)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2019-046846 2019年3月14日(14.03.2019) JP

(71) 出願人: 国立大学法人東京大学 (THE UNIVERSITY OF TOKYO) [JP/JP]; 〒1138654 東京都文京区本郷七丁目3番1号 Tokyo (JP).

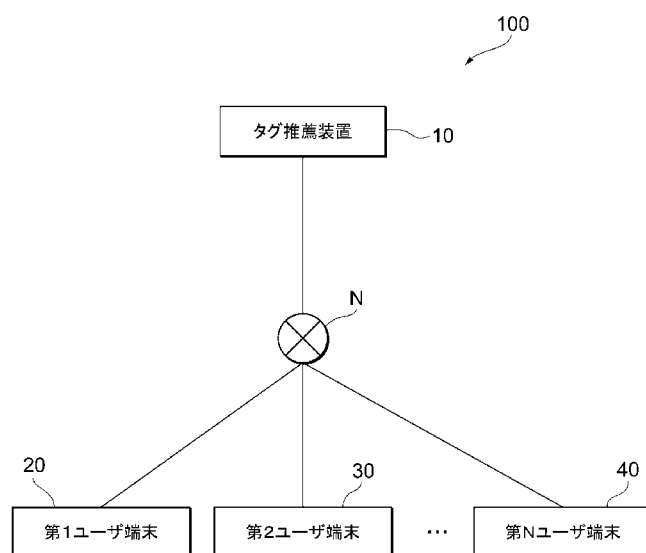
(72) 発明者: 山崎 俊彦 (YAMASAKI, Toshihiko); 〒1138654 東京都文京区本郷七丁目3番1号 国立大学法人東京大学内 Tokyo (JP).
汪 雪テイ (WANG, Xueting); 〒1138654 東京都文京区本郷七丁目3番1号 国立大学法人東京大学内 Tokyo (JP).

(74) 代理人: 稲葉 良幸, 外 (INABA, Yoshiyuki et al.); 〒1066123 東京都港区六本木6-10-1 六本木ヒルズ森タワー23階 TMI 総合法律事務所 Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,

(54) Title: TAG RECOMMENDATION DEVICE, TAG RECOMMENDATION METHOD AND TAG RECOMMENDATION PROGRAM

(54) 発明の名称: タグ推薦装置、タグ推薦方法及びタグ推薦プログラム



10 Tag recommendation device
20 First user terminal
30 Second user terminal
40 Nth user terminal

(57) Abstract: To provide a tag recommendation device that recommends tags assigned by popular users, and tags to be assigned to new contents by considering the assignment frequency. A tag recommendation device 10 is provided with: a first calculation unit 14 for calculating a third numerical value representing the co-occurrence relation between a plurality of tags on the basis of a first numerical value that gives a weight to the co-occurrence relations between a plurality of users and a plurality of tags according to the popularity of the plurality of users, and a second numerical value that gives a weight to the co-occurrence relations between a plurality of users and a plurality of tags according to the frequency the plurality of users assigned the plurality of tags to a plurality of contents; a second calculation unit 15 for calculating the

DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

- (84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

規則4. 17に規定する申立て：

- 一 不利にならない開示又は新規性喪失の例外に関する申立て（規則4. 17(v)）

添付公開書類：

- 一 国際調査報告（条約第21条(3)）

priorities of a plurality of tags on the basis of the third numerical value and one or a plurality of tags assigned to new contents; and a recommendation unit 16 for recommending, on the basis of the priorities, one or a plurality of tags to be assigned to new contents from the plurality of tags.

(57) 要約：人気度が高いユーザによって付与されたタグと、その付与頻度を考慮して新たなコンテンツに付与するタグを推薦するタグ推薦装置を提供する。タグ推薦装置10は、複数のユーザと複数のタグとの共起関係を、複数のユーザの人気度に応じて重み付けした第1数値と、複数のユーザと複数のタグとの共起関係を、複数のユーザが複数のタグを複数のコンテンツに付与した頻度に応じて重み付けした第2数値とに基づいて、複数のタグ間の共起関係を表す第3数値を算出する第1算出部14と、第3数値及び新たなコンテンツに付与された1又は複数のタグに基づいて、複数のタグの優先度を算出する第2算出部15と、優先度に基づいて、複数のタグの中から、新たなコンテンツに付与する1又は複数のタグを推薦する推薦部16と、を備える。

明 細 書

発明の名称：

タグ推薦装置、タグ推薦方法及びタグ推薦プログラム

関連出願の相互参照

[0001] 本出願は、2019年3月14日に出願された日本出願番号2019-46846号に基づくもので、ここにその記載内容を援用する。

技術分野

[0002] 本発明は、タグ推薦装置、タグ推薦方法及びタグ推薦プログラムに関する。

背景技術

[0003] 従来、タグは、画像や動画をコンテンツとするSNS（Social Networking Service）や、商品をコンテンツとする電子商取引サイトが広く用いられている。SNSや電子商取引サイトのユーザは、コンテンツに付与されたタグを用いて、コンテンツをテキスト検索する場合がある。ここで、コンテンツへのタグの付与は、コンテンツを投稿したユーザによって行われている。

[0004] タグの付与について、例えば、下記非特許文献1には、ユーザ、コンテンツ及びタグをノードとし、ノードの共起関係をリンクとする無向グラフを用いて、コンテンツに付与するタグを推薦する技術が記載されている。

また、下記非特許文献2には、SNSにおけるコンテンツの人気度とタグの共起関係に基づいて、投稿するコンテンツの人気度が高まるようにタグを推薦する技術が記載されている。

先行技術文献

非特許文献

[0005] 非特許文献1：Andreas Hotho, Robert Jaschke, Christoph Schmitz, Gerd Stumme, “FolkRank: A Ranking Algorithm for Folksonomies”, Lwa, Vol. 1, pp. 111-114 (2006)

非特許文献2：T. Yamasaki, J. Hu, S. Sano, and K. Aizawa, “Folkpopula

rityrank: Tag recommendation for enhancing social popularity using text tags in content sharing services,” Proceedings of the 26th International Joint Conference on Artificial Intelligence, pp.3231–3237, 2017

発明の概要

発明が解決しようとする課題

[0006] 先行技術によって、人気のあるコンテンツに付与されたタグと、そのタグと共起するタグを重要視して、新たなコンテンツに付与するタグを推薦することができる。一方、SNSや電子商取引サイト等のコミュニティでは、人気度の高いユーザが大きな影響力を有しており、人気度が高いユーザが付与しているタグを新しいコンテンツに付与することで、そのコンテンツの人気度が向上することがある。しかしながら、人気度が高いユーザによってどのようなタグが付与されているか網羅的に把握することは難しく、頻繁に付与されているタグと稀にしか付与されていないタグとのバランスを取ることも難しい。

[0007] そこで、本発明は、人気度が高いユーザによって付与されたタグと、その付与頻度を考慮して新たなコンテンツに付与するタグを推薦するタグ推薦装置、タグ推薦方法及びタグ推薦プログラムを提供する。

課題を解決するための手段

[0008] 本発明の一態様に係るタグ推薦装置は、複数のユーザと複数のタグとの共起関係を、複数のユーザの人気度に応じて重み付けした第1数値と、複数のユーザと複数のタグとの共起関係を、複数のユーザが複数のタグを複数のコンテンツに付与した頻度に応じて重み付けした第2数値とに基づいて、複数のタグ間の共起関係を表す第3数値を算出する第1算出部と、第3数値及び新たなコンテンツに付与された1又は複数のタグに基づいて、複数のタグの優先度を算出する第2算出部と、優先度に基づいて、複数のタグの中から、新たなコンテンツに付与する1又は複数のタグを推薦する推薦部と、を備える。

- [0009] この態様によれば、複数のタグ間の共起関係を表す第3数値を、ユーザがタグを付与した頻度と、ユーザの人気度とを用いて算出し、第3数値と、新たなコンテンツに予め付与された1又は複数のタグとに基づいて複数のタグの優先度を算出することで、人気度が高いユーザによって付与されたタグと、その付与頻度を考慮して新たなコンテンツに付与するタグを推薦することができる。
- [0010] 上記態様において、第1算出部は、複数のコンテンツと複数のタグとの共起関係を、複数のコンテンツの人気度に応じて重み付けした第4数値と、複数のコンテンツと複数のタグとの共起関係を、複数のタグが複数のコンテンツに付与された頻度に応じて重み付けした第5数値と、に基づいて、複数のタグ間の共起関係を表す第6数値を算出し、第2算出部は、第3数値、第6数値及び新たなコンテンツに付与された1又は複数のタグに基づいて、優先度を算出してもよい。
- [0011] この態様によれば、複数のタグ間の共起関係を表す第6数値を、タグがコンテンツに付与された頻度と、コンテンツの人気度とを用いて算出し、第6数値と、新たなコンテンツに予め付与された1又は複数のタグとに基づいて複数のタグの優先度を算出することで、コンテンツの人気度を考慮して新たなコンテンツに付与するタグを推薦することができる。
- [0012] 上記態様において、第3数値及び第6数値は、複数のタグの数の次元を有する正方行列であり、第2算出部は、第3数値及び第6数値の要素積並びに新たなコンテンツに付与された1又は複数のタグに基づいて、優先度を算出してもよい。
- [0013] この態様によれば、ユーザの人気度及びコンテンツの人気度両方を考慮して新たなコンテンツに付与するタグを推薦することができる。
- [0014] 上記態様において、第3数値及び第6数値は、複数のタグの数の次元を有する正方行列であり、第2算出部は、第3数値及び第6数値の和並びに新たなコンテンツに付与された1又は複数のタグに基づいて、優先度を算出してもよい。

- [0015] この態様によれば、ユーザの人気度及びコンテンツの人気度両方を考慮して新たなコンテンツに付与するタグを推薦することができる。
- [0016] 上記態様において、第1算出部は、複数のタグのうち、複数のユーザが所定の期間中に複数のコンテンツに付与した複数のタグに基づいて、第3数値を算出してもよい。
- [0017] この態様によれば、所定の期間におけるタグ付けの傾向を反映するように、新たなコンテンツに付与するタグを推薦することができる。
- [0018] 上記態様において、第1算出部は、複数のユーザのうち、所定の地域に属する複数のユーザに基づいて、第3数値を算出してもよい。
- [0019] この態様によれば、所定の地域におけるタグ付けの傾向を反映するように、新たなコンテンツに付与するタグを推薦することができる。
- [0020] 上記態様において、第1算出部は、複数のユーザのうち、所定の年代に属する複数のユーザに基づいて、第3数値を算出してもよい。
- [0021] この態様によれば、所定の年代におけるタグ付けの傾向を反映するように、新たなコンテンツに付与するタグを推薦することができる。
- [0022] 上記態様において、第1算出部は、複数のユーザのうち、所定の性別に属する複数のユーザに基づいて、第3数値を算出してもよい。
- [0023] この態様によれば、所定の性別におけるタグ付けの傾向を反映するように、新たなコンテンツに付与するタグを推薦することができる。
- [0024] 上記態様において、複数のユーザの人気度は、複数のユーザのSNSにおけるフォロワーの数、複数のユーザにより投稿された複数のコンテンツが評価された数及び複数のユーザにより投稿された複数のコンテンツが閲覧された数の少なくともいずれかに基づいて算出されてもよい。
- [0025] この態様によれば、目的に応じて適切な人気度を算出し、より魅力的なタグを推薦することができる。
- [0026] 本発明の他の態様に係るタグ推薦方法は、タグ推薦装置に、複数のユーザと複数のタグとの共起関係を、複数のユーザが複数のタグを複数のコンテンツに付与した頻度に応じて重み付けした第1数値と、複数のユーザと複数の

タグとの共起関係を、複数のユーザの人気度に応じて重み付けした第2数値とに基づいて、複数のタグ間の共起関係を表す第3数値を算出することと、第3数値及び新たなコンテンツに付与された1又は複数のタグに基づいて、複数のタグの優先度を算出することと、優先度に基づいて、複数のタグの中から、新たなコンテンツに付与する1又は複数のタグを推薦することと、を実行させる。

[0027] この態様によれば、複数のタグ間の共起関係を表す第3数値を、ユーザがタグを付与した頻度と、ユーザの人気度とを用いて算出し、第3数値と、新たなコンテンツに予め付与された1又は複数のタグとに基づいて複数のタグの優先度を算出することで、ユーザの人気度を考慮して新たなコンテンツに付与するタグを推薦することができる。

[0028] 本発明の他の態様に係るタグ推薦プログラムは、タグ推薦装置に、複数のユーザと複数のタグとの共起関係を、複数のユーザが複数のタグを複数のコンテンツに付与した頻度に応じて重み付けした第1数値と、複数のユーザと複数のタグとの共起関係を、複数のユーザの人気度に応じて重み付けした第2数値とに基づいて、複数のタグ間の共起関係を表す第3数値を算出することと、第3数値及び新たなコンテンツに付与された1又は複数のタグに基づいて、複数のタグの優先度を算出することと、優先度に基づいて、複数のタグの中から、新たなコンテンツに付与する1又は複数のタグを推薦することと、を実行させる。

[0029] この態様によれば、複数のタグ間の共起関係を表す第3数値を、ユーザがタグを付与した頻度と、ユーザの人気度とを用いて算出し、第3数値と、新たなコンテンツに予め付与された1又は複数のタグとに基づいて複数のタグの優先度を算出することで、ユーザの人気度を考慮して新たなコンテンツに付与するタグを推薦することができる。

発明の効果

[0030] 本発明によれば、人気度が高いユーザによって付与されたタグと、その付与頻度を考慮して新たなコンテンツに付与するタグを推薦するタグ推薦装置

、タグ推薦方法及びタグ推薦プログラムを提供することができる。

図面の簡単な説明

[0031] [図1]本発明の実施形態に係るタグ推薦システムのネットワーク構成を示す図である。

[図2]本実施形態に係るタグ推薦装置の機能ブロックを示す図である。

[図3]本実施形態に係るタグ推薦装置の物理的構成を示す図である。

[図4]本実施形態に係るタグ推薦装置により算出される第1数値の一例を示す図である。

[図5]本実施形態に係るタグ推薦装置により算出される第2数値の一例を示す図である。

[図6]本実施形態に係るタグ推薦装置により実行されるタグ推薦処理の第1例のフローチャートである。

[図7]本実施形態に係るタグ推薦装置により算出される第4数値の一例を示す図である。

[図8]本実施形態に係るタグ推薦装置により算出される第5数値の一例を示す図である。

[図9]本実施形態に係るタグ推薦装置により実行されるタグ推薦処理の第2例のフローチャートである。

[図10]本実施形態に係るタグ推薦装置により実行されるタグ推薦処理の第3例のフローチャートである。

[図11]本実施形態に係るタグ推薦装置により推薦されたタグをコンテンツに付与した場合の実験結果を示す図である。

発明を実施するための形態

[0032] 添付図面を参照して、本発明の実施形態について説明する。なお、各図において、同一の符号を付したものは、同一又は同様の構成を有する。

[0033] 図1は、本発明の実施形態に係るタグ推薦システム100のネットワーク構成を示す図である。タグ推薦システム100は、タグ推薦装置10、第1ユーザ端末20、第2ユーザ端末30及び第Nユーザ端末40を含む。ここ

で、第1ユーザ端末20、第2ユーザ端末30及び第Nユーザ端末40は、第1ユーザが用いるユーザ端末、第2ユーザが用いるユーザ及び第Nユーザが用いるユーザ端末を表している。以下では、第1ユーザ端末20、第2ユーザ端末30及び第Nユーザ端末40を単にユーザ端末ともいう。

[0034] タグ推薦装置10は、ユーザ端末から投稿するコンテンツを受け付けて、そのコンテンツに付与するタグを推薦する。コンテンツは、SNSに掲載する画像、動画、テキスト及びそれらの組み合わせであったり、電子商取引サイトに掲載する商品の画像、商品の動画、商品を説明するテキスト及びそれらの組み合わせであったりしてよい。以下では、SNSに画像を投稿する場合を例に説明する。

[0035] ユーザ端末は、汎用のコンピュータであり、例えば、スマートフォン、タブレットPC (Personal Computer) 又はデスクトップPCであってよい。複数のユーザ端末は、通信ネットワークNを介してタグ推薦装置10と通信可能に接続され、タグ推薦装置10より、新たに投稿するコンテンツに付与するタグの推薦を受ける。ここで、通信ネットワークNは、有線又は無線の通信網であり、インターネットやLAN (Local Area Network) であってよい。

[0036] 図2は、本実施形態に係るタグ推薦装置10の機能ブロックを示す図である。タグ推薦装置10は、取得部11、生成部12、記憶部13、第1算出部14、第2算出部15及び推薦部16を備える。

[0037] 取得部11は、ユーザ端末から、新たに投稿するコンテンツを取得する。取得部11は、特定のSNSや電子商取引サイトと連携して、ユーザ端末からSNSや電子商取引サイトにコンテンツの投稿が行われる場合に、投稿しようとしているコンテンツを取得してよい。

[0038] 生成部12は、新たなコンテンツを解析し、コンテンツの内容を説明するタグを生成する。ここで、コンテンツの内容を説明するタグは、複数のタグの候補の中から選択されてよい。例えば、コンテンツが画像である場合、生成部12は、公知の画像認識技術を用いて、画像に写されている物の名前や

画像に写されている状況を説明するタグを生成してよい。また、生成部 12 は、コンテンツを、予め定められた複数のカテゴリーの少なくともいずれかに分類し、コンテンツが属するカテゴリーに応じて、タグを生成してもよい。生成部 12 は、1つのコンテンツに対して、コンテンツの内容を説明する複数のタグを生成してよく、例えば、10個のタグを生成してよい。

[0039] 記憶部 13 は、タグ情報 13 a、ユーザ情報 13 b 及びコンテンツ情報 13 c を記憶している。タグ情報 13 a は、これまでにコンテンツに付与されたタグの履歴情報を含み、どのタグがどのコンテンツに、いつ、どの地域で、どの年代のユーザによって付与されているかを示す情報を含んでよい。また、ユーザ情報 13 b は、コンテンツを投稿したユーザの人気度を含み、ユーザの居住する地域や年代といったプロフィール情報を含んでもよい。

[0040] ここで、ユーザの人気度は、ユーザの SNS におけるフォロワーの数、ユーザにより投稿された複数のコンテンツが評価された数及びユーザにより投稿された複数のコンテンツが閲覧された数の少なくともいずれかに基づいて算出されてよい。タグ推薦を受けるユーザは、どのような指標を用いてユーザの人気度を算出するか選択できてもよい。例えば、より閲覧数が増えるようにタグを付与したい場合、ユーザの人気度をユーザにより投稿された複数のコンテンツが閲覧された数に基づいて算出することとしてよい。このように、目的に応じて適切な人気度を算出し、より魅力的なタグを推薦することができる。

[0041] コンテンツ情報 13 c は、投稿されたコンテンツの人気度を含み、いずれのユーザによって投稿されたコンテンツであるかを示す情報を含んでもよい。コンテンツの人気度は、投稿された複数のコンテンツが評価された数及び投稿された複数のコンテンツが閲覧された数の少なくともいずれかに基づいて算出されてよい。

[0042] 第 1 算出部 14 は、複数のユーザと複数のタグとの共起関係を、複数のユーザの人気度に応じて重み付けした第 1 数値と、複数のユーザと複数のタグとの共起関係を、複数のユーザが複数のタグを複数のコンテンツに付与した

頻度に応じて重み付けした第2数値とに基づいて、複数のタグ間の共起関係を表す第3数値を算出する。第1数値及び第2数値の具体例は、図4及び5を用いて詳細に説明する。

[0043] 第2算出部15は、第3数値及び新たなコンテンツに付与された1又は複数のタグに基づいて、複数のタグの優先度を算出する。ここで、新たなコンテンツに付与された1又は複数のタグは、生成部12によって生成されたタグであってよい。タグの優先度については、後に図6を用いて詳細に説明する。

[0044] 推薦部16は、第2算出部15により算出された優先度に基づいて、複数のタグの中から、新たなコンテンツに付与する1又は複数のタグを推薦する。推薦部16は、推薦する1又は複数のタグを特定する情報を、ユーザ端末に送信してよい。ユーザ端末は、推薦された1又は複数のタグ全てを新たなコンテンツに付与してもよいし、推薦された1又は複数のタグのうち、新たなコンテンツに付与するタグの選択をユーザから受け付けて、選択されたタグを新たなコンテンツに付与してもよい。

[0045] 図3は、本実施形態に係るタグ推薦装置10の物理的構成を示す図である。タグ推薦装置10は、演算部に相当するCPU (Central Processing Unit) 10aと、記憶部に相当するRAM (Random Access Memory) 10bと、記憶部に相当するROM (Read only Memory) 10cと、通信部10dと、入力部10eと、表示部10fと、を有する。これらの各構成は、バスを介して相互にデータ送受信可能に接続される。なお、本例ではタグ推薦装置10が一台のコンピュータで構成される場合について説明するが、タグ推薦装置10は、複数のコンピュータが組み合わされて実現されてもよい。また、図3で示す構成は一例であり、タグ推薦装置10はこれら以外の構成を有してもよいし、これらの構成のうち一部を有さなくてもよい。

[0046] CPU 10aは、RAM 10b又はROM 10cに記憶されたプログラムの実行に関する制御やデータの演算、加工を行う制御部である。CPU 10aは、新たなコンテンツに付与するタグを推薦するプログラム（タグ推薦プ

プログラム)を実行する演算部である。CPU 10aは、入力部10eや通信部10dから種々のデータを受け取り、データの演算結果を表示部10fに表示したり、RAM 10bやROM 10cに格納したりする。

[0047] RAM 10bは、記憶部のうちデータの書き換えが可能なものであり、例えば半導体記憶素子で構成されてよい。RAM 10bは、CPU 10aが実行するプログラム、ユーザの人気度、タグの付与頻度及びコンテンツの人気度といったデータを記憶してよい。なお、これらは例示であって、RAM 10bには、これら以外のデータが記憶されていてもよいし、これらの一部が記憶されていなくてもよい。

[0048] ROM 10cは、記憶部のうちデータの読み出しが可能なものであり、例えば半導体記憶素子で構成されてよい。ROM 10cは、例えばタグ推薦プログラムや、書き換えが行われないデータを記憶してよい。

[0049] 通信部10dは、タグ推薦装置10を他の機器に接続するインターフェースである。通信部10dは、インターネット等の通信ネットワークNに接続されてよい。

[0050] 入力部10eは、ユーザからデータの入力を受け付けるものであり、例えば、キーボード及びタッチパネルを含んでよい。

[0051] 表示部10fは、CPU 10aによる演算結果を視覚的に表示するものであり、例えば、LCD (Liquid Crystal Display) により構成されてよい。表示部10fは、タグの優先度や推薦するタグを表示してよい。

[0052] タグ推薦プログラムは、RAM 10bやROM 10c等のコンピュータによって読み取り可能な記憶媒体に記憶されて提供されてもよいし、通信部10dにより接続される通信ネットワークを介して提供されてもよい。タグ推薦装置10では、CPU 10aがタグ推薦プログラムを実行することにより、図2を用いて説明した様々な動作が実現される。なお、これらの物理的な構成は例示であって、必ずしも独立した構成でなくてもよい。例えば、タグ推薦装置10は、CPU 10aとRAM 10bやROM 10cが一体化したLSI (Large-Scale Integration) を備えていてもよい。

[0053] 図4は、本実施形態に係るタグ推薦装置10により算出される第1数値 U_p の一例を示す図である。第1数値 U_p は、コンテンツへの付与候補となる複数のタグの数の行と、複数のユーザの数の列を有する行列である。本例では、複数のタグの数が3であり、複数のユーザの数が3である場合について説明する。第1数値 U_p は、本発明の第1数値に相当し、複数のユーザと複数のタグとの共起関係を、複数のユーザの人気度に応じて重み付けした第1数値の一例である。

[0054] 第1数値 U_p は、複数のユーザの数の次元を有する行ベクトルを、複数のタグの数だけ含む。以下では、第1数値 U_p に含まれる複数の行ベクトルを第 i 行ベクトルと呼ぶ。ここで、 i は、1から M の値をとり、 M は複数のタグの数である。本例の場合、3次元の第1行ベクトルは、「a」という第1タグ T_1 に対応し、第2行ベクトルは、「b」という第2タグ T_2 に対応し、第3行ベクトルは、「c」という第3タグ T_3 に対応する。また、本例において、第1ユーザの人気度は「100」であり、第2ユーザの人気度は「200」であり、第3ユーザの人気度は「300」である。第1行ベクトルは、第1要素が $100/600$ であり、第2要素が $200/600$ であり、第3要素が $300/600$ である。また、第2行ベクトルは、第1要素が $100/400$ であり、第2要素が0であり、第3要素が $300/400$ である。また、第3行ベクトルは、第1要素が $100/300$ であり、第2要素が $200/300$ であり、第3要素が0である。各行ベクトルは、 L_1 ノルムが1となるように規格化されている。

[0055] 第 i 行ベクトルの値は、第 j ユーザ（ $j = 1 \sim N$ 、 N は複数のユーザの数）が第 i タグをコンテンツに付与している場合にゼロでない値となり、第 j ユーザが第 i タグをコンテンツに付与していない場合にゼロとなる。そして、第 i 行ベクトルのゼロでない値は、第 j ユーザの人気度に比例するように定められる。本例の場合、第 i 行ベクトルのゼロでない値は、第 j ユーザの人気度に比例し、かつ、第 i 行ベクトルの L_1 ノルムが1となるように定められている。

[0056] 本例の場合、第1タグT1は、第1ユーザ、第2ユーザ及び第3ユーザ全てによって何らかのコンテンツに付与された履歴があるため、第1行ベクトルの全ての要素がゼロでない値となっている。一方、第2タグT2は、第1ユーザ及び第3ユーザによって付与された履歴があるが、第2ユーザによっては付与されていないため、第2行ベクトルの第2要素はゼロとなっている。同様に、第3タグT3は、第1ユーザ及び第2ユーザによって付与された履歴があるが、第3ユーザによっては付与されていないため、第3行ベクトルの第3要素はゼロとなっている。

[0057] 図5は、本実施形態に係るタグ推薦装置10により算出される第2数値 U_t の一例を示す図である。第2数値 U_t は、コンテンツへの付与候補となる複数のタグの数の行と、複数のユーザの数の列を有する行列である。本例では、複数のタグの数が3であり、複数のユーザの数が3である場合について説明する。第2数値 U_t は、本発明の第2数値に相当し、複数のユーザと複数のタグとの共起関係を、複数のユーザが複数のタグを複数のコンテンツに付与した頻度に応じて重み付けした第2数値の一例である。

[0058] 第2数値 U_t は、複数のタグの数の次元を有する列ベクトルを、複数のユーザの数だけ含む。以下では、第2数値 U_t に含まれる複数の列ベクトルを第 k 列ベクトルと呼ぶ。ここで、 k は、1から N の値をとり、 N は複数のユーザの数である。本例の場合、3次元の第1列ベクトルは第1ユーザに対応し、第2列ベクトルは第2ユーザに対応し、第3行ベクトルは第3ユーザに対応する。例えば、第1列ベクトルは、第1要素が $2/6$ であり、第2要素が $1/6$ であり、第3要素が $3/6$ である。また、第2列ベクトルは、第1要素が $2/3$ であり、第2要素が0であり、第3要素が $1/3$ である。また、第3行ベクトルは、第1要素が $3/6$ であり、第2要素が $3/6$ であり、第3要素が0である。各列ベクトルは、L1ノルムが1となるように規格化されている。

[0059] 第 k 列ベクトルの値は、第 k ユーザが第 i タグをコンテンツに付与している場合にゼロでない値となり、第 k ユーザが第 i タグをコンテンツに付与し

ていない場合にゼロとなる。そして、第 k 列ベクトルのゼロでない値は、第 k ユーザが第 i タグを付与した回数に比例するように定められる。本例の場合、第 k 列ベクトルのゼロでない値は、第 k ユーザが第 i タグを付与した回数に比例し、かつ、第 k 列ベクトルの L_1 ノルムが 1 となるように定められている。

[0060] 本例の場合、第 1 ユーザは、第 1 タグを 2 回、第 2 タグを 1 回、第 3 タグを 3 回付与した履歴があるため、第 1 列ベクトルの全ての要素がゼロでない値となっている。一方、第 2 ユーザは、第 1 タグを 2 回、第 2 タグを 0 回、第 3 タグを 1 回付与した履歴があるため、第 2 列ベクトルの第 2 要素はゼロとなっている。同様に、第 3 ユーザは、第 1 タグを 3 回、第 2 タグを 3 回、第 3 タグを 0 回付与した履歴があるため、第 3 列ベクトルの第 3 要素はゼロとなっている。

[0061] 図 6 は、本実施形態に係るタグ推薦装置 10 により実行されるタグ推薦処理の第 1 例のフローチャートである。はじめに、タグ推薦装置 10 は、ユーザが投稿するコンテンツを取得する (S10)。そして、タグ推薦装置 10 は、生成部 12 によってコンテンツを説明するタグを生成し、優先ベクトル p_1 を決定する (S11)。

[0062] 優先ベクトル p_1 は、コンテンツに付与される候補となる複数のタグの数の次元を有する M 次元ベクトルであり (M は複数のタグの数)、生成部 12 によって生成されたコンテンツを説明するタグに対応する要素がゼロでない値 (例えば 1) をとり、それ以外の要素についてゼロの値をとるベクトルであってよい。また、優先ベクトル p_1 は、 L_1 ノルムが m (m は生成部 12 によって生成されたコンテンツを説明するタグの数) に規格化されていてよい。例えば、生成部 12 によって生成されたコンテンツを説明するタグが、第 1 タグ及び第 2 タグである場合、優先ベクトル p_1 は、第 1 タグに対応する要素が 1、第 2 タグに対応する要素が 1 であり、それ以外の要素がゼロであるベクトルであってよい。

[0063] その後、タグ推薦装置 10 は、第 1 数値及び第 2 数値である行列 U_p 及び

U_tを算出し、第3数値 $A_{UP} = U_p \times U_t^T$ を算出する(S12)。ここで、「T」は行列の転置を表す。第3数値 A_{UP} は、複数のタグの数の次元を有する正方行列である。第3数値 A_{UP} は、本発明の第3数値に相当し、複数のタグ間の共起関係を表す。

[0064] さらに、タグ推薦装置10は、複数のタグの優先度 r_1 を、 $r_1 = \alpha A_{UP} r_1 + (1 - \alpha) p_1$ により算出する(S13)。ここで、優先度 r_1 の算出式は再帰的に定義されている。また、 α は、0から1の間の値である。 r_1 は、L1ノルムが m (m は生成部12によって生成されたコンテンツを説明するタグの数)に規格化されてよい。なお、 r_1 の初期値は、 p_1 であってよい。

[0065] また、タグ推薦装置10は、基準優先度 r_0 を、 $r_0 = \alpha A_{UP} r_0 + (1 - \alpha) p_0$ により算出する(S14)。ここで、 p_0 は、複数のタグの数の次元を有するM次元ベクトルであり(Mは複数のタグの数)、全ての要素が等しい値であり、L1ノルムが m (m は生成部12によって生成されたコンテンツを説明するタグの数)に規格化されているベクトルであってよい。すなわち、 p_0 は、全ての要素が m/M であるベクトルであってよい。 r_0 は、L1ノルムが m に規格化されてよい。なお、 r_0 の初期値は、 p_0 であってよい。

[0066] タグ推薦装置10は、優先度 r_1 及び基準優先度 r_0 が収束したか判定する(S15)。収束したか否かは、処理S13及びS14における更新で、優先度 r_1 及び基準優先度 r_0 が所定量以下しか変化しないか否か、又は予め指定した繰り返し演算回数を上回ったかによって判定してよい。処理S13及びS14における更新で、優先度 r_1 又は基準優先度 r_0 が所定量より大きく変化し、予め指定した繰り返し演算回数を上回っていない場合、すなわち優先度 r_1 及び基準優先度 r_0 が収束していない場合(S15:NO)、タグ推薦装置10は、処理S13及びS14を再び実行する。

[0067] 一方、処理S13及びS14における更新で、優先度 r_1 又は基準優先度 r_0 が所定量以下しか変化しない場合、又は予め指定した繰り返し演算回数を上回った場合、すなわち優先度 r_1 及び基準優先度 r_0 が収束した場合(S15:YES)、タグ推薦装置10は、 $r_1 - r_0$ の値により推薦する1又は複数

のタグを決定する（S16）。 $r_1 - r_0$ は、複数のタグの数の次元を有するベクトルであり、各要素は複数のタグに対応する。タグ推薦装置10は、 $r_1 - r_0$ の要素の大きさが上位所定順位以内である1又は複数のタグを推薦してよい。以上により推薦処理の第1例が終了する。

[0068] このように、複数のタグ間の共起関係を表す第3数値 A_{UP} を、ユーザがタグを付与した頻度と、ユーザの人気度とを用いて算出し、第3数値と、新たなコンテンツに予め付与された1又は複数のタグとに基づいて複数のタグの優先度 r_1 を算出することで、人気度が高いユーザによって付与されたタグと、その付与頻度を考慮して新たなコンテンツに付与するタグを推薦することができる。

[0069] 図7は、本実施形態に係るタグ推薦装置10により算出される第4数値 B_w の一例を示す図である。第4数値 B_w は、コンテンツへの付与候補となる複数のタグの数の行と、複数のコンテンツの数の列を有する行列である。本例では、複数のタグの数が3であり、複数のコンテンツの数が3である場合について説明する。第4数値 B_w は、本発明の第4数値に相当し、複数のコンテンツと複数のタグとの共起関係を、複数のコンテンツの人気度に応じて重み付けした第5数値の一例である。

[0070] 第4数値 B_w は、複数のコンテンツの数の次元を有する行ベクトルを、複数のタグの数だけ含む。以下では、第4数値 B_w に含まれる複数の行ベクトルを第 i 行ベクトルと呼ぶ。ここで、 i は、1から M の値をとり、 M は複数のタグの数である。本例の場合、3次元の第1行ベクトルは、「a」という第1タグ T_1 に対応し、第2行ベクトルは、「b」という第2タグ T_2 に対応し、第3行ベクトルは、「c」という第3タグ T_3 に対応する。また、本例において、第1コンテンツの人気度は「40」であり、第2コンテンツの人気度は「7」であり、第3コンテンツの人気度は「4」である。そして、第1行ベクトルは、第1要素が $40/50$ であり、第2要素が $7/50$ であり、第3要素が $3/50$ である。また、第2行ベクトルは、第1要素が $40/47$ であり、第2要素が $7/47$ であり、第3要素が0である。また、第

3行ベクトルは、第1要素が0であり、第2要素が7／10であり、第3要素が3／10である。各行ベクトルは、L1ノルムが1となるように規格化されている。

[0071] 第*i*行ベクトルの値は、第*m*コンテンツ（*m*＝1～*C*、*C*は複数のコンテンツの数）に第*i*タグが付与されている場合にゼロでない値となり、第*m*コンテンツに第*i*タグが付与されていない場合にゼロとなる。そして、第*i*行ベクトルのゼロでない値は、第*m*コンテンツの人気度に比例するように定められる。本例の場合、第*i*行ベクトルのゼロでない値は、第*m*コンテンツの人気度に比例し、かつ、第*i*行ベクトルのL1ノルムが1となるように定められている。

[0072] 本例の場合、第1タグT1は、第1コンテンツ、第2コンテンツ及び第3コンテンツ全てに付与された履歴があるため、第1行ベクトルの全ての要素がゼロでない値となっている。一方、第2タグT2は、第1コンテンツ及び第2コンテンツに付与された履歴があるが、第3コンテンツには付与されていないため、第2行ベクトルの第3要素はゼロとなっている。同様に、第3タグT3は、第2コンテンツ及び第3コンテンツに付与された履歴があるが、第1コンテンツには付与されていないため、第3行ベクトルの第1要素はゼロとなっている。

[0073] 図8は、本実施形態に係るタグ推薦装置10により算出される第5数値B_tの一例を示す図である。第5数値B_tは、コンテンツへの付与候補となる複数のタグの数の行と、複数のコンテンツの数の列を有する行列である。本例では、複数のタグの数が3であり、複数のコンテンツの数が3である場合について説明する。第5数値B_tは、本発明の第5数値に相当し、複数のコンテンツと複数のタグとの共起関係を、複数のタグが複数のコンテンツに付与された頻度に応じて重み付けした第5数値の一例である。

[0074] 第5数値B_tは、複数のタグの数の次元を有する列ベクトルを、複数のコンテンツの数だけ含む。以下では、第5数値B_tに含まれる複数の列ベクトルを第*k*列ベクトルと呼ぶ。ここで、*k*は、1から*C*の値をとり、*C*は複数

のコンテンツの数である。本例の場合、3次元の第1列ベクトルは第1コンテンツに対応し、第2列ベクトルは第2コンテンツに対応し、第3行ベクトルは第3コンテンツに対応する。例えば、第1列ベクトルは、第1要素が $1/2$ であり、第2要素が $1/2$ であり、第3要素が0である。また、第2列ベクトルは、第1要素が $1/3$ であり、第2要素が $1/3$ であり、第3要素が $1/3$ である。また、第3行ベクトルは、第1要素が $1/2$ であり、第2要素が0であり、第3要素が $1/2$ である。各列ベクトルは、 L_1 ノルムが1となるように規格化されている。

[0075] 第 k 列ベクトルの値は、第 k コンテンツに第 i タグが付与されている場合にゼロでない値となり、第 k コンテンツに第 i タグが付与されていない場合にゼロとなる。本例の場合、第 k 列ベクトルのゼロでない値は、第 k 列ベクトルの L_1 ノルムが1となるように均等に定められている。

[0076] 本例の場合、第1コンテンツには、第1タグ及び第2タグが付与された履歴があるが、第3タグが付与された履歴がないため、第1列ベクトルの第3要素はゼロとなっている。一方、第2コンテンツには、第1タグ、第2タグ及び第3タグが付与された履歴があるため、第2列ベクトルの全ての要素はゼロでない値となっている。また、第3コンテンツには、第1タグ及び第3タグが付与された履歴があるが、第2タグが付与された履歴がないため、第3列ベクトルの第2要素はゼロとなっている。

[0077] 図9は、本実施形態に係るタグ推薦装置10により実行されるタグ推薦処理の第2例のフローチャートである。はじめに、タグ推薦装置10は、ユーザが投稿するコンテンツを取得する(S20)。そして、タグ推薦装置10は、生成部12によってコンテンツを説明するタグを生成し、優先ベクトル p_1 を決定する(S21)。優先ベクトル p_1 は、コンテンツに付与される候補となる複数のタグの数の次元を有する M 次元ベクトルであり(M は複数のタグの数)、生成部12によって生成されたコンテンツを説明するタグに対応する要素が1であり、それ以外の要素がゼロであるベクトルであってよい。また、優先ベクトル p_1 は、 L_1 ノルムが m (m は生成部12によって生成

されたコンテンツを説明するタグの数)に規格化されていてよい。

[0078] その後、タグ推薦装置10は、第1数値及び第2数値である行列 U_p 及び U_t を算出し、第3数値 $A_{UP} = U_p \times U_t^T$ を算出する(S22)。第3数値 A_{UP} は、複数のタグの数の次元を有する正方行列である。第3数値 A_{UP} は、本発明の第3数値に相当し、複数のタグ間の共起関係を表す。

[0079] また、タグ推薦装置10は、第4数値及び第5数値である行列 B_p 及び B_t を算出し、第6数値 $A_{FP} = B_p \times B_t^T$ を算出する(S23)。第6数値 A_{FP} は、複数のタグの数の次元を有する正方行列である。第6数値 A_{FP} は、本発明の第6数値に相当し、複数のタグ間の共起関係を表す。

[0080] タグ推薦装置10は、第3数値 A_{UP} 及び第6数値 A_{FP} に基づいて、第7数値 $A_{UFP}^{(1)}$ を以下の数式(1)により算出する(S24)。

[0081] [数1]

$$A_{UFP}^{(1)} = A_{UP} \odot A_{FP}$$

[0082] ここで、数式(1)の右辺は、第3数値 A_{UP} 及び第6数値 A_{FP} の要素積(アダマール積)を表す。

[0083] さらに、タグ推薦装置10は、複数のタグの優先度 r_1 を、 $r_1 = \alpha A_{UFP}^{(1)} r_1 + (1 - \alpha) p_1$ により算出する(S25)。ここで、優先度 r_1 の算出式は再帰的に定義されている。また、 α は、0から1の間の値である。 r_1 は、 L_1 ノルムが m (m は生成部12によって生成されたコンテンツを説明するタグの数)に規格化されていてよい。なお、 r_1 の初期値は、 p_1 であってよい。

[0084] また、タグ推薦装置10は、基準優先度 r_0 を、 $r_0 = \alpha A_{UFP}^{(1)} r_0 + (1 - \alpha) p_0$ により算出する(S26)。ここで、 p_0 は、複数のタグの数の次元を有する M 次元ベクトルであり(M は複数のタグの数)、全ての要素が等しい値であり、 L_1 ノルムが m (m は生成部12によって生成されたコンテンツを説明するタグの数)に規格化されているベクトルであってよい。 r_0

は、 L_1 ノルムが m に規格化されてよい。なお、 r_0 の初期値は、 p_0 であってよい。

[0085] タグ推薦装置 10 は、優先度 r_1 及び基準優先度 r_0 が収束したか判定する (S27)。収束したか否かは、処理 S25 及び S26 における更新で、優先度 r_1 及び基準優先度 r_0 が所定量以下しか変化しないか否か、又は予め指定した繰り返し演算回数を上回ったかによって判定してよい。処理 S25 及び S26 における更新で、優先度 r_1 又は基準優先度 r_0 が所定量より大きく変化し、予め指定した繰り返し演算回数を上回っていない場合、すなわち優先度 r_1 及び基準優先度 r_0 が収束していない場合 (S27: NO)、タグ推薦装置 10 は、処理 S25 及び S26 を再び実行する。

[0086] 一方、処理 S25 及び S26 における更新で、優先度 r_1 又は基準優先度 r_0 が所定量以下しか変化しない場合、又は予め指定した繰り返し演算回数を上回った場合、すなわち優先度 r_1 及び基準優先度 r_0 が収束した場合 (S27: YES)、タグ推薦装置 10 は、 $r_1 - r_0$ の値により推薦する 1 又は複数のタグを決定する (S28)。 $r_1 - r_0$ は、複数のタグの数の次元を有するベクトルであり、各要素は複数のタグに対応する。タグ推薦装置 10 は、 $r_1 - r_0$ の要素の大きさが上位所定順位以内である 1 又は複数のタグを推薦してよい。以上により推薦処理の第 2 例が終了する。

[0087] このように、複数のタグ間の共起関係を表す第 6 数値 A_{FP} を、タグがコンテンツに付与された頻度と、コンテンツの人気度とを用いて算出し、第 6 数値と、新たなコンテンツに予め付与された 1 又は複数のタグとに基づいて複数のタグの優先度 r_1 を算出することで、人気度が高いコンテンツに付与されたタグと、その付与頻度を考慮して新たなコンテンツに付与するタグを推薦することができる。

[0088] また、第 3 数値 A_{UP} 及び第 6 数値 A_{FP} の要素積並びに新たなコンテンツに付与された 1 又は複数のタグに基づいて、優先度 r_1 を算出することで、ユーザの人気度及びコンテンツの人気度両方を考慮して新たなコンテンツに付与するタグを推薦することができる。

- [0089] 図10は、本実施形態に係るタグ推薦装置10により実行されるタグ推薦処理の第3例のフローチャートである。はじめに、タグ推薦装置10は、ユーザが投稿するコンテンツを取得する(S30)。そして、タグ推薦装置10は、生成部12によってコンテンツを説明するタグを生成し、優先ベクトル p_1 を決定する(S31)。優先ベクトル p_1 は、コンテンツに付与される候補となる複数のタグの数の次元を有するM次元ベクトルであり(Mは複数のタグの数)、生成部12によって生成されたコンテンツを説明するタグに対応する要素が1であり、それ以外の要素がゼロであるベクトルであってよい。また、優先ベクトル p_1 は、L1ノルムが m (m は生成部12によって生成されたコンテンツを説明するタグの数)に規格化されていてよい。
- [0090] その後、タグ推薦装置10は、第1数値及び第2数値である行列 U_p 及び U_t を算出し、第3数値 $A_{UP} = U_p \times U_t^T$ を算出する(S32)。第3数値 A_{UP} は、複数のタグの数の次元を有する正方行列である。第3数値 A_{UP} は、本発明の第3数値に相当し、複数のタグ間の共起関係を表す。
- [0091] また、タグ推薦装置10は、第4数値及び第5数値である行列 B_w 及び B_t を算出し、第6数値 $A_{FP} = B_w \times B_t^T$ を算出する(S33)。第6数値 A_{FP} は、複数のタグの数の次元を有する正方行列である。第6数値 A_{FP} は、本発明の第6数値に相当し、複数のタグ間の共起関係を表す。
- [0092] タグ推薦装置10は、第3数値 A_{UP} 及び第6数値 A_{FP} に基づいて、第8数値 $A_{UFP}^{(2)}$ を $A_{UFP}^{(2)} = A_{UP} + A_{FP}$ によって算出する(S34)。
- [0093] さらに、タグ推薦装置10は、複数のタグの優先度 r_1 を、 $r_1 = \alpha A_{UFP}^{(2)} r_1 + (1 - \alpha) p_1$ により算出する(S35)。ここで、優先度 r_1 の算出式は再帰的に定義されている。また、 α は、0から1の間の値である。 r_1 は、L1ノルムが m (m は生成部12によって生成されたコンテンツを説明するタグの数)に規格化されてよい。なお、 r_1 の初期値は、 p_1 であってよい。
- [0094] また、タグ推薦装置10は、基準優先度 r_0 を、 $r_0 = \alpha A_{UFP}^{(2)} r_0 + (1 - \alpha) p_0$ により算出する(S36)。ここで、 p_0 は、複数のタグの数の

次元を有するM次元ベクトルであり（Mは複数のタグの数）、全ての要素が等しい値であり、L1ノルムがm（mは生成部12によって生成されたコンテンツを説明するタグの数）に規格化されているベクトルであってよい。 r_0 は、L1ノルムがmに規格化されてよい。なお、 r_0 の初期値は、 p_0 であってよい。

[0095] タグ推薦装置10は、優先度 r_1 及び基準優先度 r_0 が収束したか判定する（S37）。収束したか否かは、処理S35及びS36における更新で、優先度 r_1 及び基準優先度 r_0 が所定量以下しか変化しないか否か、又は予め指定した繰り返し演算回数を上回ったかによって判定してよい。処理S35及びS36における更新で、優先度 r_1 又は基準優先度 r_0 が所定量より大きく変化し、予め指定した繰り返し演算回数を上回っていない場合、すなわち優先度 r_1 及び基準優先度 r_0 が収束していない場合（S37：NO）、タグ推薦装置10は、処理S35及びS36を再び実行する。

[0096] 一方、処理S35及びS36における更新で、優先度 r_1 又は基準優先度 r_0 が所定量以下しか変化しない場合、又は予め指定した繰り返し演算回数を上回った場合、すなわち優先度 r_1 及び基準優先度 r_0 が収束した場合（S37：YES）、タグ推薦装置10は、 $r_1 - r_0$ の値により推薦する1又は複数のタグを決定する（S38）。 $r_1 - r_0$ は、複数のタグの数の次元を有するベクトルであり、各要素は複数のタグに対応する。タグ推薦装置10は、 $r_1 - r_0$ の要素の大きさが上位所定順位以内である1又は複数のタグを推薦してよい。以上により推薦処理の第3例が終了する。

[0097] このように、第3数値 A_{UP} 及び第6数値 A_{FP} の和並びに新たなコンテンツに付与された1又は複数のタグに基づいて、優先度 r_1 を算出することで、ユーザの人気度及びコンテンツの人気度両方を考慮して新たなコンテンツに付与するタグを推薦することができる。

[0098] 以上の例では、複数のタグを制限なく用いる場合について説明したが、タグ推薦装置10は、その他のタグの推薦を行ってもよい。第1算出部14は、複数のタグのうち、複数のユーザが所定の期間中に複数のコンテンツに付

与した複数のタグに基づいて、第3数値 A_{UP} を算出してもよい。同様に、第1算出部14は、複数のタグのうち、複数のユーザが所定の期間中に複数のコンテンツに付与した複数のタグに基づいて、第6数値 A_{FP} を算出してもよい。このようにして、所定の期間におけるタグ付けの傾向を反映するように、新たなコンテンツに付与するタグを推薦することができる。例えば、ある時期の流行を捉えたタグの推薦を行うことができる。

[0099] 第1算出部14は、複数のユーザのうち、所定の地域に属する複数のユーザに基づいて、第3数値 A_{UP} を算出してもよい。同様に、第1算出部14は、複数のユーザのうち、所定の地域に属する複数のユーザに基づいて、第6数値 A_{FP} を算出してもよい。ここで、所定の地域に属する複数のユーザは、ユーザの居住する地域によって判断されてよい。コンテンツにどのようなタグを付与すれば人気となるかは、国や地域によって傾向が異なることがある。ユーザの地域性を考慮することで、所定の地域におけるタグ付けの傾向を反映するように、新たなコンテンツに付与するタグを推薦することができる。

[0100] 第1算出部14は、複数のユーザのうち、所定の年代に属する複数のユーザに基づいて、第3数値 A_{UP} を算出してもよい。同様に、第1算出部14は、複数のユーザのうち、所定の年代に属する複数のユーザに基づいて、第6数値 A_{FP} を算出してもよい。ここで、所定の年代に属する複数のユーザは、ユーザのプロフィール情報によって判断されてよい。コンテンツにどのようなタグを付与すれば人気となるかは、ユーザの年代によって傾向が異なることがある。ユーザの年代を考慮することで、所定の年代におけるタグ付けの傾向を反映するように、新たなコンテンツに付与するタグを推薦することができる。

[0101] 第1算出部14は、複数のユーザのうち、所定の性別に属する複数のユーザに基づいて、第3数値 A_{UP} を算出してもよい。同様に、第1算出部14は、複数のユーザのうち、所定の性別に属する複数のユーザに基づいて、第6数値 A_{FP} を算出してもよい。ここで、所定の性別に属する複数のユーザは、

ユーザのプロフィール情報によって判断されてよい。コンテンツにどのようなタグを付与すれば人気となるかは、ユーザの性別によって傾向が異なることがある。ユーザの性別を考慮することで、所定の性別におけるタグ付けの傾向を反映するように、新たなコンテンツに付与するタグを推薦することができる。

[0102] 図11は、本実施形態に係るタグ推薦装置10により推薦されたタグをコンテンツに付与した場合の実験結果を示す図である。同図では、特定のSNSに複数の画像を投稿する実験について、縦軸に平均閲覧数（Average Number of Views）を示し、横軸に投稿からの経過日数（Days after uploaded）を示している。実験は、本実施形態に係るタグ推薦装置10により推薦されたタグを画像に付与した場合と、比較例の場合とについて行い、それぞれ異なる期間に、異なるユーザによって、複数種類の画像を投稿した場合の平均を示している。

[0103] 実線で示す第1グラフG1は、生成部12によって生成した10個のタグに基づき優先ベクトル p_1 を決定し、第7数値 $A_{UFP}^{(1)}$ を用いた更新式で優先度 r_1 を算出し、10個のタグの推薦を行い、合計20個のタグを付与して画像の投稿を行った場合の閲覧数の時間変化を示している。同図によると、第1グラフG1は最も良い成績を示し、10日間で平均閲覧数が50以上となり、増加傾向が続いている。

[0104] 破線で示す第2グラフG2は、生成部12によって生成した10個のタグに基づき優先ベクトル p_1 を決定し、第6数値 A_{FP} を用いた更新式で優先度 r_1 を算出し、推薦部16によって10個のタグの推薦を行い、合計20個のタグを付与して画像の投稿を行った場合の閲覧数の時間変化を示している。同図によると、第2グラフG2は2番目に良い成績を示し、10日間で平均閲覧数が40以上となり、増加傾向が続いている。

[0105] 一点鎖線で示す第3グラフG3は、生成部12によって生成した10個のタグに基づき優先ベクトル p_1 を決定し、第8数値 $A_{UFP}^{(2)}$ を用いた更新式で優先度 r_1 を算出し、推薦部16によって10個のタグの推薦を行い、合計

20個のタグを付与して画像の投稿を行った場合の閲覧数の時間変化を示している。同図によると、第3グラフG3は3番目に良い成績を示し、10日間で平均閲覧数が30程度となり、増加傾向が続いている。

[0106] 二点鎖線で示す第4グラフG4は、生成部12によって生成した10個のタグに基づき優先ベクトル p_1 を決定し、第3数値 A_{UP} を用いた更新式で優先度 r_1 を算出し、推薦部16によって10個のタグの推薦を行い、合計20個のタグを付与して画像の投稿を行った場合の閲覧数の時間変化を示している。同図によると、第4グラフG4は4番目に良い成績を示し、10日間で平均閲覧数が30程度となり、増加傾向が続いている。

[0107] 点線で示す第5グラフG5は、生成部12によって生成した10個のタグを付与して画像の投稿を行った場合の閲覧数の時間変化を比較例として示している。同図によると、第5グラフG5は最も悪い成績を示し、10日間で平均閲覧数が20程度となっている。なお、第5グラフG5では、画像に付与されたタグの数が10であり、本実施形態の場合の半分であるが、仮に閲覧数がタグの数に比例するとしても、第1グラフG1及び第2グラフG2は、閲覧数が第5グラフG5の倍より大きくなっており、単にタグの数を増やした以上の効果を奏していることが見て取れる。

[0108] 以上説明した実施形態は、本発明の理解を容易にするためのものであり、本発明を限定して解釈するためのものではない。実施形態が備える各要素並びにその配置、材料、条件、形状及びサイズ等は、例示したものに限定されるわけではなく適宜変更することができる。また、異なる実施形態で示した構成同士を部分的に置換し又は組み合わせることが可能である。

請求の範囲

- [請求項1] 複数のユーザと複数のタグとの共起関係を、前記複数のユーザの人気度に応じて重み付けした第1数値と、前記複数のユーザと前記複数のタグとの共起関係を、前記複数のユーザが前記複数のタグを複数のコンテンツに付与した頻度に応じて重み付けした第2数値とに基づいて、前記複数のタグ間の共起関係を表す第3数値を算出する第1算出部と、
- 前記第3数値及び新たなコンテンツに付与された1又は複数のタグに基づいて、前記複数のタグの優先度を算出する第2算出部と、
- 前記優先度に基づいて、前記複数のタグの中から、前記新たなコンテンツに付与する1又は複数のタグを推薦する推薦部と、
- を備えるタグ推薦装置。
- [請求項2] 前記第1算出部は、前記複数のコンテンツと前記複数のタグとの共起関係を、前記複数のコンテンツの人気度に応じて重み付けした第4数値と、前記複数のコンテンツと前記複数のタグとの共起関係を、前記複数のタグが前記複数のコンテンツに付与された頻度に応じて重み付けした第5数値と、に基づいて、前記複数のタグ間の共起関係を表す第6数値を算出し、
- 前記第2算出部は、前記第3数値、前記第6数値及び前記新たなコンテンツに付与された1又は複数のタグに基づいて、前記優先度を算出する、
- 請求項1に記載のタグ推薦装置。
- [請求項3] 前記第3数値及び前記第6数値は、前記複数のタグの数の次元を有する正方行列であり、
- 前記第2算出部は、前記第3数値及び前記第6数値の要素積並びに前記新たなコンテンツに付与された1又は複数のタグに基づいて、前記優先度を算出する、
- 請求項2に記載のタグ推薦装置。

- [請求項4] 前記第3数値及び前記第6数値は、前記複数のタグの数の次元を有する正方行列であり、
- 前記第2算出部は、前記第3数値及び前記第6数値の和並びに前記新たなコンテンツに付与された1又は複数のタグに基づいて、前記優先度を算出する、
- 請求項2に記載のタグ推薦装置。
- [請求項5] 前記第1算出部は、前記複数のタグのうち、前記複数のユーザが所定の期間中に前記複数のコンテンツに付与した複数のタグに基づいて、前記第3数値を算出する、
- 請求項1から4のいずれか一項に記載のタグ推薦装置。
- [請求項6] 前記第1算出部は、前記複数のユーザのうち、所定の地域に属する複数のユーザに基づいて、前記第3数値を算出する、
- 請求項1から5のいずれか一項に記載のタグ推薦装置。
- [請求項7] 前記第1算出部は、前記複数のユーザのうち、所定の年代に属する複数のユーザに基づいて、前記第3数値を算出する、
- 請求項1から6のいずれか一項に記載のタグ推薦装置。
- [請求項8] 前記第1算出部は、前記複数のユーザのうち、所定の性別に属する複数のユーザに基づいて、前記第3数値を算出する、
- 請求項1から7のいずれか一項に記載のタグ推薦装置。
- [請求項9] 前記複数のユーザの人気度は、前記複数のユーザのSNSにおけるフォロワーの数、前記複数のユーザにより投稿された前記複数のコンテンツが評価された数及び前記複数のユーザにより投稿された前記複数のコンテンツが閲覧された数の少なくともいずれかに基づいて算出される、
- 請求項1から8のいずれか一項に記載のタグ推薦装置。
- [請求項10] タグ推薦装置に、
- 複数のユーザと複数のタグとの共起関係を、前記複数のユーザが前記複数のタグを複数のコンテンツに付与した頻度に応じて重み付けし

た第1数値と、前記複数のユーザと前記複数のタグとの共起関係を、前記複数のユーザの人気度に応じて重み付けした第2数値とに基づいて、前記複数のタグ間の共起関係を表す第3数値を算出することと、

前記第3数値及び新たなコンテンツに付与された1又は複数のタグに基づいて、前記複数のタグの優先度を算出することと、

前記優先度に基づいて、前記複数のタグの中から、前記新たなコンテンツに付与する1又は複数のタグを推薦することと、

を実行させるタグ推薦方法。

[請求項11]

タグ推薦装置に、

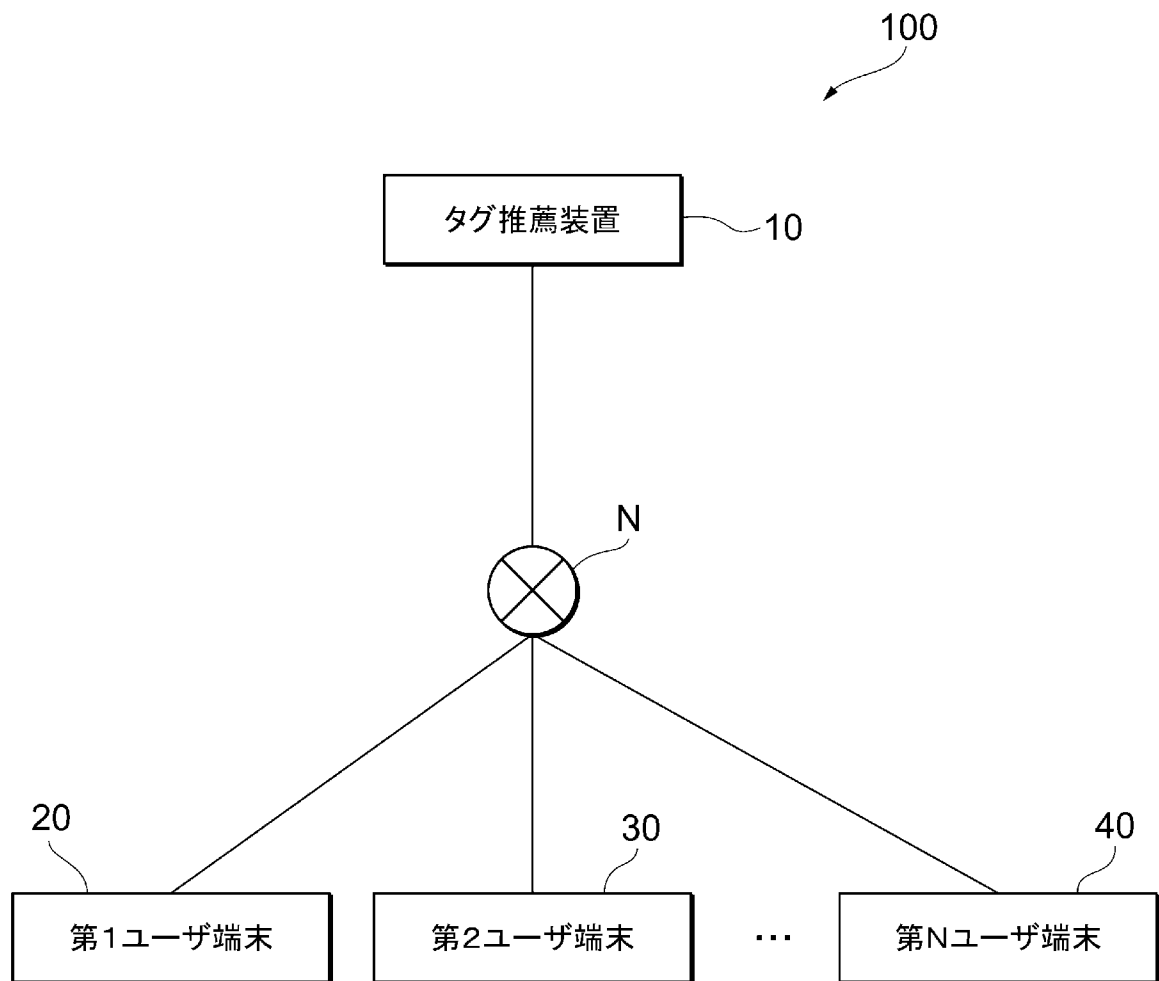
複数のユーザと複数のタグとの共起関係を、前記複数のユーザが前記複数のタグを複数のコンテンツに付与した頻度に応じて重み付けした第1数値と、前記複数のユーザと前記複数のタグとの共起関係を、前記複数のユーザの人気度に応じて重み付けした第2数値とに基づいて、前記複数のタグ間の共起関係を表す第3数値を算出することと、

前記第3数値及び新たなコンテンツに付与された1又は複数のタグに基づいて、前記複数のタグの優先度を算出することと、

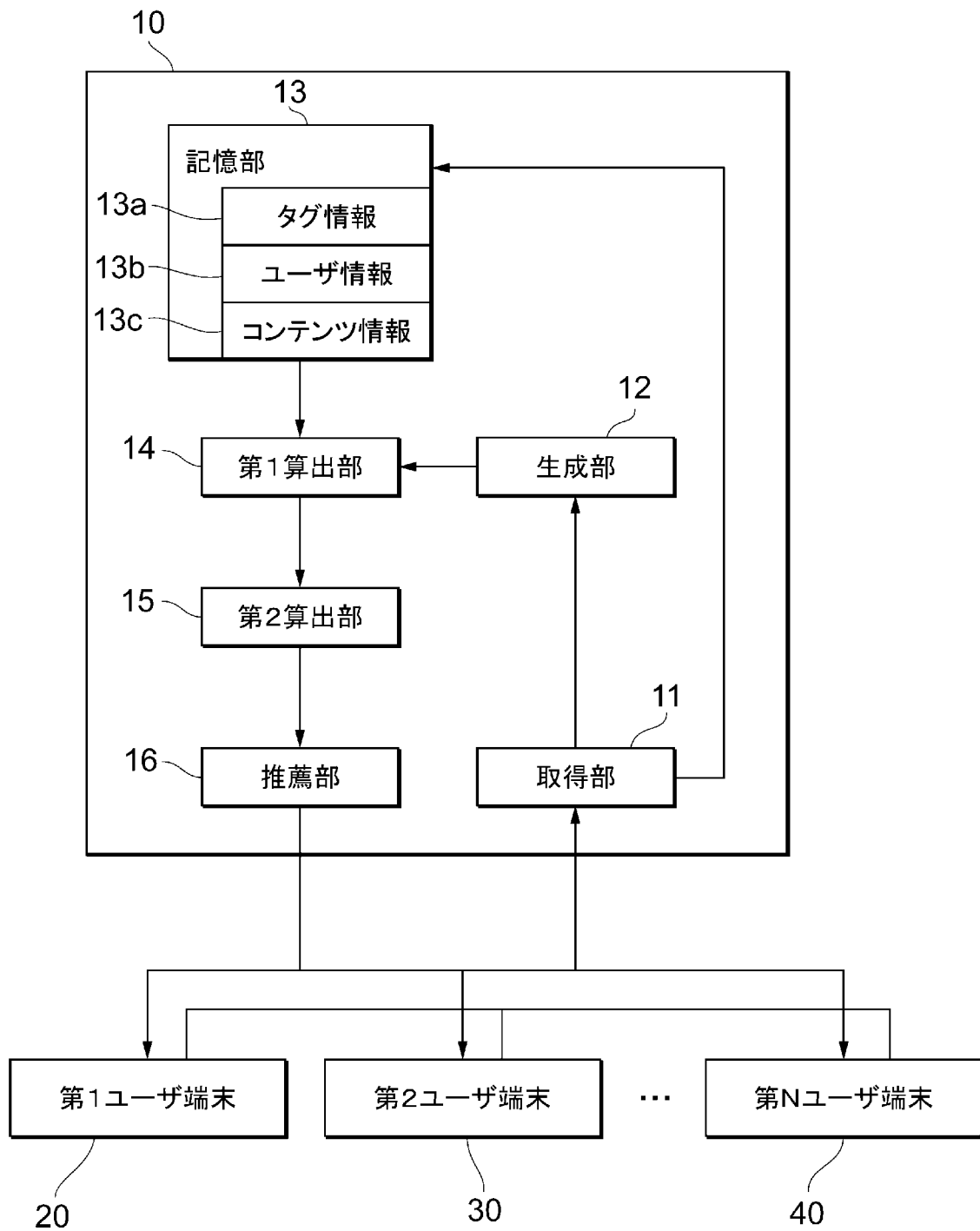
前記優先度に基づいて、前記複数のタグの中から、前記新たなコンテンツに付与する1又は複数のタグを推薦することと、

を実行させるタグ推薦プログラム。

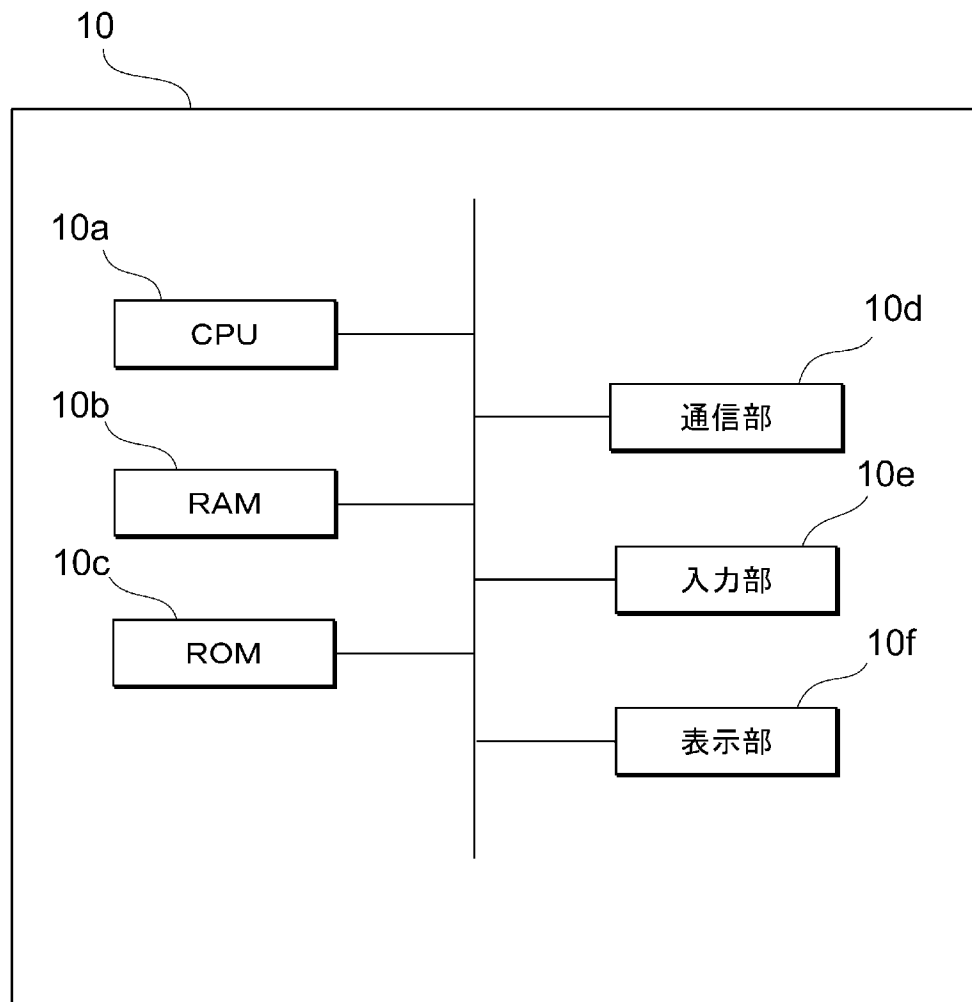
[図1]



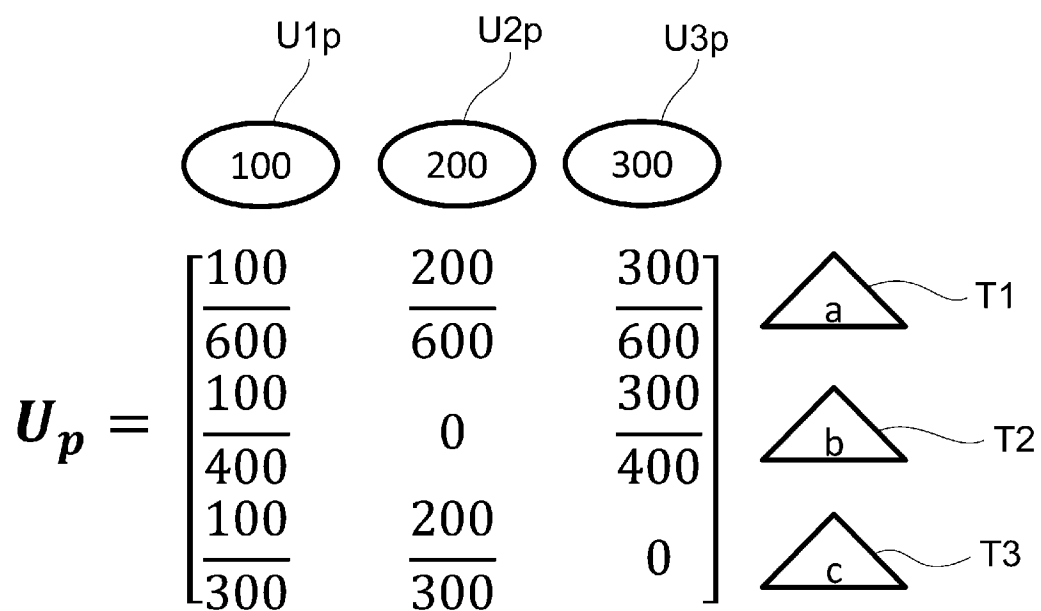
[図2]



[図3]



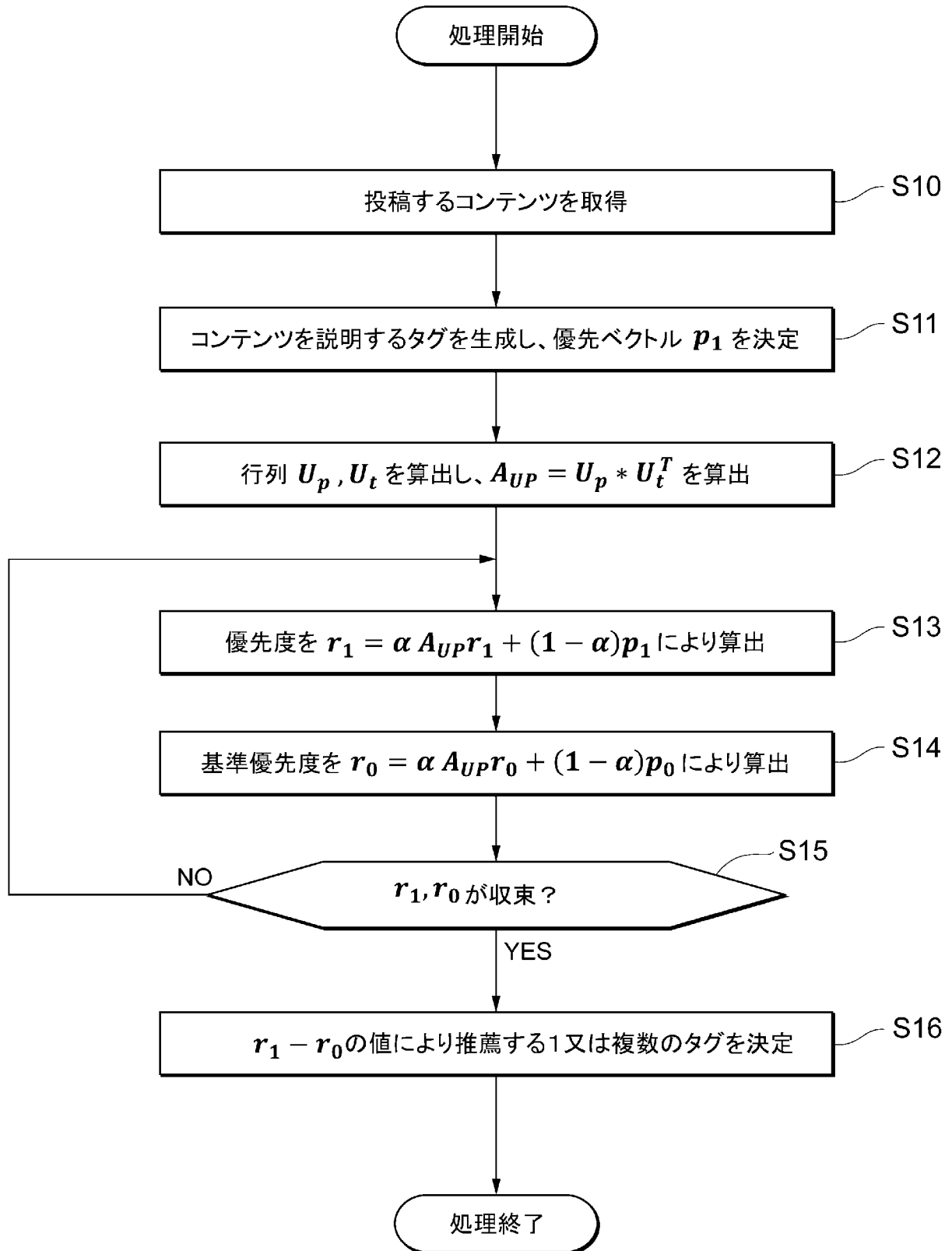
[図4]



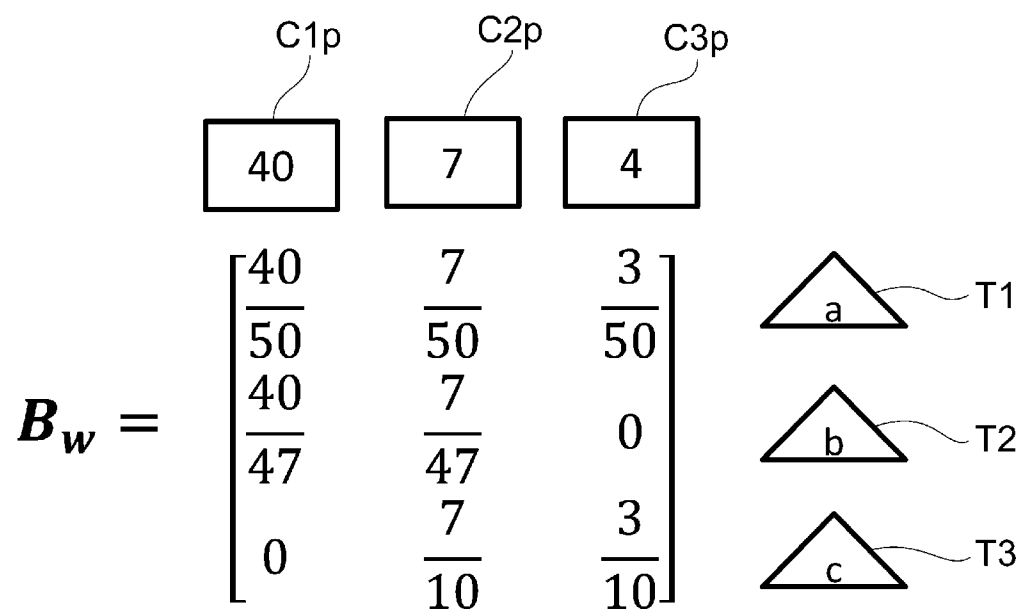
[図5]

$$\begin{array}{c}
 \begin{array}{ccc}
 \text{U1f} & \text{U2f} & \text{U3f} \\
 \text{a:2 b:1} & \text{a:2} & \text{a:3} \\
 \text{c:3} & \text{c:1} & \text{b:3}
 \end{array} \\
 \\
 \begin{array}{ccc}
 \left[\begin{array}{c} 2 \\ \frac{1}{6} \\ 1 \\ \frac{1}{6} \\ 3 \\ \frac{1}{6} \end{array} \right] & \begin{array}{c} 2 \\ \frac{2}{3} \\ 0 \\ 1 \\ \frac{1}{3} \end{array} & \left[\begin{array}{c} 3 \\ \frac{3}{6} \\ 3 \\ \frac{3}{6} \\ 0 \end{array} \right]
 \end{array}
 \end{array}
 \begin{array}{c}
 \begin{array}{c} \triangle \\ \text{a} \end{array} \text{T1} \\
 \begin{array}{c} \triangle \\ \text{b} \end{array} \text{T2} \\
 \begin{array}{c} \triangle \\ \text{c} \end{array} \text{T3}
 \end{array}
 \end{array}$$

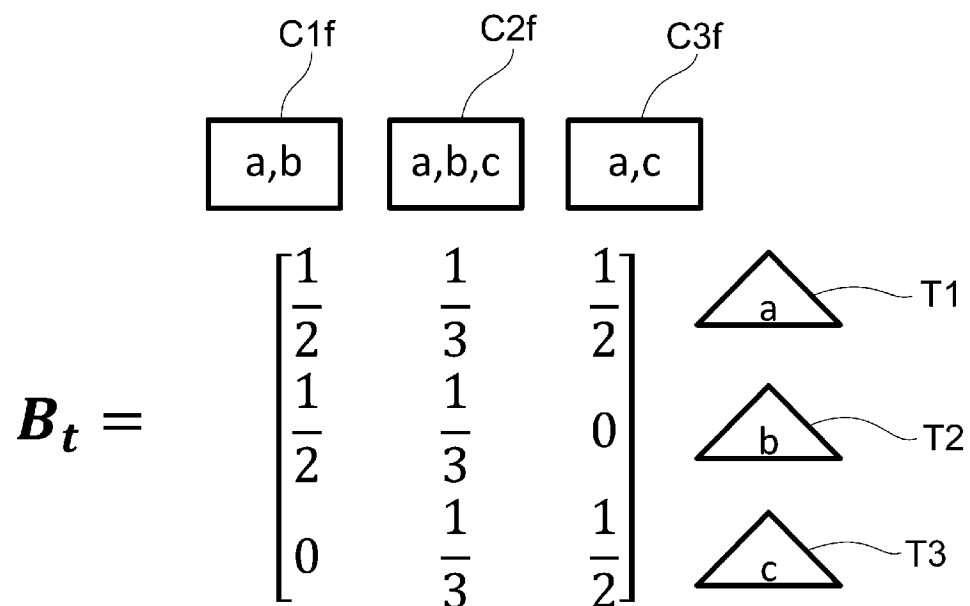
[図6]



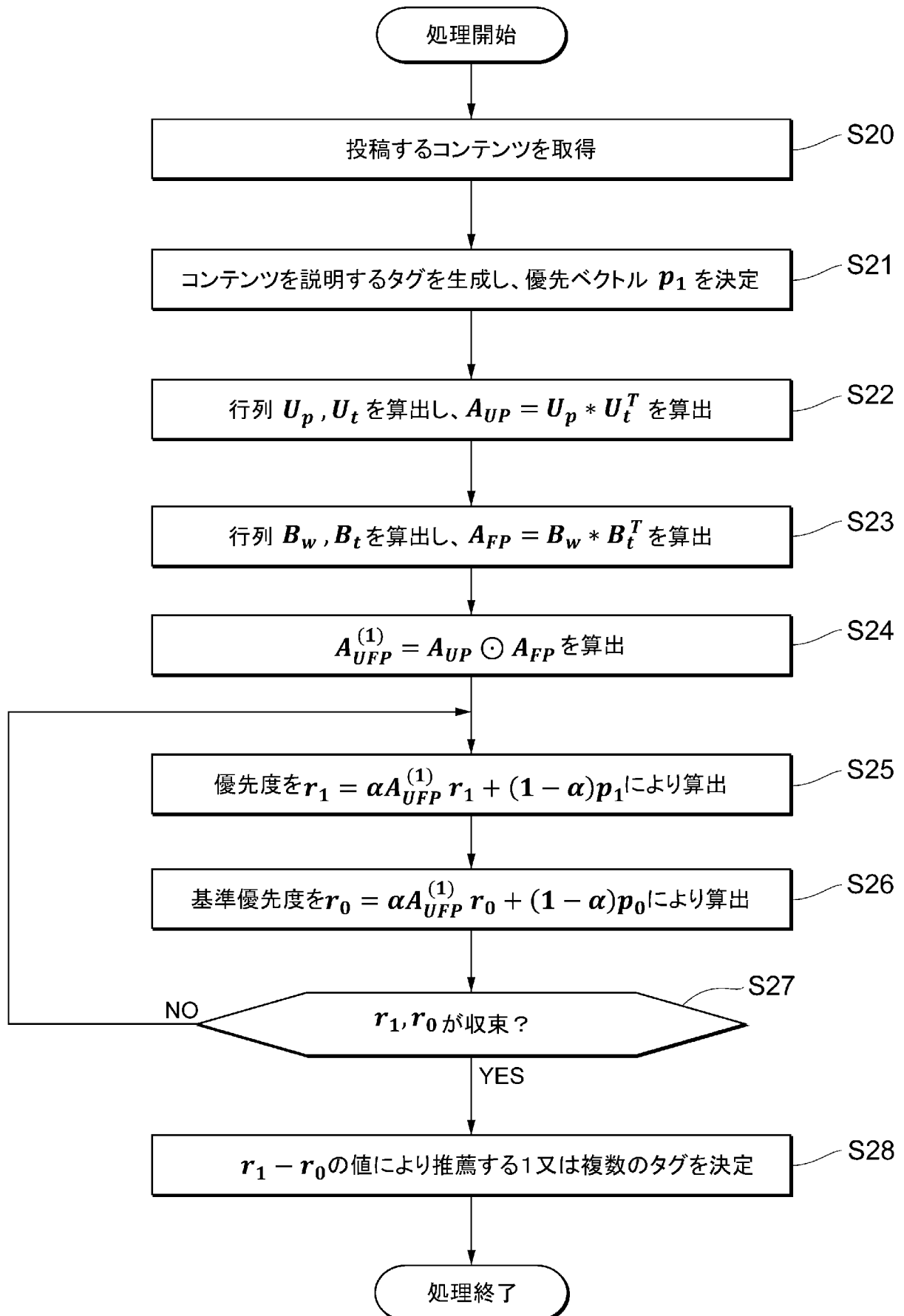
[図7]



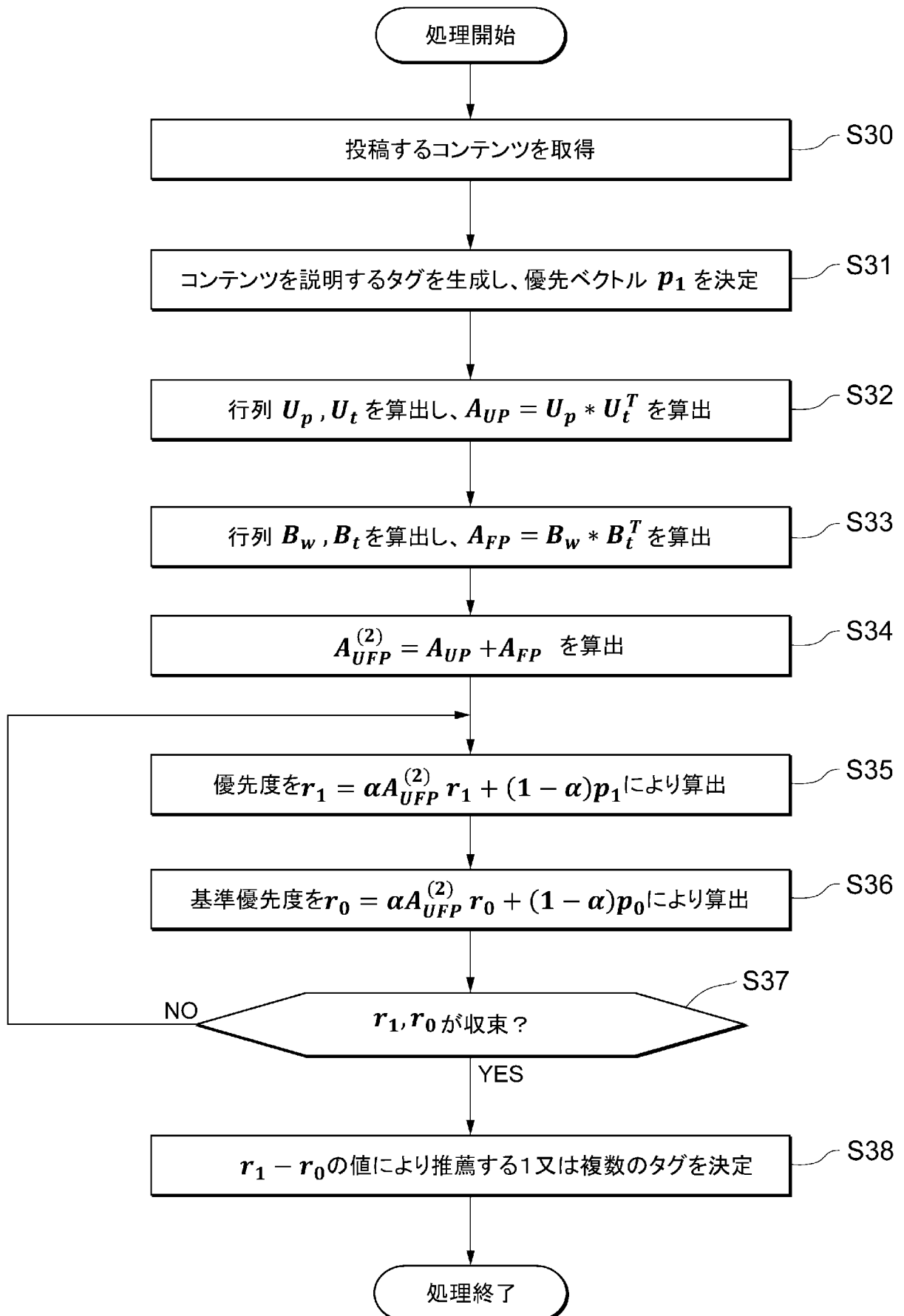
[図8]



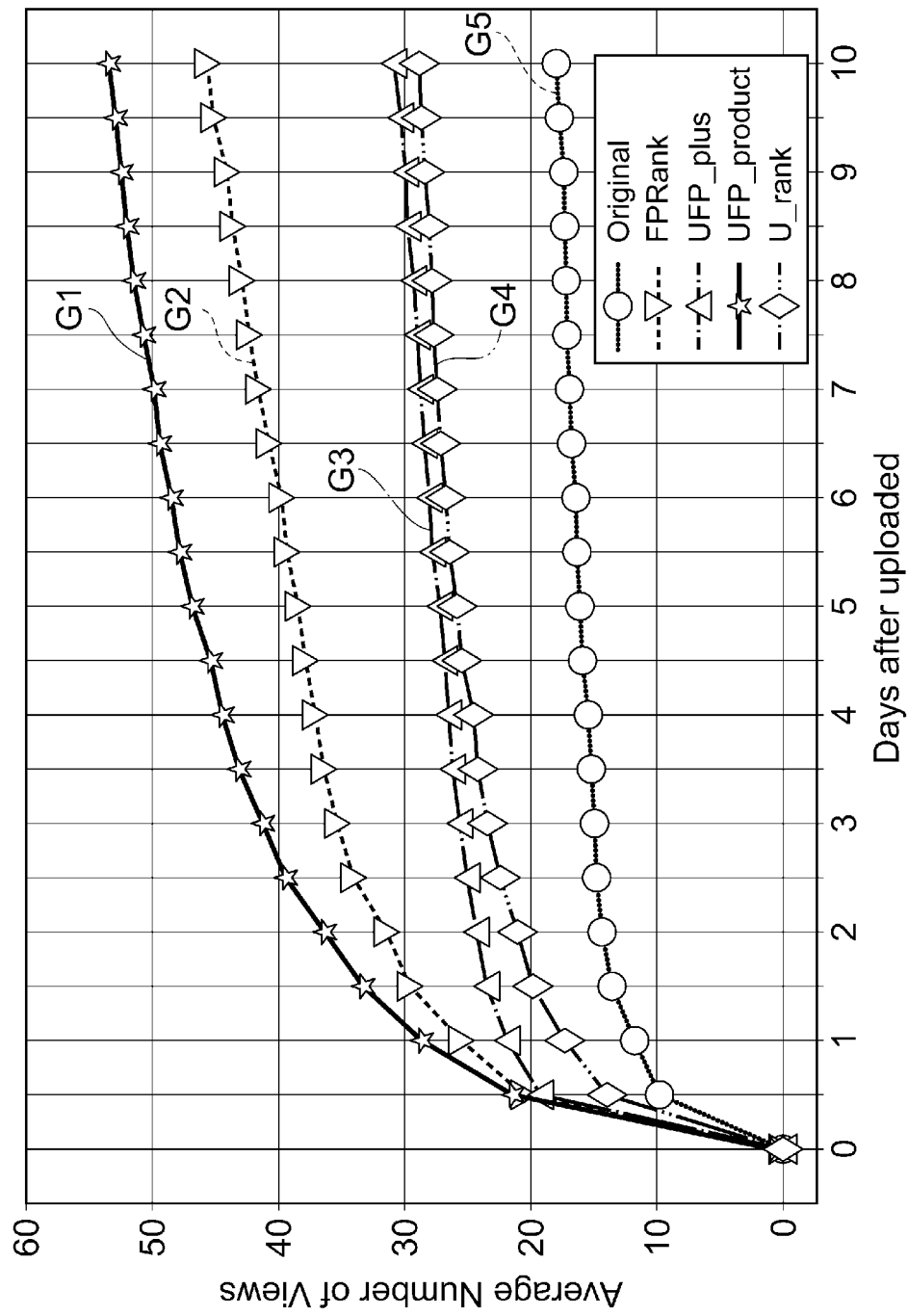
[図9]



[図10]



[図11]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2020/011225

A. CLASSIFICATION OF SUBJECT MATTER

G06Q 50/10(2012.01)i; G06F 16/9035(2019.01)i; G06F 16/9536(2019.01)i
FI: G06F16/9035; G06F16/9536; G06Q50/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06Q50/10; G06F16/9035; G06F16/9536

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan	1922-1996
Published unexamined utility model applications of Japan	1971-2020
Registered utility model specifications of Japan	1996-2020
Published registered utility model applications of Japan	1994-2020

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	YAMASAKI, Toshihiko, HU, Jiani, SANO, Shumpei, AIZAWA, Kiyoharu, "FolkPopularityRank: Tag Recommendation for Enhancing Social Popularity using Text Tags in Content Sharing Services", Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17), 19 August 2017, pp. 3231-3237 "2. Folk Popularity Rank", "3. Experimental Results"	1-11
A	US 2015/0186503 A1 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) 02.07.2015 (2015-07-02) paragraphs [0025]-[0026]	1-11
A	JP 2015-76089 A (DNA KK.) 20.04.2015 (2015-04-20) paragraph [0070]	1-11



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"I" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

07 April 2020 (07.04.2020)

Date of mailing of the international search report

28 April 2020 (28.04.2020)

Name and mailing address of the ISA/

Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/JP2020/011225

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
US 2015/0186503 A1	02 Jul. 2015	CN 103729360 A paragraphs [0036]- [0042]	
JP 2015-76089 A	20 Apr. 2015	WO 2014/056397 A1 (Family: none)	

A. 発明の属する分野の分類（国際特許分類（IPC）） G06Q 50/10(2012.01)i; G06F 16/9035(2019.01)i; G06F 16/9536(2019.01)i FI: G06F16/9035; G06F16/9536; G06Q50/10														
B. 調査を行った分野														
調査を行った最小限資料（国際特許分類（IPC）） G06Q50/10; G06F16/9035; G06F16/9536														
最小限資料以外の資料で調査を行った分野に含まれるもの <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922-1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971-2020年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996-2020年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994-2020年</td> </tr> </table>			日本国実用新案公報	1922-1996年	日本国公開実用新案公報	1971-2020年	日本国実用新案登録公報	1996-2020年	日本国登録実用新案公報	1994-2020年				
日本国実用新案公報	1922-1996年													
日本国公開実用新案公報	1971-2020年													
日本国実用新案登録公報	1996-2020年													
日本国登録実用新案公報	1994-2020年													
国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）														
C. 関連すると認められる文献														
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号												
A	YAMASAKI, Toshihiko, HU, Jiani, SANO, Shumpei, AIZAWA, Kiyoharu, FolkPopularityRank: Tag Recommendation for Enhancing Social Popularity using Text Tags in Content Sharing Services, Proceedings of the Twenty- Sixth International Joint Conference on Artificial Intelligence (IJCAI-17), 2017.08.19, p.3231-3237 [2. Folk Popularity Rank], [3. Experimental Results]	1-11												
A	US 2015/0186503 A1 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) 02.07.2015 (2015 - 07 - 02) [0025]-[0026]	1-11												
A	JP 2015-76089 A (株式会社 ディー・エヌ・エー) 20.04.2015 (2015 - 04 - 20) [0070]	1-11												
☐ C欄の続きにも文献が列挙されている。 ☒ パテントファミリーに関する別紙を参照。														
<table border="0"> <tr> <td>* 引用文献のカテゴリー</td> <td>“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの</td> </tr> <tr> <td>“A” 特に関連のある文献ではなく、一般的な技術水準を示すもの</td> <td>“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの</td> </tr> <tr> <td>“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの</td> <td>“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの</td> </tr> <tr> <td>“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）</td> <td>“&” 同一パテントファミリー文献</td> </tr> <tr> <td>“O” 口頭による開示、使用、展示等に言及する文献</td> <td></td> </tr> <tr> <td>“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献</td> <td></td> </tr> </table>			* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの	“A” 特に関連のある文献ではなく、一般的な技術水準を示すもの	“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの	“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの	“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）	“&” 同一パテントファミリー文献	“O” 口頭による開示、使用、展示等に言及する文献		“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献	
* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの													
“A” 特に関連のある文献ではなく、一般的な技術水準を示すもの	“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの													
“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの													
“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）	“&” 同一パテントファミリー文献													
“O” 口頭による開示、使用、展示等に言及する文献														
“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献														
国際調査を完了した日 07.04.2020		国際調査報告の発送日 28.04.2020												
名称及びあて先 日本国特許庁(ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号		権限のある職員（特許庁審査官） 鹿野 博嗣 5N 4063 電話番号 03-3581-1101 内線 3586												

国際調査報告
パテントファミリーに関する情報

国際出願番号

PCT/JP2020/011225

引用文献			公表日	パテントファミリー文献			公表日
US	2015/0186503	A1	02.07.2015	CN	103729360	A	
					[0036]-[0042]		
				WO	2014/056397	A1	
JP	2015-76089	A	20.04.2015	(ファミリーなし)			